United States Department of Commerce
Technology Administration
National Institute of Standards and Technology

NIST Special Publication 866

Extreme Value Theory and Applications

Proceedings of the Conference on Extreme Value Theory and Applications, Volume 3 Gaithersburg, Maryland, May 1993

Janos Galambos, James Lechner and Emil Simiu, Editors
Charles Hagwood, Technical Editor

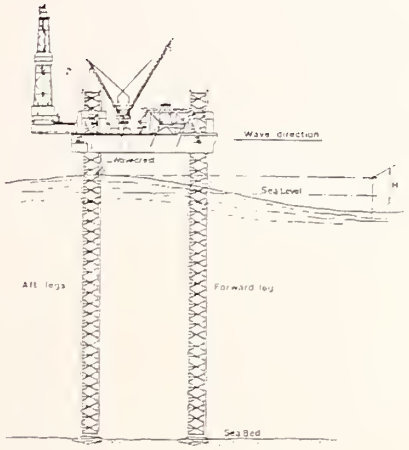


Figure 1 Jack-up platform with cantilever

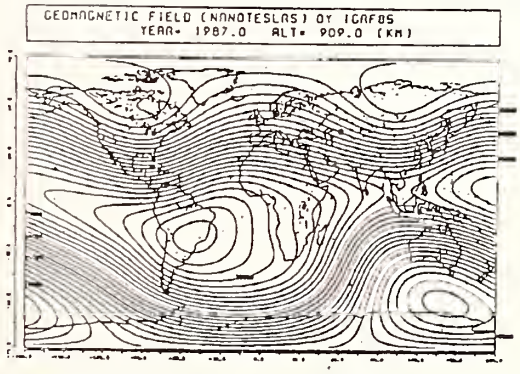
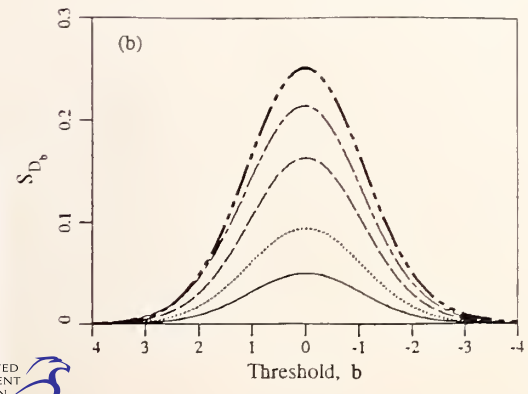
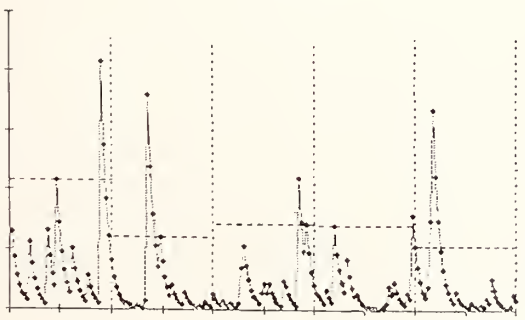


Figure 10 Geomagnetic Field Contour Map On MOS-1 Orbit

Type I (*Gumbel*): $\Lambda(x) = \exp(-\exp(-x)),\ x \in \mathbb{R}$
Type II (*Fréchet*): $\Phi_\alpha(x) = \exp(-x^{-\alpha}),\ x > 0,\ \alpha > 0$
Type III (*Weibull*): $\Psi_\alpha(x) = \exp(-(-x)^{\alpha}),\ x < 0,\ \alpha > 0$





The National Institute of Standards and Technology was established in 1988 by Congress to "assist industry in the development of technology . . . needed to improve product quality, to modernize manufacturing processes, to ensure product reliability . . . and to facilitate rapid commercialization . . . of products based on new scientific discoveries."

NIST, originally founded as the National Bureau of Standards in 1901, works to strengthen U.S. industry's competitiveness; advance science and engineering; and improve public health, safety, and the environment. One of the agency's basic functions is to develop, maintain, and retain custody of the national standards of measurement, and provide the means and methods for comparing standards used in science, engineering, manufacturing, commerce, industry, and education with the standards adopted or recognized by the Federal Government.

As an agency of the U.S. Commerce Department's Technology Administration, NIST conducts basic and applied research in the physical sciences and engineering and performs related services. The Institute does generic and precompetitive work on new and advanced technologies. NIST's research facilities are located at Gaithersburg, MD 20899, and at Boulder, CO 80303. Major technical operating units and their principal activities are listed below. For more information contact the Public Inquiries Desk, 301-975-3058.

Technology Services
- Manufacturing Technology Centers Program
- Standards Services
- Technology Commercialization
- Measurement Services
- Technology Evaluation and Assessment
- Information Services

Electronics and Electrical Engineering Laboratory
- Microelectronics
- Law Enforcement Standards
- Electricity
- Semiconductor Electronics
- Electromagnetic Fields[1]
- Electromagnetic Technology[1]

Chemical Science and Technology Laboratory
- Biotechnology
- Chemical Engineering[1]
- Chemical Kinetics and Thermodynamics
- Inorganic Analytical Research
- Organic Analytical Research
- Process Measurements
- Surface and Microanalysis Science
- Thermophysics[2]

Physics Laboratory
- Electron and Optical Physics
- Atomic Physics
- Molecular Physics
- Radiometric Physics
- Quantum Metrology
- Ionizing Radiation
- Time and Frequency[1]
- Quantum Physics[1]

Manufacturing Engineering Laboratory
- Precision Engineering
- Automated Production Technology
- Robot Systems
- Factory Automation
- Fabrication Technology

Materials Science and Engineering Laboratory
- Intelligent Processing of Materials
- Ceramics
- Materials Reliability[1]
- Polymers
- Metallurgy
- Reactor Radiation

Building and Fire Research Laboratory
- Structures
- Building Materials
- Building Environment
- Fire Science and Engineering
- Fire Measurement and Research

Computer Systems Laboratory
- Information Systems Engineering
- Systems and Software Technology
- Computer Security
- Systems and Network Architecture
- Advanced Systems

Computing and Applied Mathematics Laboratory
- Applied and Computational Mathematics[2]
- Statistical Engineering[2]
- Scientific Computing Environments[2]
- Computer Services[2]
- Computer Systems and Communications[2]
- Information Systems

[1] At Boulder, CO 80303.
[2] Some elements at Boulder, CO 80303.

NIST Special Publication 860

Extreme Value Theory and Applications

Proceedings of the Conference on Extreme Value Theory and Applications, Volume 3 Gaithersburg, Maryland, May 1993

Janos Galambos, James Lechner and Emil Simiu, Editors
Charles Hagwood, Technical Editor

Statistical Engineering Division
Computing and Applied Mathematics Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-0001

August 1994

U.S. Department of Commerce
Ronald H. Brown, Secretary

Technology Administration
Mary L. Good, Under Secretary for Technology

National Institute of Standards and Technology
Arati Prabhakar, Director

National Institute of Standards
and Technology
Special Publication 866
Natl. Inst. Stand. Technol.
Spec. Publ. 866
231 pages (Aug. 1994)
CODEN: NSPUE2

U.S. Government Printing Office
Washington: 1994

For sale by the Superintendent
of Documents
U.S. Government Printing Office
Washington, DC 20402

PREFACE

It appears that we live in an age of disasters: the Mississippi and the Missouri rivers flood millions of acres, earthquakes hit Tokyo and California, airplanes crash due to mechanical failure, and powerful windstorms cause increasingly costly damage. While these may seem to be unexpected phenomena to the man on the street, they are actually happening according to well defined rules of science known as extreme value theory. For many phenomena records must be broken in the future, so if a design is based on the worst case of the past then we are not really prepared for the future. Materials will fail due to fatigue: even if the body of an aircraft looks fine to the naked eye, it might suddenly fail if the aircraft has been in operation over an extended period of time. Extreme value theory has by now penetrated the social sciences, the medical profession, economics, and even astronomy. We believe this field has come of age. To utilize and stimulate progress in the theory of extremes and promote its application, an international conference was organized in which equal weight was given to theory and practice.

The Proceedings are published in three Volumes. Volume I, published by Klewer Academic Publishers, contains papers of general interest in extreme value theory and practice. Volume II, a special issue of the NIST Journal of Research, contains papers deemed by the Committee to be most directly relevant to NIST's mission. Volume III (this volume) contains papers selected for their important contribution to a number of specialized topics. All papers have been refereed and we are grateful to the many scientists from all over the world for serving as referees.

The conference was held on the campus of the National Institute of Standards and Technology (NIST) in Gaithersburg, Maryland, with its Statistical Engineering Division (SED) acting as host. It was organized by Temple University, Philadelphia, Pennsylvania, and NIST.

The conference had no external funding, and NIST's support was fundamental to its success. We are particularly grateful to Dr. Robert Lundegard, Chief of SED, whose support was the single most important factor in making the conference happen. The support of NIST's Building and Fire Research Laboratory is also acknowledged with thanks.

The Organizing Committee consisted of Janos Galambos (Chairman), James Lechner, Stefan Leigh (Director of Local Arrangements), James Pickands III, Emil Simiu, and Grace Yang. Stefan's enthusiasm and tireless work was essential for the success of the Conference.

The Conference included three special sessions:

The Centennial Session for Emil Gumbel. Churchill Eisenhart introduced the Session. His personal recollections of Gumbel are included in Volume I of the Proceedings. Emil Simiu then spoke on Gumbel's life and work.

The Memorial Session for Josef Tiago de Oliveira. Janos Galambos remembered Tiago, a close friend to many Conference participants, who was on the initial list of invited speakers. M. Ivette Gomes gave a detailed account of his work.

The 80th Birthday Session for B. V. Gnedenko. Janos Galambos summarized the work of Gnedenko as the founder of modern extreme value theory and his contributions to the central limit problem, limit theorems with random sample size and renewal theory.

Preceding the Conference, a Short Course was presented. Prof. Galambos gave an introductory lecture on general principles of extreme value theory, and Prof. Castillo presented a four-hour course on "Engineering Analysis of Extreme Value Data." Prof. Castillo's notes were distributed to all Conference participants.

The Conference was opened by Dr. Robert Lundegard who emphasized extreme value theory's role in several scientific and engineering fields. It ended with a panel discussion on the future of extreme value theory and its applications. The Panel was chaired by Janos Galambos, and its members were Enrique Castillo, Laurens de Haan, Lucien Le Cam and Richard L. Smith.

Finally, special thanks are extended to Shirley G. Bremer and Kaye Wade of the Statistical Engineering Division at NIST, for their tireless and efficient work on preparation for the Conference, including the typing of the Abstracts volume distributed at registration.

The Editors

TABLE OF CONTENTS

On The Record Values From Univariate Distributions

Ahsanullah, M.
Rider College, Lawerenceville, NJ

In this paper the basic concepts and properties of the records of univariate continuous distributions are presented. Inferences about the location and scale parameters of a class of univariate distributions are given. Prediction of sth record value based on the observed first m (m < s) record values are discussed.

1.0 Introduction

Suppose that $X_1, X_2,...$ is a sequence of independent and identically distributed (i.i.d.) random variables with cumulative distribution function F(x). Set Y_n = max (min)$\{X_1, ...,X_n\}$, $n \geq 1$. We say X_j is an upper (lower)record value of $\{ X_n, n \geq 1\}$, if $Y_j >$ (<)Y_{j-1}, $j>1$. By definition, X_1 is an upper as well as a lower record value. Thus the upper record values in the sequence $\{X_n, \geq 1\}$ are the successive maxima. For example, consider the weighing of objects on a scale missing it's spring. An object is placed on this scale and its weight measured. The 'needle' indicates the correct value but does not return to zero when the object is removed. If various objects are placed on the scale , only the weights greater than the previous ones can be recorded. These recorded weights are the upper record value sequence. Let X_{ij} be the highest water level of a river on the j th day of the i th location. If one is interested to study at each location the local maximum values of X_{ij}, then the local maxima are the upper record values.

Suppose we consider a sequence of products that may fail under sets. We are interested to determine the minimum failure stress of the products sequentially. We test the first product until it fails with stress less than X_1 then we record its failure stress, otherwise we consider the next product. In general we will record sets X_m of the m th product if $X_m < \min(X_1,...,X_{m-1})$, $m>1$. The recorded failure stresses are the lower record values. One can go from lower records to upper records by replacing the original sequence of rv.'s by $\{X_j, j \geq 1\}$ or if $P(X_j > 0) = 1$ by $\{ 1/X_i, i \geq 1\}$. Unless mentioned otherwise we will call the upper record values as record values. The indices at which the record values occur are given by the record times $\{U(n)\}$, $n \geq 0$, where $U(n) = \min\{j | j>U(n-1), X_j > X_{U(n-1)}, n>1\}$ and $U(1) = 1$. The record times of the sequence $\{X_n\ n \geq 1\}$ are the same as those for the sequence $\{F(X_n), \geq 1\}$. Since F(X) has an uniform distribution, it follows that the distribution of U(n) , $n \geq 1$ does not depend on F. For a given set of n observations, let $X_{1,n} \leq X_{2,n} \leq ... X_{n,n}$ be the associated order statistics.

Suppose that $P\{a_n (X_{n,n} - b_n) \leq x\} \xrightarrow{d} G(x)$ as $n \to \infty$. For necessary and sufficient conditions about this convergence for various distributions see Galambos (1987). It is well known, [Ref. [10], that

$$P(a_n (X_{n-m,n} - b_n) \leq x) \xrightarrow{d} G(x) \sum_{s=0}^{m} \frac{\{-\ln G(x)\}^s}{s!}$$

It can be shown that the right side of the above equation is the distribution function of the m th lower record value. Properties of record values of i.i.d. rvs have been extensively studied in literature, for example, see Ref. [1], Ref. [13], and Ref. {14], and

Ref. [4], for recent reviews. The conditional p.d.f. f_{1c} of $Z_{U(n)}$ given $Z_{U(n-1)} = y$ can be written as $f_{1c}(z) = z\, f(z) / (1 - F(y))$.

For many distributions including exponential, Pareto and uniform

$$E(X_{U(n)} \mid X_{U(n-1)} = y) = a + b\,y \qquad (1.1)$$

for some constants a and b . We will say a rv X with distribution function F belongs to the class C if its n th record value satisfy the condition (1.1).. In this paper, we will consider the record values of random variables belonging to class C.

2. MAIN RESULTS

RESULT 1.

If the sequence of rvs $X_1, X_2, \ldots$, belong to class C with finite variance, then

$$\mathrm{Cov}(X_{U(n)}, X_{U(m)}) = b^{n-m}\,\mathrm{Var}(X_{U(m)}),\ n > m.$$

Proof:

$$\begin{aligned} E(X_{U(m+2)}) &= EE(X_{U(m+2)} \mid X_{U(m+1)} = t) \\ &= EE(a + bt \mid X_{U(m)} = y) \\ &= E(a + b(a + b\,y)) \\ &= a + ab + b^2 E(X_{U(m)}). \end{aligned}$$

In general

$$E(X_{U(n)}) = a + ab + ab^2 + a\,b^3 + \ldots + ab^{n-m-1} + b^{n-m}.$$

$$= a\frac{b^{b-n} - 1}{b-1} + b^{n-m}E(X_{U(m)}),\ \text{if } b \neq 1$$

$$= (n-m)\ a + E(X_{U(m)}),\ \text{if } b = 1.$$

Thus

$$\mathrm{Cov}(X_{U(m)}, X_{U(n)}) = b^{n-m}\,\mathrm{Var}(X_{U(m)}),\ \text{if } b \neq 1$$

$$= \mathrm{Var}(X_{U(m)}),\ \text{if } b = 1.$$

The following result was proved by Ref. [12].

RESULT 2.

If the sequence of i.i.d. rvs $X_1, X_2, \ldots$ has an absolutely continuous distribution function F with support on [c, d), where c is finite and d may be infinite and has finite expectation. Further we assume that F belongs to the class C with $b > 0$ and with $d = \infty$ if $b \geq 1$ and $d = -\frac{a}{b-1}$ if $b > 0$.

Then $1 - F(x) = \left(\frac{a + (b-1)c}{a + (b-1)x}\right)^{\frac{b}{b-1}}$ for $b \neq 1$

and $1 - F(x) = e^{-x/a}$ for $b = 1$,

if and only if $E(X_{U(n)} \mid X_{U(n-1)} = y) = a + b\,y$.

Proof:

Writing the conditional expectation of $X_{U(n)} \mid X_{U(n-1)} = y)$ and simplifying we get

$$a + b\,y = y + \int_y^d \frac{1 - F(x)}{1 - F(y)}\,dx \qquad (2.1)$$

Differentiating both sides of the above equation with respect to y, we obtain

$$b = (a + (b-1)\,y)\,(f(y) / (1 - F(y))) \qquad (2.2)$$

Finally integrating (2.2) with respect to y from c to x, we obtain the result.

For general discussions of result 2 based on conditional expected values of rv X, see Ref. [7].

It can be shown that for the rv X having the distribution function as given in the Result 2,

$$E(X) = a + bc \text{ and } \mathrm{Var}(X)$$

$$= (b(a + (b-1)c)^2) / (2 - b).$$

For various results of the rv X based record values f or the case b =1, see Ref. [2] and [3]. The results for the case b >1 are similar to those for the case b < 1.

In this paper, we will consider the results corresponding to b>1. The distribution function corresponding to b > 1 was introduced by Pickands (1975) in connection to extreme value distribution.

For inference based on record values for the generalized extreme value distribution see REf. [3].

Let f_n be the probability density function of the n th record value, $X_{U(n)}$. Then

$$f_n(x) = \frac{\left\{\frac{b}{b-1}\ln\frac{a+(b-1)c}{a+(b-1)x}\right\}^n}{n!}\,\frac{b(a+(b-1)c^{\frac{b}{b-1}}}{(a+(b-1)x)^{\frac{2b-1}{b-1}}} \qquad (2.3)$$

It can be shown from (2.3) that

$$E(X_{U(n)}) = (1/(b-1))(b^n(a+(b-1)c) - a)$$

$$Var(X_{U(n)}) = (b-1)^{-2}(a+(b-1)c)^2\, b^n((2-b)^{-n} - b^n).$$

and $Cov(X_{U(m)}, X_{U(n)}) = b^{n-m}\, Var(X_{U(m)})$.

We will assume without the loss generality the lower bound, c of the rv X as zero and $(Y - m)/s = X$, then

$E(Y) = m + a\,s$ and $Var(Y) = a^2 b(2-b)^{-1} s^2$.

For the finite variance, b must be less than 2.

Let $T_1, T_2, \ldots, T_n$ be the record values of Y corresponding to $X_{U(1)}$,

$X_{U(2)}, \ldots, X_{U(n)}$.

It can be shown that

$$T_n = \mu - \frac{a\sigma}{b-1} + \frac{a\sigma}{b-1}\prod_{i=1}^{n} U_i$$

where $U_1, U_2, \ldots, U_n$ are independent and identically distributed with

$$P(U_i \le x) = 1 - x^{-b/(b-1)}$$

Thus

$$E(T_n) = m + a\frac{b^n - 1}{b-1}\sigma$$

and

$$Var(T_n) = a^2(b-1)^{-2} b^n \{(2-b)^{-n} - b^n\} s^2.$$

$$Cov(T_m, T_n) = b^{n-m}\, Var(T_m),\ m < n.$$

We can write the Variances and Covariances of T's as

$$Var(T_r) = a_r b_r \sigma_1^2 \text{ and } Cov(T_r, T_s) = a_r b_s \sigma_1^2,\ r \le s,$$

where $a_r = [(2-b)^{-r} - b^r]$, $b_r = b^r$, r= 1,2, ...,

and $\sigma_1^2 = a^2 (b-1)^{-2} \sigma^2$

There are other distributions see Ref. [5], for which $Cov(T_r, T_s)$ can be factored out as the product of two factors, one depends on r and the parameters and the other depends on s and the parameters.

ESTIMATORS OF m AND s .

The minimum Variance linear unbiased estimator (MVLUE) of $\hat{\mu}, \hat{\sigma}$ of m and s are

$$\hat{\mu} = T_1 - a\hat{\sigma}$$

$$\hat{\sigma} = a^{-1}\left[\left\{\frac{b-1}{b} - D^{-1}\left(\frac{2-b}{b}\right)^3\right\} T_1 + D^{-1}\frac{2-b}{b}\sum_{i=2}^{n-1}\left(\frac{2-b}{b}\right)^{i+1} T_i + D^{-1}\left(\frac{2-b}{b}\right)^{n+1} T_n\right]$$

where

$$D = \sum_{i=2}^{n-1}\left(\frac{2-b}{b}\right)^{i+1}.$$

Proof.

Let $T' = (T_1, T_2, \ldots, T_n)$, then we can write

$E(T) = m L + d\,\sigma_1$

$L' = (1,1,\ldots,1)$, $d' = (d_1, d_2, \ldots, d_n)$ and
$d_i = b^n - 1$, $i = 1,2,\ldots, n$.

Let $V(T) = \sigma_1^2 \Sigma$, $S^{-1} = W$, $W = (V^{i,j})$
It can be shown that

$$V^{i+1,i} = V^{i,i+1} = -\frac{b}{(1-b)^2}\left(\frac{2-b}{b}\right)^{i+1}$$

$$V^{i,i} = \frac{1+2b-b^2}{(1-b)^2}\left(\frac{2-b}{b}\right)^{i}$$

$$V^{n,n} = \frac{1}{(1-b)^2}\left(\frac{2-b}{b}\right)^{n}$$

$V^{i,j} = 0$, if $|i-j| > 1$

Let $W_i = a^{-1}((2-b)/b)^{1/2}(T_i - b\,T_{i-1})$, $i=1,2,\ldots,n$ and $T_0 = 0$.

Then $Var(W_i) = s^2$ and $Cov(W_i, W_k) = 0$, $i \neq k$, $1 \leq i, k \leq n$.

Suppose $W' = (W_1, W_2, \ldots, W_n)$ and $E(W) = Aq$, where $q' = (m,s)$
$A' = [A_1 A_2]$, $A_1' = (d_1, d_2, \ldots, d_n)$, $A_2' = (e_1, e_2, \ldots, e_n)$,

$d_i = ((2-b)/b)^i (1-b)$, $e_i = d_i/(1-b)$, $i = 2,3,\ldots,n$, $d_1 = (1/a)((2-b)/b)^{1/2}$
and $e_1 = a\,d_1$.

Using least squares estimation method, we get on simplification

$$\hat{u} = T_1 - a\hat{\sigma}$$

$$\hat{\sigma} = a^{-1}\left[\left\{\frac{b-1}{b} - D^{-1}\left(\frac{2-b}{b}\right)^3\right\} T_1 + D^{-1}\frac{2-b}{b}\sum_{i=2}^{n-1}\left(\frac{2-b}{b}\right)^{i+1} T_i + D^{-1}\left(\frac{2-b}{b}\right)^{n+1} T_n\right]$$

$$Var(\hat{\mu}) = \left(\frac{a^2 T}{b^2 D}\right)\sigma^2$$

$$Var(\hat{\sigma}) = \left\{\left(\frac{2-b}{b}\right)^2 + \left(\frac{b-1}{b}\right)^2 T\right\}\sigma^2/D$$

$$Cov(\hat{\mu},\hat{\sigma}) = \left\{\left(\frac{b-1}{b}\right)T - \frac{2-b}{b}\right\}\frac{a\sigma^2}{bD}$$

where

$$T = \sum_{i=1}^{n}\left(\frac{2-b}{b}\right)^i \text{ and } D = \sum_{i=2}^{n}\left(\frac{2-b}{b}\right)^{i+1}.$$

Let $a = 2$ and $b = 1.5$, then

$\hat{\mu} = (17/12)\,T_1 - (1/4)\,T_2 - (1/12)\,T_3 - (1/12)\,T_4$
and

$\hat{\sigma} = -(5/8)\,T_1 + (1/8)T_2 + (1/24)T_3 + (1/24)T_4$

The corresponding variance and covariance are

$$Var(\hat{\sigma}) = \frac{40}{81}\sigma^2 \text{ and}$$

$$Var(\hat{\sigma}) = \frac{40}{81}\sigma^2 \text{ and}$$

$$Cov(\hat{\mu},\hat{\sigma}) = -\frac{41}{9}\sigma^2$$

BEST LINEAR INVARIANCE ESTIMATORS (BLIE)

The best linear invariant (in the sense of minimum mean squared error and invariance with espect to the location parameter m) estimators

$\tilde{\mu}, \tilde{\sigma}$ of μ and σ are

$$\tilde{\mu} = \hat{\mu} - \hat{\sigma}\left(\frac{E_{12}}{1+E_{22}}\right)$$

and $\tilde{\sigma} = \hat{\sigma}(1+E_{22})^{-1}$

where $\hat{\mu}$ and $\hat{\sigma}$ are MVLUE of μ and σ and

$$\begin{pmatrix} var(\hat{\mu}) & cov(\hat{\mu},\hat{\sigma}) \\ cov(\hat{\mu},\hat{\sigma}) & var(\hat{\sigma}) \end{pmatrix} = \sigma^2 \begin{pmatrix} E_{11} & E_{12} \\ E_{12} & E_{22} \end{pmatrix}$$

The mean squared errors of these estimators are

$$MSE(\tilde{\mu}) = \sigma(E_{11} - E_{12}^2(1+E_{22})^{-1})$$

$$MSE(\tilde{\mu}) = \sigma(E_{11} - E_{12}^2(1+E_{22})^{-1})$$

Substituting the values of E_{11}, E_{12}, E_{22}, we get

$$\tilde{\mu} = \hat{\mu} - \frac{b\{b-1)T-2+b\}}{T}\hat{\sigma}$$

$$\hat{\sigma} = \hat{\sigma}\frac{Db^2}{T}$$

and

$$MSE(\tilde{\mu}) = \frac{a^2\sigma^2}{b^2 D}\left[T - \frac{\{(b-1)T-(2-b)\}^2}{T}\right]$$

$$MSE(\tilde{\sigma}) = \frac{a\sigma^2}{T}[(b-1)T-(2-b)]$$

With n = 4 and b = 1.5, we have

$$\hat{\mu} = \frac{2105}{1920}T_1 - \frac{37}{640}T_2 - \frac{37}{1920}T_3 - \frac{37}{1920}T_4$$

$$\tilde{\sigma} = \frac{1125}{1200}T_1 + \frac{9}{160}T_2 + \frac{3}{160}T_3 + \frac{3}{160}T_4$$

$$MSE(\tilde{\mu}) = \frac{127413}{10800}\sigma^2$$

$$MSE(\tilde{\sigma}) = \frac{121}{160}\sigma^2$$

PREDICTOR of T_S

We shall consider the prediction of T_S based on n observed record values for s > n.

Let $H' = (h_1, h_2, ..., h_n)$, where $s^2 h_i = Cov(T_i, T_S)$, i = 1,2,...,n and $g_S = s^{-1} E(T_S - m)$. It follows from the results of Ref. [9] that the best linear unbiased predictor (BLUP) of T_S is $\hat{T}_S$, where

$$\hat{T}_S = \hat{\mu} + \hat{\sigma}\gamma_S + H'V^{-1}(T - \hat{\mu}L - \gamma)$$

Now

$H'V^{-1} = (0,0,...., b^{s-n})$ and

$$\hat{T}_S = \hat{\mu} + \hat{\sigma}\gamma_S + b^{s-n}(T_n - \hat{\mu} + \hat{\sigma}\gamma_n$$

$$= b^{s-n}T_n + (1-b^{2-n})\hat{\mu} + (1-b^{s-n})\gamma_S\hat{\sigma}$$

The best(unrestricted) least squares predictor of T_S is $T_S^* = E(T_S | T_1, T_2, .., T_n)$.

Thus

$$T_s^* = \mu + a\frac{b^{2-n}-1}{b=1}\sigma + ab^{2-n}(T_n - \mu)$$

If we substitute the MVLUE of m and s , then T_S^* becomes $\hat{T}_S$.

Let $\tilde{T}_S$ be the best linear invariant predictor of T_S.

From the results of Ref. [11] it follows that

$$MSE(\tilde{\mu}) = \frac{127413}{10800}\sigma^2$$

where

$$c_{12}^{*} = \mathrm{Cov}(\hat{\sigma}, (1 - H'V^{-1}L)\hat{\mu} + (\gamma_s - H'V^{-1}\delta)a\hat{\sigma}$$

and

$$1 - H'V^{-1}1 - 1 - b^{s-n} \text{ and } \gamma_s - H'V^{-1}\delta = \gamma_{s=n}$$

Thus

$$\tilde{T}_s = \hat{T}_s - \gamma_{s-n}\frac{2-b}{b^2}$$

Considering the MSE of the predictor, it can be shown that

$$\mathrm{MSE}(T_s^{*}) \le \mathrm{MSE}(\tilde{T}_s) \le \mathrm{MSE}(\hat{T}_s)$$

REFERENCES

[1] Ahsanullah, M.(1988). Introduction to Record Statistics. Ginn Press, Needham Heights, MA.

[2] Ahsanullah, M. (1980).Linear prediction of record values for the two parameter exponential distribution. Ann. Inst. Stat. Math. 32,A,363-368.

[3] Ahsanullah, M. and Holland, B.(1993) On the use of record values to estimate the Location and Scale Parameters of the Generalized Extreme Value Distributions. To appear in Sankhya.

[4] Arnold, B.C., N. Balakrishnan and H.N. Nagaraja (1992). A first Course in Order Statistics, John Wiley & Sons, New York.

[5] Balakrishnan,N., M. Ahsanullah and P.S. Chan (1992). Relations for single and product moments of record values from Gumbel distribution. Statistics and Probability letters, 13,223-227.

[6] Galambos, J.(1987). The Asymptotic Theory of Extreme Order Statistics, Second edition, Krieger, Malabar, Florida.

[7] Galambos, J. and S. Kotz (1978) Characterizations of Probability Distributions. Lecture Notes in Mathematics, 675. Springer -Verlag. New York.

[8] Goldberger, A.S. (1962). Best linear unbiased prediction in the generalized linear regression model. JASA 57, 369-379.

[9] Leadbetter, M.R., G.Lindgreen and H. Rootgen (1980). Extremes. Springer-Verlag, New York, N.Y.

[10] Mann,N.R. (1969).Optimum estimators for linear functions of location and scale parameters. Ann. Math. Statist. 40,2149-2155.

[11] Nagaraja, H.N.(1970). On a characterization based on record values. Austra. J. Statist. 19, 70-73.

[12] Nagaraja, H.N. (1988). Record values and related statistics-A review, Commun.Statist.-Theor.Meth.17(7), 2223-2238.

[13] Nevzerov,V.B. (1987).Records, Theory of Probability and Its Application 32(2), 219-251.

[14] Pickands, J. (1975). Statistical Inference using extreme order statistics. Ann. Statist.3, 119-131.

Composite Sampling And Extreme Values

Argon, E.D., Gore, S.D., and Patil, G.P.
The Pennsylvania State University, University Park, PA

Issues in environmental sampling and ecological monitoring can involve extreme values as the main inferential target, or as a design tool for cost-effective sampling. Although conventional sampling methods address the problem of estimating the population mean with a desired precision, classical procedures are not always cost-effective for studies involving extreme values. In this paper, we review some procedures that allow inference on sample extreme values based on sample means while maintaining observational economy. These procedures use a common sweepout method to identify extremely large sample values when measurements on composite samples are available. These procedures are illustrated with examples in compliance monitoring and enforcement in hazardous waste site characterization. The effect of the compositing design on the performance of the sweepout method is also investigated. In conclusion, this paper highlights the need for an investigation of the statistical properties of the sweepout method.

Keywords. Compliance monitoring, Composite sampling, Concomitants of order statistics, Extreme values. Higher order statistics, Observational economy, Percentiles, Population mean, Ranked set sampling, Site characterization.

1 Introduction

Issues in environmental and ecological monitoring can involve extreme values as a primary objective for inference or as a design tool for cost-effective sampling. For example, compliance monitoring and assessment of hazardous waste sites may require both estimation of the mean and identification of "hot spots." Choice of the sampling design must take into account these objectives as well as resource limitations and any other practical constraints. Although conventional sampling methods address the problem of estimating the mean with a desired precision, they may no longer be cost-effective for inference on extreme values. In a different situation, available information on extreme values may be used advantageously for improving upon the sampling design. For instance, perceived ranks of sampling units may be used as a sample stratification tool, thereby reducing the required sample size and/or associated cost while maintaining the desired precision. In another example, it may be of interest to estimate fish abundance by sampling from known high-abundance sites.

We review some procedures that are based on composite sampling techniques and address the issue of identifying extreme sample values when measurements on composite samples are available. In Section 2, we discuss a sweepout method to identify extreme individual sample values from composite sample measurements. We illustrate this method with data on PCB concentrations in surface soil samples. In Section 3, we discuss the situation where composites are formed using two orthogonal contours, and composite samples are formed along each of the contours. In Section 4, we consider the method of ranked set sampling as a means of improving the composite sampling procedure. We evaluate the performance of the sweepout method when the ranked set sampling protocol is used to form composite samples. The Armagh site data is used to illustrate the methods.

1.1 Composite Sampling

A composite sample is formed by mixing several individual samples or subsamples. The terminology of a "sample" as used here refers to a physical sample rather than to a statistical sample. For instance, an individual sample is a single grab collected from the sampling location selected for characterization, evaluation, or monitoring. Similarly, a composite sample is a mix of subsamples drawn from several individual samples. The strength of composite sampling procedures lies mainly in the physical averaging that occurs due to homogenization of the sample material while forming the composites. Compositing, at least under ideal conditions, incurs no loss of information for estimating population means. However, the loss of information regarding individual sample values, particularly the extreme values, has been an important limitation of the method. The available choices have been either to exhaustively measure all individual samples, or to lose information on extreme individual sample values.

In section 2, we present a statistical method to recover extremely large individual sample values using composite sample measurements and a few additional measurements on carefully selected individual samples. Using available composite sample measurements, this method first identifies constituent individual samples that may potentially have large values. Obtaining measurements on these few individual samples helps recover extremely large individual sample values. The method is illustrated with data on polychlorinated biphenyl (PCB) concentration in surface soil samples at the Armagh compressor station along the gas pipeline of the Texas Eastern Gas Pipeline Company in Pennsylvania (see Ref. [1]). Reference [2] consider this problem in the context of water quality monitoring, where the maximum pollutant concentration is either an observed value or estimated from other measurements. Noting that the cost of extensive and comprehensive monitoring is prohibitively high, Ref. [2] further note that no method exists that will find the maximum concentration with certainty unless continuous monitoring is used. They use the following assumptions in their method:

1. The process of collecting samples is distinct from their measurement;
2. The cost of sample measurement is high relative to that of collection; and
3. The sample values have high positive autocorrelation.

The method of Ref. [2] first identifies the composite sample having the highest measured value and then makes measurements on all the individual samples that form this composite. The highest observed individual sample value is taken to be the predicted value of the overall sample maximum. Assuming that compositing was done along the time component, this method is based on the premise that in the presence of high positive autocorrelation, the maximum sample value will tend to appear in the composite with the highest measurement. The number, and thus the cost, of tests performed in this method is a constant and is known prior to laboratory analysis.

Under the assumption of no measurement error, Ref. [1] propose an alternative method that is certain to identify the individual sample having the largest value without measuring all individual samples.

The performance of the sweepout method of Section 2 is affected by the compositing design. In an extreme case, if composites are formed with heterogeneous individual samples, then the sweepout method may perform worse than exhaustive measurement of all the individual samples. On the other hand, if composites are internally homogeneous, then the sweepout method can be very cost effective. It then remains to determine how one forms internally homogeneous composites. Four alternatives are discussed in this paper. First, if a spatial process is known to be operative on the site to be sampled, then locational information on sampling locations may be used to form reasonably homogeneous composites. We call this situation location-based compositing. Second, if two orthogonal contours are known to exist on the site, then homogeneity may be achieved by compositing along each of the contours. In this method, every individual sample contributes to exactly two composites, and hence this compositing design is different from other designs. Third, if information is available on locations with high values, as is common with fishing activities, sampling may be concentrated only on locations that may yield high values. In this case, all the individual samples are expected to return high values, and hence composites formed from these samples are expected to be homogeneous. Finally, if sampling units can be compared without exact quantification, then selected sampling units can be grouped and ranked, so that composites of sampling units that are assigned matching ranks will be relatively homogeneous. This rank-based compositing design is expected to enhance the performance of the sweepout method of Section 2. Each of the composite designs is illustrated with data on PCB in surface soil at the Armagh Compressor Station of the Texas Eastern Gas Pipeline Company. (See Ref. [3,4]).

1.2 The Armagh Site

Location and Features. The Armagh compressor station is located in West Wheatfield Township, Indiana County, PA. The site includes one compressor building along with several other buildings on 79 acres. There are two known liquid pits. There is one wetland situated within one-half mile of the site. Richard Run, which flows to the south of the site, is classified as a cold water fishery. There are no public recreational facilities near the station. Onsite soils are defined as being within the confines of the station site fencing and are accessible only to Texas Eastern personnel and authorized site visitors.

Onsite Surface Soil Sampling. Potential sources of PCB had been identified and a rectangular grid was laid out around each such source. Four different onsite grids were identified by the alphabetic codes "A" through "D". Grid points were identified by a two-digit row number and an alphabetic column code. Sampling of the surface soil was done at selected grid points in two distinct phases. Grid "D" was not sampled during Phase I, and as such is not included in the illustration here.

The distance between consecutive rows as well as between consecutive columns was 25 feet. Soil samples were taken from a 0-inch to 6-inch depth. After removing vegetation, rocks, and other debris, the sample at each grid point was thoroughly mixed to obtain a homogeneous sample for analysis and quantification.

2 Sweepout Method to Identify Extreme Sample Values

Let $x_1, x_2, \ldots, x_k$ denote the individual sample values and let y be the composite sample measurement. Further, let $x_{(k)}$ denote the maximum of the k individual sample values. That is,

$$x_{(k)} = \max\{x_1, x_2, \ldots, x_k\}.$$

Observe that

$$y \leq x_{(k)} \leq ky.$$

This inequality implies that the measurement on a composite sample gives bounds for the largest constituent individual sample value.

Now consider two composite samples of sizes k_1 and k_2, with measurements y_1 and y_2, and having the largest individual sample values $x_{(k_1)}$ and $x_{(k_2)}$, respectively. Without loss of generality, suppose that $y_1 < y_2$. In general, this does not allow for comparison between $x_{(k_1)}$ and $x_{(k_2)}$. However, if $k_1 y_1 < y_2$, then $x_{(k_1)} < x_{(k_2)}$ and hence it is no longer necessary to consider the first composite sample when the individual sample having the largest value is of interest. In this way, a number of composites can be eliminated without any additional testing. This elimination process leads us to fewer composites which may possibly contain individual samples that have large values. Measurements on individual samples in these composites then help identify the individual sample having the largest value.

Using this reasoning for identifying the largest individual measurement we obtain the sweepout method of Ref. [5] as follows:

1. Identify the composite sample having the largest measurement, say $y_{\max}$, and of size $k_{\max}$.

2. Measure all the individual samples in this composite and identify the largest individual measurement, say $x_{\max}$.

3. Consider the next largest composite measurement, say y^*, on a composite sample of size k^*. If $x_{\max} > k^* y^*$ then $x_{\max}$ is the largest individual measurement and the search is stopped.

4. If $x_{\max} \leq k^* y^*$, measure every individual sample in this composite and identify its largest individual measurement, say x^*.

5. If $x_{\max} < x^*$ then $x_{\max}$ is redefined and assigned this new largest value x^*. Repeat from (3) until the largest individual measurement is identified.

Reference [5] illustrate the sweepout procedure with an application to simulated composite sample values of PCB concentrations in surface soil samples at the Armagh Compressor Station. Table 1 shows the individual sample values and simulated composite sample measurements. In this illustration, Gore and Patil found that only 8 additional measurements on individual samples were required to identify the largest individual sample value from among a total of 358 individual samples. There were 90 measurements already made on composite samples.

Figure 1 shows a scatterplot of individual sample values plotted against the simulated composite sample measurements. The two rays through the origin indicate the bounds on the largest individual sample values. Since 4897.5 ppm is the largest composite sample measurement (composite number 25 in Table 1), constituent individual samples of this composite are measured separately. This identifies an individual sample with a PCB concentration of 10000 ppm. A horizontal line at the height of 10000 ppm indicates that there is only one more composite (composite number 5 in Table 1) which can possibly contain an individual sample with a PCB concentration of more than 10000 ppm. Making measurements on all the individual samples constituting this composite identifies an individual sample with a PCB concentration of 10700 ppm. There is no other composite that can contain an individual sample with a PCB concentration exceeding 10700 ppm, as is evident from Figure 1 (b). Thus, making measurements on 8 individual samples constituting two composites has identified the individual sample with the largest PCB concentration.

2.1 Extensive Search of Extreme Values

The sweepout method described above can be examined further for its cost effectiveness in identifying extreme values. Note that exhaustive testing of all individual samples (without compositing) identifies all individual values. In this case,

Table 1: Individual sample values and simulated composite sample measurements.

Composite Sample	Individual Sample Values	Composite Sample Measurement	Composite Sample	Individual Sample Values	Composite Sample Measurement
01	2.9, 3.1, 22, 22	12.5	46	1.9, 1.6, 82, 390	118.9
02	21, 298, 18, 1880	554.3	47	1.4, 1, 530, 320	213.1
03	9.4, 51, 319, 1.0	95.1	48	160, 180, 19, 320	169.8
04	105, 30, 22, 67	56.0	49	5.4, 1.7, 0.0, 15	5.5
05	18, 2320, 10700, 2960	3999.5	50	7.7, 6.9, 310, 19	85.9
06	38, 2.5, 13, 154	51.9	51	27, 23, 21, 5	19
07	1.1, 12, 55, 8.7	19.2	52	7.5, 2.2, 55, 80	36.2
08	13, 1.9, 2.9, 22	10.0	53	7.7, 4.3, 24, 250	71.5
09	129, 12, 44, 22	51.8	54	4.3, 6.4, 20, 33	15.9
10	1.6, 1070, 1.0, 64	284.2	55	436, 9.5, 120. 21, 58	128.9
11	13, 3.8, 3, 6.8	6.9	56	1.5, 160, 180, 1000	335.4
12	13, 3.8, 2.8, 6.9	6.1	57	2.9, 15, 150, 12, 11	38.2
13	34, 28, 745, 3850	1164.3	58	2.9, 26, 1.2, 1.3	7.9
14	50, 18, 17, 34	29.8	59	24, 2.6, 3.5, 18	12.0
15	4.6, 22, 1.0, 42	17.4	60	3.9, 27, 5.4, 12	12.1
16	14, 3.3, 1.5, 2.6	5.4	61	72, 38, 7.1, 35	38.0
17	2.4, 1390, 3, 672	516.9	62	52, 37, 66, 38	48.3
18	8.9, 661, 20, 18	177.0	63	1.3, 2.1, 15, 4.4	5.7
19	18, 24, 26	22.7	64	60, 79, 8.7, 150	74.4
20	3.5, 16, 20	13.2	65	16, 24, 18, 160	54.5
21	97, 70, 14, 150	82.8	66	150, 210, 18, 13	97.8
22	37, 72, 40, 33	45.5	67	26, 7.8, 43, 49	31.5
23	38, 44, 83, 30	48.8	68	46, 24, 18, 12	25
24	38, 100, 140, 47	81.3	69	38, 12, 140, 60	62.5
25	590, 7100, 10000, 1900	4897.5	70	26, 14, 190, 61,33	64.8
26	670, 940, 240, 290	535	71	340, 190, 10	180
27	74, 200, 120, 220	153.5	72	0.0, 0.0, 0.0, 0.0	0.0
28	280, 260, 10, 250	200	73	0.0, 0.0, 0.0, 1.1	0.3
29	44, 110, 660, 230	261	74	1.1, 2.8, 4.2, 6.6	3.7
30	580, 1100, 1300, 4900	1970	75	6.9, 16, 7, 13	10.8
31	110, 80, 210, 12	103	76	11, 13, 6.4, 8	9.6
32	75, 890, 170, 550	421.3	77	0, 236, 7.2, 2.4	61.4
33	2300, 420, 520, 1300	1135	78	5.8, 535, 1.1, 0.0	135.5
34	0.0, 1.2, 1.67	2.5	79	0.0, 1.4, 4.9, 0.0	1.6
35	5.7, 17, 4.3, 36	15.8	80	0.0, 0.0, 5.1, 6.3	2.9
36	28, 170, 10, 62	67.5	81	7.9, 14, 20, 31	18.2
37	300, 6.4, 53	119.8	82	52, 1, 500, 46	162.3
38	16, 18, 150, 27	52.8	83	16, 5, 36, 64	30.3
39	6.2, 7.1, 31, 38	20.6	84	40, 38, 68, 7.5	38.4
40	16, 66, 61, 340, 1500	396.6	85	40, 33, 36, 17	31.5
41	1.3, 3.5, 2.1, 8.8	3.9	86	35, 4, 170	52.3
42	7.5, 2.7, 1.6, 11	5.7	87	110, 200, 4.2	104.7
43	0.0, 0.0, 17, 2.8	5.0	88	7.4, 3.3, 21, 2.3	8.5
44	1.1, 5.9, 350, 17	93.5	89	3.8, 35, 20, 17	19.0
45	3.2, 5, 11, 5.1	6.1	90	23, 17, 3, 6.8	12.5

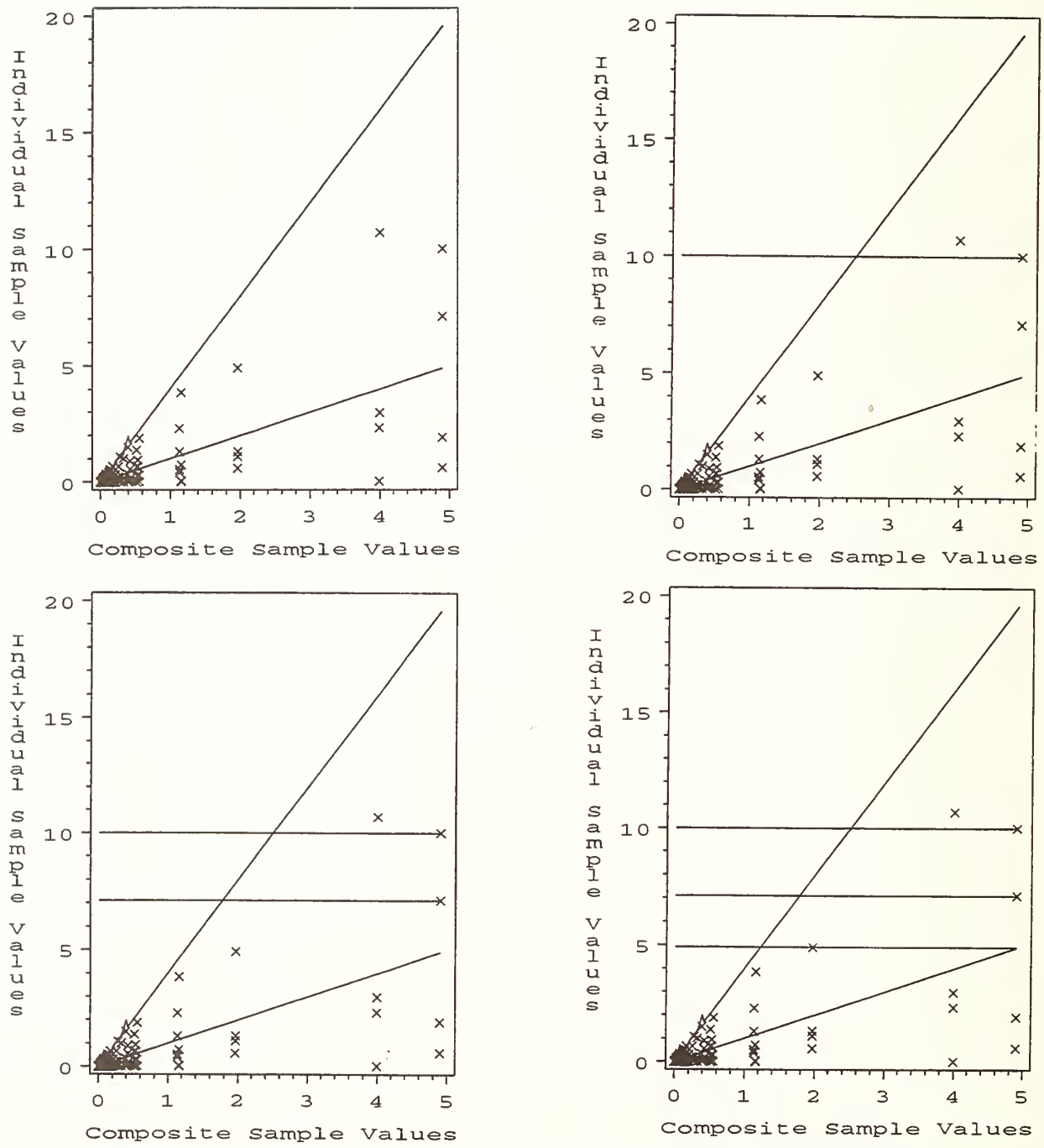


Figure 1: Illustration of the sweepout method. Individual sample values (Y axis) vs composite sample measurements (X axis) in thousand ppm. (a) The upper and lower bounds for the largest individual values. (b) Measurements on individual samples from only two composites identify two largest individual values; (c) Measurements on individual samples from only two composites identify the three largest individual values. (d) Measurements on individual samples from three composites identify the four largest individual values.

identification of extreme values is achieved simply by arranging the individual values in a descending order of magnitude. Thus, the method of exhaustive testing involves as many measurements as the number of individual samples. For example, the case of the Armagh site would require 358 measurements.

In order to investigate the relationship between the number of extreme values identified and the number of measurements made, we extend the sweepout method to all the 90 composites at the Armagh site. Figure 2 gives a graphical summary of these results. The concavity of the curve implies that identification of every additional extreme value initially requires relatively more measurements.

As a consequence of its ability to identify individual samples having large values, the sweepout method can also provide estimates of upper percentiles of the distribution of individual sample values. Reference [5] discuss the applicability of this feature of the sweepout method to compliance monitoring and to quality assurance management. See Ref. [6] and [7] for more details.

3 Two-dimensional Compositing Design

The Sweepout method discussed above assumed that every individual sample contributes to exactly one composite sample. If each individual sample is allowed to contribute to more than one composite, it is possible to expedite the search for extreme individual sample values using composite sample measurements. For instance, arranging the individual samples in a rectangular array allows us to form composites by combining all the individual samples in each row to form row-composites, and all the individual samples in each column to form column-composites. In this way, a row-column arrangement implies that every individual sample contributes to exactly two composites, a row-composite and a column-composite.

For example, consider the situation where 64 individual samples are available, and the problem is to identify the largest individual sample value using a minimum number of tests. One possible method is to form 16 composites, each of size 4, make 16 measurements on the composite samples, and then apply the sweepout method. Four additional measurements will be made for every composite sample selected for retesting. According to the row-column arrangement described above, these 64 individual samples could be arranged into a square having 8 rows and 8 columns. This arrangement will produce 8 row-composites each of size 8 and 8 column-composites each of size 8. Thus there will be 16 measurements on the 16 composite samples. Suppose $\{X_{ij},\ i = 1,\ldots,8;\ j = 1,\ldots,8\}$ denote the 64 individual sample values; $\{Y_{i\cdot}, i = 1,\ldots,8\}$ denote the 8 row-composite measurements; and $\{Y_{\cdot j}, j = 1,\ldots,8\}$ denote the 8 column-composite measurements. Suppose for some i^*, $Y_{i^*\cdot}$ is the largest row-composite measurement, and for some j^*, $Y_{\cdot j^*}$ is the largest column-composite measurement. Then the individual sample in the i^*th row and the j^*th column is likely to have the largest value. Therefore, the value of $X_{i^*j^*}$ is determined by making a measurement on this individual sample. Having known the value of $X_{i^*j^*}$, it is then possible to determine whether any other individual sample is likely to exceed $X_{i^*j^*}$ by comparing the measurements on the other row-composites and column-composites with $X_{i^*j^*/8}$. Note that every retesting stage now involves only one additional measurement, as opposed to four additional measurements in the case of the linear sweepout method of the preceding section.

We illustrate this procedure with an application to the Armagh site. For this purpose, we consider only part of the data given in Table 1, so that we have a square of 8 rows and 8 columns. Thus, selecting only the grid points from Grid A between rows 50 and 57, columns 8 and 15, we form 8 composites.

As can be observed from Table 2, the row-composite of row 54 has the largest measurement, while among the column-composites, col-

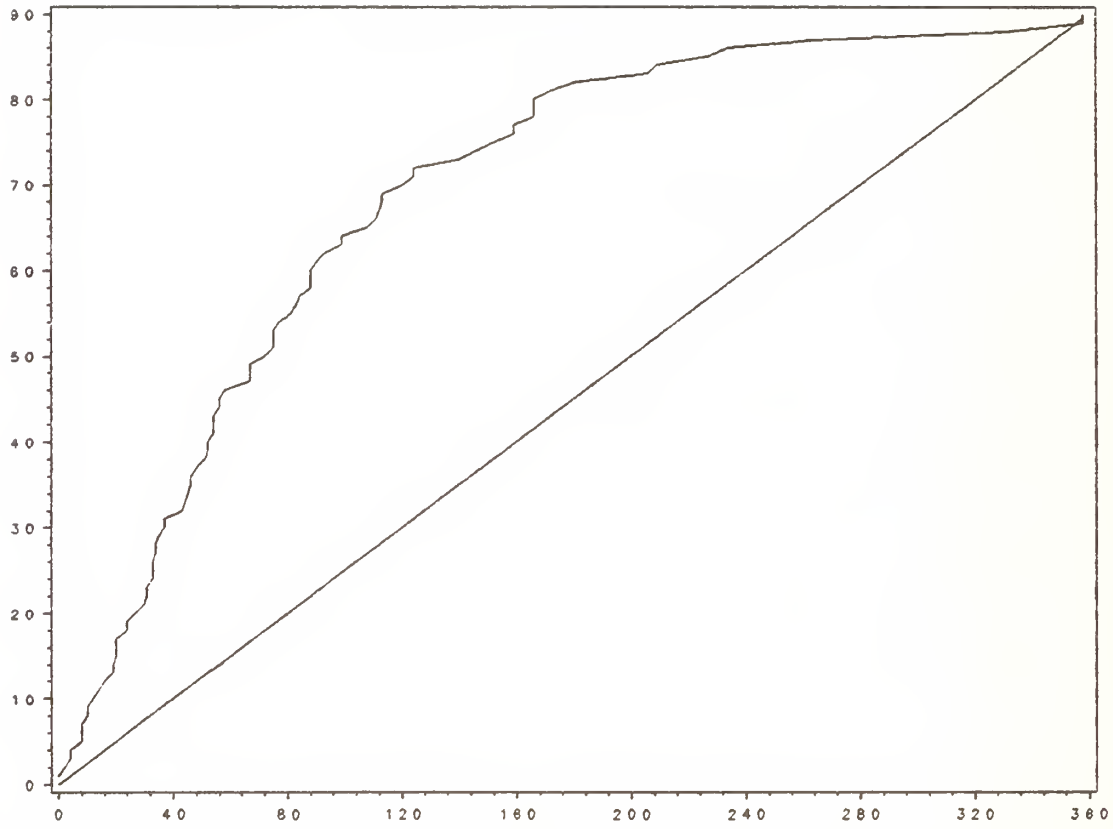


Figure 2: Number of composites retested (Y axis) vs number of extreme values identified (X axis). The diagonal line represents the optimal case in which exactly 4 extreme values are identified for every composite.

Table 2: Row and column arrangement of sampling units for compositing. Here, every individual sampling unit contributes to exactly two composites.

	Column									
Row	8	9	10	11	12	13	14	15	$Y_{i\cdot}$	$k_i Y_{i\cdot}$
50	3.0	2.9	3.1	1.9	1.6	1.4	1.0		2.129	14.9
51	16	22	22	82		530		160	138.7	832
52	20	21	298	9.4	390	320	19	320	174.7	1397
53	18	18	1880	51	319	105	18	2320	591	4729
54	24		34	1		30	10700	2960	2291	13749
55	26	28	745	3850		22	38	2	673	4711
56	50	18		11		67	13	154	89	313
57	17	34		3.8		1.1	12	55	20.48	122.9
$Y_{\cdot j}$	21.75	20.56	497	501	237	134.6	1543	785		
$k_j Y_{\cdot j}$	174	144	2982	4010	711	1077	10801	5971		

umn 14 gives the largest measurement. The proposed sweepout method suggests that the individual sample in row 54 and column 14 be measured. This gives a value of 10700 for this particular individual sample. It is then compared with the upper bounds formed from the remaining composite measurements. Since there is no other composite sample with an upper bound that exceeds this value, we conclude that 10700 is indeed the largest individual sample value among the 55 individual samples involved in the above illustration. Note that it took only one measurement on an individual sample to identify the largest individual sample value. Thus, the total number of measurements required for identification of the largest individual sample in this example is 17, with 16 measurements for the 16 composite samples and one measurement for the individual sample in row 54 and column 14.

4 Compositing a Ranked Set Sample

For predetermined positive integers m and r, ranked set sampling (RSS) involves selection and acquisition of m^2r units, of which only mr units are quantified. First, m random samples, each of size m, are randomly selected from a population (with distribution function F, mean μ, variance σ^2, say). The m selected units in each sample are ranked by a judgement process such as visual inspection or any other inexpensive method which does not require actual measurement. The unit with the smallest rank is quantified from the first sample, the unit having the second smallest rank is quantified from the second sample, and so on, until the unit with the highest rank is quantified from the *m-th* sample. Thus, m units are quantified out of the m^2 units originally selected. The process is repeated r times, thereby providing a total of mr measurements which constitute the ranked set sample. Reference [8] and [9] provide mathematical formulation for this sampling method which was introduced earlier by Ref. [10] as an improvement in simple random sampling (SRS) for estimation of mean pasture and forage yields. This method is particularly attractive when quantification of units is difficult or expensive, but ranking of a small set of units can be done with a reasonable accuracy even without making measurements.

RSS may be utilized advantageously for forming internally homogeneous composites as compared to those based on random groupings. With m samples of size m, we can form m composites by physically mixing sampling units on the basis of their ranks. Likewise, we get mr composite samples of size m from m^2r units. These samples, in turn, provide mr measurements. The standard deviation of these measurements is expected to be smaller than that of the same number of measurements obtained from composites comprising sampling units selected randomly, m at a time, out of m^2r available units in most cases. For example, in the case of 64 sampling units, 16 sets of size 4 are formed for the purpose of ranking. These 16 sets are tabulated in Table 3.

The graph in Figure 3 shows the number of measurements (on the Y axis) versus the number of extreme values identified (on the X axis). Here, the data set of Table 2 is used with three different compositing designs. First, the 64 sampling units were grouped in 16 sets of size 4 each based on the contiguity of their locations. Thus, every 2×2 square within the 8×8 arrangement of the 64 sampling units is used to identify the individual sampling units for compositing. We call this design "contiguity-based" compositing. Next, the 8 rows were used to form 8 row-composites, and similarly for 8 columns. We call this design "row-column" compositing. Finally, forming 16 sets of size 4 each, we rank the 4 sampling units within each set, then form 16 composites from sampling units that were assigned the same ranks. In other words, the compositing is based on contiguity of ranks, rather than locational contiguity, as in the first case. We call this last design "rank-based" compositing. It is easily seen that the rank-based compositing has performed better than the other two compositing designs. The difference between the three graphs is due only to the compositing

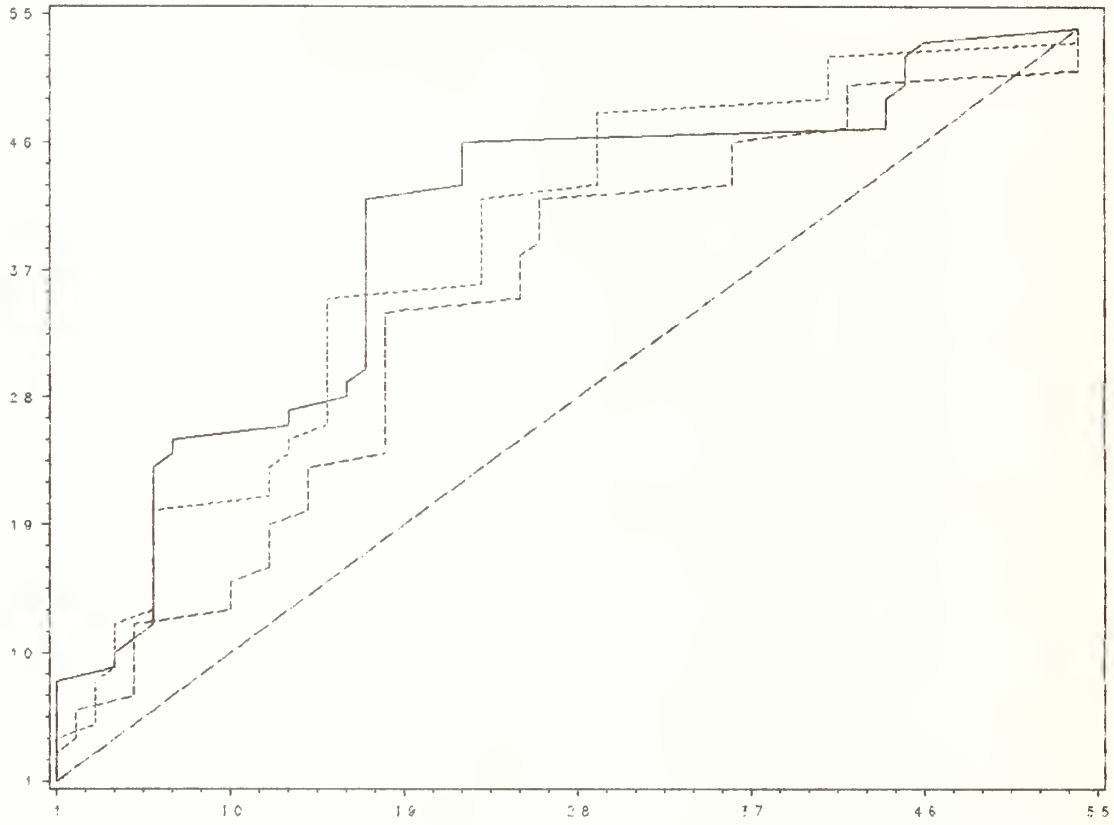


Figure 3: Number of measurements vs. Number of extreme values identified. solid line: row-column compositing; dotted line: contiguity-based compositing; and broken lines: rank-based compositing

design, since the same 64 sampling units are used in the illustration. For additional references, see Ref. [11-21].

5 In Conclusion

This paper highlights the need for an investigation of the statistical properties of the sweepout method. The performance of the sweepout method will be determined by several factors, including the autocorrelation structure among the individual sample values, the statistical distribution of these individual sample values, and the compositing plan as well as the composite sample size.

6 Acknowledgements

This paper has been prepared with partial support from the United States Environmental Protection Agency Grant Number CR-821531. The contents have not been subjected to Agency review and therefore do not necessarily reflect the views or policies of the Agency and no official endorsement should be inferred.

References

[1] Gore, S. D., Patil, G. P., and Taillie, C., Studies on the applications of composite sample techniques in hazardous waste site characterization and evaluation: II. Onsite surface soil sampling for PCB at the Armagh Site, Technical Report No. 92-0305, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, University Park, PA. 16802, 1992.

[2] Casey, D., Nemetz, P. N., and Uyeno, D., Efficient search procedures for extreme pollutant values, Environmental Monitoring and Assessment, 5 (1985), 165-176.

Table 3: Sets of sampling units for the purpose of ranking. Sampling units that are assigned the same rank are then composited together to achieve internal homogeneity of composite samples.

Set	Unit 1	Unit 2	Unit 3	Unit 4
1	2.9	18	24	17
2	3	18	26	18
3	16	20	28	34
4	22	21	*	50
5	1.9	9.4	1.0	3.8
6	3.1	51	34	11
7	22	298	745	*
8	82	1880	3850	*
9	1.4	105	22	1.1
10	1.6	319	30	67
11	530	320	*	*
12	*	390	*	*
13	1	18	2	12
14	160	19	38	13
15	*	320	2960	55
16	*	2320	10700	154

*: no sample collected from this location

[3] Texas Eastern Gas Pipeline Company, Results of the Phase II surface soil and sediment sampling activities at the Armagh site, Pennsylvania, Vol. I, Roy F. Weston, Inc., West Chester, PA, 1989a.

[4] Texas Eastern Gas Pipeline Company, Results of the Phase II surface soil and sediment sampling activities at the Armagh site, Pennsylvania, Vol. II: Appendices, Roy F. Weston, Inc., West Chester, PA, 1989b.

[5] Gore, S. D. and Patil, G. P., Identifying extremely large values using composite sample data, Environmetrics, 1 (1994), (To appear).

[6] Kahn, H., Discussion on "Identifying extremely large values using composite sample data," Environmetrics, (1994), (To appear).

[7] Warren, J., Discussion on "Identifying extremely large values using composite sample data," Journal of Environmental and Ecological Statistics, 1994, (To appear).

[8] Takahasi, K., and Wakimoto, K., On unbiased estimates of the population mean based on the sample stratified by means of ordering, Ann. Inst. Statist. Math., 20 (1968), 1-31.

[9] Dell, T. R. and Clutter, J. L., Ranked set sampling theory with order statistics background, Biometrics, 28 (1972), 545-555.

[10] McIntyre, G. A., A method of unbiased selective sampling, using ranked sets, Australian J. Agricultural Research, 3 (1952), 385-390.

[11] David, H. A., Concomitants of order statistics, Bull. Int. Statist. Inst., 45 (1973), 295-300.

[12] David, H. A., Concomitants of order statistics: Theory and applications, In: Some Recent Advances in Statistics (eds.: J. Tiago de Oliveira and B. Epstein), Academic Press, New York, 1982, pp. 89-100.

[13] Halls, L. K. and Dell, T. R., Trial of ranked set sampling for forage yields, J. Forest Science, 12 (1966), 22-66.

[14] Huber, P. J., Robust statistics: A review. (The 1972 Wald Lecture), Annals of Mathematical Statistics, 42 (1972), 1041-1067.

[15] Martin, W. L., Sharik, T. L., Oderwald, R. G., and Smith, D. W., Evaluation of ranked set sampling for estimating shrub phytomass in Appalachian oak forests, Publication Number FWS-4-80, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1980.

[16] Patil, G. P., Sinha, A. K., and Taillie, C., Ranked set sampling, In: Handbook of Statistics Volume 12: Environmental Statistics (eds.: G. P. Patil and C. R. Rao), North Holland/Elsevier Science Publishers, New York, Amsterdam, 1994, pp. 167-200.

[17] Patil, G. P. and Taillie, C., Performance of the largest order statistics relative to the sample mean for the purpose of estimating a population mean, Bull. Int. Statist. Inst., 54 (1991).

[18] Patil, G. P. and Taillie, C., Environmental sampling, observational economy, and statistical inference with emphasis on ranked set sampling, encounter sampling, and composite sampling, Bulletin of the International Statistical Institute, Proceedings of the 49th Session, (1993), 295-312.

[19] Sarhan, A. E. and Greenberg, B. G., Contributions to Order Statistics, Wiley, New York, 1962.

[20] Stokes, S. L. and Sager, T., Characterization of a ranked set sample with application to estimating distribution functions, Journal of Applied Probability, 83 (1988), 374-381.

[21] Watterson, G. A., Linear estimation in censored samples from multivariate normal populations, Annals of Mathematical Statistics, 30 (1959), 814-824.

Extremal Sojourn Times For Markov Chains

Arnold, B. C.
University of California, Riverside, CA

Consider a continuous time Markov chain with state space $\{1,2,3,\ldots\}$. Since sojourns in particular states are independent exponential random variables, it is possible to derive the asymptotic distribution of the maximal and minimal sojourn in a particular state or in any state. Discrete time analogies are described and the more challenging problem of deriving the distribution of the extreme sojourn times in a particular group of states in discrete time is introduced.

1 Extremal sojourns

Consider $X(t)$ a continuous time Markov chain with state space $\{1,2,\ldots\}$. Assume that the chain is irreducible and ergodic with intensity matrix Q satisfying

$$\underline{Q}\underline{1} = \underline{0} \tag{1}$$

and denote the long run distribution by $\underline{\pi}$. For a particular state i, sojourns in that state are i.i.d. exponential $(-q_{ii})$ random variables. During a time interval $(0,T]$, a particular state i will be visited a random number of times.

If we let $N_i(T)$ denote the number of visits to state i, it is easy to verify (using results for delayed recurrent events) that

$$N_i(T)/T \xrightarrow{a.s.} (-q_{ii})\pi_i \ . \tag{2}$$

For more detail see Ref. [1] where the finite state space case is discussed.

If we use $X_j^{(i)}$ to denote the jth sojourn in state i, then our interest is in studying the asymptotic behavior of extremal sojourn times defined as follows.

$$M_i(T) = \max_{j \le N_i(T)} X_j^{(i)}$$

$$m_i(T) = \min_{j \le N_i(T)} X_j^{(i)}$$

$$M(T) = \max_i M_i(T)$$

and

$$m(T) = \min_i m_i(T) \ .$$

To determine the asymptotic distribution of these random variables, we need a result of Barndorff-Neilson (Ref. [2]) (which may be conveniently found as part of Theorem 6.2.1 in Ref. [3]). It states that if $X_1, X_2, \ldots$ are i.i.d. in the domain of maximal attraction of some distribution Λ and if $N(T)/T \xrightarrow{P} \delta$ then $\max_{j \le N(T)} X_j$ has the same asymptotic distribution (with the same normalizing constants) as does $\max_{j \le \delta T} X_j$. We also need the observation made in Ref [1] regarding the maximum of heterogenous exponential random variables. The result in question deals with two independent sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$. The X_i's are exponential(λ_1) and the Y_i's are exponential(λ_2) where $\lambda_1 < \lambda_2$. For $\alpha \epsilon (0,1)$ we have

$$\max[\max_{i \le \alpha n} X_i, \max_{i \le (1-\alpha)n} Y_i] \approx \max_{i \le \alpha n} X_i$$

where $\approx$ denotes "has the same asymptotic distribution". Finally we recall that the minimum of m arbitrary exponential random variables is again exponential with a parameter obtained by adding the parameters of the individual exponential variables.

These observations allow us to state the following results for the maximal and minimal sojourn in a particular state.

Theorem 1: For any state i

$$\lim_{T \to \infty} P((-q_{ii})M_i(T) - \log(-q_{ii}\pi_{ii}) - logT \le z)$$

$$= exp(-e^{-z}), \quad z \epsilon \mathbf{R}$$

and

$$\lim_{T \to \infty} P((-q_{ii})^2 \pi_i T m_i(T) \le z) = 1 - e^{-z},$$

$$z > 0 \ .$$

To obtain limiting distributions for extremal sojourns in any state, some minor regularity conditions must be imposed. We assume there exists a smallest $(-q_{jj})$ and reorder the states so that $-q_{11} \le -q_{22} \le \ldots$. With this convention we find

Theorem 2: (A) If for some integer k and some $\epsilon > 0$ we have

$$(-q_{11}) = (-q_{22}) = \ldots = (-q_{kk}) \le (-q_{jj}) - \varepsilon,$$

$$\forall j > k$$

then

$$\lim_{T \to \infty} P((-q_{11})M(T) - log(-q_{11} \sum_{i=1}^{k} \pi_i) - \log T \le z)$$

$$= \exp(-e^{-z}), \quad z \epsilon \mathbf{R} \ .$$

(B) If the series $\sum_{i=1}^{\infty} (-q_{ii})^2 \pi_i$ is convergent, then

$$\lim_{T \to \infty} P([\sum_{i=1}^{\infty} (-q_{ii})^2 \pi_i] T m(T) \le z) = 1 - e^{-z},$$

$$z > 0 \ .$$

2 Matching chains

A more interesting problem involves the study of two independent Markov processes $X(t), Y(t)$ and identifying the largest

time interval during which $X(t) = Y(t)$. In continuous time this is a trivial extension of the material in the last section. We can identify $\{(i_1, i_2) : i_1 = 1, 2, \ldots, i_2 = 1, 2, \ldots\}$ as the state space of the process $\{X(t), Y(t)\}$ and we are interested first in the maximal exponential sojourn in the state (i, i) for each i and then in the maximum of these maxima. As we shall remark in the next section, the study of long matches between discrete time Markov chains is considerably more complicated.

3 Discrete time

If we study a discrete time Markov chain with state space $\{1, 2, 3, \ldots\}$ then sojourns in a particular state are i.i.d. geometric random variables. Maxima of i.i.d. geometric random variables cannot be normalized to converge in distribution (see e.g. Ref. [4]). However we can get useful approximations by using the following observation. If X is geometric(p) (i.e. $P(X = k) = p(1-p)^{k-1}$, $k = 1, 2, \ldots$) then we may introduce an exponential $(-\log(1-p))$ random variable whose integer part is identically distributed with X and consequently we have

$$W \leq X \leq W + 1$$

and we can bound maxima of X's by maxima of W's. Ref. [1] provides details in the finite state space case. Only minor regularity conditions (analogous to those in Theorem 2 above) are required to extend the results to the infinite state space case.

Now, what about long matches between two independent discrete time Markov chains? Again we may combine the two chains X_n and Y_n to form (X_n, Y_n) with state space $\{1, 2, \ldots\} \times \{1, 2, \ldots\}$. However, now a match occurs when the two dimensional chain remains in the class of states $(1, 1), (2, 2), \ldots$. Sojourns in such classes of states have more complicated distributions than sojourns in particular states (which are geometrically distributed). Exceptions occur if the chain is "lumpable" without upsetting the Markov property, but most interesting examples do not have this property.

REFERENCES

[1] Arnold, B.C. and Villaseñor, J.A. "The distribution of the maximal time to departure from a state in a Markov chain", in *Statistical Extremes and Applications*, J. Tiago de Oliveira (ed). Reidel, Dordrecht, 1984, pp 413-426.

[2] Barndorff-Nielsen, O. "On the limit distribution of the maximum of a random number of independent random variables". *Acta. Math. Acad. Sci. Hungar.* 15, 1964, 399-403.

[3] Galambos, J. *The asymptotic theory of extreme order statistics.* Wiley, New York, 1978.

[4] Anderson, C.W. "Extreme value theory for a class of discrete distributions with applications to some stochastic processes". *Journal of Applied Probability* 7, 1970, 99-113.

Bootstrapping Extremes Of I.I.D. Random Variables

Athreya, K.B. and Fukuchi, J.
Iowa State University, Ames, IA

Let $X_1, X_2, \cdots$ be i.i.d. random variables with common distribution function F. Define $X_{n:n} \equiv \max(X_1, X_2, \ldots X_n)$. Assume that there exist $a_n > 0, b_n \in \mathbf{R}, n \geq 1$ such that $G_n(x) \equiv P\{a_n(X_{n:n} - b_n) \leq x\}$ converges to one of Gnedenko's extreme value distributions. In this paper the problem of estimating $G_n(x)$ by the bootstrap technique is considered. We define different bootstrap distributions for different types of domain of attraction that F belongs to. It is shown that both when a_n and b_n are known and when a_n and b_n are estimated from the data the bootstrap distribution is weakly consistent if m=o(n) and it is strongly consistent if m=$o(\frac{n}{\log n})$. These results are applied to the problem of obtaining confidence intervals for the upper end point of the support of F.

1 Introduction

Since Efron (Ref. [1]) introduced the bootstrap method of approximating sampling distributions of statistics, many papers have investigated its asymptotic properties. One of the desired properties of this method is the consistency, namely, the limit of the bootstrap distribution is the same as that of the distribution of the original statistic. Examples of the situations where Efron's bootstrap (the simple random sampling from the original data) fail to be consistent are, among others, the sample mean of heavy tailed random variables (Ref. [2], [3]), the sample mean of weakly dependent random variables (Ref. [4], [5]), normalized maximum of i.i.d. random variables (Ref. [6]).

We study asymptotic properties of bootstrap for the distribution of normalized extremes when the underlying distribution belongs to the domain of attraction of an extreme value distribution.

[1]Research supported in part by NSF Grant DMS 92-04938, 1991 Mathematics subject classification 62G05 62G30
Keywords and phrases: extremes, bootstrap.

In Section 2, we investigate the inconsistency, weak consistency and strong consistency of bootstrappng $a_n(X_{n:n} - b_n)$ with appropriate choice of resample size m=m(n) when a_n and b_n are known.

In Section 3, the same problem as in Section 2 are investigated when a_n and b_n are unknown. In Ref. [6], it was pointed out that the naive bootstrap of the maximum of uniform i.i.d. random variables with m=n fails to be consistent. In Ref. [7], it was shown that when F belongs to the domain of attraction (in the sense of extremes) of one of the extreme value distributions, the bootstrap distribution of maximum converges to a random distribution. Recently, Deheuvels, Mason and Shorack (Ref. [8]) proved the weak consistency and the strong consistency of bootstrap for the normalized maximum when normalizing constants are estimated. They utilized von Mises's parameterization of extreme distribution and thus their method does not need the knowledge about which type of domain F belongs to.

Results in this paper were obtained independently of Ref. [8] and our approach is different

from theirs. We define different bootstrap distributions for different types of domain of attraction, so in practice we need to know which type of domain F belongs to. But our version of the bootstrap distribution is more appropriate when inferences for population parameters such as obtaining confidence intervals are concerned.

2 Limits of bootstrap distributions : when normalizing constants are known

We begin with a review of basic results from extreme value theory. Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with a common distribution function F and $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ be the corresponding order statistics. Let $F^{-1}(u) := \inf\{x : F(x) \ge u\}$ be the left continuous inverse of F and C_G be the set of continuity points of a function G. We say that $F \in D(G)$ if there exist constants $a_n > 0, b_n \in R$ and nondegenerate distribution function G such that

$$P\{a_n^{-1}(X_{n:n} - b_n) \le x\} \to G(x) \quad \forall x \in C_G. \quad (1)$$

It is known (cf. Ref. [9]) that G must be one of the following types. (up to location and scale changes)

$$G(x) = \Lambda(x) = \exp(-e^{-x}) \quad x \in \mathbf{R},$$

$$G(x) = \Phi_\alpha(x) = \begin{cases} 0 & x \le 0 \\ \exp(-x^{-\alpha}) & x > 0, \end{cases}$$

$$G(x) = \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^{\alpha}) & x \le 0 \\ 1 & x > 0, \end{cases}$$

where $\alpha > 0$.
In (1), a_n, b_n can be chosen as follows: (cf. Ref. [9])

$$F \in D(\Lambda): \quad a_n = F^{-1}(1 - \frac{1}{en}) - \gamma_n, \quad b_n = \gamma_n,$$
$$F \in D(\Phi_\alpha): \quad a_n = \gamma_n, \quad b_n = 0 \quad ,$$
$$F \in D(\Psi_\alpha): \quad a_n = \theta_F - \gamma_n, \quad b_n = \theta_F,$$

where $\gamma_n := F^{-1}(1 - \frac{1}{n})$ and $\theta_F := \sup\{x : F(x) < 1\}$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X_1, X_2, \cdots$, be a sequence of i.i.d. random variables on $(\Omega, \mathcal{F})$ with a distribution F$\in D(G)$ where G=Λ or Φ_α or Ψ_α. Let $m = m(n) \in \mathbf{N}$ be such that $m(n) \to \infty$ as $n \to \infty$. Given $\mathbf{X}_n := (X_1, X_2, \cdots, X_n)$, let $Y_1, Y_2, \cdots Y_m$ be conditionally i.i.d. random variables with the distribution

$$P(Y_1 = X_j | \mathbf{X}_n) = \frac{1}{n}, \qquad j = 1, 2, \cdots, n,$$

i.e. $(Y_1, Y_2, \cdots, Y_m)$ is a i.i.d. resample of size m from the empirical distribution of $\mathbf{X}_n$. Let $Y_{1:m} \le Y_{2:m} \le \cdots \le Y_{m:m}$ be the corresponding order statistics. Now, define

$$\begin{aligned} G_n(x) &= P(a_n^{-1}(X_{n:n} - b_n) \le x), \\ H_{n,m}(x, \omega) &= P(a_m^{-1}(Y_{m:m} - b_m) \le x | \mathbf{X}_n), \end{aligned}$$

for $\omega \in \Omega$ and call $H_{n,m}(x, \omega)$ the bootstrap distribution of $a_n^{-1}(X_{n:n} - b_n)$. The first subscript n of H represents the original sample size and the second subscript m of H represents the resample size. The next theorem shows that if m=n, $H_{n,m}(x, \omega)$ has a random limit and thus the naive bootstrap fails to approximate $G_n(x)$. Let PRM(μ) denote a Poisson random measure with mean measure $\mu(\cdot)$.

Theorem 1 *Let $F \in D(G)$ where G is an extreme value distribution and let $a_n > 0$, b_n $n \ge 1$ be such that (1) holds. (i) If $G = \Lambda$, then for any $x_i \in \mathbf{R}$, $i = 1, \cdots, r$,*

$$\begin{aligned} &(H_{n,n}(x_i, \omega), i = 1, 2, \cdots, r) \\ \xrightarrow{d} \; &(H(x_i, \omega), i = 1, 2, \cdots, r), \end{aligned} \quad (2)$$

where

$$H(x, \omega) := \exp(-T((x, \infty), \omega)),$$

and $T(\cdot, \omega)$ is a PRM(μ) on $\mathcal{B}((\infty, \infty])$ ($\mathcal{B}((\infty, \infty])$ denotes the Borel σ-algebra on $(\infty, \infty]$.) with

$$\mu(B) = \int_B e^{-x}\, dx \quad \forall B \in \mathcal{B}((\infty, \infty]). \quad (3)$$

(ii) If $G = \Phi_\alpha$, then for any $x_i > 0$, $i = 1, \cdots, r$,

$$\begin{aligned} &(H_{n,n}(x_i, \omega), i = 1, 2, \cdots, r) \\ \xrightarrow{d} \; &(H(x_i, \omega), i = 1, 2, \cdots, r), \end{aligned} \quad (4)$$

where

$$H(x,\omega) := \exp(-T((x,\infty),\omega)),$$

and $T(\cdot,\omega)$ *is a PRM*(μ) *on* $\mathcal{B}((0,\infty])$ *with*

$$\mu(B) = \int_B x^{-\alpha}\,dx \quad \forall B \in \mathcal{B}((0,\infty]). \tag{5}$$

(iii) If $G = \Psi_\alpha$, *then for any* $x_i \le 0$, $i = 1,\cdots,r$,

$$\begin{aligned}&(H_{n,n}(x_i,\omega), i = 1,2,\cdots,r)\\ \xrightarrow{d}\ &(H(x_i,\omega), i = 1,2,\cdots,r),\end{aligned} \tag{6}$$

where

$$H(x,\omega) := \exp(-T((x,0],\omega)),$$

and $T(\cdot,\omega)$ *is a PRM*(μ) *on* $\mathcal{B}((-\infty,0])$ *with*

$$\mu(B) = \int_B (-x)^\alpha\,dx \quad \forall B \in \mathcal{B}((-\infty,0]). \tag{7}$$

Proof. We can write

$$\begin{aligned}H_{n,n}(x,\omega) &= P\{a_n^{-1}(Y_{n:n} - b_n) \le x | \mathbf{X}_n\}\\ &= F_n^n(a_n x + b_n)\\ &= \left\{1 - \frac{n(1 - F_n(a_n x + b_n))}{n}\right\}^n.\end{aligned}$$

For (i), define a point process T_n on $(-\infty,\infty]$ by

$$T_n = \sum_{k=1}^n \epsilon_{a_n^{-1}(X_k - b_n)},$$

where ϵ_a is the delta measure at a. Then, by corollary 4.19 of Ref. [10], T_n converges weakly to a $PRM(\mu)$ where μ is given by (3). Therefore the continuous mapping theorem gives the result. Proofs for (ii) and (iii) are similar. □

We note that the condition $m = n$ is not necessary for the above results to hold. Even if $m/n \to c$, $0 < c < \infty$, results similar to Theorem 1 hold.

Theorem 1 can be easily extended to the bootstrap for the joint distribution of $a_n^{-1}(X_{n:n} - b_n)$, $a_n^{-1}(X_{n-1:n} - b_n), \cdots, a_n^{-1}(X_{n-r+1:n} - b_n)$. Now define

$$K := \{(k_1,\cdots,k_{r-1}) : k_i \ge 0, i = 1,\cdots,r-1,$$

$$k_1 + k_2 + \cdots + k_j \le j, j = 1,2,\cdots,r-1\},$$

and for $x_1 > x_2 > \cdots > x_r$,

$$F_r(x_1,\cdots,x_r) :=$$

$$\sum_{(k_1,\cdots k_{r-1})\in K} \frac{(\log G(x_1) - \log G(x_2))^{k_1}}{k_1!} \cdot \cdots$$

$$\cdot \frac{(\log G(x_{r-1}) - \log G(x_r))^{k_{r-1}}}{k_{r-1}!} G(x_r).$$

It can be shown that (cf. Ref. [11] for the case r=2) if (1) holds for some nondegenerate distribution function G, then

$$P\{a_n^{-1}(X_{n:n} - b_n) \le x_1, a_n^{-1}(X_{n-1:n} - b_n) \le x_2, \cdots,$$

$$a_n^{-1}(X_{n-r+1:n} - b_n) \le x_r\}$$

$$\to F_r(x_1,\cdots,x_r).$$

However, we have

Theorem 2 *Suppose that (1) holds. Then*

$$P\{a_n^{-1}(Y_{n:n} - b_n) \le x_1, a_n^{-1}(Y_{n-1:n} - b_n) \le x_2, \cdots,$$

$$a_n^{-1}(Y_{n-r+1:n} - b_n) \le x_r | \mathbf{X}_n\}$$

$$\xrightarrow{d} \sum_{(k_1,\cdots,k_r)\in \mathbf{K}} \frac{(T(x_2,\omega) - T(x_1,\omega))^{k_1}}{k_1!} \cdot \cdots$$

$$\cdot \frac{(T(x_r,\omega) - T(x_{r-1},\omega))^{k_{r-1}}}{k_{r-1}!} e^{-T(x_r,\omega)},$$

where $T(\cdot,\omega)$ *is a Poisson random measure given in Theorem 1 and* $T(x,\omega) = T((x,\infty),\omega)$ *by definition.*

But suitable choices of m make the bootstrap consistent.

Theorem 3 *Suppose that (1) holds. If* $m(n)=o(n)$,

$$\sup_{x\in\mathbf{R}} \mid H_{n,m(n)}(x,\omega) - G(x) \mid \to 0, \tag{8}$$

in probability. Moreover, if $\sum_{n=1}^\infty \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (8) holds w.p.1 . (w.p.1 stands for "with probability 1".)*

Proof. We will write m=m(n) for convinience. Let $F_n(x) := \frac{1}{n}\sum_{j=1}^{n} I_{(-\infty,x]}(X_j)$ be the empirical distribution of $X_1, X_2, \cdots, X_n$. Then

$$\begin{aligned} H_{n,m}(x,\omega) &= F_n^m(a_m x + b_m) \\ &= \left\{1 - \frac{m(1 - F_n(a_m x + b_m))}{m}\right\}^m, \end{aligned}$$

and

$$\begin{aligned} m(1 - F_n(a_m x + b_m)) &= \frac{m}{n} n(1 - F_n(a_m x + b_m)) \\ &= \frac{m}{n} T_{n,m} \quad (say). \end{aligned}$$

Let $p_m := 1 - F(a_m x + b_m)$. Then (1) implies that $mp_m \to c(x) \equiv -\log G(x)$. Thus

$$E(\frac{m}{n} T_{n,m}) = mp_m \to c(x),$$

$$Var(\frac{m}{n} T_{n,m}) = \frac{m}{n} mp_m(1 - p_m) \to 0.$$

Therefore

$$m(1 - F_n(a_m x + b_m)) \to c(x)$$

in probability $\forall x \in \mathbf{R}$. Hence

$$H_{n,m(n)}(x,\omega) \to G(x)$$

in probability $\forall x \in \mathbf{R}$.

Since $\frac{m}{n} T_{n,m} = \frac{m}{n}(T_{n,m} - np_m) + mp_m$ and $mp_m \to c(x)$, we need show that $\frac{m}{n}(T_{n,m} - np_m) \to 0$ $w.p.1$ to prove that $\frac{m}{n} T_{n,m} \to c(x)$ w.p.1 . By the Borel Cantelli lemma, it is enough to show that

$$\sum_{n=1}^{\infty} P(|\ \frac{m}{n}(T_{n,m} - np_m)\ | > \epsilon) < \infty \quad \forall \epsilon > 0.$$

Let $\varphi_m(\theta) = p_m e^{\theta} + (1 - p_m)$ be the moment generating function of Bernoulli (p_m) distribution. Then $\forall \theta > 0$,

$$\begin{aligned} & \frac{m}{n} \log P(\frac{m}{n}(T_{n,m} - np_m) > \epsilon) \\ = & \frac{m}{n} \log P(e^{\theta(T_{n,m} - np_m)} > e^{\theta \frac{n}{m} \epsilon}) \qquad (9) \\ \le & \frac{m}{n} \log e^{-\theta \frac{n}{m} \epsilon} E(e^{\theta(T_{n,m} - np_m)}) \\ = & -\theta\epsilon - \theta mp_m + \log \varphi_m(\theta)^m \\ \to & -\theta\epsilon - \theta c(x) + \log(e^{c(x)(e^{\theta} - 1)}) \\ = & -\theta\epsilon + c(x)(e^{\theta} - 1 - \theta) \\ = & f(\theta, \epsilon) \quad (say). \end{aligned}$$

By taking the derivative of $f(\theta, \epsilon)$, we can show that $\theta_0(\epsilon) := \log(\frac{c + \epsilon}{c})$ minimizes $f(\theta, \epsilon)$. Let

$$g(\epsilon) := f(\theta_0, \epsilon) = -(\epsilon + c)\log(\frac{c + \epsilon}{c}) + \epsilon.$$

Then

$$g(0+) = 0$$

and

$$\begin{aligned} g'(\epsilon) &= 1 - \log(\frac{c + \epsilon}{c}) - (\epsilon + c)\frac{1}{\frac{c+\epsilon}{c}}\frac{1}{c} \\ &= -\log(\frac{c + \epsilon}{c}) < 0, \quad \forall \epsilon > 0. \end{aligned}$$

Thus $g(\epsilon) < 0, \forall \epsilon > 0$. Define

$$g_n(\epsilon) := \frac{m}{n} \log e^{-\theta_0(\epsilon)\frac{n}{m}\epsilon} E(e^{\theta_0(\epsilon)(T_n - np_m)}),$$

then $g_n(\epsilon) \to g(\epsilon)$. Let $\theta = \theta_0(\epsilon)$ in (9), then

$$\begin{aligned} & \sum_{n=1}^{\infty} P(\frac{m}{n}(T_{n,m} - np_m) > \epsilon) \\ = & \sum_{n=1}^{\infty} e^{\log P(\frac{m}{n}(T_{n,m} - np_m) > \epsilon)} \\ \le & \sum_{n=1}^{\infty} e^{\frac{n}{m} g_n(\epsilon)}. \end{aligned}$$

Given $\epsilon > 0$, $\exists \delta_\epsilon > 0$ such that $g(\epsilon) + \delta_\epsilon < 0$ and $\exists N_\epsilon \in \mathbf{N}$ such that $g_n(\epsilon) < g(\epsilon) + \delta_\epsilon, \forall n \ge N_\epsilon$. Therefore

$$\begin{aligned} \sum_{n=1}^{\infty} e^{\frac{n}{m} g_n(\epsilon)} &= \sum_{n=1}^{N_\epsilon - 1} e^{\frac{n}{m} g_n(\epsilon)} + \sum_{n=N_\epsilon}^{\infty} e^{\frac{n}{m} g_n(\epsilon)} \\ &\le \sum_{n=1}^{N_\epsilon - 1} e^{\frac{n}{m} g_n(\epsilon)} + \sum_{n=N_\epsilon}^{\infty} e^{\frac{n}{m}(g(\epsilon) + \delta_\epsilon)} \\ &< \infty \quad (by\ assumption). \end{aligned}$$

Hence $\sum_{n=1}^{\infty} P(\frac{m}{n}(T_{n,m} - np_m) > \epsilon) < \infty, \quad \forall \epsilon > 0.$

By a similar reasoning we can show that $\sum_{n=1}^{\infty} P(\frac{m}{n}(T_{n,m} - np_m) < -\epsilon) < \infty, \quad \forall \epsilon > 0.$ □

3 Limits of bootstrap distributions : when normalizing constants are unknown

Suppose that assumptions on $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ given in section 2 hold and suppose that a_n and b_n are unknown. Let $\hat{a}_m$ and $\hat{b}_m$ are estimators of a_m and b_m based on $X_1, X_2, \cdots, X_n$. Now, define the bootstrap distribution of $a_n^{-1}(X_{n:n} - b_n)$ with estimated normalizing constants by

$$\hat{H}_{n,m}(x,\omega) := P(\hat{a}_m^{-1}(Y_{m:m} - \hat{b}_m) \leq x | \mathbf{X}_n).$$

The same choice of m(n) as in the case of known normalizing constants gives the same results on the consistency of $\hat{H}_{n,m}(x,\omega)$ if $\hat{a}_m$ and $\hat{b}_m$ are correctly chosen as shown in the next theorem. Let $k_n = [\frac{n}{m}]$ and $k_n' = [\frac{n}{em}]$.

Theorem 4 *Assume that* $F \in D(\Lambda)$. *Define*

$$\begin{aligned} \hat{a}_m &= F_n^{-1}(1 - \frac{1}{em}) - F_n^{-1}(1 - \frac{1}{m}) \\ &= X_{n-k_n':n} - X_{n-k_n:n}, \\ \hat{b}_m &= F_n^{-1}(1 - \frac{1}{m}) = X_{n-k_n:n}. \end{aligned}$$

If m=o(n), then

$$\sup_{x \in \mathbf{R}} | \hat{H}_{n,m}(x,\omega) - \Lambda(x) | \to 0, \qquad (10)$$

in probability. Moreover, if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (10) is true w.p.1 .*

Theorem 5 *Assume that* $F \in D(\Phi_\alpha)$. *Define*

$$\begin{aligned} \hat{a}_m &= F_n^{-1}(1 - \frac{1}{m}) = X_{n-k_n:n}, \\ \hat{b}_m &= 0. \end{aligned}$$

If m=o(n), then

$$\sup_{x \in \mathbf{R}} | \hat{H}_{n,m}(x,\omega) - \Phi_\alpha(x) | \to 0, \qquad (11)$$

in probability. Moreover, if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (11) is true w.p.1 .*

Theorem 6 *Assume that* $F \in D(\Psi_\alpha)$. *Define*

$$\begin{aligned} \hat{a}_m &= \theta_{F_n} - F_n^{-1}(1 - \frac{1}{m}) = X_{n:n} - X_{n-k_n:n}, \\ \hat{b}_m &= \theta_{F_n} = X_{n:n}. \end{aligned}$$

If m=o(n), then

$$\sup_{x \in \mathbf{R}} | \hat{H}_{n,m}(x,\omega) - \Psi_\alpha(x) | \to 0, \qquad (12)$$

in probability. Moreover, if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (12) is true w.p.1 .*

Note that the result of each of the above theorems implies that

$$\sup_{x \in \mathbf{R}} | \hat{H}_{n,m}(x,\omega) - G_n(x) | \to 0,$$

in probability if m=o(n) and w.p.1 if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$. Therefore $\hat{H}_{n,m}(x,\omega)$ approximates $G_n(x)$ uniformly when $n \to \infty$. Note also that m=o($\frac{n}{\log n}$) is sufficient for $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$.

The following theorem shows that the joint distribution of $a_n^{-1}(X_{n:n} - b_n), a_n^{-1}(X_{n-1:n} - b_n)$, $\cdots, a_n^{-1}(X_{n-r+1:n} - b_n)$ can be bootstrapped successfully.

Theorem 7 *Assume the hypothesis on F and choose* $\hat{a}_m$ *and* $\hat{b}_m$ *as in Theorem 4,5,or 6 according to the domain of attraction F belongs to. If m=o(n), then*

$$\begin{aligned} \sup | P\{\hat{a}_m^{-1}(Y_{m:m} - \hat{b}_m) \leq x_1, \hat{a}_m^{-1}(Y_{m-1:m} - \hat{b}_m) \leq x_2, \\ \cdots, \hat{a}_m^{-1}(Y_{m-r+1:m} - \hat{b}_m) \leq x_r | \mathbf{X}_n\} \\ - F_r(x_1, \cdots, x_r) | \to 0, \qquad (13) \end{aligned}$$

in probability, where supremum is taken among every $x_1 > \cdots > x_r$. *Moreover, if* $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (13) is true w.p.1 .*

Theorem 7 and the continuous mapping theorem gives the following.

Theorem 8 *Assume the hypothesis on F and choose* $\hat{a}_m$ *and* $\hat{b}_m$ *as in Theorem 4,5,or 6 according to the domain of attraction F belongs to. Let*

$f : \mathbf{R}^r \to \mathbf{R}^l$ *be continuous a.e. with respect to* $F_r(\cdot, \cdots, \cdot)$. *If m=o(n), then*

$$\sup_{\mathbf{y}\in\mathbf{R}^l} | P\{f(\hat{a}_m^{-1}(Y_{m:m} - \hat{b}_m), \hat{a}_m^{-1}(Y_{m-1:m} - \hat{b}_m), \cdots,$$
$$\hat{a}_m^{-1}(Y_{m-r+1:m} - \hat{b}_m)) \le \mathbf{y}|\mathbf{X}_n\}$$
$$-P\{f(a_n^{-1}(X_{n:n} - b_n), a_n^{-1}(X_{n-1:n} - b_n), \cdots,$$
$$a_n^{-1}(X_{n-r+1:n} - b_n)) \le \mathbf{y}\} | \to 0, \qquad (14)$$

in probability. Moreover, if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (14) is true w.p.1 .*

Corollary 1 *Assume that* $F \in D(\Psi_\alpha)$. *If m=o(n), then*

$$\sup_{x\in\mathbf{R}} | P(\frac{Y_{m:m} - X_{n:n}}{Y_{m:m} - Y_{m-1:m}} \le x|\mathbf{X}_n)$$
$$-P(\frac{X_{n:n} - \theta_F}{X_{n:n} - X_{n-1:n}} \le x) | \to 0, \qquad (15)$$

in probability. Moreover, if $\sum_{n=1}^{\infty} \lambda^{\frac{n}{m}} < \infty$, $\forall 0 < \lambda < 1$, *then (15) is true w.p.1 .*

Confidence intervals for θ_F based on the asymptotic distribution of $(X_{n:n} - \theta_F)/(X_{n:n} - X_{n-1:n})$ were considered by Ref. [12]. We apply above corollary to approximate the critical points of $P((X_{n:n} - \theta_F)/(X_{n:n} - X_{n-1:n}) \le x)$. Let

$$R_n = \frac{Y_{m:m} - X_{n:n}}{Y_{m:m} - Y_{m-1:m}}.$$

First we obtain a large number, say N of bootstrap replicates of size m(n) from F_n (the empirical distribution of $X_1, X_2, \cdots, X_n$) and then compute $R_{n,i}$ for $i = 1, 2, \cdots, N$ and set

$$\hat{H}_{n,m}^N(x) = \frac{1}{N} \sum_{i=1}^{N} I_{(-\infty,x]}(R_{n,i}).$$

As $N \to \infty$, $\hat{H}_{n,m}^N(x)$ approximate $P(R_n \le x|\mathbf{X}_n)$ which is close to $P((X_{n:n} - \theta_F)/(X_{n:n} - X_{n-1:n}) \le x)$ when $n \to \infty$. Thus a $100(1-\eta)\%$ approximate confidence interval for θ_F will be

$$(X_{n:n} - r_{n,1-\eta_2}(X_{n:n} - X_{n-1:n}),$$
$$X_{n:n} - r_{n,\eta_1}(X_{n:n} - X_{n-1:n})),$$

where r_{n,η_1} and $r_{n,1-\eta_2}$ are chosen such that

$$\hat{H}_{n,m}^N(r_{n,\eta_1}) = \eta_1, \quad \hat{H}_{n,m}^N(r_{n,1-\eta_2}) = 1 - \eta_2,$$

where

$$\eta_1 > 0, \ \eta_2 > 0, \ \eta_1 + \eta_2 = \eta.$$

4 Conclusions

The proofs of the results in section 3 are in Fukuchi (1994). We are currently working on extending the present work to the case of stationary sequence of random variables under appropriate mixing conditions. Also some simulation work is in progress to assess the role of the resample size m(n). We are grateful to Professor S. Lahiri and A. Vidyashankar for several useful discussion and comments.

References

[1] Efron, B., Bootstrap methods-another look at the jackknife, *Ann. Statist.*, **7** (1979), 1-26.

[2] Athreya, K. B., Bootstrap of the mean in the infinite variance case, *Ann. Statist.*, **15** (1987), 724-731.

[3] Arcones, M. and Giné, E., The bootstrap of the mean with arbitrary bootstrap sample size, *Ann. Inst. H. Poincaré.*, **25** (1989), 457-481.

[4] Singh, K., On the asymptotic efficiency of Efron's Bootstrap, *Ann. Statist.*, **9** (1981), 1187-1195.

[5] Künsch, H. R., The jacknife and the bootstrap for general stationary observations, *Ann. Statist.*, **17** (1989), 1217-1241.

[6] Bickel, P. J. and Freedman, D. A., Some asymptotic theory for the bootstrap, *Ann. Statist.*, **9** (1981), 1196-1217.

[7] Angus, J., Asymptotic theory for bootstrapping the extremes, *Commun. Statist.-Theory Meth.*,, **22(1)** (1993), 15-30.

[8] Deheuvels, P., Mason, D. and Shorack, G., Some results on the influence of extremes on the bootstrap, *Ann. Inst. H. Poincaré.*, **29** (1993), 83-103.

[9] Haan, L. de, *On Regular Variation and its Application to the Weak Convergence of Sample Extremes.* Math. Centre Tracts, **32** (1970), Mathematical Centre, Amsterdam.

[10] Resnick, S. I., *Exterme Values, Regular Variation, and Point Processes.* Springer, Berlin, (1987).

[11] Leadbetter, M. R., Lindgren, G. and Rootzen, H., *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin, (1983).

[12] Cooke, P. J., Statistical inference for bounds of random variables, *Biometrika.*, **66** (1979), 367-374.

[13] Fukuchi, J., *The Bootstrap Method for Extremes of Random Variables.* Ph.D. dissertation under preparation, Iowa State University, (1994).

Extreme Analysis Of Wave Pressure And Corrosion For Structural Life Prediction

Ayyub, B.M.
University of Maryland, College Park, MD

Extreme analysis can be used in structural life expectancy assessment. In this paper, extreme analysis was used for this purpose in two aspects of life expectancy assessment. These aspects are (1) extreme wave pressure prediction, and (2) extreme corrosion estimation. Then, they were used in a time variant reliability assessment formulation of a marine structure. The result is the reliability of a structure as a function of time, which can be viewed as the cumulative distribution function of structural life. The presented methodology was performed in an effort to assess the life expectancy of patrol boats. In the applications discussed in this paper, the underlying parent distribution in the extreme value analysis was assumed to have an infinite exponential tail. This assumption can significantly affect the resulting extreme value distributions and assessed structural reliability levels. In dealing with waves or corrosion, the tails are limited based on the physics of both problems. The former is limited by the hydrodynamics of waves, and the latter is limited by the size of a corroded element. The effects of limiting the tails of parent distributions on the results of these applications require further investigation.

1. INTRODUCTION

The factors that affect the life of a structure include design parameters, design safety factors, design methods, type of structure, structural details, materials, construction methods and quality, loads, maintenance practices, inspection methods, and other environmental factors. These factors have different types of uncertainty that can be classified as: (1) physical randomness in magnitude and time of occurrence, (2) statistical uncertainties due to using limited amount of information in estimating the characteristics of the population, (3) model uncertainties due to approximations in the prediction models, and (4) vagueness in the definition of the factors, system and/or assessing their/its effect on life. Therefore, the estimation of life expectancy is a complex process. Because of the stochastic nature of many of the uncertainties, a probabilistic approach, as opposed to a deterministic approach, is better suited for life expectancy prediction. Life expectancy associated with failure modes such as yielding, plastic deformation and buckling can be estimated using the extreme analysis, and life expectancy associated with failure modes such as fracture and fatigue are estimated using the cumulative value modeling approach. In dealing with corrosion, life expectancy assessment can be based on extreme analysis. Example applications of life expectancy prediction are provided by Ayyub et al [4], Ayyub and White [3], Harris et al [5], and Yazdani and Albrecht [12].

2. STRUCTURAL RELIABILITY ASSESSMENT

The performance function that expresses the relationship between the strength and load effects of a structural member according to a specified failure mode is given by

$$M = g(X_1, X_2, ..., X_p) = R - L \qquad (1)$$

in which the X_i, i = 1,..,p are the p basic random variables which define the loads, material properties and other structural parameters, g(.) is the functional relationship between the basic random variables and failure (or survival); R is resistance or strength; and L is load effect. The probability of failure can be evaluated by the following integral:

$$P_f = \int\int \ldots \int f_{\underline{X}}(x_1, x_2, \ldots, x_p)\, dx_1 dx_2 \ldots dx_p \quad (2)$$

where $f_{\underline{X}}$ is the joint density function of X_1, X_2, ..., X_p, and the integration is performed over the region where $M < 0$. The strength (or resistance) R of a structural component and the load effect L are generally functions of time. Therefore, the probability of failure is also a function of time. The time effect can be incorporated in the reliability assessment by considering the time dependence of one or both of the strength and load effects.

3. EXTREME ANALYSIS OF WAVE PRESSURE

Extreme values based on observational data are very important in structural safety and life assessment. The prediction of future conditions, especially extreme conditions, are necessary in engineering planning and design. The prediction is performed based on an extrapolation from previously observed data. For a set of observations (x_1, x_2, ..., x_k) from an identically distributed and independent set of random variables (X_1, X_2, ..., X_k), the distribution of X_i is called the parent (or initial) distribution. It has the cumulative probability distribution function $F_X(x)$ and the density probability function $f_X(x)$. The maximum value of the observed values is a random variable M_k which can be represented as

$$M_k = \text{Maximum}\,(X_1, X_2, \ldots, X_k) \quad (3)$$

The exact cumulative and density probability distribution functions of the maximum value are, respectively, given by

$$F_{M_k}(m) = \left[F_X(m)\right]^k \quad (4a)$$

$$f_{M_k}(m) = k\left[F_X(m)\right]^{k-1} f_X(m) \quad (4b)$$

It can be shown that for relatively large values of k, the extreme distribution approaches an asymptotic form that is not dependent on the exact form of the parent distribution; but, it depends on the tail characteristics of the parent distribution in the direction of the extreme. The central portion of the parent distribution has little influence on the asymptotic form of the extreme distribution. For parent probability distributions of exponential tails, the extreme distribution approaches an extreme value distribution of double exponential form as $k\rightarrow\infty$. For example, a normal or lognormal probability distribution approaches a Type I extreme value distribution as $k\rightarrow\infty$. In this case, the difference between an exact distribution for M_k and the Type I extreme value distribution is relatively small. The difference diminishes as $k\rightarrow\infty$. Practically, the difference is negligible for k larger than approximately 25.

For the purpose of life prediction, the mathematical model for the extreme distribution needs to be a function of k in order to relate the outcome of the analysis of extreme statistics to time. Extreme value distributions, like the Type I largest extreme value distribution, are used in this paper to model extreme load effects. Since the mathematical model is not sensitive to the type of the parent distribution, as long as it is within the same general class, the mathematical model used in this chapter is based on an parent distribution that follows the class of normal probability distributions.

For a normal parent probability distribution of the random variable X with a mean value μ and standard deviation σ, the cumulative distribution and density functions of the largest value M_k of k identically distributed and independent random variables (X_1, X_2, ..., X_k) are, respectively, given by

$$F_{M_k}(m) = \text{Exp}\left\{-\text{Exp}\left[\left(-\frac{\alpha_k}{\sigma}\right)\left(m-\mu-\sigma u_k\right)\right]\right\} \quad (5)$$

$$f_{M_k}(m) = \left(\frac{\alpha_k}{\sigma}\right)\text{Exp}\left[\left(-\frac{\alpha_k}{\sigma}\right)\left(m-\mu-\sigma u_k\right)\right] \text{Exp}\left\{-\text{Exp}\left[\left(-\frac{\alpha_k}{\sigma}\right)\left(m-\mu-\sigma u_k\right)\right]\right\} \quad (6)$$

where

$$\alpha_k = [2\ln(k)]^{0.5} \quad (7a)$$

and

$$u_k = \alpha_k - \{\ln[\ln(k)] + \ln(4\pi)\}/(2\alpha_k) \quad (7b)$$

The mean value and standard deviation of M_k can be determined approximately using the central and dispersion characteristics of Type I extreme value distribution. They are given, respectively, by the following:

$$\text{Mean value, } \overline{M}_k = \sigma u_k + \mu + \frac{\gamma\sigma}{\alpha_k} \quad (8)$$

$$\text{Standard Deviation, } \sigma_{M_k} = \frac{\pi}{\sqrt{6}}\frac{\sigma}{\alpha_k} \quad (9)$$

The constants π and γ have the values of 3.141593 and 0.577216, respectively.

3.1 Example

The method is illustrated by considering the plastic deformation failure mode of a marine vessel (Ayyub and White [3], and Ayyub et al [4]) from which this example is taken. Although this example deals with random sea loads on marine vessels, the method is equally applicable to other random loads, such as wind loads and earthquake loads. For illustration purposes the critical failure mode is assumed to be plastic plate deformation of the shell of the vessel. The objective of the analysis is to determine the cumulative distribution function of structural life for this failure mode.

The end of structural life of the vessel according to the specified failure mode is defined as having to replace more than five plate panels in a specified critical region at the end of any inspection period. It is assumed that plate panels are to be replaced when the ratio of plastic deformation to plate thickness is greater than or equal to 2.0. This assumption is usually based on the resources allocated for repair and steel replacement for the vessels in their lifetime maintenance cycle. The inspection schedule of the boat includes the warranty inspection at the end of the first year followed by regular inspections every I years, where I can be either one or two years.

In this case, the performance function takes the following general form:

g = Resistance - Still Water Load - Dynamic Load(10)

Each of the terms in the above equation are expressed in units of pressure. The still water load is the hydrostatic pressure at the depth of the critical region. It can be determined based on the design draft. The dynamic load is the extreme dynamic pressure based on the results from full scale experiments conducted on one of the vessels. The resistance term is an empirical expression developed by Hughes [6] based on elasto-plastic plate response.

In this example, only loads and load effects in head-seas are considered. No other heading is considered because reported stress records by Purcell et al [7] indicate that they result in much smaller stresses than the head-seas condition. Eight combinations of vessel's speed and sea-state for the head-seas condition are considered herein. These combinations are summarized in Table 1. For the eight combinations, strain measurements at locations of interest were performed by Purcell et al [7]. The combination of high sea-state and high speeds was not tested. The percentages that are shown in Table 1 for each combination represent the percentage usage of the vessel in the corresponding speed/sea-state combination. These percentages are based on a survey conducted by the same researchers. The total of the percentages in the table is about 20%, which is the expected percent usage in head-seas.

The performance function as given by Eq. 10 includes two components of pressure, i.e., still-water and dynamic pressure. The stress due to the still-water pressure component can be modeled using random variables. Since the stresses due to still-water pressure were not measured, the mean value of the still-water pressure was determined based on hydrostatic analysis using the vessel's draft and was found to be 2.667 psi (Purcell et al [7]). The coefficient of variation and distribution type of still-water pressure are assumed to be 0.20 and normal, respectively.

Table 1. Combinations of ship speed and sea state

Sea State [Wave Height]	Ship Speed		
	Low (12 knots)	Medium (24 knots)	High (29 knots)
Low [3 ft]	Case 1 (12 knots, 3 ft) 4.0%	Case 2 (24 knots, 3 ft) 1.7%	Case 3 (29 knots, 3 ft) 1.0%
Medium [8 ft]	Case 4 (12 knots, 8 ft) 4.7%	Case 5 (24 knots, 8 ft) 1.3%	Case 6 (29 knots, 8 ft) 0.7%
High [10 ft]	Case 7 (12 knots, 10 ft) 5.3%	Case 8 (24 knots, 10 ft) 1.0%	Not considered

The strains due to the dynamic pressure were measured, and the computed stresses should be modeled using the statistics of extremes. The parent distribution for the measured stress was assumed to be the probability distribution of a random variable that is defined as the maximum stress due to dynamic pressure in 30-second interval for all the cases in Table 1, except Case 8. For Case 8, the interval is taken as 10 seconds. The statistical characteristics of the parent distribution of stress for the eight cases were determined using the data reported by Purcell et al [7]. The mean values and coefficients of variation (COV) for Cases 6 and 8 were based on 10 and 23 maximum values taken from 10 and 23 records of stress time-history, respectively. Other cases were based on one record each. Then, plate theory and finite element analysis were used to determine the mean value of the maximum dynamic pressure, Mean(P_{max}), that causes the measured stresses. The results are summarized in Table 2. It is reasonable to assume that the COV of the maximum dynamic pressure, COV(P_{max}), is the same as the COV of the maximum measured stress. The mean value and COV of the extreme pressure were, then, determined using Eqs. 8 and 9 for an example vessel's usage period of 15 years at a rate of 3000 hours per year and according to the percent use presented in Table 1. The results are shown in Table 2. It was also assumed that the extreme pressure follows Type I largest extreme value probability distribution.

Table 2. Statistical characteristics of pressure (15 years of usage)

Case no.	Mean (P_{max}) (psi)	COV (P_{max})	No. of intervals in life, K	Mean (P_{extrm}) (psi)	COV (P_{extrm})
1	1.75	0.0993	216000	2.55	0.0177
2	1.89	0.0993	91800	2.71	0.0186
3	1.99	0.0993	54000	2.83	0.0192
4	6.17	0.0993	253800	8.99	0.0175
5	6.76	0.0993	70200	9.66	0.0189
6	3.07	0.0993	37800	4.35	0.0196
7	7.63	0.0993	286200	11.13	0.0174
8	13.37	1.0121	162000	74.30	0.0477

It is evident from the Table 2 that Case 8 is the most critical sea state/boat speed combination. Therefore, for this case the statistics of the maximum and extreme pressures were determined using the usage periods of 0.2, 0.5, 1, 2, 5, 10, 15, 50, and 100 years using Eqs. 8 and 9 as shown in Table 3.

Table 3. Statistical characteristics of pressure for case 8

Usage period (years)	Number of intervals in life, K	Mean (P_{extrm}) (psi)	COV (P_{extrm})
0.2	2160	60.49	0.0732
0.5	5400	63.70	0.0657
1	10800	66.02	0.0610
2	21600	68.24	0.0569
5	54000	71.07	0.0523
10	108000	73.13	0.0493
15	162000	74.30	0.0477
50	5400000	77.67	0.0435
100	1080000	79.54	0.0414

The statistical characteristics of the strength of the material used for the vessel and the dimensions of a

plate within the critical region were determined by Ayyub and White [3]. The mean values and COV of the yield stress and modulus of elasticity of the material were estimated to be 47.8 ksi, 29,774 ksi, 0.13 and 0.038, respectively. The mean values and COV of the thickness and the overall dimensions of the plate were estimated to be 0.161 in., 11.75 in. x 23.5 in., 0.01, 0.05 and 0.05, respectively.

The failure probabilities of a plate according to the limit state of Eq. 10 can be determined using Monte Carlo simulation with variance reduction techniques (Ayyub and Haldar [2]). The average probabilities of failure of a plate (P_{fp}), coefficients of variation of the estimate of the probability of failure COV(P_{fp}) and the numbers of simulation cycles for different usage periods of the boat are shown in Table 4.

The critical region for the vessel was defined as the region that consists of a total of 28 plates. These plates were assumed to experience the same loading and have approximately the same strength characteristics; therefore, have approximately the same probability of failure. Since the end of life is defined as failure of more than 5 plates (out of the 28 plates), the vessel (or structural system) can be considered to fail if 6 plates or more out of the 28 fail. Let us first consider a warranty period of one year and an inspection interval of one year. For a period of one year, plate failure probability (P_{fp}) is 0.06765 (from Table 4). Since end of life is defined as failure of at least 6 out of 28 plates, we can consider the n-out-of-N system with n = 6 and N = 28. Failure probability of this system depends on the statistical correlation between the plate failures. An upper bound failure probability is obtained when the correlation coefficient is unity and a lower bound is obtained when the correlation coefficient is zero. The correlation between the plate failures is assumed to be small, and so the lower bound is closer to the actual (unknown) value. The lower bound failure probability can be based on the binomial distribution. The probability of failure of at least six plates out of 28 plates ($P_{f6/28,I}$) at the end of one year of service as 0.00989. Similar calculations for inspection intervals of two years (I = 2) with P_{fp} = 0.09403 (from Table 4) yields $P_{f6/28,I}$ = 0.042719. Since the warranty period is one year (W = 1), $P_{f6/28,W}$ = 0.009895.

Table 4. Probability of failure of a plate (without inspection effect)

Usage period (years)	Number of simulation cycles	Probability of failure P_{f_p}	COV (P_{f_p})
0.2	3000	0.03004	0.0490
0.5	3000	0.05092	0.0401
1	3000	0.06765	0.0351
2	3000	0.09403	0.0294
5	2000	0.13950	0.0284
10	2000	0.17200	0.0244
15	2000	0.20310	0.0215
50	2000	0.27760	0.0155
100	2000	0.32900	0.0121

Failure probabilities at different durations of service T (years of usage) were computed at T = 1, 3, 5, 11, 21, 31, and 51 years were computed and plotted in Fig. 1. This graph of failure probability versus time is also the cumulative distribution function of structural life. It is evident from Fig. 1 that by reducing the inspection interval, expected structural life can be enlarged. This due to the fact that at the end of each inspection interval, any reported deformation damage is to be fixed before sending the vessel for the next usage period.

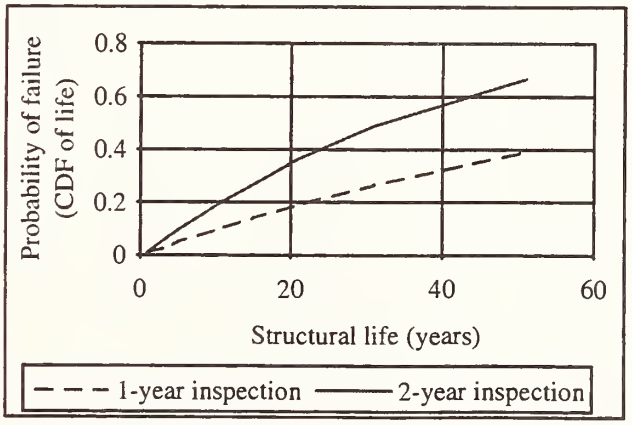


Figure 1. Probability of failure based on plate deformation

4. EXTREME ANALYSIS OF CORROSION

One of the problems facing an engineer when attempting to perform life prediction of a structure is how to adequately deal with corrosion. There are handbooks available which are full of test data for a wide variety of metals, exposure duration's, and types of corrosive attacks (Schumacher [8]). The difficulty comes in attempting to fit one of the examples to the real case at hand. Typically, the information on available corrosion rates is not the type of information needed by the engineer. The engineer needs to make a determination of remaining strength of a panel of plating based on mean thickness and a determination of the integrity of the plating from the depths of pits. White and Ayyub [9] developed an approach for both planning the number and location of measurements to be taken using semivariogram analysis (Ayyub and McCuen [1]), and for using the information obtained to do a reliability-based service life assessment of the structure (White and Ayyub [10]). In this paper, a means of determining a maximum value of pitting depth based on thickness measurements is incorporated by treating the growth of pits as a random process with some specific statistical characteristics.

When performing a service life analysis of marine structures both pitting and general wastage need to be included in assessing a number of potential failure modes. The determination of the rate of corrosion and the rate of pitting has been a major difficulty in designing cost-effective and reliable structures. The extreme depth of pits can be estimated by (1) the theory of extremes, or (2) sampling from largest pits. These two methods are described in the following sections. The methods serve different objectives. The first method can be used in cases were the corroded side of the metal is not accessible. Therefore, general sampling can be used to determine the statistical characteristics of thickness. The resulting probability distribution of thickness can be treated as an underlying parent distribution in the theory of extremes. In cases that involve accessible corroded sides, both methods can be used. However, the second method provides a direct measurement of pitting depth. Then the concept of percentile and largest depths can be used to characterize the extreme depths.

4.1 Measurements Taken Without Knowledge of Extent of Pitting

Consider a set of n observations $(x_1, x_2, ..., x_n)$ from identically distributed and independent set of random variables $(X_1, X_2, ..., X_n)$. The distribution of X_i is called the initial (or parent) distribution, that has the cumulative probability distribution function $F_X(x)$ and the probability density function $f_X(x)$. The minimum observed value is a random variable M_1 which can be represented as

$$M_1 = \text{Minimum}\,(X_1, X_2, ..., X_n) \qquad (11)$$

The exact cumulative and density probability distribution functions of the minimum value are given by, respectively:

$$F_{M_1}(m) = 1 - [1 - F_X(m)]^n \qquad (12)$$

$$f_{M_1}(m) = n[1 - F_X(m)]^{n-1} f_X(m) \qquad (13)$$

It can be shown that for relatively large values of n, the extreme value distribution approaches an asymptotic form that is not dependent on the exact form of the initial probability distribution; but, it depends on the tail characteristics of the initial distribution in the direction of the extreme. The Type I extreme value distribution is used in this paper to model extreme corrosion. Since the mathematical model is not sensitive to the type of the initial distribution, as long as it is within the same general class, the mathematical model used in this study is based on an initial distribution that follows the class of normal probability distributions.

For a log-normal initial probability distribution of the random variable X with a mean value μ and standard deviation σ, the cumulative distribution and density functions of the smallest value M_1 of n identically distributed and independent random variables $(X_1, X_2, ..., X_n)$ are of a smallest-extreme Type I distribution, and are, respectively, given by

$$F_{M_1}(m) = 1 - \text{Exp}\{-\text{Exp}[(\alpha_1/\sigma)(m - \mu - \sigma u_1)]\} \qquad (14)$$

$$f_{M_1}(m) = (\alpha_1/\sigma)\,\text{Exp}[(\alpha_1/\sigma)(m - \mu - \sigma u_1)]\,\text{Exp}\{-\text{Exp}[(\alpha_1/\sigma)(m - \mu - \sigma u_1)]\} \qquad (15)$$

where

$$\alpha_1 = [\,2 \ln(n)\,]^{1/2} \qquad (16)$$

$$u_1 = -\alpha_1 + \{\,\ln[\ln(n)] + \ln(4\pi)\,\}/(2\alpha_1) \qquad (17)$$

The mean value and standard deviation of M_1 can be determined approximately using the central and

dispersion characteristics of Type I smallest-extreme value distribution, and are , respectively, given by

$$\text{Mean value, } \overline{M}_1 = \sigma u_1 + \mu - \gamma\sigma/\alpha_1 \quad (18)$$

$$\text{Standard Deviation, } \sigma_{M_1} = (\pi/\sqrt{6})(\sigma/\alpha_1) \quad (19)$$

In using this method, n can be assumed to represent an approximate number of pits in a location of interest.

4.2 Measurements Taken in Deepest Pits

In the pervious section, thickness sampling can be performed in the form of a grid that cover a specified section of a structure. The resulting statistical characteristics were considered to constitute the moments for a parent distribution with an exponential tail (e.g. normal distribution). Then, the theory of extremes was used to determine the statistical characteristics of the smallest-extreme thickness as a measure of the deepest pit. In this section, the depth of k pits in a specified location of a structure are sampled. The statistical characteristics of these pits can then be determined using the sample of size k. Assume that the section of interest of the structure has n pits, which is sufficiently large, and also assume that the depth of a pit is a random variable X with the following probability density and distribution functions:

$$f_X(x) = \lambda \exp(-\lambda x) \qquad \text{where } x \geq 0 \quad (20)$$

$$F_X(x) = 1 - \exp(-\lambda x) \quad (21)$$

The parameter λ can be determined based on the sampled mean value $\overline{X}$ as $\lambda = 1/\overline{X}$. Then, the deepest pit P in the section of interest has, respectively, the following cumulative distribution function $F_P(p)$ and density function $f_P(p)$:

$$F_P(p) = \exp[-n \exp(-\lambda p)] \quad (22)$$

$$f_P(p) = n \lambda \exp(-\lambda p) \exp[-n \exp(-\lambda p)] \quad (23)$$

Integration or simulation can be used to determine the mean value and standard deviation of the deepest pit.

4.3 Example

This example was taken from a study performed by White and Ayyub [11]. The data used in this example were the results of an ultrasonic hull inspection of one of the vessels of the Class being studied (White and Ayyub [9]). There were over 3,000 individual thickness measurements taken on the roughly 82-foot long hull. The measurements were reported on a shell-expansion drawing with large sections marked to indicate areas of excessive corrosion or pitting. Figure 2 provides an excerpt from that drawing showing a section consisting of the plating between transverse frames covering three longitudinals. In this part of the vessel the frame spacing is 60-inches and the longitudinal spacing is 24-inches. For this 2880-sq.in. area ten thickness measurements were taken. Table 5 provides the locations of the measurement points with respect to the lower left corner of Figure 2. As can be seen from Figure 2, all of the measurements were larger than the nominal design plating thickness of the hull. This fact was noted for all parts of the hull and not just this section. The apparent discrepancy in the measured values when compared to the design thickness alone would lead one to suspect the measurements and probably discard them as not being useful. How then can the information provided by the ultrasonic survey be effectively used?

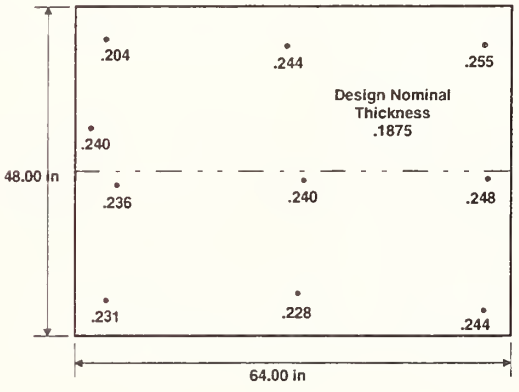


Figure 2. Measured Thicknesses on Example Bottom Plating Panel

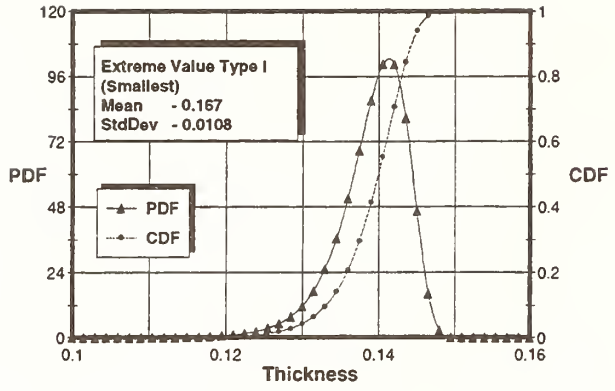


Figure 3. CDF and PDF of Extreme Smallest Thickness

Table 5. Thickness Measurements and Locations for Example Plating

Point No.	x-Location (in)	y-Location (in)	Thickness (in)
1, 2	4, 32	4, 5	0.231, 0.228
3, 4	60, 6	3, 22	0.244, 0.236
5, 6	33, 61	23, 23	0.240, 0.248
7, 8	2, 4	30, 43	0.240, 0.204
9, 10	31, 60	42, 42	0.244, 0.255

The apparent thickness increase over the design thickness may be the result of the measuring device (or operator) not being able to either get good contact with the surface, or misinterpreting the display of the returning signal, or even a mis-calibration of the measuring device. In each of these cases we can treat the difference in the thickness as a bias error. By taking the largest thickness measured and subtracting the nominal design thickness we can get an estimate of the bias. Later we could vary the size of the bias to see what effect it has on the final results. For the example case, an estimated bias of 0.07 inches was used. The ultrasonic measuring device, if properly calibrated and used by an experienced operator is capable of providing a very high accuracy. The accuracy, however, is rapidly degraded depending on surface conditions, what is behind the plating being measured, and the skill of the operator.

In performing the extreme analysis, the values for n to be used in Eqs. 14 to 19 and the confidence level desired for the confidence interval computations are needed. Because the area under investigation was identified as being "heavily pitted", a large number of pits was assumed to be present. Though value of n has little effect on the results once n exceeds about 100, we will use n = 1000 just to indicate severe pitting. Figure 3 shows the PDF and CDF (from Eqs. 14 and 15) for the extreme smallest value based on the parameters for this example.

A confidence interval can now be computed. For this example, a 90% confidence interval was desired as well as a 90th percentile extreme smallest thickness (largest pit depth). For pit depth the value is used in the inverse CDF of the extreme to find that extreme smallest thickness with only a 10% probability of being exceeded (having a pit deeper). The results of the analysis are presented in Table 6.

Table 6. Results of Analysis for Example Plating

Estimated Mean Thickness	0.1670 in.
Confidence Interval: Upper Bound	0.1733 in.
Lower Bound	0.1607 in.
Depth of Extreme Pit	0.1332 in.

5. CONCLUDING REMARKS

In the applications discussed in this paper, the underlying parent distribution in the extreme value analysis was assumed to have an infinite exponential tail. This assumption can significantly affect the resulting extreme value distributions and assessed structural reliability levels. In dealing with waves or corrosion, the tails are limited based on the physics of both problems. The former is limited by the hydrodynamics of waves, and the latter is limited by the size of a corroded element. The effects of limiting the tails of parent distributions on the results of these applications require further investigation.

REFERENCES

[1] Ayyub, B.M. and McCuen, R.C. Optimum Sampling for Structural Strength Evaluation, J. of Structural Engineering, ASCE,116(2) (1990) 518-535.

[2] Ayyub, B.M. and Haldar, A. Practical structural reliability techniques, ASCE, J. Struct. Engr., 110(8) (1984), 1707-1724.
[3] Ayyub, B.M., and White, G.J. Life expectancy assessment and durability of the Island-Class patrol boat hull structure, Report Submitted to U.S. Coast Guard R&D Center, CT (1987).
[4] Ayyub, B.M., White, G.J., and Purcell, E.S. Estimation of structural service life of ships, Naval Engineers J., 101(3) (1989), 156-166.
[5] Harris, D.O., Lim, E.Y., and Dedhia, D.D. Probability of Pipe Fracture in the Primary Coolant Loop of a PWR Plant, Volume 5, Probabilistic Fracture Mechanics Analysis, NUREG/CR-2189 (1981).
[6] Hughes, O.F. Design of laterally loaded plating - uniform pressure loads, SNAME J. Ship Res., 25(2) (1981), 77-89.
[7] Purcell, E.S., Allen, S.J. and Walker, R.J. Structural analysis of the U.S. Coast Guard Island-Class patrol boat, SNAME Trans., 96(7) (1988), 1-23.
[8] Schumacher, M. Seawater Corrosion Handbook, Noyes Data Corp., Park Ridge, NJ, (1979).
[9] White, G.J. and Ayyub, B.M. Semivariogram and Kriging Analysis in Developing Sampling Strategies, In: Proceedings, Symposium on Uncertainty Modeling and Analysis, MD, Ed. B. Ayyub, (1990), 360-365.
[10] White, G.J. and Ayyub, B.M. A Probabilistic Approach for Determining the Effect of Corrosion on the Life Expectancy of Marine Structures, In: Proceedings, Marine Structural Inspection, Maintenance, and Monitoring Symposium, SSC/SNAME, VA, (1991).
[11] White, G.J. and Ayyub, B.M. A Probabilistic-Based Methodology for Including Corrosion in the Structural Life Assessment of Marine Structures, Report EW-04-92, Engineering, U.S. Naval Academy, MD, (1992).
[12] Yazdani, N., and Albrecht, P. Risk analysis of fatigue failure of highway bridges. J. of Structural Engineering, ASCE, 113(3) (1987), 483-500.

Record Values From Rayleigh And Weibull Distributions And Associated Inference

Balakrishnan, N.
McMaster University, Hamilton, Ontario, Canada

Chan, P.S.
The Chinese University of Hong Kong, Shatin, Hong Kong

In this paper, we study the upper record values from a Rayleigh distribution and derive explicit expressions for the means, variances and covariances. We also establish some recurrence relationships for the single and product moments. These results are then used to derive explicitly the best linear unbiased estimators for the scale-parameter as well as the location-scale parameter cases. Some associated inference with regard to the prediction of a future record value and the test for spuriosity of the current record values are also developed. Next, we present two examples and illustrate all these inference procedures. Finally, we extend all these developments to the Weibull distribution and present the necessary explicit algebraic formulae.

1 Introduction

Record values and associated statistics are of importance in many real-life situations involving data relating to weather, sports, economics, and life-tests. The statistical study of record values started with Chandler [11] and since then have been pursued in different directions by several authors; for example, see Glick [15], Galambos [14], Resnick [20], Nagaraja [18], Nevzorov [19], Ahsanullah [2], Arnold and Balakrishnan [4], and Arnold, Balakrishnan and Nagaraja [5]. The record values from the exponential distribution and the best linear unbiased estimators of the location and scale parameters based on them have been discussed by Ahsanullah [1]. The prediction of future record values has been discussed for the exponential case by Dunsmore [13]. Some work of this nature has been carried out for the extreme value distribution by Nagaraja [16, 17] and Ahsanullah [3], and for the normal distribution by Balakrishnan and Chan [8].

In this paper we consider the upper record values from a Rayleigh population and derive explicit expressions for the means, variances and covariances. We also establish some simple recurrence relationships for the single and product moments. These results are similar to those established recently by Balakrishnan and Ahsanullah [6] and Balakrishnan, Ahsanullah and Chan [7] for the exponential and Gumbel distributions, respectively. The latter problem has been treated exhaustively in the order statistics context by Barnett and Lewis [10]. Next, we derive explicitly the BLUEs for the parameters in the one- and two-parameter models. The BLUEs are then used to develop prediction intervals for the future record values and also a test for spuriosity of the record value just observed. Next, we consider the data set given by Dunsmore [13] and also a simulated data set and illustrate all the necessary formulae in explicit algebraic forms. Finally, we extend all these results to the Weibull distribution and present the necessary explicit algebraic formulae.

2 Record values and properties

Let $X_{U(1)}, X_{U(2)}, \cdots$ be the upper record values arising from a sequence $\{X_i\}$ of i.i.d. Rayleigh variables with pdf

$$f(x) = xe^{-x^2/2}, \quad x > 0 \tag{2.1}$$

and distribution function

$$F(x) = 1 - e^{-x^2/2}, \quad x > 0. \tag{2.2}$$

Then it is known that the pdf of the nth upper record value $X_{U(n)}$ is given by

$$f_n(x) = \frac{1}{\Gamma(n)}\{-log\,[1-F(x)]\}^{n-1} f(x), \quad x>0,\ n=1,2,\ldots \tag{2.3}$$

and that the joint density function of $X_{U(m)}$ and $X_{U(n)}$ is given by

$$f_{m,n}(x,y) = \frac{1}{\Gamma(m)\Gamma(n-m)}\left\{-\log\left[1-F(x)\right]\right\}^{m-1} \times\frac{f(x)}{1-F(x)}\left\{-\log\left[1-F(y)\right] + \log\left[1-F(x)\right]\right\}^{n-m-1} f(y),$$
$$0<x<y<\infty,\ m=1,2,\ldots,\ m<n. \tag{2.4}$$

Let us denote $E(X^k_{U(n)})$ by $\alpha^{(k)}_n$, $\mathrm{Var}(X_{U(n)})$ by $\beta_{n,n}$, $E(X^k_{U(m)}, X^l_{U(n)})$ by $\alpha^{(k,l)}_{m,n}$, and $\mathrm{Cov}(X_{U(m)}, X_{U(n)})$ by $\beta_{m,n}$. For convenience, we will also use α_n for $\alpha^{(1)}_n$ and $\alpha_{m,n}$ for $\alpha^{(1,1)}_{m,n}$. We then have the following theorems.

Theorem 1 *For $n=1,2,\ldots$, and $k=1,2,\ldots$*

$$\alpha^{(k)}_n = 2^{k/2}\frac{\Gamma(n+\frac{k}{2})}{\Gamma(n)} \tag{2.5}$$

and for $1\le m<n$

$$\alpha_{m,n} = 2\frac{\Gamma(m+\frac{1}{2})}{\Gamma(m)}\frac{\Gamma(n+1)}{\Gamma(n+\frac{1}{2})}; \tag{2.6}$$

consequently,

$$E(X_{U(n)}) = \sqrt{2}\left\{\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}\right\}, \tag{2.7}$$

$$\mathrm{Var}(X_{U(n)}) = 2\left\{n-\left(\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}\right)^2\right\} \tag{2.8}$$

and

$$\mathrm{Cov}(X_{U(m)}, X_{U(n)}) = 2\frac{\Gamma(m+\frac{1}{2})}{\Gamma(m)} \times\left\{\frac{\Gamma(n+1)}{\Gamma(n+\frac{1}{2})} - \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}\right\}. \tag{2.9}$$

Proof: From (2.3), (2.1) and (2.2), we have

$$\begin{aligned}\alpha^{(k)}_n &= \int_0^\infty x^k\left\{-\log\left[1-F(x)\right]\right\}^{n-1} f(x)dx\\ &= \frac{1}{\Gamma(n)}\int_0^\infty x^k\left(\frac{x^2}{2}\right)^{n-1} e^{-x^2/2}x\,dx\\ &= \frac{1}{\Gamma(n)}2^{k/2}\int_0^\infty u^{k/2}u^{n-1}e^{-u}du \quad (\text{with } u=x^2/2)\\ &= 2^{k/2}\frac{\Gamma(n+\frac{k}{2})}{\Gamma(n)}.\end{aligned}$$

Next, from (2.4) we have for $1\le m<n$

$$\begin{aligned}\alpha_{m,n} &= \frac{1}{\Gamma(m)\Gamma(n-m)}\int_0^\infty\int_0^y xy\left(\frac{x^2}{2}\right)^{m-1} x\left(\frac{y^2}{2}-\frac{x^2}{2}\right)^{n-m-1} ye^{-y^2/2}dx\,dy\\ &= \frac{1}{2^{m-1}\Gamma(m)\Gamma(n-m)}\int_0^\infty y^2e^{-y^2/2}\left(\frac{y^2}{2}\right)^{n-m-1} I(y)dy,\end{aligned} \tag{2.10}$$

where

$$I(y) = \int_0^y (x^2)^m\left(1-\frac{x^2}{y^2}\right)^{n-m-1}dx. \tag{2.11}$$

By setting $u=x^2/y^2$, (2.11) becomes

$$\begin{aligned}I(y) &= \int_0^1 u^m y^{2m}(1-u)^{n-m-1}\frac{y}{2\sqrt{u}}du\\ &= \frac{1}{2}y^{2m+1}B(m+\frac{1}{2},n-m),\end{aligned} \tag{2.12}$$

where $B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the complete beta function. Substituting the expression of $I(y)$ in (2.12) into (2.10), we have

$$\begin{aligned}\alpha_{m,n} &= \frac{B(m+\frac{1}{2},n-m)}{2^m\Gamma(m)\Gamma(n-m)}\int_0^\infty e^{-y^2/2}y^{2m+2}\left(\frac{y^2}{2}\right)^{n-m-1}ydy\\ &= \frac{B(m+\frac{1}{2},n-m)}{2^m\Gamma(m)\Gamma(n-m)}\int_0^\infty e^{-v}(2v)^{m+1}v^{n-m-1}dv \quad (\text{setting } v=y^2/2)\\ &= \frac{2B(m+\frac{1}{2},n-m)}{\Gamma(m)\Gamma(n-m)}\Gamma(n+1)\\ &= 2\frac{\Gamma(m+\frac{1}{2})\Gamma(n+1)}{\Gamma(m)\Gamma(n+\frac{1}{2})}.\end{aligned}$$

Formulae in (2.7)–(2.9) then readily follow. ■

For the Rayleigh distribution, it is easily observed that

$$xf(x) = 2\left\{-\log\left[1-F(x)\right]\right\}\left[1-F(x)\right]. \tag{2.13}$$

By using this relation, we establish below some simple recurrence relations satisfied by the single and product moments of record values. Similar results for the exponential and Gumbel populations are due to Balakrishnan and Ahsanullah [6] and Balakrishnan, Ahsanullah and Chan [7], respectively.

Theorem 2 *For $n = 1, 2, \ldots$, and $k = 1, 2, \ldots$,*

$$\alpha_{n+1}^{(k)} = \frac{k+2n}{2n}\alpha_n^{(k)}. \tag{2.14}$$

Proof: Let us consider

$$\begin{aligned}\alpha_n^{(k)} &= \frac{1}{\Gamma(n)}\int_0^\infty x^k \{-\log[1-F(x)]\}^{n-1} f(x)dx \\ &= \frac{2}{\Gamma(n)}\int_0^\infty x^{k-1}\{-\log[1-F(x)]\}^n \\ &\qquad [1-F(x)]\,dx \quad \text{(using (2.13))}.\end{aligned}$$

Upon integrating by parts, we obtain

$$\begin{aligned}\alpha_n^{(k)} &= \frac{2}{k\Gamma(n)}\left[\int_0^\infty x^k\left\{-\log\left[1-F(x)\right]\right\}^n f(x)dx\right. \\ &\qquad \left. -n\int_0^\infty x^k\left\{-\log\left[1-F(x)\right]\right\}^{n-1} f(x)dx\right] \\ &= \frac{2\Gamma(n+1)}{k\Gamma(n)}\left[\alpha_{n+1}^{(k)} - \alpha_n^{(k)}\right] \\ &= \frac{2n}{k}\left[\alpha_{n+1}^{(k)} - \alpha_n^{(k)}\right]. \end{aligned} \tag{2.15}$$

Then, (2.14) follows by rearranging (2.15). ■

Theorem 3 *For $k, l = 1, 2\ldots$, and $m \geq 1$*

$$\alpha_{m,m+1}^{(k,l)} = \frac{2m}{2m+k}\alpha_{m+1}^{(k+l)} \tag{2.16}$$

and for $1 \leq m \leq n-2$

$$\alpha_{m,n}^{(k,l)} = \frac{2m}{2m+k}\alpha_{m+1,n}^{(k,l)}. \tag{2.17}$$

Proof: For proving (2.16), let us consider from (2.4)

$$\begin{aligned}\alpha_{m,m+1}^{(k,l)} &= \int_0^\infty\int_0^y x^k y^l\left\{-\log\left[1-F(x)\right]\right\}^{m-1} \\ &\qquad \times\frac{f(x)}{1-F(x)}f(y)dxdy/\Gamma(m) \\ &= \frac{1}{\Gamma(m)}\int_0^\infty y^l f(y)I(y)dy, \end{aligned} \tag{2.18}$$

where

$$\begin{aligned}I(y) &= \int_0^y x^k\left\{-\log\left[1-F(x)\right]\right\}^{m-1}\frac{f(x)}{1-F(x)}dx \\ &= 2\int_0^y x^{k-1}\left\{-\log\left[1-F(x)\right]\right\}^m dx \\ &\qquad \text{(using (2.13))}. \end{aligned} \tag{2.19}$$

Upon integrating by parts, (2.19) yields

$$\begin{aligned}I(y) = \frac{2}{k}\Bigg[&y^k\left\{-\log\left[1-F(y)\right]\right\}^m \\ &-m\int_0^y x^k\left\{-\log\left[1-F(x)\right]\right\}^{m-1}\frac{f(x)}{1-F(x)}dx\Bigg].\end{aligned}$$

Substituting the above expression of $I(y)$ into (2.18), we obtain

$$\begin{aligned}\alpha_{m,m+1}^{(k,l)} &= \frac{2}{k}\left[\int_0^\infty y^{k+l}\left\{-\log\left[1-F(y)\right]\right\}^m f(y)dy\right. \\ &\qquad -m\int_0^\infty\int_0^y x^k y^l\left\{-\log\left[1-F(x)\right]\right\}^{m-1} \\ &\qquad \left.\frac{f(x)}{1-F(x)}f(y)dxdy\right] \Big/ \Gamma(m) \\ &= \frac{2m}{k}\left[\alpha_{m+1}^{(k+l)} - \alpha_{m,m+1}^{(k,l)}\right]. \end{aligned} \tag{2.20}$$

Relation in (2.16) is simply obtained by rearranging (2.20).

Next, for proving (2.17), let us write for $1 \leq m \leq n-2$

$$\alpha_{m,n}^{(k,l)} = \frac{1}{\Gamma(m)\Gamma(n-m)}\int_0^\infty y^l f(y)I(y)dy, \tag{2.21}$$

where

$$\begin{aligned}I(y) &= \int_0^y x^k\left\{-\log\left[1-F(x)\right]\right\}^{m-1}\frac{f(x)}{1-F(x)} \\ &\qquad \left\{-\log\left[1-F(y)\right] + \log\left[1-F(x)\right]\right\}^{n-m-1} dx \\ &= 2\int_0^y x^{k-1}\left\{-\log\left[1-F(x)\right]\right\}^m \\ &\qquad \left\{-\log\left[1-F(y)\right] + \left[1-F(x)\right]\right\}^{n-m-1} dx. \end{aligned}$$

Upon integrating by parts, $I(y)$ can be written as

$$\begin{aligned} I(y) &= \frac{2}{k}\Bigg[(n-m-1)\int_0^y x^k\Big\{-\log\Big[1-F(x)\Big]\Big\}^m \\ &\quad \frac{f(x)}{1-F(x)}\Big\{-\log\Big[1-F(y)\Big] \\ &\quad +\log\Big[1-F(x)\Big]\Big\}^{n-m-2} dx \\ &\quad -m\int_0^y x^k\Big\{-\log\Big[1-F(x)\Big]\Big\}^{m-1} \\ &\quad \frac{f(x)}{1-F(x)}\Big\{-\log\Big[1-F(y)\Big] \\ &\quad +\log\Big[1-F(x)\Big]\Big\}^{n-m-1} dx\Bigg]. \end{aligned} \tag{2.22}$$

Substituting the expression of $I(y)$ in (2.22) into Eq. (2.21), we obtain

$$\alpha_{m,n}^{(k,l)} = \frac{2m}{k}\Big[\alpha_{m+1,n}^{(k,l)} - \alpha_{m,n}^{(k,l)}\Big]. \tag{2.23}$$

Relation in (2.17) is simply obtained by rearranging (2.23). ■

If we interchange the order of the double integration in the proof of Theorem 3 and proceed along the same lines, we can derive the following relations.

Theorem 4 *For $k,l = 1,2,\ldots,$ and $m = 1,2,\ldots$*

$$\alpha_{m,m+2}^{(k,l)} = \frac{l+2}{2}\alpha_{m,m+1}^{(k,l)} - m\Big(\alpha_{m+1,m+2}^{(k,l)} - \alpha_{m+1}^{(k+l)}\Big), \tag{2.24}$$

and for $1 \le m \le n-2$

$$\begin{aligned} \alpha_{m,n+1}^{(k,l)} &= \frac{l+2(n-m)}{2(n-m)}\alpha_{m,n}^{(k,l)} \\ &\quad -\frac{m}{n-m}\Big(\alpha_{m+1,n+1}^{(k,l)} - \alpha_{m+1,n}^{(k,l)}\Big). \end{aligned} \tag{2.25}$$

3 Inference for the one-parameter Rayleigh distribution

Suppose the first n record values

$$Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$$

from a one-parameter Rayleigh distribution with pdf

$$f(y;\sigma) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right), \quad x>0,\ \sigma>0, \tag{3.1}$$

are available. Then, by following the generalized least-squares approach, we may derive the BLUE of σ (Balakrishnan and Cohen, [9], pp. 74) as

$$\sigma^* = \frac{\boldsymbol{\alpha}'\,\boldsymbol{\Omega}}{\boldsymbol{\alpha}'\,\boldsymbol{\Omega}\boldsymbol{\alpha}}\mathbf{Y} = \sum_{i=1}^{n} a_i Y_{U(i)} \tag{3.2}$$

and

$$Var(\sigma^*) = \frac{\sigma^2}{\boldsymbol{\alpha}'\,\boldsymbol{\Omega}\boldsymbol{\alpha}}, \tag{3.3}$$

where

$$\begin{aligned} \mathbf{Y} &= (Y_{U(1)}, Y_{U(2)}, \ldots, Y_{U(n)})', \\ \boldsymbol{\alpha} &= (\alpha_1, \alpha_2, \ldots, \alpha_n)' \end{aligned}$$

and

$$\boldsymbol{\Omega} = ((\beta_{i,j}))_{i,j=1,2,\ldots,n}^{-1} = ((\omega_{i,j}))_{i,j=1,2,\ldots,n}.$$

Since $\beta_{i,j}$ is in the form $p_i q_j$ for $i \le j$ where

$$p_i = \sqrt{2}\left\{\frac{\Gamma(i+\frac{1}{2})}{\Gamma(i)}\right\}$$

and

$$q_j = \sqrt{2}\left\{\frac{\Gamma(j+1)}{\Gamma(j+\frac{1}{2})} - \frac{\Gamma(j+\frac{1}{2})}{\Gamma(j)}\right\},$$

$\boldsymbol{\Omega}$ can be written explicitly as

$$\omega_{i,j} = \begin{cases} \frac{p_2}{p_1(p_2q_1 - p_1q_2)}, & i=j=1, \\ \frac{p_{i+1}q_{i-1} - p_{i-1}q_{i+1}}{(p_iq_{i-1} - p_{i-1}q_i)(p_{i+1}q_i - p_iq_{i+1})}, & i=j=2,3,\ldots,n-1, \\ \frac{q_{n-1}}{q_n(p_nq_{n-1} - p_{n-1}q_n)}, & i=j=n, \\ -\frac{1}{p_{i+1}q_i - p_iq_{i+1}}, & j=i+1,\ i=1,2,\ldots,n-1, \\ 0, & |j-i|\ge 2. \end{cases} \tag{3.4}$$

That is,

$$\begin{aligned} \omega_{1,1} &= \frac{9}{2}, \\ \omega_{i,i} &= \frac{8i^2+1}{2i}, \quad i=2,3,\ldots,n-1, \\ \omega_{n,n} &= (2n-1)\frac{q_{n-1}}{q_n}, \\ \omega_{i,i+1} &= -(2i+1), \quad i=1,2,\ldots,n-1, \\ \omega_{i+1,i} &= \omega_{i,i+1}, \quad i=1,2,\cdots,n-1, \\ \omega_{i,j} &= 0, \quad |j-i|>2. \end{aligned}$$

Eqs. (3.2) and (3.3), when simplified, simply yield

$$\sigma^* = \frac{Y_{U(n)}}{\alpha_n} \tag{3.5}$$

and

$$\mathrm{Var}(\sigma^*) = \sigma^2 \left[\frac{n\Gamma^2(n)}{\Gamma^2(n+\frac{1}{2})} - 1 \right]. \qquad (3.6)$$

Prediction of the future record: Suppose the upper records $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(m)}, m = 1, 2, \ldots$ have been observed. Then the best linear unbiased predicted value of the record $Y_{U(n)}, n \geq m+1$, can be written as

$$Y^*_{U(n)} = \sigma^* \alpha_n \qquad (3.7)$$

where σ^* is the BLUE of σ based on m records. Instead of the predicted value, one might be interested in the prediction interval for $Y_{U(n)}$ with a certain confidence. The prediction interval for $Y_{U(n)}$ may be based on the scale-invariant statistic

$$T^p_{1n} = \frac{Y_{U(n)} - Y_{U(m)}}{\sigma^*} \qquad (3.8)$$

where σ^* once again is the BLUE of σ based on the first m records. Using (3.5), we can rewrite the statistic T^p_{1n} in (3.8) as

$$\begin{aligned} T^p_{1n} &= \alpha_m \left\{ Y_{U(n)} - Y_{U(m)} \right\} / Y_{U(m)} \\ &= \alpha_m \left\{ \frac{Y_{U(n)}}{Y_{U(m)}} - 1 \right\}. \end{aligned}$$

Since $Y^2/2\sigma^2$ is distributed as a standard exponential variable, we can easily show that $Y^2_{U(m)}/Y^2_{U(n)}$ is distributed as a beta $(m, n-m)$ variate (for example, see Dunsmore, [13]). Thence, the 100% prediction interval for $Y_{U(n)}$ is obtained to be

$$\left[Y_{U(m)}, Y_{U(m)}/\sqrt{b_\alpha} \right]$$

where b_α is the lower α percentage point of the beta $(m, n-m)$ distribution.

Test of spuriosity of the current record: Suppose $Y_{U(1)}, \cdots, Y_{U(n-1)}$ have been observed and a new record $Y_{U(n)}$ has just been observed. Sometimes we may be interested in testing for the spuriosity of the current record value $Y_{U(n)}$. For this purpose, we may use the scale invariant statistic

$$T^o_{1n} = \frac{Y_{U(n)}}{\sigma^*} \qquad (3.9)$$

where σ^* is the BLUE of σ based on the first $n-1$ records; using (3.5), therefore, T^o_{1n} in (3.9) becomes

$$T^o_{1n} = \alpha_{n-1} \frac{Y_{U(n)}}{Y_{U(n-1)}}. \qquad (3.10)$$

At α level of significance, we will conclude the current record $Y_{U(n)}$ to be a spurious record if $Y_{U(n)}$ is greater than $Y_{U(n-1)}/\sqrt{b_\alpha}$, where b_α is the lower α percentage point of the beta $(n-1, 1)$ distribution, namely, $b_\alpha = \alpha^{1/(n-1)}$.

4 Inference for the two-parameter Rayleigh distribution

Let us now suppose that the first n upper record values $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$ from a two-parameter Rayleigh distribution with pdf

$$f(y; \mu, \sigma) = \frac{y-\mu}{\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \quad \mu < y < \infty, \ \sigma > 0, \qquad (4.1)$$

are available. Once again, by following the generalized least-squares approach, we may derive the BLUEs of μ and σ (David, [12], pp. 130; Balakrishnan and Cohen, [9], pp. 80–81) as

$$\mu^* = \sum_{i=1}^{n} a_i Y_{U(i)} \quad \text{and} \quad \sigma^* = \sum_{i=1}^{n} b_i Y_{U(i)}, \qquad (4.2)$$

where

$$\mathrm{a} = \frac{\alpha' \, \Omega \alpha 1' \, \Omega - \alpha' \, \Omega 1 \alpha' \, \Omega}{(\alpha' \, \Omega \alpha)(1' \, \Omega 1) - (\alpha' \, \Omega 1)^2} \qquad (4.3)$$

and

$$\mathrm{b} = \frac{1' \, \Omega 1 \alpha' \, \Omega - 1' \, \Omega \alpha 1' \, \Omega}{(\alpha' \, \Omega \alpha)(1' \, \Omega 1) - (\alpha' \, \Omega 1)^2}, \qquad (4.4)$$

with

$$1' = (1 \, 1 \cdots 1)_{1 \times n}$$

and α and Ω as defined in the last section. The variances and covariance of the above estimators are (David, [12], pp. 130; Balakrishnan and Cohen, [9], pp. 81) given by

$$\frac{\mathrm{Var}(\mu^*)}{\sigma^2} = \frac{\alpha' \, \Omega \alpha}{(\alpha' \, \Omega \alpha)(1' \, \Omega 1) - (\alpha' \, \Omega 1)^2}, \qquad (4.5)$$

$$\frac{\mathrm{Var}(\sigma^*)}{\sigma^2} = \frac{1' \, \Omega 1}{(\alpha' \, \Omega \alpha)(1' \, \Omega 1) - (\alpha' \, \Omega 1)^2}, \qquad (4.6)$$

and

$$\frac{\mathrm{Cov}(\mu^*, \sigma^*)}{\sigma^2} = -\frac{\alpha' \, \Omega 1}{(\alpha' \, \Omega \alpha)(1' \, \Omega 1) - (\alpha' \, \Omega 1)^2}. \qquad (4.7)$$

Omitting the intervening algebra, we get the coefficients of the BLUEs of μ and σ from (4.3) and (4.4) to be

$$\begin{aligned} a_1 &= \frac{3}{2} \frac{\alpha_n q_n}{\alpha_n q_n \Delta - 1}, \\ a_i &= \frac{\alpha_n q_n}{(2i)(\alpha_n q_n \Delta - 1)}, \ i = 2, 3, \ldots, n-1 \\ a_n &= 1 - \frac{\alpha_n q_n}{2(\alpha_n q_n \Delta - 1)} \left[3 + \sum_{i=2}^{n-1} \frac{1}{i} \right], \end{aligned}$$

and

$$b_1 = -\frac{3}{2}\frac{q_n}{\alpha_n q_n \Delta - 1},$$

$$b_i = -\frac{1}{2i}\frac{q_n}{\alpha_n q_n \Delta - 1},\ i = 2, 3, \ldots, n-1,$$

$$b_n = \frac{q_n}{2(\alpha_n q_n \Delta - 1)}\left\{3 + \sum_{i=2}^{n-1}\frac{1}{i}\right\},$$

where

$$\Delta = \left\{\frac{3}{2} + \sum_{i=2}^{n-1}\frac{1}{2i} + (2n-1)\left[\frac{q_{n-1}}{q_n} - 1\right]\right\}.$$

The variances and covariance of these estimators are obtained from (4.5)–(4.7) to be

$$\frac{\mathrm{Var}(\mu^*)}{\sigma^2} = \frac{\alpha_n q_n}{\alpha_n q_n \Delta - 1},$$

$$\frac{\mathrm{Var}(\sigma^*)}{\sigma^2} = \frac{q_n^2 \Delta}{\alpha_n q_n \Delta - 1}$$

and

$$\frac{\mathrm{Cov}(\mu^*, \sigma^*)}{\sigma^2} = -\frac{q_n}{\alpha_n q_n \Delta - 1}.$$

Prediction of the future record: Suppose the upper records $Y_{U(1)}, Y_{U(2)}, \ldots, Y_{U(n-1)}, (n \geq 3)$ have been observed. Then the best linear unbiased predicted value of the record $Y_{U(n)}$ can be written as

$$Y^*_{U(n)} = \mu^* + \sigma^* \alpha_n, \tag{4.8}$$

where μ^* and σ^* are the BLUEs of μ and σ based on the first $n-1$ record values.

Suppose we are interested in giving an interval for $Y_{U(n)}$ with a certain confidence. This prediction interval for $Y_{U(n)}$ may be based on the location and scale invariant statistic

$$T^p_{2n} = \frac{Y_{U(n)} - Y_{U(n-1)}}{\sigma^*} \tag{4.9}$$

where σ^* is once again the BLUE of σ based on the first $n-1$ upper record values. For aiding the users, we have determined some percentage points of the statistic T^p_{2n} in (4.9) through Monte Carlo simulations (based on 10,001 runs). These simulated percentage points of T^p_{2n} are presented in Table 1 for $n = 3(1)11$. With the help of this table, one could easily construct $100(1-\alpha)\%$ prediction intervals for the future record value $Y_{U(n)}$.

Testing for spuriosity of the current record: Suppose the upper records $Y_{U(1)}, Y_{U(2)}, \ldots, Y_{U(n-1)}$ have been observed, and a new record $Y_{U(n)}$ has just been observed. We may sometimes be interested in testing for the spuriosity of the current record value $Y_{U(n)}$. For this purpose, we may use the location and scale invariant statistic

$$T^o_{2n} = \frac{Y_{U(n)} - \mu^*}{\sigma^*}, \tag{4.10}$$

where μ^* and σ^* are once again the BLUEs of μ and σ based on the first $n-1$ records. Large values of T^o_{2n} will support the spuriosity of $Y_{U(n)}$.

For assisting the users, we have simulated some critical points of the statistic T^o_{2n} in (4.10) through Monte Carlo simulations (based on 10,001 runs). These simulated percentage points of T^o_{2n} are presented in Table 2 for $n = 3(1)11$. By using this table, one could easily test for the spuriosity of the current record value $Y_{U(n)}$ at the desired level of significance.

5 Illustrative Examples

Example 1: Dunsmore [13] has given the size of rock crushed by a rock crushing machine. The machine has to be reset if, at any operation, the size of rock being crushed is larger than that has been crushed before. The following are the records of the sizes dealt with up to the third time that the machine has been reset:

$$9.3, 24.4, 33.8.$$

Suppose for illustration that the sizes of the rocks to be crushed can be represented by independent one-parameter Rayleigh random variables. A simple plot of these three upper record values against the expected values of the Rayleigh upper record values indicates a strong correlation (correlation coefficient as high as 0.984). Hence, the assumption that these record values have come from the Rayleigh model seems quite reasonable. From formula (3.2), the BLUE of σ is simply obtained to be

$$\sigma^* = \frac{33.8}{\alpha_3} = 14.38.$$

The 90% prediction interval for the fourth upper record is obtained to be

$$\left[33.8, 33.8/\sqrt{0.1^{1/3}}\right] = \left[33.8, 49.6\right].$$

Furthermore, if the fourth upper record has been observed and it is not in [33.8, 49.6], we can conclude

that the observed record is a spurious record at 10% level of significance.

Now, suppose the upper records are assumed to have come from the two-parameter Rayleigh model. We then compute the BLUEs of μ and σ to be

$$\begin{aligned}\mu^* &= (2.000 \times 9.3) + (0.333 \times 24.4) - (1.333 \times 33.8) \\ &= -18.33\end{aligned}$$

and

$$\begin{aligned}\sigma^* &= -(0.8511 \times 9.3) - (0.1418 \times 24.2) \\ &\quad +(0.9929 \times 33.8) \\ &= 22.21.\end{aligned}$$

From Table 1, we determine the 90% one-sided prediction interval for the fourth upper record to be

$$\Big[33.8, 69.1\Big].$$

If the fourth record has been observed and it is greater than $\mu^* + \sigma^* \times 3.93 = 68.96$, we can conclude that the record is a spurious record at 10% level of significance.

It is of interest to mention here that the above given prediction interval [33.8, 69.1] is very close to the prediction interval based on the exponential distribution given by Dunsmore [13]. This is not surprising as the Rayleigh distribution is also seen to fit the data very well.

Example 2: For the purpose of illustration, we simulated a set of record values from the Rayleigh distribution with $\mu = 50$ and $\sigma = 10$. The following are the simulated upper record values:

$$66.42, 72.27, 78.07, 81.82, 86.33, 87.42, 90.05.$$

The BLUEs of μ and σ are computed in this case to be

$$\begin{aligned}\mu^* &= 1.2245 \times 66.24 + 0.2041 \times 72.27 \\ &\quad +0.1361 \times 78.07 + 0.1020 \times 81.82 \\ &\quad +0.0817 \times 86.33 + 0.0680 \times 87.42 \\ &\quad -0.8163 \times 90.05 \\ &= 54.54\end{aligned}$$

and

$$\begin{aligned}\sigma^* &= -0.3332 \times 66.24 - 0.0555 \times 72.27 \\ &\quad -0.0370 \times 78.07 - 0.0277 \times 81.82 \\ &\quad -0.0222 \times 86.33 - 0.0185 \times 87.42 \\ &\quad +0.4942 \times 90.05 \\ &= 9.67.\end{aligned}$$

The 90% one-sided prediction interval for the next upper record (eighth record) is obtained as

$$[90.05, 90.05 + 9.67 \times 0.733] = [90.05, 97.14].$$

6 Results for the Weibull Distribution

All results for the Rayleigh distributions developed in the previous sections can be extended to the Weibull distribution with pdf

$$f(x) = x^{c-1}e^{-x^c/c}, \quad x > 0,\ c > 0 \tag{6.1}$$

and cdf

$$F(x) = 1 - e^{-x^c/c}, \quad x > 0, c > 0, \tag{6.2}$$

where c is the known shape parameter. Analogous to Theorem 1, we then have the following theorem for the Weibull distribution.

Theorem 5 *For $n = 1, 2, \ldots,$ and $k = 1, 2, \ldots$*

$$\alpha_n^{(k)} = c^{k/c}\frac{\Gamma(n + \frac{k}{c})}{\Gamma(n)} \tag{6.3}$$

and for $1 \le m < n$

$$\alpha_{m,n} = c^{2/c}\frac{\Gamma(m + \frac{1}{c})}{\Gamma(m)}\frac{\Gamma(n + \frac{2}{c})}{\Gamma(n + \frac{1}{c})}; \tag{6.4}$$

consequently,

$$E(X_{U(n)}) = c^{1/c}\left\{\frac{\Gamma(n + \frac{1}{c})}{\Gamma(n)}\right\}, \tag{6.5}$$

$$\mathrm{Var}(X_{U(n)}) = c^{2/c}\left\{\frac{\Gamma(n + \frac{2}{c})}{\Gamma(n)} - \left(\frac{\Gamma(n + \frac{1}{c})}{\Gamma(n)}\right)^2\right\} \tag{6.6}$$

and

$$\mathrm{Cov}(X_{U(m)}, X_{U(n)}) = c^{2/c}\frac{\Gamma(m + \frac{1}{c})}{\Gamma(m)}\left\{\frac{\Gamma(n + \frac{2}{c})}{\Gamma(n + \frac{1}{c})} - \frac{\Gamma(n + \frac{1}{c})}{\Gamma(n)}\right\}. \tag{6.7}$$

For the Weibull distribution, the following relation is easily observed:

$$xf(x) = c\Big\{-\log\Big[1 - F(x)\Big]\Big\}\Big[1 - F(x)\Big]. \tag{6.8}$$

By using (6.8) and proceeding along the same lines as in Theorems 2–4, we can establish the following recurrence relations for the Weibull distribution.

Theorem 6 *For $n = 1, 2, \ldots$, and $k = 1, 2, \ldots$*

$$\alpha_{n+1}^{(k)} = \frac{k+nc}{nc}\alpha_n^{(k)}. \tag{6.9}$$

Theorem 7 *For $k, l = 1, 2 \ldots$, and $m = 1, 2, \ldots$*

$$\alpha_{m,m+1}^{(k,l)} = \frac{mc}{k+mc}\alpha_{m+1}^{(k+l)} \tag{6.10}$$

and for $1 \le m \le n-2$

$$\alpha_{m,n}^{(k,l)} = \frac{mc}{k+mc}\alpha_{m+1,n}^{(k,l)}. \tag{6.11}$$

Theorem 8 *For $k, l = 1, 2, \ldots$, and $m \ge 1$*

$$\alpha_{m,m+2}^{(k,l)} = \frac{l+c}{c}\alpha_{m,m+1}^{(k,l)} - m\left(\alpha_{m+1,m+2}^{(k,l)} - \alpha_{m+1}^{(k+l)}\right) \tag{6.12}$$

and for $1 \le m \le n-2$

$$\begin{aligned}\alpha_{m,n+1}^{(k,l)} &= \frac{l+c(n-m)}{c(n-m)}\alpha_{m,n}^{(k,l)} \\ &\quad - \frac{m}{n-m}\left(\alpha_{m+1,n+1}^{(k,l)} - \alpha_{m+1,n}^{(k,l)}\right).\end{aligned} \tag{6.13}$$

From Theorem 5, we observe that the variance-covariance matrix $((\beta_{i,j}))$ of the upper record values can be written as $p_i q_j, i \le j$, where

$$p_i = c^{1/c}\frac{\Gamma(i+\frac{1}{c})}{\Gamma(i)}$$

and

$$q_j = c^{1/c}\left\{\frac{\Gamma(j+\frac{2}{c})}{\Gamma(j+\frac{1}{c})} - \frac{\Gamma(j+\frac{1}{c})}{\Gamma(j)}.\right\}$$

Therefore, using (3.4), the inverse of $((\beta_{i,j}))$ can be explicitly written as

$$\begin{aligned}\omega_{1,1} &= c^{-2/c}\frac{(c+1)^2}{\Gamma(1+\frac{2}{c})}, \\ \omega_{i,i} &= c^{-2/c}\frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})} \\ &\quad \left[c^2(2i^2-2i+1) + c(4i-2) + 1\right], \\ &\quad i = 2, 3, \ldots, n-1, \\ \omega_{n,n} &= c^{-2/c}\frac{\Gamma(n)}{\Gamma(n+\frac{2}{c})}\frac{q_{n-1}}{q_n} \\ &\quad \times \left[(nc-c+1)(nc-c+2)\right], \\ \omega_{i,i+1} &= -c^{-2/c}\frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})}ic(ic+1), \\ &\quad i = 1, 2, \ldots, n-1, \\ \omega_{i,j} &= 0, \quad |j-i| > 2.\end{aligned}$$

One-parameter Weibull model: Suppose the first n upper record values $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$ from a one-parameter Weibull distribution with pdf

$$f(y;\sigma) = \frac{y^{c-1}}{\sigma^c}\exp\left(-\frac{y^c}{c\sigma^c}\right), \; y > 0, \; \sigma > 0, \tag{6.14}$$

are available. The BLUE of σ can be derived, using (3.2), as

$$\sigma^* = \frac{Y_{U(n)}}{\alpha_n} \tag{6.15}$$

with its variance as

$$\mathrm{Var}(\sigma^*) = \sigma^2\left[\frac{\Gamma(n)\Gamma(n+\frac{2}{c})}{\Gamma^2(n+\frac{1}{c})} - 1\right]. \tag{6.16}$$

Two-parameter Weibull model: Suppose the first n upper record values $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$ from a two-parameter Weibull distribution with pdf

$$\begin{aligned}f(y;\mu,\sigma) &= \frac{(y-\mu)^{c-1}}{\sigma^c}\exp\left(-\frac{1}{c\sigma^c}(y-\mu)^c\right), \\ &\quad \mu < y < \infty, \sigma > 0,\end{aligned} \tag{6.17}$$

are available. The BLUEs of μ and σ can be derived, using (4.3) and (4.4), as

$$\mu^* = \sum_{i=1}^{n} a_i Y_{U(i)} \tag{6.18}$$

and

$$\sigma^* = \sum_{i=1}^{n} b_i Y_{U(i)}, \tag{6.19}$$

where

$$\begin{aligned}a_1 &= \frac{\alpha_n q_n}{\alpha_n q_n - 1}\frac{c^{-2/c}(c+1)}{\Gamma(1+\frac{1}{c})}, \\ a_i &= \frac{\alpha_n q_n c^{-2/c}(c-1)}{\alpha_n q_n \Delta - 1}\frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})}, \\ &\quad i = 2, 3, \cdots, n-1, \\ a_n &= 1 - \frac{\alpha_n q_n c^{-2/c}}{\alpha_n q_n \Delta - 1} \\ &\quad \times \left[\frac{(c+1)}{\Gamma(1+\frac{2}{c})} + (c-1)\sum_{i=2}^{n-1}\frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})}\right],\end{aligned}$$

and

$$\begin{aligned}b_1 &= -\frac{c^{-2/c}q_n(c+1)}{(\alpha_n q_n \Delta - 1)\Gamma(1+\frac{2}{c})}, \\ b_i &= -\frac{c^{-2/c}q_n(c-1)}{\alpha_n q_n \Delta - 1}\frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})}, \; i = 2, 3, \cdots, n-1,\end{aligned}$$

$$b_n = \frac{c^{-2/c} q_n}{\alpha_n q_n \Delta - 1} \left(\frac{c+1}{\Gamma(1+\frac{2}{c})} + (c-1) \sum_{i=2}^{n-1} \frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})} \right),$$

with

$$\Delta = c^{-2/c} \left\{ \frac{c+1}{\Gamma(1+\frac{2}{c})} + (c-1) \sum_{i=2}^{n-1} \frac{\Gamma(i)}{\Gamma(i+\frac{2}{c})} + \frac{\Gamma(n)}{\Gamma(n+\frac{2}{c})} (nc-c+1)(nc-c+2) \left[\frac{q_{n-1}}{q_n} - 1 \right] \right\}.$$

The variances and covariance of these estimators are obtained from (4.5)–(4.7) to be

$$\frac{\mathrm{Var}(\mu^*)}{\sigma^2} = \frac{\alpha_n q_n}{\alpha_n q_n \Delta_n - 1},$$
$$\frac{\mathrm{Var}(\sigma^*)}{\sigma^2} = \frac{q_n^2 \Delta}{\alpha_n q_n \Delta - 1},$$

and

$$\frac{\mathrm{Cov}(\mu^*, \sigma^*)}{\sigma^2} = -\frac{q_n}{\alpha_n q_n \Delta - 1}.$$

Inference procedures for the Weibull distribution can be developed along the lines of Section 3 and 4 by making use of the explicit forms of the BLUEs presented above. For brevity, we have not pursued this here.

Acknowledgments

The first author would like to thank the Natural Sciences and Engineering Research Council of Canada for funding this research.

References

[1] Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Ann. Inst. Statist. Math.* **32**, 363–368.

[2] Ahsanullah, M. (1988). *Introduction to Record Values*, Ginn Press, Needham Heights, Massachusetts.

[3] Ahsanullah, M. (1990). Estimation of the parameters of the Gumbel distribution based on the m record values, *Comput. Statist. Quart.* **3**, 231–239.

[4] Arnold, B. C. and N. Balakrishnan (1989). *Relations, Bounds and Approximations for Order Statistics*, Lecture Notes in Statistics 53, Springer-Verlag, New York.

[5] Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (1992). *A First Course in Order Statistics*, John Wiley & Sons, New York.

[6] Balakrishnan, N. and M. Ahsanullah (1993). Relations for single and product moments of record values from exponential distribution, *J. Appl. Statist. Sci.* (To appear).

[7] Balakrishnan, N., M. Ahsanullah and P. S. Chan (1992). Relations for single and product moments of record values from Gumbel distribution, *Statist. Probab. Lett.* **15**, 223–227.

[8] Balakrishnan, N. and P. S. Chan (1993). On the normal record values and associated inference, *Submitted for publication.*

[9] Balakrishnan, N. and A. C. Cohen (1991). *Order Statistics and Inference: Estimation Methods*, Academic Press, Boston.

[10] Barnett, V. and T. Lewis (1984). *Outliers in Statistical Data*, Second edition, John Wiley & Sons, New York.

[11] Chandler, K. N. (1952). The distribution and frequency of record values, *J. Roy. Statist. Soc.*, Ser. B **14**, 220–228.

[12] David, H. A. (1981). *Order Statistics*, Second edition, John Wiley & Sons, New York.

[13] Dunsmore, I. R. (1983). The future occurrence of records, *Ann. Inst. Statist. Math.* **35**, 267–277.

[14] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, John Wiley & Sons, New York. (1987). Second edition, Krieger, Florida.

[15] Glick, N. (1978). Breaking records and breaking boards, *Amer. Math. Monthly* **85**, 2–26.

[16] Nagaraja, H. N. (1982). Record values and extreme value distributions, *J. Appl. Probab.* **19**, 233–239.

[17] Nagaraja, H. N. (1984). Asymptotic linear prediction of extreme order statistics, *Ann. Inst. Statist. Math.* **36**, 289–299.

[18] Nagaraja, H. N. (1988). Record values and related statistics – A review, *Commun. Statist. – Theor. Meth.* **17**(7), 2223-2238.

[19] Nevzorov, V. B. (1987). Records, *Teoriya Veroyatnostei i ee Premeneniya* (Theory of Probability and Its Applications) **32**(2), 219–251 (in Russian).

[20] Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*, Springer-Verlag, New York.

Table 1. Simulated percentage points of the statistic T_n^p for the two-parameter Rayleigh distribution

n	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
3	0.0048	0.0123	0.0238	0.0520	3.7321	7.6296	14.8732	38.4384
4	0.0041	0.0098	0.0199	0.0413	1.5834	2.4319	3.7907	6.5884
5	0.0033	0.0086	0.0177	0.0380	1.1380	1.6498	2.2644	3.3609
6	0.0034	0.0085	0.0166	0.0332	0.9045	1.2637	1.6990	2.3337
7	0.0032	0.0071	0.0137	0.0273	0.7326	0.9934	1.2950	1.7487

Table 2. Simulated percentage points of the test statistic T_n^o for the two-parameter Rayleigh distribution

n	0.900	0.950	0.975	0.990	0.995
3	5.6121	9.5095	16.7532	40.3184	75.8375
4	3.9335	4.7821	6.1409	8.9401	12.6876
5	3.8797	4.3915	5.0061	6.1027	7.0952
6	3.9888	4.3480	4.7833	5.4181	6.1317
7	4.1750	4.4947	4.8057	5.2405	5.7279

On The Estimation Of The Pareto Tail-Index Using k-Record Values

Berred, A.M.
Universitè du Havre, Le Havre Cedex, France

Let $\{X_n,\ n \geq 1\}$ be an i.i.d. sequence of positive random variables with a continuous distribution function F having a regularly varying upper tail. In this paper we consider the k-record versions of two statistics introduced in [1] and study their asymptotic behavior. Such statistics can be used as alternative estimates for the exponent of regular variation of the tail $1 - F$.

1. Introduction

Let $\{X_n,\ n \geq 1\}$ be an i.i.d. sequence of positive random variables having a continuous distribution function F with regularly varying upper tail. Namely

(F) $1 - F(x) = x^{-1/\alpha} L(x)$, for $x \geq a > 0$,

where $\alpha > 0$ and L is a slowly varying function at infinity. Denote by $X_{1,n} \leq \cdots \leq X_{n,n}$ the order statistics associated to the sample $X_1, \ldots, X_n$. Let k be a positive integer and define the sequences of the *k-record times* and *values* (Ref. [2], [3]) by

$$\begin{aligned} \tau^{(k)}(1) &= k, \\ \tau^{(k)}(i) &= \min\{j > \tau^{(k)}(i-1), \\ &\qquad R(j) \geq j - k + 1\}, \\ X^{(k)}(i) &= X^{(k)}_{\tau^{(k)}(i)-k+1,\tau^{(k)}(i)},\ i \geq 1, \end{aligned}$$

where $R(n)$ is the sequential rank of X_n in the sample $X_1, \ldots, X_n$, i.e., $X_n = X_{R(n),n}$, for $n \geq 1$ (for the general theory of records Ref., e.g., chap. 6 of [4], chap. 4 of [5], [6], [7], [8] and the references therein). We consider the k-record versions of two statistics (for other statistics based on record values Ref. [9], [10]) introduced in [1],

$$\alpha_n = \frac{k}{m}\left\{\log X^{(k)}(n) - \log X^{(k)}(n-m)\right\},$$

$$\beta_n = \frac{k}{n(m)} \sum_{i=1}^{m} \log X^{(k)}(n-i+1),$$

where $1 \leq k < n$, $1 \leq m < n$ and $n(m) = nm - m(m-1)/2$.

The statistics α_n and β_n can be used as estimates of α. For the estimation of α based on extreme values, Ref., e.g., [11], [12], [13], [14], [15], [16], [17], [18], [19], [20].

The paper is organized as follows. Section 2 is devoted to the main results concerning the consistency and the asymptotic normality of α_n and β_n. The proofs of these results are given in section 3. Section 4 contains some examples which illustrate the results of section 2. Finally, section 5 presents some numerical results concerning the behavior of the estimates α_n and β_n in practice.

2. Results

In the sequel, we will impose some assumptions on the increase of the sequences $\{m_n,\ n \geq 1\}$ and $\{k_n,\ n \geq 1\}$.

(M1) $m = m_n \longrightarrow \infty$ is a sequence of integers such that $1 \leq m < n$ and $m/n \longrightarrow 0$ as $n \longrightarrow \infty$,

(M2) $m = m_n \longrightarrow \infty$ is a sequence of integers

such that $1 \le m < n$ and $m/\log n \longrightarrow \infty$ with $m/n \longrightarrow 0$ as $n \longrightarrow \infty$.

(K) $k = k_n \longrightarrow \infty$ is a sequence of integers such that $1 \le k < n$ and $k/n \longrightarrow 0$ as $n \longrightarrow \infty$.

Whenever one of the assumption (M1), (M2) or (K) is made, we write k and m instead of k_n and m_n for the sake of simplicity.

The notations "$\stackrel{d}{=}$", "$\stackrel{d}{\longrightarrow}$", "$\stackrel{p}{\longrightarrow}$" and "$\longrightarrow$ a.s." stand respectively for equality in distribution, convergence in distribution, convergence in probability and almost sure convergence.

Denote by $h(x) = -\log(1 - F(x))$ the cumulative hazard function associated to F. The function h is nondecreasing on $(-\infty, +\infty)$, so that its inverse function may be defined by

$$H(x) = h^{\leftarrow}(x) = \inf\{t : h(t) \ge x\} \text{ on } (0, +\infty).$$

We now state our main results concerning the limiting behavior of the above statistics.

THEOREM 1. *Assume that* (F) *and* (K) *hold. Then*
(i) *under* (M1) *with* $k = O(m)$,

$$\alpha_n \stackrel{p}{\longrightarrow} \alpha \text{ as } n \longrightarrow \infty;$$

(ii) *under* (M2) *with* $k = O(m)$,

$$\alpha_n \longrightarrow \alpha \text{ a.s. as } n \longrightarrow \infty.$$

THEOREM 2. *Assume that* (F) *and* (K) *hold. Then for* $1 \le m < n$ *fixed or for* $m = m_n$ *satisfying* (M1),

$$\beta_n \longrightarrow \alpha \text{ a.s. as } n \longrightarrow \infty.$$

Furthermore we need the following assumptions on the slow variation of L to ensure the asymptotic normality of α_n and β_n.

(SR1) $\forall \lambda > 1,\ \dfrac{L(\lambda t)}{L(t)} - 1 = O(g(t))$ as $t \longrightarrow \infty$,

(SR2) $\forall \lambda > 1,\ \dfrac{L(\lambda t)}{L(t)} - 1 \sim K(\lambda) g(t)$ as $t \longrightarrow \infty$,

(SR3) $\forall \lambda > 1,\ \dfrac{L(\lambda t)}{L(t)} - 1 = o(g(t))$ as $t \longrightarrow \infty$,

where g is a positive function such that $g(t) \longrightarrow 0$ as $t \longrightarrow \infty$. The assumptions of slow variation with a remainder term (SR1-3) were introduced in [21] (for a general review on slow variation with a remainder and its applications Ref. [22]).

THEOREM 3. *Assume that* (F), (M1) *and* (K) *hold with* $k = O(m)$ *and* $m/\sqrt{n} \longrightarrow 0$. *Then if* L *is* (SR1-3) *with* g *nonincreasing and* $\sqrt{m} g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \stackrel{p}{\longrightarrow} 0$ *as* $n \longrightarrow \infty$, *then*

$$(2.1)\quad \frac{\sqrt{m}}{\alpha}(\alpha_n - \alpha) \stackrel{d}{\longrightarrow} \mathcal{N}(0,1) \text{ as } n \longrightarrow \infty.$$

THEOREM 4. *Assume that* (F) *and* (K) *hold. If* $1 \le m < n$ *and* $k/\sqrt{n} \longrightarrow 0$ $(n \longrightarrow \infty)$ *or* $m = m_n$ *satisfies* (M1) *with* $k = O(m)$ $(n \longrightarrow \infty)$, L *is* (SR1-3) *with* g *nonincreasing and* $\sqrt{n}\, g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \stackrel{p}{\longrightarrow} 0$ *as* $n \longrightarrow \infty$, *then*

$$(2.2)\quad \frac{\sqrt{n}}{\alpha}(\beta_n - \alpha) \stackrel{d}{\longrightarrow} \mathcal{N}(0,1) \text{ as } n \longrightarrow \infty.$$

3. Proofs

Let $\{e_n,\ n \ge 1\}$ be a sequence of i.i.d. unit exponential random variables and denote by $S_n = e_1 + \cdots + e_n$, $n \ge 1$, their partial sums. The following lemma gives a representation of the exponential k-record values in terms of the partial sums S_n, $n \ge 1$.

LEMMA 1. *Let* k *be a positive integer, then* $\{e^{(k)}(n),\ n \ge 1\} \stackrel{d}{=} \{S_n/k,\ n \ge 1\}$.

PROOF. The Theorem 1 in [3] (Ref. also Lemma 1 in [2]) implies that the sequence $\{e^{(k)}(n),\ n \ge 1\}$ is a Markov chain with transition probabilities

$$P\left(e^{(k)}(n+1) > x | e^{(k)}(n) = y\right) = \begin{cases} exp\{-k(x-y)\}, & x > y, \\ 1, & x \le y. \end{cases}$$

and initial probability

$$P\left(e^{(k)}(1) > x\right) = exp(-kx),\ x > 0.$$

Since the same holds for the sequence $\{S_n/k,\ n \ge 1\}$, it follows that

$$\left\{e^{(k)}(n),\ n \ge 1\right\} \stackrel{d}{=} \{S_n/k,\ n \ge 1\}.$$

This establishes our lemma. □

We may assume, without loss of generality, that the original probability space carries, in addition to the sequence $\{X_n,\ n \geq 1\}$, a sequence $\{e_n,\ n \geq 1\}$ of i.i.d. unit exponential random variables (since h is continuous, we may take $e_n = h(X_n)$). Let $\{k_n,\ n \geq 1\}$ be a sequence of positive integers such that $1 \leq k_n < n$, we consider the double array $\left\{X^{(k_j)}(i),\ i \geq 1, j \geq 1\right\}$ of k-records. Now by Lemma 1,

$$\left\{e^{(k_n)}(n-r),\ 0 \leq r < n\right\} \stackrel{d}{=} \left\{\frac{S_{n-r}}{k_n},\ 0 \leq r < n\right\}.$$

Since h is continuous it follows that H is strictly increasing and

$$\left\{X^{(k_n)}(n-r),\ 0 \leq r < n\right\} \stackrel{d}{=} \left\{H\left(\frac{S_{n-r}}{k_n}\right),\ 0 \leq r < n\right\}.$$

So from now on we shall assume without loss of generality that $X^{(k_n)}(n-r) = H(S_{n-r}/k_n)$, for $n \geq 1$ and $0 \leq r < n$.

we state two lemmas related to the slowly varying function L in (F).

LEMMA 2. *For every slowly varying function l*

$$\frac{\log l(x)}{\log x} \longrightarrow 0 \text{ as } x \longrightarrow \infty.$$

PROOF. Ref., e.g., [23], Proposition 1.3.6, p. 16. □

LEMMA 3. *Assume that* (F) *holds. Then*
(i) $\log H(x) = \alpha x + \log L'(e^x)$, *for* $0 < x < \infty$;
(ii) $\log L'(x) \sim \alpha \log L(R^{\leftarrow}(x))$

PROOF. Suppose that (F) holds. We prove (i). Set $R(x) = x^{\frac{1}{\alpha}}/L(x)$. Note that $h(x) = \log R(x)$. It follows that R is regularly varying with a positive exponent $\frac{1}{\alpha}$. Hence (Ref., e.g., [23], Theorem 1.5.12 p. 28) its generelized inverse $R^{\leftarrow}(x) = x^{\alpha}L'(x)$, where L' is a slowly varying function. Now $H(x) = R^{\leftarrow}(e^x)$ and

$$\log H(x) = \alpha x + \log L'(e^x), \text{ for } 0 < x < \infty.$$

We show (ii). Noting that $R(R^{\leftarrow}(x)) \sim x$ as $x \longrightarrow \infty$, we have $L'^{\frac{1}{\alpha}}(x)/L(R^{\leftarrow}(x)) \longrightarrow 1$ as $x \longrightarrow \infty$ and

$$\frac{\log L'(x)}{\alpha \log L(R^{\leftarrow}(x))} \longrightarrow 1,$$

as $x \longrightarrow \infty$. Our assertions are established. □

To simplify the proofs, we introduce some notations:

$$A_n^1 = \alpha \frac{S_n - S_{n-m}}{m},$$

$$A_n^2 = \frac{k}{m}\left\{\log L'(\exp \frac{S_n}{k}) - \log L'(\exp \frac{S_{n-m}}{k})\right\},$$

$$B_n^1 = \frac{\alpha}{n(m)} \sum_{i=1}^{m} S_{n-i+1},$$

$$B_n^2 = \frac{k}{n(m)} \sum_{i=1}^{m} \log L'(\exp \frac{S_{n-i+1}}{k}).$$

It follows by (i) of Lemma 3 that

$$\alpha_n = A_n^1 + A_n^2, \tag{3.1}$$

$$\beta_n = B_n^1 + B_n^2. \tag{3.2}$$

we implicitly make use of (3.1) and (3.2) in the proofs of Theorems 1–4.

PROOF OF THEOREM 1. Assume that (F) and (K) hold with $k = O(m)$. We first prove that $A_n^2 \xrightarrow{p} 0$ $(n \longrightarrow \infty)$ under (M1) and $A_n^2 \longrightarrow 0$ a.s. $(n \longrightarrow \infty)$ under (M2). The Karamata Representation Theorem for slowly varying functions (Ref., e.g., [23], Theorem 1.3.1, p. 12) implies that

$$A_n^2 = \frac{k}{m}\left\{\eta\left(\exp \frac{S_n}{k}\right) - \eta\left(\exp \frac{S_{n-m}}{k}\right)\right\} + \frac{k}{m}\int_{\exp \frac{S_{n-m}}{k}}^{\exp \frac{S_n}{k}} \frac{\epsilon(t)}{t} dt, \tag{3.3}$$

$$= \zeta_n + \xi_n,$$

where η is a bounded function on $[a, \infty)$ such that $\eta(x) \longrightarrow c$ $(x \longrightarrow \infty,\ |c| < \infty)$, and ϵ is a continuous function on $[a, \infty)$ such that $\epsilon(x) \longrightarrow 0$ as $x \longrightarrow \infty$. Under (M1), the law of large numbers yields

$$\frac{S_{n-m}}{m} \longrightarrow \infty \text{ a.s. as } n \longrightarrow \infty, \tag{3.4}$$

$$\frac{S_n}{m} \longrightarrow \infty \text{ a.s. as } n \longrightarrow \infty. \tag{3.5}$$

Hence $\zeta_n \longrightarrow 0$ a.s. as $n \longrightarrow \infty$. Therefore, it remains to show that $\xi_n \xrightarrow{p} 0$ under (M1) and $\xi_n \longrightarrow 0$ a.s. under (M2). write ξ_n as

$$|\xi_n| = \frac{k}{m}\left|\int_{\frac{S_{n-m}}{k}}^{\frac{S_n}{k}} \epsilon(e^t) dt\right|, \tag{3.6}$$

$$\leq \frac{S_n - S_{n-m}}{m} \sup_{\frac{S_{n-m}}{k} \leq t \leq \frac{S_n}{k}} \left|\epsilon(e^t)\right|.$$

Now (3.4) imply that $\sup_{\frac{S_{n-m}}{k} \le t \le \frac{S_n}{k}} |\epsilon(e^t)| \longrightarrow 0$, a.s. as $n \longrightarrow \infty$. Since

$$(3.7) \qquad \frac{S_n - S_{n-m}}{m} \stackrel{d}{=} \frac{S_m}{m}, \text{ for } 1 \le m < n,$$

the assumption (M1) and the law of large numbers imply again that

$$(3.8) \qquad \frac{S_n - S_{n-m}}{m} \xrightarrow{p} 1 \text{ as } n \longrightarrow \infty.$$

Denote by $\{\theta_n = e_n - 1,\ n \ge 1\}$ the centred sequence of $\{e_n,\ n \ge 1\}$. The moment generating function of θ_1 is

$$\begin{aligned} \Phi(t) &= E(\exp(t\theta_1), \\ &= \frac{\exp(-t)}{1-t}, \text{ for } t < 1. \end{aligned}$$

Hence the assumption (M2) and the Theorem 2.1 in [24] give

$$(3.9) \qquad \frac{S_n - S_{n-m}}{m} \longrightarrow 1 \text{ a.s. as } n \longrightarrow \infty.$$

Finally, relations (3.8) and (3.9) yield respectively (i) and (ii). The theorem is established. □

Proof of Theorem 2. Assume that (F) and (K) hold. We first suppose that $1 \le m < n$ is fixed. Note that $n(m) \sim nm$ $(n \longrightarrow \infty)$, the strong law of large numbers yields for $1 \le i \le m$,

$$(3.10) \qquad \frac{S_{n-i+1}}{n(m)} \longrightarrow \frac{1}{m} \text{ a.s. as } n \longrightarrow \infty.$$

Consequently

$$B_n^1 \longrightarrow \alpha \text{ a.s. as } n \longrightarrow \infty.$$

Therefore it remains to prove that $B_n^2 \longrightarrow 0$ a.s. as $n \longrightarrow \infty$. Since $1 \le m < n$ is fixed, this reduces to showing that

$$\frac{k \log L'(\exp \frac{S_{n-i+1}}{k})}{n(m)} \longrightarrow 0 \text{ a.s. as } n \longrightarrow \infty,$$

for $1 \le i \le m$. Rewriting the argument of the last limit as

$$(3.11) \qquad \frac{\log L'(\exp \frac{S_{n-i+1}}{k})}{\log \exp \frac{S_{n-i+1}}{k}} \frac{S_{n-i+1}}{n(m)},$$

the first term converges to 0 almost surely by Lemma 2, the second one converges to $\frac{1}{m}$ almost surely by (3.10).

Assume (M1). Since

$$B_n^1 = \frac{\alpha}{m} \sum_{i=1}^{m} \frac{m S_{n-i+1}}{n(m)},$$

is a Cezaro mean with the summand tending a.s. to 1 by (3.10), consequently $B_n^1 \longrightarrow \alpha$ a.s. as $n \longrightarrow \infty$. Now write

$$B_n^2 = \frac{1}{m} \sum_{i=1}^{m} \frac{\log L'(\exp \frac{S_{n-i+1}}{k})}{\log \exp \frac{S_{n-i+1}}{k}} \frac{m S_{n-i+1}}{n(m)}.$$

We see again that B_n^2 is a Cezaro mean with the summand converging to 0 by (3.11), it follows that $B_n^2 \longrightarrow 0$ a.s. as $n \longrightarrow \infty$. □

Proof of Theorem 3. Assume that (M1) holds. By (3.7) and the central limit theorem for i.i.d. random variables

$$\frac{\sqrt{m}}{\alpha} \left(A_n^1 - \alpha \right) \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \longrightarrow \infty.$$

Now assume that L is (SR1). It is well known that (Ref. [22], Theorem 2.2.2) $\log L$ has representation

$$\log L(x) = \eta(x) + \int_1^x O(g(t)) \frac{dt}{t} \text{ as } x \longrightarrow \infty,$$

where $\eta(t) = c + O(g(t)$ $(|c| < \infty, x \longrightarrow \infty)$. In view of (ii) of Lemma 3), $\log L'$ can be represented as

$$\log L'(x) = \eta'(x) + \int_1^{R^{-}(x)} O(g(t)) \frac{dt}{t} \text{ as } x \longrightarrow \infty,$$

where $\eta'(x) = c' + O(g(R^{-}(x)))$ $(|c'| < \infty, x \longrightarrow \infty)$. Hence

$$\begin{aligned} \log L'(e^x) &= c' + O(g(H(x))) + \int_1^{H(x)} O(g(t)) \frac{dt}{t}, \\ &= c' + O(g(H(x))) + \int_{h(1)}^{x} O(g(H(t))) d \log H(t), \end{aligned}$$

as $x \longrightarrow \infty$. It follows by (3.3) that

$$\zeta_n = \frac{k}{m}\left\{O\left[g\left(H\left(\frac{S_n}{k}\right)\right)\right] - O\left[g\left(H\left(\frac{S_{n-m}}{k}\right)\right)\right]\right\},$$

$$\xi_n = \frac{k}{m}\int_{\frac{S_{n-m}}{k}}^{\frac{S_n}{k}} O(g(H(t)))d\log H(t),$$

as $n \longrightarrow \infty$. Therefore our assertion will be proved if we show that $\sqrt{m}\ \zeta_n \xrightarrow{p} 0$ and $\sqrt{m}\ \xi_n \xrightarrow{p} 0$ as $n \longrightarrow \infty$. Since $k = O(m)$, g is nonincreasing and

$$\sqrt{m} g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \xrightarrow{p} 0,$$

we have $\sqrt{m}\zeta_n \xrightarrow{p} 0$ as $n \longrightarrow \infty$. Now we take care of $\sqrt{m}\ \xi_n$. We have

$$\sqrt{m}\ |\xi_n| \le O(1)\alpha_n \sup_{\frac{S_{n-m}}{k} \le t \le \frac{S_n}{k}} \sqrt{m}\ g(H(t)),$$

Under the assumption that g is nonincreasing, we obtain

$$\sqrt{m}\ |\xi_n| \le O(1)\alpha_n \sqrt{m}\ g(H(\frac{S_{n-m}}{k})),$$

as $n \longrightarrow \infty$. Now the assumption

$$\sqrt{m}\ g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \xrightarrow{p} 0, \text{ as } n \longrightarrow \infty,$$

and Theorem 1 imply that $\sqrt{m}\ \zeta_n \xrightarrow{p} 0$ as $n \longrightarrow \infty$. The proof in the cases (SR2–3) is similar to that under (SR1), We therefore omit it. □

To prove the asymptotic normality of β_n we approximate the main term in (3.2) by $\frac{S_n}{n}$.

LEMMA 4. *Assume that* (F) *holds. Then*
(i) *if* $1 \le m < n$ *is fixed*

$$B_n^1 = \frac{\alpha}{n}S_n + O_p\left(\frac{1}{n}\right) \text{ as } n \longrightarrow \infty;$$

(ii) *if* (M1) *holds*

$$B_n^1 = \frac{\alpha}{n}S_n + O_p\left(\frac{m}{n}\right) \text{ as } n \longrightarrow \infty.$$

PROOF. This amounts to evaluating for $1 \le i \le m < n$,

$$\left|\frac{S_n}{mn} - \frac{S_{n-i+1}}{n(m)}\right| \le \frac{S_n - S_{n-m}}{mn} + \frac{(m-1)}{2n\,n(m)}S_n = O_p\left(\frac{1}{n}\right) \text{ as } n \longrightarrow \infty,$$

by Markov's inequality. Hence

$$\left|B_n^1 - \frac{\alpha}{n}S_n\right| \le \alpha\sum_{i=1}^{m}\left|\frac{S_n}{mn} - \frac{S_{n-i+1}}{n(m)}\right|,$$
$$= \alpha\left(\frac{S_n - S_{n-m}}{n} + \frac{m(m-1)}{2n\,n(m)}S_n\right),$$
$$= O_p\left(\frac{m}{n}\right),$$
$$= O_p\left(\frac{1}{n}\right) \text{ if m is fixed },$$

as $n \longrightarrow \infty$. This completes the proof of the lemma. □

PROOF OF THEOREM 4. Applying Lemma 4 and the central limit theorem for i.i.d. random variables, we obtain for $1 \le m < n$ fixed or $m = m_n$ satisfying (M1) with $\sqrt{m}/n \longrightarrow 0$ as $n \longrightarrow \infty$,

$$\frac{\sqrt{n}}{\alpha}\left(B_n^1 - \alpha\right) \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n \longrightarrow \infty.$$

We will be done if we show that

$$\sqrt{n}\ B_n^2 \xrightarrow{p} 0 \text{ as } n \longrightarrow \infty. \tag{3.12}$$

Assume that L is (SR1), $1 \le m < n$ is fixed and $k/\sqrt{n} \longrightarrow 0$ $(n \longrightarrow \infty)$. Since $n(m) \sim nm$ $(n \longrightarrow \infty)$, it is sufficient to prove that

$$\frac{k}{\sqrt{n}}\log L'(\exp\frac{S_{n-i+1}}{k}) \xrightarrow{p} 0 \text{ as } n \longrightarrow \infty,$$

for $1 \le i < m$. In view (3.12) and the fact that g is nonincreasing, we have

$$\frac{k}{\sqrt{n}}\log L'(\exp\frac{S_{n-i+1}}{k}) =$$
$$\frac{k}{\sqrt{n}}\left\{c' + O\left[g\left(H\left(\frac{S_{n-i+1}}{k}\right)\right)\right]\right\} +$$
$$\frac{k}{\sqrt{n}}\int_{h(1)}^{\frac{S_{n-i+1}}{k}} O(g(H(t)))d\log H(t) \le$$
$$\frac{k}{\sqrt{n}}\left\{c' + O\left[g\left(H\left(\frac{S_{n-m}}{k}\right)\right)\right]\right\} +$$

$$O_p(1)\frac{1}{\log H(S_{n-m}/k)} \times \int_0^{\log H(S_{n-m}/k)} O(\sqrt{n}\, g(t))dt,$$

as $n \longrightarrow \infty$. Now the assumption

$$\sqrt{n}\, g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \xrightarrow{p} 0, \text{ as } n \longrightarrow \infty$$

implies that

$$\frac{k}{\sqrt{n}} \log L'(\exp \frac{S_{n-i+1}}{k}) \xrightarrow{p} 0 \tag{3.13}$$

as $n \longrightarrow \infty$ and for $1 \leq i < n$. When $m = m_n$ satisfies (M1), $\sqrt{n}B_n^2$ is a Cezaro mean with the summand tending to 0 in probability by (3.13), consequently (3.12) is true. The proof in the cases (SR2-3) is similar. □

4. Examples

We give some examples of distributions satisfying (F) and demonstrate the applicability of the previous results.

EXAMPLE 1. Let $1 - F(x) = cx^{-1/\alpha}(\log x)^\theta$, for x large and $c > 0$, $\theta \neq 0$. It follows that $L(x) = c(\log x)^\theta$ and $\log L(x) \sim \theta \log\log x$ as $x \longrightarrow \infty$. Therefore (2.2) is valid for $1 \leq k < n$ fixed or sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o(\sqrt{n})$ as $n \longrightarrow \infty$.

Now L is (SR2) with $g(x) = 1/\log x$ and $K(\lambda) = \theta \log \lambda$. Hence by (i) of Lemma 3, Lemma 2 and the law of large numbers, $g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \sim \frac{k}{\alpha n}$. Consequently (2.1) is true for sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o(\sqrt{n})$ as $n \longrightarrow \infty$.

EXAMPLE 2. Let $1 - F(x) = x^{-1/\alpha}(c + dx^{-\theta})$ as $x \longrightarrow \infty$; $c, d, \theta > 0$. In this case $L(x) = c + dx^{-\theta}$ and $\log L(x) \longrightarrow \log c$. Hence (2.2) is true for $1 \leq k < n$ fixed or sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o(\sqrt{n})$ as $n \longrightarrow \infty$.

Here L is (SR1) with $g(x) = x^{-\theta}$ but it is more convenient to consider directly $\sqrt{k}\, A_n^2$. It follows that $\sqrt{k}\, A_n^2 = O_p\left(\frac{k^{3/2}}{n}\right)$ $(n \longrightarrow \infty)$, by the law of large numbers. Therefore (2.1) is true for sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o(\sqrt{n})$ as $n \longrightarrow \infty$.

EXAMPLE 3. Let $1 - F(x) = cx^{-1/\alpha} \exp\left[(\log x)^\theta\right]$, for x large and $c > 0$, $0 < \theta < 1$. Thus $L(x) = c\exp\left[(\log x)^\theta\right]$. We see that $\log L(x) \sim (\log x)^\theta$ as $x \longrightarrow \infty$. Hence if $0 < \theta < 1/2$, (2.2) is valid for $1 \leq k < n$ fixed or sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o(\sqrt{n})$ as $n \longrightarrow \infty$.

Clearly L is (SR2) with $g(x) = 1/(\log x)^{1-\theta}$ and $K(\lambda) = \theta \log \lambda$. Hence $g\left(H\left(\frac{S_{n-m}}{k}\right)\right) \sim \left(\frac{k}{\alpha n}\right)^{1-\theta}$ a.s. as $n \longrightarrow \infty$ by (i) of Lemma 3, Lemma 2 and the law of large numbers. It follows that (2.1) is true for sequences $k = k_n \longrightarrow \infty$ satisfying (K1) with $k = o\left(n^{\frac{1-2\theta}{2-2\theta}}\right)$ $(n \longrightarrow \infty, 0 < \theta < \frac{1}{2})$.

5. Simulation and numerical results

To validate the theoritical results of section 2, we give some numerical results to show how the estimates α_n and β_n behave in practice. The k-record values are obtained by generating exponential random variables and then applying the function H to their sums. We consider here the Pareto distributions of the form $1 - F(x) = cx^{1/\alpha}\,(1 + x^{-1})$, where $\alpha, c > 0$.

5.1. Number of k-records

The expected number of k-record values and its variance are related to the sequence $k = k_n$ in the following manner. Set $m^{(k_n)}(n) = k_n \sum_{i=k_n}^{n} 1/i$ and $v^{(k_n)}(n) = m^{(k_n)}(n) - k_n^2 \sum_{i=k_n}^{n} 1/i^2$ for $n \geq 1$. Denote by $N^{(k_n)}(n)$ the number of k-record values in the sequence $X_1, \ldots, X_n$. Then (Ref., e.g., Theorem 3.2 in [6] and Lemma 2.1 in [8]),

$$E(N^{(k_n)}(n)) = m^{(k_n)}(n),$$
$$Var(N^{(k_n)}(n)) = v^{(k_n)}(n) \text{ for } n \geq 1.$$

TABLE 1. *The expected number of k-records in a sample of size n for a given sequence $k = k_n$. The notation $[x]$ stands for the integer part of x.*

n	$k = k_n$	$m^{(k)}(n)$	$v^{(k)}(n)$
10	$[\log n]$	3.86	1.29
	$[n^{0.5}]$	4.29	1.26
	$[n^{0.8}]$	3.87	0.88
100	$[\log n]$	13.42	3.01
	$[n^{0.5}]$	23.58	3.75
	$[n^{0.8}]$	37.42	3.61
1000	$[\log n]$	31.21	4.97
	$[n^{0.5}]$	108.20	8.81
	$[n^{0.8}]$	347.58	12.61

5.2. Consistency of α_n and β_n

The tables 2–4 show that the estimates α_n and β_n behave quite well for a reasonable number of k-records (see Table 1). For $c = 1$, the statistic β_n is more precise than α_n in estimating α. In the other hand, when $c > 1$, the estimate β_n tends to over-evaluate the value of α for $n \leq 10$.

In the tables 2–4 below, n represents the number of k-records, α_n the first estimate of α for a given n, β_n the second estimate of α for a given n, σ_n the theoritical standard deviation for a given n and $\sigma(\alpha_n$ or $\beta_n)$ the standard deviation of 5000 estimates of α.

TABLE 2. *The sequences* $k_n = m_n = [\log n]$.

n	c	α	α_n	σ_n	$\sigma(\alpha_n)$
5	1	0.5	0.469	0.500	0.479
10			0.464	0.353	0.332
15			0.491	0.353	0.349
5	1	1.0	0.973	1.000	0.978
10			0.989	0.707	0.695
15			1.008	0.707	0.719
5	2	1.0	0.988	1.000	1.005
10			0.982	0.707	0.706
15			1.013	0.707	0.731
5	1	2.0	2.047	2.000	2.076
10			1.998	1.414	1.436
15			1.977	1.414	1.376

n	c	α	β_n	σ_n	$\sigma(\beta_n)$
5	1	0.5	0.508	0.223	0.211
10			0.511	0.158	0.152
15			0.502	0.129	0.131
5	1	1.0	1.004	0.447	0.442
10			1.011	0.316	0.318
15			1.003	0.258	0.265
5	2	1.0	1.135	0.447	0.469
10			1.138	0.316	0.344
15			1.097	0.258	0.276
5	1	2.0	2.009	0.894	0.899
10			1.992	0.632	0.632
15			2.002	0.516	0.520

TABLE 3. *The sequences* $k_n = m_n = [n^{0.25}]$.

n	c	α	α_n	σ_n	$\sigma(\alpha_n)$
5	1	0.5	0.471	0.500	0.491
10			0.491	0.500	0.501
15			0.496	0.500	0.501
5	3	0.5	0.481	0.500	0.491
10			0.505	0.500	0.515
15			0.506	0.500	0.510
5	1	1.0	0.964	1.000	0.996
10			0.969	1.000	0.969
15			1.009	1.000	0.984
5	1	2.0	1.979	2.000	1.988
10			2.028	2.000	2.015
15			2.006	2.000	1.983

n	c	α	β_n	σ_n	$\sigma(\beta_n)$
5	1	0.5	0.509	0.223	0.214
10			0.501	0.158	0.163
15			0.499	0.129	0.129
5	3	0.5	0.613	0.223	0.246
10			0.557	0.158	0.170
15			0.538	0.129	0.135
5	1	1.0	1.006	0.447	0.436
10			0.995	0.316	0.314
15			0.999	0.258	0.257
5	1	2.0	2.006	0.894	0.889
10			2.001	0.632	0.627
15			2.004	0.516	0.517

TABLE 4. *The sequences* $k_n = m_n = [n^{0.49}]$.

n	c	α	α_n	σ_n	$\sigma(\alpha_n)$
5	1	0.5	0.579	0.353	0.325
10			0.530	0.228	0.0262
15			0.469	0.228	0.276
5	1	1.0	1.068	0.707	0.667
10			1.015	0.577	0.547
15			0.996	0.577	0.583
5	1	2.0	2.014	1.414	1.362
10			1.987	1.154	1.120
15			1.984	1.154	1.140

n	c	α	β_n	σ_n	$\sigma(\beta_n)$
5	1	0.5	0.540	0.223	0.209
10			0.522	0.158	0.155
15			0.508	0.129	0.125
5	1	1.0	1.034	0.447	0.437
10			1.009	0.316	0.313
15			1.002	0.258	0.255
5	1	2.0	2.000	0.894	0.884
10			1.994	0.632	0.635
15			1.997	0.516	0.513

References

[1] Berred, M., On record values and the exponent of a distribution with a regularly varying upper tail, *J. Appl. Probab.*, **29** (1992a), 575–586.

[2] Dziubdziela, W. and Kopociński, B., Limit properties of the k-th record values, *Zastos. Mat.*, **15** (1976), 187–190

[3] Deheuvels, P., The characterization of distributions by order statistics and record values — a unified approach, *J. Appl. Probab.*, **21** (1984), 326–334.

[4] Galambos, J., *The asymptotic Theory of Extreme Order Statistics*, John Wily & Sons, 1978.

[5] Resnick, S. I., *Extreme Values, Regular variation and Point Processes*, Springer-Verlag, New York, 1987.

[6] Nevzorov, V.B., Records, *Theory Probab. Appl.*, **32** (1987), 201–228.

[7] Nevzorov, V.B., On the kth record times and their generalisations, *Zap. Nauchn. S[illegible] Leningrad. Otdel. Mat. Inst. Steklov.*, **153** (1986), 115–121.

[8] Deheuvels, P. and Nevzorov, V. B., Limit laws for k-record times, (1992), *Preprint.*

[9] Berred, M., On record values and the estimation of the Weibull tail-coefficent, *C. R. Acad. Paris, Série I*, **312** (1991), 943–946.

[10] Berred, M., K-record values and the extreme-value index, presented at *XIII^{es} Rencontres Franco-Belges de Statisticiens*, (1992b), submitted to *JSPI.*

[11] Hill, B.M., A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, **3** (1975), 1163–1174.

[12] De Haan, L. and Resnick, S.I., A simple asymptotic estimate for the index of a stable distribution, *J. Roy. Statist. Soc.*, **42** (1980), 83–87.

[13] Teugels, J.L., Limit theorems on order statistics, *Ann. Probab.*, **9** (1981), 868–880.

[14] Hall, P., On some simple estimates of an exponent of regular variation, *J. Roy. Statist. Soc.*, 44 (1982), 37–42.

[15] Mason, D.M., Laws of large numbers for sums of extreme values, *Ann. Probab.*, **10** (1982), 754–764.

[16] Haeusler, E. and Teugels, J.L., On asymptotic normality of Hill estimator for the exponent of a regular variation, *Ann. Statist.*, **13** (1985), 743–757.

[17] Csörgő, S., Deheuvels, P. and Mason, D.M., Kernel estimate of the tail index of a distribution, *Ann. Statist.*, **13** (1985), 1050–1077.

[18] Deheuvels, P. and Mason, D.M., The asymptotic behavior of sums of exponential extreme values, *Bull. Sc. Math.*, 2^e *Série*, **112** (1988), 211–233.

[19] Deheuvels, P., Haeusler, E. and Mason, D.M., Almost sure convergence of the Hill estimator, *Math. Proc. Camb. Phi. Soc.*, **104** (1988), 371–381.

[20] Deheuvels, P., Haeusler, E. and Mason, D.M., On the almost sure behavior of sums of extreme values from a distribution in the domain of attraction of a Gumbel law, *Bull. Sc. Math.*, 2^e Série, **114** (1990), 61–95.

[21] Smith, R.L., Estimating tails of probability distributions, *Ann. Statist.*, **15** (1987), 1174–1207.

[22] Goldie, C. M. and Smith, R. L., Slow variation with remainder : Theory and applications, *Quart. J. Math., Oxford (2)*, **38** (1987), 45–71.

[23] Bingham, N.H., Goldie C.M. and Teugels J.L., *Regular Variation*, Cambrige University Press, 1987.

[24] Hanson, D.L. and Russo, R.P., On the law of large numbers, *Ann. Probab.*, **9** (1981), 513–519.

The Point-Process Approach To The Directional Analysis Of Extreme Wind Speeds

Bortot, P.
Università di Padova, Padova, Italia

In this paper the problem of directional modeling of extreme wind speeds is discussed. An adapted version of the method developed by Smith in 1989 [1], based on a point-process view on extreme value problems, is proposed. Techniques are considered to solve difficulties deriving from serial correlation and angular dependence. The procedure is illustrated with an application to real data.

1. Introduction

The problem of directional modeling of extreme wind speeds, although rarely discussed in literature, plays an important role in civil engineering. In fact, the directional analysis of the extremal behaviour of winds provides engineers with useful information for an accurate choice of building orientation and leads potentially to considerable savings.

The problem has already been dealt with by Coles and Walshaw [2]. They employ a modified version of the r largest annual events model in which the parameters of the Generalized Extreme Value (GEV) distribution are expressed as functions of direction. In this paper our aim is to use recent developments in the methodology of univariate analysis of extreme values, suitably adapted, to obtain a model which takes into account directional aspects of wind process. In doing this we will partly follow the ideas proposed by Coles and Walshaw.

2. Description of data

The data analyzed were collected at the meteorological Military Air Force station in Trieste, Italy. The station is situated at 8m above sea level; the anemometer reaches a height of 39m above ground level. The data consists of measurements of the direction and the average intensity of wind: averages are calculated over the ten minutes preceding the recording which is limited to the so-called synoptic hours (00, 03, 06, 09, 12, 15, 18, 21). The records cover a period of 23 years: from the 1st of January 1951 to the 30th of December 1973, nominally 67200 observations. Of these values, 1361 are missing from periods when the equipment was out of service. Wind speed is measured in knots. There are 36 directional sectors: the angle is recorded to the nearest 10° in clockwise orientation starting from North.

Like most environmental data, the time series of wind speed departs from iid sequences in two respects: first, in being heavily seasonal and second, in exhibiting short-range dependence due to the persistence of the weather leading to clustering of high-level exceedances. Seasonality and serial correlation compel us to adjust the tools derived from classical extreme value theory, since it concerns maxima of independent, identically distributed random variables.

To assess the connection between wind intensity and sector of origin, a boxplot of wind mean speed has been constructed for each direction (Figures 1a and 1b). As we are only interested in extreme values, we have discarded all the observations below 10 knots. It seems clear that the extremal behaviour of wind is strongly influenced by direction: the North-East sector being the most affected.

3. The point-process approach

The method employed in this analysis is an adapted version of the one developed by Smith [1] in the study

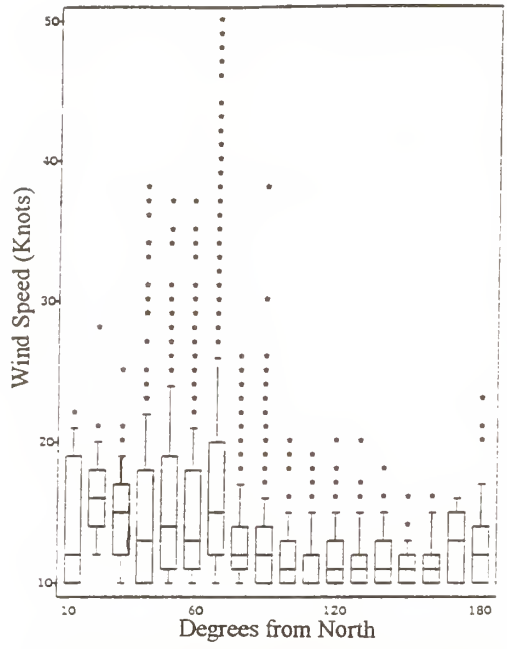


Fig. 1a. Boxplot for wind speeds in each direction (10°-180°)

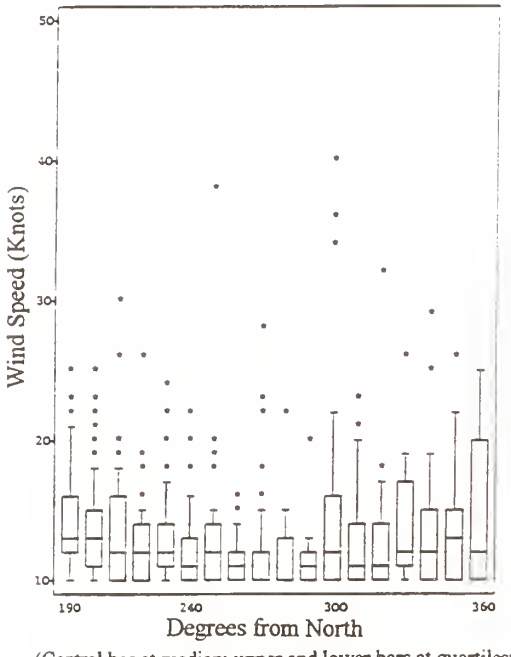


Fig. 1b. Boxplot for wind speeds in each direction (190°-360°)

of ground-level ozone concentrations. It is based on a point-process approach to extreme value problems, which was originally introduced by Pickands [3] and, in more recent years, emphasized in the books of Leadbetter, Lindgren and Rootzén [4] and Resnick [5]. In its application to ozone data, the method has shown a considerable versatility and a wide applicability. Moreover, it includes all the other methods of analysis of extremes (the traditional method, the *Peaks over Threshold* method and the method based on the r largest annual events) as special cases.

Let $X_1, X_2, \ldots$ denote an iid sequence with common distribution function F and $M_n = \max(X_1, \ldots, X_n)$. Suppose that there exist normalizing sequences $a_n > 0$, b_n such that, as $n \to \infty$,

$$P\{(M_n - b_n)/a_n \le x\} = F^n(a_n x + b_n) \to H(x),$$

where H is the Generalized Extreme Value distribution, i.e.

$$H(x;\mu,\sigma,k) = \exp[-\{1-k(x-\mu)/\sigma\}^{1/k}]$$

valid over the range $\{x: 1-k(x-\mu)/\sigma > 0\}$; $\sigma > 0$, $-\infty < \mu < \infty$ and $-\infty < k < +\infty$.

Let $Y_{n,i} = (X_i - b_n)/a_n$ and P_n denote the point process on the plane with points at $(i/(n+1), Y_{n,i})$, $i = 1, \ldots, n$. Then, under a topology which essentially excludes points whose ordinates approach the lower endpoint of distribution H, P_n converges, as $n \to \infty$, to a nonhomogeneous Poisson process P with intensity measure

$$\Lambda\{(t_1,t_2) \times (x,\infty)\} = (t_2 - t_1)[1-k(x-\mu)/\sigma]^{1/k} \quad (3.1)$$

whenever $0 \le t_1 \le t_2 \le 1$ and $1-k(x-\mu)/\sigma > 0$.

This result may be regarded as fundamental in yielding all relevant asymptotic distributional properties. For instance, the limiting conditional probability that $Y_{n,i} > u+y$ given $Y_{n,i} > u$ is given by the ratio between the mean number of points that P has on $(0,1) \times (u+y,\infty)$ and the mean number of points on $(0,1) \times (u,\infty)$. By using (3.1) we obtain

$$\frac{[1-k(u+y-\mu)/\sigma]^{1/k}}{[1-k(u-\mu)/\sigma]^{1/k}} = [1 - \frac{ky}{\sigma - ku + k\mu}]^{1/k} \quad (3.2)$$

which is the Generalized Pareto distribution with shape parameter k and scale parameter $\sigma - ku + k\mu$.

It should be noted that the parameters of the intensity measure of the process P are the parameters of the maximum limiting distribution. In an application to real data this result enables us to use the Poisson process P to represent the time series and, in

this way, to obtain more precise estimates of the parameters of the annual maximum distribution. Nevertheless, seasonal variation and serial correlation, inherent in wind data, make it impossible to apply the approach above described directly, since, so far, it has been confined to iid sequences.

4. Seasonality and serial correlation

Smith [1] suggests overcoming the seasonal problem by splitting up the year into a number of periods, each of which is modelled separately: that is, allowing all the parameters of the process P to be seasonally dependent. However, to avoid further complicating the analysis, we choose not to take seasonality into account, believing that this simplification can be justified.

With reference to serial dependence, since extreme values tend to occur in clusters, a technique extensively employed in this kind of application is to try to identify clusters of high-level exceedances, with the intention of concentrating on cluster maxima for the rest of the analysis. There is no universally accepted method for identifying clusters. The one followed by Smith and adopted here consists of fixing a threshold u and a cluster interval z^*. Two adjacent exceedances of u are deemed to lie in the same cluster if the interval between them is less then z^*. If the time interval between them is longer than z^*, it is assumed that the old cluster has finished and a new one begun. In this way clusters are defined and only the largest observation within each of them is retained for fitting. Cluster interval and threshold are chosen empirically. The technique that we suggest for this choice is to assess the goodness of fit of the model for different values of z^* and u over a reasonable range and, finally, select the smallest of the couples which yield satisfactory results. In fact, if a certain value of z^* ensures independence between clusters, then this independence is also ensured for all higher values of z^*. Nevertheless, the increment of z^* reduces the number of independent observations available, with the twofold effect of: 1) raising questions about the validity of the asymptotic arguments justifying approximations based on the Poisson process P; 2) reducing the estimation precision. Similar arguments are valid for u.

The properties of the point-process approach can be employed to test the goodness of fit of the model on varying z^* and u. A first test is based on equation (3.2) and consists of graphically assessing how closely the excesses over the threshold u fit a Generalized Pareto distribution with parameters k and $\sigma - ku + k\mu$. It is possible to carry out an alternative graphical test using GEV distribution to transform annual maxima to uniformity before plotting them against the empirical distribution function.

5. Directional components

Following ideas by Coles and Walshaw, we choose to calculate the directional components of velocity of each recorded speed and then to model wind components. This is because each recorded speed, although associated to a certain direction, has a contribution from all directions. This procedure is also justified by the fact that we work with mean speeds, each of which is the result of speeds coming from a number of different sectors.

So, let Y_α denote a mean speed of magnitude Y in direction α. The component of velocity of Y_α in direction ϕ is $Y\cos(\alpha - \phi)$ if $|\alpha - \phi|$ modulo $\pi < \pi/2$; otherwise it is zero.

Using the resolution into directional components we obtain 36 complete series of wind intensities: one for each sector. As with the original time series, the 36 sequences feature a high short range autocorrelation.

6. Direction as a covariate

We could easily apply the point-process approach to the directional study of Trieste data by fitting the model described in section 3 separately to the observations of each of the 36 sequences. Nevertheless, since data are recorded on a fine directional scale, we can obtain a considerable gain in efficiency modelling the parameters of the process P (k, μ and σ) as functions of direction. This also has the advantage of smoothing annual maximum speed distributions across directions in accordance with the features of the physical process.

Suppose that the annual maximum wind components in direction α ($\alpha \in A = \{10°, 20°, ..., 360°\}$) have GEV distribution with shape parameter k_α, location parameter μ_α and scale parameter σ_α. As suggested by Coles and Walshaw, a natural choice is to express each of the three parameters as a sum of the first terms in a Fourier series. Therefore, let

$$k_\alpha = a_1 + \sum_{t=1}^{n_1} b_{1,t}\cos(t\alpha - c_{1,t}),$$

$$\mu_\alpha = a_2 + \sum_{t=1}^{n_2} b_{2,t} \cos(t\alpha - c_{2,t}),$$

$$\sigma_\alpha = a_3 + \sum_{t=1}^{n_3} b_{3,t} \cos(t\alpha - c_{3,t}). \qquad (6.1)$$

The parameters in the model are now a_m, $b_{m,t}$ and $c_{m,t}$; $t = 1, \ldots, n_m$; $m = 1, 2$ and 3. In order for the model to be well-defined we have to restrict $b_{m,t} \geq 0$ and $0° \leq c_{m,t} \leq 360°$. We have also to exclude models in which $b_{m,t} = 0$ for some m and t with the corresponding $c_{m,t} \neq 0°$, since $b_{m,t} = 0$ corresponds to the absence of the t-th harmonic term and, in this case, the value of $c_{m,t}$ does not influence the model.

7. Angular dependence

Since we have chosen to model GEV parameters as functions of direction, we can not ignore the dependence of extreme wind speeds across directions. This dependence, which Coles and Walshaw call angular dependence to distinguish it from temporal dependence discussed in section 4, is a consequence of the fact that storms tend to give successive high observations in a number of different directions. Moreover, the resolution into directional components itself induces dependence across directions.

To solve the problem of temporal dependence we filter each of the 36 sequences of wind components following the rule described in section 4 and using a cluster interval $z^* = 24$ hours and a threshold $u = 11$ knots. These values derive from a previous analysis which was carried out adopting the point-process approach but ignoring directional aspects. They are the smallest couple ensuring a good fit of the point-process model to Trieste data. From this operation we obtain 36 series of cluster maxima, each being independent temporally; however, dependence across directions remains.

We have overcome this further obstacle following ideas proposed by Smith [6] for dealing with spatially dependent data. It consists of constructing the likelihood function as if there were independence across directions. To account for angular dependence, standard errors of parameter estimates and likelihood ratio test are suitably modified.

Let G be the observed information matrix under the model which assumes independence. If the independence assumption were valid, G^{-1} could be used to approximate the covariance matrix of maximum likelihood estimates. Smith shows that, to take account of dependence, this approximation should be replaced by $G^{-1}VG^{-1}$, where V is the covariance matrix of log-likelihood derivatives. If years are independent, V may be obtained empirically. Similar arguments are applied to adjust the usual asymptotic distribution of likelihood ratio test for model discrimination.

8. Directional model

In adapting the method proposed by Smith to our study we assume stationarity from year to year, since the previous analysis, ignoring direction, proved that there was no (linear) trend in the data.

Let M_i denote the length of observation in days in year i ($i = 1, \ldots, 23$), $N_{i\alpha}$ the number of cluster maxima in year i and direction α ($\alpha \in A$) and $Y_{i\alpha j}$ ($j = 1, \ldots, N_{i\alpha}$) the cluster maxima in year i for direction α. Let u denote the fixed threshold ($u = 11$ knots). Following the discussion of previous sections, we assume that for direction α, in any given year, the exceedance times of threshold u and cluster maxima form a nonhomogeneous Poisson process with intensity measure given by (3.1) with $\mu_\alpha, \sigma_\alpha$ and k_α replacing μ, σ and k. Then, the likelihood function in direction α is:

$$L_\alpha = \prod_{i=1}^{23} [\exp\{-\frac{M_i}{365}(1 - k_\alpha(u - \mu_\alpha)/\sigma_\alpha)^{1/k_\alpha}\} \cdot \prod_{j=1}^{N_{i\alpha}} \{\frac{1}{\sigma_\alpha}(1 - k_\alpha(y_{i\alpha j} - \mu_\alpha)/\sigma_\alpha)^{1/k_\alpha - 1}\}] \qquad (8.1)$$

where k_α, μ_α, and σ_α are defined by equations (6.1).

Assuming independence across directions, the log-likelihood is $l = \sum_{\alpha \in A} \ln L_\alpha$. Maximum likelihood estimates of a_m, $b_{m,t}$ and $c_{m,t}$ ($t = 1, \ldots, n_m$; $m = 1, 2, 3$) are obtained by maximizing numerically l.

The quantities of greatest interest for engineers are return levels. The T-year return level in direction α is the $1 - T^{-1}$ quantile of the annual maximum distribution in direction α; it is given by

$$q_{T,\alpha} = \mu_\alpha + \frac{\sigma_\alpha}{k_\alpha}[1 - \{-\ln(1 - T^{-1})\}^{k_\alpha}].$$

It should be noted that $q_{T,\alpha}$ is the quantile of the distribution of the annual maximum component in direction α.

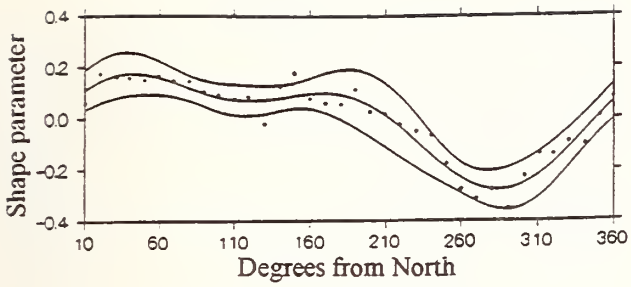


Fig. 2a. Shape Parameter Estimates

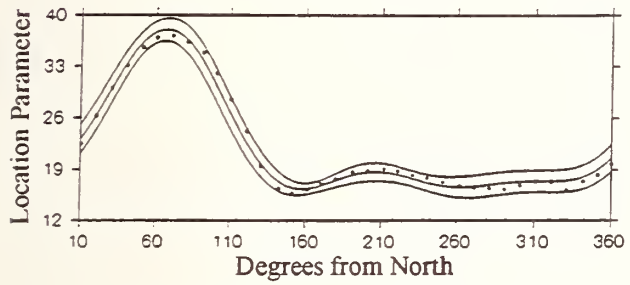


Fig. 2b. Location Parameter Estimates

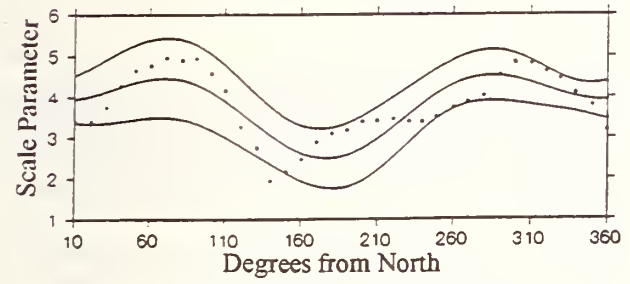


Fig. 2c. Scale Parameter Estimates

9. Results

The procedure followed for identifying the model which yields the best fit to Trieste data consists of varying the number of harmonic terms in equations (6.1) and comparing fitted models using likelihood ratio tests modified in the manner advocated by Smith.

The model chosen with this procedure has three harmonic terms for the location parameter, two for the scale parameter and for the shape parameter. Results are shown in Table 1. Standard errors of parameter estimates (shown in parentheses) are adjusted for angular dependence. The parameters $c_{m,t}$ ($t = 1, ..., n_m$; $m = 1, 2$ and 3) are expressed in radians.

It is not possible to reduce the number of harmonic terms for the three parameters, and, in particular, for the shape parameter, without suffering a heavy loss in terms of goodness of fit of the model.

The last question concerns assessing the fit of the chosen model to data. For each of the 36 sectors separately we have employed the two tests outlined in section 4: one based on the Generalized Pareto distribution and the other based on the GEV distribution. The plots (not shown) indicate a good agreement between expected values and observed values.

An alternative technique is proposed by Coles and Walshaw. It consists of examining how closely maximum likelihood estimates of k_α, μ_α and σ_α follow the corresponding values when each sector is considered separately. Figures 2a, 2b and 2c contain such comparisons. In each plot lines join the maximum likelihood estimates and the upper and lower bounds of the 95% confidence interval, while the plotted points represent the parameter estimates when (8.1) is maximized on each sector separately. A similar comparison has been made for the 50-year return level (not shown). It is important to recall that maximum likelihood estimates derived from the *separate sectors* analysis have larger standard errors than those of the covariate model. For this reason one or more points outside the confidence interval do not necessarily indicate a lack of fit of the model. The plots show no systematic deviation of covariate model estimates from *separate sectors* analysis estimates. Therefore, we can conclude that the chosen model seems able to capture the variations induced by direction on annual maximum wind component distribution.

10. Conclusion

In this application to wind data the method proposed by Smith has confirmed the qualities shown in the study of ozone concentrations. Besides ensuring accurate results and requiring very mild assumptions, it is a flexible tool and can be easily adapted for handling complex features of real data.

The analysis has led to the conclusion that direction heavily influences the extremal behaviour of wind. As before mentioned, the proposed model seems to accurately describe variations across sectors. Therefore, it can be of great utility to civil engineers for a correct assessment of wind impact on structures.

As a final comment, we recall that we have only adjusted the point-process method in order to take into account angular dependence, with no attempt to model such a dependence. This would require the use of tools derived from multivariate extreme value theory.

Table 1. Maximum likelihood estimates and standard errors of the selected model

	Parameters	Estimates	
Shape parameter	a_1	-0.001	(0.009)
	$b_{1,1}$	0.180	(0.017)
	$c_{1,1}$	1.659	(0.201)
	$b_{1,2}$	0.099	(0.034)
	$c_{1,2}$	0.637	(0.182)
Location parameter	a_2	22.225	(0.585)
	$b_{2,1}$	8.338	(0.233)
	$c_{2,1}$	1.137	(0.052)
	$b_{2,2}$	5.260	(0.174)
	$c_{2,2}$	2.221	(0.096)
	$b_{2,3}$	2.287	(0.202)
	$c_{2,3}$	3.703	(0.084)
Scale parameter	a_3	3.791	(0.281)
	$b_{3,1}$	0.711	(0.113)
	$c_{3,1}$	6.189	(0.343)
	$b_{3,2}$	0.590	(0.137)
	$c_{3,2}$	3.047	(0.255)

ACKNOWLEDGEMENTS

I am grateful to Prof. Alessandra Salvan and Dr. Matteo Grigoletto, of the University of Padua (Italy), for all their help and support throughout this project.

REFERENCES

[1] Smith, R.L., Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone, *Statistical Science*, 4 (1989), 367-393.

[2] Coles, S.G. and Walshaw, D., Directional modelling of extreme wind speeds, Submitted, 1992.

[3] Pickands, J., The two-dimensional Poisson process and extremal processes, *J. Appl. Probab.*, 8 (1971), 745-756.

[4] Leadbetter, M.R., Lindgren, G. and Rootzén, H., Extremes and *Related Properties of Random Sequences and Series*, Springer, New York, 1983.

[5] Resnick, S., *Extreme Values, Point Processes and Regular Variation*, Springer, New York, 1987.

[6] Smith, R.L., Regional estimation from spatially dependent data, Submitted, 1991.

High Boundary Excursions Of Locally Stationary Gaussian Processes

Bräker, H.U.
Institut für Mathematische und Versicherungslehre, Bern, Germany

Let $\{X(t),\ t \in T\}$ be a locally stationary Gaussian process and $f(t),\ t \in T$ a continuous function, where T is a finite or infinite interval. An asymptotic estimate for small probabilities $\mathrm{P}(X(t) \geq f(t),\ \text{some } t \in T)$ is derived by approximating the density of the first passage time and integrating over T. This work extends a result proven by J. Cuzick, Ref. [5], for stationary Gaussian processes.

keywords: locally stationary Gaussian process, boundary crossing, first passage time

Let $\{X(t),\ t \in T\}$ be a Gaussian process ($T = [0, r),\ r \leq \infty$) and $f(t),\ t \in T$ a continuous function. We are interested in

$$\mathrm{P}(X(t) \geq f(t),\ \text{some } t \in T). \qquad (1)$$

The asymptotic behavior of this probability has been studied during the last thirty years.

J. Pickands, Ref. [8], proved the following

Theorem 1 *Let $\{X(t),\ t \geq 0\}$ be a separable stationary Gaussian process with*

$EX(t) \equiv 0$ *and covariance function*

$$\rho(h) = 1 - 1/2R^2|h|^\alpha + o(|h|^\alpha) \quad (h \to 0)$$

$(0 < R < \infty,\ 0 < \alpha \leq 2)$ such that $\rho(h) \log h \to 0$ as $h \to \infty$ (Berman's condition).
If $u = u_\tau \to \infty$ as $\tau \to \infty$ such that

$$\tau H_\alpha (Ru)^{2/\alpha} \psi(u) \to \tau \in (0, \infty)$$

then

$$\mathrm{P}(X(t) \leq u,\ t \leq r) \underset{r \to \infty}{\longrightarrow} e^{-\tau},$$

where $\psi(x) = (2\pi)^{-1/2} x^{-1} e^{-x^2/2}$ and H_α is a constant depending on α only.

(Here we use the definition of H_α given by Qualls and Watanabe, Ref. [9], which is also used by Cuzick and differs slightly from that given by Pickands.)

The probability (1) is analyzed by splitting the interval T into 'small' intervals $T_i,\ i = 0, 1, \ldots$ and approximating the probability for an exceedance of $f(t)$ in T_i. This suggests that only the local behavior of the process is important and that the assumption of stationarity might be relaxed. Therefore S. M. Berman, Ref. [2], introduced the concept of local stationarity:

Definition 2 *A real valued separable Gaussian process* $\{X(t), t \in T\}$ *is said to be locally stationary if*

(i) $EX(t) \equiv 0$ *and* $EX^2(t) \equiv 1$

(ii) There exist a continuous function $R(t), t \in T$ *with* $0 < \inf\{R(t) : t \in T\} \leq \sup\{R(t) : t \in T\} < \infty$ *and a strictly increasing continuous function* $K(h), 0 \leq h < h_0$ $(h_0 > 0)$ *with* $K(0) = 0$ *such that*

$$\lim_{h \to 0} \frac{E\,(X(t+h) - X(t))^2}{K^2(|h|)} = R^2(t)\ ,$$

uniformly in $t \in T$.

So the covariance function of a locally stationary Gaussian process has the form

$$\rho(t, t+h) = 1 - 1/2R^2(t)K^2(|h|) + o(K^2(|h|)). \tag{2}$$

We will restrict ourselves to the case where $K^2(h)$ is regularly varying at zero with index α $(0 < \alpha \leq 2)$. It can be shown that such a process has continuous sample paths with probability one.

Theorem 1 was generalized by J. Hüsler, Ref. [7], for locally stationary Gaussian processes and nonconstant boundaries.

In Theorem 1 both the length of the time interval (r) and the boundary $u = u_r$ must tend to ∞ in order to obtain a nondegenerate limit for (1). If $r(< \infty)$ is fixed and $u \to \infty$ or if $r \to \infty$ and the boundary is very high, then $P(X(t) \geq u$, some $t \leq r) \to 0$ and the question about its convergence rate arises.

Assume for the moment that $X(\cdot)$ is stationary. Then by Theorem 1

$$P(X(t) \geq u, \text{ some } t \leq r) \approx rH_\alpha(Ru)^{2/\alpha}\psi(u) \tag{3}$$

for large u $(r < \infty)$. Cuzick showed that (3) is indeed the correct convergence rate. He also obtained the convergence rate for the case, where the boundary is a function of t. His result can be extended for locally stationary Gaussian processes in the following way:

Theorem 3 *Let* $\{X(t),\ t \in T\}$ *be a separable locally stationary Gaussian process with covariance function* (2), *where* $R^2(\cdot)$ *is uniformly continuous on* T *and* $K^2(\cdot)$ *is regularly varying at zero with index* α, $0 < \alpha \leq 2$ *such that* $K^{-1}(\cdot)$ *exists in a neighborhood of zero. Let* $(f_n(t),\ t \in T)_{n \in \mathbb{N}}$ *be a sequence of continuous functions satisfying (f1)-(f3):*

(f1)

$$\inf_{t \in T} f_n(t) \underset{n \to \infty}{\longrightarrow} \infty$$

(f2)

$$\int_T \frac{\psi(\epsilon f_n(t))}{\Delta_n(t)}\, dt \underset{n \to \infty}{\longrightarrow} 0\ , \qquad \forall \epsilon > 0,$$

where $\Delta_n(t) = K^{-1}(1/R(t)f_n(t))$.

(f3)

$$(f_n(t + \tau\Delta_n(t))) - f_n(t))f_n(t) \underset{n \to \infty}{\longrightarrow} g(t, \tau)\ ,$$

uniformly in $t \in T$ *and* τ *in compact sets of* $\mathbb{R}$, *where* $g(\cdot, \cdot)$ *is a function satisfying*

$$\sup_{\substack{t \in T \\ |\tau| \leq \theta}} |g(t, \tau)| < \infty\ , \qquad \forall \theta,\ 0 \leq \theta < \infty.$$

Then

$$\frac{1}{\Lambda_n} P(X(t) \geq f_n(t), \text{ some } t \in T) \underset{n \to \infty}{\longrightarrow} 1$$

with

$$\begin{aligned} \Lambda_n \ = \ & \int_T H_\alpha(g(t, \cdot)) \frac{\psi(f_n(t))}{\Delta_n(t)}\, dt \\ & + \Phi^*(f_n(0)) \cdot 1(g(0,1) > 0) \\ & + \Phi^*(f_n(r)) \cdot 1(r < \infty \text{ and } g(r, -1) > 0). \end{aligned} \tag{4}$$

In (4) $\Phi^*(x)$ denotes the tail probability of the standard normal law and the functional $H_\alpha(w)$ $(0 < \alpha \leq 2)$ is defined as follows:

Let $\{X_\alpha(\tau),\ \tau \geq 0\}$ be a Gaussian process with $X_\alpha(0) = 0$ a.s., $EX_\alpha(\tau) = -\tau^\alpha/2$ and

$\mathrm{Var}(X_\alpha(\tau) - X_\alpha(\mu)) = |\tau - \mu|^\alpha$. For continuous functions $w(\tau)$, $\tau \geq 0$ define

$$H_\alpha^a(w,\theta) = \frac{\int_0^\infty P_\alpha^a(w,\theta,s)e^s\,ds}{\int_0^\theta e^{-|w(\tau)|}\,d\tau},$$

where

$$P_\alpha^a(w,\theta,s) = \mathrm{P}\Big(X_\alpha(\tau) \geq s + |w(\tau)|, \text{ some } \tau \in [0,\theta] \cap I_a\Big)$$

with

$$I_a = \begin{cases} \{0,a,2a,\ldots\}, & a > 0 \\ [0,\infty), & a = 0. \end{cases}$$

$H_\alpha^a(w)$ is defined as

$$H_\alpha^a(w) = \limsup_{\theta\to\infty} H_\alpha^a(w,\theta).$$

Let

$$C_0 = \{w : w \text{ continuous and monotone on } [0,\infty) \text{ and } w(0) = 0\}.$$

The following Lemma was proved by J. Cuzick.

Lemma 4 *If $w \in C_0$ then*

a)

$$H_\alpha^a(w,\theta) \underset{\theta\to\infty}{\longrightarrow} H_\alpha^a(w),$$

uniformly in $0 \leq a \leq 1$.

b) $0 < H_\alpha^a(w) < \infty$, $a \geq 0$, $\quad H_\alpha^a(0) = H_\alpha^a$

c) *$H_\alpha^a(w,\theta)$ and $H_\alpha^a(w)$ are jointly continuous in a and w, where on C_0 a sequence $(w_n)_{n\in\mathbb{N}}$ is said to converge to w iff*

1) *$w_n(\tau) \longrightarrow w(\tau)$ uniformly on compact sets*

2) $\Big(\int_0^\infty e^{-|w_n(\tau)|}\,d\tau\Big)^{-1} \longrightarrow \Big(\int_0^\infty e^{-|w(\tau)|}\,d\tau\Big)^{-1}$.

The conditions on $(f_n)_{n\in\mathbb{N}}$ imply that $g(t,\tau)$ is linear in τ, i.e. $g(t,\tau) = C(t)\tau$ with some continuous and bounded function $C(t)$, $t \in T$. Therefore $H_\alpha^a(g(t,\cdot))$ is well defined, continuous in t and $0 < \inf_t H_\alpha^a(g(t,\cdot)) \leq \sup_t H_\alpha^a(g(t,\cdot)) < \infty$. For $w(\tau) = c\tau$ ($c \in \mathbb{R}$) one can show that $H_1(w) = 1/2$ and $H_2(w) = \phi(c) - |c|\Phi^*(|c|)$.

Remarks:

- Cuzick's result can be obtained as a special case of Theorem 3 by letting $X(\cdot)$ be stationary with $R(t) \equiv 1$.
- If $X(\cdot)$ is stationary with $R(t) \equiv R$, $K^2(h) = h^\alpha$ and if $\tau < \infty$, $f_n(t) \equiv u_n$, then $\Lambda_n = \tau H_\alpha (Ru_n)^{2/\alpha}\psi(u_n)$.
- Note that Berman's condition is not necessary, since we are only interested in small exceedance probabilities.

Let us consider two examples of sequences $(f_n)_{n\in\mathbb{N}}$ which satisfy conditions *(f1)-(f3)*. Suppose that $K^2(\cdot)$ has the special form $K^2(h) = h^\alpha$ and let $\tau < \infty$. *(f2)* follows then from *(f1)*.

Example 1: $f_n(t) := n + f(t)$ with $f(\cdot)$ positive and continuously differentiable on $[0,\tau]$. Then

$$g(t,\tau) = \begin{cases} 0, & \alpha < 2 \\ \tau\frac{f'(t)}{R(t)}, & \alpha = 2 \end{cases}$$

Example 2: $f_n(t) := n^\beta f(t)$ ($\beta > 0$) with $f(t) \geq \delta > 0$, $\forall t \leq \tau$ and continuously differentiable. Then

$$g(t,\tau) = \begin{cases} 0, & \alpha < 1 \\ \tau\frac{f'(t)}{R^2(t)f(t)}, & \alpha = 1 \\ \infty, & 1 < \alpha \leq 2 \end{cases}$$

Here *(f3)* is satisfied for $\alpha \leq 1$ only.

Sketch of the proof

As already mentioned, the excursion probability (1) is analyzed by splitting the interval T

into subintervals $T_i = [t_i, t_{i+1}]$ of 'appropriate' length. The split points are chosen as follows:

$$\begin{aligned} t_0 &= 0 \\ t_{i+1} &= t_i + \theta_n \Delta_n(t_i), \quad i \geq 1, \end{aligned}$$

where $\theta_n \to \infty$ 'slowly' such that for instance $\theta_n \Delta_n(t) \to 0$ uniformly in $t \in T$.

Let for $i \geq 0$

$$A_i = \{X(\underline{t_i}) < f_n(\underline{t_i}),\ X(t) \geq f_n(t),\ \text{some } t \in T_i\},$$

where $\underline{t_i} = \operatorname{argmin}\{f_n(t_i), f_n(t_{i+1})\}$.
Then

$$\begin{aligned} &\mathrm{P}(X(t) \geq f_n(t),\ \text{some } t \in T) \\ &\quad = \mathrm{P}\Big(\bigcup_i \big(\{X(\underline{t_i}) \geq f_n(\underline{t_i})\} \cup A_i\big)\Big). \end{aligned}$$

One can show that $\{X(\underline{t_i}) \geq f_n(\underline{t_i})\}$ $(1 \leq i \leq I)$ is either a subset of A_{i-1} or of A_{i+1} or its probability is of smaller order than $\mathrm{P}(A_i)$ $(I = \sup\{i \geq 1 : t_i < \tau\})$. However, $\{(X(0) \geq f_n(0)\}$ and $\{X(\tau) \geq f_n(\tau)\}$ for $\tau < \infty$ are not negligible if $f_n(0) < f_n(t_1)$ and $f_n(\tau) < f_n(t_I)$, respectively (i.e. if $g(0,1) > 0$ or $g(\tau, -1) > 0$).

Thus

$$\begin{aligned} &\mathrm{P}(X(t) \geq f_n(t),\ \text{some } t \in T) \\ &\quad \approx \mathrm{P}\Big(\bigcup_i A_i\Big) + \Phi^*(f_n(0)) \cdot 1(g(0,1) > 0) \\ &\qquad + \Phi^*(f_n(\tau)) \cdot 1(g(\tau, -1) > 0, \tau < \infty). \end{aligned}$$

The proof consists of two major parts

1) Showing that

$$\mathrm{P}(A_i) \sim \int_{t_i}^{t_{i+1}} H_\alpha(g(t,\cdot)) \frac{\psi(f_n(t))}{\Delta_n(t)}\, dt.$$

This step is built upon the local behavior of the process.

2) Showing that

$$\mathrm{P}\Big(\bigcup_i A_i\Big) \sim \sum_i \mathrm{P}(A_i).$$

Here the problem is to find a lower estimate.

To 1)
Suppose $f_n(t_i) \leq f_n(t_{i+1})$. (The case $f_n(t_i) > f_n(t_{i+1})$ is treated similarly.)
Write $g_n(t,\tau) = (f_n(t + \tau\Delta_n(t)) - f_n(t))f_n(t)$.
Using an idea of Pickands, Ref. [8], one can show that

$$\begin{aligned} &\mathrm{P}(A_i) \\ &\quad \approx \psi(f_n(t_i)) \int_0^\infty P_\alpha(g_n(t_i,\cdot), \theta_n, s) e^s\, ds, \\ &\quad \approx \psi(f_n(t_i)) \int_0^{\theta_n} e^{-g(t_i,\tau)}\, d\tau H_\alpha(g(t_i,\cdot)) \\ &\quad \approx \frac{1}{\Delta_n(t_i)} \int_{t_i}^{t_{i+1}} \psi(f_n(t))\, dt H_\alpha(g(t_i,\cdot)). \end{aligned}$$

Application of the mean value theorem gives the desired expression.

To 2)
Using Bonferroni's second inequality we get

$$\begin{aligned} &\mathrm{P}\Big(\bigcup_i A_i\Big) \\ &\quad \geq \sum_i \mathrm{P}(A_i)\Big(1 - \sum_{i \neq j} \mathrm{P}(A_i \cap A_j) / \sum_i \mathrm{P}(A_i)\Big). \end{aligned}$$

Unfortunately, it is not possible to show directly that the ratio term is asymptotically negligible. Therefore one has to use a discrete time approximation first. The interval T_i is split into N_n equally spaced intervals $[t_{ij}, t_{i,j+1}]$, $0 \leq j \leq N_n$, where

$$t_{ij} = t_i + j a_n \Delta_n(t_i) \qquad (a_n = \theta_n / N_n).$$

Writing

$$\begin{aligned} A_i^{a_n} = \{&X(\underline{t_i}) < f_n(\underline{t_i}),\ X(t_{ij}) \geq f_n(t_{ij}), \\ &\text{some } 0 \leq j \leq N_n\}, \end{aligned}$$

we have

$$\mathrm{P}\Big(\bigcup_i A_i\Big) \geq \sum_i \mathrm{P}(A_i^{a_n})(1 - Q_n),$$

where

$$Q_n = \sum_{i \neq j} \mathrm{P}(A_i^{a_n} \cap A_j^{a_n}) / \sum_i \mathrm{P}(A_i^{a_n}).$$

One can show that

$$Q_n = O\Big(\max\Big(\frac{1}{\sqrt{\theta_n}a_n^2}, \frac{\sqrt{\theta_n}}{a_n^2}\int_T \frac{\psi(\epsilon f_n(t))}{\Delta_n(t)}\,dt\Big)\Big)$$

(some $\epsilon > 0$), which tends to 0 if $\theta_n \to \infty$ and $a_n \to 0$ slowly enough. At this point the assumption *(f2)* about $(f_n)_{n\in\mathbb{N}}$ is needed. In this part of the proof the asymptotic linearity of $f_n(t)$ on the intervals T_i is used several times.
As above one can show that

$$\mathrm{P}(A_i^{a_n}) \sim \int_{t_i}^{t_{i+1}} H_\alpha^{a_n}(g(t,\cdot))\frac{\psi(f_n(t))}{\Delta_n(t)}\,dt.$$

Since $H_\alpha^a(g(t,\cdot))$ tends to $H_\alpha(g(t,\cdot))$ as $a \to 0$ (Lemma 4), the proof is complete.

Application to first zeros of empirical characteristic functions

Let Y be a random variable with $\mathrm{E}|Y|^\beta < \infty$ some $\beta > 0$ and write $u(t) = \mathrm{E}\cos tY$ (real part of the characteristic function) and $\sigma^2(t) = \mathrm{Var}\cos tY$. Let $Y_1, \ldots, Y_n$ be a random sample of Y and write $U_n(t) = (1/n)\sum_j \cos tY_j$ (real part of the empirical characteristic function). Denote by τ_0 and R_n the first zero of $u(t)$ and $U_n(t)$ respectively. Heathcote and Hüsler, Ref. [6], showed that if $1 - u(h)$ is regularly varying at zero with index α $(0 < \alpha \le 2)$, then for $\tau > 0$

$$\begin{aligned}&\mathrm{P}(R_n \le \tau)\\ &\quad\approx\ \mathrm{P}(X(t) \ge n^{1/2}u(t)/\sigma(t),\ \text{some } t \le \tau),\end{aligned}$$

where $X(\cdot)$ is a locally stationary Gaussian process with index α (covariance function $\rho(t, t+h) = 1 - 1/2(1-u(h))(1/\sigma^2(t) + o(1))$ as $h \to 0$). If $\alpha \le 1$ and $\tau < \tau_0$, Theorem 3 can be used. The case $\alpha \le 1$ includes for example the Cauchy distribution ($\alpha = 1$), whereas distributions with finite expectation ($\alpha = 2$) are not included.
Suppose $Y \sim$ Cauchy. Then $u(t) = \exp\{-|t|\}$, i.e. $\tau_0 = \infty$ and $\sigma^2(t) = 1/2(1 - \exp\{-2|t|\})$. Application of Theorem 3 yields for $0 < \tau < \infty$

$$\begin{aligned}&\mathrm{P}(R_n \le \tau)\\ &\quad\approx\ (n/\pi)^{1/2}\int_0^\tau e^{-t}/(1-e^{-2t})^{3/2}\\ &\qquad\qquad \exp\{-n/(e^{2t}-1)\}\,dt \quad \text{as } n \to \infty\\ &\quad=\ 1/(2\sqrt{\pi})\int_{n/(e^{2\tau}-1)}^\infty v^{-1/2}e^{-v}\,dv\end{aligned}$$

References

[1] Berman, S. M., Maxima and high level excursions of stationary Gaussian processes, Trans. Amer. Math. Soc., 160 (1971), 67-85.

[2] Berman, S. M., Sojourns and extremes of Gaussian processes, Ann. Prob., 2 (1974), 999-1026.

[3] Berman, S. M., Sojourns and extremes of stochastic processes, Wadsworth and Brooks, Pacific Grove, 1992.

[4] Bräker, H. U., High boundary excursion of locally stationary Gaussian processes, Ph.D. thesis, University of Berne, 1993.

[5] Cuzick, J., Boundary crossing probabilities for stationary Gaussian processes and Brownian motion, Trans. Amer. Math. Soc., 263 (1981), 469-492.

[6] Heathcote, C. R. and Hüsler, J., Estimating the first zero of an empirical characteristic function, Stoch. Proc. Appl., 35 (1990), 347-360.

[7] Hüsler, J., Extreme values and high boundary crossings of locally stationary Gaussian processes, Ann. Prob., 18 (1990), 1141-1158.

[8] Pickands, J. III, Upcrossing probabilities for stationary Gaussian processes, Trans. Amer. Math. Soc., 145 (1969), 51-73.

[9] Qualls, C. and Watanabe, H., Asymptotic properties of Gaussian processes, Ann. Math. Stat., 43 (1972), 580-596.

Asymptotic Approximations For The Crossing Rates Of Poisson Square Waves

Breitung, K.
Sem. F.A. Stochastik, Akademiestr. 1/IV, Munich, Germany

For the extreme value distribution of functions of random vector processes it is difficult to derive exact expresssions; therefore approximations are needed. A mdel commonly used for loads in structural reliability is the Poisson square wave processes $x(t)$. Such a process is defined by a Poisson point process and an additional sequence of i.i.d. random variables $Y_0, Y_1, Y_2, \ldots$ such that between two points t_n and t_{n+1} of the Poisson process the value of $x(t)$ is defined by $x(t) = Y_n$. In this paper the asymptotic Poissonian behavior of the point process of level crossings of functions of independent Poisson square wave processes is shown. This can be used to approximate the extreme value distribution of such functions.

1 Introduction

In many reliability problems it is necessary to calculate the distribution of the maximum of functions of vector random processes. A survey of random processes used as models in load combination problems is given in [1]. For many models it is difficult or impossible to derive the exact distribution of the maximum. Therefore it is of interest to obtain asymptotic approximations for high levels, since especially such results are needed in reliability problems.

Here for a special model, Poisson square wave processes, the asymptotic Poissonian behavior of the point process of level crossings of functions of such processes will be described.

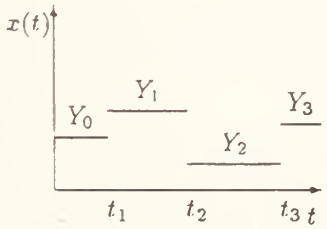


Figure 1: Poisson square wave process

A Poisson square wave process consists of an homogeneous Poisson point process $N(A)$ with intensity λ and a sequence $Y_0, Y_1, Y_2, \ldots$ of i.i.d. random variables, which are independent of the point process.

The Poisson square wave process $x(t)$ is then defined by

$$x(t) = Y_{N(0,t]} = Y_j, \text{ if } N(0,t] = j. \tag{1}$$

The process is often used to describe loads wich change in time (see [2], [1] and [3]). The model can be generalized to a vector process by taking random vectors instead of random variables. In none of the references above a complete proof for the asymptotic Poissonian character of the level crossing point processes is given.

We consider now an n-dimensional Poisson square wave process $\boldsymbol{x}(t) = (x_1(t), \ldots, x_n(t))$. The i-th component is defined by an one-dimensional Poisson square wave process $x_i(t)$ with intensity λ_i and standard normally distributed amplitudes. All component processes are assumed to be independent of each other. The process $\boldsymbol{x}(t)$ changes its value at the jumps of the component processes $x_i(t)$. As described for example in [4] and [1], p. 73, all random variables X with c.d.f. $F(x)$ and with a continuous p.d.f. $f(x)$ can be transformed

into standard normal random variables U by the transformation $U = \Phi^{-1}(F(X))$ ($\Phi(x)$ the standard normal integral); therefore the assumption is not too restrictive in the sense that we need this transformation only to derive a simple proof for Poisson convergence.

Further a function $g : \mathbb{R}^n \to \mathbb{R}$ is given. In the reliability context, this function describes the state of the engineering system under consideration. If $g(\boldsymbol{x}) \geq 0$, the system is intact and if $g(\boldsymbol{x}) < 0$, the system is defect. The problem is to determine the probability

$$\Pr\left(g(\boldsymbol{x}(t)) \geq 0,\ 0 \leq t \leq T\right). \tag{2}$$

This is the probability that the system remains intact during the time interval $[0, T]$. The simplest example for such a function is the sum of the processes, i.e. $g(\boldsymbol{x}(t)) = \beta - \sum_{i=1}^{n} x_i(t)$ with $\beta \in \mathbb{R}$.

To find approximations for this probability, we consider the point process of the outcrossings out of the domain $S = \{\boldsymbol{x}; g(\boldsymbol{x}) \geq 0\}$ into the domain $F = \{\boldsymbol{x}; g(\boldsymbol{x}) < 0\}$. The boundary of F is defined by $G = \{\boldsymbol{x}; g(\boldsymbol{x}) = 0\}$.

Firstly, the point process of jumps of $\boldsymbol{x}(t)$ is defined

$$N(A) = \#\left\{t \in A;\ \begin{array}{c}\text{A jump of one of the com-}\\ \text{ponent processes } x_i(t) \text{ at } t.\end{array}\right\}\#.$$

Then, the point process of outcrossings $U(A)$ out of S into F is defined by

$$U(A) = \#\{t \in A; g(\boldsymbol{x}(t-0)) \geq 0 > g(\boldsymbol{x}(t))\}\#. \tag{3}$$

Now, since $\boldsymbol{x}(t)$ changes its value at the jump times only, we obtain

$$\begin{aligned} &1 - \Pr(g(\boldsymbol{x}(t)) \geq 0; 0 \leq t \leq T) \\ &= 1 - \Pr(g(\boldsymbol{x}(0)) \geq 0, U(0,T) = 0) \\ &\leq \Pr(g(\boldsymbol{x}(0)) < 0) + \Pr(U(0,T) > 0). \end{aligned} \tag{4}$$

Since $U(0,T)$ is a non-negative random variable, we get

$$\begin{aligned} &1 - \Pr(g(\boldsymbol{x}(t)) \geq 0; 0 \leq t \leq T) \\ &\leq \Pr(g(\boldsymbol{x}(0)) < 0) + \mathbb{E}(U(0,T)). \end{aligned} \tag{5}$$

Therefore, we obtain an upper bound for the probability in equation 2. In the following we show that then under some regularity conditions, $U(A)$ is asymptotically an homogeneous Poisson process. This gives

$$\Pr(U(0,T) = 0) \approx e^{-\lambda_U \cdot T} \tag{6}$$

with λ_U the intensity of the point process $U(A)$.

Then we get

$$\begin{aligned} &1 - \Pr(g(\boldsymbol{x}(t)) \geq 0; 0 \leq t \leq T) \\ &\approx \Pr(g(\boldsymbol{x}(0)) < 0) + (1 - e^{-\lambda_U \cdot T}). \end{aligned} \tag{7}$$

2 The asymptotic behavior of the outcrossing point process

Consider a Poisson square wave vector process $\boldsymbol{x}(t)$ as defined in the last paragraph. Given is further a continuos limit state function $g(\boldsymbol{x})$ with $\min_{g(\boldsymbol{x})=0} |\boldsymbol{x}| = 1$ and there is a unique point $\boldsymbol{x}_0$ on G with $|\boldsymbol{x}_0| = 1$ such that near this point g is twice continously differentiable. This means that the function $|\boldsymbol{x}|$ has a minimum with respect to G at $\boldsymbol{x}_0$. We assume that this minimum is regular, i.e. the main curvatures $\kappa_1, \ldots, \kappa_{n-1}$ of G at $\boldsymbol{x}_0$ are less than unity. (This follows the differential geometry of G, see for example [5], chap. 12.)

We define two sequences of domains $S_\beta = \{\boldsymbol{x}; g(\beta^{-1}\boldsymbol{x}) \geq 0\}$ and $F_\beta = \{\boldsymbol{x}; g(\beta^{-1}\boldsymbol{x}) < 0\}$. Now we consider the point processes $U_\beta(A)$ of outcrossings of the process $\boldsymbol{x}(t)$ out of S_β into F_β, defined by

$$U_\beta(A) = \#\{t \in A; g_\beta(\boldsymbol{x}(t-0)) \geq 0 > g(\boldsymbol{x}(t))\}\#. \tag{8}$$

We will study the asymptotic behavior of the standardized point processes $U^s_\beta(A)$, defined by

$$U^s_\beta(A) = U_\beta(E_\beta^{-1} \cdot A), \text{ with } E_\beta = \mathbb{E}(U_\beta(0,1)). \tag{9}$$

These processes converge towards a Poisson point process under some regularity conditions.

It will be assumed that

$$\boldsymbol{x}_0 = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i,\ \sum_{i=1}^{n} \alpha_i^2 = 1,\ |\alpha_i| > 0,\ i = 1, \ldots, n. \tag{10}$$

Here $\boldsymbol{e}_i$ is the unit vector in the direction of the n-th component. This means that all direction cosines of $\boldsymbol{x}_0$ are not zero. We define

$$\alpha_0 = \min_{i=1,\ldots,n} |\alpha_i|. \tag{11}$$

Now, due to the definition of $\boldsymbol{x}_0$, there exists a $\delta > 0$ and an $\epsilon > 0$ such that for all $\boldsymbol{y}$ with $g(\boldsymbol{y}) \leq 0$ and $\sum_{i=1}^{n} y_i \alpha_i = \boldsymbol{y}^T \boldsymbol{x}_0 \leq (1-\delta)|\boldsymbol{y}|$ (i.e. the cosine of the angle between $\boldsymbol{y}$ and $\boldsymbol{x}_0$ is less than $1-\delta$)

$$|\boldsymbol{y}| > (1+\epsilon)|\boldsymbol{x}_0|. \tag{12}$$

Elsewhere there would be another point on the surface $G = \{\boldsymbol{x}; g(\boldsymbol{x}) = 0\}$ with unit distance to the origin.

Let be defined

$$\tilde{F}_\beta = \{\boldsymbol{x}; g_\beta(\boldsymbol{x}) < 0. \boldsymbol{x}^T \boldsymbol{x}_0 \leq (1-\delta)|\boldsymbol{x}|\} \tag{13}$$

$$F^*_\beta = F_\beta \setminus \tilde{F}_\beta. \tag{14}$$

see figure 2.

The number of the points of the outcrossing point process $U_\beta(A)$ is bounded from above by the number

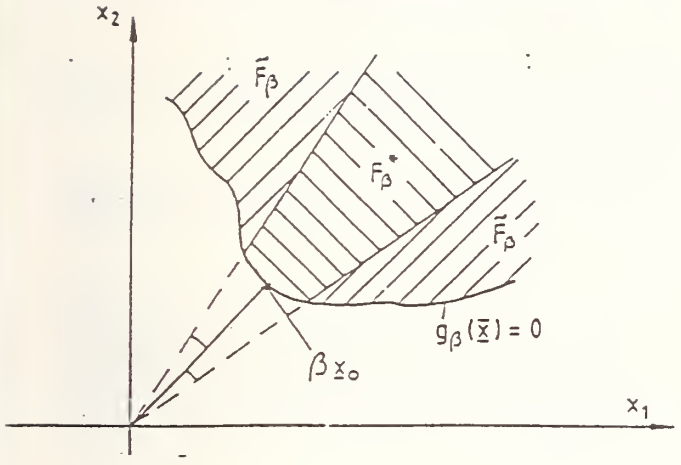


Figure 2: The sets F_β^* and $\tilde{F}_\beta$.

of the points of the point process $\tilde{U}_\beta(A)$, which counts all jumps of the process $\boldsymbol{x}(t)$ into a point in F_β

$$\tilde{U}_\beta(A) = \#\{t \in A; \boldsymbol{x}(t-0) \neq \boldsymbol{x}(t), g_\beta(\boldsymbol{x}(t)) < 0\}\#, \quad (15)$$

i.e. $\tilde{U}_\beta(A) \geq U_\beta(A)$. Further, for the point processes

$$\tilde{U}_\beta(A) = \#\{t \in A; g_\beta(\boldsymbol{x}(t-0)) \geq 0, \boldsymbol{x}(t) \in \tilde{F}_\beta\}\#, \quad (16)$$

we see by bounding from above that

$$\begin{aligned} &\mathbb{E}(\tilde{U}_\beta(A)) \quad (17)\\ &\leq |A| \left(\sum_{i=1}^n \lambda_i\right) \Pr(g_\beta(\boldsymbol{x}(t-0)) \geq 0, \boldsymbol{x}(t) \in \tilde{F}_\beta)\\ &\leq |A| \left(\sum_{i=1}^n \lambda_i\right) \Pr(\boldsymbol{x}(t) \in \tilde{F}_\beta)\\ &\leq |A| \left(\sum_{i=1}^n \lambda_i\right) \Pr(|\boldsymbol{x}(t)| \geq \beta(1+\epsilon))\\ &= o(\Phi(-\beta)), \beta \to \infty. \quad (18) \end{aligned}$$

The last inequality follows, since $|\boldsymbol{x}(t)|^2$ has a χ^2-distribution with n degrees of freedom and using equation 26.4.5 in [6] we get $\Pr(|\boldsymbol{x}(t)| \geq \beta(1+\epsilon)) \asymp (\beta(1+\epsilon))^{n-2} \exp(-\beta(1+\epsilon)) = o(\Phi(-\beta))$. Therefore $\mathbb{E}(U_\beta(A) - \tilde{U}_\beta(A)) = o(\mathbb{E}(U_\beta(A)), \beta \to \infty$, i.e. asymptotically the point process of jumps which leads to points in $\tilde{F}_\beta$ is negligible in comparison with the point process $\tilde{U}_\beta(A)$.

For the point processes $\tilde{U}_\beta(A)$ we get

$$\mathbb{E}(\tilde{U}_\beta(0,1)) = \left(\sum_{i=1}^n \lambda_i\right) \cdot \Pr(g_\beta(\boldsymbol{x}(t)) \leq 0). \quad (19)$$

Using the results of [7] and [8] about asymptotic approximations for the probability content of domains, we get asymptotically

$$\Pr(g_\beta(\boldsymbol{x}) \leq 0) \sim \Phi(-\beta) \prod_{j=1}^{n-1} (1-\kappa_j)^{-1/2}, \ \beta \to \infty, \quad (20)$$

and therefore

$$\begin{aligned} &\mathbb{E}(\tilde{U}_\beta(0,1)) \quad (21)\\ &\sim \left(\sum_{i=1}^n \lambda_i\right) \Phi(-\beta) \prod_{j=1}^{n-1} (1-\kappa_j)^{-1/2}, \ \beta \to \infty. \end{aligned}$$

Since $\tilde{U}_\beta(A) \geq U_\beta(A)$, this yields an upper bound for $\mathbb{E}(U_\beta(0,1))$.

We show that this upper bound is in fact an asymptotic approximation for this expected value as $\beta \to \infty$. This is done by proving that the expected value $\mathbb{E}(\tilde{U}_\beta(0,1) - U_\beta(0,1))$ is asymptotically negligible as $\beta \to \infty$. The point process $\tilde{U}_\beta(A) - U_\beta(A)$ consists of all points t of $\tilde{U}_\beta(A)$ with $g_\beta(\boldsymbol{x}(t-0)) < 0$, i.e. the points where the process was immediately before the jump in F_β and afterwards again in this domain.

The probability of obtaining a point in the domain $\tilde{F}_\beta$ after a jump is of order $o(\Phi(-\beta))$, as shown in equation 17. If the jump results in a point in the domain F_β^*, it is obvous that this is due to the occurrence of a new component amplitude, which has a value larger than $\alpha_0 \cdot \beta/2$, since $F_\beta^* \subset \{\boldsymbol{x}; \min_{i=1,\ldots,n} |x_i| > \alpha_0 \cdot \beta/2\}$. Therefore the expected number $\mathbb{E}(\tilde{U}_\beta(0,1) - U_\beta(0,1))$ is less than the expected number of jumps multiplied by the probability $\Pr(g_\beta(\boldsymbol{x}(t)) < 0)$ and the probability of a new component amplitude larger than $\alpha_0\beta/2$.

An upper bound for the first probability is one and the last probability is $\Phi(-\alpha_0\beta/2)$. Therefore

$$\begin{aligned} &\mathbb{E}(U_\beta(0.1) - U_\beta(0,1)) \quad (22)\\ &\leq \left(\sum_{i=1}^n \lambda_i\right) \Pr(g_\beta(\boldsymbol{x}(t-0)) < 0)\\ &\quad \times [\Phi(-\alpha_0\beta/2) + o(\Phi(-\beta))]. \end{aligned}$$

Asymptotically we get using equation 20 for the expected value

$$\begin{aligned} &\mathbb{E}(\tilde{U}_\beta(0.1) - U_\beta(0,1)) \quad (23)\\ &\sim \left(\sum_{i=1}^n \lambda_i\right) \Phi(-\beta) \prod_{j=1}^{n-1} (1-\kappa_j)^{-1/2}\\ &\quad \times [\Phi(-\alpha_0\beta/2) + o(\phi(-\beta))]\\ &\quad = o(\Phi(-\beta)), \beta \to \infty. \end{aligned}$$

This shows the asymptotic equivalence of the two point processes and gives finally

$$\begin{aligned} &\mathbb{E}(U_\beta(0,1)) \quad (24)\\ &\sim \left(\sum_{i=1}^n \lambda_i\right) \Phi(-\beta) \prod_{j=1}^{n-1} (1-\kappa_j)^{-1/2}, \ \beta \to \infty. \end{aligned}$$

If there is instead of one minimal distance point x_0 a finite number of points $x_1, \ldots, x_k$ with regular minimal distance points, an analogous result is obtained by splitting up the domain F into k disjoint sets $F_1, \ldots, F_k$ with $\cup_{i=1}^{k} F_i = F$ such that for each F_i exactly one point x_i lies on the boundary of F_i, by treating each set seperately and then adding the results. We get then

$$\mathbb{E}(U_\beta(0,1)) \sim \left(\sum_{i=1}^{n} \lambda_i\right) \Phi(-\beta) \left[\sum_{m=1}^{k} \prod_{j=1}^{n-1} (1-\kappa_{m,j})^{-1/2}\right], \ \beta \to \infty.$$

with the $\kappa_{m,j}$ the main curvatures of the surface G at x_m.

We prove the convergence to a Poisson process for the standardized processes $U_\beta^s(A)$ in the next paragraph. To this purpose we replace $U_\beta(A)$ by an approximating point process

$$\hat{U}_\beta(A) = \#\{t \in A; g(x(t)) < 0, |x(t-0)| \leq \beta - \log(\beta)\}\#$$

This is the point process of all outcrossings from S_β into F_β, which start from a point in the sphere around the origin with radius $\beta - \log(\beta)$; since this sphere is inside S_β, we have $\hat{U}_\beta(A) \leq U_\beta(A)$ (see figure 3).

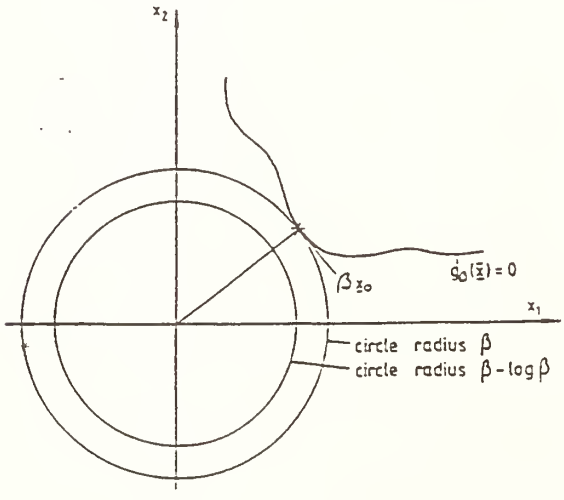


Figure 3: Approximating point process $\hat{U}_\beta(A)$

Using similar arguments as before it can be shown that the two point processes $U_\beta(A)$ and $\hat{U}_\beta(A)$ are asymptotically equivalent. Therefore, if $\hat{U}_\beta^s(A)$, i.e. in standardized form, converges to a Poisson process as $\beta \to \infty$, this is valid for $\tilde{U}_\beta^s(A)$ and $U_\beta^s(A)$ too. In the next paragraph it is shown that for $\hat{U}_\beta^s$ such a convergence can be proved.

3 Convergence to a Poisson process

To show the convergence of the standardized processes $U_\beta^s(A)$ to a homogeneous Poisson process, the following result is used.

Given is a sequence $N_\beta(A)$, $\beta \geq 1$ of stationary and orderly point processes with the following properties:

1. $E_\beta = \mathbb{E}(N_\beta(0,1)) \to 0$, $\beta \to \infty$.
2. There exists a function $\alpha : (0,\infty) \to \mathbb{R}$ with $\alpha(\tau) \to 0$ as $\tau \to \infty$ such that for all events A_β and B_β, where A_β depends only on the behavior of the point process $N_\beta(A)$ until the time t and B_β only on the behavior of $N_\beta(A)$ after the time $t + \tau E_\beta^{-1}$

   $$|\Pr(A_\beta \cap B_\beta) - \Pr(A_\beta) \cdot \Pr(B_\beta)| \leq \alpha(\tau E_\beta^{-1}). \quad (25)$$

   i.e. the standardized processes $N_\beta^s(A)$ are uniformly mixing with mixing coefficient $\alpha(\tau)$.
3. $\int_0^{E_\beta^{-1/2}} \int_0^{E_\beta^{-1/2}} p_\beta(t_1,t_2)dt_1\ dt_2 = o(E_\beta^{1/2})$, $\beta \to \infty$, where $p_\beta(t_1,t_2)$ denotes the two-dimensional product density of the point processes $N_\beta(A)$.

Under the conditions $1-3$ above the standardized processes $N_\beta^s(A) = \tilde{N}_\beta(E_\beta^{-1} \cdot A)$ converge to an homogeneous Poisson point process with intensity 1.

As outlined in [9], p. 37 for a stationary and orderly point process $N_\beta(A)$ the factorial moment $\mathbb{E}(N_\beta(0,t))(N_\beta(0,t)-1))$ is given by

$$\begin{aligned} &\mathbb{E}(N_\beta(0,t))(N_\beta(0,t)-1)) \quad (26)\\ &= \int_0^t \int_0^t p_\beta(t_1,t_2)dt_1\ dt_2 \\ &= 2 \cdot \int_0^t \int_0^{t_2} p_\beta(t_1,t_2)dt_1\ dt_2. \end{aligned}$$

This result can be shown by using theorem 1.3 in a paper of Volkonskii and Ronzanov ([10]). This theorem says that stationary and orderly point processes converge to a homogeneous Poisson point process, if the conditions 1 and 2 are fulfilled and further for $t \to 0$ and $\beta \to \infty$ always

$$\Pr(N_\beta(0, E_\beta \cdot t) > 0) \to 0. \quad (27)$$

But this condition can be replaced by the condition given in equation 1.38 in the paper of Volkonskii and Rozanov, i.e.

$$\mathbb{E}(N_\beta^2(0, E_\beta t))/\mathbb{E}(N_\beta(0, E_\beta t) \to 1,\ t \to 0,\ \beta \to \infty. \quad (28)$$

Due to the stationary of the processes it is sufficient to show this for the intervals $(0, E_\beta^{-1/2})$.

The third condition above says that

$$\mathbb{E}\left(N_\beta(0,E_\beta^{-1/2})(N_\beta(0,E_\beta^{-1/2})-1)\right) \quad (29)$$
$$= o(\mathbb{E}(N_\beta(0,E_\beta^{-1/2}))),\ \beta\to\infty$$

and therefore, since $\mathbb{E}(N_\beta(0,E_\beta^{-1/2}))\to 0$ as $\beta\to\infty$, we have

$$\mathbb{E}(N_\beta^2(0,E_\beta^{-1/2}))/\mathbb{E}(N_\beta(0,E_\beta^{-1/2}))\to 1,\ \beta\to\infty. \quad (30)$$

Therefore then relation 28 holds. So the conditions 1-3 are sufficient for convergence to a Poisson process.

We show that the processes $\hat{U}_\beta(A)$ fulfill the conditions; from this follows the result then for $U_\beta(A)$ too. To prove this we note that equation 24 shows that the first condition is fulfilled, since $E_\beta = \Phi(-\beta)\left(\sum_{i=1}^n \lambda_i\right)\prod_{i=1}^{n-1}(1-\kappa_i)^{-1/2}\to 0$ as $\beta\to\infty$.

The second condition is fulfilled, if we take as strong mixing function $\alpha(\tau)=\sum_{i=1}^n e^{-\lambda\tau}$. This follows from the fact that the underlying point processes $N_i(A)$ are homogeneous Poisson point processes with intensities λ_i and therefore $\Pr(N_i(0,\tau)=0)=e^{-\lambda_i\tau}$.

To show the third condition, we split the joint intensity function, denoted by $f_\beta(t_1,t_2)$, into two functions:

$$f_\beta(t_1,t_2) \quad (31)$$
$$= P_1(t_2-t_1)f_\beta^1(t_1,t_2)+P_2(t_2-t_1)f_\beta^2(t_1,t_2)$$

with

$$P_1(t_2-t_1) \quad (32)$$
$$= \Pr\underbrace{\left(\begin{array}{c}\text{All point processes } N_i,\ i=1,\dots,n\\ \text{have at least one point in } (t_1,t_2).\end{array}\right)}_{=I(t_1-t_2)}$$
$$P_2(t_2-t_1)=1-P_1(t_2-t_1)=\Pr(I(t_1-t_2)^c).$$

Further $f_\beta^1(t_1,t_2)$ (resp. $f_\beta^2(t_1,t_2)$) denotes the conditional density of the point process $\hat{U}_\beta(A)$ under condition $I(t_2-t_1)$ (resp. under condition $I(t_2-t_1)^c$). In the first case all point processes N_i, $i=1,\dots,n$ have at least one point between the two time points t_1 and t_2. The probability for this event is $P(t_2-t_1)=1-\prod_{i=1}^n \Pr(N_i(t_2,t_1)=0)=1-\prod_{i=1}^n e^{-\lambda_i(t_2-t_1)}\le 1$.

The condition means that all loads have changed and that therefore the behavior of the point processes $\hat{U}_\beta(A)$ at the points t_1 and t_2 is independent under this condition. Therefore the conditional joint intensity $f_\beta^1(t_1,t_2)$ is just the product of the constant one-dimensional intensity λ_β of the process at these points. Since the process asymptotically equivalent to the process $U_\beta(A)$, we get from equation 24 $\lambda_\beta\sim E_\beta$ as $\beta\to\infty$. Therefore, we have

$$P_1(t_2-t_1)f_\beta^1(t_1,t_2)\le E_\beta^2+o(E_\beta^2). \quad (33)$$

Integrating over this function gives

$$2\cdot\int_0^{E_\beta^{-1/2}}\int_0^{t_2} P_1(t_2-t_1)f_\beta^1(t_1,t_2)dt_1\ dt_2 \quad (34)$$
$$\le 2\cdot\int_0^{E_\beta^{-1/2}}\int_0^{t_2}(E_\beta^2+o(E_\beta^2))dt_1\ dt_2 = o(E_\beta^{1/2}),\ \beta\to\infty.$$

In the second case, at least one of the point processes N_i had no point in the time interval between t_1 and t_2 and therefore at least one of the amplitudes did not change. Since these processes are independent Poisson point processes, the probability for this is bounded by $0\le P_2(t_2-t_1)\le\sum_{i=1}^n e^{-\lambda_i(t_2-t_1)}$.

The conditional intensity in this case is less than the intensity of the point process of occurrences of a component larger than $\log(\beta)$ at t_2, since only in this case a jump of the point process $U_\beta^*(A)$ will be at t_2, multiplied by the one-dimensional intensity of $\bar{U}_\beta(A)$; this intensity is asymptotically equal to E_β.

Therefore, as $\beta\to\infty$

$$P_2(t_2-t_1)f_\beta^2(t_1,t_2) \quad (35)$$
$$\le\left(\sum_{i=1}^n\lambda_i\right)\Phi(-\log(\beta))2E_\beta\left(\sum_{i=1}^n e^{-\lambda_i(t_2-t_1)}\right).$$

Integrating over this function gives

$$2\cdot\int_0^{E_\beta^{-1/2}}\int_0^{t_2} P_2(t_2-t_1)f_\beta^2(t_1,t_2)dt_1\ dt_2$$
$$\le 2\left(\sum_{i=1}^n\lambda_i\right)\Phi(-\log(\beta))2E_\beta$$
$$\cdot\int_0^{E_\beta^{-1/2}}\int_0^{t_2}\left(sum_{i=1}^n e^{-\lambda_i(t_2-t_1)}\right)dt_1\ dt_2$$
$$= K_0(E_\beta^{1/2}\Phi(-\log(\beta))+o(E_3^{1/2}))=o(E_\beta^{1/2}),\ \beta\to\infty,$$

with K_0 a constant.

Together with the upper bound for the first term in equation 34 this yields

$$\mathbb{E}(U_\beta(0,E_3^{-1/2})(U_\beta(0,E_\beta^{-1/2})-1)) \quad (36)$$
$$= 2\cdot\int_0^{E_\beta^{-1/2}}\int_0^{t_2} f_\beta(t_1,t_2)dt_1\ dt_2 = o(E_\beta^{1/2}),\ \beta\to\infty.$$

So the third condition holds too. Therefore we have for the standardized point process U_β^s the convergence to a Poisson point process, i.e.

$$\Pr(U_\beta^s(A)=0)\sim e^{-|A|},\ \beta\to\infty. \quad (37)$$

4 Summary

In this paper a Poisson convergence theorem for the point process of level crossings of functions of independent Poisson square wave processes is shown. The idea is to split the two-dimensional joint intensity of the point process into two conditional intensities, one describing the behavior under independence of the processes at the two time points and one under dependence. Then by estimating it from above, it is shown that the second intensity can be neglected asymptotically; this gives the Poisson convergence.

5 Acknowledgement

The main results of this paper are based on research done in a project supported by the Deutsche Forschungsgemeinschaft.

References

[1] H. Madsen, S. Krenk, and N.C. Lind. *Methods of Structural Safety*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1986.

[2] K. Breitung und R. Rackwitz. Nonlinear combination of load processes. *Journal of Structural Mechanics*, 10(2):145–166, 1982.

[3] K. Schrupp and R. Rackwitz. Outcrossing rates of marked Poisson cluster processes in structural reliability. *Applied Mathematical Modelling*, 12:482–490, 1988.

[4] M. Hohenbichler and R. Rackwitz. Non-normal dependent vectors in structural safety. *Journal of the Engineering Mechanics Division ASCE*, 107(6):1227–1241, 1981.

[5] J.A. Thorpe. *Elementary Topics in Differential Geometry*. Springer, New York, 1979.

[6] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1965.

[7] K. Breitung. Asymptotic approximations for multinormal integrals. *Journal of the Engineering Mechanics Division ASCE*, 110(3):357–366, 1984.

[8] K. Breitung and M. Hohenbichler. Asymptotic approximations for multivariate integrals with an application to multinormal probabilities. *Journal of Multivariate Analysis*, 30:80–97, 1989.

[9] D.R. Cox and V. Isham. *Point Processes*. Chapman and Hall, London, 1980.

[10] V.A. Volkonskii and Yu.A. Rozanov. Some li
theorems for random functions. I. *Theory of Pr
ability and its Applications*, 4(2):178–197, 1959

Meso-Scale Estimation Of Expected Extreme Values

Burton, R.M., Goulet, M.R., and Yim, S.C.S.
Oregon State University, Corvallis, OR

We consider algorithms for estimating the expected maximum value of a time series for a period in the future given past observations. This is a "mid-range" problem in which the long term asymptotics of extreme value theory do not apply. There are essentially two approaches, estimating an "extremal index" and the "Poisson clumping heuristic". Variations on these methods are tested with simulated Gaussian data. Similarities in performance are explained rigorously.

INTRODUCTION

We consider the following problem. Given a time series from a stationary process $\{X_i\}_{i=1}^{\infty}$, define the expected maximum $E[M_{N,N'}]$, where $N' > N$ and

$$M_{N,N'} = \max_{N+1 \le i \le N'} X_i.$$

The problem is to find a good estimator of $E[M_{N,N'}]$ based on observations $\{X_i\}_{i=1}^{N}$. We will alway consider Gaussian time series but it will be clear that our methods apply more generally.

Here we describe several estimators of $E[M_{N,N'}]$, present some empirical results and give some theoretical explanations of our results.

DESCRIPTION OF THE ESTIMATORS

Time Rescaling. The idea is to estimate an *extremal index* of the process for this time scale. We say that ρ is the extremal index for $\{X_i\}_{i=1}^{\infty}$ on the scale of N if M_N has approximately the same distribution as the maximum of $[\rho N]$ independent random variables with the same distribution as X_1, i.e Gaussian.

There are essentially three choices to be made in approach.

The first is whether to use the time series itself or an enveloped version of it. Given the data, $\{X_i\}_{i=1}^{N}$ we may construct the discrete Hilbert transform $\{Y_i\}_{i=1}^{N}$. The process $\{R_i\}_{i=1}^{N}$ defined by

$$R_i = \sqrt{X_i^2 + Y_i^2}$$

is called the *analytic envelope*. It covers the "surface" of the time series, smoothing out the oscillations. The maximum of the envelope is close to that of the original process, especially in the narrow band case. It has the further computational advantage of being Rayleigh distributed. This is described in detail in Ref. [1, 2]. We call these choices *direct* and *enveloped*.

The second choice is how carefully to compute the expected maximum of n independent Gaussian random variables (or Rayleigh in the enveloped case) as a function of n. One could either use a good but computationally intensive numerical approximation or an asymptotic formula, $\sqrt{2 \log n}$. We will refer to these choices as *strong* and *weak* and call this approximation we use $L(n)$.

The third choice is how carefully to fit the empirical expected maxima as computed from the data to $L(n)$. One possibility is to use one value of n_0, say 50. Find the average of the maximum value in the data for non-overlapping windows of length n_0. This value $\hat{L}(n_0)$ is the empirical expected maximum at n_0. To estimate the extremal index then find ρ so that $\hat{L}(n_0) = L(\rho n_0)$.

The other possibility is to compute multiple window lengths, i.e. to compute $\hat{L}(n) = L(n)$ for various n. If we take n to be powers of 2 then the computation time is not large because we may "nest" the computations of the maxima. We will refer to these choices as *single window* and *multiple window* methods.

We note that the prevailing method among ocean engineers was the enveloped, strong, single window method.

Poisson Clumping. Another possible estimator is suggested by Aldous' use of the Poisson clumping heuristic, Ref. [3]. This heuristic assumes that the set of t for which $X_t > b$ is given by random sets distributed as a Poisson process. We make the further assumption that these random sets are intervals.

Consider, for b relatively large

$$\{t | X_{t-1} < b, X_t \geq b\}$$

to be distributed as a Poisson process with rate λ_b. The following fundamental relation is assumed

$$P[X_1 \geq b] = \lambda_b E[C_b],$$

where C_b is the random length of an interval (clump) in which the time series spends above a given value b. The event $[M_N < b]$ is equivalent to

$$\{t | X_{t-1} < b, X_t \geq b\} = \emptyset.$$

So by the Poisson assumption,

$$P[M_N < b] = e^{-\lambda_b N},$$

and by the fundamental identity

$$P[M_N < t] = e^{-P[X_t \geq b]N/E[C_b]}.$$

Hence we have

$$E[M_N] = \int_0^\infty (1 - e^{-P[X_1 \geq b]N/E[C_b]})\, db. \qquad (1)$$

The work now reduces to estimating $E[C_b]$. To do this we fix a value of b and average the length of the intervals where the time series is above b. Varying b and plotting $E[C_b]$ versus b yields data which is well fit by a curve of the form

$$y = b^{-\gamma}/A.$$

Substituting this curve into (1) yields our estimator,

$$\hat{E}_P[M_{N,N'}] = \int_0^\infty (1 - e^{-AP[X_1 \geq b](N'-N)b^\gamma})\, db.$$

As before we may use either the original or enveloped data. Note that the above analysis assumes that the clumps are intervals so one guesses that enveloping narrow band data would be advantageous.

EMPIRICAL RESULTS

We have described 10 possible algorithms in all.

In earlier work, Ref. [2], we investigated several algorithms. The algorithm used by most ocean engineers was due to Pierce, Ref. [1]. In our terminology this was an enveloped, strong, single window time rescaling method.

We proposed to modify this by removing the envelope, that is to use instead the direct, strong single window time rescaling method. In cases where computation ease was paramount we proposed the direct, weak, multiple window rescaling method. These were compared with the direct Poisson clumping algorithm. Here we describe the results of these simulations. In subsequent versions of this paper we will also include study of the enveloped Poisson clumping algorithm. Work continues on the other variations.

In this study two types of Gaussian time series are used. The first is a second order autoregressive moving average.

$$X_n = aX_{n-1} + bX_{n-2} + Z_n,$$

where the Z_n is are independent identically distributed Gaussian random variables.

The second type is intended to simulate random waves in the ocean and are obtained by superposition of sinusoids, with amplitudes specified by the Pierson-Moskowitz and JONSWAP spectrums, Ref. [4]. One thousand cosines with unequal frequency spacings and uniformly random phases are employed. More detail on these processes is given in Ref. [2].

These time series are run for various parameters and the expected maximum are estimated by the algorithms. The mean relative errors are computed.

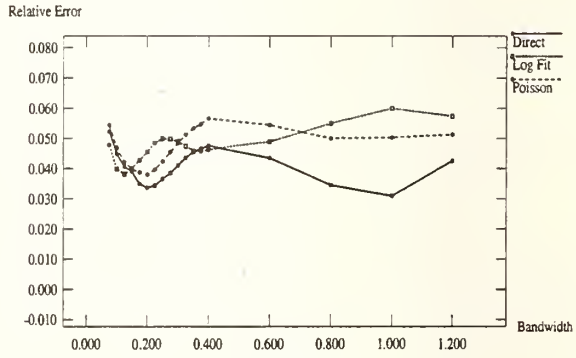


Figure 1

Our results are pictured in Figures 1-3. To summarize the ARMA experiments (Figure 1), the Direct Method consistently gives the estimator with minimal relative error, while the Poisson clumping and Log Fit methods yield estimators with relative errors under 6%. It is interesting to note that the results for Poisson Clumping and the direct, strong, single window time rescaling method follow each other.

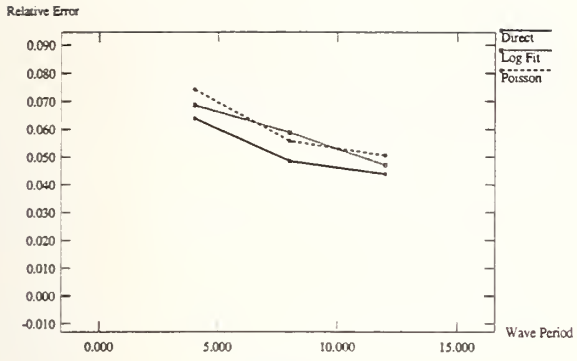


Figure 2

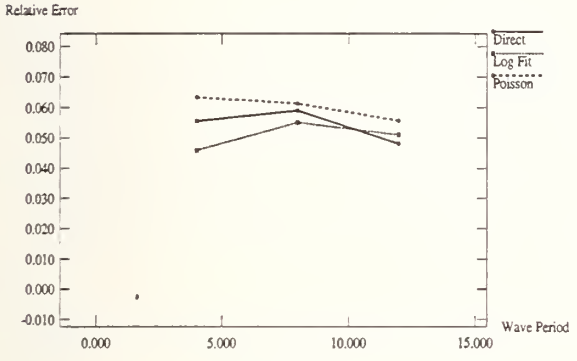


Figure 3

For the simulated ocean waves the results are similar. In the case of the Pierson-Moskowitz spectrum (Figure 2) the Direct Method consistently provides the best estimator, regardless of the dominant wave length. The JONSWAP spectrum (Figure 3) provides a narrow band case and all the techniques yield relative errors between 4% and 7%, while the Log Fit provides the best estimator in two instances.

CONCLUSION

The theoretical underpinning for these algorithms is given by theorems of O'Brien, Ref. [5], and independently by Rootzen, Ref. [6]. There it is shown for processes satisfying a strong mixing condition (as ours do) that for long enough time scales there is an extremal index which in turn gives the Poisson clump structure of the exceedance process. Thus it is not surprising to find the direct, strong, single window estimator and the direct Poisson clumping estimator in close agreement.

Perhaps the most striking result in this study is the performance of the direct, weak, multiple window estimator. It is the simplest conceptually and algorithmically, and gives relative errors near 6%.

References.

[1] R.D. Pierce. Extreme value estimates for arbitrary bandwidth gaussian processes using the analytic envelope. *Ocean Engineering*, 12(6):493, 1985.

[2] R.M. Burton S.C.S. Yim and M.R. Goulet. Practical methods of extreme value estimation based on measured time series for ocean systems. *Ocean Engineering*, 19(3):219–238, 1992.

[3] D. Aldous. *Probability Approximations via the Poisson Clumping Heuristic.* Springer-Verlag, New York, 1989.

[4] T. Sarpkaya and M. Isaacson. *Mechanics of Wave Forces on Offshore Structures.* Van Nostrand Reinhold, New York, 1981.

[5] G.L. O'Brien. Extreme values for stationary and markov sequences. *Annals of Probability*, 15:281–291, 1987.

[6] H. Rootzén. Maxima and exceedences of stationary markov chains. *Advances in Applied Probability*, 20:371–390, 1988.

An Expert System Prototype For The Analysis Of Extreme Value Problems

Castillo, E., Alvarez, E., Cobo, A. and Herrero, M.T.
University of Cantabria, Santander, Spain

This paper presents an expert system prototype for the analysis of extreme value problems. The system includes a package of computer aided instruction for the most common concepts in Extreme Value Theory and illustrative examples of applications. The user can navigate through the information at wish. The system incorporates a computer program to simulate, estimate, draw samples on extreme probability papers and determine the domain of attraction of a parent from samples, based on the Pickands' and/or the curvature methods. A set of rules controls the selection of probability papers and estimation methods adequate to given problems.

Key words: Extreme value problems, expert system, simulation, estimation.

1 Introduction

Extreme value problems are very frequent to engineers. In fact, in many engineering situations, design is based on the probability of occurence of extreme values of single or combined random variables. When dealing with extreme value problems, engineers and scientists find some difficulties due to the following facts:

- Extreme value theory is complicated.
- Extreme value theory is not easily available.
- In general, technicians have not been prepared to deal with this problem.

In spite of its importance, extreme value theory has not been included in standard curricula. Even in some specific fields, such as statistics, for example, a very large part of the student population ignore fundamental aspects of this theory. This problem is even more important in engineering areas where a large amount of technical material must be covered.

On the other hand, extreme value theory is not easily available to those who need it. Most of the advances have been published in journals and books mainly addressed to probability and statistic specialists and using a langauge difficult to understand for engineers and scientists.

The consequence of all the above is that important errors and inconsistencies have been made in the past, such as: the use of incorrect models or probability papers, the use of non-stable models, the use of wrong estimation methods, etc. It is easy to find examples of limit models for minima that are used for maxima and vice versa, or examples of incorrect use of probability papers. In other cases, non-stable models, either in extreme or truncation operations, lead to confusions and lack of consistency.

All the above justifies the need for expert systems. We must remind the reader that expert systems are useful mainly when (see Ref[1]):

- there is a lack of human experts
- there is lack of knowledge among those who need it
- one needs more reliable solutions
- one needs cost reduction.

To our knowledge no expert system exists covering all these needs. This paper addresses this problem and presents a simple prototype showing some of the excellences an expert system should have.

2 Steps in the development of an extreme value expert system

Expert system design must be carefully programmed if success is desired. Some of the main steps to be followed are (Ref[2]):

1. Statement of the problem to be solved.
2. Searching for human experts and/or data.
3. Design of the expert system.
4. Selection of the development tool, shell or programming language.
5. Development of a prototype.
6. Prototype checking.
7. Refinement and generalization. Final expert system.
8. Maintenance.
9. Updating.

The first step consists of defining the problem to be solved. Under no circunstances should time spent on this period be curtailed, and work should be done with rigor and precision. All extra time dedicated to this step will be saved in the following steps. This step implies identifying all or a large part of the possible difficulties to be encountered when dealing with extreme values and the possible solutions to be employed. When this is clear, a decision about which of them are to be solved by the expert system must be taken.

Once the problem has been completely defined, one must look for human experts able to solve it with a reasonable chance of success.

The third step is the design of the expert system, which includes the structures for knowledge storage, the inference engine, the explanation subsystem, the user interface and so on.

In the following step we must decide whether to use a shell or a programming language. It is important to avoid useless efforts that are also expensive. The final steps cover the prototype development, checking, refinement and updating.

3 Minimal requirements

In this section we discuss some of its minimal requirements. The expert system should include at least:

- Computer aided instruction on extremes.
- Tools for doing statistics of extremes.
- Tools for gaining experience and expertise.
- Bibliographic information.

The first part should cover the most important concepts relevant to extreme values by means of: theory, illustrative examples of applications, interactive methods, such as, animations, examples, hypertext (user driven navigation through information), hypermedia, guided tours, etc.

The new computer aided instruction techniques, based on hypertext and hypermedia, allow an easy and quick development of a teaching module. A guided tour guarantees the most important concepts to be covered and aprehended by the user.

Determination of design values or probability assessments are the result of an iterative method based on a combination of different steps as drawing data, selecting models, estimating parameters, etc. Thus, the tools for doing statistics of extremes should cover:

- Drawing data on probability papers and other graphic representations.
- Model selection.
- Estimation.
- Determination of domains of attraction.
- Determination of design values.

Another interesting role of an expert system is its possible contribution to the user in gaining experience and expertise. For this to be possible, the system should include:

- Access to real cases (data base).
- Simulation.

The easy access to previous experience and methods facilitates the solution of many real problems. The system could include practical cases and different alternative solutions given to some typical problems.

A simulation system allows for gaining experience and testing the appropriateness of some temptative methods. Note that testing of several alternatives can lead to a very useful information thus facilitating the final decision.

Finally, a complete bibliographic information including cross references should be available. This facilitates the access and navigation of the interested reader through the information.

4 Prototype description

The implementation of an expert system for the analysis of extremal problems is a complicated task, which must involve a group of people. In this section we present a simple prototype including only some of the above possibilities to show the convenience of such a system.

We shall divide the exposition in three parts: the computer aided instruction subsystem, the statistical tools subsystem and the proper expert system.

4.1 Computer aided instruction subsystem

This package covers the material indicated on the menu card (see figure 1), where different options can be selected.

The user can select one of the topics by just clicking the mouse on it and the system shows the relevant information associated with it. In some cases this produces a new menu, as in the case of figure 2, where new options can be selected.

With the purpose of introducing some concepts some animations are used, such as those illustrated in figure 3, where samples are drawn at random from the population and in figure 4, where the whole process of simulating order statistics is animated.

In other cases, graphical information is used to illustrate concepts, such as that indicated in figure 5, where the curvature of tails is used to decide about domains of attractions. Other cards are used for definitions of relevant concepts (see figure 6). Information is structured in such a way that the user can go back and forth at wish, consulting examples, definitions, graphics, animations, etc.

4.2. Statistical tools subsystem

The expert system incorporates a set of computer programs to do a statistical analysis of data.

If we choose reading data from a file, a dialog with all available files is shown (see figure 7). Then, the file is opened and data appears in the text window.

Once the statistical program is launched, a new document is opened and two empty windows on the computer screen are obtained. They hold the sample data and its associated drawing, respectively.

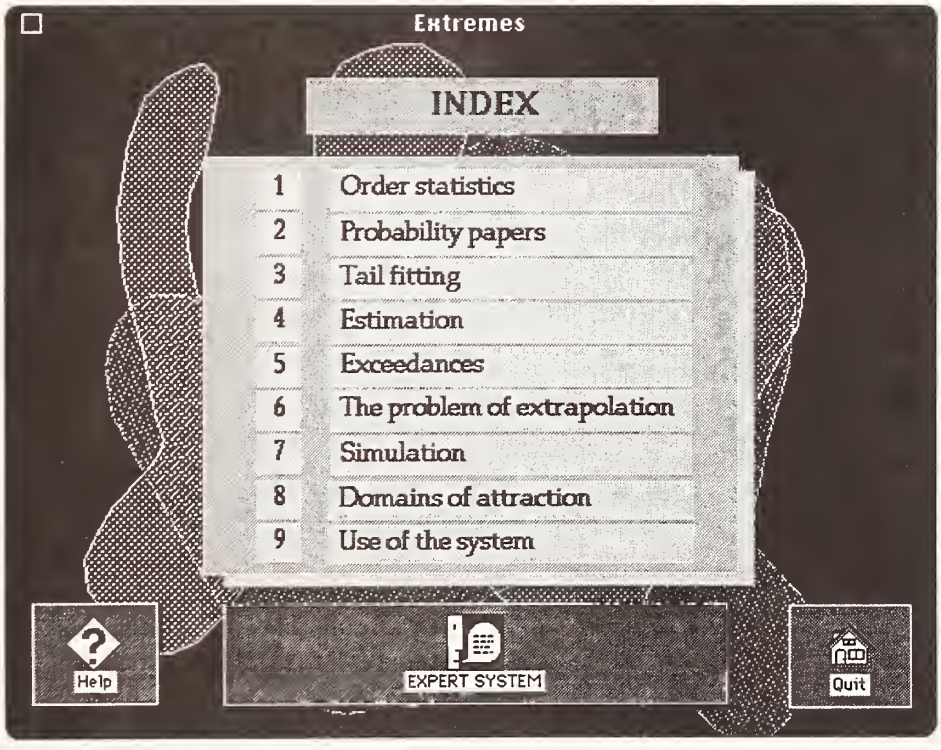


Figure 1: The main menu.

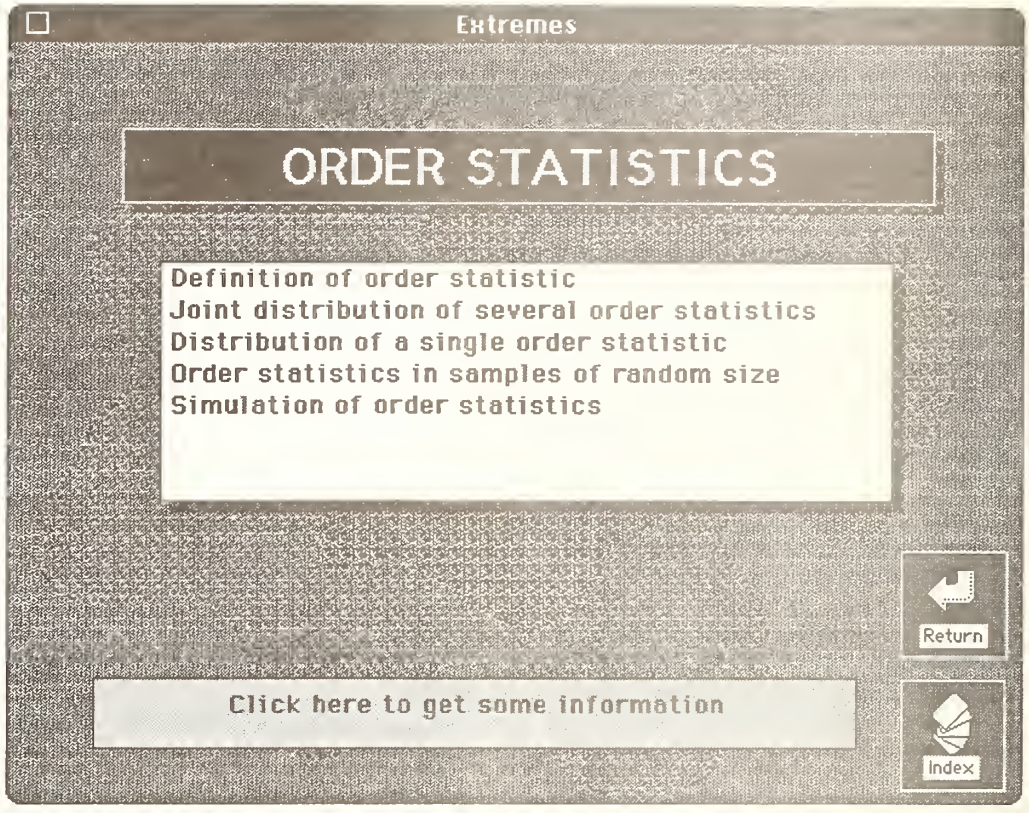


Figure 2: The order statistics menu.

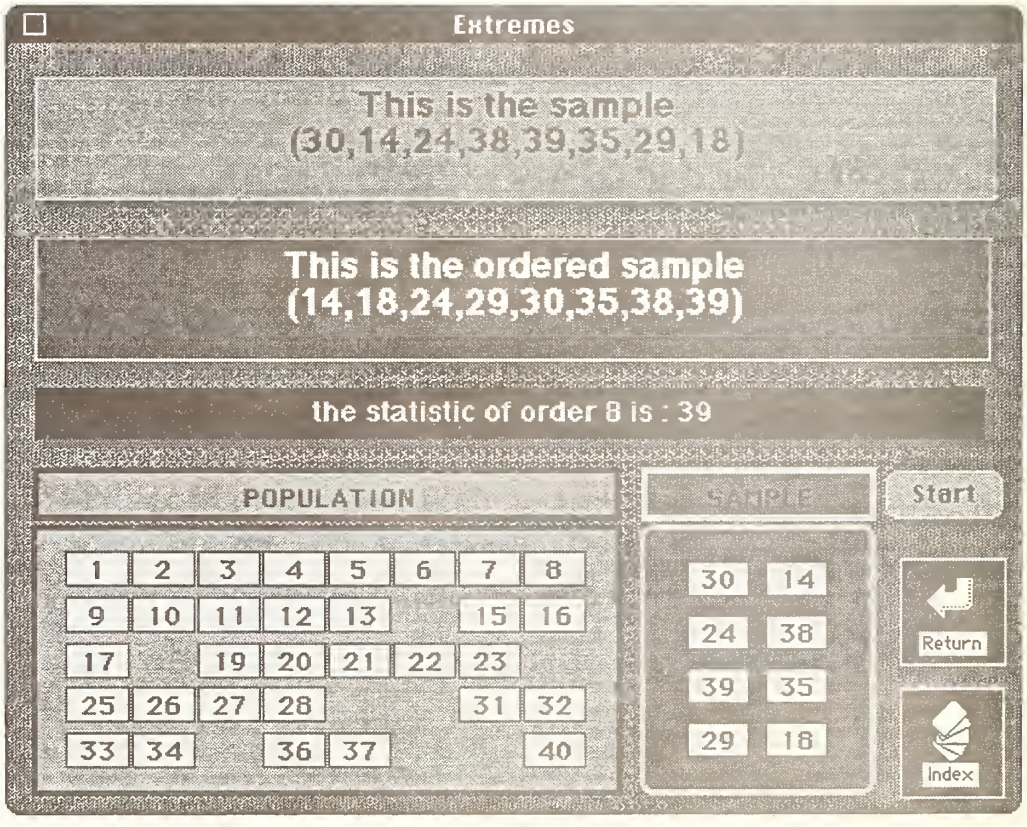


Figure 3: The concept of order statistic.

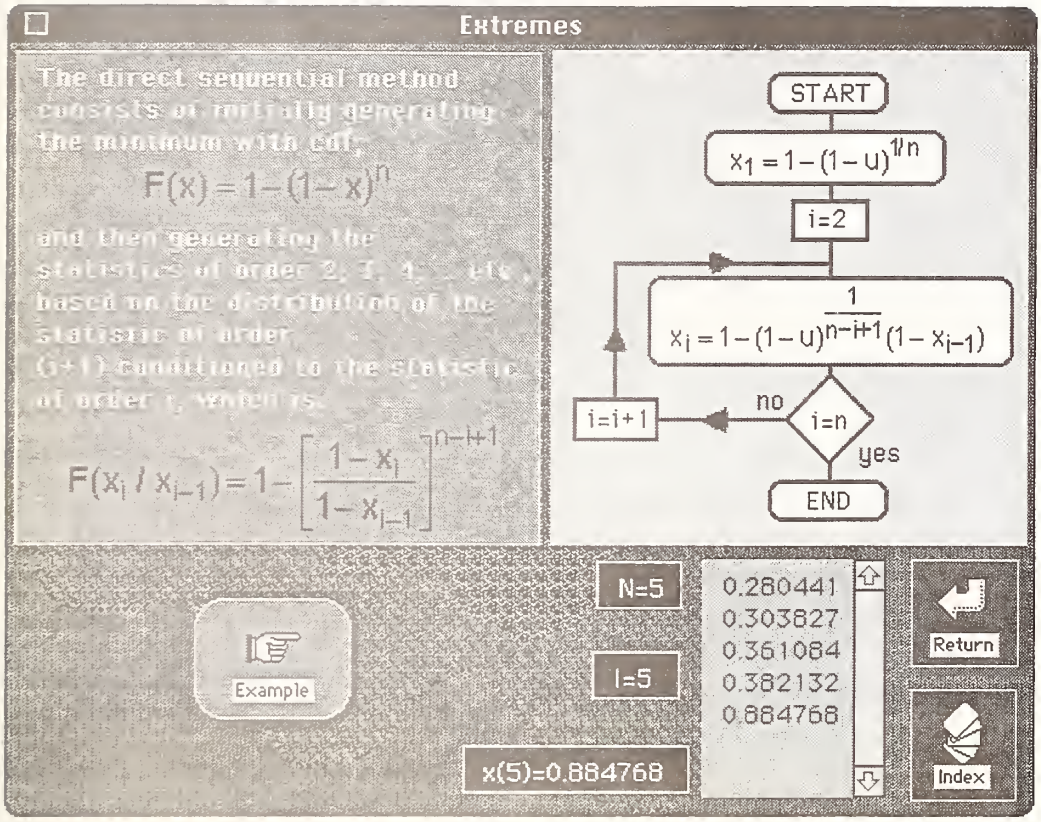


Figure 4: Simulation of order statistics.

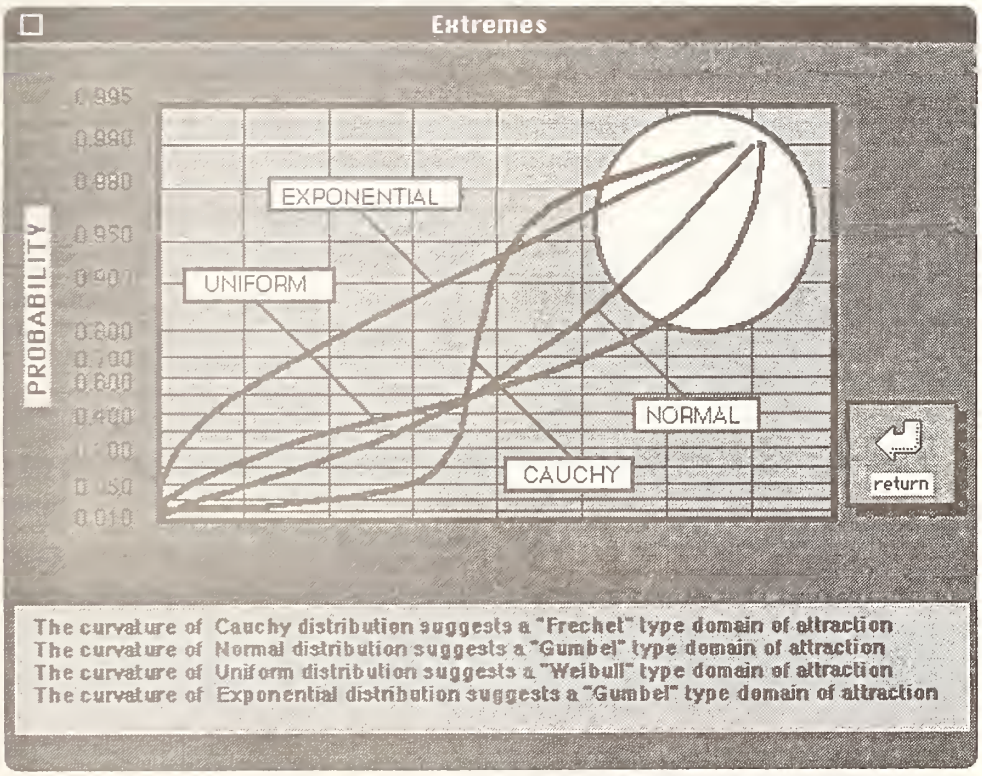


Figure 5: Determining maximal domains of attraction.

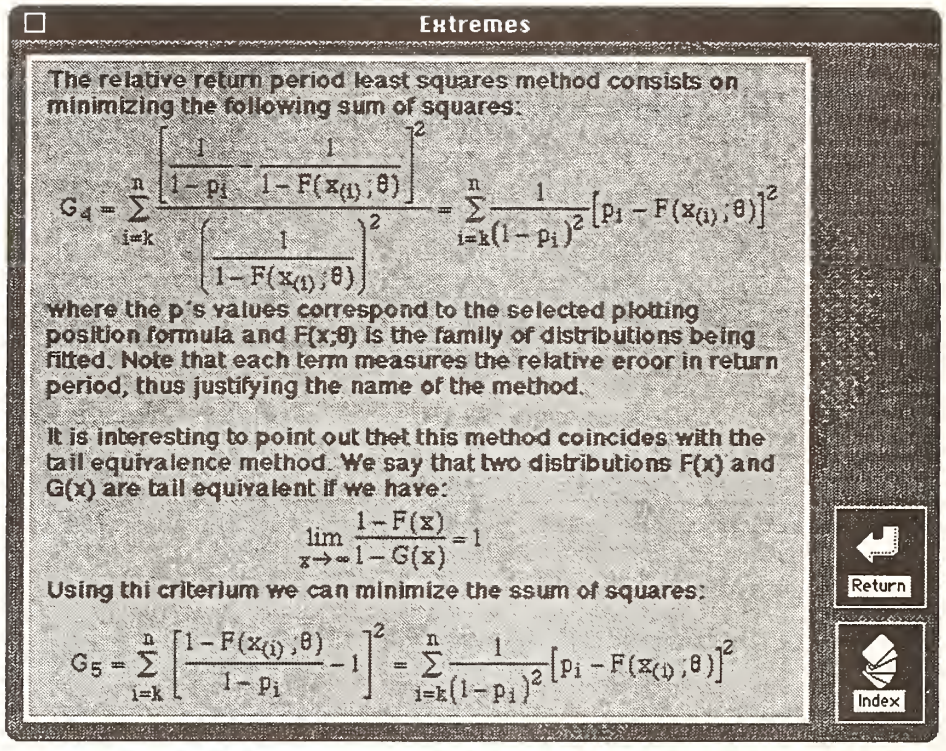


Figure 6: Least squares methods.

The program allows for the four operations shown in the main menu of figure 7.

Initially, due to the lack of data, only the "*simulate*" option is available and the rest appear as dimmed. At this step we can choose either reading data from an external file, typing data directly on the text window or simulate data.

If we choose the "*simulate*" option, the dialog in figure 8 appears to allow us to choose the distribution, its corresponding parameters, sample sizes and range of order statistics to be simulated. In any case, the text window ends up with the working data. At this step all the options above become available.

Selection of the "*draw*" option of the main menu leads to the dialogs in figures 9 and 10, where the desired probability paper and plotting point position formula are selected. Then, the drawing of the indicated probability paper and the sample is obtained (see figure 11).

Selection of the "*estimate*" option leads to the dialogs in figures 12 and 13, where the method of estimation and the family of distributions to be used in the fitting procedure are selected. Then, the range of order statistics is given and the system initiates the estimation process and informs the user of its progress. At the end, it gives the estimates and, in some cases, the variance-covariance matrix of estimates as is shown in figure 14.

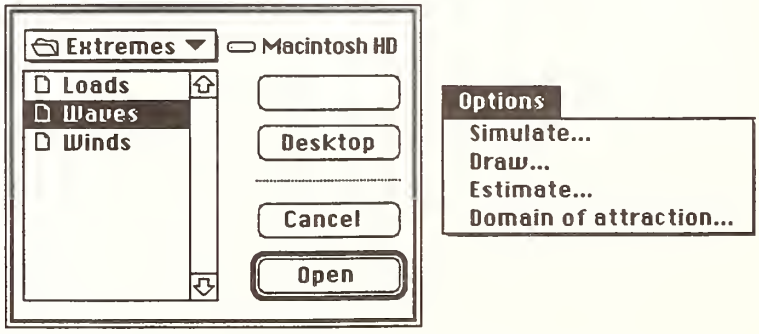


Figure 7: Reading from a file and the options menu.

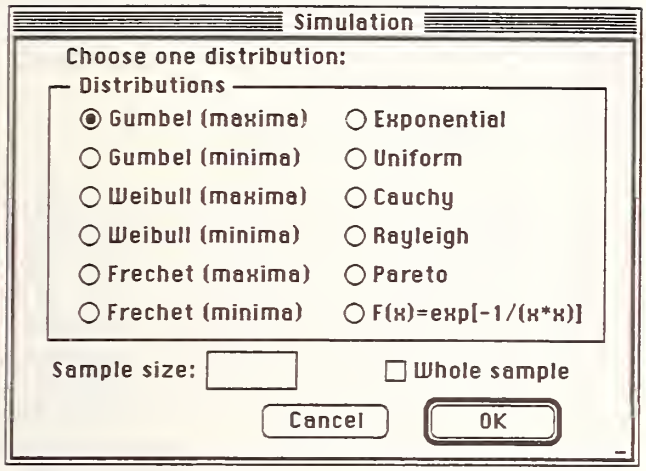


Figure 8: List of distributions to be simulated.

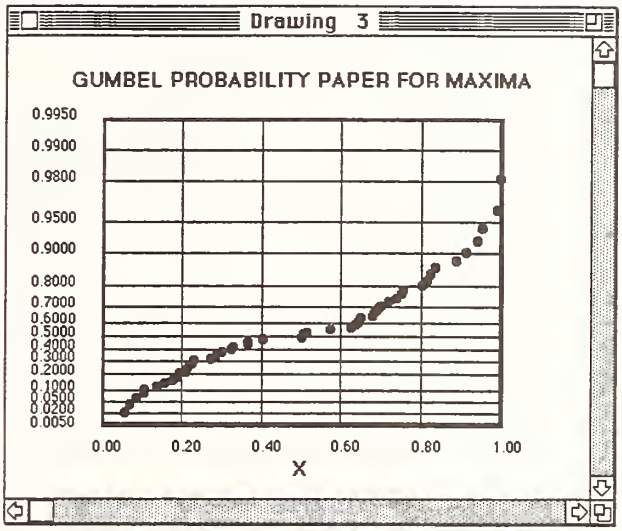


Figure 11: Sample drawn on probability paper.

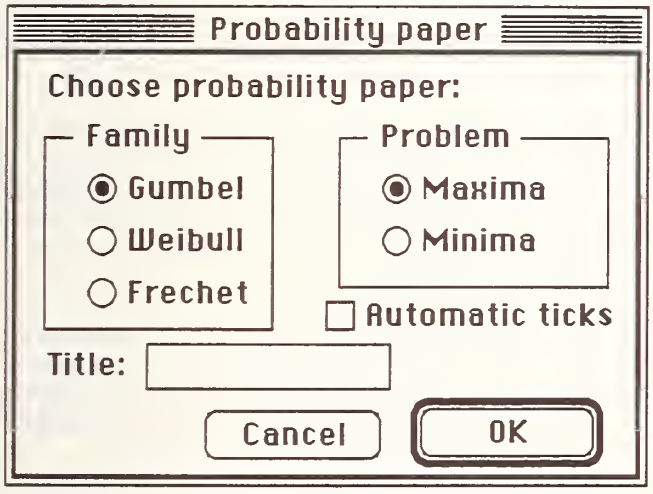


Figure 9: Choosing probability papers.

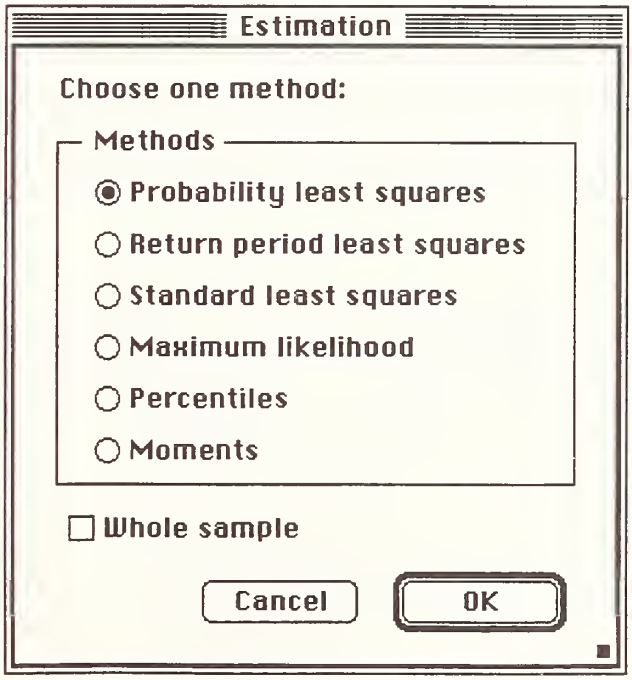


Figure 12: Estimation methods dialog.

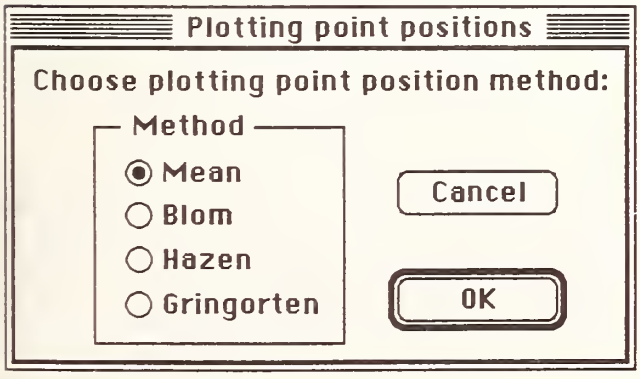


Figure 10: Choosing plotting point position formulas.

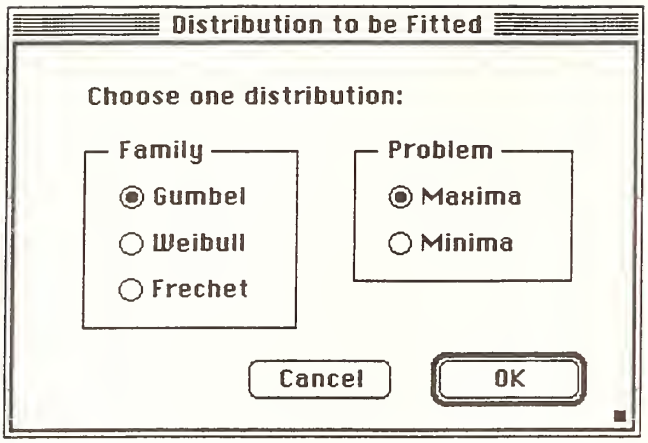


Figure 13: Choosing limit model to be fitted.

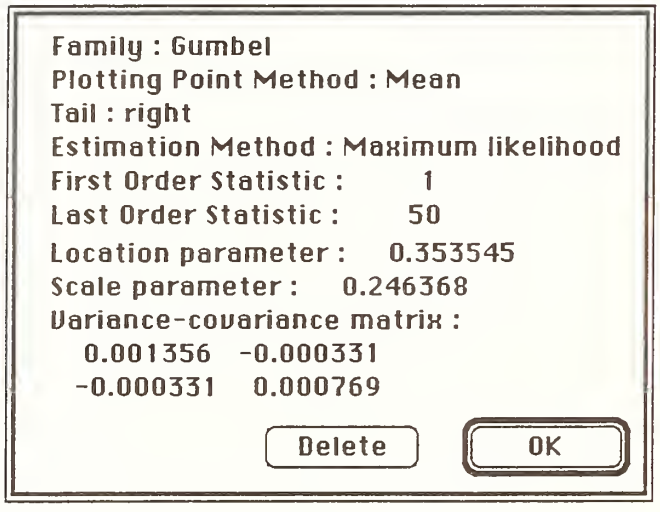


Figure 14: Estimates.

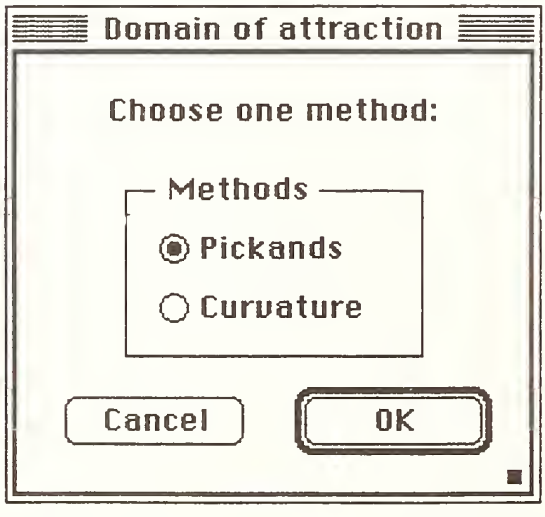


Figure 15: Choosing one method to determine domain of attraction.

Finally, selection of the "*domain of attraction*" option of the main menu leads to the dialog in figure 15 which allows us to choose the Pickands' or the curvature method.

4.3 The expert system

The system can be used by inexperienced users to be guided in all the process of determining design values, for example. The system starts by giving some information to the user who must answer questions to the system. Depending on the answers, the progress is conducted in different directions. As one example, we have included the diagram in figure 16. Initially, the user is asked about whether or not he is going to extrapolate available data. By extrapolation we mean that the required values to predict the random variable are out of the range of the observed values, This can occur either because we are interested in large or small values of the random variable or because we deal with very large or very small associated probabilities. If the user is not dealing with extrapolation, the problem is not an extreme value problem (limit) and standard statistical methods can be used (see figure 16).

Next, we decide about dependence, independence or asymptotic independence. The latter is handled by means of the determination of a critical threshold value above which independence can be assumed. In the first case the system is unable to solve the problem and informs the user.

In the case of independence, we decide about which is the tail of interest (left or right) and the system draws the sample on the corresponding Gumbel probability paper to make a decision about domains of attraction. Depending on the pattern of the drawing, the system recommends one of the classical models or some alternatives.

A whole collection of decision trees, similar to the one above, can be easily incorporated to the system to solve different extreme value problems. With the guide of the system, the possibility of errors is greatly reduced.

4.4 Software implementation

Tne computer aided instruction package was initially implemented on a Macintosh computer using SuperCard. The statistical tools (simulation, estimation, drawing and determination of domains of attraction) were implemented as external commands in Pascal language to allow for a direct access from SuperCard. Later, the convenience of a

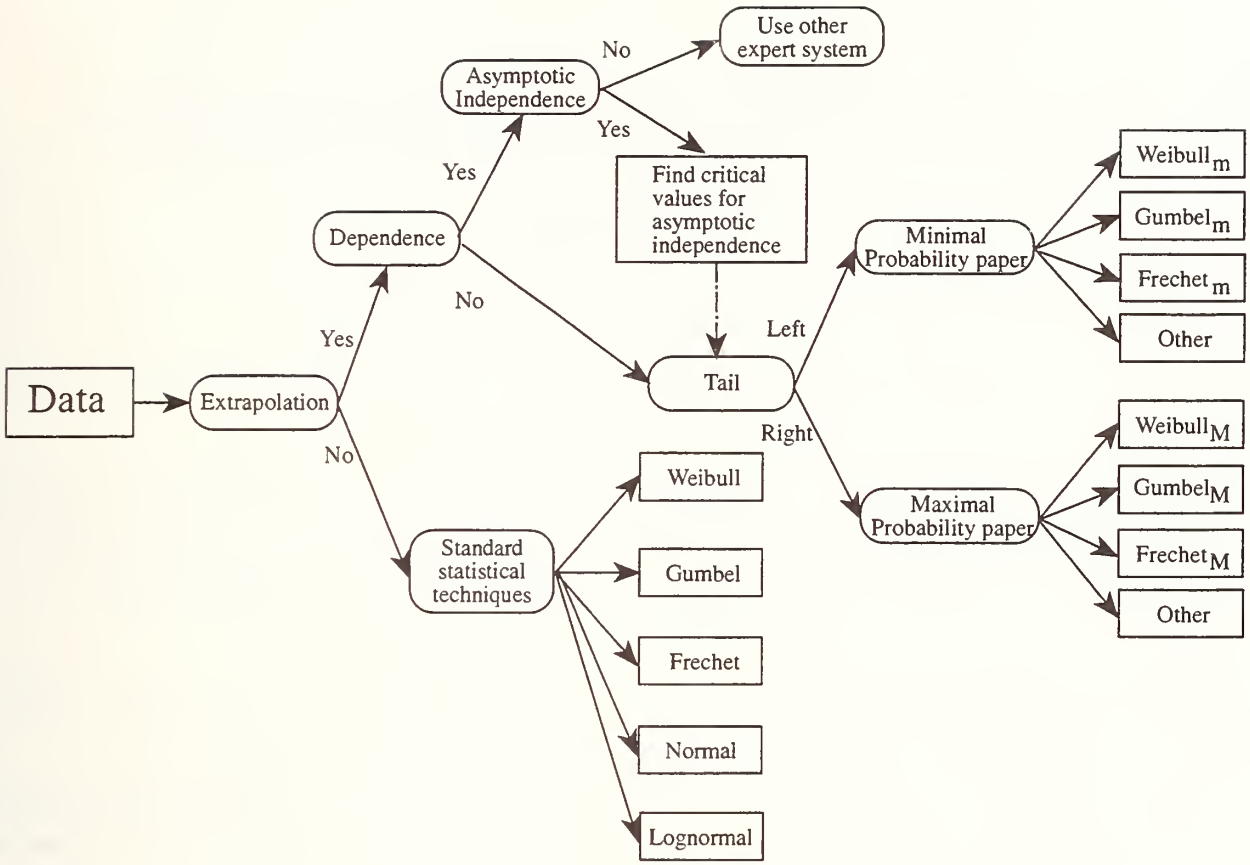


Figure 16: Decision tree.

separate module was recognized and a different program was developed using the Think Pascal object oriented library. Communication between both programs can be solved via Apple events.

5 Conclusions

From all the above we can conclude the following:

- Expert systems can be a good help to applied scientists and engineers to deal with extreme value problems.
- Control of usual errors must be implemented in expert systems to avoid risks and/or waste of money.
- Recent developments of computer aided instruction allows for an easy implementation of teaching modules.
- Use of a set of decision trees leading to a guided determination of design values or other extreme value problems can avoid errors and leads to an important quality improvement.
- Simulation tools can be used to gain experience and expertise on extreme value behaviour of random variables.

6 Acknowledgements

We thank the University of Cantabria and the Dirección General de Investigación Científica y Técnica (DGICYT) (proyect PB91-0302), for partial support of this work.

7 References

[1] Castillo, E. and Alvarez, E. Expert systems. Uncertainty and learning. Computational Mechanics Publications and Elsevier Applied Science. London, New York. (1991). 331 pages.

[2] Weiss, S. M. and Kulikowski, C. A. A practical guide to designing expert systems. Kowman and Allanheld, Publishers. (1984).

Poisson Approximation Of Point Processes Of Exceedances Under von Mises Conditions

Drees, H.
Universität zu Köln, Köln, Germany

Kaufmann, E.
Universität-Gh Siegen, Siegen, Germany

Bounds on the Hellinger distance between certain truncated empirical processes and certain Poisson processes are derived. These bounds depend, roughly speaking, on the rate at which a fairly general von Mises condition holds. Applying these results, also approximations of the joint distribution of the k largest order statistics w.r.t. the variational distance are established.

1 Introduction

Let $X_1, \ldots, X_n$ be i.i.d. random variables (r.v.'s) with common distribution function (d.f.) F. Classical extreme value theory deals with the distributional theory and the asymptotic behaviour of sample maxima $M_n := \max(X_1, \ldots, X_n)$. From Gnedenko [7] one knows that, if the distribution of the standardized maximum $\mathcal{L}(a_n^{-1}(M_n - b_n))$ converges weakly to a nondegenerated distribution for some constants $a_n > 0$ and $b_n \in \mathbb{R}$, then the limiting d.f. must be of the following type (up to a scale and location parameter):

$$G_\beta(x) := \begin{cases} \exp(-x^{-1/\beta}) & x > 0,\ \beta > 0, \\ \exp(-(-x)^{-1/\beta}) & \text{if } x < 0,\ \beta < 0, \\ \exp(-e^{-x}) & \beta = 0. \end{cases}$$

Taking the logarithm and the first derivative, one obtains the functions

$$\Psi_\beta(x) := \log G_\beta(x) \quad \text{if} \quad G_\beta(x) > 0,$$

$$\psi_\beta(x) := \begin{cases} \frac{1}{\beta} x^{-1/\beta - 1} 1_{(0,\infty)}(x) & \beta > 0, \\ \frac{1}{|\beta|}(-x)^{-1/\beta - 1} 1_{(-\infty,0)}(x) & \text{if } \beta < 0, \\ e^{-x} & \beta = 0 \end{cases}$$

that will serve as mean value functions and intensity functions of point processes. Moreover, if $\Psi_\beta(x) \geq -1$, $1 + \Psi_\beta(x)$ defines a generalized Pareto d.f. with shape parameter β and Lebesgue density $\psi_\beta(x)$. The importance of generalized Pareto d.f.'s in extreme value theory was first pointed out by Pickands [15]. These d.f.'s play a central role in the present paper.

More generally, one may deal with the k largest order statistics, where $k \in \{1, \ldots, n\}$, or with those observations that exceed a given threshold. In the second case, one has to deal with point processes. For details about point processes, we refer to Resnick [19] and Reiss [18].

The empirical point process based on the standardized r.v.'s $a_n^{-1}(X_i - b_n)$, $1 \leq i \leq n$, is defined by

$$N_n := \sum_{i=1}^{n} \varepsilon_{a_n^{-1}(X_i - b_n)}$$

where ε_x denotes the Dirac measure with mass 1 at x and 0 elsewhere. Given a threshold $t \in \mathbb{R}$, the empirical point process of exceedances and a truncated Poisson process are defined by

$$N_{n,t} := N_n(\cdot \cap [t, \infty)) \tag{1.1}$$

and

$$N_t^* := N^*(\cdot \cap [t, \infty))$$

where N^* denotes a Poisson point process.

In the following, we will prove results concerning the strong convergence of distributions of truncated empirical processes, that is, convergence w.r.t. the Hellinger distance. Recall that the Hellinger distance H and the variational distance $\| \cdot - \cdot \|$ between two probability measures Q_0 and Q_1 on a σ-field $\mathcal{C}$ are defined by

$$H(Q_0, Q_1) := \left(\int \left(q_0^{1/2} - q_1^{1/2} \right)^2 d\mu \right)^{1/2}$$

and

$$\begin{aligned} \|Q_0 - Q_1\| &:= \sup_{C \in \mathcal{C}} |Q_0(C) - Q_1(C)| \\ &= \frac{1}{2} \int |q_0 - q_1| \, d\mu \end{aligned}$$

where q_i is a μ-density of Q_i, $i = 0, 1$, and μ is a measure dominating Q_0 and Q_1. Note that the variational and Hellinger distances are topologically equivalent, yet the rates of convergence in terms of these distances can be of different order.

The basic tool for dealing with the strong approximation of empirical processes is given by the following two theorems.

Theorem 1.1 *Let $N_{n,D} := N_n(\cdot \cap D)$ be a truncated empirical process and $N^*_{n,D}$ be a Poisson process having the same intensity measure as $N_{n,D}$, $D \in \mathcal{B}$. Then*

(i) $\|\mathcal{L}(N_{n,D}) - \mathcal{L}(N^*_{n,D})\| \le P\{X_1 \in D\}$,

(ii) $H(\mathcal{L}(N_{n,D}), \mathcal{L}(N^*_{n,D})) \le 3^{1/2} P\{X_1 \in D\}$.

For a proof of that result, see Theorem 1.4.2 in [18].

Theorem 1.2 *Let N^*_1, N^*_2 be Poisson processes with finite intensity measures ν^*_1 and ν^*_2, respectively. Then*

(i) $\|\mathcal{L}(N^*_1) - \mathcal{L}(N^*_2)\| \le 3\|\nu^*_1 - \nu^*_2\|$,

(ii) $H(\mathcal{L}(N^*_1), \mathcal{L}(N^*_2)) \le H(\nu^*_1, \nu^*_2)$.

For a proof we refer to [14], Proposition 1.12.1, or [18], Theorem 3.2.2 and Theorem 3.2.1.

It was proved in [19] that weak convergence of the sample maximum to an extreme value r.v. holds if, and only if, weak convergence of the empirical processes N_n to a certain Poisson process N^* is valid. It was shown in [6] that the corresponding result holds w.r.t. the variational distance for the point processes of exceedances $N_{n,t}$ and the Poisson point process N^*_t truncated left of $t > \inf\{x : G_\beta(x) > 0\}$. Moreover, certain bounds for the accuracy of such approximations were established in [17], [6] and [18].

The weak joint behaviour of several intermediate order statistics was studied in [1]. The strong asymptotic normality of single intermediate order statistics under von Mises conditions was proved in [4]. The approximation of intermediate empirical point processes $N_n(\cdot \cap [t_1, t_2])$, truncated about the $(1 - s/n)$-quantile, by a sequence of Poisson processes $N^{*,s}_{t_1,t_2}$, including the homogeneous Poisson process, was investigated in [10]. There the accuracy of approximations was measured by the rate at which a von Mises condition holds (see (2.1)–(2.3)). This idea was also fruitfully utilized in [5] for d.f.'s which are tail equivalent to a generalized Pareto d.f.

It was shown in [11] that multivariate maxima, defined by taking the maxima in each component, converge w.r.t. the variational distance if, and only if, certain truncated multivariate empirical processes converge. In that context, random thresholds are permitted. Approximation rates in the bivariate case were established in [12].

The paper is organized as follows: In Section 2, the density of F will be represented as the product of a generalized Pareto density and a term depending only on a von Mises function.

In Section 3, the truncated empirical process $N_{n,t}$ will be approximated by a Poisson process $N^{*,s}_t$ with mean value function $x \to s\Psi_\beta(x)$, for $x \ge t$. That goal will be achieved in two steps by applying Theorem 1.1 and 1.2: first, we establish a rate of convergence of order s/n for the Hellinger distance between $N_{n,t}$ and the Poisson process $N^*_{n,t}$ having the same intensity measure as $N_{n,t}$. In a second step, $N^*_{n,t}$ will be approximated by a 'limiting' process $N^{*,s}_t$. Our main result in Section 3 will be the following:

Corollary 1.1 *For $t \in \mathbb{R}$, let*

$$t_\beta := \begin{cases} 1 + \beta t & \beta > 0, \\ -(1+\beta t) & \text{if } \beta < 0, \\ t & \beta = 0. \end{cases}$$

Then, under the conditions of Theorem 3.1 and 3.2, we have

$$H(\mathcal{L}(N_{n,t_\beta}), \mathcal{L}(N^{*,s}_{t_\beta})) = O\Big(\frac{s}{n} + s^{1/2}\Delta_{n,\beta,s,t_\beta}\Big)$$

for every fix $t \in \mathbb{R}$ with $0 < G_\beta(t_\beta) < 1$ uniformly for $|\beta| \le C < D^{-1}$.

The term $\Delta_{n,\beta,s,t_\beta}$, that is defined in (3.2), measures the rate at which a von Mises conditon holds. Moreover, the term s serves as a parameter of the standardizing constants; for example, in case $\beta = 0$, the r.v.'s $X_1, \ldots, X_n$ will be centered about the $(1 - s/n)$-quantile. If the threshold t is fix, the expected number of exceedances is proportional to s.

In Section 4, we establish a bound on the variational distance between the joint distribution of the k largest order statistics and a 'limiting distribution' if $k := [s] \ge \log n$.

2 Representation of Density

Subsequently, assume that the underlying d.f. F possesses a density f on $(u(F), \omega(F))$ for some $u(F) < \omega(F) := \sup\{t : F(t) < 1\}$. Denote by $\alpha(F) := \inf\{t : F(t) > 0\}$ the left endpoint of the d.f. F. It is well known (see, e.g. [17]) that F belongs to the strong domain of attraction of an extreme value distribution G_β if one of the following von Mises conditions is satisfied (for $x \uparrow \omega(F)$):

(i) $\omega(F) = \infty$ and

$$\rho_\beta(x) := \frac{f(x)\,x}{1 - F(x)} \longrightarrow \frac{1}{\beta} > 0 \text{ if } \beta > 0, \qquad (2.1)$$

(ii) $\omega(F) < \infty$ and

$$\rho_\beta(x) := \frac{f(x)\,(\omega(F)-x)}{1-F(x)} \longrightarrow -\frac{1}{\beta} > 0 \text{ if } \beta < 0, \tag{2.2}$$

(iii) $\int_{u(F)}^{\omega(F)} 1-F(u)\,du < \infty$ and

$$\rho_\beta(x) := \frac{f(x)\int_x^{\omega(F)} 1-F(u)\,du}{(1-F(x))^2} \longrightarrow 1 \text{ if } \beta = 0. \tag{2.3}$$

Moreover, put $\rho_\beta(x) := \infty$, $x < u(F)$, and $\rho_\beta(x) := 0$, $x > \omega(F)$. The functions ρ_β will be addressed as von Mises terms.

A necessary and sufficient condition for a d.f. F to belong to the weak domain of attraction of G_0 is the existence of an auxiliary function $\tilde{U}$ such that

$$\frac{1-F(x_0+\tilde{U}(x_0)x)}{1-F(x_0)} \to \exp(-x)$$

for every x and $x_0 \uparrow \omega(F)$ (cf. [8]). In this case, we may choose

$$\tilde{U}(x) := U(x) := \frac{\int_x^{\omega(F)} 1-F(t)\,dt}{1-F(x)} \tag{2.4}$$

for $x < \omega(F)$ if $\int_x^{\omega(F)} 1-F(t)\,dt < \infty$. Notice that $U(y)-U(x) = \int_x^y \rho_0(u)-1\,du$ if $u(F) < x, y < \omega(F)$.

Moreover, define

$$\begin{aligned}\tilde{\rho}_{0,x_0}(x) &:= \frac{\rho_0(x_0+U(x_0)x)}{1+\int_0^x \rho_0(x_0+U(x_0)t)-1\,dt} \\ &= \rho_0(x_0+U(x_0)x)\frac{U(x_0)}{U(x_0+U(x_0)x)}\end{aligned}$$

if $u(F) < x_0, x_0+U(x_0)x < \omega(F)$ and $\tilde{\rho}_{0,x_0}(x) := 0$ elsewhere.

The density possesses the following representation.

Lemma 2.1 *(i) Let $\beta = 0$, $x_0 \in (\alpha(F),\omega(F))$, and $x_0+U(x_0)x > u(F)$. Then*

$$\begin{aligned}&U(x_0)f(x_0+U(x_0)x) \\ &= \psi_0(x)\tilde{\rho}_{0,x_0}(x)\exp\Big(-\int_0^x \tilde{\rho}_{0,x_0}(t)-1\,dt\Big) \\ &\quad\times(1-F(x_0)).\end{aligned} \tag{2.5}$$

(ii) Let $\beta > 0$, $x_0 > \alpha(F)$, and $xx_0 > u(F)$. Then

$$\begin{aligned}&x_0 f(x_0 x) \\ &= \psi_\beta(x)(\beta\rho_\beta(x_0 x))\exp\Big(-\int_1^x \frac{\rho_\beta(x_0 t)-1/\beta}{t}\,dt\Big) \\ &\quad\times(1-F(x_0)).\end{aligned} \tag{2.6}$$

(iii) Let $\beta < 0$, $x_0 \in (\alpha(F),\omega(F))$, and $\omega(F)-(\omega(F)-x_0)x > u(F)$. Then

$$\begin{aligned}&(\omega(F)-x_0)f(\omega(F)+(\omega(F)-x_0)x) \\ &= \psi_\beta(x)(-\beta)\rho_\beta(\omega(F)+(\omega(F)-x_0)x) \\ &\quad\times\exp\Big(-\int_{-1}^x \frac{\rho_\beta(\omega(F)+(\omega(F)-x_0)t)+1/\beta}{-t}\,dt\Big) \\ &\quad\times(1-F(x_0)).\end{aligned} \tag{2.7}$$

PROOF. We restrict our attention on (i), because (ii) and (iii) can be dealt with in an analogous way. We have

$$\begin{aligned}&\frac{1-F(x_0+U(x_0)x)}{1-F(x_0)} \\ &= \exp\Big(-\int_{x_0}^{x_0+U(x_0)x} \frac{f(v)}{1-F(v)}\,dv\Big) \\ &= \exp\Big(-\int_{x_0}^{x_0+U(x_0)x} \frac{\rho_0(v)}{U(v)}\,dv\Big) \\ &= \exp\Big(-\int_0^x \frac{\rho_0(x_0+U(x_0)v)U(x_0)}{U(x_0+U(x_0)v)}\,dv\Big) \\ &= \exp(-x)\exp\Big(-\int_0^x \tilde{\rho}_{0,x_0}(v)-1\,dv\Big)\end{aligned} \tag{2.8}$$

if $x_0, x_0+U(x_0)x \in (u(F),\omega(F))$ and hence

$$\begin{aligned}&U(x_0)f(x_0+U(x_0)x) \\ &= U(x_0)\frac{\rho_0(x_0+U(x_0)x)}{U(x_0+U(x_0)x)}(1-F(x_0+U(x_0)x)) \\ &= \exp(-x)\tilde{\rho}_{0,x_0}(x)\exp\Big(-\int_0^x \tilde{\rho}_{0,x_0}(v)-1\,dv\Big) \\ &\quad\times(1-F(x_0)).\end{aligned}$$

□

Lemma 2.1 immediately yields an expansion for f in terms of a generalized Pareto density and a remainder term h, that is, $f(x) = \psi_\beta(x)(1+h(x))$ for sufficiently large x. Moreover, conditions for tail equivalence may be deduced. For example, in case of $\beta = 0$, tail equivalence of the density f to an exponential density holds if $\tilde{\rho}_{0,x_0}(x) \to a \in (0,\infty)$ and $\exp\big(-\int_0^x \tilde{\rho}_{0,x_0}(t)-a\,dt\big)$ converges in $(0,\infty)$ if $x \to \infty$.

In the following section, it turns out that the factorization in (2.5)–(2.7) in a generalized Pareto density and a factor depending only on the von Mises term ρ_β is an useful tool for developing rates of convergence in extreme value theory.

3 Point Processes of Exceedances

In this section, rates of convergence are established for the Hellinger distance between distributions of point

processes of exceedances $N_{n,t}$ and distributions of certain Poisson processes.

First let us point out the trade off between the accuracy of the approximation and the efficiency of statistical inference in the Poisson process model. If the expected number of exceedances increases and the sample size is fix, then the information contained in the limiting model increases, but the accuracy of approximation decreases, and vice versa. For that reason it is of interest to study the accuracy of approximation for several thresholds and standardizing constants such that the expected number of exceedances increases when the sample size n tends to infinity.

Given a d.f. F with density f on $(u(F),\omega(F))$, for some $u(F) < \omega(F)$, define

$$b_n := \begin{cases} 0 & \beta > 0, \\ \omega(F) & \text{if } \beta < 0, \\ F^{-1}(1-s/n) & \beta = 0 \end{cases}$$

and

$$a_n := \begin{cases} F^{-1}(1-s/n) & \beta > 0, \\ \omega(F) - F^{-1}(1-s/n) & \text{if } \beta < 0, \\ U(F^{-1}(1-s/n)) & \beta = 0 \end{cases}$$

where the function U is defined in (2.4) and $s = s(n) \in (0,n)$. Hence the expected number of exceedances $E(N_{n,t}(I\!R)) = n(1-F(b_n + a_n t))$ depends on s and t.

Applying Lemma 2.1, we see that the intensity measure of $N_{n,t}$ possesses the Lebesgue density

$$na_n f(b_n + a_n x) = s\psi_\beta(x) r_{\beta,n}(x) \times \tag{3.1}$$
$$\begin{cases} \beta \exp\Big(-\int_1^x (r_{\beta,n}(u) - 1/\beta)/u\, du\Big), & \beta > 0, \\ |\beta| \exp\Big(-\int_{-1}^x (r_{\beta,n}(u) + 1/\beta)/(-u)\, du\Big), & \beta < 0, \\ \exp\Big(-\int_0^x r_{0,n}(u) - 1\, du\Big), & \beta = 0 \end{cases}$$

for $b_n + a_n x > u(F)$, where $r_{\beta,n}(x) := \rho_\beta(b_n + a_n x)$, $\beta \neq 0$, and $r_{0,n}(x) := \tilde{\rho}_{0,b_n}(x)$.

Denote by

$$\Delta_{n,\beta,s,t} := \tag{3.2}$$
$$\begin{cases} \sup_{x \in [b_n + a_n \min(t,1), \infty)} |\rho_\beta(x) - 1/\beta|, & \beta > 0, \\ \sup_{x \in [b_n + a_n \min(t,-1), \omega(F))} |\rho_\beta(x) + 1/\beta|, & \beta < 0, \\ \sup_{x \in [b_n + a_n \min(t,0), \omega(F))} |\rho_0(x) - 1|, & \beta = 0 \end{cases}$$

the distance of the von Mises term ρ_β and its limit. Observe that $\Delta_{n,\beta,s,t} \to 0$ if $s/n \to 0$ and the von Mises condition is satisfied for β. In the following, the accuracy of the approximations will be measured in terms of $\Delta_{n,\beta,s,t}$. Notice that the term $\Delta_{n,\beta,s,t}$ is related to the term Δ_n defined in [10], yet the normalizing constants are different.

In the sequel, the following notation is used for the approximating processes. Let $N_n^* = N_n^{*,s,F}$ denote the Poisson process having the same intensity measure as N_n, and $N^{*,s} = N^{*,s,\beta}$ the Poisson process with mean value function $x \to s\Psi_\beta(x)$.

In Theorem 3.1, $N_{n,t}$ is approximated by $N_{n,t}^*$ w.r.t. the Hellinger distance, where the error is, up to a constant factor, the expected number of exceedances divided by the sample size n. The conditions in Theorems 3.1 and 3.2 are introduced to keep the terms $C_{i,\beta,t}$, $i = 1,2$, defined in these theorems independent of s, n, and F.

To simplify the notation, let $(\omega(F) - b_n)/a_n := \infty$ if $\omega(F) = \infty$.

Theorem 3.1 *If $\beta \in I\!R$, $t \in (\alpha(G_\beta), \omega(G_\beta))$, $D \in (0,1)$, $\Delta_{n,\beta,s,t} \le D$ and, in addition, $|t| < 2^{-1}\Delta_{n,0,s,t}^{-1/2}$ if $\beta = 0$, then*

$$H(\mathcal{L}(N_{n,t}), \mathcal{L}(N_{n,t}^*)) \le C_{1,\beta,t}\frac{s}{n}$$

where

$$C_{1,\beta,t} := 3^{1/2} \begin{cases} t^{-1/\beta + sign(t-1)D} & \beta > 0, \\ (-t)^{-1/\beta - sign(t+1)D} & \text{if } \beta < 0, \\ e^{-t+5/4} & \beta = 0. \end{cases}$$

Proof. It follows from Theorem 1.1 that

$$H(\mathcal{L}(N_{n,t}), \mathcal{L}(N_{n,t}^*)) \le 3^{1/2}\,(1 - F(b_n + a_n t))\,.$$

First we are going to prove the case $\beta = 0$. Let $x_n := 2^{-1}\Delta_{n,0,s,t}^{-1/2}$. Notice that in the case $\omega(F) < \infty$ we have $b_n + a_n x_n \le \omega(F)$. This follows from

$$U(x) = \frac{\int_x^{\omega(F)} 1 - F(u)\, du}{1 - F(x)} \le \omega(F) - x \to 0$$

as $x \uparrow \omega(F)$ and, thus, $U(b_n) = -\int_{b_n}^{\omega(F)} \rho_0(x) - 1\, dx \le (\omega(F) - b_n)\Delta_{n,0,s,t}$. Hence $(\omega(F) - b_n)/a_n \ge \Delta_{n,0,s,t}^{-1} \ge x_n$. We have for $x \in [t, x_n)$ and some $\vartheta(x) \in [-1,1]$

$$\begin{aligned} &|\tilde{\rho}_{0,b_n}(x) - 1| \\ &= \Big|\frac{\rho_0(b_n + a_n x)}{1 + \int_0^x \rho_0(b_n + a_n u) - 1\, du} - 1\Big| \\ &= \Big|\frac{\rho_0(b_n + a_n x)}{1 + \vartheta(x) x \Delta_{n,0,s,t}} - 1\Big| \\ &\le 2(1 + |x|)\Delta_{n,0,s,t} \le 3D^{1/2} \end{aligned} \tag{3.3}$$

and hence

$$\Big|\int_0^x \tilde{\rho}_{0,b_n}(u) - 1\, du\Big| \le D^{1/2} + 1/4 \le 5/4. \tag{3.4}$$

Together with $s/n = 1 - F(b_n)$ and (2.8) one obtains

$$\begin{aligned} &1 - F(b_n + a_n t) \\ &= \frac{s}{n} e^{-t} \exp\Big(-\int_0^t \tilde{\rho}_{0,b_n}(x) - 1\, dx\Big) \\ &\le \frac{s}{n} e^{-t+5/4}. \end{aligned}$$

In the case $\beta > 0$, use the identity $1 - F(a_n) = s/n$ to show that for appropriate $\vartheta \in [-1, 1]$

$$\begin{aligned}
&1 - F(b_n + a_n t) \\
&= \frac{s}{n} t^{-1/\beta} \exp\Big(- \int_1^t \frac{\rho_\beta(b_n + a_n x) - 1/\beta}{x}\, dx \Big) \\
&= \frac{s}{n} t^{-1/\beta} \exp\Big(- \vartheta \Delta_{n,\beta,s,t} \int_1^t \frac{1}{x}\, dx \Big) \\
&= \frac{s}{n} t^{-1/\beta - \vartheta \Delta_{n,\beta,s,t}} \\
&\le \frac{s}{n} t^{-1/\beta + \mathrm{sign}(t-1) D}.
\end{aligned}$$

The proof of the case $\beta < 0$ can be carried out by similar arguments. □

In a second step, the Poisson process $N^*_{n,t}$ will be replaced by the Poisson process $N^{*,s}_t$.

Theorem 3.2 *If $\beta \in \mathbb{R}$, $t \in (\alpha(G_\beta), \omega(G_\beta))$, $D \in (0,1)$, $\Delta_{n,\beta,s,t} \le D$, $D^{-1} > |\beta|$ and, in addition, $|t| < 2^{-1} \Delta_{n,0,s,t}^{-1/2}$, if $\beta = 0$, then*

$$H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(N^{*,s}_t)) \le C_{2,\beta,t}\, s^{1/2} \Delta_{n,\beta,s,t}$$

where

$$\begin{aligned}
C_{2,\beta,t} :=\ & |\beta|(-\Psi_\beta(t))^{1/2} + (1 + |\beta| D)^{1/2} \times \qquad (3.5) \\
& \begin{cases}
\Big(\int_t^{\omega(G_\beta)} \Big(\frac{\log |x|}{2} \Big)^2 \psi_\beta(x) \\
\quad \times \max(|x|^D, |x|^{-D})\, dx \Big)^{1/2}, & \beta \ne 0, \\
\Big(\int_t^\infty 4(1 + |x|)^2 e^{-x}\, dx + (8/e)^4 \Big)^{1/2} \\
+ \Big(\int_t^\infty (1 + |x|)^2 e^{5/4} (1 + 3D^{1/2}) e^{-x}\, dx \\
\quad + (8/e)^4 e^{5/4} \Big)^{1/2}, & \beta = 0.
\end{cases}
\end{aligned}$$

Proof. Let

$$L_\beta(x) := \begin{cases}
\int_1^x (\rho_\beta(b_n + a_n u) - 1/\beta)/u\, du, & \beta > 0, \\
\int_{-1}^x (\rho_\beta(b_n + a_n u) + 1/\beta)/(-u)\, du, & \beta < 0, \\
\int_0^x \tilde\rho_{0,b_n}(u) - 1\, du, & \beta = 0
\end{cases}$$

if $t \le x < (\omega(F) - b_n)/a_n$.

First we prove the case $\beta = 0$. Using Theorem 1.2, one obtains that the Hellinger distance between Poisson processes is bounded by the Hellinger distance of the corresponding intensity measures. Let $\tilde N$ be a Poisson process whose intensity measure has the density $x \to s \tilde\rho_{0,b_n}(x) e^{-x} 1_{[t,x_n)}(x)$ with $x_n := 2^{-1} \Delta_{n,0,s,t}^{-1/2}$. Applying the triangle inequality and (3.1), one obtains

$$\begin{aligned}
&H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(N^{*,s}_t)) \\
&\le H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(\tilde N)) + H(\mathcal{L}(\tilde N), \mathcal{L}(N^{*,s}_t)) \\
&\le \Big(\int_t^\infty \Big((n a_n f(b_n + a_n x))^{1/2} \\
&\qquad - (s \tilde\rho_{0,b_n}(x) e^{-x} 1_{[t,x_n)}(x))^{1/2} \Big)^2 dx \Big)^{1/2} \\
&\qquad + \Big(\int_t^\infty \Big((s \tilde\rho_{0,b_n}(x) e^{-x} 1_{[t,x_n)}(x))^{1/2} \\
&\qquad - (s e^{-x})^{1/2} \Big)^2 dx \Big)^{1/2} \\
&= \Big(s \int_t^{x_n} \Big(\exp\Big(- L_0(x) \Big)^{1/2} - 1 \Big)^2 \tilde\rho_{0,b_n}(x) e^{-x}\, dx \\
&\qquad + n(1 - F(b_n + a_n x_n)) \Big)^{1/2} \\
&\qquad + \Big(s \int_t^{x_n} (\tilde\rho_{0,b_n}(x)^{1/2} - 1)^2 e^{-x}\, dx + s e^{-x_n} \Big)^{1/2} \\
&=: s^{1/2} I_1^{1/2} + s^{1/2} I_2^{1/2}, \text{ say.}
\end{aligned}$$

Taking into account that $(x^{1/2} - 1)^2 \le (x - 1)^2$ and $e^{-x} \le (ex/4)^{-4}$, $x \ge 0$, we obtain from (3.3)

$$\begin{aligned}
I_2 &\le \int_t^{x_n} (\tilde\rho_{0,b_n}(x) - 1)^2 e^{-x}\, dx + e^{-x_n} \\
&= \int_t^{x_n} (2(1 + |x|) \Delta_{n,0,s,t})^2 e^{-x}\, dx + (e x_n/4)^{-4} \\
&\le \Delta_{n,0,s,t}^2 \Big(\int_t^\infty 4(1 + |x|)^2 e^{-x}\, dx + (8/e)^4 \Big) \qquad (3.6)
\end{aligned}$$

Combining (2.8) and (3.4) leads to

$$\begin{aligned}
&n(1 - F(b_n + a_n x_n)) \\
&\quad = s e^{-x_n} \exp\Big(- \int_0^{x_n} \tilde\rho_{0,b_n}(u) - 1\, du \Big) \\
&\quad \le s \Delta_{n,0,s,t}^2 (8/e)^4 e^{5/4}. \qquad (3.7)
\end{aligned}$$

A Taylor expansion yields

$$\begin{aligned}
&I_1 - (n/s)(1 - F(b_n + a_n x_n)) \\
&= \int_t^{x_n} \Big(\exp(-L_0(x)/2) - 1 \Big)^2 \tilde\rho_{0,b_n}(x) e^{-x}\, dx \\
&= \int_t^{x_n} \Big(- L_0(x)/2 \Big)^2 \Big(\exp\Big(- \vartheta(x) L_0(x)/2 \Big) \Big)^2 \\
&\qquad \times \tilde\rho_{0,b_n}(x) e^{-x}\, dx \\
&\le \Delta_{n,0,s,t}^2 \int_t^{x_n} (1 + |x|)^2 e^{5/4} (1 + 3D^{1/2}) e^{-x}\, dx \\
&\qquad (3.8)
\end{aligned}$$

for some $\vartheta(x) \in [0, 1]$ where the last inequality follows from (3.3) and (3.4). Combining (3.6), (3.7), and (3.8), we get $H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(N^{*,s}_t)) \le C_{2,0,t} s^{1/2} \Delta_{n,0,s,t}$ with $C_{2,0,t}$ defined in (3.5).

Next, we turn to the case $\beta \ne 0$. Check that

$$|L_\beta(x)| \le \Delta_{n,\beta,s,t} |\log |x|| \qquad (3.9)$$

holds for $\beta \ne 0$ if $t \le x < \omega(G_\beta)$.

Let $\tilde{N}$ be a Poisson process whose intensity measure has the density $x \to s\rho_\beta(b_n + a_n x)|\beta| \times \psi_\beta(x) 1_{[t,\omega(G_\beta))}(x)$. Arguing as in the case $\beta = 0$, one obtains

$$\begin{aligned}
&H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(N^{*,s}_t)) \\
&\le H(\mathcal{L}(N^*_{n,t}), \mathcal{L}(\tilde{N})) + H(\mathcal{L}(\tilde{N}), \mathcal{L}(N^{*,s}_t)) \\
&\le \Big(\int_t^{\omega(G_\beta)} \Big((na_n f(b_n + a_n x))^{1/2} \\
&\qquad - (s\rho_\beta(b_n + a_n x)|\beta|\psi_\beta(x))^{1/2} \Big)^2 dx \Big)^{1/2} \\
&\quad + \Big(\int_t^{\omega(G_\beta)} \Big((s\rho_\beta(b_n + a_n x)|\beta|\psi_\beta(x))^{1/2} \\
&\qquad - (s\psi_\beta(x))^{1/2} \Big)^2 dx \Big)^{1/2} \\
&= \Big(s \int_t^{\omega(G_\beta)} \Big(\exp\Big(-L_\beta(x) \Big)^{1/2} - 1 \Big)^2 \\
&\qquad \times \rho_\beta(b_n + a_n x)|\beta|\psi_\beta(x)\, dx \Big)^{1/2} \\
&\quad + \Big(s \int_t^{\omega(G_\beta)} ((|\beta|\rho_\beta(b_n + a_n x))^{1/2} - 1)^2 \\
&\qquad \times \psi_\beta(x)\, dx \Big)^{1/2} \\
&= s^{1/2} I_{1,\beta}^{1/2} + s^{1/2} I_{2,\beta}^{1/2}, \text{ say.}
\end{aligned}$$

Moreover,

$$\begin{aligned}
I_{2,\beta} &\le \int_t^{\omega(G_\beta)} (|\beta|\rho_\beta(b_n + a_n x) - 1)^2 \psi_\beta(x)\, dx \\
&\le \beta^2 \Delta^2_{n,\beta,s,t} \int_t^{\omega(G_\beta)} \psi_\beta(x)\, dx \\
&= \beta^2 \Delta^2_{n,\beta,s,t}(-\Psi_\beta(t)).
\end{aligned}$$

Applying a Taylor expansion and (3.9), we obtain

$$\begin{aligned}
I_{1,\beta} &= \int_t^{\omega(G_\beta)} \Big(\exp\Big(-L_\beta(x)/2 \Big) - 1 \Big)^2 \\
&\qquad \times \rho_\beta(b_n + a_n x)|\beta|\psi_\beta(x)\, dx \\
&= \int_t^{\omega(G_\beta)} \Big(-L_\beta(x)/2 \Big)^2 \\
&\qquad \times \Big(\exp\Big(-\vartheta(x) L_\beta(x)/2 \Big) \Big)^2 \\
&\qquad \times \rho_\beta(b_n + a_n x)|\beta|\psi_\beta(x)\, dx \\
&\le \Delta^2_{n,\beta,s,t} \int_t^{\omega(G_\beta)} \Big(-(\log|x|)/2 \Big)^2 \\
&\qquad \times \exp\Big(\Delta_{n,\beta,s,t} |\log|x|| \Big) \\
&\qquad \times \Big(\frac{1}{|\beta|} + D \Big) |\beta|\psi_\beta(x)\, dx \\
&\le \Delta^2_{n,\beta,s,t}(1 + |\beta| D) \\
&\quad \times \int_t^{\omega(G_\beta)} \Big(\frac{\log|x|}{2} \Big)^2 \psi_\beta(x) \max(|x|^D, |x|^{-D})\, dx
\end{aligned}$$

for some $\vartheta(x) \in [0,1]$. □

The rates established in Theorem 3.1 and 3.2 are sharp if $s \ge 5 \log n$. This follows from results in Section 4, where it is shown that for the k largest order statistics ($k = [s/5]$) the same rates of approximation hold as for the point processes and that the rates obtained for the k largest order statistics are sharp. In the case $\beta = 0$, similar bounds to that in the Theorems 3.1 and 3.2 are established in [2] for different normalizing constants.

Proof of Corollary 1.1. We have to show that there exists a constant $C > 0$ such that $C_{i,\beta,t} \le C$, $i = 1, 2$. But this is immediate from the well-known formula $(1+\beta x)^{-1/\beta} \to e^{-x}$ as $\beta \to 0$ and some straightforward calculations. □

Using Theorems 3.1 and 3.2 (respectively Corollary 1.1), one may establish rates of convergence for point processes of exceedances if the d.f. F fulfills one of the von Mises conditions (2.1)–(2.3). The proofs of the following examples may be carried out by elementary calculations (cf. [10]).

Example 3.1 If $F(x) := 1 + \Psi_\beta(x)$, $\Psi_\beta(x) \ge -1$, denotes a generalized Pareto d.f. for some $\beta \in \mathbb{R}$, then

$$H(\mathcal{L}(N_{n,t}), \mathcal{L}(N^{*,s}_t)) \le 3^{1/2}(-\Psi_\beta(t))\frac{s}{n}$$

for $t \in (\alpha(G_\beta), \omega(G_\beta))$.

Since for the generalized Pareto distribution the approximating Poisson process in Theorem 3.1 equals the process $N^{*,s}_t$, the bound in Example 3.1 is, up to a constant factor, equal to the expected number of exceedances $EN_t(\mathbb{R})$ divided by n.

The following example was first proved in [6], Theorem 4.

Example 3.2 (δ-condition)
Assume that for some $\beta \in \mathbb{R}$, $\delta > 0$, and $L > 0$ the density f has the form

$$f(x) = \psi_\beta(x) e^{h(x)} \tag{3.10}$$

where $|h(x)| \le L(-\Psi_\beta(x))^\delta$. Then

$$H(\mathcal{L}(N_{n,t}), \mathcal{L}(N^{*,s}_t)) = O\Big(\frac{s}{n} + s^{1/2} \Big(\frac{s}{n} \Big)^\delta \Big) \tag{3.11}$$

for every fix $t \in (\alpha(G_\beta), \omega(G_\beta))$.

Example 3.3 (Normal distribution)
Let F be the normal d.f. Then

$$H(\mathcal{L}(N_{n,t}), \mathcal{L}(N^{*,s}_t)) = O\Big(\frac{s^{1/2}}{\log(n/s)} \Big)$$

for every fix $t \in \mathbb{R}$.

4 Joint Distribution of Order Statistics

In this section, we study an application of the preceding results on point process approximations to derive rates of convergence for the joint distribution of the k largest order statistics. Expansions of the d.f. of sample maxima under von Mises conditions may be found in Radtke [16]. Uniform convergence of order statistics under von Mises conditions was studied by Falk [3]. Sweeting [20] has shown that the von Mises conditions (2.1)–(2.3) are equivalent to the uniform convergence of densities of maxima on finite intervalls. For a comprehensive treatment of order statistics and, in particular, approximations of intermediate and extreme order statistics w.r.t. the Hellinger distance, we refer to Reiss [17]. There a method was introduced for establishing rates of convergence of point processes from rates for the k largest order statistics. We use that method in the converse direction.

In the following, we define for a point measure μ on $\mathbb{R}$ and $k \in \{1, \dots, n\}$ the terms

$$m_{k,t}(\mu) := \inf\{x > t : \mu((x, \infty)) < k\}$$

and

$$M_{k,t}(\mu) := (m_{1,t}(\mu), \dots, m_{k,t}(\mu)).$$

Notice that, for the order statistics $X_{1:n} \le \cdots \le X_{n:n}$ of $X_1, \dots, X_n$ and for the empirical process $N_{n,t}$ defined in (1.1), the identity

$$\Big(a_n^{-1}(X_{n-i+1:n} - b_n)\Big)_{i=1}^{k} = M_{k,t}(N_{n,t})$$

holds if $a_n^{-1}(X_{n-k+1:n} - b_n) \ge t$.

To obtain sharp bounds from Theorems 3.1 and 3.2, one has to find conditions such that $P\{a_n^{-1}(X_{n-k+1:n} - b_n) < t\}$ is of the same order as the error of the point process approximation. For that purpose, we use an exponential bound for single order statistics, which is a consequence of [17], Lemma 3.1.1. For a proof of Lemma 4.1 see [9].

Lemma 4.1 *Let $U_1, \dots, U_n$ be independent and uniformly distributed on $(0,1)$, $k \in \{1, \dots, n\}$ and $C > 1$. Then*

$$P\{U_{n-k+1:n} < 1 - Ck/n\} \le \exp\Big(-k\frac{(C-1)^2 n^2}{3C(n+1)^2}\Big).$$

Notice that in Lemma 4.1 a bound of order $O(n^{-1})$ may be achieved by choosing $k \ge \log n$ and $C \ge 5$.

Denote by $Q_{k,\beta} := \mathcal{L}(M_{k,-\infty}(N^{*,s}))$ the joint distribution of the k largest order statistics of the Poisson process $N^{*,s}$ with mean value function $x \to s\Psi_\beta(x)$, $x > \alpha(G_\beta)$. Our main result in this section is the following:

Theorem 4.1 *Let $k \in \{1, \dots, n\}$, $k \ge \log n$, $C > 1$, $s = Ck$, and $t_\beta := sign(\beta)$. Then, under the conditions of Theorem 3.1 and 3.2, we have*

$$\begin{aligned} &\|\mathcal{L}(a_n^{-1}(X_{n-i+1:n} - b_n))_{i=1}^{k} - Q_{k,\beta}\| \\ &\le 2\Big(n^{-\frac{(C-1)^2 n^2}{3C(n+1)^2}} + 3^{1/2} e^{5/4}\frac{Ck}{n} \\ &\qquad + C_{2,\beta,t_\beta} C^{1/2} k^{1/2} \Delta_{n,\beta,Ck,t_\beta}\Big) \end{aligned} \tag{4.1}$$

where C_{2,β,t_β} is defined in (3.5).

Proof. Recall that F possesses a density on $(u(F), \omega(F))$. The condition $\Delta_{n,\beta,Ck,t_\beta} \le D$ implies $u(F) < a_n t_\beta + b_n = F^{-1}(1 - s/n)$. Hence F is continuous at $F^{-1}(1 - s/n)$. Denote by $U_{n-k+1:n}$ the kth largest order statistic of n independent r.v.'s which are uniformly distributed on $(0,1)$. Since $\mathcal{L}(X_{n-k+1:n}) = \mathcal{L}(F^{-1}(U_{n-k+1:n}))$, we may write

$$\begin{aligned} &P\{a_n^{-1}(X_{n-k+1:n} - b_n) \le t_\beta\} \\ &= P\{X_{n-k+1:n} \le F^{-1}(1 - s/n)\} \\ &= P\{F^{-1}(U_{n-k+1:n}) \le F^{-1}(1 - s/n)\} \\ &= P\{U_{n-k+1:n} \le 1 - s/n\}. \end{aligned}$$

Using Lemma 4.1, we get

$$P\{a_n^{-1}(X_{n-k+1:n} - b_n) \le t_\beta\} \le n^{-\frac{(C-1)^2 n^2}{3C(n+1)^2}}. \tag{4.2}$$

Recall that the Hellinger distance dominates the variational distance. The triangle inequality and the monotonicity theorem (see, e.g. [13] or [18], Lemma 1.4.2) yield

$$\begin{aligned} &\|\mathcal{L}(a_n^{-1}(X_{n-i+1:n} - b_n))_{i=1}^{k} - Q_{k,\beta}\| \\ &\le \|\mathcal{L}(a_n^{-1}(X_{n-i+1:n} - b_n))_{i=1}^{k} - \mathcal{L}(M_{k,t_\beta}(N_{n,t_\beta}))\| \\ &\quad + \|\mathcal{L}(M_{k,t_\beta}(N_{n,t_\beta})) - \mathcal{L}(M_{k,t_\beta}(N_{t_\beta}^{*,s}))\| \\ &\quad + \|\mathcal{L}(M_{k,t_\beta}(N_{t_\beta}^{*,s})) - Q_{k,\beta}\| \\ &\le P\{a_n^{-1}(X_{n-k+1:n} - b_n) \le t_\beta\} \\ &\quad + \|\mathcal{L}(N_{n,t_\beta}) - \mathcal{L}(N_{t_\beta}^{*,s})\| + P\{N_{t_\beta}^{*,s}(\mathbb{R}) < k\}. \end{aligned}$$

Since $P\{N_{t_\beta}^{*,s}(\mathbb{R}) < k\} \le P\{N_{n,t_\beta}(\mathbb{R}) < k\} + \|\mathcal{L}(N_{n,t_\beta}) - \mathcal{L}(N_{t_\beta}^{*,s})\|$, the assertion follows from (4.2) and Theorem 3.1 and 3.2. □

Notice that $m_{k,-\infty}(N^{*,s})$ possesses the d.f.

$$\begin{aligned} &P\{m_{k,-\infty}(N^{*,s}) \le x\} \\ &= P\{N^{*,s}((x,\infty)) \le k-1\} \\ &= G_\beta^s(x) \sum_{i=0}^{k-1} \frac{(-\log G_\beta^s(x))^i}{i!} \end{aligned}$$

for $x > \alpha(G_\beta)$ (cf. [18], E.6.2).

If $k \geq \log n$, the rate obtained in (4.1) is sharp for distributions treated in Examples 3.1 and 3.2 with $\delta = 1$. That follows from [17], Theorem 5.4.4 and Example 5.5.6. Sharp rates for every $k \in \{1, \ldots, n\}$ may be derived by direct calculations using Theorem 5.5.4 in [17].

In the case of normal r.v.'s Theorem 4.1 yields more accurate bounds than those known in literature if $k \geq \log n$ (cf. [3], Example 4.53).

Example 4.1 *Let $X_1, \ldots, X_n$ be independent, normally distributed r.v.'s and $k \geq \log n$. Then*

$$\|\mathcal{L}(a_n^{-1}(X_{n-i+1:n} - b_n))_{i=1}^k - Q_{k,0}\| = O\Big(\frac{k^{1/2}}{\log(n)}\Big)$$

where a_n and b_n are chosen as in Section 3 with $s := 5k$.

PROOF. The proof follows from Theorem 4.1 and Example 3.3. □

References

[1] Cooil, B., Limiting multivariate distributions of intermediate order statistics. Ann. Probab., 13 (1984), 469–477.

[2] Drees, H., *Refined estimation of the extreme value index.* Ph.D. thesis, University of Siegen, 1993.

[3] Falk, M., *Uniform convergence of extreme order statistics.* Habilitatonsschrift, University of Siegen, 1985.

[4] Falk, M., A note on uniform asymptotic normality of intermediate order statistics. Ann. Inst. Statist. Math., 41 (1990), 19–29.

[5] Falk, M. and Marohn, F., Von Mises conditions revisited. Ann. Probab., 29 (1993), 1310–1328.

[6] Falk, M. and Reiss, R.-D., Poisson approximation of empirical processes. Statist. Probab. Letters, 14 (1992), 39–48.

[7] Gnedenko, B., Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math., 44 (1943), 423–453.

[8] Haan, L. de, *On Regular Variation and its Application to the Weak Convergence of Sample Extremes.* Math. Centre Tracts 32, Amsterdam, 1975.

[9] Kaufmann, E., *Contributions to Approximations in Extreme Value Theory.* Ph.D. thesis, University of Siegen, 1992.

[10] Kaufmann, E. and Reiss, R.-D., Poisson approximation of intermediate empirical processes. J. Appl. Prob., 29 (1992), 825–837.

[11] Kaufmann, E. and Reiss, R.-D., Strong convergence of multivariate point processes of exceedances. Ann. Inst. Statist. Math., 45 (1993), 433–444.

[12] Kaufmann, E. and Reiss, R.-D., Approximation rates for exceedances processes. To appear in: J. Statist. Plann. Inference (1994).

[13] Liese, F. and Vajda, I., *Convex Statistical Distances.* Teubner-Texte zur Mathematik, Bd. 95. Teubner, Leipzig, 1987.

[14] Matthes, K., Kerstan, J. and Mecke, J., *Infinitely Divisible Point Processes.* Wiley, Chichester, 1978.

[15] Pickands, J. III, Statistical inference using extreme order statistics. Ann. Statist., 3 (1975), 119–131.

[16] Radtke, M., *Konvergenzraten und Entwicklungen unter von Mises Bedingungen in der Extremwerttheorie.* Ph.D. thesis, University of Siegen, 1988.

[17] Reiss, R.-D., *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics.* Springer, New York, 1989.

[18] Reiss, R.-D., *A Course on Point Processes.* Springer, New York, 1993.

[19] Resnick, S.I., *Extreme Values, Regular Variation, and Point Processes.* Springer, New York, 1987.

[20] Sweeting, T.J., On domains of uniform local attraction in extreme value theory. Ann. Probab., 13 (1985), 196–205.

Estimating The Extremal Index Under A Local Dependence Condition By The Reciprocal Of The Average Length Of Successive Runs

Duarte, L.C.C.
University of Lisbon, Lisbon, Portugal

Abstract: Whenever a strictly stationary, strongly mixing sequence, satisfies the local dependence condition D"(u_n) of Ref. [1], the point process $\tilde{N}_n$ defined by the upcrossings of a high level u_n has an important contribution to the characterization of the limiting compound Poisson process of exceedances.

For each n let $N(n,u_n) = \sum_{i=1}^{n} \mathbb{1}(X_i>u_n)$ be the random variable that represents the number of exceedances of u_n in a sample of size n, and $\tilde{N}(n,u_n)=\sum_{i=1}^{n} \mathbb{1}(X_{i-1}\le u_n<X_i)$ the number of upcrossings of the same level. Then, if we consider levels $u_n(\tau)$ such that $nP[X_1>u_n(\tau)]\to\tau$, when $n\to\infty$, the extremal index θ verifies

$$\lim_{n\to\infty} \frac{nP[X_1\ge u_n(\tau)]}{nP[X_0\le u_n(\tau)<X_1]} = \lim_{n\to\infty} \frac{E[N(n,u_n(\tau))]}{E[\tilde{N}(n,u_n(\tau))]} = \frac{1}{\theta}\ .$$

Now, if the sequence of levels is such that $E[\tilde{N}(n,u_n)]\to 1$, when $n\to\infty$, then the reciprocal of the extremal index will merely be the limit of the mean number of exceedances. Based on this result we have developed a method of estimation of θ that consists on dividing the sample in k_n blocks of size r_n and taking the average number of exceedances of the level u_{ni}, suitably defined for the i-th block, so that the mean number of upcrossings of this level in that block is approximately 1. More precisely, we present here some properties of the estimator

$$\tilde{\tilde{\theta}}_n = \left(\frac{\sum_{i=1}^{k_n} N_i(r_n,u_{ni})}{k_n} \right)^{-1}\ .$$

1. Introduction

In this paper we study an estimator for the extremal index which has been motivated by Ref. [1] and for this reason we will assume for the underlying stationary process, conditions analogous to those used by these authors.

In order to make clear the origin of this estimator we start with the presentation of some known results, related to this subject.

Given a stationary sequence $\{X_i\}$, $i\ge 1$, and denoting by $\mathcal{F}_i^j(u)$ the σ-field generated by the events $\{(X_k\le u): i\le k\le j\}$, the following mixing coefficients will be used

$$\alpha_n(\ell,u_n) = \sup\left\{ \left| P(A\cap B) - P(A)P(B) \right| : A\in\mathcal{F}_1^k(u_n),\ B\in\mathcal{F}_{k+\ell}^n(u_n) \right\}$$

These are related to the well known long range dependence condition $\Delta(u_n)$. More precisely, $\{X_i\}$ is said to satisfy $\Delta(u_n)$ if, for some $\ell_n=o(n)$, $\alpha_n(\ell_n,u_n)\to 0$ as $n\to\infty$. This condition is stronger than $D(u_n)$ defined in Ref. [2], but it will be needed to validate the convergence of some point processes important in this work.

Concerning the local dependence structure of $\{X_i\}$ we assume henceforth the validity of the condition D" defined in Ref. [1], which, in some way, restricts the local occurrence of two or more upcrossings of high levels.

Let k_n be a sequence of integers, with $k_n \to \infty$ and such that $k_n \alpha_n(\ell_n, u_n) \to 0$, $k_n \ell_n n \to 0$ and $k_n P(X_1 > u_n) \to 0$. We say that $\{X_i\}$ verifies $D''(u_n)$ if for $r_n = \left[\frac{n}{k_n}\right]$ we have

$$\lim_{n\to\infty} n \sum_{j=2}^{r_n - 1} P[X_1 > u_n, X_j \le u_n < X_{j+1}] = 0.$$

Note that we say that $\{X_i\}$ has an upcrossing of the level u at j if $X_{j-1} \le u < X_j$. So, if we represent by $\mu(u)$ the probability of such an event (which is independent of j by stationarity) then $\mu(u)$ can be interpreted as the mean number of upcrossings of u per unit time. Considering that for a stationary sequence

$$\mu(u) = P(X_1 \le u < X_2) = P(X_2 \le u \mid X_1 > u)\, P(X_1 > u)$$

we can conclude that for sequences with the same marginal distribution, the lesser the value of μ for a fixed level u, the stronger the tail dependence structure near that level.

Notice that the extremal index of $\{X_i\}$ is equal to θ if for each sequence of levels $u_n(\tau)$ such that $nP[X_1 > u_n(\tau)] \to \tau$ the limit of $P[M_n \le u_n(\tau)]$ is $e^{-\theta\tau}$, (here M_n denotes the random variable $\max_{i=1}^{n}\{X_i\}$). This parameter θ takes values in the interval [0,1] and measures the strength of dependence of a stationary sequence. The stronger the dependence the lesser the value of θ so that for an i.i.d. sequence we have $\theta=1$ whereas the value $\theta=0$ corresponds to a long memory sequence. In our study we admit a weak dependent structure for the sequence in such a way that its extremal index, when exists, is strictly positive.

The connection between the last two paragraphs is quite clear and it can be proved that for a suitably chosen sequence of levels u_n, for which both conditions Δ and D" hold, the extremal index can be obtained as the limit of $P(X_2 \le u_n \mid X_1 > u_n)$ when n goes to infinity. This result comes straightforward from the following proposition established in Ref. [1]:

Proposition 1. Suppose that $\Delta(u_n)$ and $D''(u_n)$ hold for some sequences $\{u_n\}$, $\{k_n\}$ and $\{r_n\}$ satisfying the conditions mentioned above. Then

$$P[M_n \le u_n] \to e^{-\nu} \quad \text{if and only if} \quad n\mu(u_n) \to \nu .$$

Hence, when the process has an extemal index θ, we have

$$n\mu(u_n) \to \nu \quad \text{if and only if} \quad n\,P[X_1 > u_n] \to \nu/\theta \qquad (1)$$

In the sequel we will also be interested in two point processes relevant to the development of an estimator of the extremal index. The first one is the (time normalized) point process N_n of exceedances of a high level u_n:

$$N_n(B) = \sum_{i=1}^{n} \varepsilon_{i/n}(B)\, \mathbb{1}(X_i > u_n), \quad B \subset [0,1] .$$

This point process has been fully studied in Ref. [3], and it is shown, for example, that under $\Delta(u_n)$ any existing limit of N_n must be a compound Poisson point process. When $u_n = u_n(\tau)$ is such that $nP[X_1 > u_n(\tau)] \to \tau$, it is proved, under general conditions, that the underlying Poisson point process has intensity $\theta\tau$ and the mean value of the multiplicities is θ^{-1}. From the proof we can infere that the Poisson points can be regarded as positions of clusters of exceedances and the number of exceedances in each cluster corresponds to the multiplicities.

Let $\pi_n(j)$ be the distribution of the number of exceedances in a cluster given that there is at least one exceedance

$$\pi_n(j) = P\Big\{ \sum_{i=1}^{r_n} \mathbb{1}(X_i > u_n) = j \;\Big|\; \sum_{i=1}^{r_n} \mathbb{1}(X_i > u_n) > 0 \Big\}, \quad j = 1, 2, \ldots$$

The result mentioned above can be more precisely stated as follows:

Proposition 2. Assume that $\Delta(u_n)$ holds for the stationary sequence $\{X_i\}$ and that, for some $\nu > 0$,

$$\lim_{n\to\infty} P[M_n \le u_n] = e^{-\nu}.$$

Suppose there exists a probability distribution $\pi(j)$ such that $\pi(j) = \lim_{n\to\infty} \pi_n(j)$, $j = 1, 2, \ldots$, where $\pi_n(j)$ is the conditional probability distribution of the number of exceedances of u_n defined for blocks of size $r_n = [n/k_n]$, with k_n going to infinity in the conditions stated above. Then N_n converges in distribution to a compound Poisson process with intensity ν and multiplicity distribution $\pi(.)$.

As we have seen, for a stationary sequence with extremal index θ it is possible to consider normalized levels $u_n(\tau)$ in such a way that we have $\lim_{n\to\infty} P[M_n \le u_n] = e^{-\theta\tau}$. For this kind of levels the

mean number of exceedances in a sample of size n is assimptotically τ, since

$$E\{N_n([0,1])\} = \sum_{i=1}^{n} \varepsilon_{i/n}([0,1])\, E\{\mathbb{1}(X_i > u_n(\tau))\}$$
$$= n\, P[X_1 > u_n(\tau)],$$

and in turn, if $\Delta(u_n(\tau))$ holds, it is proved in Ref. [4] that the mean cluster size is approximately θ^{-1}:

$$\lim_{n\to\infty} \sum_{j\geq 1} j\, \pi_n(j) = \theta^{-1}.$$

Now, the size of a cluster induced by $u_n(\tau)$ is just the number of exceedances of that level in a block with size r_n, conveniently chosen. A natural estimator of θ^{-1} would then be obtained by dividing the total number of exceedances of a previously defined high level by the total number of clusters. This procedure is quite general and has been proposed in Ref. [5] . It has the (not so minor) problems of choosing a convenient block size and a convenient high level. As we will see, the first problem can be avoided if we assume a local dependence condition like D". This is due to the important role played by the point process of upcrossings in the characterization of the limiting compound Poisson process of exceedances. According to Ref. [1] the point process of upcrossings

$$\tilde{N}_n(B) = \sum_{i=1}^{n} \varepsilon_{i/n}(B)\, \mathbb{1}(X_{i-1} \leq u_n < X_i)), \quad B \subset [0,1],$$

converges to a Poisson process whose intensity depends on the chosen sequence of normalized levels and on the value of the extremal index.

Proposition 3: Suppose $\Delta(u_n)$ and $D''(u_n)$ hold for a sequence of levels u_n such that the mean number of uppcrossings is approximately ν, i.e., $n\mu(u_n) \to \nu$.
Then $\tilde{N}_n \to \tilde{N}$, where $\tilde{N}$ is a Poisson process in [0,1], with intensity ν.
If, in addition, $\{X_i\}$ has an extremal index θ, (1) holds and, consequently, $\theta\tau$ is the intensity of the limiting Poisson process $\tilde{N}$, generated by the levels $u_n(\tau)$ such that $nP[X_1 > u_n(\tau)] \to \tau$.

Under the conditions mentioned above the equivalence (1) gives us two ways of constructing normalized levels. In the following we denote by $\tilde{u}_n(\delta)$ levels such that $n\mu(\tilde{u}_n(\delta)) \to \delta$, and by $u_n(\delta)$ levels such that $nP[X_1 > u_n(\delta)] \to \delta$ so, in case of existence of the extremal index, $u_n(\delta) = \tilde{u}_n(\theta\delta)$).
Another consequence of assuming D" is that it enables us to identify clusters of exceedances with runs of consecutive exceedances.

Given a sequence of levels u_n, define for each n the random variables $Y_j = Y(j, u_n)$, $1 \leq j \leq n$, which represent the number of consecutive exceedances of the level u after time j,

$$\{Y_j = 0\} = \{X_j \leq u_n\}$$
$$\{Y_j = k\} = \{X_j > u_n, X_{j+1} > u_n, \ldots, X_{j+k-1} > u_n, X_{j+k} \leq u_n\} \quad k \geq 1.$$

Denote by $Z_j(u_n)$ the length of a run of consecutive exceedances after the occurrence of an upcrossing at time j, and represent by

$$\tilde{\pi}_n(k) = P[Z_j(u_n) = k]$$
$$= P[Y_j = k \,/\, X_{j-1} \leq u_n < X_j],\ k \geq 1,$$

its probability distribution (which does not depence on j, due to stationarity).

The following statements established in Ref. [4] summarize a few results that are essential to our work.

Proposition 4. Let $\{X_i\}$ be a stationary sequence with extremal index θ. If $u_n = u_n(\tau)$ is such that $D(u_n)$ and $D''(u_n)$ hold then

$$\lim_{n\to\infty} E[Z_1(u_n)] = \theta^{-1}$$

Proof:
In order to simplify the notation, let Z denote the random variable $Z_1(u_n)$. Since Z takes only positive integer values,

$$E(Z) = \sum_{k\geq 1} (1 - F_Z(k-1)) = \sum_{k\geq 1} P[Z \geq k]. \qquad (2)$$

But Z represents the length of a successive run, so

$$P[Z \geq k] = P[X_1 > u_n; X_2 > u_n; \ldots; X_k > u_n \,|\, X_0 \leq u_n < X_1]$$
$$= \frac{P[X_0 \leq u_n;\, X_1 > u_n;\, \ldots;\, X_k > u_n]}{P[X_0 \leq u_n < X_1]}$$
$$= \frac{P[X_1 > u_n; \ldots; X_k > u_n] - P[X_0 > u_n; \ldots; X_k > u_n]}{P[X_0 \leq u_n < X_1]}$$
$$= \frac{P[Y_1 \geq k] - P[Y_1 \geq k+1]}{P[X_0 \leq u_n < X_1]}.$$

Hence the mean value of Z can be easily calculated from (2)

$$E(Z) = \frac{\sum_{k\geq 1}\left(P[Y_1\geq k]-P[Y_1\geq k+1]\right)}{P[X_0\leq u_n<X_1]} = \frac{P[Y_1\geq 1]}{\mu(u_n)}.$$

The result follows immediately since $P[Y_1\geq 1] = P[X_1\geq u_n]$ and $u_n=u_n(\tau)$ $(=\tilde{u}_n(\theta\tau))$. ♦

Proposition 5. If u_n is a sequence of normalized levels for which both conditions $D(u_n)$ and $D''(u_n)$ are verified then, for each k=1,2,...,

$$\lim_{n\to\infty}[\pi_n(k) - \tilde{\pi}_n(k)] = 0.$$

After this result the following proposition, analogous to proposition 2, becomes quite apparent:

Proposition 6. Assume that $\Delta(u_n)$ and $D''(u_n)$ hold for a sequence of levels $u_n=\tilde{u}_n(\nu)$, with $\nu>0$. Then, if $Z_1(u_n)$ converges in distribution to some non degenerated random variable Z, the point process of exceedances N_n converges vaguely to a point process N such that

$$N(B) = \sum_{i=1}^{\tilde{N}(B)} Z_i, \quad B\subset[0,1],\ \{Z_i\} \text{ i.i.d. with } Z,$$

where $\tilde{N}$ is the existing Poisson limit of $\tilde{N}_n$. Hence the point process N is a compound Poisson process with intensity ν and multiplicity Z.

The estimation of the extremal index presented in Ref. [4] is based in these last results. From proposition 4 we see that in the limit the cluster centers can be identified with the upcrossings whereas, by proposition 1, the mean size of each cluster is aproximately θ^{-1}. The suggested estimator for θ^{-1} is then constructed by dividing the total number of exceedances of a conveniently chosen level by the total number of upcrossings of that level.

2. An estimator for the extremal index, under D''

Given a sequence of levels u_n define for each n the random variable $N(n,u_n)$ which represents the number of exceedances of u_n in a sample of size n, i.e.,

$$N(n,u_n) = \sum_{i=1}^{n} \mathbb{1}(X_i>u_n).$$

Note that $N(n,u_n)$ is nothing but the measure of the interval [0,1] through the time normalized point process N_n.

In a similar way we denote by $\tilde{N}(n,u_n)$ the number of upcrossings of u_n in a sample of size n, i.e.,

$$\tilde{N}(n,u_n) = \sum_{i=1}^{n} \mathbb{1}(X_{i-1}\leq u_n<X_i).$$

Now, the mean values of these variables are $E(N(n,u_n))=nP(X_1>u_n)$ and $E(\tilde{N}(n,u_n))=n\mu(u_n)$, respectively. So, under $D''(u_n)$ and if the extremal index exists, we have, by (1),

$$\lim_{n\to\infty}\frac{E[N(n,u_n)]}{E[\tilde{N}(n,u_n)]} = \frac{1}{\theta}$$

for any sequence of normalized levels. If we consider the sequence of levels in such a way that $E[\tilde{N}(n,u_n)]\to 1$, when $n\to\infty$, then the reciprocal of the extremal index will merely be the limit of the mean number of exceedances.

For i=1,2, ..., k_n, let u_{ni} be a sequence such that $r_n\mu(u_{ni})\to 1$ (recall that k_n is the number of blocks of size r_n in which we subdivide the sample). Represent by $N_i(r_n,u_{ni})$ the number of exceedances of u_{ni} in block i, and consider the following estimator for the extremal index

$$\tilde{\tilde{\theta}}_n = \left(\frac{1}{k_n}\sum_{i=1}^{k_n} N_i(r_n,u_{ni})\right)^{-1}.$$

To prove the consistency of $\tilde{\tilde{\theta}}_n$ we use the assimptotic independence of $a_nN_i(r_n,u_n)$, i=1,2, ..., k_n, for any sequence of real numbers $\{a_n\}$, under the validity of the condition

$$k_n[\alpha_n(\ell_n-2,u_n) + P(M_{\ell_n}>u_n)] \to 0 \qquad (3)$$

for some ℓ_n (Lema 5.2.1, Ref. [4]).

Proposition 7. Let k_n be a sequence of integers such that $k_n\to\infty$ and let $r_n=[n/k_n]$. Suppose that the stationary sequence $\{X_i\}$ has an extremal index θ and that for a sequence of levels u_n verifying $r_n\mu(u_n)\to 1$, $\Delta(u_n)$, $D''(u_n)$ and (3) hold. Then

$$\frac{\sum_{i=1}^{k_n} N_i(r_n,u_n)}{k_n} \xrightarrow{p} \frac{1}{\theta},$$

under the condition that

$$\lim_{n\to\infty} E[N^2(r_n,u_n)] = c^2 < \infty.$$

Proof:

Consider a sequence $\{N_i^*(r_n,u_n)\}$, $i\geq 1$, of independent random variables identically distributed with $N_1(r_n,u_n)$. Taking $a_n = 1/k_n$ in Lema 5.2.1, Ref. [4], we have the assimptotic independence of $N_i(r_n,u_n)/k_n$ and the following convergenge

$$\frac{\sum_{i=1}^{k_n} N_i(r_n,u_n)}{k_n} - \frac{\sum_{i=1}^{k_n} N_i^*(r_n,u_n)}{k_n} \xrightarrow{p} 0.$$

On the other hand, since for this kind of levels we have $\lim_{n\to\infty} E[N_1(r_n,u_n)] = \theta^{-1}$, the result follows immediately from the convergence in probability of $\sum_{i=1}^{k_n} N_i^*(r_n,u_n)/k_n$ to θ^{-1}. Indeed

$$P\left[\left| \frac{1}{k_n}\sum_{i=1}^{k_n} N_i^*(r_n,u_n) - E\left[N_1(r_n,u_n)\right]\right| > \varepsilon\right]$$

$$= P\left[\frac{1}{k_n}\left|\sum_{i=1}^{k_n}\left(N_i^*(r_n,u_n) - E\left[N_i^*(r_n,u_n)\right]\right)\right| > \varepsilon\right]$$

$$\leq \frac{\operatorname{var}\left[\sum_{i=1}^{k_n} N_i^*(r_n,u_n)\right]}{k_n^2\varepsilon^2} = O\left(\frac{1}{k_n}\right)$$

by the asymptotic boundness of $\operatorname{var}[N_i^*(r_n,u_n)]$. ♦

Proposition 8. If conditions of proposition 7 hold for levels u_{ni}, such that $r_n\mu(u_{ni})\to 1$, $i=1,2,\ldots,k_n$, then $\tilde{\tilde{\theta}}_n$ is consistent.

Proof:

Let u_n and u_n' be two sequences of levels such that $r_n\mu(u_n)\to 1$ and $r_n\mu(u_n')\to 1$. Then, using (1) we derive

$$P[N(r_n,u_n) \neq N(r_n,u_n')]$$

$$\leq r_n\left|P[X_1\leq u_n] - P[X_1\leq u_n']\right|$$

$$= r_n\left(\frac{1}{\theta r_n}(1+o(1)) - \frac{1}{\theta r_n}(1+o(1))\right) = o(1)$$

when n goes to infinity.

Consequentely $N(r_n,u_n) - N(r_n,u_n') \xrightarrow{p} 0$ and the consistency of $\tilde{\tilde{\theta}}_n$ follows from the identity

$$\frac{\sum_{i=1}^{k_n} N_i(r_n,u_n)}{k_n} = \frac{\sum_{i=1}^{k_n}\left(N_i(r_n,u_n) - N_i(r_n,u_{ni})\right)}{k_n} + \left(\tilde{\tilde{\theta}}_n\right)^{-1}$$

where the right hand side converges in probability to $1/\theta$, by proposition 7, and the first term in the left hand side goes to zero, in probability, by the above remark. ♦

The Lindberg condition for the asymptotic normality of $\tilde{\tilde{\alpha}}_n = 1/\tilde{\tilde{\theta}}_n$ is easily established from the fact that $\tilde{\tilde{\alpha}}_n$ is the arithmetic mean of approximately independent random variables.

Proposition 9. Under the conditions of proposition 7 if, for each $\varepsilon>0$,

$$\lim_{n\to\infty} k_n E\left[N^2(r_n,u_n)\,\mathbb{1}(N(r_n,u_n)>\varepsilon)\right] = 0$$

(Lindberg condition)

then

$$k_n^{-1/2}\left(\sum_{i=1}^{k_n} N_i(r_n,u_n) - k_n E(N(r_n,u_n))\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0; \frac{1}{\theta}\sqrt{\theta^2c^2-1}\right) \qquad (4)$$

Proof:

Let $\{N_i^*(r_n,u_n)\}$, $i\geq 1$, be a sequence of i.i.d. random variables with the distribution of $N(r_n,u_n)$. This sequence verifies the conditions of Theorem 3 of Ref. [6], p.101, from which, using the convergence of $E(N(r_n,u_n))$ to $1/\theta$, we conclude that

$$k_n^{-1/2}\left(\sum_{i=1}^{k_n}\left(N_i^*(r_n,u_n) - E(N_i^*(r_n,u_n))\right)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0; \frac{1}{\theta}\sqrt{\theta^2c^2-1}\right).$$

The asymptotic distribution (4) follows immediately since

$$\left(\sum_{i=1}^{k_n} N_i(r_n,u_n) - k_n E(N(r_n,u_n))\right)$$

$$= \left(N(n,u_n) - \sum_{i=1}^{k_n} N_i^*(r_n,u_n)\right) +$$

$$+ \left(\sum_{i=1}^{k_n} \left(N_i^*(r_n,u_n) - E(N_i^*(r_n,u_n))\right)\right)$$

and by Lema 5.2.1, from Ref. [4],

$k_n^{-1/2}\left(N(n,u_n) - \sum_{i=1}^{k_n} N_i^*(r_n,u_n)\right)$ converges in distribution to zero. ♦

Proposition 10. Suppose in addition to the conditions of proposition 9 that

$$\sqrt{k_n}\,(N_i(r_n,u_n) - N_i(r_n,u_{ni})) \xrightarrow{p} 0. \qquad (5)$$

for levels u_{ni}, such that $r_n\mu(u_{ni}) \to 1$, i=1, 2,..., k_n. Then

$$\sqrt{k_n}\left(\tilde{\tilde{\theta}}_n - \theta_n\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0; \theta\sqrt{\theta^2c^2-1}\right), \qquad (6)$$

where $\theta_n = (E[N(r_n,u_n)])^{-1}$.

If in addition $\theta_n = \theta + o\left(\frac{1}{\sqrt{k_n}}\right)$, then the following convergence also occurs

$$\sqrt{k_n}\left(\tilde{\tilde{\theta}}_n - \theta\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0; \theta\sqrt{\theta^2c^2-1}\right). \qquad (7)$$

Proof:

Let $\tilde{\tilde{\alpha}}_n = 1/\tilde{\tilde{\theta}}_n$ and $\alpha_n = E(N(r_n,u_n))$. After some rearrangements we have

$$\sqrt{k_n}\left(\tilde{\tilde{\alpha}}_n - \alpha_n\right)$$

$$= k_n^{-1/2}\left(\sum_{i=1}^{k_n} N_i(r_n,u_n) - k_n\alpha_n\right) -$$

$$- k_n^{-1/2}\left(\sum_{i=1}^{k_n}\left(N_i(r_n,u_n) - N_i(r_n,u_{ni})\right)\right),$$

from which, using (4) and (5) we first conclude that

$$\sqrt{k_n}\left(\tilde{\tilde{\alpha}}_n - \alpha_n\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0; \frac{1}{\theta}\sqrt{\theta^2c^2-1}\right).$$

The asymptotic distribution established in (6) can easily be obtained using for instance the δ-method (Ref. [7]).

The convergence in (7) is trivial after the decomposition of the first member of (6) into the terms $\sqrt{k_n}\left(\tilde{\tilde{\theta}}_n - \theta\right)$ and $\sqrt{k_n}\left(\theta - \theta_n\right)$. ♦

Remark: Both the consistency and the asymptotic normality of this estimador strongly depend on the applicability of Lema 5.2.1 from Ref. [4], in which condition (3) is assumed. It is easy to see that, for normalized levels u_n, $\Delta(u_n)$ is sufficient to (3), but for practical proposes it is worth while to remark further that, if $r_n\mu(u_n) \to 1$, then

$$k_n\,P(M_{\ell_n} > u_n) \geq k_n\,P(X_1 > u_n) = \frac{k_n}{r_n}\,(\theta + o(1)).$$

So, in this case, $k_n/r_n \to 0$ is a necessary condtion to (3), that is, the number of blocks must be significantly smaller than the size of each block.

3. Simulation results

In this section the results of a simulation study are presented, in which we choose for $\{X_i\}$ the max-autorregressive process studied in Ref. Ref. [8]. More precisely, let $X_i = \max(cX_{i-1}, Y_i)$, $c \in (0,1)$, where the r.v.'s Y_i are i.i.d. with a Fréchet distribution function with parameter α. This process has a stationary distribution, verifies conditions D and D" and has an extremal index $\theta = 1-c^\alpha$. Further, it can be proved that the limiting cluster distribution is Geometric with parameter θ and so, if we denote this r.v. by Z, (in accordance with the notation of section 1), we have $E(Z) = 1/\theta$, $var(Z) = (1-\theta)\theta^2$.

Now, in most practical cases, the levels u_n satisfying the condition $r_n\,P[X_1 \leq u_n < X_2] \sim 1$ are typically unknown since they depend on the knowlege of the joint distribution of (X_1,X_2). As in other methods we will consider the replacement of these deterministic levels by random ones in accordance with the above relation. So we suggest the following strategy:

— Choose k_n (the number of blocks) in such a way that $[n/k_n] = r_n > k_n$.
— In each block pick up all the maximum terms within the points j such that $X_{j-1} \geq X_j$, $X_{j+1} > X_j$.
— For block i, consider U_{ni} as the second greatest value among these maximum terms.

In this way there will be in each block at most one upcrossing of this level which will be as low as possible. In most cases the level U_{ni} corresponds to the second highest peak in block i, as illustrated in fig.1.

As n goes to infinity it is reasonable to think that the exceedances of U_{ni} occur in a cluster that is asymptoticaly independent of the cluster that contains U_{ni}. Thus, the next result supports our choice of U_{ni}.

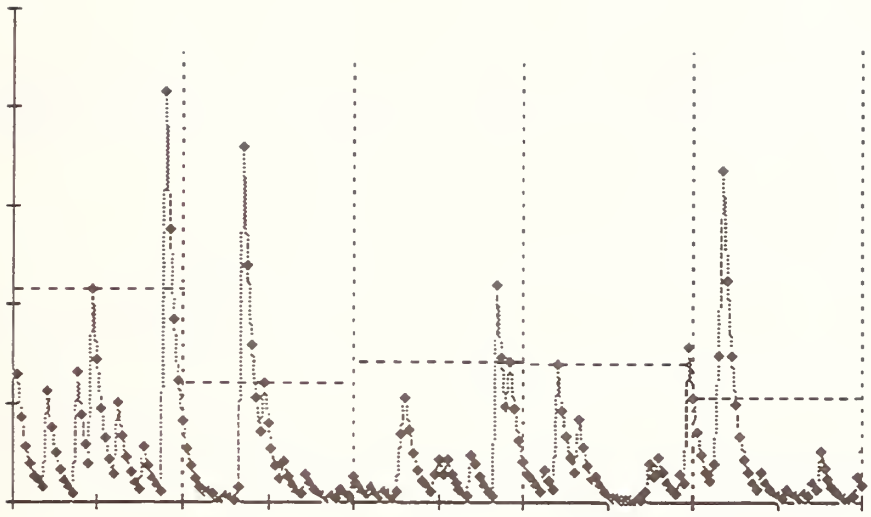


Figure 1: Construction of the random levels U_{ni}

Proposition 11. Let $\{X_i\}$ be a stationary sequence with extremal index θ and a marginal continuous distribution F. Let $M_n=\max(X_1,X_2,\ldots,X_n)$ and consider a r.v. X with the same distribution F, independent of $\{X_i\}$. Then $\lim_{n\to\infty} nP(X>M_n) = \theta^{-1}$.

Proof:

Denote by W_n the r.v. $n\bar{F}(M_n)$, where $\bar{F}=1-F$. By the definition of the extremal index it follows that

$$\lim_{n\to\infty} P\{W_n>x\} = \lim_{n\to\infty} P\{M_n<\bar{F}^{-1}(x/n)\} = e^{-\theta x},\ x>0,$$

since $\bar{F}^{-1}(x/n)$ is a sequence of normalized levels. This means that the distribution of W_n is asymptoticaly exponential. Using this fact we have

$$\begin{aligned} nP\{X>M_n\} &= nP\{n\bar{F}(X)<W_n\} \\ &= \int_0^{+\infty} nP\{n\bar{F}(X)<x\}dF_{W_n}(x) \\ &= \int_0^{+\infty} n\frac{x}{n}dF_{W_n}(x) \\ &= E(W_n) \end{aligned}$$

and the result follows.♦

In our simulation procedures the estimate of θ was computed based on levels U_{ni} determined in accordance to the above described algorithm.

At a first step, sequences of size n=1000 of the max-autorregressive process $\{X_i\}$ with α=0.5, 1.0 and c=0.1, 0.5, 0.9 (i.e., θ = 0.05, 0.1, 0.29, 0.5, 0.8, 0.9) were generated and estimates of θ were computed for r_n= 5, 10, 15, 20, 25, 40, 50, 100, 200, 250. This procedure was repeated 500 times and, for each pair (α,c) and for each r_n,the average and the mean square error of the corresponding estimates of the extremal index were computed.

In fig.2, r_n is plotted against the estimates of $E(\tilde{\tilde{\theta}}_n)$ for the different values of θ considered above. From its observation we can conclude that for r_n ranging between 40 and 100 this procedure seems to work reasonably well for all values of θ. Notice in particular, the apparent agreement with the latter remark about r_n being significantly bigger than k_n.

The mean square errors represented in figure 3, do not contradict the good behaviour of these estimates for values of r_n in the same region.

In order to examine the empirical probability distribution of the number of exceedances of the random levels U_{ni}, which we expect to be approximately geometric, another simulation procedure that generates sequences of n=1000 of $\{X_i\}$, was made for the same pairs (α,c). This procedure was run 1000 times and the empirical distribution of Z, its sample mean and its sample variation were computed for different choices of r_n. Some of these results are presented in table 1. Once again the simulation results suggest that the behaviour of this estimator is more sensitive to the ratio between the size of each block and the total number of blocks than to the value of θ.

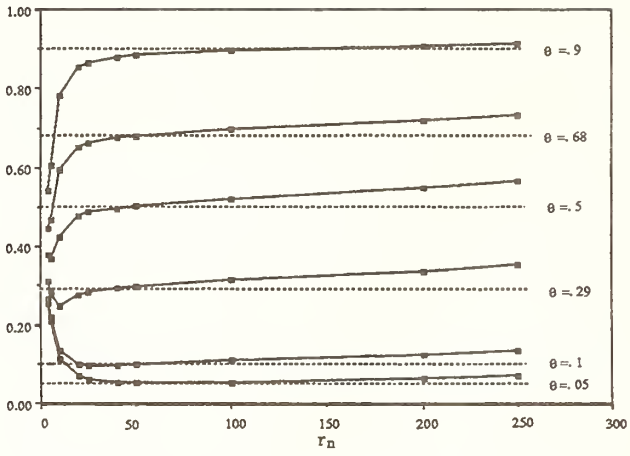


Figure 2. Estimates of $E(\tilde{\tilde{\theta}}_n)$, for n=1000 and r_n = 5,10,15,20,25,50,100,200,250.

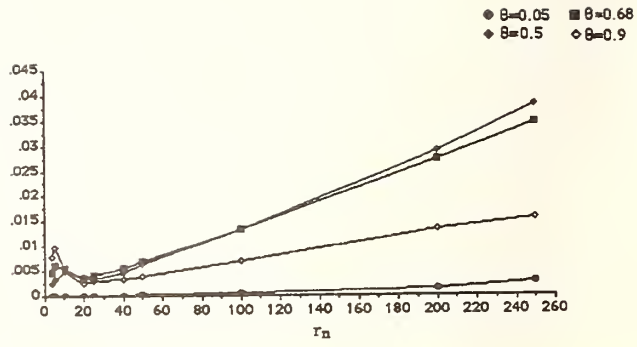


Figure 3. Estimates of MSE$(\tilde{\tilde{\theta}}_n)$, for n=1000 and r_n = 5,10,15,20,25,50,100,200,250.

Table 1. Estimates of $E(\tilde{\tilde{\theta}}_n)$, E(Z) and var(Z), for n=1000 and r_n = 5,10,15,20,25,50,100,200,300,400,500,1000

c α	θ E(Z) Var(Z)	r_n	5	10	15	20	25	50	100	200	300	400	500	1000
0.90	0.05								*0.056*	*0.055*	*0.053*	*0.053*	*0.053*	*0.050*
0.5	*19.5*								*17.72*	*18.30*	*18.71*	*18.82*	*18.69*	*20.01*
	360								299	313	340	358	377	440
0.90	0.1							*0.11*	*0.106*	*0.103*	*0.102*	*0.101*	*0.102*	*0.098*
1.0	*10*							*9.4*	*9.4*	*9.7*	*9.8*	*9.9*	*9.75*	*10.2*
	90							78.6	83.7	83	88	101	96.34	108.2
0.50	0.29					*0.29*	*0.29*	*0.3*	*0.3*	*0.29*	*0.29*	*0.3*	*0.299*	*0.29*
0.5	*3.45*					*3.43*	*3.46*	*3.37*	*3.31*	*3.39*	*3.40*	*3.36*	*3.35*	*3.45*
	8.24					8.16	8.56	7.61	7.56	8.03	8.51	9.17	8.86	9.84
0.50	0.50				*0.49*	*0.49*	*0.49*	*0.51*	*0.50*	*0.50*	*0.50*	*0.50*	*0.505*	*0.495*
1.0	*2*				*2.1*	*2.1*	*2.0*	*2.0*	*2.0*	*2.0*	*2.0*	*2.0*	*1.98*	*2.0*
	2				2.0	2.1	2.1	2.1	2.1	2.0	2.1	2.3	2.2	2.4
0.10	0.68			*0.64*	*0.66*	*0.66*	*0.67*	*0.68*	*0.69*	*0.68*	*0.69*	*0.69*	*0.70*	*0.67*
0.5	*1.46*			*1.56*	*1.52*	*1.52*	*1.50*	*1.48*	*1.44*	*1.46*	*1.46*	*1.45*	*1.44*	*1.48*
	0.68			0.86	0.71	0.74	0.72	0.69	0.62	0.67	0.69	0.75	0.69	0.84
0.10	0.90		*0.71*	*0.82*	*0.86*	*0.87*	*0.87*	*0.90*	*0.91*	*0.90*	*0.90*	*0.90*	*0.91*	*0.89*
1.0	*1.11*		*1.40*	*1.22*	*1.16*	*1.15*	*1.14*	*1.11*	*1.11*	*1.10*	*1.11*	*1.11*	*1.11*	*1.12*
	0.124		0.63	0.30	0.20	0.17	0.16	0.16	0.11	0.125	0.12	0.14	0.13	0.16

References:

[1] Leadbetter, M.R. and Nandagopalan, S., On Exceedance Point Processes for Stationary Sequences under Mild Oscillation Restrictions. In: Extreme Value Theory. (eds.:J. Hüsler and R.-D. Reiss), Springer-Verlag, 1989, p.69-80.

[2] Leadbetter, M.R. et al, Extremes and Related Properties of Random Sequences and Processes, Springer Verlag, New-York, 1983.

[3] Hsing, T. et al, On the Exceedance Point Process for a Stationary Sequence, Prob. Th. Rel. Fields, **78**, 1988, 97-112.

[4] Nandagopalan, S., Multivariate Extremes and Estimation of the Extremal Index, Ph. D. Thesis. Technical Report N°315, Center for Stochastic Processes, University of Norh Carolina, 1990.

[5] Hsing, T., Estimating the Parameters of Rare Events. Preprint, Texas A&M University, ,1990.

[6] Gnedenko, B.V., Kolmogorov, A.N., Limit Distributions for Sums of Independent Random Variables, Addison-Wesley, Cambridge,1954.

[7] Cramér, H., Mathematical Methods of Statistics, Princeton Univ. Press, Princeton, 1946.

[8] Alpuim, M.T., An Extremal Markovian Sequence. J.Appl.Probab., **26**, 1989, 219-232.

Approximate Extreme Value Analysis For A Rigid Block Under Seismic Excitation

Facchini, L. and Spinelli, P.
University of Florence, Florence, Italy

The problem of the collapse risk (due to earthquakes) of a rigid block resting on a rigid foundation is dealt with in this work; and three hypotheses are introduced to describe the series of the seismic events:

1. the arrival rates of the single earthquakes during the structural life can be described by means of a counting process (specifically, a Poisson random process);
2. the yearly maxima for the peak acceleration of the soil can be described by a Fisher-Typpet II distribution;
3. the single event can be described by a random non stationary process.

The behavior of the rigid block has been investigated via the statistical linearization method, which can provide satisfactory approximations of the response of the system.

By means of a reasonable combination of the model for the seismic events and the approximation of the peak rotation of the block, an approximation was obtained of the p.d.f. for the extreme value of the rotation of the block due to earthquakes during a given period of time.

This is of crucial interest in the evaluation of the collapse probability of the block, as it can conveniently be combined with material resistance distribution; the collapse mechanism for such a system involves two distinct characteristics: the former is the overturning condition, and the latter concerns the resistance of the block material (that is, the material can collapse before the overturning condition is reached, thus causing the structural failure).

Introduction

The mechanical system taken into consideration is sketched in fig. 1; it is a rigid block free to rock without sliding on either the base corners; its foundation is a rigid horizontal plane which moves in the x-direction according to a given function of time $x_g(t)$ which, in this particular case, will be assumed to be a realization of the random process $X_g(t)$.

In addition, let $R = \frac{1}{2}\sqrt{H^2 + B^2}$ and $\theta(t)$ be the angle measuring the tilting of the block; positive or negative θ means rocking about corner O or corner O' respectively.

The equations of motion about each one of the base corners [5] are:

$$\begin{aligned} &\ddot{\varphi} - \varphi + 1 = f(\tau) \\ &\ddot{\varphi} - \varphi - 1 = f(\tau) \end{aligned} \qquad 2.1$$

These equations describe the evolution of the system via the lagrangian coordinate $\varphi = \theta/\theta_{cr}$, where θ_{cr} is the critical toppling angle (see fig. 1) and f is a function of non-dimensional time parameter τ;

besides, the non dimensional time parameter $\tau = \mu t$ was introduced, where μ is the frequency of small oscillations of the block, when it is suspended from one of the base corners; specifically,

$$\mu = \sqrt{\frac{MgR}{I_O}} \qquad 2.2$$

where M is the mass of the block, g is the gravity acceleration and I_O is the moment of inertia about the corner O.

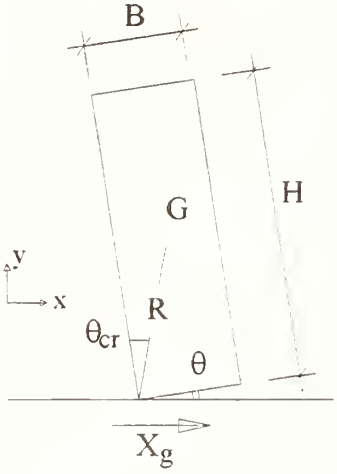


fig. 1

Impact occurs between the block and its foundation whenever there is a transition from rocking about one corner to rocking about the other; the associated energy loss is accounted for by reducing the angular velocity of the block after impact.

Specifically, it is assumed that

$$\lim_{\tau \to \tau^+} \dot{\varphi}(\tau) = e \cdot \lim_{\tau \to \tau^-} \dot{\varphi}(\tau) \qquad 2.3$$

where e is defined as the coefficient of restitution of the angular velocity, and τ^+ and τ^- are respectively the nondimensional time parameters just after and just before impact.

The statistical linearization of the system

The two equations of motion (2.1) and the linking condition (2.3) can be conveniently summarized into the expression

$$\ddot{\varphi} + (1-e)\delta(\varphi)\dot{\varphi}^- \left|\dot{\varphi}^-\right| - \varphi + sign(\varphi) = f(\tau) \qquad 3.1$$

Here, $\delta(\cdot)$ is the Dirac Delta function.

It is clear that the given system, even when it is slender enough and the equations of motion can be piecewise linearized, is still strongly non-linear.

When $f(\tau)$ is a deterministic function of time, the behavior of the block can be investigated easily enough because a closed form solution can generally be obtained for each half cycle; on the other hand, if $f(\tau)$ is a realization of a random process, it is generally necessary to make use of approximate methods, such as statistical linearization [1] [4] is.

Before proceeding, we briefly discuss the assumptions on the forcing process $f(\tau)$.

Following Kanai-Tajimi [3] [6] theory, the power spectral denity of the baseline excitation was assumed:

$$S_{\ddot{X}\ddot{X}}(\omega) = \\ = S_0 \frac{1 + 4\xi_g^2 (\omega/\omega_g)^2}{\left[1 - (\omega/\omega_g)^2\right]^2 + 4\xi_g^2 (\omega/\omega_g)^2} \qquad 3.2$$

where $\xi = 0.60$ and $\omega_g = 5\pi$; S_0 is the intensity function of a non stationary white noise obtained from a Gaussian stationary white noise multiplied by a deterministic function of (actual) time $\chi(t)$; in particular, $\chi(t)$ is made up by:

1. a quadratic build-up:
   ($\chi(t) = t^2/16$ for $t \le t_1 = 4$ secs.);
2. a constant term equal to unity (for 4 secs.$= t_1 < t \le t_2 = 15$ secs.);
3. an exponential decay:
   $(\chi(t) = exp(-0.0992(t - t_2))$
   for 15 secs.$= t_2 < t$) (fig. 2).

It can be shown that the power spectrum of the forcing process f(τ) is linked to Kanai-Tajimi model by means of the relation:

$$S_{FF}(\omega) = \frac{\mu}{(g\theta_{cr})^2} S_{XX}(\omega) \qquad 3.3$$

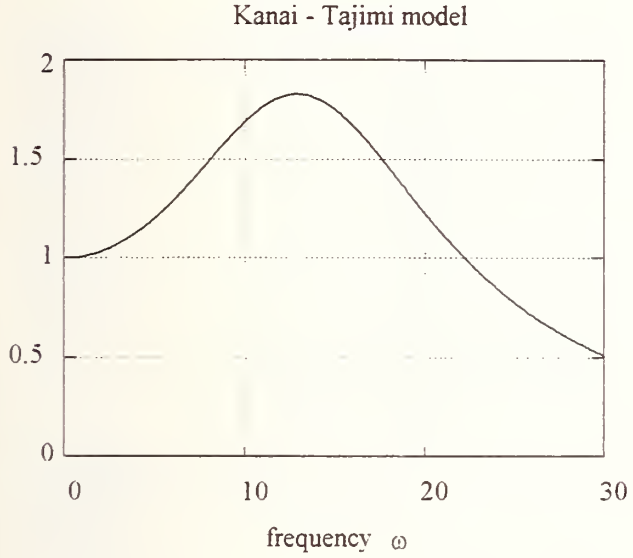
Thus, one has to investigate the response of a linear system whose equation of motion is

$$\ddot{\varphi}_e + c_e\dot{\varphi}_e + (k_e - 1)\varphi_e = f(\tau) \qquad 3.4$$

where

$$c_e = \frac{4}{\pi}\frac{1-e}{1+e}\frac{\sigma_{\dot{\Phi}_e}}{\sigma_{\Phi_e}} \quad k_e = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma_{\Phi_e}} \qquad 3.5$$

and σ_{Φ_e} and $\sigma_{\dot{\Phi}_e}$ are respectively the standard deviations of the two processes Φ_e and $\dot{\Phi}_e$.



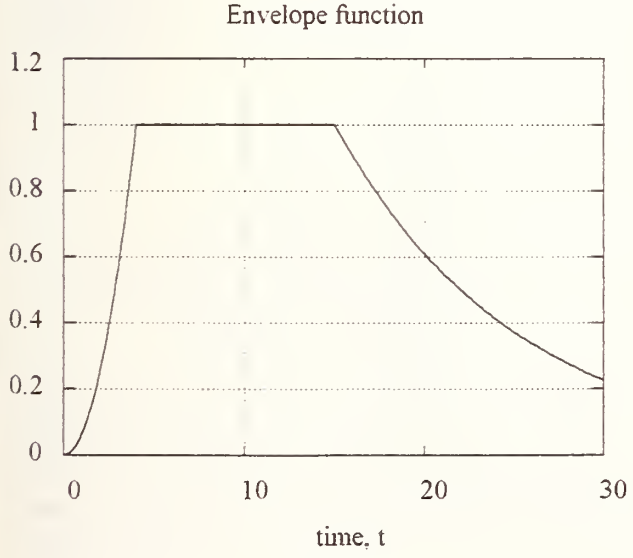


fig. 2

It has been assumed that, when $\chi(t) = 1$, i.e. from $t_1 = 4$ secs. to $t_2 = 15$ secs. (or at least during a portion of this time period), both the excitation and the system response could be regarded as if they were stationary. This is a rough assumption, and studies are at present being carried out to evaluate its adequacy. However, numerical results from [8] (see next paragraph), suggest it yelds satisfactory results.

As usual, the variances of the two processes, tilt angle and angular velocity, were estimated via the following relations:

$$\sigma^2_{\dot{\Phi}_e} = \int_{-\infty}^{\infty} \omega^2 S_{\Phi_e\Phi_e}(\omega)d\omega \qquad 3.6a$$

$$\sigma^2_{\Phi_e} = \int_{-\infty}^{\infty} S_{\Phi_e\Phi_e}(\omega)d\omega \qquad 3.6b$$

where

$$S_{\Phi_e\Phi_e}(\omega) = \frac{S_{FF}(\omega)}{\left[\omega - (k_e - 1)^2\right]^2 + c_e^2\omega^2} \qquad 3.7$$

Extreme value distribution for a single earthquake

Once the excitation intensity is given via the parameter S_0, the extreme value distribution during the time period 4 secs.$\le t \le$ 15 secs., i.e. when the response is assumed to be stationary, can be evaluated for the equivalent linear system by means of Vanmarcke's [7] relation

$$F_{\overline{\Phi}_e}(\overline{\varphi}_e | a_p) =$$

$$= exp\left[-\nu_e T \frac{1 - exp\left(-\sqrt{\frac{\pi}{2}}\frac{q\overline{\varphi}_e}{\sigma_{\Phi_e}}\right)}{exp\left(\frac{\overline{\varphi}_e^2}{2\sigma^2_{\Phi_e}}\right) - 1}\right] \qquad 4.1$$

where:

$$\nu_e = \frac{1}{2\pi}\frac{\sigma_{\dot{\Phi}_e}}{\sigma_{\Phi_e}} \quad ; \quad q = \sqrt{1 - \frac{\lambda_1^2}{\lambda_0\lambda_2}} \qquad 4.2a$$

$$\lambda_i = \int_0^\infty \omega^i S_{\Phi_e \Phi_e}(\omega) d\omega \qquad 4.2b$$

The variable a_p does not affect directly F_{Φ_e}; λ_i and q are functions of $S_{\Phi_e \Phi_e}$ and therefore of a_p.

By means of relation 4.1, it has been possible to check the validity of the approximations introduced by the equivalent linearization; some numerical results were available from [8].

Reference of [8] studied the behavior of a rigid block with aspect ratio $R = 10$ ft. and slenderness $H/B = 5$; it was subjected to several ground accelerations histories which followed the Kanai-Tajimi model. The mean peak acceleration was $a_p = 0.4g$.

The coefficient of restitution was varied from $e = 0.90$ to $e = 0.95$: twenty histories for each one of the values of e were numerically integrated, and the empirical cumulative distributions were evaluated.

Fig. 3 shows the comparison between the empirical frequencies obtained in Ref. [8] (the crosses) and the curve obtained by Vanmarcke's relation (4.1) for the equivalent linear systems (computed according to the proposed formulas) for four values of e (continuous line); the dashed lines are the boundary of the confidence interval of a Kolmogorov-Smirnov goodness-of-fit test.

Since each cross is contained in the confidence interval, this gives evidence that the approximation obtained was satisfactory.

Then, as the values of e did not differ much from each other, all the empirical c.d.f. were consolidated so as to obtain a sample population of 120 samples: an equivalent linear system was computed imposing $e = 0.925$ (that is, the mean of the previous values taken for e).

Fig. 4 shows the extreme value distribution of Φ_e together with the empirical frequencies and the confidence interval of a Kolmogorov Smirnov test.

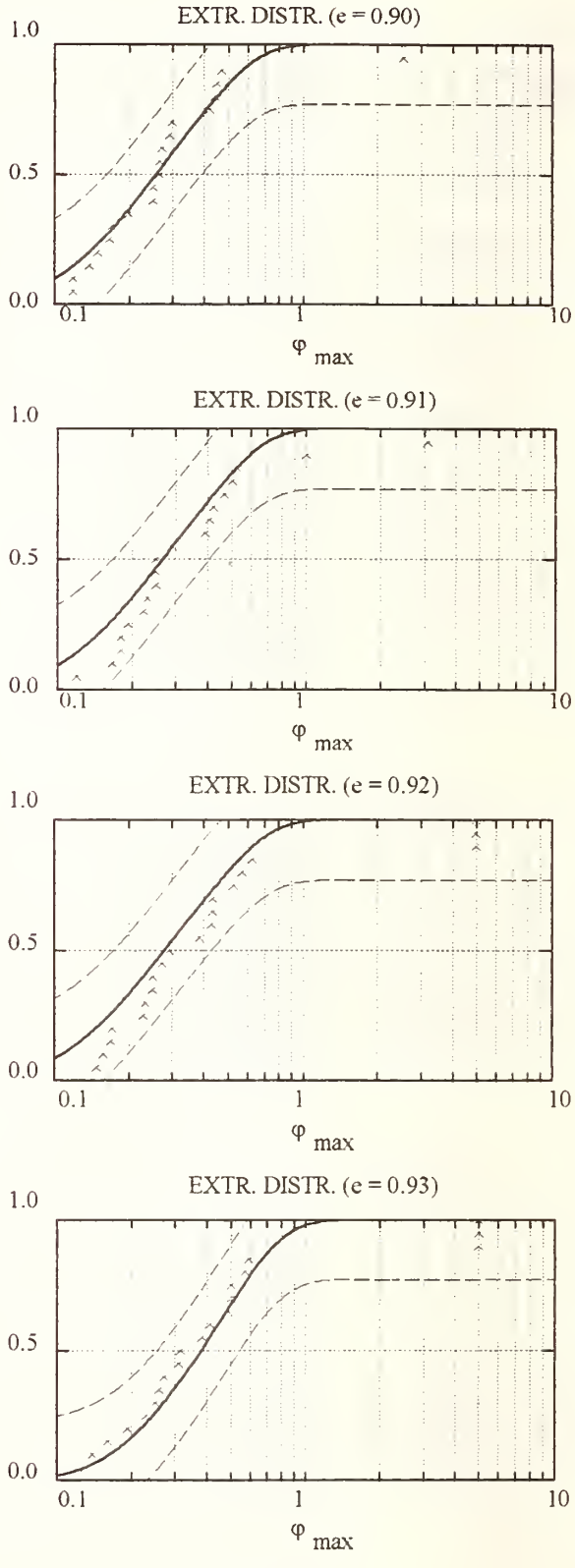


fig. 3

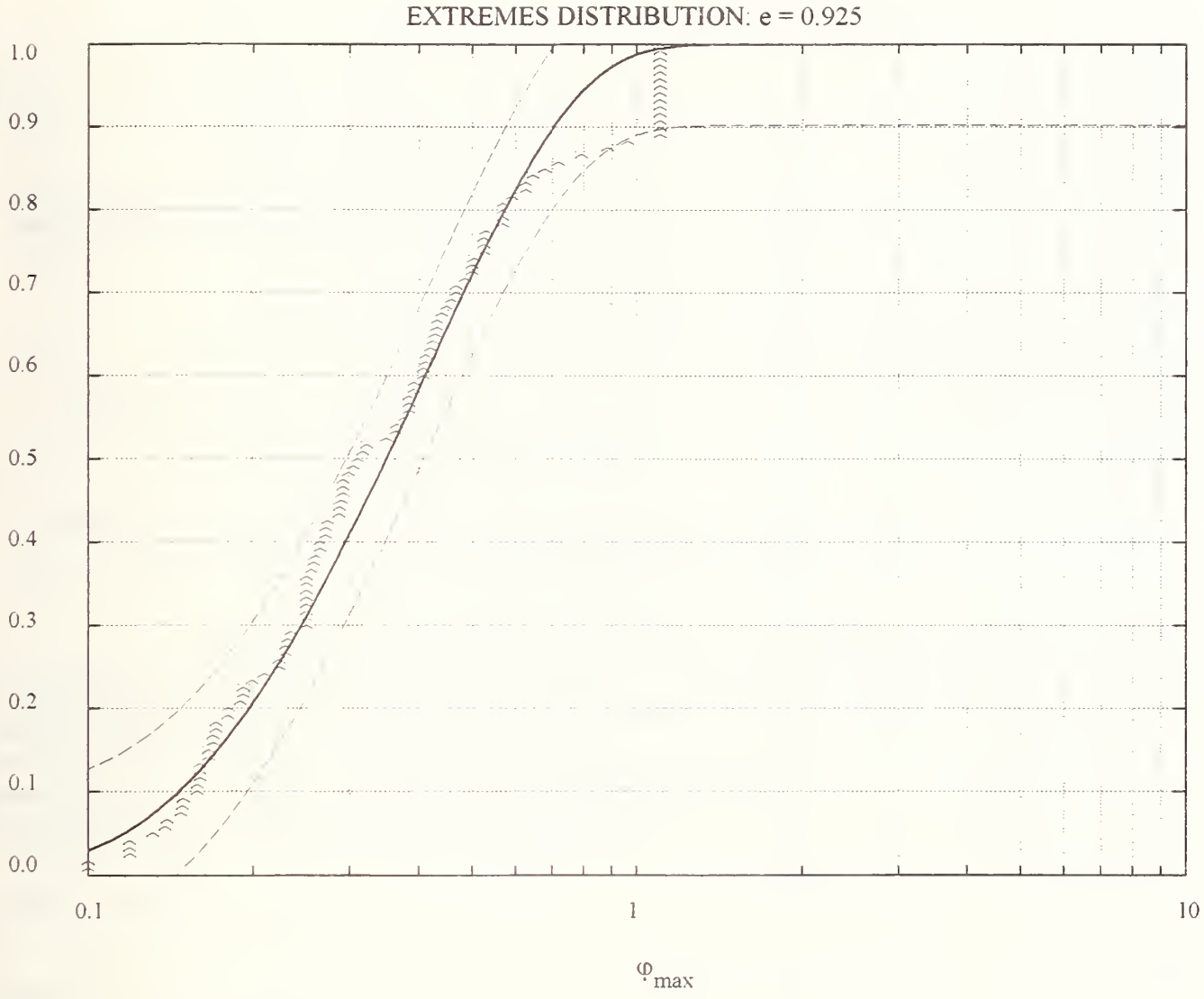


fig. 4

Extreme value distribution

during the structural life

In the previous section the method for obtaining the cumulative distribution function for the extreme rotation of the block has been described. The starting point was the shape of the power spectral density of the forcing process, coupled to the mean of its peak acceleration.

If an upper bound is imposed for the extreme rotation of the block, for instance the toppling condition $|\varphi| \leq \overline{\varphi} = 1$, a safety region can be determined and the curve which expresses the collapse probability conditional on a given event having occurred can be plotted.

In order for the seismic vulnerability to be estimated, the collapse probability conditional on a given event having occurred must be combined with the probability that this event takes place during the structure's life.

The upper bound for the rotation will be denoted by $\overline{\varphi}$ and $S \equiv [-\overline{\varphi}; \overline{\varphi}] \in \mathbf{R}$ will stand for the safety region of the dynamic system.

Let $\mathcal{B}_S(\varphi | a_p)$ be the safety likelihood and $\mathcal{B}_C(\varphi | a_p)$ the collapse probability for the block

subjected to a seismic motion whose mean peak acceleration is a_p:

$$\mathcal{B}_c(\varphi|a_p) = 1 - \mathcal{F}_s(\varphi|a_p) = \\ = Prob\left[|\varphi| < \overline{\varphi}|a_p\right] \quad 5.1$$

We will assume that the time series of the seismic events can be modeled as a Poisson process and that the safety probability of the system does not change with time: then the safety probability of the system during its life T_{str} is [2]

$$S(T_{str}) = \\ = exp\left\{-T_{str}\int_0^\infty \mathcal{B}_c(\varphi|a_{py})p_{A_{py}}(a_{py})da_{py}\right\} \quad 5.2$$

where $p_{A_{py}}$ is the probability distribution function of the yearly maximum of mean peak acceleration of seismic motion for the considered site.

From eqs. 4.1 and 5.1

$$\mathcal{B}_c(\varphi|a_p) = 1 - F_{\overline{\Phi}_e}(\overline{\varphi}, a_p) \quad 5.3$$

Once the probability density function (p.d.f.) of yearly peak acceleration a_{py} is given, eq. 5.2 may be viewed as the probability of the rotation being less than an arbitrarily fixed value φ_s during the system life T_{str}; from this new point of view, eq. 3.1 defines the cumulative probability function for the extreme rotation during the whole structural life.

This new function takes into account the random structure of the seismic events in the considered site by means of the p.d.f. of yearly maxima of ground acceleration $p_{A_{py}}(a_{py})$:

$$P_{\Phi_s}(\varphi_s) = \\ = exp\left\{-T_{str}\int_0^\infty \mathcal{I}(\varphi_s, a_{py})da_{py}\right\} = \\ = Prob\left[\varphi \le \varphi_s\right] \quad 5.4$$

where

$$\mathcal{I}(\varphi_s, a_{py}) = \\ = \left[1 - F_{\overline{\Phi}_e}(\varphi_s, a_{py})\right]p_{A_{py}}(a_{py})$$

during the whole structural life T_{str}.

The corresponding p.d.f. can be obtained by a derivation of this last equation; it can be found that

$$p_{\Phi_s}(\varphi_s) = \frac{\partial}{\partial\varphi_s}P_{\Phi_s}(\varphi_s) = \\ = -T_{str}\,exp\left\{-T_{str}\int_0^\infty \mathcal{I}(\varphi_s, a_{py})da_{py}\right\}\cdot \\ \cdot\int_0^\infty \frac{\partial}{\partial\varphi_s}\mathcal{I}(\varphi_s, a_{py})da_{py} \quad 5.5$$

Applications

The proposed method was used to evaluate the p.d.f. of the angle of rotation of Foca's column in Rome during a period $T_{str} = 50$ years; the structure can be roughly sketched as a 12.83 mt. cylindrical rigid block whose medium diameter is approximately 1.27 mt (see fig. 5).

It shows a critical angle $\theta_{cr} = 0.0984$ rads. and a natural frequency $\mu = 1.069\ s^{-1}$.

Several forcing processes were considered (following Kanai-Tajimi model with mean peak accelerations ranging from $0.001g$ to $1.000\ g$) and an equivalent linear system was computed for each of them; in fig. 6 the extreme value c.d.f.'s are reported for ten processes, whose mean peak acceleration ranges from $0.1g$ to $1.0g$.

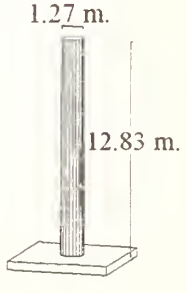


fig. 5

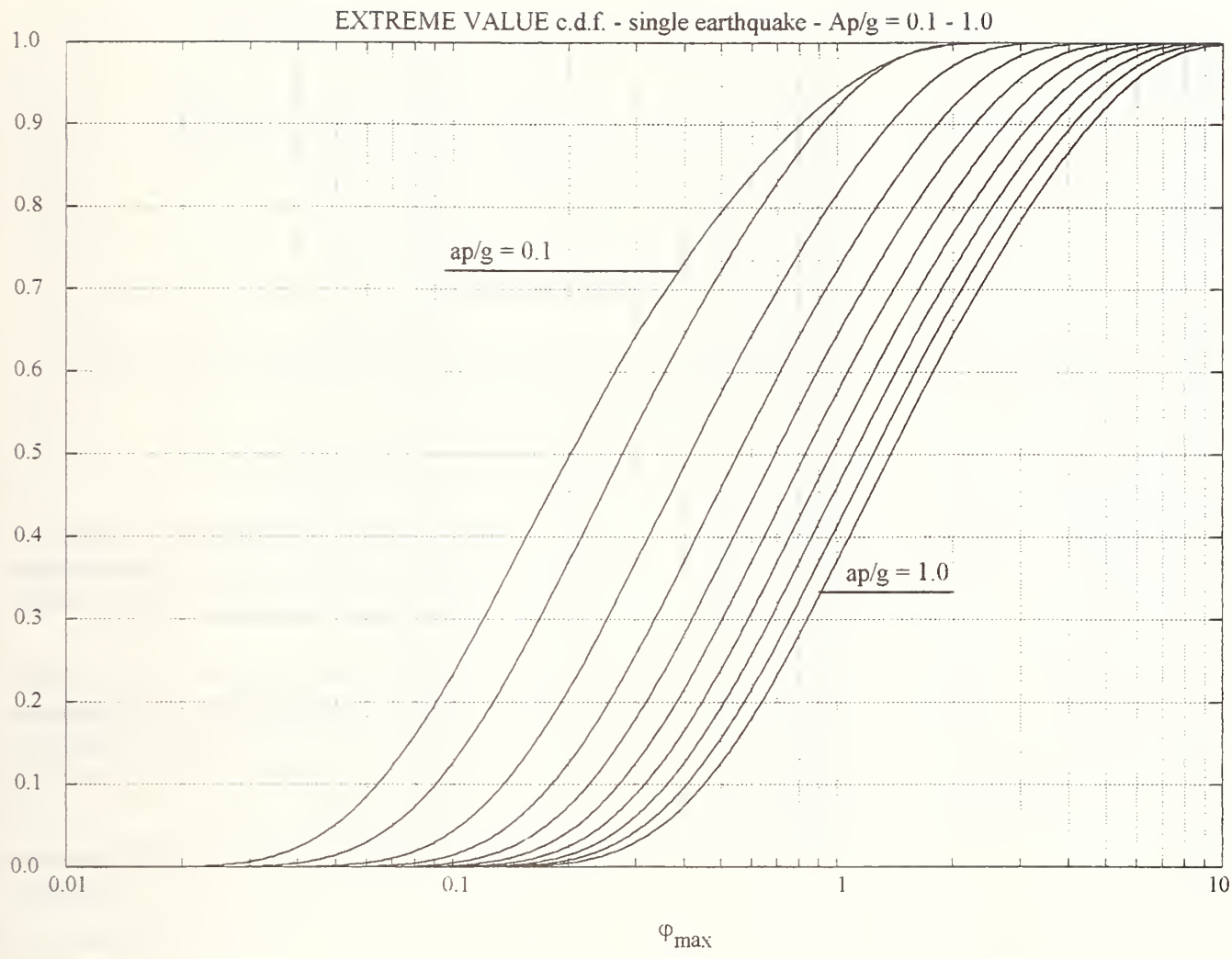

EXTREME VALUE c.d.f. - single earthquake - Ap/g = 0.1 - 1.0
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
ap/g = 0.1
ap/g = 1.0
0.01
0.1
1
10
φmax

fig. 6

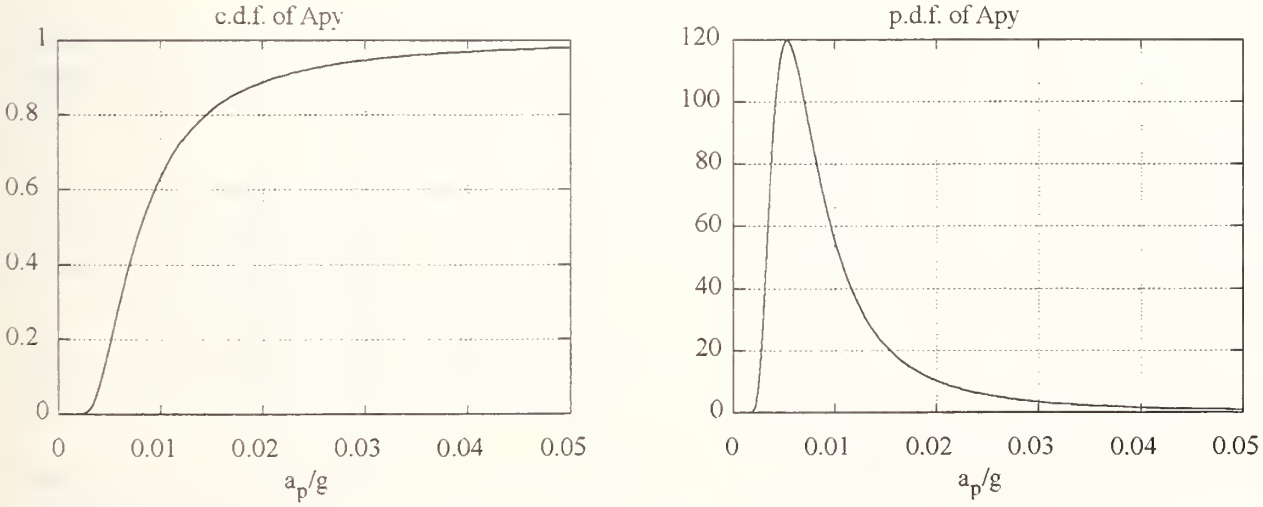
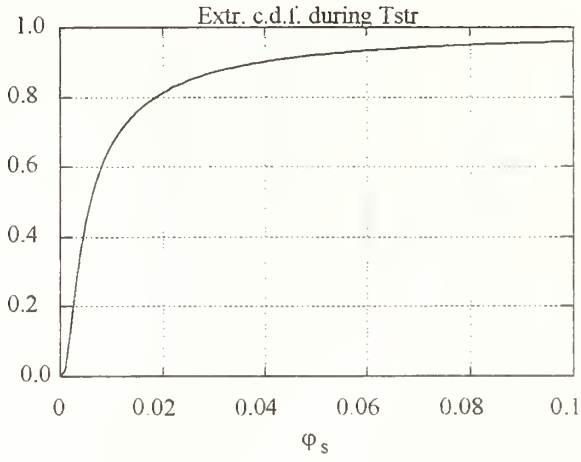

c.d.f. of Apy
1
0.8
0.6
0.4
0.2
0
0
0.01
0.02
0.03
0.04
0.05
ap/g
p.d.f. of Apy
120
100
80
60
40
20
0
0
0.01
0.02
0.03
0.04
0.05
ap/g

fig. 7



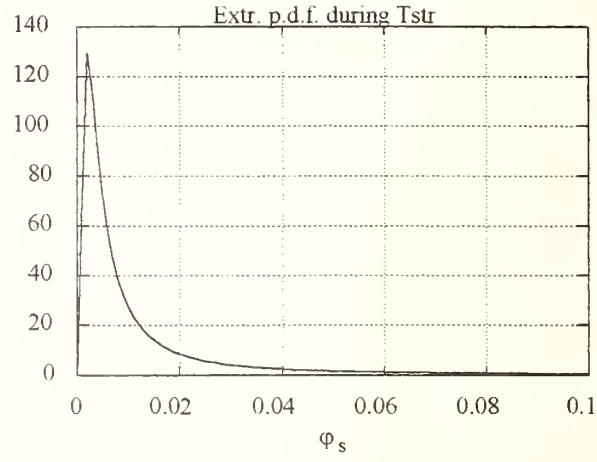


fig. 8

This results were combined with the probability distribution function of the yearly maximum of mean peak acceleration of the ground; a Fisher-Typpet II type c.d.f. for the yearly peak acceleration was used:

$$P_{A_{py}}(a_{py}) = exp\left[-(u_1/a_{py})^{u_2}\right]$$

$$p_{A_{py}}(a_{py}) = \frac{u_2}{u_1}\left(\frac{u_1}{a_{py}}\right)^{u_2-1} exp\left\{-\left(\frac{u_1}{a_{py}}\right)^{u_2}\right\}$$

where $u_1 = 6.6552 \cdot 10^{-3}$ and $u_2 = 1.9313$ (see fig. 7).

The p.d.f. for the extreme rotation of the column was computed (see fig. 8) by means of equations 5.4 and 5.5 assuming $T_{str} = 50$ years.

Conclusions

A method to evaluate the extreme value distribution for the response of a SDOF non linear system was proposed; this method makes a massive use of the statistical linearization technique, and is based on the assumption that both the excitation and the response are stationary for a certain period of time, even though they are not.

An approximation of the extreme value distribution of the response of a nonlinear system during a given time period (in our case, 50 years) was obtained.

References

[1] L. Facchini, Strutture monolitiche sottoposte ad azioni aleatorie: applicazione al comportamento sismico di opere monumentali, Graduate Thesis, Univ. of Florence, 1990.

[2] V. Gusella. Structural failure and stochastic discrete process of random events: an application to historic building seismic vulnerability analysis, Dept. of Civil Engineering report, Univ. of Florence, 1990.

[3] K. Kanai, An empirical formula for the spectrum of strong earthquake motions, Bull. Earthquake Research Inst., Univ. of Tokyo, 39, 1961.

[4] J. B. Roberts, P. D. Spanos, Random Vibration & Statistical Linearization, John Wiley & Sons, 1990.

[5] P. D. Spanos, A. S. Koh, Rocking of rigid blocks due to harmonic shaking, Journal of Engineering Mechanics, Vol. 110, No. 11, November 1984, pages 1627 - 1642.

[6] H. Tajimi, A statistical method of determining the maximum response of a building during an earthquake, Proc. 2nd World Conf. Earthquake Eng., vol. II, Tokyo & Kyoto, 781-798, 1960.

[7] E. H. Vanmarcke, On the distribution of the first-passage time for normal stationary random processes, J. Appl. Mechanics, 42, 1975.

[8] C.-S. Yim, A. K. Chopra, J. Penzien, rocking response of rigid blocks to earthquakes, Earthquake Engineering & Structural Dynamics, 8, 565-587, 1980.

The Rate Of Convergence Or Divergence For Percentiles Of Gamma Distributions And Its Application To Sample Extremes

Gan, G. and Bain, L.J.
University of Missouri, Rolla, MO

For a sequence (α_n,β_n) of positive constants with $\alpha_n\downarrow 0$ and $\beta_n\downarrow 0$ as $n\uparrow\infty$, it is shown that if $\lim_{n\uparrow\infty}\frac{\ln(\alpha_n)}{\ln(\beta_n)}=a\geq 0$, then $\lim_{n\uparrow\infty}\frac{\eta_{1-\alpha_n}(\theta_1,k_1)}{\eta_{1-\beta_n}(\theta_2,k_2)}=a\frac{\theta_1}{\theta_2}$, where $\eta_p(\theta,k)$ is the 100pth percentile of the gamma distribution with mean $k\theta$ and variance $k\theta^2$, and if $\lim_{n\uparrow\infty}\frac{\alpha_n}{\beta_n}=a\geq 0$, then $\lim_{n\uparrow\infty}\frac{\eta_{\alpha_n}(\theta_1,k_1)}{\eta_{\beta_n}(\theta_2,k_2)}=0$ if $k_1<k_2$, $=\infty$ if $k_2<k_1$, and $=a^{1/k}\theta_1/\theta_2$ if $k_1=k_2$. An example of its application to sample extremes is given.

1. INTRODUCTION

It is well known that the normalizing constants to any domain of attraction are all related to the percentiles of the underlying distribution, and that the gamma distribution belongs to the domain of attraction of the exponential type for the maximum and is attracted to the Weibull distribution for the minimum. Also, by Ref. [1], the moments of normalized extremes from the gamma distribution tend to the moments of the appropriate limiting distribution, and so the approximation for the moments of sample extremes based on the limiting extreme value distribution is considered.

If one is interested in asymptotic expectations of sample extremes from gamma distributions and their ratio, the rate of convergence or divergence for percentiles of gamma distributions will be helpful. This paper will derive the rate in Section 2, and give an example of its application to sample extremes in Section 3.

Throughout this paper, $\eta_p(\theta,k)$, $\theta>0$ and $k>0$, will denote the 100pth percentile of the gamma distribution $Gam(\theta,k)$ whose density function is given by

$$f(x;\theta,k) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k}, \quad x>0,$$

and distribution function is denoted by $F(x;\theta,k)$, "$x\to\psi(x)$" will represent a function ψ of x, and "$\zeta(x)\downarrow 0$ $(\zeta(x)\uparrow 0)$ as $x\to\infty$" will represent that ζ is strictly decreasing (increasing) and $\lim_{x\to\infty}\zeta(x)=0$.

2. MAIN RESULT

LEMMA 1. Let $\{\alpha_n\}$ be a sequence of positive constants with $\alpha_n\downarrow 0$ as $n\uparrow\infty$. Then

$$(1) \qquad \lim_{n\uparrow\infty}\frac{\eta_{1-\alpha_n}(\theta_1,k)}{\eta_{1-\alpha_n}(\theta_2,1)} = \frac{\theta_1}{\theta_2},$$

and

$$(2) \qquad \lim_{n\uparrow\infty}\frac{\eta_{\alpha_n}(\theta_1,k_1)}{\eta_{\alpha_n}(\theta_2,k_2)} = \begin{cases} 0 & 0<k_1<k_2 \\ \infty & 0<k_2<k_1 \\ \theta_1/\theta_2 & 0<k_2=k_1 \end{cases}.$$

PROOF. Since $\alpha_n\to\eta_{1-\alpha_n}(\theta,k)$ is decreasing and differentiable for $\alpha_n\in]0,1[$, by L'Hospital's rule,

$$\lim_{\alpha_n\downarrow 0}\frac{\eta_{1-\alpha_n}(\theta_1,k)}{\eta_{1-\alpha_n}(\theta_2,1)} = \lim_{\alpha_n\downarrow 0}\frac{\eta_{1-\alpha_n}(\theta_1,k)}{-\theta_2\ln[1-F(\eta_{1-\alpha_n}(\theta_1,k);\theta_1,k)]} = \frac{\theta_1}{\theta_2},$$

and (1) holds because $\alpha_n\downarrow 0$ as $n\uparrow\infty$. Next for $\alpha_n\in]0,1[$, since

$$(3) \qquad \alpha_n = F[\eta_{\alpha_n}(\theta_i,k_i);\theta_i,k_i], \quad i=1,2,$$

and $\alpha_n\to\eta_{\alpha_n}(\theta_i,k_i)$ is increasing and differentiable,

$$1 = f[\eta_{\alpha_n}(\theta_i,k_i);\theta_i,k_i]\frac{\partial}{\partial\alpha_n}\eta_{\alpha_n}(\theta_i,k_i), \quad i=1,2.$$

Hence, for $\alpha_n\in]0,1[$,

$$\frac{\frac{\partial}{\partial\alpha_n}[\eta_{\alpha_n}(\theta_1,k_1)]^{k_1}}{\frac{\partial}{\partial\alpha_n}[\eta_{\alpha_n}(\theta_2,k_2)]^{k_1}} = \frac{\Gamma(k_1)\theta_1^{k_1}}{\Gamma(k_2)\theta_2^{k_2}}[\eta_{\alpha_n}(\theta_2,k_2)]^{k_2-k_1} \cdot \exp\left[\frac{1}{\theta_1}\eta_{\alpha_n}(\theta_1,k_1) - \frac{1}{\theta_2}\eta_{\alpha_n}(\theta_2,k_2)\right].$$

Finally, (2) holds from a backward application of L'Hospital's rule.

□

THEOREM 1. Let $\{\alpha_n,\beta_n\}$ be a sequence of positive constants with $\alpha_n\downarrow 0$ and $\beta_n\downarrow 0$ as $n\uparrow\infty$ and $\lim_{n\uparrow\infty}\frac{\ln(\alpha_n)}{\ln(\beta_n)} = a\geq 0$. Then for $k_1,k_2>0$,

$$(4) \qquad \lim_{n\uparrow\infty}\frac{\eta_{1-\alpha_n}(\theta_1,k_1)}{\eta_{1-\beta_n}(\theta_2,k_2)} = a\frac{\theta_1}{\theta_2}.$$

PROOF. Since $\lim_{n\uparrow\infty}\frac{\eta_{1-\alpha_n}(\theta_1,1)}{\eta_{1-\beta_n}(\theta_2,1)} = a\frac{\theta_1}{\theta_2}$, (4) follows from (1).

□

LEMMA 2. Let $\alpha:]0,\infty[\to]0,1[$ and $\beta:]0,\infty[\to]0,1[$ both be differentiable functions

with $\alpha(x)\downarrow 0$ and $\beta(x)\downarrow 0$ as $x\uparrow\infty$ and $\lim_{x\uparrow\infty} \frac{\alpha(x)}{\beta(x)} = a \geq 0$, and let $\alpha_n = \alpha(n)$ and $\beta_n = \beta(n)$ for $n=1,2,\ldots$. Then

$$\lim_{n\uparrow\infty} \frac{\eta_{\alpha_n}(\theta_1,k)}{\eta_{\beta_n}(\theta_2,k)} = a^{1/k}\frac{\theta_1}{\theta_2}. \tag{5}$$

PROOF. Refer to formula (3). $\frac{d}{dx}\alpha(x) = f[\eta_{\alpha(x)}(\theta_1,k);\theta_1,k]\frac{\partial}{\partial x}\eta_{\alpha(x)}(\theta_1,k)$ and $\frac{d\beta(x)}{dx} = f[\eta_{\beta(x)}(\theta_2,k);\theta_2,k]\frac{\partial}{\partial x}\eta_{\beta(x)}(\theta_2,k)$. Hence, $\frac{\frac{d}{dx}\alpha(x)}{\frac{d}{dx}\beta(x)} = \exp[\frac{1}{\theta_2}\eta_{\beta(x)}(\theta_2,k) - \frac{1}{\theta_1}\cdot\eta_{\alpha(x)}(\theta_1,k)]\frac{\theta_2^k}{\theta_1^k}\frac{\frac{\partial}{\partial x}[\eta_{\alpha(x)}(\theta_1,k)]^k}{\frac{\partial}{\partial x}[\eta_{\beta(x)}(\theta_2,k)]^k}$, and (5) holds from a forward and a backward application of L'Hospital's rule.

□

By applying (2) and (5), the following theorem holds and the proof is omitted.

THEOREM 2. Let $\{\alpha_n,\beta_n\}$ be defined as in Lemma 2. Then

$$\lim_{n\uparrow\infty} \frac{\eta_{\alpha_n}(\theta_1,k_1)}{\eta_{\beta_n}(\theta_2,k_2)} = \begin{cases} 0 & 0<k_1<k_2 \\ \infty & 0<k_2<k_1 \\ a^{1/k}\theta_1/\theta_2 & 0<k_2=k_1=k \end{cases}.$$

3. AN EXAMPLE

EXAMPLE. Suppose an electrical company's charges are based on the maximum power demanded over a time period of n days. A company has r meters, and some of r meters will be charged at a discount rate if kept separate, but r meters will be charged at the same rate if pooled. If X_{ij} represents the power demand for the jth meter on the ith day, and $S_i = X_{i1}+\ldots+X_{ir}$, then $S_{(n)} = \max\{S_1,\ldots,S_n\} \leq X_{(n)1}+\ldots+X_{(n)r}$, where $X_{(n)j} = \max\{X_{1j},\ldots,X_{nj}\}$. So the question is when will the reduction in the peak demand obtained by pooling meters offset the added cost incurred by pooling. We need an indication of the relative effect of maximizing a sum compared to sums of maxima. This information should be useful in other types of applications as well. For example, a greater peak load capacity per work station is required if the work stations are kept separate, than if they are pooled.

There are many potential applications where analogous results for minima would also be useful, so both cases are discussed below for the gamma distribution.

The gamma distribution is a flexible two-parameter model, and results for the gamma distribution should provide a guide as to what effects might

generally be expected.

Let $X_{1j},\ldots,X_{nj}$, $j=1,\ldots,r$, be r independent random samples selected from $Gam(\theta,k)$, $X_{(i)j}$= the ith smallest order statistic of $X_{1j},\ldots,X_{nj}$, $\xi_n(k) = \frac{1}{\theta} E[X_{(n)j}]$, and $\delta_n(k) = \frac{1}{\theta}E[X_{(1)j}]$. Then the relative effect on maxima is given by $R_n(rk,k) = \frac{\xi_n(rk)}{r\xi_n(k)}$, and the relative effect on minima is given by $Q_n(rk,k) = \frac{r\delta_n(k)}{\delta_n(rk)}$. The approximations for $R_n(rk,k)$ and $Q_n(rk,k)$ based on the limiting extreme value distribution are given by

$$R_n(rk,k) \approx R'_n(rk,k) = \frac{\gamma\chi^2_{1-\frac{1}{ne}}(2rk) + (1-\gamma)\chi^2_{1-\frac{1}{n}}(2rk)}{r[\gamma\chi^2_{1-\frac{1}{ne}}(2k) + (1-\gamma)\chi^2_{1-\frac{1}{n}}(2k)]},$$

and

$$Q_n(rk,k) \approx Q'_n(rk,k) = \frac{r^2\chi^2_{1/n}(2k)\,\Gamma(\frac{1}{k})}{\chi^2_{1/n}(2rk)\,\Gamma(\frac{1}{rk})}.$$

where $\gamma\approx 0.5772157$ denotes Euler's constant and $\chi^2_p(v)$ denotes the 100pth percentile of the Chi-square distribution with v degrees of freedom.

It can be seen numerically from Ref. [2] that the respective approximations for $R_n(rk,k)$ and $Q_n(rk,k)$ based on $R'_n(rk,k)$ and $Q'_n(rk,k)$ are surprisingly good.

By applying Theorems 1 and 2, the following result holds.

$\lim_{n\to\infty} R'_n(rk,k) = \frac{1}{r}$ and $\lim_{n\to\infty} Q'_n(rk,k) = 0$, which provides information for $R_n(rk,k)$ and $Q_n(rk,k)$ when n is large.

REFERENCES

[1] Pickands, J., Moment Convergence of Sample Extremes, Ann. Math. Statist., 39 (1968), 881-889.

[2] Bain, L. J. and Gan, G., Comparison of Expectations of the Extremes of Sums and the Sum of Extremes from Gamma Distributions, J. Statist. Comput. Simul., 47 (1993), 219-225.

Application Of Extreme-Value Theory To Reliability Physics Of Electronic Parts (On-Orbit Single Event Phenomena)

Goka, T.
National Space Development Agency of Japan, Tokyo, Japan

Some models, for example weakest link model and bundle of fiber model used to be applied to reliability physics in the fields of reliability and destructive engineering. To the test data and the statistical analysis of the distribution of the smallest and the largest values that can be explained with these models, extreme-value theory (particularly doubly exponential distribution) can be applied. The purpose of this paper is to examine the application of extreme-value theory to the on-orbit data on single event phenomena of memory IC under the space radiation environment. The application of extreme-value theory is compared with that of the conventional Poisson distributions to verify the effectiveness of the application of extreme-Value theory (doubly exponential distribution).

1.INTRODUCTION

Since the destruction of material is thought to take place at the weakest point of the component (the weakest link model), the material property has stochastic characteristics. Therefore, to correctly determine the tensile strength and/or the life of a material, the distribution of smallest values is appropriate. On the other hand, the distribution of largest values is significant to the leak current failure of electronic parts or the corrosion pit depth for the rupture strength. The selection of the smallest or the largest values of subject data, in other words, will approach the extreme-value distribution. Therefore analytical methods have been established on the basis of extreme-value theory.

At present, a method based on reliability physics (a method applied to developing high reliability parts that are completely fault-free by thoroughly pursuing the cause of actually generated faults, detecting and correcting latent faults through accelerated test) is drawing attention as the ultimate decisive factor for reliability assurance. Under the circumstances, there are high possibilities for utilizing extreme-value distributions, particularly the doubly exponential distribution, in the field of reliability data analysis of electronic parts.

Single event phenomena are the well known interactions between high-energy particles in the space environment and electronic devices on spacecraft. These phenomena are caused by high energy proton or heavier particles such as helium, carbon, nitrogen, oxygen, iron and so forth in galactic cosmic rays, the trapped Van Allen belt particles and solar flares. These phenomena can be classified into Single Event Upset (SEU) and Single Event Latchup (SEL).

SEU is a reversible soft error that the information (digit 1 or 0) which are maintained in the memory or the microprocessor unit (MPU) of the spacecraft are upset (1→0, 0→1) by the particles in space. Especially, the upset caused by protons is termed as proton upset, and that occurs even in the low-altitude orbit owing to the protons trapped by the magnetic field of the earth. These protons are most abundant above the south-Atlantic ocean, so called South-Atlantic Anomaly(SAA).

SEL is an irreversible hard error which is caused by the high-energy particles to the electronic parts of the CMOS technology.

2 MEASUREMENT OF SEU AND SEL BY ETS-V

2.1 RAM SOFT-ERROR MONITOR EXPERIMENT

Engineering Test Satellite-V (ETS-V) has been launched by NASDA on August 27,1987 and has been put into geo-stationary orbit at 150° east longitude. This spacecraft has a Technical Data Acquisition Equipment (TEDA) aiming at obtaining technical data which is necessary to develop spacecraft. TEDA includes a RAM Soft-error Monitor (RSM) that makes a measurement of the SEU or SEL occurring at eight 64-kbit CMOS static RAM devices (NEC, μ PD4464D-20). The thickness of the shield of these devices is estimated to be 21.5mm in Al. This monitor was developed by NASDA in collaboration with NTT, Japan.[1]

2.2 SEU AND SEL DATA ON ETS-V

(1) IN-ORBIT MONITORING RESULTS

Figure 1 shows the SEL data (sum of 8 devices) acquired by the ETS-V TEDA RSM. The abscissa and the ordinate axis in this figure indicate passing days and the number of SEL that occurred in a week respectively. The period of data acquisition is about 4 years from November 22,1987 to Jan.31,1992. The solar activity became intensive since September, 1989 and the 4B-class solar flares were observed on September 29 and October 1989. The number of SEL drastically increased during these solar Flares. The number of SEU (sum of 8 devices) measured by RSM each week is also plotted in Figure 2. From this figure one can see that the number of SEU is less than that of SEL, and it also increased remarkably when the solar flare occurred.[2]

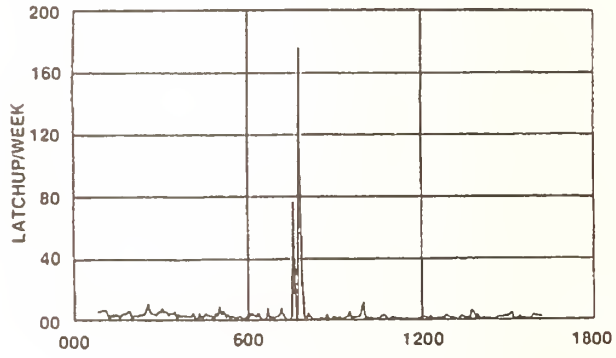


Figure 1 Measured Latchup rate as a Function of Time(ETS-V)

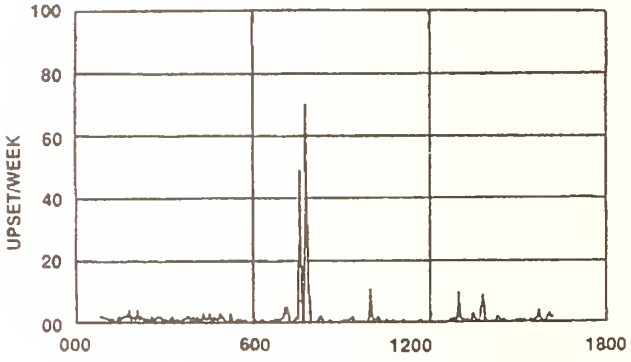
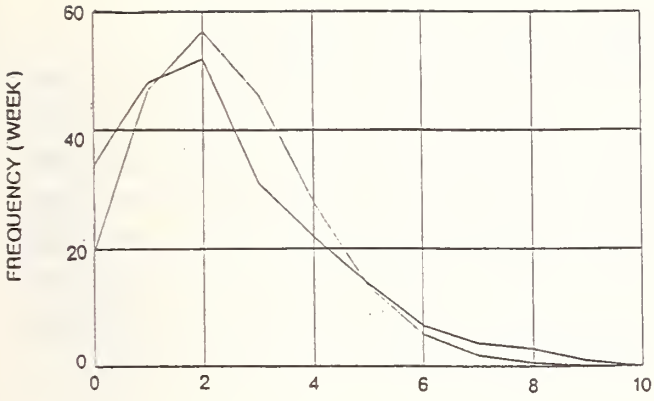


Figure 2 Measured Upset rate as a Function of Time(ETS-V)

2.3 POISSON DISTRIBUTION

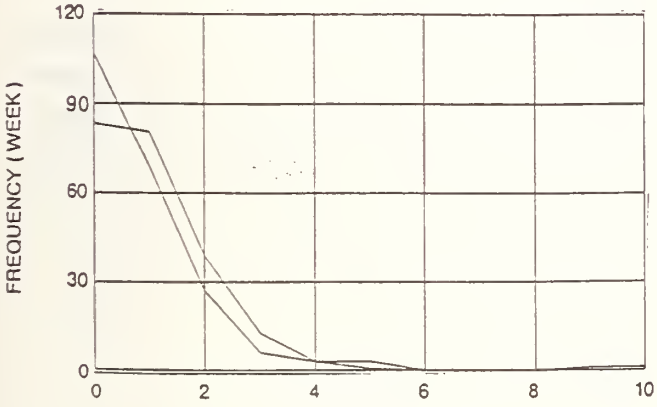
The distribution of the frequency versus the number of SEL and SEU in a week are plotted in Figure 3 and Figure 4 respectively.

The average number of SEL and SEU in a week are ML=2.4(/week) and MU=0.76(/week) respectively. If the phenomena are perfectly random and uniform process, these distributions will agree to the Poisson distribution. In Figure 3 and Figure 4 dotted lines are the results of substitution of these values in the formula of the Poisson distribution,



NUMBER OF LATCHUP IN A WEEK

Figure 3 The Distribution of the Frequency Versus the number of SEL(ETS-V)



NUMBER OF UPSET IN A WEEK

Figure 4 The Distribution of the Frequency Versus the number of SEU(ETS-V)

$$P(K) = \frac{M^K}{K!} e^{-M} \qquad (1)$$

where,

K:The number of SEL or SEU

P(K):The probability of observation of K SELs or K SEUs

M:The average number of SEL or SEU (ML or MU, namely)

These figures disclose that the observed data do not fit the Poisson distribution very well. This is owing to a change in the tendency of data before and after the solar flare which occurred in the latter half of the total observation period.

Figure 5 and Figure 6 indicate the distribution in the first (September 1 , 1987 to June 11 , 1989 : Solar Minimum) and latter harf (June 11 , 1989 to January 31 , 1992 ; Solar Maximum) of the total period as to SEL and SEU respectively.

Figure 5 and Figure 6 indicate the distribution of the frequency in the first and latter half of the total period as to number of SEL and SEU respectively.There are good agreements between the data and the Poisson distributions in these figures,and the number of SEL and SEU decreased apparently after the solar flare.

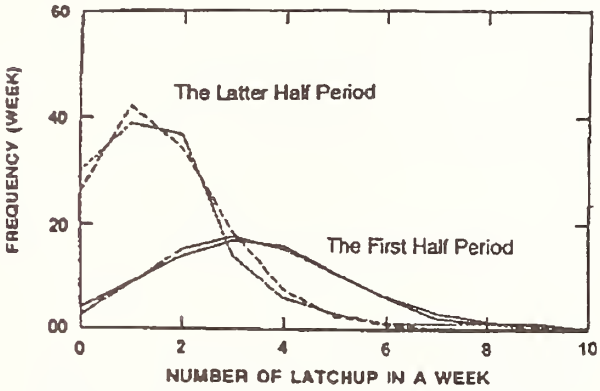


Figure 5 The Distribution of the Frequency Versus the number of SEL (ETS-V)

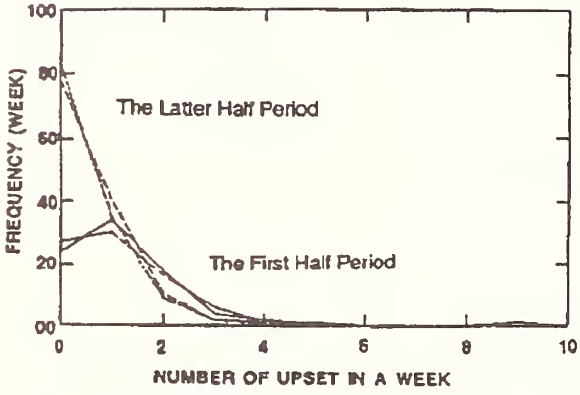


Figure 6 The Distribution of the Frequency Versus the number of SEU (ETS-V)

The data during the solar flares,namely August 12,September 29,October 19,1989 and May 21-25,1990 are removed from these figures.It is apparent that the Poisson distribution is inapplicable to the data when the solar flare occurred.

According to Figure 5-Figure 6,the number of SEL was about 3 times as much as that of SEU.Generally,when SEL and SEU occured simultaneously,only SEL will be detected by RSM because SEL is a hard-error.This means that the condition for the detection of SEU is that SEL should not

occur at the same time. Namely,
The condition for the detection of SEL:

$$LL<L$$

The condition for the detection of SEU:

$$LU<L<LL$$

where L is the LET (Linear Energy Transfer) of the incident particle, LL is the threshold LET for SEL and LU is the threshold LET for SEU of the RAM devices. It is expected that SEU would be comparatively hard to observe when the value of LU is close to that of LL.

2.4 DOUBLY EXPONENTIAL DISTRIBUTION (EXTREME-VALUE THEORY)

Extreme-value theory is introduced to analyze the maximum single event rate data inclusive of data during solar flares. Suppose we have a random sample from a probability density function which has a tail that decreases as exponential type (Poisson, normal, log normal, logistic, etc.) and we are interested in the upper tail of the probability density function (largest extreme values). The distribution function and probability density function of the so-called Type-I asymptotic distribution of largest values (double exponential distribution) are, respectively,

$$F(y)=\exp(-\exp(-y)), \qquad (2)$$

$$f(y)=\exp(-y-\exp(-y)). \qquad (3)$$

Where $y=(x-\lambda)/\alpha$, $-\infty<x<\infty$, $-\infty<\lambda<\infty$, $\alpha>0$, and location parameter λ and scale parameter α are unknown. Taking the natural logarithm of the distribution function twice,
we have

$$-\ln(-\ln F(x))=(x-\lambda)/\alpha \qquad (4)$$

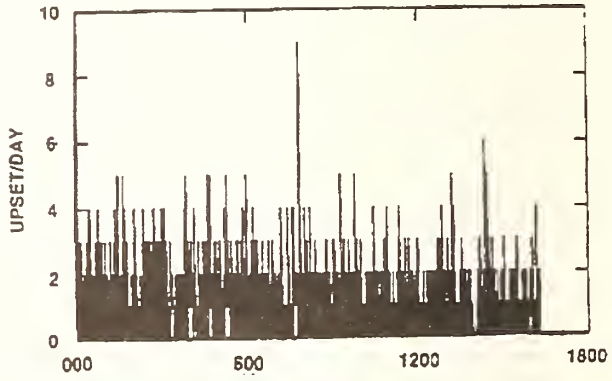
Which stands for the equation of a straight line on the extreme-value probability paper. Assuming that single event phenomena comply with the Poisson distribution, the distribution of the maximum values of the number of single events will agree with the doubly exponential distribution. The cumulative probability of the maximum single event rates (events/week) in a month are plotted in the left side of Figure 7 (extreme-value probability paper).

For this analysis, the data during the solar flare namely August 12,September 29,October 19,1989 and May 21-25,1990 are included to these figures. Figure 7 indicate the distribution in the first half (1.7 years) and the latter half (2.3 years) of the total observation period. The distributions become linear for the first half period, while the slope of the line changes for the latter half period. Apparently this discrepancy depends on the effect of the solar flares.

3 SEU DATA ON MARINE OBSERVATION SATELLITE-1(MOS-1)

3.1 SEU DATA ON MARINE OBSERVATION SATELLITE-1(MOS-1)

Marine Observation Satellite-1(MOS-1) was launched by NASDA on February 19,1987. Figure 8 shows the number of SEU that occured in a day at the stored command memory used in the command decoder of MOS-1. The period of data acquisition is about 4 year and a half from September 1,1987



DAY(The day of the origin -September 1,1987)

Figure 8 Measured Upset rate as a Function of

to January 31,1992.
This memory consists of three bi-polar Static Random Access Momory(SRAM) devices(93419, 64 9bit SRAM, Fairchild). The shield thickness is assumed to be about 2.7mm in Al. Though MOS-1 is a Low Earth Orbit(LEO, 909km) altitude and 99 degree inclination), one can see apparently from this figure that the number of SEU (9 upsets a day) increased by the effect of solar flare on October 19, 1989. The upset bits data are reported each second to the ground. From this information the place where the upset has occurred can be determined. Figure 9 is the upset map, in

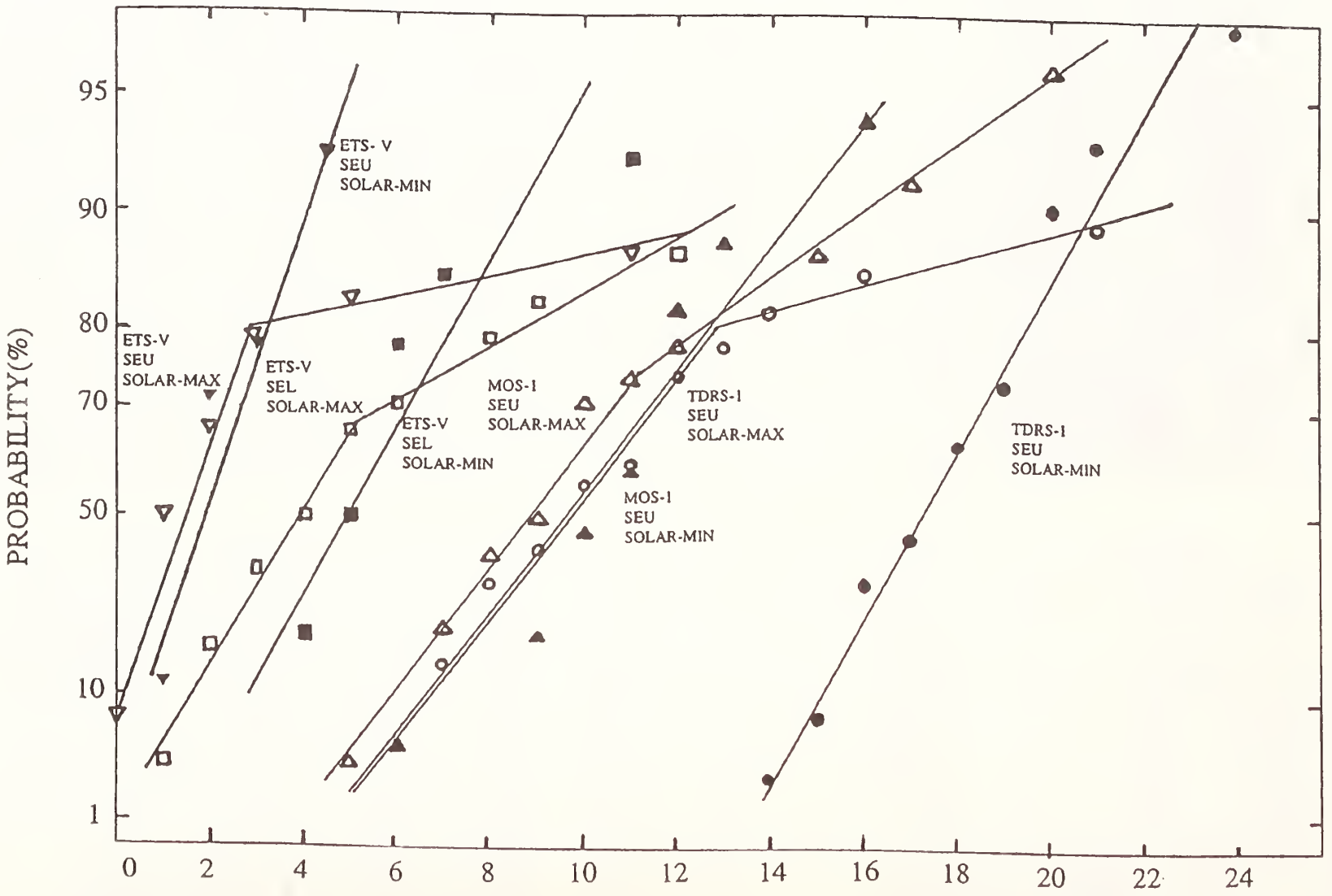


Figure 7 Doubly Exponential Plots for Mos-1(SEU),ETS-V(SEL,SEU),TDRS-1(SEU)

which the places where upset occurred from Feburary 1987 to August 1988 are indicated. This figure shows that the upset places concentrated on the so called South-Atlantic Anomaly.

Figure 10 shows geomagnetic field contour map on the MOS-I orbit spherical surface using IGRF85. One can see from Figure 9 and Figure 10 MOS-1 SEU occurred mainly the region which corresponds to the area less than 20000 nano Tesla (geomagnetic total intensity) contour. A few upsets occurred at both polar regions which correspond to about 70° north latitude and 70° south latitude. Ground testing of the SRAM was carried out using high energy proton at the cyclotron facility. Comparison of predicted upset rate using NASA AP-8 model and orbital data was carried out. There is quite good correlation between AP-8 prediction and orbital data.[3]

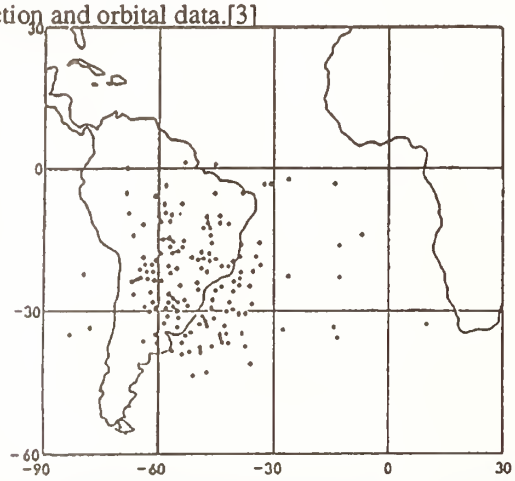


Figure 9 Upset Map For Three 93419's On MOS-1 Spacecraft

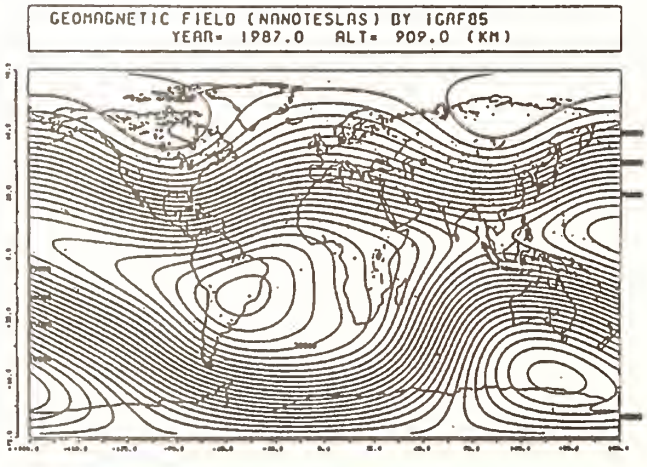


Figure 10 Geomagnetic Field Contour Map On MOS-1 Orbit

3. 2 STATISTICAL ANALYSIS OF MOS-1 SEU DATA

(1) POISSON DISTRIBUTION

The distribution of the frequency versus the number of SEU in a day during the total period is plotted in Figure 11. The distribution of the frequency versus the number of SEU in a day during the first half period of the total period (September 1,1987 to June 11,1989:Solar Minimum) and during the latter half period (June 11,1989 to January 31, 1992:Solar Maximum) are plotted in Figure 12 and Figure 13 respectively. These average number of SEU in a day are MU_1=1.13, MU_2=0.82 respectively.

In Figure 12 and Figure 13 dotted lines are the results of substitution of these values in the formula of the Poisson distribution. These are good agreements between the data and the Poisson distributions in these figures. The number of SEU decreased slightly in the solar maximum period even

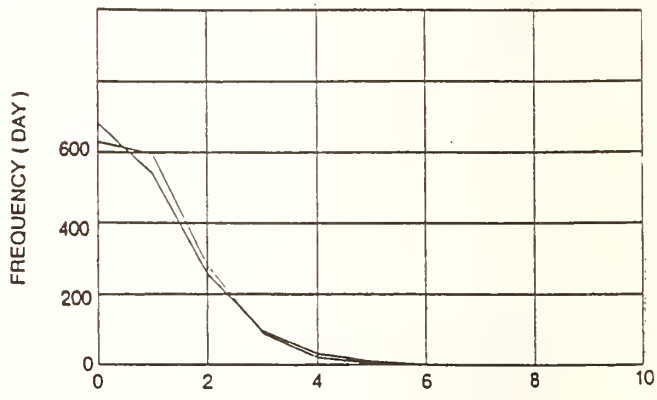


Fig.11.The Distribution of the Frequency Versus the Number of SEU (Total period)

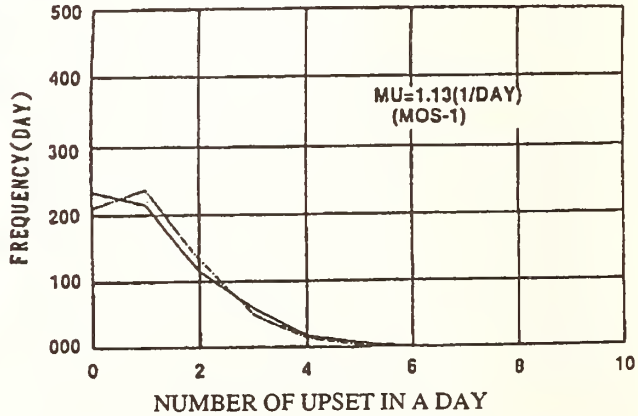


Figure 12 The Distribution of the Frequency Versus the Number of SEU (First half period)

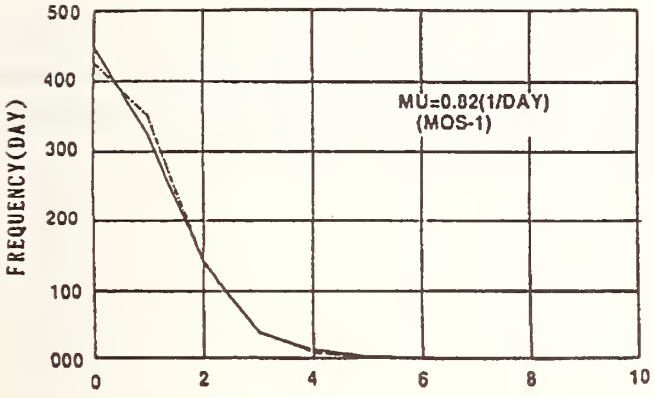


Figure 13 The Distribution of the Frequency Versus the Number of SEU (Latter half period)

for the MOS-1 spacecraft.

(2) DOUBLY EXPONENTIAL DISTRIBUTION (EXTREME - VALUE THEORY)

The cumulative probability of the maximum single event rates (events/week) in a month are plotted in the center part of Figure 7 (extreme - value probability paper).

3.3 SEU DATA ON THE TRACKING AND DATA RELAY SATELLITE(TDRS-1) [4]

The TDRS-1 was designed by NASA, and was launched from the space shuttle, *Challenger* in April 1983, and was put into a geostationary orbit in July 1983. The SEUs were observed in the Attitude Control System (ACS). The ACS contains four pages of RAM, 256 bytes per page. Each page consists of two static bi-polar Fairchild 93L422(256k X 4bit)RAM chips.

The TDRS-1 weekly SEU count shown in Figure 14 begins in April 1986. Some SEUs go unobserved. The spikes in August, September and October 1989 are responses to solar flares. The off-scale responses in September and October are for weekly SEU total of 88 and 157 respectively. [4]

The cumulative probability of the maximum single event rates (event/week) in a month of the TDRS-1 SEU data are plotted in the right side of Figure 7 (extreme-value probability paper) even for MOS-1 SEU, ETS-V SEU and SEL. Figure 7 indicates the distribution in the solar Minimum and the Solar-Maximum observation period. Solar Minimum corresponds to the period of MOS-1; September 1987-November 1988, ETS-V; November 1987-November 1988, TDRS-1; May 1986- November 1988. Solar Maximum corresponds to the period; November 1988- January 1991 in all satellites. The distributions become linear for the Solar-Minimum period, while the slope of the line changes for the Solar-Maximum period. Apparently this discrepancy depends on the effect of the

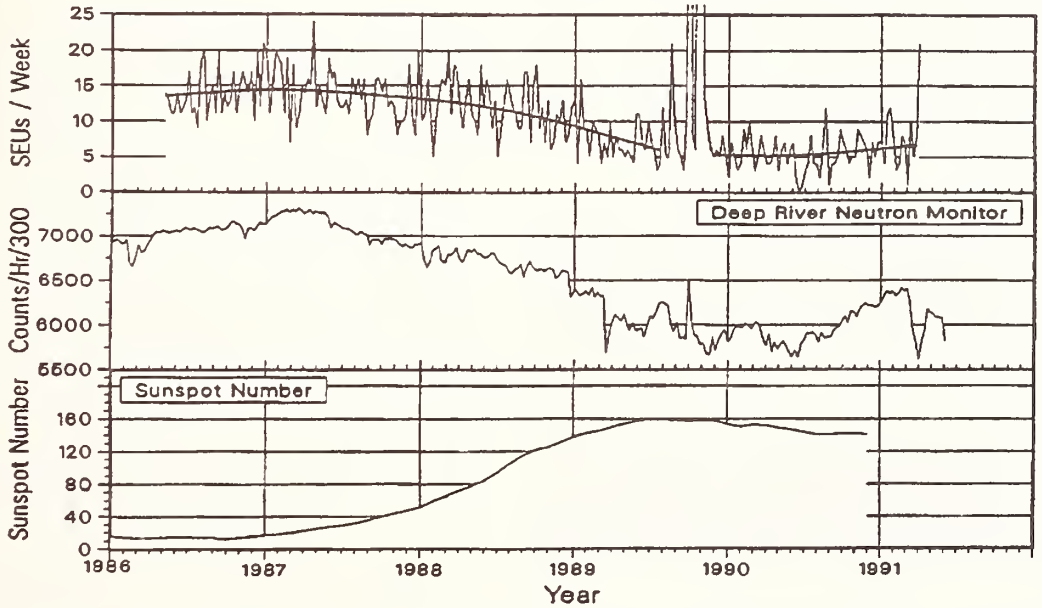


Figure 14 The SEUs showing that the envelope of the TDRS-1 SEUs clearly follows the modulation of the galactic cosmic rays. The smoothed line through the TDRS data was created by using a cubic spline function . The spikes in September and October 1989 reach 88 and 157 SEUs per week respectively. (Wilkinson, 1991)

solar flares. From all of ther analysis a decrease of the number of single events could be found during the SolarMaximum. This phenomena are due to the screen effect, that is to say the solar flare particles screen heavy ions from the Galactic cosmic ray.

4. CONCLUSIONS

Application of extreme-value theory (doubly expotential distribution of the largest values) to the observation data of Single event phenomena about 4 years by the geostationary satellite and the medium altitude satellite are examined.
In comparison with the Poisson distribution, the doubly exponential distribution has the advantage to be able to analyze the data include the big solar flare and to enable us to discriminate clearly from the effect of solar flares.

A judgement from above point comes to the conclusion that the doubly exponential distribution is preferable to the Poisson distribution.

5. ACKNOWLEDGMENTS

The author wish to acknowledge Dr.Shigeo Kase, University of the Air for the helpful discussions concerning extreme value theory, and the support of Mr.Kazuhiro Kiyono for the analyzing of the data.
The work described in this paper was carried out at the Tsukuba Space Center, NASDA.

6. REFERENCES

[1] Shiono N, Sakagawa Y, Sekiguchi M, Sato K, Sugai I, Hattori T, Hirao Y, Single Event Effects in High Density CMOS SRAMs, IEEE Trans. Nucl. Sci. NS-33PP1632-1636, 1986

[2] Goka T, Kuboyama S, Shimano Y, Kawanishi T, The On-Orbit Measurements of Single Event Phenomena by ETS-V Spacecraft, IEEE Trans. Nucl. Sci. NS-38 PP1693-1699, 1991

[3] Shimano Y, Goka T, Kuboyama S, Kawai K, Kanai T, Takami Y, The Measurements and the Prediction of Proton Upset, IEEE Trans. Nucl. Sci. NS-36 PP2344-2348, 1989

[4] wilkinson D, Daughtridge S,Stone J, Sauer H, Darling P, TDRS-1 Single Event Upset and the Effect of the Space Environment, IEEE Trans. Nucl. Sci. Vol 38 No. 6, PP 1708 - 1712 ; 1991

Certain Identities In Expectations Of Functions Of Order Statistics And Characterization Of Distributions

Govindarajulu, Z.
University of Kentucky, Lexington, KY

A method of generating identities in expectations of functions of order statistics defined on the positive real axis is obtained. These identities are specialized to the exponential, the folded normal, the folded logistic and the uniform distributions. The specialized identities are used to characterize the exponential, the generalized truncated normal, the folded logistic and the uniform distributions.

Key words order statistics, identities, characterization of distributions.
AMS classification 62G30, 62E10.

1 Certain Identities in Expectations of Functions of order Statistics.

Let X be a random variable having distribution function $F(x)$ and probability density function $f(x)$ where the latter is zero for $x < 0$. Let g be a continuous differentiable function such that differentiation of $g(x)$ with respect ot its argument and expectation of $g(x)$ with respect to an absolutely continuous distribution are interchangable. Furthermore let $X_{1N} \leq X_{2N} \leq \ldots \leq X_{NN}$ denote the order statistics in a random sample of size N drawn from $f(x)$. Then we have the following results.

Proposition 1. If f is differentiable, then for $1 < i \leq N$ and $N = 2, 3, \ldots$ we have

$$E[g'(X_{iN})] = -Nf(0)Eg(X_{i-1,N-1}) - \sum_{k=1}^{N} E[g(X_{iN})f'(X_{kN})/f(X_{kN})] .$$

PROOF. For $1 < i \leq N$

$$\begin{aligned} E[g(X_{iN}+t)] &= N! \int\cdots\int_{0<x_1<\ldots<x_N<\infty} g(x_i+t)\pi_{j=1}^{N} f(x_j)dx_j \\ &= N! \int\cdots\int_{t<y_1<\ldots<y_N<\infty} g(y_i)\pi_{j=1}^{N} f(y_j - t)dy_j \\ &= N! \int_t^{\infty} f(y_1 - t)h(y_1, t)dy_1 \end{aligned}$$

where

$$h(y_1,t) = \int\cdots\int_{y_1<y_2<\leq<y_N<\infty} g(y_i)\pi_{j=2}^N f(y_j-t)dy_j\,.$$

Then

$$E[g'(X_{iN})] = \frac{\partial}{\partial t}E[g(X_{iN}+t)]|_{t=0} = -N!f(0)h(0,0) - N!\int_0^\infty f'(y_1)h(y_1,0)dy_1 + N!\int_0^\infty f(y_1)h'(y_1,0)dy_1 \quad (1.1)\ (1.2)$$

where

$$h(0,0) = \int\cdots\int_{0<y_2<y_N<\infty} g(y_i)\pi_{j=2}^N f(y_j)dy_j = \frac{1}{(N-1)!}Eg(X_{i-1,N-1}) \quad (1.3)$$

$$N!\int_0^\infty f'(y_1)h(y_1,0)dy_1 = N!\int_\infty^\infty f'(y_1)\left[\int\cdots\int_{y_1<y_2<\ldots<y_N<\infty} g(y_i)\pi_{j=2}^N f(y_j)dy_j\right] \quad (1.4)$$

$$= E\left[g(X_{iN})\frac{f'(X_{1N})}{f(X_{1N})}\right]$$

$$N!\int_\sigma^\infty f(y_1)h'(y_1,0)dy_1 = N!\int_0^\infty f(y_1)\left[\int\cdots\int_{y_1<y_2<\ldots<y_N<\infty} g(y_i)\left[-\sum_{k=2}^N\frac{f'(y_k)}{f(y_k)}\right]\pi_{j=2}^N f(y_j)dy_j\right]dy_j$$

$$= -\sum_{k=2}^N E\left[g(X_{iN})\frac{f'(X_{kN})}{f(X_{kN})}\right]\,. \quad (1.5)$$

Using (1.2) (1.3) and (1.4) in (1.1) we obtain the desired result.

Remark 1. This method of deriving identities in expectations of functions of order statistics when the support of the distribution is the real line has been used by Seal [10] and Govindarajulu ([5], p. 638).

Corollary 1.1 Let $g(x) = x$ in Proposition 1. Then we obtain

$$1 = -Nf(0)E(X_{i-1,N-1}) - \sum_{k=1}^N E\left\{X_{iN}\frac{f'(X_{kN})}{f(X_{kN})}\right\} \quad (1.6)$$

Special cases. Now if $f(x) = \exp(-x)$, for $x > 0$. Then (1.5) becomes

$$1 = N\{E(X_{iN}) - E(X_{i-1,N-1})\}\,. \quad (1.7)$$

If $f(x) = (2/\pi)^{1/2}\exp(-x^2/2)$ for $0 < x < \infty$, then for $1 < i \leq N$

$$1 = -N(2/\pi)^{1/2}E(X_{i-1,N-1}) + \sum_{k=1}^N E(X_{iN}X_{kN})\,. \quad (1.8)$$

If $f(x) = 2e^{-x}/(1+e^{-x})^2$ for $x > 0$, then for $1 < i \leq N$

$$1 = -NE(X_{i-1,N-1}) + \sum_{k=1}^N E\{X_{iN}F(X_{kN})\}\,. \quad (1.9)$$

Using analogous methods one can prove the following proposition.

Proposition 2. If f is differentiable then we have

$$E\{g'(X_{1N})\} = -Ng(0)f(0) - \sum_{k=1}^N E\{g(X_{1N})\frac{f'(X_{kN})}{f(X_{kN})}\}\,. \quad (1.10)$$

Remark 2. If $X_{0,N-1}$ is taken to be zero, then Proposition 2 can be included in Proposition 1.

Corollary 2.1 Let $g(x) = x$. Then (1.9) becomes

$$1 = -\sum_{k=1}^N\left(X_{1,N}\frac{f'(X_{k,N})}{f(X_{k,N})}\right)\,. \quad (1.11)$$

Special cases.

(a) Let $f(x) = e^{-x}$ for $x \geq 0$ and zero elsewhere. Then (1.10) becomes

$$1/N = EX_{1,N}\,. \quad (1.12)$$

(b) Let $f(x) = (2/\pi)^{1/2}\exp(-x^2/2)$ for $x \geq 0$. Then (1.10) takes the form of

$$1 = \sum_{k=1}^N E(X_{1,N}X_{k,N})\,. \quad (1.13)$$

(c) Let $f(x) = 2e^{-x}/(1+e^{-x})^2$ for $x \geq 0$.

Then since $-f'(x)/f(x) = F(x) = \{2/(1+e^{-x})\} - 1$, (1.10) takes the form of

$$1 = \sum E\{X_{1N}F(X_{k,N})\}. \qquad (1.14)$$

Remark 3. Since $f'(x) \equiv 0$ the above propositions are not applicable to the uniform density on $(0,1)$. In the following we obtain the specific results for the uniform density.

Proposition 3. If f is the standard uniform density, for $1 < i < N$, and $N = 2, 3 \ldots$

$$E\{g'(X_{iN})\} = N\{Eg(X_{i,N-1}) - Eg(X_{i-1,N-1})\} \qquad (1.15)$$

PROOF. For $1 < i < N$, consider

$$\begin{aligned} E[g(X_{iN}+t)] &= N! \int\cdots\int_{0<x_1<\ldots<x_N<1} g(x_i+t)dx_1\ldots dx_N \\ &= N! \int\cdots\int_{t\le y_1<\ldots<y_N\le 1+t} g(y_i)dy_1\ldots dy_N \\ &= \frac{N!}{(i-1)!(N-i)!}\int_t^{1+t} (y-t)^{i-1}(1+t-y)^{N-i}g(y)dy. \end{aligned}$$

Thus

$$E[g'(X_{iN}+t)] = \frac{N!}{(i-1)(N-i)'}$$
$$\int_t^{1+t}\left[(-1)(i-1)(y-t)^{i-2}(1+t-y)^{N-i}\right.$$
$$\left.+(N-i)(y-t)^{i-1}(1+t-y)^{N-1-i}\right]g(y)dy.$$

Hence

$$Eg'(X_{iN}) = \frac{N!}{(i-1)!(N-i)!}$$
$$\left[\int_0^1 g(y)y^{i-2}(1-y)^{N-1-i}\right.$$
$$\left.\{(N-i)y-(i-1)(1-y)\}dy\right]$$
$$= \frac{N!}{(i-1)!(N-i-1)!}$$
$$\int_0^1 g(y)y^{i-1}(1-y)^{N-1-i}dy$$
$$-\frac{N!}{(i-2)!(N-i)!}\int_0^1 g(y)y^{i-2}(1-y)^{N-i}dy.$$

Thus

$$Eg'(X_{iN}) = N\left[Eg(X_{i,N-1}) - Eg(X_{i-1,N-1})\right]$$
$$1 < i < N, N = 2, 3, \ldots$$

Special Case 1 Let $g(x) = x$ in (1.14) and obtain

$$1 = N[E(X_{i,N-1}) - E(X_{i-1,N-1})]. \qquad (1.16)$$

By proceeding analogously we obtain Propositions 4 and 5.

Proposition 4. If f is the standard uniform density, then for $N = 2, 3, \ldots$

$$E(g'(X_{1N})) = NEg(X_{1,N-1}) - Ng(0). \qquad (1.17)$$

Proposition 5. If f is the uniform density, then for $N = 2, 3, \ldots$

$$Eg'(X_{NN}) = Ng(1) - NE\{g(X_{N-1,N-1})\} \qquad (1.18)$$

Special cases. Setting $g(x) = x$ in (1.16) and (1.17) we have

$$1 = E(X_{1,N-1}), N = 2, 3, \ldots \qquad (1.19)$$

and

$$\frac{N-1}{N} = E(X_{N-1,N-1}), N = 2, 3, \ldots. \qquad (1.20)$$

Remark 4. If $X_{0,N-1}$ is taken to be zero, then Proposition 4 can be included in Proposition 3. If $X_{N,N-1}$ is considered to be unity, then Propostion 5 can be included in Proposition 3.

2 Characterization of the Exponential Distribution

In this section we characterize the exponential distribution using identities (1.6) and (1.11). Toward, this we need the following notation. Let

$$H(y) = \inf\{x : F(x) \ge u\}, 0 < u < 1. \qquad (2.1)$$

Take F to be right continuous. Then for $0 < u < 1$, we have

$$H(u) \le x \Longleftrightarrow u \le F(x).$$

Let $U_1, \ldots U_N$ be a random sample from the standard uniform distribution. Then the distribution

of $H(U_1), H(U_2), \ldots, H(U_N)$ is the same as that of $X_1, X_2, \ldots, X_N$. Also $H(U_{1N}) = \min_{1 \le k \le N} H(U_k)$ has the same distribution as X_{1N} etc. Throughout, we assume that F is absolutely continuous. That is, $H'(u)$ exists almost everywhere for $0 < u < 1$. Then we have the following propositions.

Proposition 6. Let $F(A) = 0$. Then for $2 \le i \le N$ and $N = 2, 3, \ldots, E(X_{iN}) - E(X_{i-1,N-1}) = 1/N$ if and only if $F(x) = 1 - \exp(-(x - A)), x > A$.

PROOF. From the proof of Proposition 1, we can write

$$E(X_{iN}) = \frac{N!}{(i-1)!(N-i)!}$$

$$\int_0^1 H(u) u^{i-1} (1-u)^{N-i} du .$$

Then writing $N = (N - i + 1) + (i - 1)$ we have

$$E(X_{i,N}) - E(X_{i-1,N-1}) = \frac{(N-1)!}{(i-2)!(N-i)!}$$

$$\int_0^1 H(u) u^{i-2} (u-1)(1-u)^{N-i} du$$

$$+ \frac{(N-1)!(N-i+1)}{(i-1)!(N-i)!}$$

$$\int_0^1 H(u) u^{i-1} (1-u)^{N-i} du .$$

Performing integrating by parts in the first integral by writing $u^{i-2}du$ as $d(u^{i-1}/(i-1))$ we obtain (after cancelling out terms and noting that $H(0) = A$)

$$E(X_{i,N}) - E(X_{i-1,N-1}) = \frac{(N-1)!}{(i-1)!(N-i)!} \int_0^1$$

$$H'(u) u^{i-1} (1-u)^{N-i+1} du .$$

Now writing

$$\frac{1}{N} = \frac{1}{N} \frac{N!}{(i-1)!(N-i)!} \int_0^1 u^{i-1} (1-u)^{N-i} du$$

we have

$$E(X_{i,N}) - E(X_{i-1,N-1}) = 1/N \text{ for } 2 \le 1 \le N \text{ and}$$

$$N = 2, 3, \ldots$$

imply that

$$\frac{(N-1)!}{(i-1)!(N-i)!} \int_0^1 u^{i-1} (1-u)^{N-i} \{H'(u)(1-u) - 1\} d$$

$$u = 0 \text{ for } 2 \le i \le N$$

and $N = 2, 3, \ldots$. Now the only continuous function which is orthogonal to $u^{i-1}(1-u)^{N-i}$, a linear combination of $u^0, u, \ldots u^{N-1}$ for $N = 2, 3, \ldots$ is the zero function itself.

Hence $H'(u)(1-u) - 1 = 0$ for almost all u in $(0, 1)$. Integrating on both sides we obtain $H(u) + c = -\ln(1-u)$.

Now $H(0) = A$ implies that $c = -A$. Also $H'(u)$ exists for all u in $(0, 1)$.

Now it follows that $1 - F(x) = \exp(-(x - A)), x \ge A$.

Proceeding in an analogous manner one can prove the following.

Proposition 7. Let $F(A) = 0$. Then for $N = 1, 2, \ldots E(X_{1N}) = A + 1/N$ if and only if $F(x) = 1 - \exp(-(x - A)), x \ge A$.

Remark 5. By defining $X_{0,N-1} = A$, Proposition 7 can be included in Proposition 6.

3 Characterization of the Generalized Truncated Normal Distributions

In this section we characterize the folded normal and the generalized truncated normal distributions, using identities (1.7) and (1.12).

Proposition 8. If $F(0) = 0$, then for $N = 2, 3, \ldots, \sum_{k=1}^{N} E(X_{1N} X_{kN}) = 1$ if and only if $F(x) = 2\Phi(x) - 1$ for $0 < x < \infty$.

Proof. See Theorem 3.2 of Govindarajulu ([6], p. 1013) or Theorem 6 of Lin ([9], p. 403).

Proposition 9. If $F(A) = 0$, then for $2 \le i \le N - 1$ and $N = 2, 3, \ldots$

$$\sum_{k=1}^{N} E(X_{iN} X_{kN}) = 1 + N(EX) E(X_{i-1,N-1})$$

if and only if

$$F(x) = \frac{\Phi(x) - \Phi(A)}{1 - \Phi(A)} \text{ for } A < x < \infty$$

where Φ denotes the standard normal distribution function.

PROOF. One can write

$$\sum_{j=1}^{N} E(X_{iN} X_{jN}) = \sum_{j=1}^{i-1}$$

$$+E(X_{iN}^2)+\sum_{j=i+1}^{N}E(X_{iN}X_{jN})$$
$$=\frac{N!}{(i-2)!(N-i)!}\iint_{z\le w} zwF^{i-2}(w)[1-F(w)]^{N-i}dF(z)dF(w)$$
$$+\frac{N!}{(i-1)!(N-i)!}\int_0^\infty z^2 F^{i-1}(z)[1-F(z)]^{N-i}dF(z)$$
$$+\frac{N!}{(i-1)!(N-i-1)!}\iint_{z\le w} zwF^{i-1}(z)[1-F(z)]^{N-i-1}dF(t)dF(w).$$

After performing integration by parts with respect to w once in the first double integral we obtain

$$\sum_{j=1}^{N}E(X_{iN}X_{jN})=\frac{-N!}{(i-1)!(N-i)!}\iint_{z\le w} z F^{i-1}(w)[1-F(w)]^{N-i}dF(z)dw$$
$$+\frac{N!}{(i-1)!(N-i-1)!}\iint_{z\le w} zw\left\{F^{i-1}(w)[1-F(w)]^{N-i}\right.$$
$$\left.+F^{i-1}(z)[1-F(z)]^{N-i-1}\right\}dF(z)dF(w)$$
$$=\frac{-N!}{(i-1)!(N-i)!}\iint_{z\le w} zF^{i-1}(w)[1-F(w)]^{N-i}dF(z)dw$$
$$+\left(\int_A^\infty z dF(z)\right)\left(\frac{N!}{(i-1)!(N-i-1)!}\int_0^\infty wF^{i-1}(w)[1-F(w)]^{N-i-1}dF(w)\right).$$

Now when f is the truncated normal density, the right hand side simplifies to

$$1-f(A)N\{-\mu_{i-1,N-1}+\mu_{i,N-1}\}+NE(X)\mu_{i,N-1}$$
$$=1+NE(X)\mu_{i-1,N-1}(\text{since } f(A)=EX).$$

On the other hand $\sum_{j=1}^{N}E(X_{iN}X_{jN})=1+NE(X)\mu_{i-1,N-1}$ implies that

$$\frac{-N!}{(i-1)!(N-i)!}\iint_{0<u<v<1}H(u)H'(v)v^{i-1}(1-v)^{N-i}dudv$$
$$+NE(X)\mu_{i,N-1}=1+NE(X)\mu_{i-1,N-1}. \qquad (3.1)$$

Since

$$\mu_{i,N-1}=\frac{(N-1)!}{(i-1)!(N-1)!}\int_0^1 H(v)v^{j-1}(1-v)^{N-i-1}dv$$

and

$$\mu_{i-1,N-1}=\frac{(N-1)!}{(i-2)!(N-i)!}\int_0^1 H(v)v^{i-2}(1-v)^{N-i}dv=\frac{(N-1)!}{(N-i)!(i-1)!}\int_0^1 H(v)(1-v)^{N-i}d(v^{i-1})$$
$$=\frac{-(N-1)!}{(N-i)!(i-1)!}\int_0^1 H'(v)v^{i-1}(1-v)^{N-i}dv+\frac{(N-1)!}{(i-1)!(N-i-1)!}\int_0^1 H(v)v^{i-1}(1-v)^{N-i-1}dv.$$

So

$$N(\mu_{i,N-1}-\mu_{i-1,N-1})=\frac{N!}{(i-1)!(N-i)!}\int_0^1 H'(v)v^{i-1}(1-v)^{N-i}dv$$

Hence (3.1) can be written as

$$\frac{N!}{(i-1)!(N-i)!}\left[\int_0^1 v^{i-1}(1-v)^{N-i}[-H'(v)\int_0^v H(u)du]+(EX)\int_0^1 H'(v)v^{i-1}(1-v)^{N-i}dv-\int_0^1 v^{i-1}(1-v)^{N-i}dv\right]$$
$$=0, i=2,\dots,N-1, N=2,3\dots,$$

That is, for all most all v in $(0,1)$

$$-H'(v)\int_0^v H(u)du+H'(v)EX-1=0;$$

ie

$$H'(v)\int_v^1 H(u)du=1 \text{ since } EX=\int_0^1 H(u)du. \qquad (3.2)$$

Now, using the arguments in Govindarajulu ([6], p.1012) we can establish that $F(x) = \{\Phi(x) - \Phi(A)\}/\{1 - \Phi(A)\}$ for $A < x < \infty$.

Proposition 10. If $F(A) = 0$, then for $N = 2, 3, \ldots$

$$\sum_{j=1}^{N} E(X_{jN}X_{NN}) = 1 + NE(X)E(X_{N-1,N-1})$$

if and only if $F(x) = \{\Phi(x) - \Phi(A)\}/\{1 - \Phi(A)\}$ for $A < x < \infty$.

PROOF

$$L.H.S = \sum_{j=1}^{N-1} + E(X_{NN}^2) = N(N-1) \iint_{z \le w} zw$$

$$F^{N-2}(w)dF(z)dF(w) + E(X_{NN}^2). \qquad (3.3)$$

Now

$$E(X_{NN}^2) = N \int_A^\infty z \left[\int_A^z \frac{d}{dF(w)} NF^{N-1}(w)dF(w) \right] dF(z)$$

$$= N \iint_{w \le z} zF^{N-1}(w)dwdF(z)$$

$$+N(N-1) \iint_{w \le z} wzF^{N-2}(w)dF(z)dF(w).$$

Now use this in (3.3) and combine the two symmetrical double integrals and proceed as in the proof of Proposition 9.

4 Characterization of the Folded Logistic Distribution

In this section we characterize the folded logistic distribution using the identities in (1.8) and (1.13).

Proposition 11. Let $F(A) = 0$. Then $E\{XF(X)\} = 1 - \frac{A}{2}$ if

$$1 + F(x) = 2[1 + \bar{e}^{(x-A)}]^{-1} - 1 \text{ for } x \le A.$$

PROOF After performing integration by parts once we obtain

$$E\{XF(X)\} = \int_A^\infty xFdF = \int F(1-F)dx$$

$$+ \int x(1-F)dF.$$

That is

$$\begin{aligned} 2E\{XF(X)\} &= \int F(1-F)dx + \int (1-F)dx + A \\ &= \int \{2f - (1-F)\}dx + \int (1-F) \\ &\quad dx + A \\ &= 2 + A. \end{aligned}$$

Proposition 12. If $F(A) = 0$, then for $1 < i < N, N = 2, 3, \ldots$

$\sum_{k=1}^{N} E\{X_{i,N}F(X_{k,N})\} = \frac{N}{2}E(X_{i-1,N-1}) + 1$ if and only if

$$1 + F(x) = 2(1 + \bar{e}^{(x-A)})^{-1}, x \ge A.$$

PROOF. Consider

$$\begin{aligned} &\sum_{k=1}^{N} E[X_{i,N}F(X_{k,N})] \\ &= \sum_{k=1}^{i-1} + E\{X_{iN}F(X_{iN})\} \\ &\quad + \sum_{k=i+1}^{N} E\{X_{i,N}F(X_{k,N})\} \\ &= \frac{N!}{(i-2)!N-i)!} \iint_{z \le w} wF^{i-2} \\ &(w)[1 - F(w)]^{N-i}F(z)dF(z)dF(w) \\ &\quad + \frac{N!}{(i-1)!(N-i)!} \int_0^\infty zF^i \\ &\quad (z)[1 - F(z)]^{N-i}dF(z) \\ &+ \frac{N!}{(i-1)!(N-i-1)!} \iint_{z<w} zF^{i-1} \\ &(z)[1 - F(z)]^{N-i-1}F(w)dF(z)dF(w). \end{aligned}$$

We can write the last integral (without the constant multiplier) as

$$\frac{1}{N-i} \int_A^\infty F(w) \left[\int_A^w zF^{i-1}(z)d\{(1-F)^{N-i}\} \right]$$

$$dF(w) = \frac{1}{N-i} \left[-\int_A^\infty wF^i(1-F)^{N-i}dF(w) + \right.$$

$$(i-1) \iint_{z \le w} zF^{i-2}(z)(1-F)^{N-i}F(w)dF(w)$$

$$dF(z)+\int\int_{z\le w} F^{i-1}(z)$$

$$(1-F(z))^{N-i}\,F(w)dF(w)dz\Big]\,.$$

Thus

$$\sum_{k=1}^{N} E\{X_{iN}F(X_{k,N})\} = \frac{N!}{(i-2)!(N-i)!}$$
$$\int\int_{z\le w}\{wF^{i-2}(w)[1-F(w)]^{N-i}F(z)$$
$$+zF^{i-2}(z)[1-F(z)]^{N-i}F(w)\}$$
$$\cdot dF(z)dF(w)$$
$$+\frac{N!}{(i-1)!(N-i)!}\int\int_{z\le w} F^{i-1}(z)$$
$$(1-F(z))^{N-i}F(w)dF(w)dz$$
$$=\frac{1}{2}\frac{N!}{(i-2)!(N-i)!}\int_A^\infty zF^{i-2}$$
$$(z)[1-F(z)]^{N-i}dF(z)$$
$$+\frac{1}{2}\frac{N!}{(i-1)!(N-i)!}\int_A^\infty F^{i-1}$$
$$(z)[1-F(z)]^{N-i}\{1-F^2(z)]dz$$
$$=\frac{N}{2}E(X_{i-1,N-1})+\frac{1}{2}\frac{N!}{(i-1)!(N-i)!}\int_A^\infty$$
$$F^{i-1}(z)[1-F(z)]^{N-i}[1-F^2(z)]dz\,.$$

Now for the logistic distribution on (A,∞), $f(z)=\frac{1}{2}(1-F^2(z))$; hence the last integral on the right side reduces to unity. Now $\sum_{k=1}^{N}E\{X_{iN}F(X_{k,N})\}=\frac{N}{2}E(X_{i-1,N-1})+1$ for $1<i<N$ and $N=2,3,\ldots$ imply that

$$\frac{N!}{(i-1)!(N-i)!}\int_0^1 u^{i-1}(1-u)^{N-i}\{H'(u)\frac{(1-u^2)}{2}-1\}$$
$$du=0,\text{ for } 1<i<N \text{ and } N=2,3\ldots.$$

That is, $H'(u)=\frac{2}{1-u^2}$ for almost all u in $(0,1)$.

Integrating on both sides and using the fact that $H(0)=A$, we obtain

$H(u)-A=\ln\{(1+u)/(1-u)\}$ and $H'(u)$ exists for all u.

Now letting $u=F(z)$ and $x=H(u)$, we obtain the result.

We will give the following propositions without proofs which are analogous to that of Proposition 12.

Proposition 13. If $F(A)=0$, then for $N=1,2,\ldots$, $\sum_{k=1}^{N}E\{X_{NN}F(X_{k,N})\}=1+\frac{N}{2}E(X_{N-1,N-1})$ if and only if

$$F(x)=2\{1+\bar{e}^{(x-A)}\}^{-1}-1 \text{ for } x\ge A\,.$$

Proposition 14. If $F(0)=0$, then

$1=\sum_{k=1}^{N}E\{X_{1N}F(X_{k,N})\}$ for $N=1,2,\ldots$ if and only if

$$F(x)=2(1+e^{-x})^{-1}-1 \text{ for} x\ge 0\,.$$

5 Characterization of the Uniform Distribution

In this section we characterize the uniform distribution using the identities in (1.15), (1.18) and (1.19).

Proposition 15. For $1<i<N, N=2,3$.

$1=N[E(X_{i,N-1}-E(X_{i-1,N-1})]$ if and only if

$$F(x)=x, 0<x<1\,.$$

Proof. Consider

$$N\{E(X_{i,N-1})-E(X_{i-1,N-1})\}$$
$$=\frac{N!}{(i-1)!(N-1-i)!}\int_0^1 H(u)u^{i-1}(1-u)^{N-1-i}du$$
$$-\frac{N!}{(i-2)!(N-i)!}\int_0^1 H(u)u^{i-2}(1-u)^{N-i}du\,.$$

Now

$$-\frac{1}{N-i}\int_0^1 H(u)u^{i-1}d\{(1-u)^{N-i}\}$$
$$=-\frac{1}{N-i}\,[H(u)u^{i-1}(1-u)^{N-i}|_0^1$$
$$-\int_0^1 H'(u)u^{i-1}(1-u)^{N-i}du$$
$$-(i-1)\int_0^1 H(u)u^{i-2}(1-u)^{N-i}du\,]$$
$$=\frac{1}{N-i}\int_0^1 H'(u)u^{i-1}(1-u)^{N-i}$$
$$du+\frac{i-1}{N-i}$$
$$\int_0^1 H(u)u^{i-2}(1-u)^{N-i}du\,.$$

Hence

$$N\{E(X_{i,N-1}) - E(X_{i-1,N-1})\}$$
$$= \frac{N!}{(i-1)!(N-i)!}\int_0^1 H'(u)u^{i-1}(1-u)^{N-i}du\,.$$

$L.HS = 1$, for $1 < i < N, N = 2, 3, \ldots$ imply that

$$\frac{N!}{(i-1)!(N-i)!}\int_0^1 u^{i-1}(1-u)^{N-i}\{H'(u)-1\}du = 0 \text{ for } 1 < i < N, N = 2.3, \ldots .$$

That is

$H'(u) = 1$ for almost all u in $(0,1)$. Thus $H'(u)$ exists for all u in $(0,1)$and $H'(u) = \frac{1}{f(F^{-1}(u)}$.

Hence

$$\frac{1}{f(x)} = 1 \text{ or } f(x) = 1\,.$$

Proposition 16. $F(0) = 0$ and $E(X_{1N}) = 1/(N+1)$ for $N = 1, 2, \ldots$ if and only if $F(x) = x, 0 < x < 1$.

Proposition 17. $F(0) = 0$ and $E(X_{NN}) = N/(N+1)$ for $N = 1, 2, \ldots$ if and only if $F(x) = x, 0 < x < 1$.

Proofs of Propositions 16 and 17 are analogous to the proof of Proposition 15. (For Proposition 16, see also Galambos and Kotz ([4], pp. 55-57) or Lin ([9], Theorem 1, p. 398)).

Remark 6 Ahsanullah [1] has characterized the uniform distribution using the identical distribution of $X_{NN} - X_{1,N}$ and $X_{N-1,N}$ in the class of all super or subadditive and absolutely continuous distributions. Note that a distribtuion F is said to be super (sub) additive if $F(x+y) \geq F(x)+F(y)\{F(x+y) \leq F(x)+F(y)\}$ for all $x, y > 0$.

References

[1] Ahsanullah, M. On characterizations of the uniform distribution based on functions of order statistics. *Aligarh Journal of Statistics* 9 (1989) 1-6.

[2] Bennett, C. A. *Asymptotic Properties of Ideal Estimators* Ph.D. Thesis, University of Michigan (1952).

[3] Govindarajulu, Z. Exact lower moments of order statistics in samples from the chi-distribution (l.d.f.) *Ann. Math. Statist.* **33** (1962) 1292-1305.

[4] Galambos, J. and Kotz, S. *Characterizations of Probability Distributions* (1978) Lecture Notes in Mathematics, Springer, Berline-New York.

[5] Govindarajulu, Z. On moments of order statistics and quasi-ranges from normal populations. *Ann. Math. Statist.* **34** (1963) 633-651.

[6] Govindarajulu, Z. Characterization of normal and generalized truncated normal distributions using order statistics *Ann. Math Statist.* **37** (1966) 1011-1015.

[7] Govindarajulu, Z. Characterization of the exponential distribution using lower moments of order statistics pp. 117-129 in *A Modern Course on Statistical Distributions in Scientific Work* (1974) (Editors G. P. Patil, S. Kotz and J. K. Ord). D. Reidel Publishing Co. Dordrecht. Holland.

[8] Govindarajulu, Z. and Lindqvist, B.H. Asymptotic efficiency of the Spearman estimator and Characterizations of distributions. *Ann. Inst. Statist. Math.* **39** (1987) 349-361.

[9] Lin, G. D. The product moments of order statistics with applications to characterizatoins of distributions. *J. Statist. Planning and Inference* **21** (1989). 395-406.

[10] Seal, K. C. On minimum variance among certain linear functions of order statistics. *Ann. Math. Statist.* **27** (1956) 854-855.

Investigating The Bias And MSE Of Exceedance Based Tail Estimators For The Cauchy Distribution

Grimshaw, S.D.
Brigham Young University, Provo, UT

One approach to estimating the tails of the cumulative distribution function, quantile function and probability density function is to use only those observations in the sample which exceed a high threshold. This investigation for the Cauchy probability model will indicate how the threshold selection affects the bias and MSE of tail estimators.

Key words and phrases: Generalized Pareto distribution, Hill's estimator.

1. Introduction

Suppose that the possible observed values from a population can be characterized by a random variable X whose probability model is estimated using a sample from the population. The properties of this estimated probability model which correspond to the population characteristics of interest are the foundation of statistical analysis.

Three important functions of a probability model for a continuous random variable are the absolutely continuous distribution function $F(x) = \mathrm{P}[X \leq x]$, the quantile function $Q(u) = F^{-1}(u)$ and the density function given by $f(x)$ which represents $\mathrm{P}[a < X < b] = \int_a^b f(x)dx$ for $a < b$. The significance of these three functions in statistical analysis follows from their interpretation as key properties of the population.

This work focuses on the problem of estimating the tails of $F(x)$, $Q(u)$ and $f(x)$ from a random sample. The sample (empirical) distribution function, sample (empirical) quantile function and nonparametric density estimates are typically used in early stages of statistical analysis when minimal assumptions are made on the underlying probability model. However, these estimators prove unsatisfactory for values in the tail since they are confined to the observed sample values and ignore the possibility of more extreme values than the observed sample in future observations. For example, the empirical quantile function in the case of insurance claims would not estimate any insurance claim larger those already observed in the sample, that is, it assumes the largest insurance claims have already been filed and any future insurance claims will not exceed the sample extremes. This restriction to tail estimation is unacceptable.

The classical approach to tail estimation is to assume the underlying probability model belongs to some known class $\mathcal{P}$ whose elements are indexed by a parameter θ taking values in a set Θ, that is $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}$. The distribution function, quantile function and density function then have parametric representations $F(x;\theta)$, $Q(u;\theta)$ and $f(x;\theta)$. Tail estimates are given by $F(x;\hat{\theta})$, $Q(u;\hat{\theta})$ and $f(x;\hat{\theta})$, where $\hat{\theta}$ denotes an estimate of the parameter θ based on the sample information. For example, if the underlying probability model is assumed to follow the normal probability law, then $\theta = (\mu, \sigma)$, the mean and standard deviation, with estimator $\hat{\theta} = (\overline{X}, s)$, the sample mean and sample standard deviation. The tail estimator is formed by replacing the unknown parameters with the estimators in the normal distribution function, quantile function and

density function.

The beauty of this classical parametric approach is tarnished by what Fisher [1] called the problem of specification. Often it is difficult to select a single parametric family for the population. Several candidates may appear reasonable judging from their fit to the observed values. To demonstrate this complication, suppose that a random sample is taken from a population characterized by a symmetric unimodal probability model. Two possible parametric families are the normal and the Cauchy. Inference on the central values of the random variable will be similar for either of these parametric models. However, the focus of this work is on tail values, not central values, and inference at the tails is quite different under the two parametric models. Extremely small and extremely large values are much more likely under the Cauchy modeling. The distribution function $F(x)$ approaches zero and one much more rapidly under the normality assumption. The quantile function $Q(u)$ for the Cauchy model decreases more rapidly in a neighborhood of zero and increases more rapidly in a neighborhood of one. The density function $f(x)$ for the Cauchy model has much more area in the tail.

It is very difficult to discriminate between the different possible parameterizations even when the possible parametric models specify very different tail properties. Very large sample sizes are required for a goodness of fit test to have sufficient power to detect differences in the observed tail and the fitted tail under the normal and Cauchy modeling.

2. Exceedance Based Tail Estimators and Their Properties

The main objective of this work is to obtain estimators of $F(x)$, $Q(u)$ and $f(x)$ which allow the data, not the parametric family, to dictate the tail behavior of the underlying probability model. These estimators can be used in applications to validate tail behavior properties in probability modeling applications. In this discussion, the upper tail is discussed without loss of generality since the lower tail results follow immediately after one observes that the lower tail becomes the upper tail if the data is negated.

Grimshaw [2] proposed the following paradigm for tail estimation:

1. From a random sample $X_1, \ldots, X_n$, choose, as a function of n, a threshold percentile $t_n = k/n$ close to zero, for some integer k.
2. Estimate the corresponding threshold $T_n = X(n-k;n)$, the $(n-k)^{th}$ order statstic.
3. Obtain parameter estimates $(\hat{\rho}, \hat{a})$ from the exceedances $X_i - T_n$ for all $X_i > T_n$.
4. Estimate the tails of the quantile function, distribution function, and density function by

$$\hat{Q}(u) = T_n + \hat{a}\left[-g\left(\frac{1-u}{t_n}; -\hat{\rho}\right)\right]$$

$$\text{for } 1 - t_n < u < 1,$$

$$\hat{F}(x) = 1 - t_n \cdot \left[g^{-1}\left(-\frac{1}{\hat{a}}[x - T_n]; -\hat{\rho}\right)\right]$$

$$\text{for } T_n < x < Q(1),$$

$$\hat{f}(x) = t_n \frac{1}{\hat{a}} \cdot (g^{-1})'\left(-\frac{1}{\hat{a}}[x - T_n]; -\hat{\rho}\right)$$

$$\text{for } T_n < x < Q(1).$$

where

$$g(z;\lambda) = \begin{cases} \dfrac{z^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln z, & \lambda = 0 \end{cases}$$

$$g^{-1}(z,\lambda) = \begin{cases} (1+\lambda z)^{1/\lambda}, & \lambda < 0,\ z < 0 \\ e^z, & \lambda = 0,\ z < 0 \\ (1+\lambda z)^{1/\lambda}, & \lambda > 0, \\ & -1/\lambda < z < 0 \end{cases}$$

and

$$(g^{-1})'(z,\lambda) = \begin{cases} (1+\lambda z)^{(1/\lambda)-1}, & \lambda < 0,\ z < 0 \\ e^z, & \lambda = 0,\ z < 0 \\ (1+\lambda z)^{(1/\lambda)-1}, & \lambda > 0, \\ & -1/\lambda < z < 0 \end{cases}$$

Grimshaw [2] has shown that these estimators are asymptotically normal if the estimators $(\hat{\rho}, \hat{a})$ are asymptotically normal as $nt \to \infty$.

Two options are available for estimating the parameters (ρ, a). One approach suggested by Pickands [3] is to model the k exceedances of the threshold T_n as a sample from a generalized Pareto distribution (GPD) whose parameters can be estimated by maximum likelihood. One example of this approach is given by Smith [4] to estimate extreme ozone levels. Grimshaw [5] has proposed an algorithm for computing the GPD maximum likelihood estimates. Let $(\hat{\rho}_{\text{GPD}}, \hat{a}_{\text{GPD}})$ denote the maximum likelihood estimates from the GPD model for the exceedances.

A second approach was suggested by Hill [6] using

$$\hat{\rho}_{\text{Hill}} = \frac{1}{k}\sum_{i=1}^{k} \ln\left[\frac{X(n-i+1;n)}{X(n-k;n)}\right]$$
$$\hat{a}_{\text{Hill}} = \hat{\rho}_{\text{Hill}} T_n.$$

It has been shown by many authors that for these estimators based on the exceedances of a high threshold, as $nt \to \infty$,

$$\begin{bmatrix}\hat{\rho}_{\text{GPD}} \\ \hat{a}_{\text{GPD}}\end{bmatrix} \quad \text{is} \quad \text{AN}\left(\begin{bmatrix}\rho \\ a\end{bmatrix}, (nt)^{-1}V_{\text{GPD}}\right)$$

where

$$V_{\text{GPD}} = \begin{bmatrix}(\rho+1)^2 & -a(\rho+1) \\ -a(\rho+1) & 2a^2(\rho+1)\end{bmatrix}$$

and

$$\begin{bmatrix}\hat{\rho}_{\text{Hill}} \\ \hat{a}_{\text{Hill}}\end{bmatrix} \quad \text{is AN}\left(\begin{bmatrix}\rho \\ a\end{bmatrix}, (nt)^{-1}V_{\text{Hill}}\right)$$

where

$$V_{\text{Hill}} = \begin{bmatrix}\rho^2 & a\rho \\ a\rho & a^2\end{bmatrix}.$$

3. Effect of Threshold Selection for the Cauchy Distribution

This investigation focuses on the Cauchy probability model with

$$F(x) = \tfrac{1}{2} + \tfrac{1}{\pi}\tan^{-1} x, \qquad x \in \mathbb{R}$$
$$Q(u) = \tan\pi(u - \tfrac{1}{2}), \qquad 0 < u < 1$$
$$f(x) = \frac{1}{\pi(1+x^2)}, \qquad x \in \mathbb{R}.$$

According to the tail behavior classification of Parzen [7], the Cauchy distribution has a tail exponent of $\rho = 1$ and therefore is classified as a long tailed distribution.

Figure 1 shows the bias of the proposed tail estimators for $t = 0.25,\ 0.15,\ 0.10,\ 0.05$ for the distribution function, quantile function and density function, respectively. As anticipated, as $t \to 0$ the bias is reduced. For both the distribution function and density function it is interesting to note that the largest bias occurs immediately following the threshold. Figures 2(a) and 2(b) compare the MSE of the distribution function tail estimator for the GPD estimators and Hill's estimators, Figures 3(a) and 3(b) compare the MSE of the quantile function tail estimator for the GPD estimators and Hill's estimators, and Figures 4(a) and 4(b) compare the MSE of the density function tail estimator for the GPD estimators and Hill's estimators. A sample size such that $nt = 30$ is assumed in each of these figures. In all cases it is clear that the superior precision of Hill's estimators translate into superior tail estimators.

From these figures, it is obvious that a small value for t results in better estimates. However, the evaluation of a choice of t must be based relative to the desired region for the tail estimate. For example, the choice $t = 0.25$ provides an excellent estimator of the quantile function for $0.75 < u < 0.95$. But if a more extreme tail estimate is desired, say from $0.90 < u < 0.98$, then $t = 0.10$ is a more appropriate choice.

A simulation study was performed to investigate the effect of threshold selection on tail parameter estimators. In this simulation, threshold values of $t = 0.25,\ 0.15,\ 0.10,\ 0.05,\ 0.01$ were chosen with the number of exceedences of the selected threshold of $nt = 30,\ 50,\ 100$. Table I contains the estimated mean and estimated variance of the tail estimator parameters for the Cauchy distribution based on 100 simulations.

Notice both the GPD and Hill estimators are biased for the tail parameter $\rho = 1$. This bias is not due to any failure of the methodology. Grimshaw [2] showed that estimators of the tail exponent are asymptotically unbiased as $nt \to \infty$, but the rate of convergence can be very slow. Further, it can be shown that for $\rho > 0$ Hill's estimators are asymptotically superior. From the simulation, it appears that Hill's estimators are superior in finite samples and that for the Cauchy probability model the bias appears quite small.

It appears from the simulation that threshold selection affects the GPD and Hill estimators differently. For $\hat{\rho}_{\text{GPD}}$, as t decreases and nt increases at each step the estimator improves incrementally. In contrast, $\hat{\rho}_{\text{Hill}}$ has a large improvement changing from $t = .25$ to $t = .15$, but the remaining choices of t yield nearly equivalent estimators. Further, there is little change in the bias as nt is increased. This indicates that $\hat{\rho}_{\text{Hill}}$ may be more robust to threshold selection than $\hat{\rho}_{\text{GPD}}$.

Also notice that the parameter a depends on the choice of t. Grimshaw [2] showed that the parameter a is a function of t and must be asymptotically equivalent to $t/fQ(1-t)$ as $t \to 0$, where $fQ(u)$ denotes the density-quantile function. For the Cauchy distribution, at $t = 0.25$, $a \sim 1.571$; at

$t = 0.15$, $a \sim 2.286$; at $t = 0.10$, $a \sim 3.290$; at $t = 0.05$, $a \sim 6.419$; and at $t = 0.01$, $a \sim 31.841$. In estimating a it appears that $\hat{a}_{\text{GPD}}$ provides a good estimator for all choices of t but $\hat{a}_{\text{Hill}}$ only performs well for small values of t and large values of nt.

Therefore, it appears that Hill's estimator of ρ is rather robust to threshold selection but the corresponding estimator of a is poor except as $t \to 0$ and $nt \to \infty$. The GPD estimator of ρ is inferior compared to Hill's estimator, but the pair $(\hat{\rho}_{\text{GPD}}, \hat{a}_{\text{GPD}})$ demonstrate a steady incremental improvement over the values of t and nt.

REFERENCES

[1] Fisher, R.A., *Statistical Methods for Research Workers*, Oliver and Boyd, London, 1948.

[2] Grimshaw, S.D., "A unified approach to estimating tail behavior," *Ph.D. Dissertation*, Texas A&M University, (1989).

[3] Pickands, J., "Statistical inference using extreme order statistics," *Annals of Statistics*, 3 (1975) 119–131.

[4] Smith, R.L., "Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone," *Statistical Science*, 4 (1989), 367–393.

[5] Grimshaw, S.D., "Computing maximum likelihood estimates for the generalized Pareto distribution," *Technometrics*, 35 (1993), 185–191.

[6] Hill, B.M., "A simple general approach to inference about the tail of a distribution," *Annals of Statistics*, 3 (1975), 1163–1174.

[7] Parzen, E.,"Nonparametric statistical data modeling," *Journal of the American Statistical Association*, 74 (1979), 105-131.

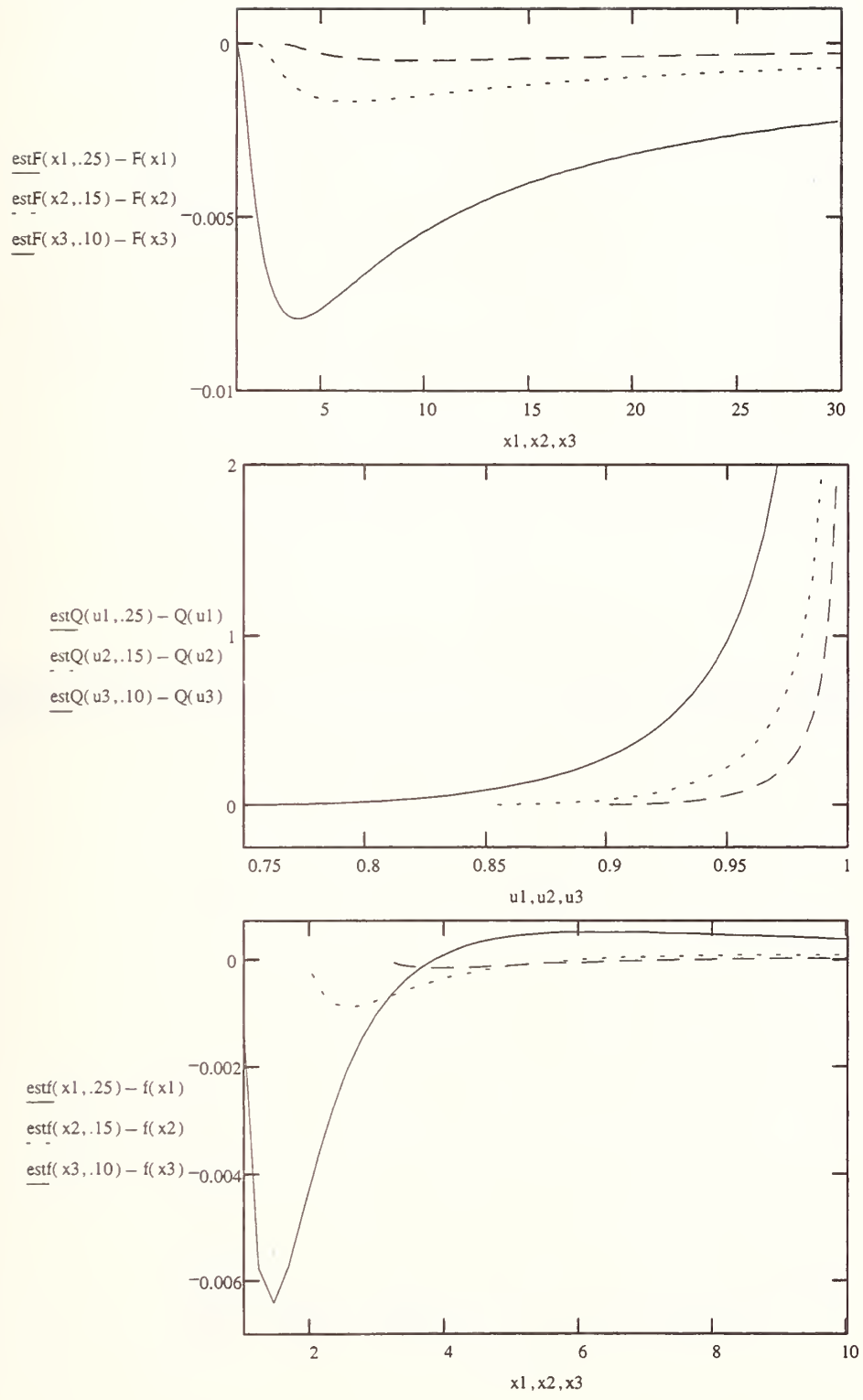


Figure 1. Bias of the Tail Estimators of $F(x)$, $Q(u)$ and $f(x)$, respectively, for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

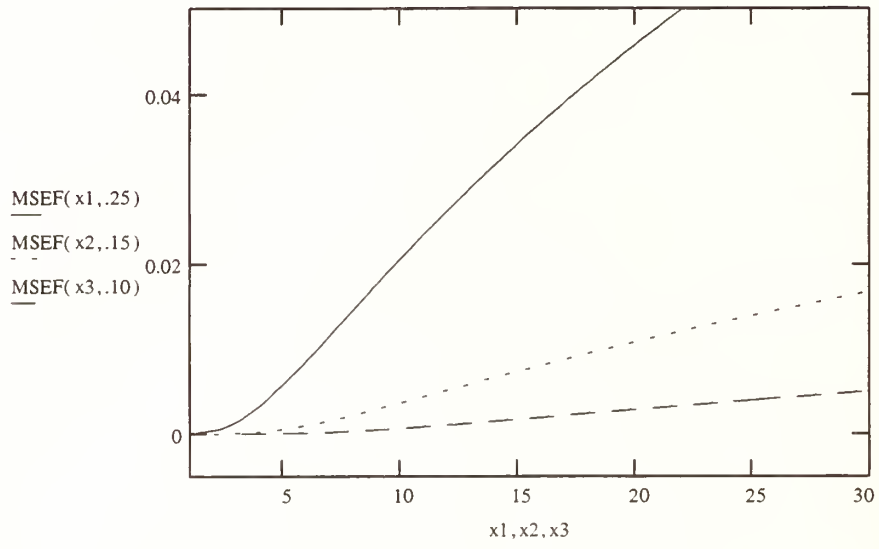


Figure 2(a). Mean Square Error of the Tail Estimators of $F(x)$ using the GPD parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

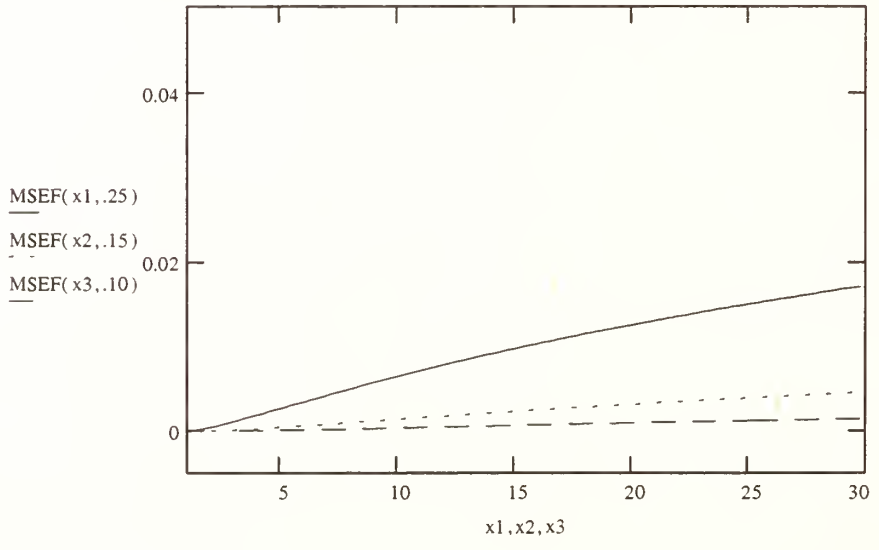


Figure 2(b). Mean Square Error of the Tail Estimators of $F(x)$ using the Hill parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

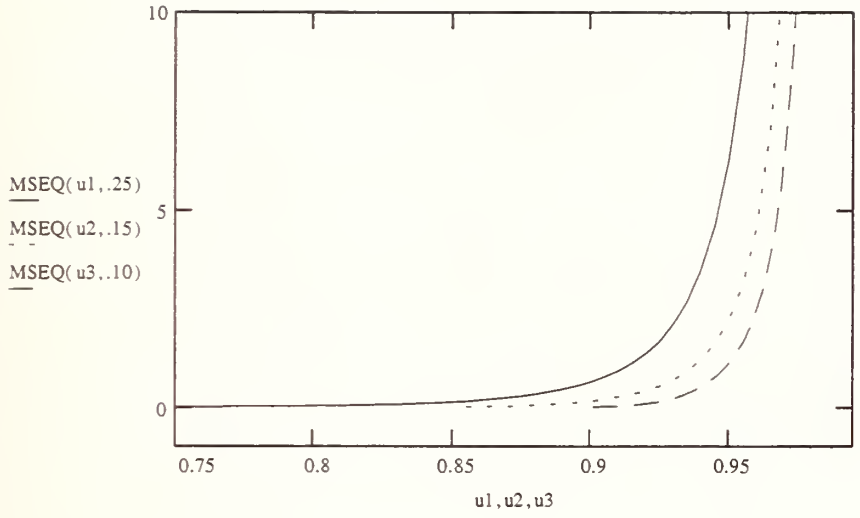


Figure 3(a). Mean Square Error of the Tail Estimators of $Q(u)$ using the GPD parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

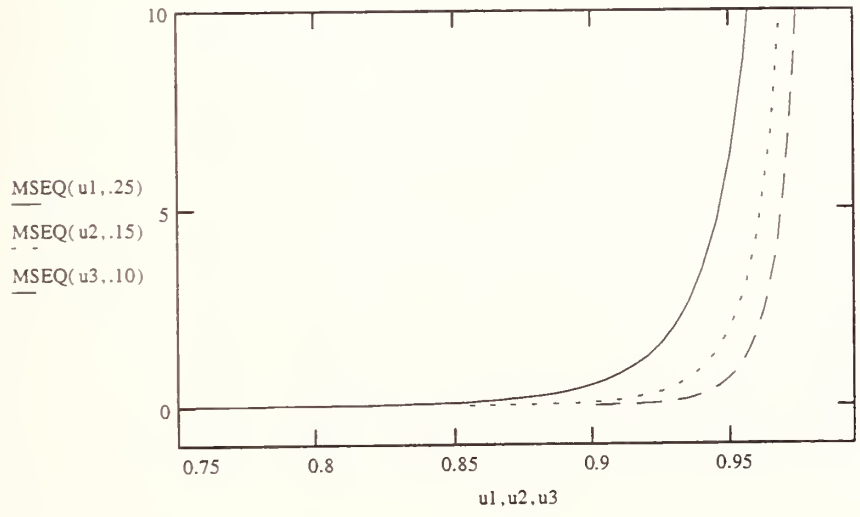


Figure 3(b). Mean Square Error of the Tail Estimators of $Q(u)$ using the Hill parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

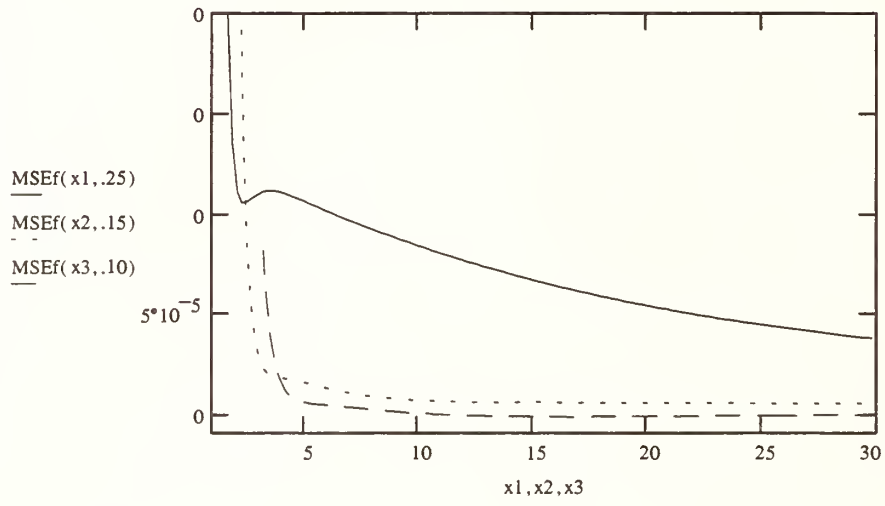


Figure 4(a). Mean Square Error of the Tail Estimators of $f(x)$ using the GPD parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

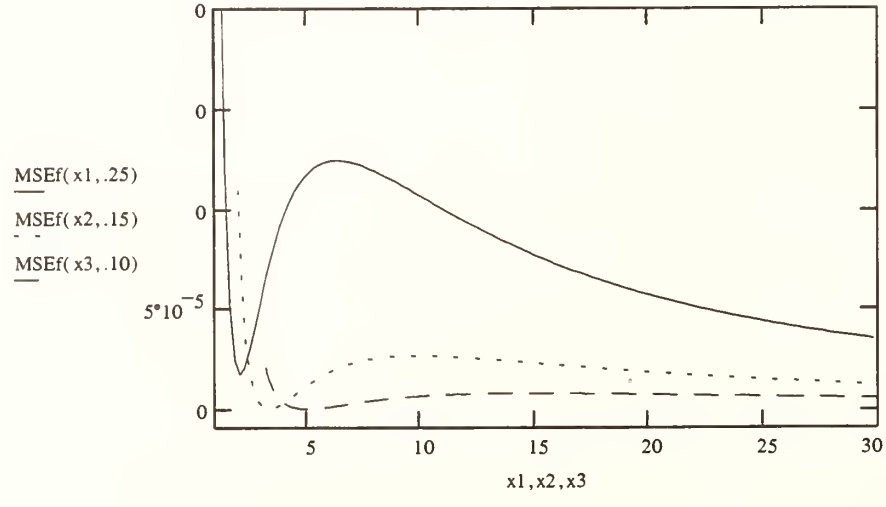


Figure 4(b). Mean Square Error of the Tail Estimators of $f(x)$ using the Hill parameter estimators for the Cauchy Distribution for threshold percentiles $t = 0.25,\ 0.15,\ 0.10$.

TABLE I

Estimated Mean and Estimated Variance of Tail Estimator Parameters
from a Cauchy Distribution Based on 100 Simulations

		$t = 0.25$		$t = 0.15$		$t = 0.10$		$t = 0.05$		$t = 0.01$	
		GPD	Hill	GPD	Hill	GPD	Hill	GPD	Hill	GPD	Hill
$nt = 30$	ρ	0.818456	1.133128	0.919654	1.044839	0.974232	1.052050	0.987711	1.036878	0.906788	1.004992
		(0.131044)	(0.037187)	(0.125110)	(0.036518)	(0.167713)	(0.034125)	(0.170044)	(0.032416)	(0.191387)	(0.035137)
	a	1.544909	1.146493	2.390553	2.149118	3.638868	3.245504	7.090819	6.473356	36.38259	32.28838
		(0.338616)	(0.120682)	(0.709719)	(0.353321)	(2.916572)	(0.647649)	(7.156168)	(2.776208)	(184.4395)	(82.55424)
$nt = 50$	ρ	0.894691	1.151105	0.895877	1.046539	0.985744	1.036116	0.940233	0.983823	0.914454	0.974548
		(0.065732)	(0.030686)	(0.080714)	(0.023244)	(0.075546)	(0.017996)	(0.069046)	(0.021497)	(0.112587)	(0.022282)
	a	1.481984	1.145476	2.437840	2.100748	3.400780	3.120634	6.538807	6.051194	33.91855	31.53411
		(0.184154)	(0.041350)	(0.514333)	(0.180620)	(0.821246)	(0.366974)	(3.510222)	(1.319512)	(117.2677)	(35.81257)
$nt = 100$	ρ	0.947562	1.160791	1.002174	1.069873	0.967803	1.023114	1.005176	1.005949	0.983477	1.007008
		(0.042786)	(0.014231)	(0.043653)	(0.010437)	(0.035851)	(0.010236)	(0.038494)	(0.009977)	(0.043505)	(0.010141)
	a	1.444734	1.160812	2.250815	2.086513	3.308631	3.088761	6.431612	6.315314	33.41112	32.24122
		(0.117365)	(0.022773)	(0.226079)	(0.087592)	(0.277570)	(0.195727)	(1.724948)	(0.702704)	(60.79857)	(18.15151)

Estimating Quantiles For A Type III Domain Of Attraction Based On The k Largest Observations

Hasofer, A.M.
The University of New South Wales, New South Wales Australia

Wang, J.Z.
University of Western Sydney, New South Wales, Australia

A method of estimating the high quantiles of a distribution belonging to the domain of attraction of type III extreme value distribution (reversed Weibull) is proposed by means of the W statistic, which is used to determine the extreme value domain of attraction. The procedure is based on the k largest order statistics from a sequence of n observations. The problem is treated by a three-parameter model and the endpoint of the distribution is estimated. A test of hypothesis is used to eliminate cases where the endpoint does not exist. The estimators are shown to be asymptotically consistent. Simulation results are provided.

1. Introduction and summary

Suppose a distribution F(x) is in the domain of attraction of a distribution H(x), which has been identified to be the same, up to location and scale, as one of the extreme value distributions, whose types are given by

Type I (Gumbel):

$$H_0(x) = \exp\{-\exp\{-x\}\},\ \infty>x>-\infty;$$

Type II (Frechet):

$$H_{1,\gamma}(x) = \exp\{-x^{-\gamma}\},\ x>0;$$

Type III (reversed Weibull):

$$H_{2,\gamma}(x) = \exp\{-(-x)^{\gamma}\},\ x<0,$$

where γ is some positive constant (see Ref. [2]). The problem of estimating large quantiles of the distribution has been addressed by several authors.

In Ref. [6] an estimator for large quantiles of a distribution based on the k largest observations is derived. It is assumed that the distribution belongs to the domain of attraction of $H_0(x)$ and the derivation is based on the asymptotic distribution of the k largest order statistics of the sample as the sample size goes to infinity. For distributions in the domain of attraction of $H_{1,\gamma}(x)$ and $H_{2,\gamma}(x)$, however, there are three parameters and therefore a different approach is needed.

For the three-parameter problem Hasofer and Wang (Ref. [4]) suggested estimating high quantiles of a distribution in the domain of attraction of $H_{1,\gamma}(x)$ by a local maximum likelihood method based on the extremal distribution of the k largest order statistics. This procedure works for all values of γ, provided n, the sample size, and k are large enough.

The three-parameter problem for distributions in the domain of attraction of $H_{2,\gamma}(x)$ was studied by Smith and Weissman (Ref. [5]). The method is based on a local maximum estimation procedure. However, the approach sometimes fails to yield an estimator because the likelihood function fails to have a needed local

maximum, when $\gamma<1$.

In this paper, we suggest that the three-parameter problem for a distribution in the domain of attraction of $H_{2,\gamma}(x)$ can be solved by means of the *W* statistic. The endpoint of the underlying distribution is estimated by solving a simple equation based on the *W* statistic and the relation between $H_0(x)$ and $H_{2,\gamma}(x)$. We shall show that the estimator is asymptotically consistent. The *W* statistic was proposed by Hasofer and Wang (Ref. [3]) for testing the extreme value domain of attraction. The statistical properties of the *W* statistic were studied and the asymptotic distribution was determined.

2. The *W* statistic

The statistic *W* introduced by Hasofer and Wang (Ref. [3]) is a function of the top k order statistics of a sample of size n: $X_{1n} \geq \cdots \geq X_{kn}$, and is given by

$$W(X_{1n}, \ldots, X_{kn}) = \frac{k}{k-1} \frac{(\bar{X} - X_{kn})^2}{\sum_{i=1}^{k} (X_{in} - \bar{X})^2}$$

where $\bar{X} = (\sum_{i=1}^{k} X_{in})/k$. The critical values for the null hypothesis that the distribution of X is in the domain of attraction of $H_0(x)$ are given in Table VIII of Ref. [3]. It was shown there that a value of *W* lower than the lower critical point indicates that the distribution belongs to the domain of attraction of $H_{1,\gamma}(x)$, while a value higher than the higher critical point indicates that it belongs to the domain of attraction of $H_{2,\gamma}(x)$. It was also shown that under the null hypothesis the *W* statistic is asymptotically normally distributed with mean k^{-1} and variance $4k^{-3}$. The power of the *W* test was studied by extensive simulation.

3. Estimating the endpoint

Suppose that the distribution of X, F(x), is determined to be in the domain of attraction of $H_{2,\gamma}(x)$. Then the support of F(x) must have a finite upper bound ω_0, say. Let $X_{1n} \geq \cdots \geq X_{kn}$ be the *k* top order statistics of a sample of size n. In this case the limiting distribution of the $Y_{in} = -\ell n(\omega_0 - X_{in})$, i = 1, ..., k, is, after a transformation of scale and origin by a suitable pair of sequences, the extremal distribution corresponding to a distribution in the domain of attraction of $H_0(x)$ (Ref. [6]). Note that the value of *W* is invariant with respect to the above linear transformation.

Our proposal for dealing with the estimation problem for distributions in the domain of attraction of $H_{2,\gamma}(x)$ is to seek an estimator of ω_0, $\hat{\omega}$, such that, for a given sample,

$$W\left(-\ell n(\hat{\omega} - X_{1n}), \ldots, -\ell n(\hat{\omega} - X_{kn})\right) = E\left(W(U_1, \ldots, U_k)\right)$$

which is asymptotic to 1/k as $k \to \infty$. Here $E(\cdot)$ is the expected value and the U_i's have the joint density

$$h_0(u_1, \ldots, u_k) = \exp\left\{-\exp\{-u_k\} - \sum_{i=1}^{k} u_i\right\},$$

for $u_1 \geq \ldots \geq u_k$. In the following discussion, we denote X_{in} by X_i, i = 1, ..., k, and

$$W(\omega) = W\left(-\ell n(\omega - X_1), \ldots, -\ell n(\omega - X_k)\right).$$

The estimation of ω enables one to make the transformation and then to estimate the quantile as for the case of a distribution in the domain of attraction of $H_0(x)$ (Ref. [6]).

3.1. The limits of $W(\omega)$

We look at the limits of *W* as a function of ω, when ω tends to X_1 and ∞.

Lemma 1. Suppose $X_1 > X_2$. Then

$$\lim_{\omega \to X_1} W(\omega) = \frac{1}{(k-1)^2}.$$

Proof. Let $Y_i = -\ell n(\omega - X_i)$, $i = 1, \ldots, k$, and

$$G(\omega) = \frac{1}{k} + \frac{1}{k(k-1)W(\omega)}.$$

Then

$$G(\omega) = \frac{\sum_{i=1}^{k-1} (Y_i - Y_k)^2}{\left[\sum_{i=1}^{k-1} (Y_i - Y_k)\right]^2}.$$

Note that $Y_1 > Y_2$. Since $Y_1 \to \infty$ and all other Y_i's remain finite, then

$$\lim_{\omega \to X_1} G(\omega) = 1$$

and

$$\lim_{\omega \to X_1} W(\omega) = \frac{1}{(k-1)^2}. \quad \blacksquare$$

Lemma 2.

$$\lim_{\omega \to \infty} W(\omega) = W(X_1, \ldots, X_k).$$

Proof. This can be shown by observing that

$$\lim_{\omega \to \infty} \omega(Y_i - Y_k) = X_i - X_k,$$

for $i = 1, \ldots, k-1$. $\blacksquare$

3.2. The monotonicity of $W(\omega)$

We are going to show that $W(\omega)$ is an increasing function for all $\omega > X_1 \geq \ldots \geq X_k$.

Lemma 3. The function

$$f(x) = \frac{(x-c)/(cx)}{\ell n(c/x)}$$

is a strictly monotone increasing function for all $0 < x \leq c$.

Proof.

$$\frac{\partial f(x)}{\partial x} = \frac{x^{-2}[\ell nc - \ell nx + ((x-c)/c)]}{[\ell n(c/x)]^2}.$$

By noting that

$$\ell nx = \ell nc + ((x-c)/c) - ((x-c)^2/2\xi^2), \quad 0 < \xi < c,$$

the lemma can be directly proved. $\blacksquare$

Theorem 1. $W(\omega)$ is a strictly monotone increasing function for all $\omega > X_1 \geq \ldots \geq X_k$, where the X_i's are not all equal.

Proof. Let

$$0 \leq y_i = Y_i - Y_k = \ell n \frac{\omega - X_k}{\omega - X_i}$$

and

$$y_i' = \partial y_i / \partial \omega, \quad i = 1, \ldots, k-1.$$

Then

$$\partial G(\omega)/\partial \omega = N/D$$

where

$$D = 2^{-1}\left(\sum_{i=1}^{k-1} y_i\right)^3$$

$$N = \sum_{i=1}^{k-1} \sum_{j=1}^{i-1} (y_i - y_j) y_i y_j [(y_i'/y_i) - (y_j'/y_j)]$$

(see Ref. [3]). Note that

$$y_i'/y_i = \frac{[(\omega - X_i) - (\omega - X_k)]/[(\omega - X_i)(\omega - X_k)]}{\ell n[(\omega - X_k)/(\omega - X_i)]}.$$

Then by lemma 3,

$$y_i'/y_i \geq y_j'/y_j, \text{ for } i > j,$$

since $\omega - X_i \geq \omega - X_j$. Moreover we have $y_i \leq y_j$, for $i > j$. Therefore $N \leq 0$. And, since the X_i are not all equal, then $N < 0$ and $D > 0$. So $\partial G(\omega)/\partial \omega < 0$ and $\partial W(\omega)/\partial \omega > 0$. $\blacksquare$

Now it is clear that the proposed estimator of ω_0 inevitably exists when the W test rejects the null hypothesis at some level less than 20%, say, in favor of H_A: the random variable is in the domain of attraction of type III, since $W(\omega)$ varies from $1/(k-1)^2$, which is less than $E\left(W(U_1, \ldots, U_k)\right)$, to $W(X_1, \ldots, X_k)$, which is greater than $E\left(W(U_1, \ldots, U_k)\right)$, for $k \geq 3$.

3.3. Simulation results

The simulation is based on the limiting distribution of the k top order statistics of a sample from $H_{2,\gamma}(x)$. The joint density is given by $h_{2,\gamma}(x_1, \ldots, x_k)$

$= \gamma^k[(-x_1) \ldots (-x_k)]^{\gamma-1}\exp\{-(-x_k)^\gamma\}$

for $0 > x_1 \geq \ldots \geq x_k$ (see Ref. [6]). Here we set $k = 50$ and

$$E\left(W(U_1, \ldots, U_{50})\right) = 1/50.$$

Figure 1 shows a successful case when the estimator of ω_0 exists. As ω increases from X_1, the value of W increases over 1/50. Figure 2 shows a failure case when the value of $W(X_1, \ldots, X_{50})$ is lower than 1/50. $W(\omega)$ increases slowly up to $W(X_1, \ldots, X_{50})$ as ω tends to ∞.

Table 1 is a simulation result with 2000 replications on estimating endpoint of $h_{2,\gamma}(x_1, \ldots, x_{50})$ with an endpoint $\omega_0 = 0$. **N.L.** is the number of samples for which the solutions of $\hat{\omega}$ do not exist. The differences in **N.L.** for each γ are apparently because of the power of the W test: with lower value of γ the power of the test is higher.

Table 2 is a simulation result with 2000 replications on estimating endpoint of $h_{2,\gamma}(x_1, \ldots, x_{100})$ with the endpoint $\omega_0 = 0$.

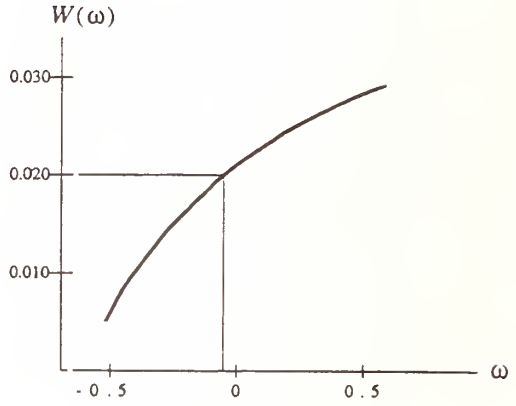


Figure 1

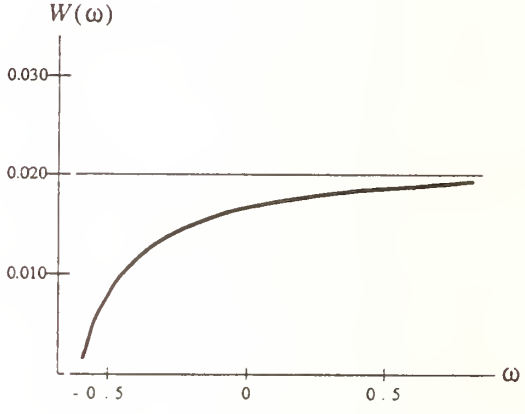


Figure 2

Table 1

γ		0.6	0.8	1	2	3	4
ESTIMATED VALUES	MEAN	-0.223	-0.177	-0.130	0.052	0.156	0.169
	SD	4.195	2.077	1.535	1.354	1.334	1.313
N.L.		0	0	0	22	33	104

Table 2

γ		0.6	0.8	1	2	3	4
ESTIMATED VALUES	MEAN	-0.313	-0.239	-0.144	-0.085	0.025	0.140
	SD	3.819	2.014	1.416	0.807	0.954	1.144
N.L.		0	0	0	0	0	12

3.4. Consistency of the estimator of endpoint.

Assume that $X_1 \geq \ldots \geq X_k$ have the joint density $h_{2,\gamma}(x_1, \ldots, x_k)$, up to location and scale with the end-point ω_0. We have the following.

Theorem 2. The proposed estimator is asymptotically consistent.

The proof is given in the appendix.

4. Estimating quantiles

Let the upper ε-quantile, denoted by $q_{1-\varepsilon}$, of a random variable X be defined by

$$P(X \leq q_{1-\varepsilon}) = 1 - \varepsilon,$$

where ε is some positive small number. Suppose that the distribution function of X is in the domain of attraction of $H_{2,\gamma}(x)$ and the endpoint ω_0 has been estimated with value $\hat{\omega}$ by the proposed method based on the top k statistics from a sample of size n. Then the estimator of quantile is given by

$$\hat{q}_{1-c/n} = \hat{\omega} - \exp\left\{\hat{b} - \hat{a}\ell nc\right\}$$

where

$$\hat{a} = \frac{k}{k-1}(\bar{Y} - Y_k)$$

and

$$\hat{b} = \hat{a}(S_k - \gamma_0) + Y_k$$

with

$$Y_i = -\ell n(\hat{\omega} - X_i),\ i = 1, \ldots, k,$$

$$S_k = \Sigma_{i=1}^{k-1} i^{-1}$$

and

$$\gamma_0 = 0.5772\cdots$$

(Ref. [6]).

As an example of estimation of quantiles, simulation based on $\mathcal{U}(0,1)$ with replication 3000 is carried out (Table 3). The sample size is 500 and the top 55 is selected.

In order to compare our method with some classical methods, the LINT procedure for percentile estimation (see Ref. [1]) is carried out. Table 4 is obtained by simulation with 3000 replications.

Another example of estimating quantiles by the proposed procedure is based on the reversed exponential distribution with replication 3000 (Table 5). The sample size is 500 and the top 55 is selected.

Table 3

p		0.950	0.975	0.990	0.995	0.998
EXACT QUANTILE		0.950	0.975	0.990	0.995	0.998
ESTIMATED VALUES	MEAN	0.949	0.975	0.990	0.994	0.997
	SD	0.00866	0.00626	0.00369	0.00256	0.00216

Table 4

p		0.950	0.975	0.990	0.995	0.998
EXACT QUANTILE		0.950	0.975	0.990	0.995	0.998
ESTIMATED VALUES	MEAN	0.948	0.973	0.988	0.993	0.996
	SD	0.00989	0.00718	0.00487	0.00355	0.00276

Table 5

p		0.950	0.975	0.990	0.995	0.998
EXACT QUANTILE		-0.0513	-0.0253	-0.0101	-0.00501	-0.00200
ESTIMATED VALUES	MEAN	-0.0517	-0.0257	-0.0106	-0.00570	-0.00276
	SD	0.00905	0.00636	0.00370	0.00258	0.00214

Appendix: Proof of Theorem 2

Note that the asymptotic solution of $W(\omega)=1/k$ is the same as the asymptotic solution of $kG(\omega)=2$ (see the proof of Lemma 1). Write now

$$h = (\omega-\omega_0)k^{-\alpha} \geq 0$$

and

$$X_i = \omega_0 - e^{-\alpha U_i}$$

where $\alpha = 1/\gamma$ and the U_i's have the joint density

$$h_0(u_1, \ldots, u_k) = \exp\left\{-\exp\{-u_k\} - \sum_{i=1}^{k} u_i\right\},$$

for $u_1 \geq \ldots \geq u_k$. Then

$$Y_i = -\ell n(\omega - X_i) = -\ell n(\omega - \omega_0 + e^{-\alpha U_i})$$

and

$$Y_i - Y_k = -\ell n \frac{(ke^{U_k})^{-\alpha} e^{-\alpha T_i} + h}{(ke^{U_k})^{-\alpha} + h}$$

where $T_i = U_i - U_k$. Following the same steps as in Ref. [4] (proof of consistency of $\hat{a}$), we obtain that

$$\lim_{k\to\infty} kG(\omega) = f(h)/[g(h)]^2 = \zeta(h),$$

say, almost surely, where

$$f(h) = \int_0^\infty e^{-x}\, \ell n^2\left(\frac{e^{-\alpha x}+h}{1+h}\right) dx$$

$$g(h) = \int_0^\infty e^{-x}\, \ell n\left(\frac{e^{-\alpha x}+h}{1+h}\right) dx.$$

Since $W(\omega)$ is a strictly monotone increasing function of ω, $kG(\omega)$ is a strictly monotone decreasing function, and thus $\zeta(h)$ must be non-increasing. Since $\zeta(0) = 2$ and, as will be shown, $\zeta'(0) < 0$, it follows that the equation $\zeta(h) = 2$ has a unique solution: $h = 0$. Hence the unique solution $\hat{\omega}$ of $kG(w) = 2$ will tend almost surely to ω_0.

We now show that $\zeta'(0) < 0$ for all $\alpha > 0$. By direct calculation we have the following:

$$f(0) = 2\alpha^2,$$

$$f'(h) = \int_0^\infty 2\left(\ell n(e^{-\alpha x}+h) - \ell n(1+h)\right)\left((e^{-\alpha x}+h)^{-1} - (1+h)^{-1}\right)e^{-x}dx,$$

$$g(0) = -\alpha,$$

$$g'(h) = \int_0^\infty \left((e^{-\alpha x}+h)^{-1} - (1+h)^{-1}\right)e^{-x}dx,$$

When $0 < \alpha < 1$,

$$f'(0) = -2\alpha\left((1-\alpha)^{-2} - 1\right),$$

$$g'(0) = (1-\alpha)^{-1} - 1,$$

$$\zeta'(0) = \frac{g^2(0)f'(0) - 2f(0)g(0)g'(0)}{g^4(0)}$$

$$= -2\alpha^3/\alpha^2(1-\alpha)^2 < 0.$$

When $\alpha = 1$, we have that

$$\zeta'(h) \sim f'(h) + 4g'(h), \text{ as } h \to 0.$$

It is easy to show that

$$f'(h) \sim -(\ell nh)^2 + 2(\ell nh)(\ell n(1+h)), \text{ as } h \to 0,$$

$$g'(h) \sim -\ell nh, \text{ as } h \to 0.$$

Hence

$$\zeta'(0) = \lim_{h\to 0} \zeta'(h) = -\infty$$

When $\alpha > 1$,

$$\alpha^2\zeta'(h) \sim f'(h) + 4\alpha g'(h).$$

It can be shown, by direct calculation, that

$$f'(h) \sim [2B+2h^{1-1/\alpha}\ell nh-2A\ell n(1+h^{-1})]/h^{1-1/\alpha},$$

as $h \to 0$, and

$$g'(h) \sim A/h^{1-1/\alpha},$$

as $h \to 0$, where A and B are some positive constants. Hence

$$\alpha^2\zeta'(h) \sim -2A\ell n(1 + 1/h)]/h^{1-1/\alpha}, \text{ as } h \to 0,$$

and

$$\zeta'(0) = \lim_{h\to 0} \zeta'(h) = -\infty. \quad \blacksquare$$

References

[1] Boos, D. D., "Using extreme value theory to estimate percentiles," *Technometrics*, Vol. 26, No. 1 (1984), 33-39.

[2] Galambos, J., "*The asymptotic theory of order statistics*," 2nd ed. Robert E. Krieger Publishing Co., Malaber Fla, 1982.

[3] Hasofer, A. M. and Wang, Z., "A test for extreme value domain of attraction," *J. Am. Stat. Ass.* Vol. 87, No. 417 (1992), 171-177.

[4] Hasofer, A. M. and Wang, Z., Estimating quantiles for a Type II domain of attraction," (Submitted for publication, 1992).

[5] Smith, R. L. and Weissman, I., "Maximum likelihood estimation of the lower tail of a probability distribution," *J. R. Statist. Soc.* B, 47, No. 2 (1985), 285-298.

[6] Weissman, I., "Estimation of parameters and large quantiles based on the *k* largest observations," *J. Am. Stat. Ass.* Vol. 73, No. 364 (1978), 812-815.

Extreme Values Of Monotonic Functions And Evaluation Of Catastrophic Flood Loss

Lambert, J.H., Li, D. and Haimes, Y.Y
University of Virginia, Charlotttesville, VA

A univariate monotonic function is often a useful model in engineering risk assessment, e.g., in relating the magnitude of flood discharge to the consequent economic losses. An approach is developed in this paper to determine the domain of attraction for a monotonic function of an underlying random variable. Using von Mises' criteria, sufficient conditions are derived to find the domain of attraction of a transformed variable for situations where only incomplete knowledge concerning the underlying random variable and the monotonic function is available. The sufficient conditions, along with relationships for the transformation of the extremal statistical parameters, lead to a practical methodology for estimating the expected loss, conditional on the exceedance of a threshold level or percentile. The conditional expected value can serve as a measure of the risk of extreme events, such as catastrophic floods.

INTRODUCTION

In engineering assessment of the risk of extremes, a univariate monotonic function is often a useful model relating an underlying random variable and the consequence. Knowledge of the tail of the probability distribution of the underlying random variable can be applied to derive the distribution of the system outcome. There is considerable literature on applications of the statistics of extremes in engineering, including Refs. [1]–[4]. However, it is difficult to quantify the risk associated with the extreme outcomes when the distribution of the underlying variable is uncertain due to scarce data and lack of knowledge of the physical process. Using statistics of extremes, this paper studies the tail of the distribution of a monotonic function of an underlying random variable whose exact form of probability distribution is unknown.

Background

Let f_X and F_X be the probability density function and cumulative distribution function, respectively, of an underlying random variable X. The largest sample value, X_n^{max}, from n independent observations of X is itself a random variable with the following cumulative distribution function

$$F_{X_n^{max}}(x) = [F_X(x)]^n \tag{1}$$

As n approaches to infinity, the distribution of the largest sample value from X usually converges to one of the three particular forms, or domains of attraction: the Gumbel, which is of a double exponential form; the Frechet, which is of an exponential form; and the Weibull, which is of an exponential form with an upper bound (Refs. [2], [3], and [5]). There exist in the literature (Refs. [3]–[6]) various forms of necessary and sufficient conditions for determining the domain of attraction of X. A simple set of sufficient conditions for determining the form of the asymptotic distribution is von Mises' criteria (Refs. [2] and [7]).

From (Ref. [2]), the distribution of the largest value from X converges to a Gumbel form if X is unlimited in the direction of the largest value and

$$\lim_{x\to\infty} \frac{d}{dx}\left[\frac{1 - F_X(x)}{f_X(x)}\right] = 0 \tag{2}$$

In the Appendix of this paper it is shown that Eq. (2) implies that

$$\lim_{x\to\infty} \frac{1 - F_X(x)}{x\, f_X(x)} = 0 \tag{3}$$

and that Eq. (3) implies Eq. (2) if the function $(1 - F_X)/f_X$ is a monotone function (increasing or decreasing) for large x. The monotonicity condition is satisfied by most, if not all, common distributions in the Gumbel domain of attraction. Following Ref. [2], we consider in this paper only distributions of Gumbel forms that are unlimited in the direction of the largest value.

From (Ref. [2]), the distribution of the largest value from X converges to a Frechet form if X is unlimited in the direction of the largest value and there exists a strictly positive constant k such that

$$\lim_{x\to\infty} \frac{x\, f_X(x)}{1 - F_X(x)} = k \quad k > 0 \tag{4}$$

From (Ref. [2]), the distribution of the largest value from X converges to a Weibull form if X has a finite upper bound w which satisfies $w = \sup\{x : F_X(x) < 1\}$ and there exists a strictly positive constant k such that

$$\lim_{x \to w} \frac{(w-x)f_X(x)}{1 - F_X(x)} = k \quad k > 0 \tag{5}$$

Two important parameters in the statistics of extremes are the characteristic largest value and the inverse measure of dispersion. For the underlying random variable X, the characteristic largest value, u_n^X, is defined in (Ref. [2]) by

$$F_X(u_n^X) = 1 - \frac{1}{n} \tag{6}$$

and the inverse measure of dispersion, δ_n^X, is defined in (Ref. [2]) by

$$\delta_n^X = n\, f_X(u_n^X) \tag{7}$$

or equivalently,

$$\frac{1}{\delta_n^X} = \frac{du_n^X}{d\ln(n)} \tag{8}$$

Ref. [2] shows that for many distributions of engineering interest, the characteristic largest value and inverse measure of dispersion as defined here are sufficient to parameterize each of the three extremal forms.

The hazard function $f_X(x)/[1 - F_X(x)]$ (Ref. [2]) is equal to the corresponding inverse measure of dispersion $\delta_n^X(x)$, where n is determined by Eq. (6) with u_n^X equal to x.

Objectives of the Paper

A strictly monotone increasing function represents situations in which the higher the realization of the underlying random variable, the higher the outcome. A function, Y = g(X), of an underlying random variable X is considered in the following, where g is assumed to be a strictly monotone increasing function. It is often possible from the observation data to determine for X a domain of attraction, the characteristic largest value, u_n^X, and the inverse measure of dispersion, δ_n^X, by means such as Gumbel's extremal probability paper and the method of moments or the method of order statistics (Refs. [2] and [3]). However, the exact form of the underlying distribution of X is probably never certain. The problem is to study the extremes of the function Y with only limited knowledge of the probabilistic description of X. In this direction, Ref. [3] demonstrates conditions under which a monotone transformation preserves the Weibull asymptote in the *smallest* value.

Sufficient conditions are derived in this paper to identify for the random variable Y its domain of attraction based on the von Mises criteria. Expressions are also obtained for the characteristic largest value, u_n^Y, and the inverse measure of dispersion, δ_n^Y. The derivation does not assume or require knowledge of the specific form of the distribution of X. These results are then used to evaluate a conditional expected value that is a measure of the risk associated with extreme events. The organization of the paper is as follows. Results in the form of nine sets of mapping routes and conditions are first derived that identify the domain of attraction of the function Y. Expressions are then developed to derive the characteristic largest value and the inverse measure of dispersion for the function Y. These results are used to assist in the calculation of the estimate of a conditional expected value of Y in the extremal range. An application of the developed method in assessing catastrophic flood losses is presented.

PRESERVATION AND TRANSFORMATION OF DOMAIN OF ATTRACTION

Preservation and Transformation of the Gumbel Domain of Attraction

Assume the distribution of the largest value from the initial variate X to be of the Gumbel asymptotic form. Sufficient conditions are derived here under which the asymptotic distribution of the largest value from Y is preserved in the Gumbel asymptotic form or is transformed to the forms of Frechet or Weibull.

Theorem 1: Assume Y = g(X) is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function Y = g(X) is of a Gumbel form if (i) $[1/\delta_n^X(x)]\,[dg(x)/dx]/g(x)$ is monotone for large x and its limit is 0 as x approaches infinity, or (ii) the limit of $d[\ln g(x)]/d(\ln x)$ is finite as x approaches infinity, X satisfies the condition in Eq. (3), and $[1/\delta_n^X(x)]\,[dg(x)/dx]/g(x)$ is monotone for large x.

Proof: We consider convergence criterion of the Gumbel domain of attraction in Eq. (3) for Y:

$$\begin{aligned}
&\lim_{y \to \infty} \frac{1 - F_Y(y)}{y\, f_Y(y)} \\
&= \lim_{y \to \infty} \frac{1 - F_X[g^{-1}(y)]}{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy} \\
&= \lim_{x \to \infty} \frac{1 - F_X(x)}{g(x)\, f_X(x)\, dx/dg(x)} \\
&= \lim_{x \to \infty} \frac{1 - F_X(x)}{x\, f_X(x)} \frac{x\, dg(x)/dx}{g(x)}
\end{aligned} \tag{9}$$

The above limit is equal to zero if either the limit of $[1/\delta_n^X(x)]\,[dg(x)/dx]/g(x)$ is 0 as x approaches infinity, or if $d[\ln g(x)]/d(\ln x)$ is finite as x approaches infinity when X satisfies the condition in Eq. (3). Then it can be concluded from Eq. (3) that the asymptotic distribution of the largest

value of the function Y is of a Gumbel form if $[1/\delta_n^X(x)]$ $[dg(x)/dx]/g(x)$ is monotone for large x. ◊

Theorem 2: Assume Y = g(X) is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function Y = g(X) is of a Frechet form if the limit of $\delta_n^X(x)\, g(x)/[dg(x)/dx]$ exists and is strictly positive as x approaches infinity.

Proof: We consider the von Mises convergence criterion of the Frechet domain of attraction for Y:

$$\lim_{y\to\infty} \frac{y\, f_Y(y)}{1 - F_Y(y)} = \lim_{y\to\infty} \frac{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy}{1 - F_X[g^{-1}(y)]} = \lim_{x\to\infty} \frac{f_X(x)}{1 - F_X(x)} \frac{g(x)}{dg(x)/dx} \qquad (10)$$

If g(x)/[dg(x)/dx] and $[1 - F_X(x)]/f_X(x)$ are of the same order of magnitude as x approaches infinity, then the limit in the last expression is strictly positive and it can be concluded from Eq. (4) that the asymptotic distribution of the largest value of the function Y is of a Frechet form. ◊

Remark. Note that if X satisfies the convergence criterion of the Gumbel form in Eq. (3), a necessary condition for Y = g(X) to satisfy the von Mises convergence criterion of Frechet is that the limit of d(lnx)/d[lng(x)] exists and is equal to 0 as x approaches infinity.

Theorem 3: The asymptotic distribution for the largest value of the function Y = g(X) is of a Weibull form if there exists a finite ω such that ω = sup{y: $F_Y(y) < 1$} and the limit of $\delta_n^X(x)[\omega - g(x)]/[dg(x)/dx]$ exists and is strictly positive as x approaches infinity.

Proof: We consider the von Mises convergence criterion of the Weibull domain of attraction for Y:

$$\lim_{y\to\omega} \frac{(\omega - y) f_Y(y)}{1 - F_Y(y)} = \lim_{y\to\omega} \frac{(\omega - y) f_X[g^{-1}(y)] dg^{-1}(y)/dy}{1 - F_X[g^{-1}(y)]} = \lim_{x\to\infty} \frac{f_X(x)}{1 - F_X(x)} [\omega - g(x)] \frac{dx}{dg(x)} \qquad (11)$$

If $\frac{\omega - g(x)}{dg(x)/dx}$ and $[1 - F_X(x)]/f_X(x)$ are of the same order of magnitude, then the limit in the last expression is strictly positive and we conclude from Eq. (5) that the largest value from Y is asymptotically of the Weibull form. ◊

Remark. Note that if X satisfies the convergence criterion of the Gumbel form in Eq. (3), a necessary condition for Y = g(X) to satisfy the von Mises convergence criterion of the Weibull form is that the limit of $[\omega - g(x)]/\{x[dg(x)/dx]\}$ exists and is equal to zero as x approaches infinity.

Preservation and Transformation of Frechet Domain of Attraction

Assume the distribution of the largest value from the initial variable X to be of the Frechet form and that F_X satisfies the von Mises convergence condition in Eq. (4). Sufficient conditions are derived here under which the asymptotic distribution of the largest value from Y is preserved as Frechet or is transformed to Gumbel or Weibull.

Theorem 4: Assume Y = g(X) is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function Y = g(X) is of a Gumbel form if the limit of d[ln g(x)]/d(ln x) exists and is equal to zero as x approaches infinity and $[1/\delta_n^X(x)][dg(x)/dx]/g(x)$ is monotone for large x.

Proof: We consider the convergence criterion of the Gumbel domain of attraction in Eq. (3) for Y:

$$\lim_{y\to\infty} \frac{1 - F_Y(y)}{y\, f_Y(y)} = \lim_{y\to\infty} \frac{1 - F_X[g^{-1}(y)]}{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy} = \lim_{x\to\infty} \frac{1 - F_X(x)}{g(x)\, f_X(x)\, dx/dg(x)} = \frac{1}{k} \lim_{x\to\infty} \frac{x\, dg(x)/dx}{g(x)} \qquad (12)$$

When the limit of d[ln g(x)]/d(ln x) is equal to zero as x approaches infinity, from Eq. (3) the distribution of the largest value from Y = g(X) is asymptotically of a Gumbel form if $[1/\delta_n^X(x)][dg(x)/dx]/g(x)$ is monotone for large x. ◊

Theorem 5: Assume Y = g(X) is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function Y = g(X) is of a Frechet form if the limit of d[ln g(x)]/d(ln x) exists and is strictly positive as x approaches infinity.

Proof: We consider the von Mises convergence criterion of the Frechet domain of attraction for Y:

$$\lim_{y\to\infty} \frac{y\, f_Y(y)}{1 - F_Y(y)} = \lim_{y\to\infty} \frac{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy}{1 - F_X[g^{-1}(y)]} = \lim_{x\to\infty} \frac{g(x)\, f_X(x)\, dx/dg(x)}{1 - F_X(x)} = k \lim_{x\to\infty} \frac{g(x)}{x\, dg(x)/dx} \qquad (13)$$

If the limit of d[ln g(x)]/d(ln x) exists and is strictly positive as x approaches infinity, from Eq. (4) the distribution of the largest value from Y = g(X) is asymptotically of a Frechet form. ◊

Theorem 6: The asymptotic distribution for the largest value of the function Y = g(X) is of a Weibull form if there exists a finite ω such that ω = sup{y : $F_Y(y) < 1$}

and the limit of $[dx/dg(x)][\omega - g(x)]/x$ exists and is strictly positive as x approaches infinity.

Proof: We consider the von Mises convergence criterion of the Weibull domain of attraction for Y:

$$\lim_{y\to\omega} \frac{(\omega - y)f_Y(y)}{1 - F_Y(y)}$$

$$= \lim_{y\to\omega} \frac{(\omega - y)f_X[g^{-1}(y)]dg^{-1}(y)/dy}{1 - F_X(g^{-1}(y))}$$

$$= \lim_{x\to\infty} \frac{x\, f_X(x)}{1 - F_X(x)} \frac{\omega - g(x)}{x} \frac{dx}{dg(x)}$$

$$= k \lim_{x\to\infty} \frac{\omega - g(x)}{x} \frac{dx}{dg(x)} \qquad (14)$$

If the limit of $[dx/dg(x)][\omega - g(x)]/x$ exists and is strictly positive as x approaches infinity, then we conclude from Eq. (5) that the distribution of the largest value from $Y = g(X)$ is asymptotically of the Weibull domain of attraction. ◊

Preservation and Transformation of the Weibull Domain of Attraction

Assume the random variable X to have an upper bound w such that $w = \sup\{x: F_X(x) < 1\}$, that the distribution of the largest value from X converges to a Weibull form, and that X satisfies the von Mises convergence condition in Eq. (5). Sufficient conditions are derived here under which the asymptotic distribution of the largest value from Y is preserved as Weibull or is transformed to Gumbel or Frechet.

Theorem 7: Assume $Y = g(X)$ is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function $Y = g(X)$ is of a Gumbel form if the limit of $[(w - x)/g(x)]\, dg(x)/dx$ exists and is equal to zero as x approaches w and $[1/\delta_n^X(x)][dg(x)/dx]/g(x)$ is monotone for x close to w.

Proof: We consider the convergence criterion of the Gumbel domain of attraction in Eq. (3) for Y:

$$\lim_{y\to\infty} \frac{1 - F_Y(y)}{y\, f_Y(y)}$$

$$= \lim_{y\to\infty} \frac{1 - F_X[g^{-1}(y)]}{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy}$$

$$= \lim_{x\to w} \frac{1 - F_X(x)}{g(x)\, f_X(x)\, dx/dg(x)}$$

$$= \lim_{x\to w} \frac{1 - F_X(x)}{(w - x)\, f(x)} \frac{(w - x)\, dg(x)/dx}{g(x)}$$

$$= \frac{1}{k} \lim_{x\to w} \frac{(w - x)}{g(x)} \frac{dg(x)}{dx} \qquad (15)$$

If the limit of $[(w - x)/g(x)]\, dg(x)/dx$ exists and is equal to zero as x approaches its upper bound w, then from Eq. (3) we conclude that the distribution of the largest value from $Y = g(X)$ is asymptotically of a Gumbel form if $[1/\delta_n^X(x)][dg(x)/dx]/g(x)$ is monotone for x close to w. ◊

Theorem 8: Assume that $Y = g(X)$ is unlimited in the direction of the largest value. The asymptotic distribution for the largest value of the function $Y = g(X)$ is of a Frechet form if the limit of $g(x)/[(w - x)\, dg(x)/dx]$ is strictly positive as x approaches its upper bound w.

Proof: We consider the von Mises convergence criterion of the Frechet domain of attraction for Y:

$$\lim_{y\to\infty} \frac{y\, f_Y(y)}{1 - F_Y(y)}$$

$$= \lim_{y\to\infty} \frac{y\, f_X[g^{-1}(y)]\, dg^{-1}(y)/dy}{1 - F_X[g^{-1}(y)]}$$

$$= \lim_{x\to w} \frac{g(x)\, f_X(x)\, dx/dg(x)}{1 - F_X(x)}$$

$$= \lim_{x\to w} \frac{(w - x)\, f_X(x)}{1 - F_X(x)} \frac{g(x)}{(w - x)\, dg(x)/dx}$$

$$= k \lim_{x\to w} \frac{g(x)}{(w - x)\, dg(x)/dx} \qquad (16)$$

If the limit of $g(x)/[(w - x)\, dg(x)/dx]$ exists and is strictly positive as x approaches its upper bound w, then from Eq. (4) the distribution of the largest value from $Y = g(X)$ is asymptotically of a Frechet form. ◊

Theorem 9: The asymptotic distribution for the largest value of the function $Y = g(X)$ is of a Weibull form if there exists a finite ω such that $\omega = \sup\{y : F_Y(y) < 1\}$ and the derivative $dg(x)/dx$ exists as x approaches w.

Proof: We consider the von Mises convergence criterion of the Weibull domain of attraction for Y:

$$\lim_{y\to\omega} \frac{(\omega - y)f_Y(y)}{1 - F_Y(y)}$$

$$= \lim_{y\to\omega} \frac{(\omega - y)f_X(g^{-1}(y))dg^{-1}(y)/dy}{1 - F_X(g^{-1}(y))}$$

$$= \lim_{x\to w} \frac{(w - x)\, f_X(x)}{1 - F_X(x)} \frac{\omega - g(x)}{(w - x)} \frac{dx}{dg(x)}$$

$$= k \lim_{x\to w} \frac{\omega - g(x)}{w - x} \frac{dx}{dg(x)}$$

$$= k \lim_{x\to w} \frac{dg(x)}{dx} \lim_{x\to w} \frac{dx}{dg(x)}$$

$$= k \qquad (17)$$

where l'Hôpital's rule is used in the second step from the last. It can be concluded from Eq. (5) that the distribution of the largest value from $Y = g(X)$ is asymptotically of a Weibull form. ◊

Summary of Conditions for Preservation and Transformation of the Domain of Attraction

We emphasize here that the exact form of the monotonic transformation is not always needed to obtain the domain of attraction of F_Y. Consider, for example, the cases where F_X is any distribution that satisfies the von

Mises convergence criteria to be of a Frechet domain of attraction, and g(X) is of the polynomial form, such that

$$g(X) = p(X) = \sum_{i=0}^{N} a_i X^i \tag{18}$$

and satisfying

$$d(\sum_{i=0}^{N} a_i x^i)/dx > 0 \tag{19}$$

Since d[ln p(x)]/d(ln x) is equal to N as x approaches infinity, the polynomial function p will always preserve the domain of attraction of the underlying variable X, by Theorem 5.

Theorems 5, 6, 8, and 9 require only knowledge of the transformation g when the von Mises condition is assumed to be satisfied by the underlying distribution F_X. The remaining theorems require knowledge of both g and δ_n^X where the second alternative (ii) of Theorem 1, Theorems 4 and 7 assume the satisfaction of the von Mises condition while the first alternative (i) of Theorem 1 and Theorems 2 and 3 do not necessarily assume satisfaction of the von Mises condition for the underlying random variable X. For engineering application, a table of the order of magnitude of the function $\delta_n^X(x)$ for large x for some common distributions is useful (Ref. [8]).

ESTIMATION OF THE CONDITIONAL EXPECTED VALUE

From the definition of the characteristic largest value of Y,

$$n[1 - F_Y(u_n^Y)] = 1 \tag{20}$$

u_n^Y can be derived as a function depending only on u_n^X (Ref. [9]),

$$u_n^Y = g(u_n^X) \tag{21}$$

From Eq. (7), δ_n^Y can be expressed as a function of both u_n^X and δ_n^X:

$$\begin{aligned}\delta_n^Y &= nf_Y(u_n^Y) \\ &= nf_X(u_n^X) \left.\frac{dx}{dg(x)}\right|_{u_n^X} \\ &= \delta_n^X / \left.\frac{dg(x)}{dx}\right|_{u_n^X}\end{aligned} \tag{22}$$

Note that the formulas for u_n^Y and δ_n^Y do not rely on knowledge of the exact form of the underlying distribution F_X.

Reference [10] proposes a conditional expectation $E[Y \mid Y > F_Y^{-1}(\alpha)]$ to represent the extreme risk realized from the tail of the distribution, where F_Y^{-1} is the inverse of the cumulative distribution function and α is a partitioning probability that is chosen to bound from below the range of extreme events. Building on this concept, Ref. [11] presents a result in approximating the conditional expectation based only on the knowledge of the domain of attraction, the characteristic largest value u_n^Y, and the inverse measure of dispersion δ_n^Y. The conditional expectation $E[Y = g(X) \mid Y > F_Y^{-1}(\alpha)]$ can be obtained based on the identified domain of attraction of F_Y without knowledge of the exact form of the probability distribution of X. These approximations are nearly exact for large values of n, and, equivalently, large values of the partition probability α, since there is the relationship

$$n = \frac{1}{1 - \alpha} \tag{23}$$

between the selected partitioning probability α and the corresponding value of n (Ref. [9]).

For a random variable Y of the Gumbel domain of attraction, unlimited in the direction of the largest value, and of an exponential tail (Ref. [2]), the approximation to $E[Y \mid Y > F_Y^{-1}(\alpha)]$ is given by (Ref. [11])

$$E[Y \mid Y > F_Y^{-1}(\alpha)] = u_n^Y + \frac{1}{\delta_n^Y} \tag{24}$$

For a random variable Y of the Frechet domain of attraction and of a polynomial tail (Ref. [2]), the approximation to $E[Y \mid Y > F_Y^{-1}(\alpha)]$ is given by (Ref. [11])

$$E[Y \mid Y > F_Y^{-1}(\alpha)] = u_n^Y + \frac{1}{\delta_n^Y} + (\frac{1}{\delta_n^Y})^2 (u_n^Y - \frac{1}{\delta_n^Y}) \tag{25}$$

For many (see the qualification in Ref. [2]) distributions of the Weibull domain of attraction, the approximation to $E[Y \mid Y > F_Y^{-1}(\alpha)]$ is given by (Ref. [11])

$$E[Y \mid Y > F_Y^{-1}(\alpha)] = u_n^Y + \frac{1}{\delta_n^Y} - \frac{\frac{1}{\delta_n^Y}}{[(\omega - u_n^Y)\delta_n^Y + 1]} \tag{26}$$

where ω is the upper limit of Y. These approximations are important in that no assumption of an exact distribution F_Y is needed.

EXAMPLE--FLOOD LOSSES

To illustrate the use of the derived results, we consider the evaluation of extreme monetary flood loss when the underlying probability distribution of peak discharges is uncertain. Denote X to be the peak discharge in unit of m^3/sec and assume that statistical analysis of the peak discharge record yields the following estimates of the characteristic largest value and the inverse measure of dispersion:

$$u_{100}^X = 6{,}000 \text{ m}^3/\text{sec} \tag{27}$$

$$\frac{1}{\delta_{100}^X} = 1{,}500 \text{ m}^3/\text{sec} \quad (28)$$

and that the distribution F_X of peak discharges could be of the Frechet domain of attraction. The stage–discharge relationship is assumed to be of the following form for large flows (Ref. [12]):

$$Y = g(X) = 3.92 \left(\frac{X}{10{,}000}\right)^{0.30} \quad \text{(m)}; \quad X \geq 1{,}000 \text{ m}^3/\text{sec} \quad (29)$$

where Y is the stage in unit of m. The stage–damage relationship h is assumed to be of the following form for high stages (Ref. [12]):

$$Z = h(Y) = 15{,}000{,}000 \left(1 - \frac{1.10}{Y}\right) \quad (\$); \quad Y \geq 1.10 \quad (30)$$

where Z is the monetary flood loss in unit of dollar.

Assume that the peak discharge F_X satisfies the von Mises criterion to be of the Frechet domain of attraction. It can be concluded from Theorem 5 that the Frechet asymptotic form is preserved in the stage Y since the stage–discharge function g is of the general form $Y = X^a$, where $a > 0$. Since the stage–damage function h is of the general form $\omega - c/Y$ with $c > 0$, it can be concluded from Theorem 6 that the Frechet asymptotic form in stage Y is transformed to a Weibull type asymptote in the loss Z. Although here the forms of the two monotonic functions are given exactly, the domain of attraction of the resulting functions $Y = g(X)$ and $Z = h(Y)$ could be obtained if they were known only by their general forms.

In two successive applications of Eqs. (21) and (22) the characteristic largest value and inverse measure of dispersion of the stage Y and the loss Z are obtained as follows:

$$u_{100}^Y = g(u_{100}^X) = 3.92 \left(\frac{6{,}000}{10{,}000}\right)^{0.30} = 3.36304 \text{ m} \quad (31)$$

$$\frac{1}{\delta_{100}^Y} = \frac{1}{\delta_{100}^X} \left.\frac{dg(x)}{dx}\right|_{u_{100}^X} = (1{,}500) \frac{3.92}{(10{,}000)^{0.30}} (0.30)(6{,}000)^{-0.70} = 0.252228 \text{ m} \quad (32)$$

$$u_{100}^Z = h(u_{100}^Y) = (15{,}000{,}000) \left(1 - \frac{1.10}{3.363}\right) = \$10{,}093{,}666 \quad (33)$$

$$\frac{1}{\delta_{100}^Z} = \frac{1}{\delta_{100}^Y} \left.\frac{dh(y)}{dy}\right|_{u_{100}^Y} = (0.252)(15{,}000{,}000) \frac{1.10}{(3.363)^2} = \$367{,}646.77 \quad (34)$$

These results enable us to use Eq. (26) for F_Z of the Weibull domain of attraction to calculate the conditional expected value with $\alpha = 0.99$ (and consequently $n = 100$) as follows

$$E[Z \mid Z > F_Z^{-1}(0.99)] = u_{100}^Z + \frac{1}{\delta_{100}^Z} - \frac{\frac{1}{\delta_{100}^Z}}{[(\omega - u_{100}^Z)\delta_{100}^Z + 1]}$$
$$= 10093666 + 367647 - \frac{367647}{[(15000000 - 10093666)/367647 + 1]}$$
$$= \$10{,}435{,}684 \quad (35)$$

The interpretation of this measure of extreme events is the expected flood loss conditional on either exceedance of the 99th percentile of loss or, equivalently, exceedance of the 100–year discharge. The conditional expected value of loss can be used in addition to the expected value of loss (and other criteria) for selecting an optimal design (Refs. [13] and [14]).

CONCLUSIONS

This paper has studied the characteristics of extreme realizations of a monotonic function of an underlying random variable with an unknown distribution. The results make possible analyzing the extreme values of a univariate monotonic function with limited knowledge of the underlying distribution and of the exact form of the function itself. Estimation of the conditional expected value of the system outcome, a measure of the risk of extreme events, has been demonstrated when exact knowledge of the underlying distribution is unavailable. The von Mises convergence criteria used as the basis for these results are sufficient conditions for determining the domain of attraction (Ang and Tang 1984). One future extension of this research is to derive rules governing the transformation and preservation of domains of attraction using conditions that are both necessary and sufficient.

ACKNOWLEDGMENTS

The authors are grateful to Professor Loren Pitt for assistance with the derivation in the Appendix and to Professor Enrique Castillo for consultation during the review of literature. This research was supported in part by the National Science Foundation under grant No. BCS–8912630.

APPENDIX. EQUIVALENCE OF TWO CONVERGENCE CRITERIA FOR GUMBEL DOMAIN OF ATTRACTION

Assume that X is unlimited in the direction of the largest value and define

$$H(x) = \frac{1 - F_X(x)}{f_X(x)} \quad (36)$$

where $F_X(x)$ and $f_X(x)$ are the cumulative distribution function and the probability density function of X, respectively.

Theorem A1: If H(x) is continuous, differentiable, and monotone for large x, then

$$\lim_{x \to \infty} \frac{H(x)}{x} = 0 \quad (37)$$

implies that

$$\lim_{x\to\infty} \frac{d}{dx} H(x) = 0 \qquad (38)$$

Proof: Note that H(x) is always nonnegative. We have

$$H(2x) = H(x) + \int_x^{2x} H'(\xi)\, d\xi \geq \int_x^{2x} H'(\xi)\, d\xi \qquad (39)$$

It follows from Eq. (39) and the mean value theorem that there exists some $y \in [x,2x]$ such that

$$H(2x) \geq (2x - x)\, H'(y) \qquad (40)$$

In the limit as x approaches infinity, dividing Eq. (40) by 2x yields

$$\lim_{x\to\infty} \frac{H(2x)}{2x} \geq \lim_{x\to\infty} \frac{H'(y)}{2} \qquad y \in [x,2x] \qquad (41)$$

If $H'(x) \geq 0$ for large x, then Eq. (38) follows from Eqs. (37) and (41).

It remains to consider the case of $H'(x) \leq 0$ for large x. Rearranging Eq. (39), we have

$$H(x) = H(2x) - \int_x^{2x} H'(\xi)\, d\xi \geq \int_x^{2x} -[H'(\xi)]\, d\xi \qquad (42)$$

It follows from Eq. (42) and the mean value theorem that there exists some $y \in [x,2x]$ such that

$$H(x) \geq -(2x - x)\, H'(y) \qquad (43)$$

In the limit as x approaches infinity, dividing Eq. (43) by x yields

$$\lim_{x\to\infty} \frac{H(x)}{x} \geq \lim_{x\to\infty} -H'(y) \qquad y \in [x,2x] \qquad (44)$$

If $H'(x) \leq 0$ for large x, then Eqs. (37) and (44) lead to Eq. (38). ◊

Theorem A2: If H(x) is continuous and differentiable for large x, then

$$\lim_{x\to\infty} \frac{d}{dx} H(x) = 0 \qquad (45)$$

implies that

$$\lim_{x\to\infty} \frac{H(x)}{x} = 0 \qquad (46)$$

Proof: If Eq. (45) holds, then for any $\varepsilon > 0$, there exists a number N such that $|H'(x)| < \varepsilon$ for $x \geq N$. We have from the triangle inequality

$$\frac{H(x)}{x} = [H(N) + \int_N^x H'(\xi)\, d\xi]/x$$

$$\leq |H(N)|/x + \left| \int_N^x H'(\xi)\, d\xi \right| /x$$

$$\leq |H(N)|/x + \varepsilon\,(x - N)/x \qquad (47)$$

Taking the limit of Eq. (47) as x approaches infinity gives

$$\lim_{x\to\infty} \frac{H(x)}{x} \leq \varepsilon \qquad (48)$$

Since ε can be chosen arbitrarily small (N arbitrarily large), the Eq. (46) follows immediately from Eq. (48). ◊

REFERENCES

[1] Gumbel, E.J., Statistics of Extremes, Columbia University Press, New York, 1958.

[2] Ang, A. H–S. and W. H. Tang, Probability Concepts in Engineering Planning and Design, Volume II: Decision, Risk, and Reliability, John Wiley and Sons, New York, 1984.

[3] Castillo, E., Extreme Value Theory in Engineering, Academic Press, Boston, 1988.

[4] Johnson, N.L. and S. Kotz, Continuous Univariate Distributions, John Wiley and Sons, New York, 1970.

[5] Galambos, J., The Asymptotic Theory of Extreme Order Statistics, 2nd edition, John Wiley and Sons, New York, 1987.

[6] de Haan, L., On Regular Variation and its Application to the Weak Convergence of Sample Extremes, Mathematical Centre Tracts, Amsterdam, Vol. 32, 1970.

[7] von Mises, R. von, La distribution de las plus grande de n valeurs, Mathematique, Interbalkanique. 1 (1936), 141–160.

[8] Lambert, J.H. and D. Li, Evaluating risk of extreme events: results for monotonic functions, To appear in Journal of Water Resources Planning and Management, May 1994.

[9] Karlsson, P–O. and Y.Y. Haimes, Risk–based analysis of extreme events, Water Resources Research, 24 (1988), 9–20.

[10] Asbeck, E.L. and Y.Y. Haimes, The partitioned multiobjective risk method, Large Scale Systems, 6 (1984), 13–38.

[11] Mitsiopoulis, J.A., Y.Y. Haimes, and D. Li, Approximating catastrophic risk through statistics of extremes, Water Resources Research, 27 (1991), 1223–1230.

[12] Haimes, Y.Y., D. Li, and V. Tulsiani. Integration of Structural Measures and Flood Warning Systems for Flood

Damage Reduction, Report to The Institute of Water Resources, U.S. Army Corps of Engineers, Fort Belvoir, Virginia, Prepared by Environmental Systems Modeling, Inc., Charlottesville, Virginia, 1992.

[13] Karlsson, P–O. and Y.Y. Haimes, Risk assessment of extreme events: An application, Journal of Water Resources Planning and Management, 115 (1989), 299–320.

[14] Haimes, Y.Y., J.H. Lambert, and D. Li, Risk of extreme events in a multiobjective framework, Water Resources Bulletin. 28 (1992), 201–209.

Second Order Behavior Of Domains Of Attraction And The Bias Of Generalized Pickands' Estimator

Pereira, T.T.
University of Lisbon, Lisbon, Portugal

The domain of attraction of the generalized extreme value distribution is studied with respect to second order conditions using the tail quantile function. The particularly important case of the differentiable domain of attraction is emphasized. Next the weak consistency of a generalization of Pickands' estimator for the main parameter of an extreme value distribution is proved. Moreover, under quite general conditions on the underlying distribution function, that include second order behaviour and being in the differentiable domain of attraction, the asymptotic normality of the estimator is proved and the asymptotic bias that can occur is determined. A result concerning the minimization of the asymptotic mean squared error of the estimator is given which leads to an optimal choice of the number of intermediate upper order statistics involved in the definition of the estimator. Several examples, including all the usual continuous distributions, illustrate the results. Suggestions on how to choose the parameters in practical applications are made.

1. Introduction

The classical extreme value theory is primarily concerned with the asymptotic distribution of the maximum of independent and identically distributed (i.i.d.) random variables. Let X_i, $i\geq 1$, be a sequence of i.i.d. random variables with distribution function F and let $X_{n:n}=\max(X_1,\ldots,X_n)$, $n\geq 1$. Gnedenko, Ref. [1], proved that, if there exist sequences of real constants a_n and b_n, $n\geq 1$, with $a_n>0$ and a nondegenerate distribution function G such that

$$\lim_{n\to\infty} P((X_{n:n}-b_n)/a_n \leq x) = \lim_{n\to\infty} F^n(a_n x+b_n) = G(x) \quad (1.1)$$

for all x at which G(x) is continuous, then G(x) belongs to one of the three types of extreme value distribution functions. Using a parametrization of von Misès, Ref. [2], these limiting types can be written in a unified way, known as Generalized Extreme Value (GEV) distribution,

$$G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\},\ 1+\gamma x >0, \quad \gamma\in\mathbb{R}.$$

We say that F belongs to the domain of attraction of G_γ, notation $F\in D(G_\gamma)$, if (1.1) holds for some sequences a_n and b_n. The characterization of the domains of attraction of extreme value distributions has been dealt with by several authors. Von Misès, Ref. [2], found sufficient conditions for a distribution function to be in the domain of attraction of each of the three extreme value distributions and Gnedenko, Ref. [1], gave necessary and sufficient conditions. More recently, de Haan, Ref. [3], and Pickands, Ref. [4], presented a unified characterization of the domain of attraction of the GEV distribution. Let the function U be defined by $U(x) = (1/(1-F))^{\leftarrow}(x)$, $x\geq 1$, where the arrow denotes the (generalized) inverse function. Note that $U(x) = Q(1/x)$ with Q the quantile function of the upper tail of F. A necessary and sufficient condition for a distribution function F to be in the domain of attraction of G_γ, for some $\gamma\in\mathbb{R}$, is the existence of a positive function a(.) such that, for $x>0$,

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{a(t)}=\frac{x^{\gamma}-1}{\gamma}\quad\text{(read logx, if }\gamma=0) \qquad (1.2)$$

(Ref. [3]). Moreover the auxiliar function a(.) is regularly varying with index γ, i.e., $a\in RV_{\gamma}$ (cf. Ref. [5], th.1.9.). Pickands, Ref. [4], considered the inverse of the hazard cumulative function, $H^{-1}(x)=(1/(1-F))^{\leftarrow}(e^x)=U(e^x)$, to give a characterization of the domain of attraction which is easily seen, through a logarithmic transformation, to be equivalent to the one of de Haan. The necessary and sufficient condition (1.2) for the domain of attraction of G_{γ} can of course be written

$$\frac{U(tx)-U(t)}{a(t)}=\frac{x^{\gamma}-1}{\gamma}+R_{\gamma,x}(t)\quad\text{with } R_{\gamma,x}(t)=o(1),\ t\to\infty. \qquad (1.3)$$

If there exists a positive function R(t) such that R(t)=o(1), $t\to\infty$, and $R_{\gamma,x}(t)=h_{\gamma}(x)R(t)+o(R(t))$, we can speak of second order behaviour of U. We then have for x>0,

$$\lim_{t\to\infty}\frac{U(tx)-U(t)-a(t)\frac{x^{\gamma}-1}{\gamma}}{a(t)R(t)}=h_{\gamma}(x) \qquad (1.4)$$

and our purpose in section 3 is to find the possible limit functions $h_{\gamma}(x)$ with $h_{\gamma}(x)$ finite and not constant. Besides we will find the function R(t) which describes the rate of convergence of the limit (1.4) to be a regularly varying function with index ρ for some $\rho\le 0$. We also analyse the second order behaviour of U for a distribution function in the differentiable domain of attraction of G_{γ}, and we identify the generalized Pareto distribution as the only family of distribution functions with positive derivative in this subdomain of attraction for which $R_{\gamma,x}(t)$ is identically zero. Related papers are Ref. [6], Ref. [7] and Ref. [8].

Section 4 deals with the problem of estimating the extreme value index γ from a finite sample $X_1,X_2,\ldots,X_n$. A semi-parametric approach is due to Pickands, Ref. [9]; he proposed the estimator

$$\hat{\gamma}_n^{P}=(\log 2)^{-1}\log\frac{Z_m^{(n)}-Z_{2m}^{(n)}}{Z_{2m}^{(n)}-Z_{4m}^{(n)}},\quad 1\le m\le[n/4]$$

where $Z_1^{(n)}\ge Z_2^{(n)}\ge\ldots\ge Z_n^{(n)}$ are the descending order statistics of $X_1,X_2,\ldots,X_n$ and m=m(n) is an intermediate sequence of integers, i.e., $m\to\infty$ and $m/n\to 0$ ($n\to\infty$). Pickands proved that his estimator (based on the 4m largest observations) is weakly consistent and Dekkers and de Haan (1989) proved this same result and showed that if the sequence m(n) increases suitably rapidly the estimator is also strong consistent. Moreover, under additional conditions on the distribution (differentiability and Π-variation of $\pm t^{1-\gamma}U'(t)$ for real γ or second order regular variation conditions on U for $\gamma\neq 0$), they proved the asymptotic normality of the estimator for intermediate sequences m(n) which increase at certain rates. Note that with a reparametrization the estimator can be written

$$\hat{\gamma}_n^{P}=(\log 2)^{-1}\log\frac{Z_{[m/4]}^{(n)}-Z_{[m/2]}^{(n)}}{Z_{[m/2]}^{(n)}-Z_m^{(n)}}$$

based on the m largest observations. We consider the following generalization of Pickands' estimator

$$\hat{\gamma}_{n,\theta}^{P}=(-\log\theta)^{-1}\log\frac{Z_{[m\theta^2]}^{(n)}-Z_{[m\theta]}^{(n)}}{Z_{[m\theta]}^{(n)}-Z_m^{(n)}},\quad 0<\theta<1,$$

where we adopt the convention $[x]=\begin{cases}1 & ,\ 0<x\le 1\\ \text{largest integer}\le x & ,\ x\ge 1\end{cases}$. Note that $\hat{\gamma}_n^{P}=\hat{\gamma}_{n,1/2}^{P}$.

This estimator is weakly consistent for any θ in the interval]0,1[and for any intermediate sequence m(n). Under quite general conditions on F (F in the differentiable domain of attraction of G_{γ} and second order behaviour of U), we shall prove the asymptotic normality of $\hat{\gamma}_{n,\theta}^{P}$ for a certain rate of growth of the intermediate sequence. These conditions are more general than the ones considered in Ref. [7]; namely, they include the case of second order regular variation behaviour for $\gamma=0$ which was not taken into account in Ref. [7]. Also our proofs do not use the arguments in Ref. [7] but are based on the asymptotic joint distribution of a fixed number of intermediate order statistics associated to the same intermediate sequence established by Cooil, Ref. [10] and Ref. [11]. Moreover we shall give the asymptotic bias of the estimator that occurs if the sequence m(n) is allowed to increase at a

faster rate, as well as a theoretical result concerning the minimization of the asymptotic mean squared error of generalized Pickands' estimator.

For a review of the different approaches (parametric and semi-parametric) to the estimation of the tail of a distribution see Ref. [12].

In section 5 we will ilustrate the results with the continuous distribution functions that are typically used in statistical applications.

Finally, in section 6 we make some considerations on how to choose the parameters θ and m in pratical applications, and we suggest the use of the particular value of θ (corresponding to one of the possible generalizations of Pickands' estimator) which minimizes the variance of the estimator when $\gamma=0$.

2. Preliminary results

The importance of regular variation theory in the characterization of domains of attraction is shown in the next theorem, where $U(\infty) = \lim_{t\to\infty} U(t)$ and $x_\infty = \sup\{x: F(x)<1\}$.

THEOREM 2.1.(Gnedenko, Ref. [1]; de Haan, Ref. [3])
(1) For $\gamma>0$ are equivalent: (i) $F\in D(G_\gamma)$, (ii) $U(\infty)=\infty$ and $U\in RV_\gamma$, (iii) $x_\infty=\infty$ and $1-F\in RV_{-1/\gamma}$.
(2) For $\gamma<0$ are equivalent: (i) $F\in D(G_\gamma)$, (ii) $U(\infty)<\infty$ and $U(\infty)-U(x)\in RV_\gamma$, (iii) $x_\infty<\infty$ and $1-F(x_\infty-x^{-1})\in RV_{1/\gamma}$.
(3) For $\gamma=0$ are equivalent: (i) $F\in D(G_0)$, (ii) there exists a positive function f such that for real x, $\lim_{t\uparrow x_\infty} \frac{1-F(t+xf(t))}{1-F(t)} = e^{-x}$ (i.e., $1/(1-F)\in\Gamma(f)$), (iii) $U\in\Pi(a)$.

When $F\in D(G_\gamma)$ with $\gamma\neq0$, we say that F has a first order regular variation tail behaviour and that U has a first order regular variation behaviour. When $F\in D(G_0)$, we say that U has a first order Π–variation behaviour; moreover, the functions a(.) and f(.) are related by $a(t) = f(U(t))$.

The differentiable domains of attraction were introduced by Pickands, Ref. [4]. We say that F belongs to the (once) differentiable domain of attraction of G_γ, notation $F\in D_{dif}(G_\gamma)$, if the distribution function F is differentiable in a left neighborhood of x_∞ and if there exist sequences $a_n>0$ and b_n such that

$$\lim_{n\to\infty} \frac{d}{dx}[F^n(a_nx+b_n)] = (G_\gamma)'(x) \qquad (2.1)$$

uniformly for all x in any finite interval. Clearly, $F\in D_{dif}(G_\gamma)$ implies $F\in D(G_\gamma)$ for the same attraction coefficients $a_n>0$ and b_n. Condition (2.1) can also be written

$$\lim_{n\to\infty} na_nF'(a_nx+b_n) = (1+\gamma x)^{-1/\gamma-1}$$

and, in particular for x=0 we have $\lim_{n\to\infty} na_nF'(b_n) = 1$, that is, $a_n \sim 1/(nF'(b_n))$. If F' is positive and we take $b_n=U(n)$ this allows us to consider $a_n = nU'(n)$. The following theorem characterizes the differentiable domain of attraction of G_γ.

THEOREM 2.2. (Pickands, Ref. [4]) A distribution function $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ if and only if $H^{-1}(t)$ is differentiable for all sufficiently large t and, for real x,

$$\lim_{t\to\infty} \frac{(H^{-1})'(t+x)}{(H^{-1})'(t)} = e^{\gamma x}. \qquad (2.2)$$

It is easily verified that condition (2.2) is equivalent to $(H^{-1})'(\log t)$ being a γ–regularly varying function and, as $H^{-1}(t) = U(e^t)$, (2.2) can be written $tU'(t)\in RV_\gamma$. Hence, the above theorem can be restated in terms of the tail quantile function U.

COROLLARY 2.1. $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ if and only if U(t) is differentiable for all sufficiently large t and $t^{1-\gamma}U'(t)\in RV_0$.

Cooil, Ref. [10] and [11], proved that if $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ then for any intermediate sequence m=m(n) and $\theta>0$, the stochastic process

$$\omega_m^{(n)}(\theta) = \sqrt{m}\,\frac{Z_{[m\theta]}^{(n)} - b_{n/m\theta}}{a_{n/m}},$$

where $b_{n/m\theta} = F^{\leftarrow}(1-m\theta/n) = U(n/m\theta)$, converges in distribution, $n\to\infty$, to the gaussian stochastic process $\omega(\theta)$ characterized by

$$E(\omega(\theta)) = 0, \quad \theta>0,$$
$$Cov(\omega(\theta_1),\omega(\theta_2)) = \theta_1^{-\gamma}\theta_2^{-\gamma-1}, \quad 0<\theta_1\le\theta_2.$$

From now on we shall consider the normalizing sequence $b_n = U(n)$.

3. Second order conditions for domains of attraction

The next theorem identifies the generalized Pareto distribution, Ref. [9], as the only family of distribution functions belonging to $D_{dif}(G_\gamma)$ and having a positive derivative for which $R_{\gamma,x}(t)$ can be identically zero.

THEOREM 3.1. Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and suppose F has a positive derivative F'. Then are equivalent:

(i) there exists a positive function a(t) such that, for x>0,

$$\frac{U(tx)-U(t)}{a(t)}=\frac{x^\gamma-1}{\gamma} \text{ for sufficiently large t}$$

(i.e., $R_{\gamma,x}(t)\equiv 0$).

(ii) F is the generalized Pareto distribution function

$$F(x)=\left(1+\gamma\frac{x-c_2}{c_1}\right)^{-1/\gamma},\ x\geq c_2,\ 1+\gamma\frac{x-c_2}{c_1}>0,\ c_1>0,\ c_2\in\mathbb{R}.$$

PROOF. The existence of a positive derivative U' of U allows us to say that (i) is equivalent to (i') "there exists a positive function a(t) such that $tU'(tx)/a(t)=x^{\gamma-1}$ for all x>0" and taking x=1 in (i') we obtain a(t) = tU'(t).So, the functional equation to be solved is $U'(tx)/U'(t))=x^{\gamma-1}$ for x>0 and $t\geq\max\{1,1/x\}$. For $x\geq 1$ we can take in particular t=1 which leads to $U'(x)=x^{\gamma-1}U'(1)$, $x\geq 1$, which is equivalent to $U(x)=c_1(x^\gamma-1)/\gamma+c_2$ for some $c_1>0$, $c_2\in\mathbb{R}$. For $0<x\leq 1$ we can take in particular t=1/x and get $U'(1/x)=U'(1)x^{1-\gamma}$, $0<x\leq 1$, which is equivalent to $U'(x)=x^{\gamma-1}U'(1)$, $x\geq 1$. ♦

Note that for any other function a(t) asymptotically equal to tU'(t) the generalized Pareto distribution verifies (1.3) with $R_{\gamma,x}(t)\not\equiv 0$. What the theorem implies is that $R_{\gamma,x}(t)\not\equiv 0$ for any other distribution function $F\in D_{dif}(G_\gamma)$ with positive derivative, whatever the possible function a(t) considered. As can be seen in the proof of the theorem, (i) is also equivalent to $t^{1-\gamma}U'(t)\equiv c_1$, $c_1>0$.

LEMMA 3.1. Let $F\in D(G_\gamma)$ for some $\gamma\in\mathbb{R}$. If there exists a positive function b(t) such that, for x>0,

$$\lim_{t\to\infty}\frac{(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)}{b(t)}=h_\gamma^*(x) \qquad (3.1)$$

with $h_\gamma^*(x)$ finite and not constant, then

$$h_\gamma^*(x)=c\,\frac{x^\rho-1}{\rho},\ x>0,\ \text{for some } c\neq 0,\ \rho\leq 0, \qquad (3.2)$$

the case $\rho<0$ being possible only if $\lim_{t\to\infty}t^{-\gamma}a(t)=d>0$. Moreover $b(t)\in RV_\rho$ and $b(t)=o(t^{-\gamma}a(t))$.

PROOF. Remember that $F\in D(G_\gamma)$ implies $t^{-\gamma}a(t)\in RV_0$. The finite and not constant limit function of $[(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)]/b(t)$ is $c(x^\rho-1)/\rho$ for some $\rho\in\mathbb{R}$ and $c\neq 0$ (cf. th. 1.9, Ref.[5]). But $\rho>0$ implies $t^{-\gamma}a(t)\in RV_\rho$ (cf. th. 1.10, Ref.[5]) which contradicts the fact of $F\in D(G_\gamma)$. If $\rho<0$, then $\lim_{t\to\infty}t^{-\gamma}a(t)$ exists and moreover $\pm((\lim_{t\to\infty}t^{-\gamma}a(t))-t^{-\gamma}a(t))\in RV_\rho$, with the plus sign if c>0 and the minus sign if c<0. Now, if $\lim_{t\to\infty}t^{-\gamma}a(t)=0$ we will have $t^{-\gamma}a(t)\in RV_\rho$ with $\rho<0$ and, again, this is not possible; if $\lim_{t\to\infty}t^{-\gamma}a(t)=d>0$, then $\pm(d-t^{-\gamma}a(t))\in RV_\rho$ which implies that $t^{-\gamma}a(t)\in RV_0$ and only this last situation is possible if $\rho<0$ and for a distribution function F in the conditions of the theorem. ♦

Note that if (3.2) holds with $\rho=0$, $\pm t^{-\gamma}a(t)$ belongs to the class Π ($\pm t^{-\gamma}a(t)\in\Pi(a_1)$ with $a_1(t)=|c|b(t)$) and if it holds with $\rho<0$, $t^{-\gamma}a(t)$ is a slowly varying function but not a Π-varying one and $\pm(d-t^{-\gamma}a(t))\in RV_\rho$. A particularly important case of the above lemma is the one of $F\in D_{dif}(G_\gamma)$ with U having a positive derivative U' and a(t) = tU'(t): if $\rho=0$ we will have $\pm t^{1-\gamma}U'(t)\in\Pi$ and if $\rho<0$ then $\lim_{t\to\infty}t^{1-\gamma}U'(t)$ exists and is positive and $\pm((\lim_{t\to\infty}t^{1-\gamma}U'(t))-t^{1-\gamma}U'(t))\in RV_\rho$. The following theorem shows that U will then have a second order behaviour and identifies the function R(t) and the possible limiting functions $h_\gamma(x)$.

THEOREM 3.2. Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and suppose that U admits a positive derivative U'. If (3.1) holds for a(t) = tU'(t), then for x>0,

$$\lim_{t\to\infty}\frac{U(tx)-U(t)-tU'(t)\frac{x^\gamma-1}{\gamma}}{t^\gamma b(t)}=h_\gamma(x)$$

$$\text{with } h_\gamma(x)=\begin{cases}\frac{c}{2}\log^2 x & \gamma=0 \quad \rho=0\\ \frac{c}{\gamma}\left[x^\gamma\log x-\frac{x^\gamma-1}{\gamma}\right] & \gamma\neq 0 \quad \rho=0\\ \frac{c}{\rho}\left[\frac{x^{\gamma+\rho}-1}{\gamma+\rho}-\frac{x^\gamma-1}{\gamma}\right] & \gamma\in\mathbb{R} \quad \rho<0\end{cases}$$

for some $c\neq 0$.

PROOF. For $x>0$,

$$\frac{U(tx)-U(t)-tU'(t)\frac{x^\gamma-1}{\gamma}}{t^\gamma b(t)} = \int_1^x \frac{(ts)^{1-\gamma}U'(ts)-t^{1-\gamma}U'(t)}{b(t)}\, s^{\gamma-1}ds.$$

If $\rho=0$ this integral converges to $c\int_1^x s^{\gamma-1}\log s\,ds$, $t\to\infty$ (th. 1.14., Ref. [5]). If $\rho<0$ we easily see that $\lim_{t\to\infty}(t^{1-\gamma}U'(t)-d)/b(t) = c/\rho$ and as the above integral is equal to $\frac{t^{1-\gamma}U'(t)-d}{b(t)}\int_1^x\left(\frac{(ts)^{1-\gamma}U'(ts)-d}{t^{1-\gamma}U'(t)-d}-1\right)s^{\gamma-1}ds$ it converges to $\frac{c}{\rho}\int_1^x(s^\rho-1)s^{\gamma-1}ds$, $t\to\infty$ (th. 1.3., Ref. [5]). ♦

For $\rho=\gamma=0$ the converse statement is also true (cf. Ref. [6]).

In a wider context a related result concerning the possible limiting functions can be obtained if we assume a second order behaviour for U. Again the important limit will be the one of (3.1). A similar result was obtained in Ref. [8] for the function logU.

THEOREM 3.3. Let $F\in D(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and suppose that (1.4) holds with $h_\gamma(x)$ finite and not constant. If, in adittion, one of the following conditions holds

(a) $\lim_{t\to\infty}\frac{(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)}{t^{-\gamma}a(t)R(t)}$ exists for all $x>0$ and the limit function is not constant,

(b) $\lim_{t\to\infty}\frac{(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)}{t^{-\gamma}a(t)R(t)}=0$, R is measurable and $\lim_{t\to\infty}\frac{R(tx)}{R(t)}$ exists and is finite for all $x>0$,

then $h_\gamma(x)$ belongs to one of the following classes

$$\begin{cases} c_1\frac{\log^2 x}{2}+c_2\log x, & \gamma=0,\ \rho=0,\ (1)\\ \frac{c_1}{\rho}\left(\frac{x^\rho-1}{\rho}-\log x\right)+c_2\frac{x^\rho-1}{\rho}, & \gamma=0,\ \rho<0,\ (2)\\ \frac{c_1}{\gamma}\left(x^\gamma\log x-\frac{x^\gamma-1}{\gamma}\right)+c_2\frac{x^\gamma-1}{\gamma}, & \gamma\neq0,\ \rho=0,\ (3)\\ \frac{c_1}{\rho}\left(\frac{x^{\rho+\gamma}-1}{\rho+\gamma}-\frac{x^\gamma-1}{\gamma}\right)+c_2\frac{x^{\rho+\gamma}-1}{\rho+\gamma}, & \gamma\neq0,\ \rho<0,\ (4)\end{cases} \tag{3.3}$$

with $c_1\neq0$ and $c_2\in\mathbb{R}$ or $c_1=0$ and $c_2\neq0$.

PROOF. For $x,y>0$,

$$\frac{U(txy)-U(t)-a(t)\frac{(xy)^\gamma-1}{\gamma}}{a(t)R(t)} =$$

$$= \frac{U(txy)-U(tx)-a(tx)\frac{y^\gamma-1}{\gamma}}{a(tx)R(tx)}\frac{a(tx)R(tx)}{a(t)R(t)} +$$

$$+\frac{U(tx)-U(t)-a(t)\frac{x^\gamma-1}{\gamma}}{a(t)R(t)} + x^\gamma\frac{y^\gamma-1}{\gamma}\frac{(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)}{t^{-\gamma}a(t)R(t)} \tag{3.4}$$

(a) If the limit of $\frac{(tx)^{-\gamma}a(tx)-t^{-\gamma}a(t)}{t^{-\gamma}a(t)R(t)}$, as $t\to\infty$, exists for all $x>0$ and the limit function is not constant, it will be of the form $c_1(x^\rho-1)/\rho$, for some $c_1\neq0$ and $\rho\leq0$ ($\rho<0$ only if $\lim t^{-\gamma}a(t)=d>0$, $t\to\infty$) and $t^{-\gamma}a(t)R(t)$ is a ρ–regularly varying function (Lemma 3.1). Hence $R(t)\in RV_\rho$. Taking limits, as $t\to\infty$, in (3.4) one obtains the following functional equation

$$h_\gamma(xy) = h_\gamma(y)x^{\gamma+\rho}+h_\gamma(x)+c_1x^\gamma\frac{y^\gamma-1}{\gamma}\frac{x^\rho-1}{\rho} \tag{3.5}$$

Assuming that $h_\gamma(x)$ is a differentiable function, it follows by differentiation with respect to y that $xh_\gamma'(xy) = h_\gamma'(y)x^{\gamma+\rho}+c_1x^\gamma\frac{x^\rho-1}{\rho}y^{\gamma-1}$ and setting $y=1$ yields $h_\gamma'(x) = h_\gamma'(1)x^{\gamma+\rho-1}+c_1x^{\gamma-1}(x^\rho-1)/\rho$ which, noting that $h_\gamma(1)=0$ and $h_\gamma'(1)\in\mathbb{R}$, leads to

$$h_\gamma^1(x)=\begin{cases} c_1\frac{\log^2 x}{2}+c_2\log x & \gamma=0,\ \rho=0\\ \left(\frac{c_1}{\rho}+c_2\right)\frac{x^\rho-1}{\rho}-\frac{c_1}{\rho}\log x, & \gamma=0,\ \rho<0\\ \left(c_2-\frac{c_1}{\gamma}\right)\frac{x^\gamma-1}{\gamma}+\frac{c_1}{\gamma}x^\gamma\log x, & \gamma\neq0,\ \rho=0\\ \left(\frac{c_1}{\rho}+c_2\right)\frac{x^{\rho+\gamma}-1}{\rho+\gamma}-\frac{c_1}{\rho}\frac{x^\gamma-1}{\gamma}, & \gamma\neq0,\ \rho<0\end{cases}$$

with $c_1\neq0$ and $c_2\in\mathbb{R}$. Let $h_\gamma^r(x)$ be the difference between the general solution of the equation (3.5) and the differentiable one, i.e., $h_\gamma^r(x)=h_\gamma(x)-h_\gamma^1(x)$. Then $h_\gamma^r(x)$

satisfies the equation $h^r_\gamma(xy) = h^r_\gamma(y)x^{\rho+\gamma} + h^r_\gamma(x)$. Also by symmetry $h^r_\gamma(xy) = h^r_\gamma(x)y^{\gamma+\rho} + h^r_\gamma(y)$ and subtracting we get $h^r_\gamma(y)(x^{\gamma+\rho}-1) = h^r_\gamma(x)(y^{\gamma+\rho}-1)$ for $x,y>0$. Hence $h^r_\gamma(x)/(x^{\gamma+\rho}-1)$ is constant, i.e., $h^r_\gamma(x) = c(x^{\rho+\gamma}-1)/(\rho+\gamma)$, $c\neq 0$ for all $x>0$ (read $\log x$ for $\rho+\gamma=0$) or $h^r_\gamma(x)\equiv 0$ (because $h^1_\gamma(x)$ is a solution of (3.5)). Then $h^r_\gamma(x) = c(x^{\rho+\gamma}-1)/(\rho+\gamma)$, $c\in\mathbb{R}$, and $h_\gamma(x)\equiv h^1_\gamma(x)$.

(b) Suppose now that $\lim_{t\to\infty} \frac{(tx)^{-\gamma}a(tx) - t^{-\gamma}a(t)}{t^{-\gamma}a(t)R(t)} = 0$ for all $x>0$ and let $f(x) = \lim_{t\to\infty} \frac{R(tx)}{R(t)}$ which exists and is non negative for all $x>0$. Since $\frac{R(txy)}{R(t)} = \frac{R(txy)}{R(tx)}\frac{R(tx)}{R(t)}$ we have $f(xy) = f(y)f(x)$ for all $x,y>0$. As R is measurable the function $f(x)$ is also measurable and the only measurable solutions of Cauchy functional equation are $f(x)=x^\rho$ for some $\rho\in\mathbb{R}$ and $f(x)=0$ for $x>0$. However this last solution is not compatible with $f(1)=1$. Hence $f(x)=x^\rho$, for $x>0$, for some $\rho\in\mathbb{R}$ and $R(t)\in RV_\rho$. But as $R(t)=o(1)$, $t\to\infty$, it has to be $\rho\leq 0$. This leads to the functional equation $h_\gamma(xy)=h_\gamma(y)x^{\gamma+\rho}+h_\gamma(x)$. Hence, $h_\gamma(x) = c(x^{\rho+\gamma}-1)/(\rho+\gamma)$, $c\neq 0$ for all $x>0$ (read $\log x$ for $\rho+\gamma=0$). Note that these are the classes of (3.3) with $c_1=0$ and $c_2\neq 0$. ♦

In both theorem 3.2 and theorem 3.3 we have for the function R(t) describing the rate of convergence of the limit (1.4) that $R(t)\in RV_\rho$ for some $\rho\leq 0$. If U has a second order behaviour and the limit (3.1) exists finite and not constant with $\rho<0$ ($\rho=0$) in (3.2) we say that U has a second order regular variation (Π-variation) behaviour.

4. Asymptotic normality and bias of generalized Pickands' estimator

THEOREM 4.1. (Weak consistency of $\hat{\gamma}^P_{n,\theta}$)
If $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$, $m=m(n)\to\infty$ and $m=o(n)$, $n\to\infty$, then $\hat{\gamma}^P_{n,\theta} \xrightarrow{P} \gamma$, $n\to\infty$.
($\xrightarrow{P}$ stands for convergence in probability)

PROOF. The asymptotic distribution of $Z^{(n)}_{[m\theta]} - Z^{(n)}_m$, $0<\theta<1$ (Ref. [10]),

$$\frac{Z^{(n)}_{[m\theta]} - Z^{(n)}_m - (b_{n/m\theta} - b_{n/m})}{a_{n/m}}\sqrt{m} \xrightarrow{W} N(0, 1-2\theta^{-\gamma}+\theta^{-2\gamma-1}),$$

$n\to\infty$, implies that

$$(Z^{(n)}_{[m\theta]} - Z^{(n)}_m)/a_{n/m} = (b_{n/m\theta} - b_{n/m})/a_{n/m} + O_p(1/\sqrt{m})$$
$$= (\theta^{-\gamma}-1)/\gamma + o_p(1),$$

$$\text{i.e., } (Z^{(n)}_{[m\theta]} - Z^{(n)}_m)/a_{n/m} \xrightarrow{P} (\theta^{-\gamma}-1)/\gamma,\ n\to\infty. \quad (4.1)$$

Hence, for $0<\theta<1$,

$$A_{m,\theta} = (Z^{(n)}_{[m\theta^2]} - Z^{(n)}_{[m\theta]})/(Z^{(n)}_{[m\theta]} - Z^{(n)}_m) =$$
$$= (Z^{(n)}_{[m\theta^2]} - Z^{(n)}_m)/(Z^{(n)}_{[m\theta]} - Z^{(n)}_m) - 1 \xrightarrow{P} \theta^{-\gamma}$$

and $\hat{\gamma}^P_{n,\theta} = \log A_{m,\theta}/(-\log\theta) \xrightarrow{P} \log\theta^{-\gamma}/(-\log\theta) = \gamma$. ♦

However, the weak consistency of $\hat{\gamma}^P_{n,\theta}$ for any intermediate sequence m(n), as well as the strong consistency for an intermediate sequence m(n) such that $m(n)/\log\log n\to\infty$, can be proved for F in $D(G_\gamma)$ by using the argument in Ref.[7].

THEOREM 4.2. (Asymptotic normality of $\hat{\gamma}^P_{n,\theta}$)
Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$, $m=m(n)$ be an intermediate sequence ($m=m(n)\to\infty$ and $m=o(n)$, $n\to\infty$) and $0<\theta<1$.
(A) Suppose that for some normalizing function a(.),

$$\frac{U(t/\theta) - U(t)}{a(t)} = \frac{\theta^{-\gamma}-1}{\gamma},\ \text{i.e.,}\ R_{\gamma,\theta^{-1}}(t) \equiv 0. \quad (4.2)$$

Then $\sqrt{m}(\hat{\gamma}^P_{n,\theta} - \gamma)$ has asymptotically a normal distribution with mean value zero and variance $\sigma^2(\gamma,\theta)$ for any intermediate sequence m(n).
(B) Being $R_{\gamma,\theta^{-1}}(t)$ not identically zero, suppose that U has a second order behaviour, that is, there exists a positive function R such that,

$R_{\gamma,\theta^{-1}}(t) = h_\gamma(\theta^{-1})R(t) + o(R(t))$ and $R(t)=o(1)$, $t\to\infty$ (4.3)

Then $\sqrt{m}(\overset{\wedge}{\gamma}^{P}_{n,\theta} - \gamma)$

(i) has asymptotically a normal distribution with mean value zero and variance $\sigma^2(\gamma,\theta)$ for intermediate sequences satisfying $m=o(n/g^{\leftarrow}(n))$,

(ii) has asymptotically a normal distribution with mean value $b_c(\gamma,\theta)$ and variance $\sigma^2(\gamma,\theta)$ for intermediate sequences satisfying $m \sim n/(g^{\leftarrow}(n/c^2))$, $c>0$,

where $g(t)=t/R^2(t)$ and $g^{\leftarrow}$ is the asymptotic inverse function of g,

$$\sigma^2(\gamma,\theta) = \frac{\gamma^2(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{(1-\theta^{-\gamma})^2\log^2\theta} \qquad \left(\sigma^2(0,\theta) = \frac{\theta^{-2}-1}{\log^4\theta}\right)$$

$$b_c(\gamma,\theta)=cb(\gamma,\theta)=c\,\frac{\gamma H(\gamma,\theta)}{\theta^{-\gamma}(\theta^{-\gamma}-1)(-\log\theta)} \quad \left(b_c(0,\theta)=c\,\frac{H(0,\theta)}{\log^2\theta}\right)$$

and $H(\gamma,\theta) = h_\gamma(\theta^{-2})-(\theta^{-\gamma}+1)h_\gamma(\theta^{-1})$.

REMARK. In view of Corollary 2.1, Lemma 3.1 and Theorem 3.2 the conditions of Theorem 4.2 include the following ones (with $\gamma\in\mathbb{R}$),

(A') $t^{1-\gamma}U'(t)\in RV_0$ and $\dfrac{U(tx)-U(t)}{tU'(t)} = \dfrac{x^\gamma-1}{\gamma}$.

(B') $t^{1-\gamma}U'(t)\in RV_0$ and for $x>0$,

$$\lim_{t\to\infty}\frac{(tx)^{1-\gamma}U'(tx) - t^{1-\gamma}U'(t)}{b(t)} = c\,\frac{x^\rho-1}{\rho}$$

for some $c\neq0$, $\rho\leq0$ and some positive function b(.).

(B'') $t^{1-\gamma}U'(t)\in RV_0$ and $\pm t^{1-\gamma}U'(t)\in\Pi$ or $\pm(d-t^{1-\gamma}U'(t))\in RV_\rho$ with $d = \lim_{t\to\infty} t^{1-\gamma}U'(t)>0$ for some $\rho\leq0$.

PROOF. We begin by investigating in what conditions $\sqrt{m}(A_{m,\theta} - \theta^{-\gamma})$ has asymptotically a normal distribution. By (4.1)

$$\sqrt{m}(A_{m,\theta}-\theta^{-\gamma}) \sim \sqrt{m}\left\{\frac{Z^{(n)}_{[m\theta^2]}-Z^{(n)}_{[m\theta]}-\theta^{-\gamma}(Z^{(n)}_{[m\theta]}-Z^{(n)}_{m})}{a_{n/m}(\theta^{-\gamma}-1)/\gamma}\right\}$$

in probability $(n\to\infty)$ (4.4)

Now we introduce for $0<\theta<1$,

$$N_{m,\theta} = \sqrt{m}\,\frac{Z^{(n)}_{[m\theta]} - Z^{(n)}_{m} - (b_{n/m\theta} - b_{n/m})}{a_{n/m}}$$

$$N_{m,\theta^2} = \sqrt{m}\,\frac{Z^{(n)}_{[m\theta^2]} - Z^{(n)}_{[m\theta]} - (b_{n/m\theta^2} - b_{n/m\theta})}{a_{n/m}}$$

$$T_{m,\theta} = \frac{\gamma}{\theta^{-\gamma}-1}\left\{N_{m,\theta^2} - \theta^{-\gamma}N_{m,\theta}\right\}$$

and note that $T_{m,\theta}$ has asymptotically a normal distribution with mean value zero and variance $V_A = \dfrac{\gamma^2(\theta^{-1}-1)(1+\theta^{-2\gamma-1})\theta^{-2\gamma}}{(1-\theta^{-\gamma})^2}$. The right hand side of (4.4) is equal to

$$T_{m,\theta}+\frac{\gamma}{\theta^{-\gamma}-1}\sqrt{m}\,\frac{b_{n/m\theta^2}-b_{n/m\theta}-\theta^{-\gamma}(b_{n/m\theta}-b_{n/m})}{a_{n/m}} \quad (4.5)$$

If the second term is zero or if we are able to make the second term negligible by letting the sequence m(n) increase appropriately, the asymptotic distribution of $\sqrt{m}(A_{m,\theta} - \theta^{-\gamma})$ will be the one of $T_{m,\theta}$. If the sequence converges to infinity in such a way that the second term converges to a constant different from zero, then the asymptotic distribution will still be normal with the same variance but with mean value different from zero.

Case (A) – If the distribution function F is such that for some sequence of attraction coefficients a_n we have

$$\frac{b_{n/m\theta}-b_{n/m}}{a_{n/m}} = \frac{U(n/m\theta)-U(n/m)}{a_{n/m}} = \frac{\theta^{-\gamma}-1}{\gamma}, \quad n\geq1, \text{ then}$$

$$\sqrt{m}\left\{\frac{b_{n/m\theta^2} - b_{n/m}}{a_{n/m}} - (\theta^{-\gamma}+1)\frac{b_{n/m\theta} - b_{n/m}}{a_{n/m}}\right\} = 0, \; n\geq1$$

and $\sqrt{m}(A_{m,\theta} - \theta^{-\gamma})$ has aymptotically a normal distribution with mean value zero and variance V_A for any intermediate sequence m(n). Note that for any other sequence of attraction coefficients a_n^*, we have $a_n^*/a_n=1+o(1)$ and the second term is also zero.

Case (B) – If $R_{\gamma,\theta^{-1}}(n/m)$ is not identically zero for any possible choice of the normalizing sequence a_n we obtain

$$\frac{b_{n/m\theta^2}-b_{n/m}}{a_{n/m}} - (\theta^{-\gamma}+1)\frac{b_{n/m\theta}-b_{n/m}}{a_{n/m}} =$$

$$= R_{\gamma,\theta^{-2}}(n/m) - (\theta^{-\gamma}+1)R_{\gamma,\theta^{-1}}(n/m)$$

$$= [h_\gamma(\theta^{-2}) - (\theta^{-\gamma}+1)h_\gamma(\theta^{-1})]R(n/m) + o(R(n/m))$$

$$= H(\gamma,\theta)R(n/m) + o(R(n/m)).$$

Hence the second term of (4.5) is equal to

$$\frac{\gamma}{\theta^{-\gamma}-1} H(\gamma,\theta)\sqrt{m}R(n/m) + \sqrt{m}o(R(n/m)). \qquad (4.6)$$

If the intermediate sequence m(n) is such that

$$\sqrt{m}\, R(n/m) \to 1,\ n\to\infty, \qquad (4.7)$$

which is equivalent to $n \sim g(n/m)$, with $g(t) = t/R^2(t)$, which in turn is equivalent to

$$m \sim n/g^{\leftarrow}(n),\ n\to\infty, \qquad (4.8)$$

with $g^{\leftarrow}$ the asymptotic inverse function of g, the second term of (4.5) converges to $b_A(\gamma,\theta) = \gamma H(\gamma,\theta)/(\theta^{-\gamma}-1)$ and $\sqrt{m}(A_{m,\theta} - \theta^{-\gamma})$ has asymptotically a normal distribution with mean $b_A(\gamma,\theta)$ and variance V_A. Note that $g(t)\to\infty$, $t\to\infty$.

Let $m_0(n)$ be a sequence satisfying the above condition, that is,

$$m_0 \sim n/g^{\leftarrow}(n),\quad n\to\infty \qquad (4.9)$$

and let m(n) be any other intermediate sequence of smaller order than m_0,

$$m=o(m_0),\ n\to\infty. \qquad (4.10)$$

We are going to show that for the sequence m(n) satisfying (4.10) the second term converges to zero. In fact, we know that for the sequence m_0 we have $\sqrt{m_0}R(n/m_0) \to 1$, $n\to\infty$. Now

$$\sqrt{m}R(n/m) = \sqrt{\frac{m}{m_0}}\sqrt{m_0}R(n/m_0)\frac{R(n/m)}{R(n/m_0)}$$

Noting that $m=o(m_0)$ is equivalent to $n/m_0=o(n/m)$ and that $R(t)\to 0$, $t\to\infty$, we conclude that R(n/m) can not be of a greater order than $R(n/m_0)$ and hence $R(n/m)/R(n/m_0)$ can not converge to infinity. Then $\sqrt{m}R(n/m)\to 0$, $n\to\infty$. Hence, for any intermediate sequence $m=o(m_0)=o(n/g^{\leftarrow}(n))$, $\sqrt{m}(A_{m,\theta}-\theta^{-\gamma}) \xrightarrow{W} N(0,V_A)$, $n\to\infty$ (case (B)–(i)).

Assume now that the intermediate sequence m=m(n) is such that

$$\sqrt{m}R(n/m) \to c,\ n\to\infty,\ c>0, \qquad (4.11)$$

that is, $n \sim c^2 g(n/m)$, $n\to\infty$, which is equivalent to

$$m \sim \frac{n}{g^{\leftarrow}(n/c^2)},\ n\to\infty. \qquad (4.12)$$

In this case $\sqrt{m}(A_{m,\theta}-\theta^{-\gamma})$ has asymptotically a normal distribution with mean $cb_A(\gamma,\theta)$ and variance V_A (case (B)–(ii)).

Now to obtain the asymptotic distribution of $\hat{\gamma}^P_{n,\theta}$ we just have to expand it about $\theta^{-\gamma}$ since the estimator is equal to $\log A_{m,\theta}/(-\log\theta)$. From

$$\hat{\gamma}^P_{n,\theta} = \gamma + \frac{A_{m,\theta}-\theta^{-\gamma}}{\theta^{-\gamma}(-\log\theta)} + o(A_{m,\theta}-\theta^{-\gamma})$$

we obtain

$$\sqrt{m}(\hat{\gamma}^P_{n,\theta} - \gamma) = \sqrt{m}\frac{A_{m,\theta}-\theta^{-\gamma}}{\theta^{-\gamma}(-\log\theta)} + o(\sqrt{m}(A_{m,\theta}-\theta^{-\gamma}))$$

$$= \sqrt{m}\frac{A_{m,\theta}-\theta^{-\gamma}}{\theta^{-\gamma}(-\log\theta)} + o_p(1)$$

and finally we conclude that $\sqrt{m}(\hat{\gamma}^P_{n,\theta}-\gamma)$ has an asymptotic distribution which is,

in case (A), $N(0,V_A/\theta^{-2\gamma}\log^2\theta) = N(0,\sigma^2(\gamma,\theta))$
for any sequence $m=m(n)\to\infty$ and $m=o(n)$.

in case (B)–(i), $N(0,V_A/\theta^{-2\gamma}\log^2\theta) = N(0,\sigma^2(\gamma,\theta))$
for sequences $m=m(n)\to\infty$ and $m=o(n/g^{\leftarrow}(n))$.

in case (B)–(ii), $N(cb_A(\gamma,\theta)/\theta^{-\gamma}(-\log\theta), V_A/\theta^{2\gamma}\log^2\theta) = N(b_c(\gamma,\theta),\sigma^2(\gamma,\theta))$
for sequences $m=m(n)\to\infty$ and $m \sim n/g^{\leftarrow}(n/c^2)$. ♦

The first question which naturally arises is the one of knowing what are the distribution functions belonging to the differentiable domain of attraction of the GEV distribution for which the asymptotic normality of the estimator is valid for any intermediate sequence. Corollary 4.1 gives an answer to this question showing that from among the distribution functions belonging to the differentiable domain of attraction of GEV distribution which admit a positive derivative only the generalized Pareto distribution verifies case (A) of theorem 2.2.

COROLLARY 4.1. Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and suppose that F admits a positive derivative F'. Then $\sqrt{m}(\hat{\gamma}^P_{n,\theta}-\gamma)$ has asymptotically a normal distribution

with mean value zero and variance $\sigma^2(\gamma,\theta)$ for any intermediate sequence m(n) if and only if F is the generalized Pareto distribution function.

PROOF. Note that for a distribution function F in the conditions of the theorem $\sqrt{m}(\hat{\gamma}^P_{n,\theta}-\gamma)$ is asymptotically $N(0,\sigma^2(\gamma,\theta))$ for any intermediate sequence m(n) if and only if

$$U(n/m\theta^2)-U(n/m)-(\theta^{-\gamma}+1)(U(n/m\theta)-U(n/m)) = 0, \quad (4.13)$$

as is easily seen from the proof of theorem 4.2. We are going to solve the functional equation

$$U(tx^2)-U(t) = (x^\gamma+1)(U(tx)-U(t)) \quad \text{for } t\geq 1 \text{ and } x\geq 1. \quad (4.14)$$

This equation is equivalent to $\int_{tx}^{tx^2} U'(y)dy = x^\gamma \int_t^{tx} U'(y)dy$ or $\int_t^{tx} U'(xz)xdz = x^\gamma \int_t^{tx} U'(z)dz$ which in turn is equivalent to $U'(xz)-x^{\gamma-1}U'(z)=0$. In particular for z=1, this equation reads $U'(x) = x^{\gamma-1}U'(1)$ which is equivalent to $U(x) = c_1(x^\gamma-1)/\gamma + c_2$, $c_1>0$, $c_2\in\mathbb{R}$ ($= c_1\log x+c_2$, $\gamma=0$) which is the tail quantile function of the generalized Pareto distribution. Moreover equation (4.14) is equivalent to $(U(tx)-U(t))/tU'(t) = (x^\gamma-1)/\gamma$ which means that the normalizing sequence a_n such that (4.2) holds is $a_n=nU(n)$. ♦

If we assume R(t) to be a regularly varying function (which appears to be a natural restriction in view of previous section results) the conditions on the sequence m(n) stated in the above theorem admit a slight simplification and furthermore it is possible in most cases to minimize the mean squared error of the estimator when the bias is different from zero.

COROLLARY 4.2. If the function R(t) of theorem 4.2 is ρ–regularly varying at infinity then $\rho\leq 0$ and $g^{\leftarrow}$ is also a regularly varying function at infinity with index $\rho^*=(1-2\rho)^{-1}$ ($0<\rho^*\leq 1$). Moreover condition (4.12) of theorem 4.2 is equivalent to

$$m \sim c^{2\rho^*}\frac{n}{g^{\leftarrow}(n)}, \quad c>0. \qquad (4.15)$$

PROOF. Let $R(t)\in RV_\rho$. As $\lim_{t\to\infty} R(t)=0$ it has to be $\rho\leq 0$ (cf. corollary 1.2.1, property 1, Ref. [13]). This is equivalent to saying that $g(t) = t/R^2(t)$ is regularly varying at infinity with index $1-2\rho$, with $1-2\rho\geq 1$. On the other hand, there exists a function $V(t)\in RV_{1-2\rho}$ and strictly monotone which is asymptotically equal to g(t) (cf. corollary 1.2.1, property 7, Ref. [13]). It follows from $g(t)\to\infty$, $t\to\infty$, that $V(\infty)=\infty$ and hence V is a strictly increasing function. This implies (cf. corollary 1.2.1, property 5, Ref. [13]) that the generalized inverse of V, $V^{\leftarrow}$, is regularly varying with index $\rho^*=(1-2\rho)^{-1}$. Furthermore, as $g\sim V$ and $V^{\leftarrow}$ is regularly varying we have $g^{\leftarrow}\sim V^{\leftarrow}$. Then g admits an asymptotic inverse $g^{\leftarrow}$ which is regularly varying with index $\rho^*=(1-2\rho)^{-1}$ where $0<\rho^*\leq 1$. This means that $g^{\leftarrow}(n/c^2) \sim (1/c^2)^{\rho^*} g^{\leftarrow}(n)$ for all $c>0$, which implies that $m \sim n/g^{\leftarrow}(n/c^2)$ is equivalent to $m \sim c^{2\rho^*}n/g^{\leftarrow}(n)$, $0<\rho^*\leq 1$, $c>0$. ♦

Note that $n/g^{\leftarrow}(n)\in RV_{1-\rho^*}$, $0\leq 1-\rho^*<1$, that is, $n/g^{\leftarrow}(n)\in RV_{-2\rho/(1-2\rho)}$, $\rho\leq 0$.

Following Hall, Ref. [14], the next theorem shows that if $\rho<0$ something can be added to the results of theorem 4.2.

THEOREM 4.3. Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and assume that (4.3) holds with $R(t)\sim t^\rho$, $\rho<0$. Then, for intermediate sequences m=m(n) such that

(**i**) $m = o(n^{-2\rho/(1-2\rho)})$, $n\to\infty$, $\sqrt{m}(\hat{\gamma}^P_{n,\theta}-\gamma)$ has asymptotically a normal distribution with mean value zero and variance $\sigma^2(\gamma,\theta)$.

(**ii**) $m \sim c^{2/(1-2\rho)}n^{-2\rho/(1-2\rho)}$, $n\to\infty$, $\sqrt{m}(\hat{\gamma}^P_{n,\theta}-\gamma)$ has asymptotically a normal distribution with mean value $b_c(\gamma,\theta)$ and variance $\sigma^2(\gamma,\theta)$.

(**iii**) $m/n^{-2\rho/(1-2\rho)}\to\infty$, $n\to\infty$, $(n/m)^{-\rho}(\hat{\gamma}^P_{n,\theta}-\gamma) \xrightarrow{P} b(\gamma,\theta)$.

PROOF. Noting that $R(t)\sim t^\rho$, for some $\rho<0$, implies $t/g^{\leftarrow}(t)) \sim t^{-2\rho/(1-2\rho)}$, (i) and (ii) follow immediatly from theorem 4.2. In what concerns (iii), from (4.4) and (4.5) we obtain the following representation

$$\left(\frac{n}{m}\right)^{-\rho}(A_{m,\theta}-\theta^{-\gamma}) = \frac{n^{-\rho}}{m^{1/2-\rho}}T_{m,\theta}+\gamma\frac{H(\gamma,\theta)}{\theta^{-\gamma}-1}\left(\frac{n}{m}\right)^{-\rho}R(n/m) + o((n/m)^\rho R(n/m))$$

$$= \gamma\frac{H(\gamma,\theta)}{\theta^{-\gamma}-1} + o_p(1)$$

since $n^{-\rho}/m^{1/2-\rho}\to 0$ and $T_{m,\theta}\xrightarrow{W} N(0,V_A)$, $n\to\infty$ for any intermediate sequence m(n). Hence,

$$(\frac{n}{m})^{-\rho}(\overset{\wedge}{\gamma}^{P}_{n,\theta}-\gamma)=(\frac{n}{m})^{-\rho}\frac{(A_{m,\theta}-\theta^{-\gamma})}{\theta^{-\gamma}(-\log\theta)}+o((n/m)^{-\rho}(A_{m,\theta}-\theta^{-\gamma}))$$

$$=\frac{\gamma H(\gamma,\theta)}{\theta^{-\gamma}(\theta^{-\gamma}-1)(-\log\theta)}+o_p(1). \quad \blacklozenge$$

If the number of upper order statistics involved in the estimation of γ is small the variance of the estimator is large. But if we increase the number of upper order statistics used in order to obtain a smaller variance, the estimator will have a bias different from zero. However in this last situation the mean squared error of the estimator can be minimized if $\rho<0$ giving an optimal criterion to choose m.

THEOREM 4.4. (Minimization of mean squared error) Let $F\in D_{dif}(G_\gamma)$ for some $\gamma\in\mathbb{R}$ and suppose U has a second order behaviour with $R(t)\in RV_\rho$, $\rho\leq 0$. Then, for sequences $m=m(n)\to\infty$ and $m\sim c^{2\rho^*}n/g^{\leftarrow}(n)$, with $c>0$ and $\rho^*=(1-2\rho)^{-1}$, $0<\rho^*\leq 1$, the mean squared error of the estimator $\overset{\wedge}{\gamma}^{P}_{n,\theta}$ is asymptotically equal to

$$\frac{\sigma^2(\gamma,\theta)/c^{2\rho^*}+c^{2(1-\rho^*)}b^2(\gamma,\theta)}{n/g^{\leftarrow}(n)}.$$

If $0<\rho^*<1$ the mean squared error is minimum for $c_0=(\rho^*\sigma^2(\gamma,\theta)/(1-\rho^*)b^2(\gamma,\theta))^{1/2}$ and if $\rho^*=1$ the mean squared error is a decreasing function of c (the bias of the estimator remains constant but the variance decreases as m increases).

PROOF. We have seen in theorem 4.2 and corollary 4.2 that if $m\sim c^{2\rho^*}n/g^{\leftarrow}(n)$, $\sqrt{m}(\overset{\wedge}{\gamma}^{P}_{n,\theta}-\gamma)$ is asymptotically normal with mean value $b_c(\gamma,\theta)=cb(\gamma,\theta)$ and variance $\sigma^2(\gamma,\theta)$. Hence, the (asymptotic) mean squared error of $\overset{\wedge}{\gamma}^{P}_{n,\theta}$ is

$$MSE_\infty(\overset{\wedge}{\gamma}^{P}_{n,\theta})=[\sigma^2(\gamma,\theta)+b_c^2(\gamma,\theta)]/m$$

$$\sim[\sigma^2(\gamma,\theta)/c^{2\rho^*}+c^{2(1-\rho^*)}b^2(\gamma,\theta)]/(n/g^{\leftarrow}(n))$$
$$=f(c)/(n/g^{\leftarrow}(n)).$$

If $0<\rho^*<1$ the value of c that minimizes the $MSE_\infty(\overset{\wedge}{\gamma}^{P}_{n,\theta})$ is the zero of f'(c) which is easily seen to be $c_0=[\rho^*\sigma^2(\gamma,\theta)/(1-\rho^*)b^2(\gamma,\theta)]^{1/2}$. If $\rho^*=1$, we have $f(c)=\sigma^2(\gamma,\theta)/c^2+b^2(\gamma,\theta)$ and hence the bias remains constant and the variance decreases as c increases. $\blacklozenge$

So if $0<\rho^*<1$ we should consider the m largest observations of the sample with $m=[c_0^{2\rho^*}n/g^{\leftarrow}(n)]$ to evaluate the estimator $\overset{\wedge}{\gamma}^{P}_{n,\theta}$. The situation $\rho^*=1$ is more complicated: it corresponds to slowly varying functions R(t) and $t/g^{\leftarrow}(t)$ and the above theorem does not give a definite answer to the problem of choosing m.

5. Examples

In this section we illustrate the above results with the continuous distribution functions that are typically used in statistical applications. In what concerns the differentiable domain of attraction of Gumbel distribution we will see that logistic distribution and Gumbel distribution itself have a second order regular variation behaviour while the Gamma(r), $r\neq 1$, and Normal distribution functions have a second order Π-variation behaviour, as well as the distribution function $F(x)=1-\exp(-x^\alpha)$, $\alpha>0$, $\alpha\neq 1$. For the differentiable domain of attraction of GEV, $\gamma\neq 0$, we shall see that GEV($\gamma\neq 0$) and Cauchy distribution functions have a second order regular variation behaviour. Furthermore we consider the asymptotic normality of the estimator and determine the asymptotic bias for each one of the continuous distributions considered as well as the theoretical optimal value of m for the "polynomial rate" distributions.

EXAMPLE 1. For the logistic distribution we have $F(t)=(1+e^{-t})^{-1}$, $t\in\mathbb{R}$, and $U(t)=\log(t-1)$. The function $tU'(t)=t/(t-1)\in RV_0$ is strictly decreasing with $\lim_{t\to\infty}tU'(t)=1$ and $tU'(t)-1\in RV_{-1}$. We also have, for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)}=\log x+(1-x^{-1}-\log x)t^{-1}+o(t^{-1}),\ t\to\infty.$$

Here $h_0(x)$ is a function of class (2) with $c_1=\rho=-1$, $c_2=0$ and $R(t)=t^{-1}$. Hence logistic distribution has a second order regular variation behaviour. If $m=o(n^{2/3})$ the estimator has asymptotically a normal distribution with mean zero and variance $\sigma^2(0,\theta)/m$ (theorem 4.2.(B)–(i)). If $m\sim dn^{2/3}$, $d>0$, $\sqrt{m}\overset{\wedge}{\gamma}^{P}_{n,\theta}$ has a bias equal to $-d^{3/2}(\theta-1)^2/\log^2\theta$ (theorem 4.2.(B)–(ii)) and the asymptotic MSE of $\overset{\wedge}{\gamma}^{P}_{n,\theta}$ is minimum for $m\sim d_0n^{2/3}$ with $d_0=[(1+\theta)/2\theta^2(1-\theta)^3]^{1/3}$ (theorem 4.4.).

EXAMPLE 2. For the Gumbel distribution we have $U(t) = -\log(-\log(1-t^{-1}))$. The slowly varying function $tU'(t) = [(1-t)\log(1-t^{-1})]^{-1}$ is strictly decreasing with $\lim_{t\to\infty} tU'(t)=1$ and $tU'(t)-1\in RV_{-1}$. Also, for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x + \frac{1}{2}(1-x^{-1}-\log x)t^{-1} + o(t^{-1}),\ t\to\infty.$$

It follows that $h_0(x)$ is a function of class (2) with $c_1=-1/2$, $\rho=-1$, $c_2=0$ and $R(t) = t^{-1}$ and hence Gumbel distribution has a second order regular variation behaviour. If $m=o(n^{2/3})$ the bias of $\hat{\gamma}^P_{n,\theta}$ is zero (Theorem 4.2. (B)-(i)) and if $m \sim dn^{2/3}$, $d>0$, the bias is equal to $-d^{3/2}(\theta-1)^2/(2\log^2\theta\sqrt{m})$ (Theorem 4.2.(B)-(ii)). Moreover MSE $(\hat{\gamma}^P_{n,\theta})$ is minimum for $m \sim d_0 n^{2/3}$ with $d_0 = 2(1+\theta)^{1/3}/[\theta^{2/3}(1-\theta)]$ (Theorem 4.4.).

EXAMPLE 3. For the distribution functions $F(x) = 1-\exp(-x^\alpha)$, $\alpha>0$, $\alpha\neq1$, we have $U(t) = (\log t)^{1/\alpha}$ and $tU'(t) = (1/\alpha)(\log t)^{1/\alpha-1}\in RV_0$ is a strictly increasing function for $\alpha<1$ and a strictly decreasing function for $\alpha>1$. One obtains, for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x+\frac{1-\alpha}{\alpha}\frac{\log^2 x}{2}(\log t)^{-1}+o((\log t)^{-1}),\ t\to\infty.$$

Hence $F\in D(G_0)$. Here $h_0(x)$ is a function of class (1) with $c_1=(1-\alpha)/\alpha$, $c_2=0$ and $R(t) = (\log t)^{-1}$. It follows that these distribution functions have a second order Π-variation behaviour. Note that $\pm tU'(t)\in\Pi(a_1)$ with $a_1(t) = (|1-\alpha|/\alpha^2)(\log t)^{1/\alpha-2}$. When $m = o(\log^2 n)$, $\sqrt{m}\hat{\gamma}^P_{n,\theta}$ will be asymptotically normal with mean value zero and variance $\sigma^2(0,\theta)$; when $m \sim d\log^2 n$, $d>0$, the distribution will have a bias equal to $d^{1/2}(1/\alpha-1)$ and the MSE$(\hat{\gamma}^P_{n,\theta})$ is a decreasing function of d (the bias remains constant and the variance decreases).

EXAMPLE 4. For the Cauchy distribution we have

$$F(t) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}(t),\ t\in\mathbb{R},$$

$$U(t) = \operatorname{tg}(\frac{\pi}{2} - \frac{\pi}{t}) = \frac{t}{\pi}\left[1 - \frac{\pi^2}{3t^2} + o(t^{-2})\right],\ t\to\infty,$$

and $tU'(t) = \frac{\pi}{t}\{1+\frac{t^2}{\pi^2}[1+\frac{2\pi^2}{3t^2}+o(t^{-2})]\}$, $t\to\infty$. One obtains, for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = x-1 + \frac{\pi^2}{3}(2-x^{-1}-x)t^{-2} + o(t^{-2}),\ t\to\infty$$

and $h_1(x)$ is a function of class (4) with $\rho=-2$, $c_1=-(2/3)\pi^2$, $c_2=0$ and $R(t) = t^{-2}$. Moreover, $U'(t)$ is a strictly decreasing function with $\lim_{t\to\infty} U'(t) = \pi^{-1}$ and $U'(t)-\pi^{-1}\in RV_{-2}$. It follows that Cauchy distribution function has a second order regular variation behaviour. For sequences $m=o(n^{4/5})$ the asymptotic distribution of $\sqrt{m}(\hat{\gamma}^P_{n,\theta} - 1)$ is $N(0,\sigma^2(1,\theta))$ whereas for sequences $m \sim dn^{4/5}$ the distibution will have a bias equal to $d^{5/2}\pi^2\theta(1-\theta^2)/3\log\theta$. The mean squared error of the estimator is minimum for $m \sim d_0 n^{4/5}$ where $d_0 = [9(1+\theta^3)/(4\pi^4\theta^4(1-\theta)^3(1+\theta)^2)]^{1/5}$.

EXAMPLE 5. For the Generalized Extreme Value distribution, $G_\gamma(x)=\exp\{-(1+\gamma x)^{-1/\gamma}\}$, $1+\gamma x>0$, $\gamma\in\mathbb{R}$, we have $U(t) = \{[-\log(1 - (1/t))]^{-\gamma} - 1\}/\gamma$, $t>1$, and $t^{1-\gamma}U'(t) = t^{-(\gamma+1)}(1 - (1/t))[(-\log(1 - (1/t))]^{-(\gamma+1)}$. For $\gamma\le1$, $t^{1-\gamma}U'(t)$ is a strictly decreasing function and for $\gamma>1$ there exists $t_0>1$ such that $t^{1-\gamma}U'(t)$ is strictly increasing for $t>t_0$. In both cases we have $\lim_{t\to\infty} t^{1-\gamma}U'(t)=1$. For $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = \frac{x^\gamma-1}{\gamma} + \frac{1-\gamma}{2}\left(\frac{x^{\gamma-1}-1}{\gamma-1} - \frac{x^\gamma-1}{\gamma}\right)t^{-1} + o(t^{-1}),$$

$t\to\infty$, $\gamma\neq1$.

$$\frac{U(tx)-U(t)}{tU'(t)} = x-1 + \frac{1}{12}(2-x-x^{-1})t^{-2} + o(t^{-2}),\ t\to\infty,\ \gamma=1.$$

Here $h_\gamma(x)$ is a function of class (4) with $c_1=(\gamma-1)/2$, $c_2=0$, $\rho=-1$ and $R(t) = t^{-1}$ for $\gamma\neq1$ and with $c_1=1/12$, $c_2=0$, $\rho=-2$ and $R(t) = t^{-2}$ for $\gamma=1$. Hence GEV distribution function has a second order regular variation behaviour for any $\gamma\in\mathbb{R}$. If $\gamma\neq1$ the conclusion (B)-(i) of Theorem 4.2. holds for sequences $m = o(n^{2/3})$ whereas for sequences $m \sim dn^{2/3}$, $d>0$, holds conclusion (B)-(ii) of Theorem 4.2. with $b_d(\gamma,\theta) = d^{3/2}\gamma(\theta-1)(\theta^{-\gamma+1}-1)/[2(\theta^{-\gamma}-1)\log\theta]$; the minimum mean squared error of $\hat{\gamma}^P_{n,\theta}$ is attained for $m\sim d_0 n^{2/3}$ with $d_0 = [2(\theta^{2\gamma+1}+1)/(\theta^4(1-\theta)(1-\theta^{\gamma-1})^2)]^{1/3}$. If $\gamma=1$ and for sequences $m=o(n^{4/5})$ part (B)-(i) of theorem 4.2. holds whereas for sequences $m \sim dn^{4/5}$, $d>0$, part (B)-

(ii) holds with the bias $b_d(1,\theta) = d^{5/2}\theta(1-\theta^2)/(12\log\theta)$. Conclusion of Theorem 4.4. holds for sequences $m \sim d_0 n^{4/5}$ with $d_0 = [36(\theta^3+1)/\theta^4(1-\theta)^3(\theta+1)^2]^{1/5}$.

EXAMPLE 6. For Gamma distribution,

$$1-F(t) = \frac{1}{\Gamma(r)}\int_t^{+\infty} s^{r-1}e^{-s}ds, \quad r>0,\ r\neq 1,$$

and using the expansion

$$\int_t^{+\infty} s^{r-1}e^{-s}ds = e^{-t}t^{r-1}[1+(r-1)t^{-1}+(r-1)(r-2)t^{-2}+o(t^{-2})],$$

$t>0$, one obtains, after some calculations,

$$\lim_{t\to\infty}\left[\frac{1-F(t+xf_0(t))}{1-F(t)} - e^{-x}\right]/\beta(t) = \pm\frac{x^2}{2}e^{-x}$$

with the plus sign for $r<1$ and the minus sign for $r>1$, $f_0(t) = (1-F(t))/F'(t)$ and $\beta(t) = \mp(r-1)t^{-2}$. By theorem A.10 in Ref. [7] this is equivalent to $\pm tU'(t)\in\Pi(a)$ with $a(t) = f_0(U(t)).\beta(U(t)) = tU'(t).\beta(U(t))$ which is equivalent (theorem A.5, Ref. [7]) to

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x - (r-1)\frac{\log^2 x}{2}(U(t))^{-2} + o((U(t))^{-2}),\ t\to\infty,$$

for $x>0$, and we have to find an asymptotic expression for $U(t)$. As $1-F(t) \sim F'(t)$ one obtains $U(t) \sim (\Gamma(r)t^{1-r}e^t)^{\leftarrow} \sim$ $\sim \log(t/\Gamma(r))$. Hence, for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x - (r-1)\frac{\log^2 x}{2}(\log t)^{-2} + o((\log t)^{-2}),$$
$$t\to\infty,\ r\neq 1,$$

with $h_0(x)$ a function of class (1) with $c_1=r-1$, $c_2=0$ and $R(t) = (\log t)^{-2}$ and gamma distribution has a second order Π-variation behaviour. For sequences $m=o(\log^4 n)$, $\sqrt{m}\hat{\gamma}^P_{n,\theta}$ has asymptotically a normal distribution with mean value zero and variance $\sigma^2(0,\theta)$, and for sequences $m \sim d\log^4 n$, $d>0$, the asymptotic distribution has a bias equal to $d^{1/2}(1-r)$; as d increases the bias of the estimator remains constant and the variance decreases.

EXAMPLE 7. For normal distribution,

$$1-F(t) = \frac{1}{\sqrt{2\pi}}\int_t^{+\infty} e^{-s^2/2}ds$$

and using the expansion

$$\int_t^{+\infty} e^{-s^2/2}ds = e^{-t^2/2}t^{-1}\{1-t^{-2}+o(t^{-2})\},\ t\to\infty,$$

one obtains

$$\lim_{t\to\infty}\left[\frac{1-F(t+xf_0(t))}{1-F(t)} - e^{-x}\right]/\beta(t) = -(x^2/2)e^{-x}$$

with $f_0(t) = t^{-1}\{1-t^{-2}+o(t^{-2})\}$ and $\beta(t) = t^{-2}$. This is equivalent (theorem A.10, Ref. [7]) to $-tU'(t)\in\Pi(a)$ with $a(t) = tU'(t).\beta(U(t))$ and this statement is still equivalent (theorem A.5, Ref. [7]) to

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x - \frac{\log^2 x}{2}(U(t))^{-2} + o((U(t))^{-2}),\ t\to\infty,$$

for $x>0$. As $1-F(t) \sim F'(t)/t$ we have $U(t) \sim (\sqrt{2\pi}te^{t^2/2})^{\leftarrow} \sim$ $\sim \sqrt{2}(\log t)^{1/2}$. Hence for $x>0$,

$$\frac{U(tx)-U(t)}{tU'(t)} = \log x - \frac{\log^2 x}{4}(\log t)^{-1} + o((\log t)^{-1}),\ t\to\infty,$$

with $h_0(x)$ a function of class (1) with $c_1=1/2$, $c_2=0$ and $R(t) = (\log t)^{-1}$ and normal distribution has a second order Π-variation behaviour. For sequences $m=o(\log^2 n)$, $\sqrt{m}\hat{\gamma}^P_{n,\theta}$ has asymptotically a normal distribution with mean zero and variance $\sigma^2(0,\theta)$. For sequences $m \sim d\log^2 n$, $d>0$, the distribution has a bias equal to $-d^{1/2}/2$ and as d increases the bias of the estimator remains constant whereas the variance decreases.

6. On choosing the parameters

The variance of the generalized Pickands' estimator, $\sigma^2(\gamma,\theta)$, does not have the same behaviour, as a function of θ, for all real γ. Thus it is not possible to choose a value of θ in $]0,1[$ that minimizes the variance for any real γ. Anyway a value of θ can be choosen in order to improve Pickands' estimator. We can look for the value of θ that minimizes the variance of the estimator when $\gamma=0$ because of the central role the Gumbel distribution plays in extreme value theory. It is easily shown that when F is in $D_{dif}(G_0)$, $\mathrm{var}(\hat{\gamma}^P_{n,\theta})$ has a minimum for θ_0 the unique zero of the function $s(\theta) = (2+\log\theta)/\theta^2-2$ in $]0,1[$, i.e., $\theta_0 \approx 0.14$. The estimator $\hat{\gamma}^P_{n,0.14}$ is asymptotically more efficient than Pickands' estimator for (approximately)

$\gamma<1.3$. In what concerns the fraction of the sample to be used in the definition of the estimator, it does not seem possible to find an optimal criterion for choosing the value of m independently of the underlying distribution. However the estimator is expected to present always the same kind of behaviour: a great variability for small values of m, a more or less constant value for moderate values of m and, after, a significant increase of the bias for great values of m. In practical applications, estimates of γ for the different values of m should be calculated and then be considered for γ the value more or less constant corresponding to the relative stability phase of the estimates.

Acknowledgements. I am indebted to Professor J. Tiago de Oliveira for suggesting me the study of generalized Pickands' estimator and for many interesting and valuable discussions on extreme value theory. I would also like to thank Professor Ivette Gomes for her important comments and encouragement.

REFERENCES

[1] Gnedenko, B.V., Sur la distribution limite du terme maximum d'une séries aléatoire, Ann. of Math., 44 (1943), 423-453.

[2] Von Mises, R., La distribution de la plus grande de n valeurs, (1936). Reprinted in Selected Papers II. Amer. Math. Soc., Providence R.I. (1954) 271-294.

[3] De Haan, L., Slow variation and characterization of domains of attraction. In: Statistical Extremes and Applications (ed.: J. Tiago de Oliveira), Reidel, Dordrecht, 1984, pp 31-38.

[4] Pickands III, J., The continuous and differentiable domains of attraction of the extreme value distributions, Ann. Probab., 14 (1986), 996-1004.

[5] Geluk, J. L. and de Haan, L., Regular Variation, Extensions and Tauberian Theorems, Math. Centre Tracts 40. Centre for Mathematics and Computer Science, Amsterdam, 1987.

[6] Omey, E. and Willekens, E., Π-variation with remainder, J. London Math. Soc., 37(2) (1987), 105-118.

[7] Dekkers, A.L.M. and de Haan, L., On the estimation of the extreme–value index and large quantile estimation, Ann. Statist., 17 (1989), 1795-1832.

[8] Dekkers, A. L. M., On Extreme–Value Estimation, Ph.D. thesis, Erasmus University, Rotterdam, 1991.

[9] Pickands III, J., Statistical inference using extreme order statistics, Ann. Statist., 3 (1975), 119-131.

[10] Cooil, B., Limiting multivariate distributions of intermediate order statistics, Ann. Probab., 13 (1985), 469-477.

[11] Cooil, B., When are intermediate processes of the same stochastic order?, Statist.&Probab. Letters, 6 (1988), 159-162.

[12] Tiago de Oliveira, J., Extreme values and tail estimation. In: Order Statistics and Nonparametrics; Theory and Applications (ed.: I. A. Salama), Alexandria., 1991, pp 1-16.

[13] De Haan, L., On Regular Variation and its Application to the Weak Convergence of Sample Extremes, Math. Centre Tracts 32, Mathematical Centre, Amsterdam, 1970.

[14] Hall, P., On some simple estimates of an exponent of regular variation, J. Roy. Statist. Soc. Ser. B, 44 (1982), 37-42.

Normal Sample Range: Asymptotic Distribution, Approximations And Power Comparisons

Rukhin, A.L.
UMBC, Baltimore, MD

Abstract

The largest interpoint Euclidean distance gives a quick estimator of the unknown variance of a two-dimensional normal sample and also provides a short-cut test statistic for this parameter. Its asymptotic distribution and bounds for moderate sample sizes are discussed. In particular the power comparison of the optimal test and the test based on the sample range is reported.

1 Asymptotic Distribution of the Sample Range

This work was inspired by a quality control problem which arises in the manufacturer's testing of handguns. A handgun is placed in a vice and fired ten times at a target with a grid on it after which the largest interprojectile distance is determined. If the distance exceeds 4 inches the gun is rejected. The advantage of this method is that this determination is an easy task whereas the calculation of the sum of squares needed for the optimal test about the dispersion of projectiles is time consuming and not always feasible.

Thus the bivariate sample range is an

important statistic in gun quality control and it also appears rather naturally in other accuracy related problems of vector observations. Wilks obtained by the Monte Carlo method the first four moments of the bivariate normal sample range for some values of the sample size n, which are reproduced in [1] where also a chi-approximation is suggested.

In general the sample range could be used for detecting outliers (cf. Ch. 9.3 of [2]) or to provide a quick estimate of the dispersion. The monograph of Grubbs [3] discusses further statistical measures of precision which are also summarized in Section 7.5 of ref [4]. where this statistic is described as "intriguing". Despite obvious interest in the distribution of the bivariate normal sample range its asymptotic form has not been determined until recently by Matthews and Rukhin [5]. Let $X_1, X_2, \ldots$ be independent k-dimensional normal random vectors with zero mean vector and the identity covariance matrix. The sample range R is defined as the largest interpoint distance between the first n observations

$$R = R_n = \max_{1 \le i < j \le n} |X_i - X_j|.$$

The exact and asymptotic distribution of R is well known in the special case $k = 1$. Namely for any positive r with $l_2 n = \log \log n$

$$Pr\left(R \le 2[2\log n - l_2 n - \log 4\pi + r]^{1/2}\right)$$

$$\sim Pr\Big(\sqrt{2\log n}\Big[R - 2\sqrt{2\log n} + \frac{l_2 n + \log 4\pi}{\sqrt{2\log n}}\Big] \le r\Big)$$

$$\to e^{-r} \int_{-\infty}^{\infty} \exp(-e^{t-r} - e^{-t})\, dt, \quad (1)$$

i.e. the asymptotic distribution of the range is the convolution of the limiting distributions for the extreme order statistics.

A related result on the asymptotic behavior of random points with specified nearest neighbour relations was obtained in [6].

Theorem 1 *For $k \ge 2$ and positive r*

$$\lim_{n\to\infty} Pr\Big(R \le 2\Big[2\log n + \frac{1}{2}(k-3) l_2 n + l_3 n + a + r\Big]^{1/2}\Big)$$

$$= \lim_{n\to\infty} Pr\Big(\sqrt{2\log n}\Big[R - 2\sqrt{2\log n} - \frac{0.5(k-3) l_2 n + l_3 n + a}{\sqrt{2\log n}}\Big] \le r\Big)$$

$$= \exp(-e^{-r}).$$

Here

$$l_3 n = \log l_2 n$$

and

$$a = a(k) = \log \frac{(k-1) 2^{\frac{k-7}{2}}}{\Gamma(k/2) \pi^{1/2}}.$$

Let C_n denote the cardinality of the set $\{(i,j) : 1 \le i < j \le n, |X_i - X_j| \ge 2r_1\}$, i.e. the number of exceedances by the

interpoint distances of the given level $2r_1$ where

$$r_1 = [2\log n + \frac{1}{2}(k-3)l_2 n + l_3 n + a + r]^{1/2}. \quad (2)$$

The equivalence of the events $C_n > 0$ and $R > 2r_1$, and the following more general result imply Theorem 1.

Theorem 2 *As n tends to infinity, C_n converges in distribution to a Poisson random variable with parameter e^{-r}.*

The proof of Theorem 2 is based on the relationship of exceedance count C_n to the Poisson clumping heuristic argument. It turns out that the sample size n can be replaced by a Poisson number of points with mean n without affecting the asymptotic distribution. This fact allows to transform the problem into the one about a Poisson process on the k-dimensional Euclidean space possessing intensity function $n(2\pi)^{-k/2}exp(-|x|^2/2)$ with independent number of points in disjoint regions of space. A coupling argument establishes the asymptotic equivalence.

If r_1 is given by (2),

$$r_2 = [2\log n + \frac{1}{2}(k-3)l_2 n + \frac{kl_3 n}{2} + 2(a+r)]^{1/2},$$

$$r_3 = [2\log n + (k-2)l_2 n + 2l_3 n]^{1/2}$$

and

$$r_0 = 2r_1 - r_2$$

then one can show that radii r_1, r_2, r_3, r_4 and the points leading to exceedances are in a narrow annulus at $(2\log n)^{1/2} + O(l_2 n(\log n)^{-1/2})$. (These values are slightly different from the ones given in [6] where a superfluous factor 2 appears in the denominator on the third line on p. 456. The author is grateful to H. Henze who noticed that the original radii were incorrectly specified.) There are no points with $|X_i| > r_3$ with probability tending to one and that there are no pairs of points, whose lengths are between r_2 and r_3, with distance between them exceeding $2r_1$. Intuitively, points in this range are sparse enough so that their angular separations are not likely to be close enough to π (which would lead to a sufficiently large interpoint distance). It can be proven that large interpoint distances with both points within the radius r_2 and between r_0 and r_2 are also nonexistent with probability tending to one. Finally a coupling argument shows that the number of interpoint distances exceeding $2r_1$, asymptotically has a Poisson distribution.

Intuitively, a vector X_i with an exceptionally large norm could lead to a clump of exceedances of $2r_1$ when combined with vectors of large length and nearly opposite direction. Indeed it looks that this possibility prevents the moments of C_n from converging to the moments of the limiting Poisson distri-

bution. For example when $k = 2$

$$\lim EC_n = \infty.$$

2 Lower and Upper Bounds

In this section we discuss some approximations to the distribution function $F_n(r) = P(R_n \leq r)$ of the bivariate normal sample range R_n based on the random sample $X_1 = (Z_1, Y_1), \ldots, X_n = (Z_n, Y_n)$.

Notice first of all that

$$R_n = \max_{0\leq\theta<2\pi} \max_{1\leq i,j\leq n} \Big[(Z_i - Z_j)\cos\theta + (Y_i - Y_j)\sin\theta\Big]$$

$$= \max_{0\leq\theta<2\pi} \max_{1\leq i,j\leq n} \Big[(Z_i\cos\theta + Y_i\sin\theta) - (Z_j\cos\theta + Y_j\sin\theta)\Big]. \quad (3)$$

Therefore

$$R_n \geq \max\Big\{\max_{1\leq i\leq n} U_i - \min_{1\leq i\leq n} U_i, \max_{1\leq i\leq n} V_i - \min_{1\leq i\leq n} V_i\Big\}$$

where U_i, V_i, $i = 1, \ldots, n$ are independent standard normal random variables.

As was mentioned in Section 1 the distribution of the scalar normal sample range is well known. Its distribution function G_n has the form

$$G_n(w) = P\left(\max_{1\leq i\leq n} U_i - \min_{1\leq i\leq n} U_i \leq w\right)$$

$$= \frac{n}{\sqrt{2\pi}} \times \int_{-\infty}^{\infty} e^{-x^2/2}\left[\Phi(x+w) - \Phi(x)\right]^{n-1}\, dx \quad (4)$$

with the limiting distribution specified in (1).

Thus

$$F_n(r) \leq G_n(r)^2$$

and better upper estimates can be derived from (3) by considering a larger nuber of angle θ values.

Let ρ_n denote the radius of the smallest circle containing the two-dimensional normal sample. Then

$$R_n \leq 2\rho_n. \quad (5)$$

The distribuion of ρ_n is known (see refs [8,9]). Namely

$$H_n(z) = P(2\rho_n \leq z)$$

$$= n\left(1 - e^{-z^2/8}\right)^{n-1}$$

$$(n-1)\left(1 - e^{-z^2/8}\right)^n. \quad (6)$$

In other terms ρ_n has the distribution of the $(n-1)$-th order statistic in a sample of n independent $\chi(2)$ distributed random variables. Thus (5) provides a sharper bound than the maximum of such a sample which corresponds to inequality $R_n \leq 2\max_i ||X_i||$.

By combining the inequalities above one can formulate the folllowing result.

Theorem 3 *With H_n and G_n defined by (6) and (4) one has for all positive r*

$$H_n(r) \leq F_n(r) \leq G_n^2(r).$$

Bounds of Theorem 3 also lead to bounds on the moments of R_n. For instance

$$ER_n^2 < 4E\rho_n^2 = 8\left[\Psi(n+1) - \Psi(2)\right]$$

$$< 8\left[\log(n+1) - 1 + C - \frac{1}{2(n+1)}\right]$$

and for sufficiently large n

$$ER_n < \sqrt{8\log n} - \frac{2(1-C)}{\sqrt{2\log n}}.$$

Here Ψ denotes the log-derivative of gamma-function so that $\Psi(2) = 1 - C = 1 - 0.5772...$ These bounds turn out to be reasonably tight for moderate values of n.

However the question about the convergence of the moments of R_n to these of the extreme value distribution remain open. For instance it is not known if

$$\lim_{n\to\infty} \frac{ER_n^2 - 8\log n}{l_2 n} = -2.$$

The approximations of Theorem 3 show that this limit is between -4 and 0.

Figure 1. The distribution functions of F_n (solid line), H_n (dashdotted line) and G_n^2 (dotted line) for $n = 10$.

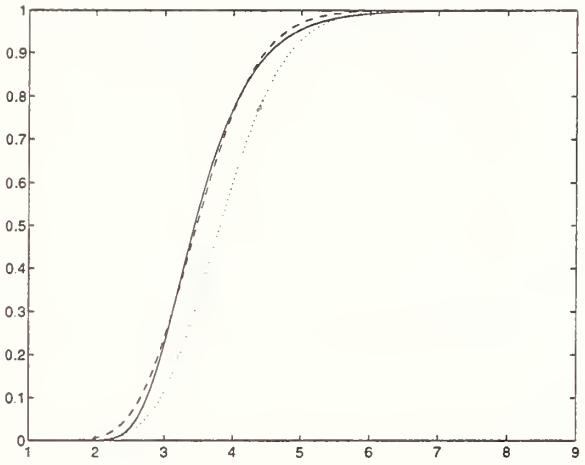


Figure 2. The distribution functions F_n (dotted line) and E_n (solid line) for $n = 10$.

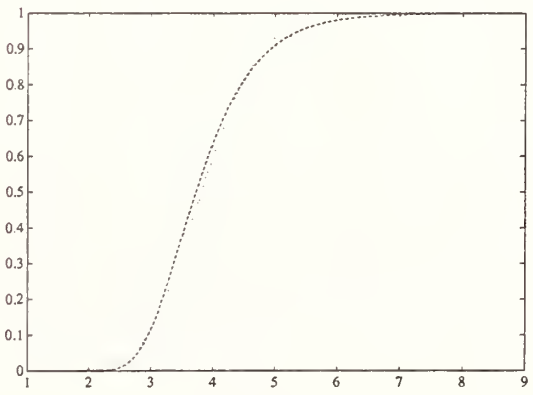


Figure 3. The distribution function of R_n^2 (dotted line) and E_n (solid line) for $n = 10$.

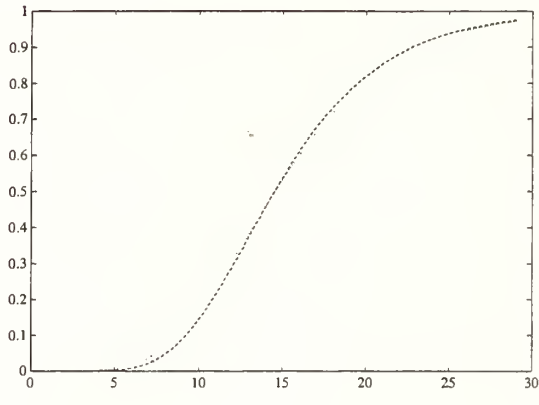


Figure 1 show for $n = 10$ the distribution functions of the lower and upper bounds H_n and G_n^2 along with the Monte-Carlo simulated empirical distribution function F_n for 10000 repetitions. Figure 2 shows F_n and the distribution function E_n of the best fitted extreme value distribution,i.e.

$$E_n(r) = \exp\{-e^{(\beta-r)/\alpha}\}$$

for α and β chosen so as to match the first two moments. Figure 3 provides the same graphs for the distribution of R_n^2. Our simulations suggest that the distribution of the squared range R_n^2 allows a better approximaton by an extreme value distribution.

The fact that extreme value distribution approximations are often more accurate for squares of extreme order statistics is actually well known in the classical asymptotic theory of normal order statistics (see [10,11]).

3 Power comparisons

Let $X_1 = (Z_1, Y_1), \ldots, X_n = (Z_n, Y_n)$ be a random sample of two-dimensional random normal vectors with independent coordinates (Z_i, Y_i) with zero means and the same unknown variance σ^2. As indicated in Section 1 the quality control problem for handguns leads to the hypothesis testing $H_0 : \sigma \leq \sigma_0$ versus the alternative $H_1 : \sigma > \sigma_0$. The optimal (uniformly most powerful) test has the critical region of the form $\{S_n^2 \geq \chi_\alpha^2(2n-2)\}$ where

$$S_n^2 = \sum_1^n \frac{\left(Z_i - \bar{Z}\right)^2 + (Y_i - \bar{Y})^2}{(2n-2)\sigma_0^2}.$$

Here $\bar{Z}, \bar{Y}$ are the coordinatewise sample means and $\chi_\alpha^2(m)$ denotes the critical point of χ^2 distribution with m degrees of freedom.

In the situation mentioned in Section 1 a handgun is rejected if in consequtive 10 shots the largest inertprojectile distance exceeds 4 inches. This procedure corresponds to a test of level 0.05 for $\sigma_0 = .79$. Indeed the Monte-Carlo simulation for $n = 10$ gives 95-th percentile of the distribution of thwe sample range to be about 5.14.

The results of numerical comparison of powers of tests based on S_n and R_n are given in Figures 4-6 for $n = 5, 10$ and 15. The power function of the test based on the sample range never falls below 65% of the power of the optimal

test.

Figure 4. Power function of the optimal test (dashed line) and the test based on the sample range (dotted line) for $n = 5$.

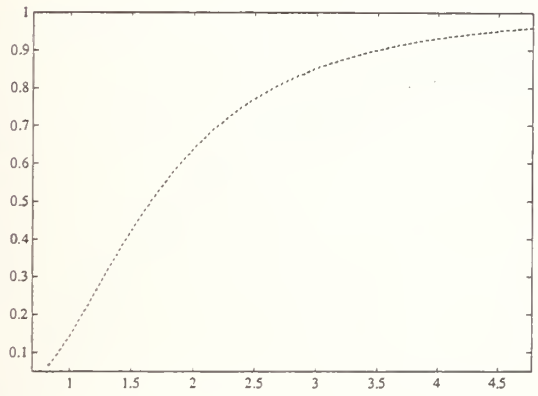


Figure 5. Power function of the optimal test (dashed line) and the test based on the sample range (dotted line) for $n = 10$.

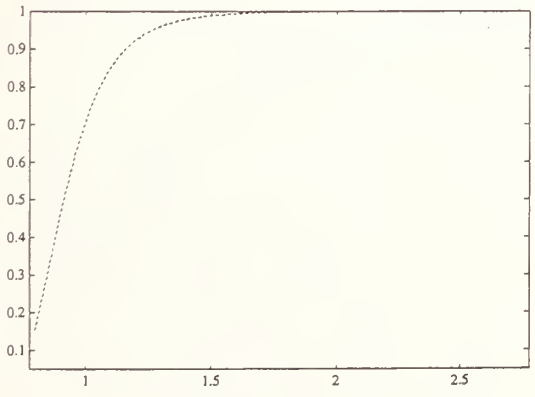
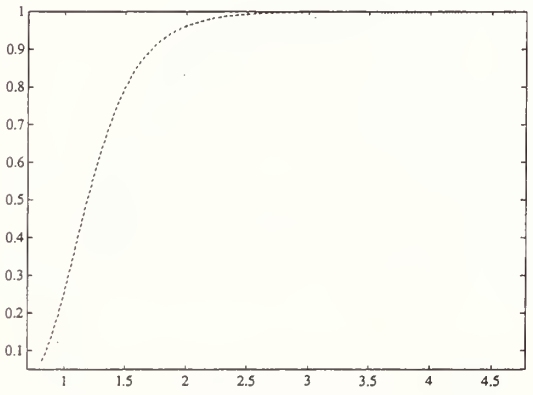


Figure 6. Power function of the optimal test (dashed line) and the test based on the sample range (dotted line) for $n = 15$.



References

[1] T. Cacoullos and H. DeCiccio. On the distribution of the bivariate range. *Technometrics*, 9:476–480, 1967.

[2] V. Barnett and T. Lewis. *Outliers in Statistical Data.* Wiley, New York, 1984, 2nd ed.

[3] F. E. Grubbs. *Statistical Measures of Accuracy for Riflemen and Missile Engineers.* Edwards Broths, Ann Arbor, 1964.

[4] H. A. David. *Order Statistics.* Wiley, New York, 1981, 2nd ed.

[5] P. Matthews and A. L. Rukhin. Asymptotic distribution of the normal sample range. *Ann. Appl. Probab*, 3:454–466, 1993.

[6] N. Henze. On the fraction of random points with specified nearest-neighbour interrelations and degree of attraction. *Adv. Appl. Prob.*, 19:873–895, 1987.

[7] H. Daniels. The covering circle of a sample from a circlar normal distribution. *Biometrika*, 39:137–143, 1952.

[8] E. J. Halteman. The Chebyshev center: a multidimensional estimate of location. *Journ. Statist. Planning Inf.*, 13:389–394, 1986.

[9] J. B. Haldane and S. D. Jayakar. The distribution of extremal and nearly extremal values in samples from a normal distribution. *Biometrika*, 50:89–94, 1963.

[10] S.B. Weinstein. Theory and applications of some classical and generalized asymptotic distributions of extreme values. *IEEE Trans. Inf. Theory*, 19:148–154, 1973.

Estimation Of Extreme Sea Levels At Major Ports In Korea

Shim, J.S., Oh, B.C. and Jun, K.C.,
Korea Ocean Research & Development Institute, Seoul, Korea

The design of coastal structures requires knowledge of the probability of extreme sea levels, as well as of extreme wave heights for safety. Two methods for computing extreme sea levels, the annual maxima method and the joint probability method, are examined for major ports (Incheon, Cheju, Yeosu, Pusan, Mukho) in Korea. The annual maxima method estimates the extreme sea levels from three different probability distributions of Gumbel, Weibull and generalized extreme value(GEV) using the least square method(LSM), the conventional moment method(CMM) and the probability weighted moment(PWM) method, respectively.

The results show that the extreme sea levels estimated by the Gumbel distribution or the least square method appear, in general, higher than those calculated by other distributions or methods. The extreme values estimated by the extreme probability method are approximately 5-10cm lower than the values estimated by the joint probability method.

1. Introduction

The rise and fall of sea level is caused by the repetitive combination of astronomical tide and storm surge. The Office of Hydrographic Affairs in Korea has observations of the sea level since 1960 at the major ports of the country.

The extreme sea levels obtained from relative long-term tidal data play a very important role not only in planning the overall layout of coastal structures, but also for fixing the positions of the intake pipes of nuclear power plants. The information about the extreme sea levels is especially important in Korea where there are large tidal ranges in the southern and western coasts of the country.

The best way to design a structure must be based upon the proper analysis of field data obtained at the possible construction site, together with the appropriate consideration of the functional and financial constraints and the life time of the structure. In general, we often perform the preliminary design of the structure through the extreme statistical analysis of hindcast data since we may not frequently accumulate enough observated data.

There are two kinds of estimation methods of the extreme sea level, that is, the annual maxima method and the joint probability method. The first method makes use of the distribution function of maximum sea level.[1], [2], [3], [4], [5], [6], [7] Therefore for a place of interest, the annual maximum for each year is extracted from hourly observed sea level and is used to estimate the parameters of the probability distributions. The latter method calculates the extreme sea levels by convoluting probability density functions of the tide and surge components on the assumption that the two components are independent of each other.[7], [8], [9], [10]

In this study, the extreme sea levels are computed by the two methods at major ports (Incheon, Cheju, Yeosu, Pusan, Mukho) of the

country with relatively long-term observation data. The annual maxima method estimates extreme sea levels from the three different probability papers of Gumbel[11], Weibull[12] and generalized extreme value(GEV)[13], each of which is prepared by applying three different methods for estimating parameters; the least square method (LSM), the conventional moment method(CMM) and the probability weighted moment(PWM) method. The joint probability method, compared with the annual maxima method, is more useful when only few years observations of sea level are available. In the annual maxima method, however, long-term data is needed. In this study, the long-term data is used in both the joint probability method and the annual maxima method in order to compare their results under the same condition. Considering the data requirements of each method likewise, we analyze and compare all the results estimated by the above methods.

2. Annual Maxima Method

The annual maxima method has been a classical method for analyzing extreme values, applied to sea level estimation since Ref. [1], [2]. In particular, this was the method used in the comprehensive study by Ref. [5].

The assumptions made in using this method are namely that hourly sea level heights are (1) independent, (2) identically distributed and (3) that the number of hours in a year is large enough for the asymptotic approximation to hold.

Table 1 gives the expressions for $F(x)$ defining the different distributions considered here, and also includes expressions for their means and variances. We estimated the parameters of the distributions by means of LSM, CMM and PWM.

2.1. Least Square Method (LSM)

This method is to provide a straight line fit to the data when it is plotted on a pertinent probability paper. This gives a slope(a) and an intercept(b) of the best-fit line $y = ax + b$ in terms of the coordinates(x_i, y_i) of all data points. The corresponding estimated values of the distribution parameters, if required, may then be obtained from the slope and the intercept by the expressions given in Table 2.

In estimating the parameters by LSM, plotting position is necessary. In this paper, the Gumbel plotting position is used as follow;

$$F(x_i) = \frac{i}{n+1} \tag{1}$$

where i denotes the rank of data, with $i = 1$ for the smallest value and n for the largest value, and n is the number of data.

2.2. Conventional Moment Method (CMM)

In this method, parameters are estimated

Table 1. Asymtotic probability distributions function

Distribution	Range	Cumulative probability	Mean	Variance
Gumbel	$-\infty < x < \infty$ $-\infty < \varepsilon < \infty$ $0 < \theta < \infty$	$\exp\left[-\exp\left\{-\left(\frac{x-\varepsilon}{\theta}\right)\right\}\right]$	$\varepsilon + \gamma^{*}\theta$ $(\approx \varepsilon + 0.58\theta)$	$\frac{\pi^2}{6}\theta^2$ $(\approx 1.64\theta^2)$
Weibull	$\varepsilon < x < \infty$ $0 < \theta < \infty$ $0 < \alpha < \infty$	$1-\exp\left\{-\left(\frac{x-\varepsilon}{\theta}\right)^{\alpha}\right\}$	$\varepsilon + \theta\Gamma\left(1+\frac{1}{\alpha}\right)$	$\theta^2\left\{\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right\}$
GEV**	$\varepsilon < x < \infty$ $0 < \theta < \infty$ $-\infty < \varepsilon < \infty$	$\exp\left[-\left\{1-\frac{\alpha}{\theta}(x-\varepsilon)\right\}^{\frac{1}{\alpha}}\right]$	$\varepsilon + \theta\{1-\Gamma(1+\alpha)\}/\alpha$	$\theta^2\{\Gamma(1+2\alpha) - \Gamma^2(1+\alpha)\}/\alpha^2$

* : γ is Euler's constant equal to 0.5772
** : if α=0, GEV equals Gumbel

Table 2. Scale relationships for probability distributions

Distribution	Abscissa scale(x)	Ordinate scale(y)	Slope(a)	Intercept(b)
Gumbel	x	$-\ln[-\ln\{F(x)\}]$	$1/\theta$	$-\varepsilon/\theta$
Weibull	$\ln(x-\varepsilon)$	$\ln[-\ln\{1-F(x)\}]$	α	$-\alpha\ln\theta$
	x	$[-\ln\{1-F(x)\}]^{\frac{1}{\alpha}}$	$1/\theta$	$-\varepsilon/\theta$
GEV	x	$[1-\{-\ln F(x)\}^{\alpha}]/\alpha$	$1/\theta$	$-\varepsilon/\theta$

from the moment of the probability density function for the distribution. The moments which are the first, second or third ones, are estimated from the sample. This method often leads to an acceptable model, since the lower moments have the stronger influence on the shape of the distribution.[14] The estimated values of the parameters are expressed in terms of $\overline{x}$, $\overline{x^2}$ and $\overline{x^3}$ as indicated in Table 3. Here $\overline{x}$, $\overline{x^2}$ and $\overline{x^3}$ are obtained directly from the data and are defined as follows;

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\quad \overline{x^2}=\frac{1}{n}\sum_{i=1}^{n}x_i^2,\quad \overline{x^3}=\frac{1}{n}\sum_{i=1}^{n}x_i^3 \qquad (2)$$

Because the Gumbel distribution has two parameters, these can be easily obtained as given in Table 3. The Weibull and the GEV distributions involve three parameters, and of these, $\hat{\alpha}$ is first estimated by equating the skewness($\sqrt{\beta}$) of the sample to those of the model. The remaining parameters can then be estimated from the first and second moments. $\sqrt{\beta}$ can be expressed as follows;

$$\sqrt{\beta}=\mu_3/\mu_2^{\frac{3}{2}}=\frac{\overline{x^3}-3\overline{x}\,\overline{x^2}+2(\overline{x})^3}{\{\overline{x^2}-(\overline{x})^2\}^{\frac{3}{2}}} \qquad (3)$$

in which μ_2 and μ_3 are the second and third central moments of the distribution. Eq. (3) is replaced by the function of $\hat{\alpha}$;

$$\sqrt{\beta}=\left\{\Gamma(1+\frac{3}{\hat{\alpha}})-3\Gamma(1+\frac{1}{\hat{\alpha}})\Gamma(1+\frac{2}{\hat{\alpha}})+2\Gamma^3(1+\frac{1}{\hat{\alpha}})\right\}\Big/\left\{\Gamma(1+\frac{2}{\hat{\alpha}})-\Gamma^2(1+\frac{1}{\hat{\alpha}})\right\}^{\frac{3}{2}} \qquad (4)$$

2.3. Probability Weighted Moment (PWM) Method

A new class of moment, called probability weighted moment, was introduced by Ref. [15]. It was indicated to be of potential interest for distributions that may be written in inverse form, that is, if X is a random variable and F is the cumulative distribution function for X, the value of X may be written as a function of F : $x = x(F)$. Reference [15] defined a probability weighted moment as;

Table 3. Parameters of distributions as estimated by conventional moment method

Distribution	Estimated parameters		
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\varepsilon}$
Gumbel	-	$\frac{\sqrt{6}}{\pi}\{\overline{x^2}-(\overline{x})^2\}^{\frac{1}{2}}$	$\overline{x}-\gamma\theta$
Weibull	$\sqrt{\beta}=f(\hat{\alpha})$	$\left\{\frac{\overline{x^2}-(\overline{x})^2}{\Gamma(1+2/\hat{\alpha})-\Gamma^2(1+1/\hat{\alpha})}\right\}^{\frac{1}{2}}$	$\overline{x}-\hat{\theta}\Gamma(1+1/\hat{\alpha})$
GEV	$\sqrt{\beta}=f(\hat{\alpha})$	$\frac{1}{\hat{\alpha}}\left\{\frac{\overline{x^2}-(\overline{x})^2}{\Gamma(1+2/\hat{\alpha})-\Gamma^2(1+1/\hat{\alpha})}\right\}^{\frac{1}{2}}$	$\overline{x}-\hat{\alpha}\hat{\theta}\{1-\Gamma(1+1/\hat{\alpha})\}$

$$M_{l,j,k} = E[X^l F^j (1-F)^k] = \int_0^1 \{x(F)\}^l F^j (1-F)^k \, dF \quad (5)$$

where l, j, k are real numbers.

If an interesting probability distribution function is substituted for F, it is possible to make an integral of Eq. (5) about the fixed set (l, j, k), and the result of the integral is expressed as the function of F's parameters. Therefore as we calculate PWM about as many sets (l, j, k) as the unknown parameters, we can get simultaneous equations for these unknown parameters.

To practice the above method, we have to decide the set (l, j, k) first of all. In the case of PWM method, since either $l=1$, $j=0$ or $l=1$, $k=0$ uses, x certainly has the first power and either F or $1-F$ is excluded. That is, the following conventional equations are adopted;

$$M_{(k)} = M_{1,0,k} = \int_0^1 x(1-F)^k \, dF \quad (6)$$

$$M_j = M_{1,j,0} = \int_0^1 x F^j \, dF \quad (7)$$

There are a lot of differences between the above equations and the conventional moment below.

$$M_r = \int_{-\infty}^{\infty} x^r f(x) \, dx \quad (8)$$

One of the differences is that the Eq. (8) has the operation of the rth power about x so that the observation error or abnormal values are amplified as the power gets higher, but since PWM of Eq. (6) and (7) places its operation in cumulative distribution function F, the sampling errors or abnormal values become smaller. This effect is, however, estimated from the definition of the moment merely, and there has been no definitive evaluation of this approach.[16]

If j and k are nonnegative integers, then

$$M_{l,0,k} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j M_{l,j,0}$$
$$M_{l,j,0} = \sum_{k=0}^{j} \binom{j}{k} (-1)^k M_{l,0,k} \quad (9)$$

In the special case where l, j and k are nonnegative integers, $M_{l,j,k}$ is proportional to $E(X^l_{j+1,k+j+1})$, the lth moment about the origin of the $(j+1)$th order statistic for a sample of size $k+j+1$.[17] More specifically,

$$M_{l,j,k} = B(j+1, k+1) \; E[X^l_{j+1,\,k+j+1}] \quad (10)$$

where $B(\cdot,\cdot)$ denotes the beta function. For $j=0$ and $l=1$ the convention

$$M_{(k)} = M_{1,0,k} = B(1,k+1) E[X_{1,k+1}] \quad (11)$$

is adopted. $\widehat{M}_{(k)}$, unbiased estimate of $M_{(k)}$ from a sample size of n and where k is a nonnegative integer, is obtained as follows;

$$\widehat{M}_{(k)} = \frac{1}{k+1} \sum_{i=1}^{n} x_i \binom{n-i}{k} \Big/ \binom{n}{k+1} = \frac{1}{n} \sum_{i=1}^{n-k} x_i \binom{n-i}{k} \Big/ \binom{n-1}{k} \quad (12)$$

And also, for $k=0$ and $l=1$, the Eq. (10) becomes

$$M_j = M_{1,j,0} = B(j+1, 1) \; E[X_{j+1,\,j+1}] \quad (13)$$

To get the unbiased estimate $\widehat{M}_j$ of M_j, we need the $E(X_{j+1,j+1})$, the first moment about the origin of the $(j+1)$th order statistic for a sample of size $j+1$. In drawing randomly $j+1 (n > j+1)$ of the sample $x_1, x_2, \cdots, x_n$, the probability of which the maximum value is x_i is $\binom{i-1}{j} / \binom{n}{j+1}$. Therefore $E(X_{j+1,j+1})$ is

$$E[X_{j+1,\,j+1}] = \lim_{n\to\infty} \sum_{i=1}^{n} x_i \binom{i-1}{j} \Big/ \binom{n}{j+1} \quad (14)$$

$\widehat{M}_j$ can be estimated from the Eq. (13) and (14) as follows;

$$\widehat{M}_j = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i \frac{(i-1)(i-2)\cdots(i-j)}{(n-1)(n-2)\cdots(n-j)} & (j \geq 1) \\ \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i & (j=0) \end{cases} \quad (15)$$

When the sample of size n is arranged from x_1 to x_n in ascending order, either an estimate $\widehat{M}_{(k)}$ from the Eq. (12) or $\widehat{M}_j$ from the Eq. (15) can be made. And then the solutions of PWM are obtained by substituting them into the simultaneous equations of the parameters.

The probability weighted moments, $M_{l,0,k}$ or $M_{l,j,0}$ of three distributions are given in Table 4, and the parameters of each distribution are shown as M_j and $M_{(k)}$ in Table 5. From Table 3 and Table 5, Gumbel's parameters are defined explicitly as the functions of both conventional moment and PWM. On the other hand, those of Weibull and GEV can be explicitly expressed as

the function of PWM only, not as conventional moment.

Table 4. Expressions of probability weighted moment

Distribution	PWM $M_{1,j,k}$ (real $j,k \geq 0$)
Gumbel	$M_{1,j,0} = M_j = \frac{\varepsilon}{1+j} + \frac{\theta\{\ln(1+j)+\gamma\}}{1+j}$
Weibull	$M_{1,0,k} = M_{(k)} = \frac{\varepsilon}{1+k} + \frac{\theta\Gamma(1+1/\alpha)}{(1+k)^{1+1/\alpha}}$
DEV	$M_{1,j,0} = M_j = \frac{\varepsilon+\theta\{1-(j+1)^{-\alpha}\Gamma(1+\alpha)\}/\alpha}{j+1}$

Table 5. Parameters expressions of probability weighted moment

Distribution	Parameter	PMW $M_{(k)}, M_j$
Gumbel	$\hat{\varepsilon}$	$M_0 - \gamma\theta$
	$\hat{\theta}$	$(2M_1 - M_0) / \ln(2)$
Weibull	$\hat{\varepsilon} = 0$	0
	$\hat{\theta}$	$M_{(0)} / \Gamma[\ln\{M_{(0)} / M_{(1)}\} / \ln(2)]$
	$\hat{\alpha}$	$\ln(2) / \ln\{M_{(0)} / 2M_{(1)}\}$
	$\hat{\varepsilon} \neq 0$	$\frac{4(M_{(3)}M_{(0)} - M_{(1)}^2)}{4M_{(3)} + M_{(0)} - 4M_{(1)}}$
	$\hat{\theta}$	$\frac{M_{(0)} - \varepsilon}{\Gamma\left\{\ln\left(\frac{M_{(0)} - 2M_{(1)}}{M_{(1)} - 2M_{(3)}}\right) \Big/ \ln 2\right\}}$
	$\hat{\alpha}$	$\ln(2) / \ln\left\{\frac{M_{(0)} - 2M_{(1)}}{2(M_{(1)} - 2M_{(3)})}\right\}$
DEV	$\hat{\alpha}$	$\frac{\ln\{(M_0 - 2M_1)/2(M_1 - 2M_3)\}}{\ln 2}$
	$\hat{\theta}$	$\alpha(2M_1 - M_0) / \{\Gamma(1+\alpha)(1-2^{-\alpha})\}$
	$\hat{\varepsilon}$	$M_0 - \theta\{1 - \Gamma(1+\alpha)\} / \alpha$

3. Joint Probability Method

In this method, the hourly observed data is separated into mean sea level, tide and surge components. Then the probability density functions of the tide and surge components are convoluted to obtain the probability of a particular sea level, incorporating return period levels. This method has advantages over the annual maxima method in that the method can estimate the extreme sea levels even with short-term observed sea level data and can get the low extreme sea level as well. In addition, this method can evade difficulties for establishing a suitable distribution function and for estimating unbiased parameters in the annual maxima method.

Any instantaneous value of sea level $\zeta(t)$ measured from a defined datum may be considered to be a sum of three independent components; mean sea level $Z_0(t)$, tidal level $X(t)$ and residual or surge level $Y(t)$.

$$\zeta(t) = Z_0(t) + X(t) + Y(t) \qquad (16)$$

Annual mean sea level for a particular year can be determined from hourly observed data, and is removed from the data. However, annual mean sea levels are not constant.

The tidal component of the sea level data is directly or indirectly affected by astronomical forcing. It is also removed from the observed sea level data expressed as the finite sum of harmonic constants which have the following form;

$$X(t) = \sum_{n=1}^{n} f_n H_n \cos[\sigma_n t + (v_n + \mu_n) + g_n] \qquad (17)$$

where H_n is the amplitude of the nth constituent, σ_n is its angular speed defined astronomically, v_n is its equilibrium tidal phase at $t = 0$, g_n is the phase lag of the constituent on the equilibrium tide, and f_n and μ_n are the nodal corrections. Removal of the tidal component as above does not require an excessive length of record as a satisfactory tidal analysis can be obtained from even one year observation.

Once both the tide and the mean sea level are removed from the observed data, only the surge or non-tidal component remains. Over a sufficiently long period the surge is a random variable. Obviously over a short period like a month, however, very few surges are likely to occur and produce random phases.

If the tide and surge components are then independent each other at any time t, the sea level relative to the mean sea level, i.e. $w = \zeta(t) - Z_o(t)$, may be regarded as the sum of two independent components $x = X(t)$ and $y = Y(t)$. Thus if the probability density functions of the tidal and surge components are $f_t(x)$ and $f_s(y)$ respectively, then the probability

density function $f(w)$ of w is

$$f(w) = \int_{-\infty}^{\infty} f_t(x)\, f_s(y)\, dy = \int_{-\infty}^{\infty} f_t(w-y)\, f_s(y)\, dy \quad (18)$$

The omission of dependence on time, t, when replacing $X(t)$ by x etc. implies an assumption of stationarity for the series $X(t)$ involved. The hourly predicted tide series is generally considered to be stationary, but the hourly residual series is, to some degree, nonstationary since seasonal and meteorological effects like storm surge will give rise to series of residual not randomly distributed in time.[10]

The probability of exceedance of a particular sea level η may be evaluated from the corresponding cumulative distribution function $F_\zeta(\eta) = \mathrm{Prob}(\zeta \le \eta)$ defined as follows;

$$1 - F_\zeta(\eta) = \int_{\eta}^{\infty} f(w)\, dw = \int_{\eta}^{\infty} \int_{-\infty}^{\infty} F_t(w-y)\, f_s(y)\, dy\, dw \quad (19)$$

The alternative form;

$$1 - F_\zeta(\eta) = 1 - \int_{-\infty}^{\infty} F_t(\eta - y)\, f_s(y)\, dy \quad (20)$$

where $F_t(\cdot)$ is the cumulative distribution function of tide.

Reference [8] suggested that the return period in year of a particular sea level η is expressed as;

$$R_p = 1/[\{1 - F_\zeta(\eta)\}\lambda] \quad (21)$$

where λ is the number of data used per unit time, and has 1.0(annual maximum value in a year), 8766(average number of hourly values in a year) for the annual maxima method and the joint probability method, respectively.

4. Evaluation of Extreme Sea Levels

The Korean Peninsula is enclosed by the Yellow(West) Sea, the South Sea and the Sea of Japan(East Sea). In the Yellow Sea, the bottom topography is very flat and the average water depth is about 40m. Its tidal range is very large and increases to the north, reaching about 8.0m at Incheon. In the South Sea, the average water depth is about 100m and the bottom is fairly flat. The tidal range increases to the west recording about 3.0m at Yeosu. The Sea of Japan whose average depth is about 1500m has monotonous shorelines and very steep shelves. The tidal range is about 0.3m only and increases to the south, reaching about 1.2m at Pusan.

The lengths of records are 29 years at Incheon, 27 years at Cheju, 25 years at Yeosu, 30 years at Pusan and 22 years at Mukho, respectively. Their locations are shown in Fig. 1. Editing tidal data is necessary to eliminate erroneously recorded data prior to evaluating the extreme sea levels. The editing method chosen to examine the sea level records fundamentally consist of plotting the observed sea level and the surge as a function of time and of examining the plotted values by eye for detecting errors marked as irregularities and spikes. The errors are then corrected by referring to the original tide gauge chart.

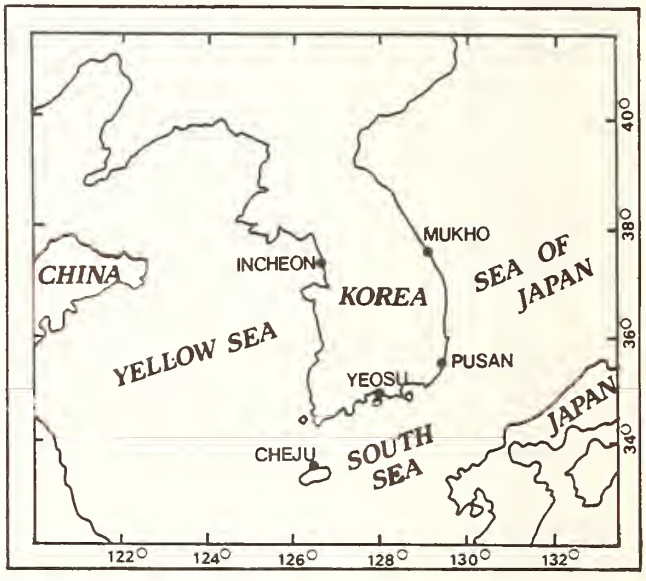


Fig. 1 Location map

The extreme sea levels are evaluated by the above-mentioned various methods, and are expressed in Table 6. The result of the annual maxima method at Incheon is shown in Fig. 2. In this figure, the lines representing CMM and PWM are plotted in the probability paper after estimating unbiased parameters and have no connection with plotting position.

As given in Table 6, the extreme sea levels calculated by the joint probability method are

Table 6. Extreme sea levels with return periods estimated with various methods
(a)Incheon, (b)Cheju, (c)Yeosu, (d)Pusan, (e)Mukho

(a) Incheon (Sea level relative to defined datum, Unit : cm)

Return period (yr)	Annual maxima method									Joint probability method		Remark
	Gumbel			Weibull			GEV			High extreme sea level	Low extreme sea level	Max.(min.) observed
	LSM	CMM	PWM	LSM	CMM	PWM	LSM	CMM	PWM			
29	986.0	982.1	984.0	980.3	978.9	978.9	980.9	978.3	979.6	982.6	-128.5	984.0 (-102.0)
50	991.8	987.2	989.5	983.3	981.9	981.7	984.0	982.3	982.6	987.6	-135.3	
100	999.2	993.8	996.6	986.8	985.3	985.0	987.3	985.7	985.9	993.9	-143.6	
200	1006.2	1000.3	1003.5	989.9	988.4	987.9	990.2	988.6	988.7	1000.2	-151.4	
300	1010.9	1004.1	1007.6	991.6	990.1	989.5	991.7	990.1	990.1	1003.8	-155.8	

(b) Cheju

Return period (yr)	Annual maxima method									Joint Probability method		Remark
	Gumbel			Weibull			GEV			High extreme sea level	Low extreme sea level	Max.(min.) observed
	LSM	CMM	PWM	LSM	CMM	PWM	LSM	CMM	PWM			
27	320.3	318.2	318.6	319.9	317.9	317.4	319.9	317.9	318.3	330.5	-52.4	324.0 (-48.0)
50	323.8	321.3	321.7	322.8	320.3	319.5	323.1	320.6	321.1	334.3	-55.7	
100	327.6	324.6	325.2	325.8	322.8	321.6	326.6	323.4	324.2	338.7	-59.6	
200	331.5	328.0	328.7	328.8	325.1	323.6	330.0	327.6	328.9	342.9	-63.5	
300	333.7	330.0	330.7	330.4	326.4	324.7	332.0	328.6	330.1	345.3	-65.7	

(c) Yeosu

Return period (yr)	Annual maxima method									Joint probability method		Remark
	Gumbel			Weibull			GEV			High extreme sea level	Low extreme sea level	Max.(min.) observed
	LSM	CMM	PWM	LSM	CMM	PWM	LSM	CMM	PWM			
25	419.3	416.5	417.6	418.5	415.5	417.6	419.2	415.8	417.5	423.5	-54.9	416.0 (-57.0)
50	424.5	421.0	422.4	422.7	418.6	421.7	424.3	419.0	422.3	428.5	-58.4	
100	429.6	425.4	427.2	426.6	421.4	425.6	429.4	422.0	427.0	434.2	-61.9	
200	434.7	429.9	431.9	430.2	424.0	429.3	434.4	424.7	431.0	441.1	-65.4	
300	437.7	432.5	434.7	432.3	425.5	431.4	437.3	426.2	434.3	445.6	-67.5	

(d) Pusan

Return period (yr)	Annual maxima method									Joint probability method		Remark
	Gumbel			Weibull			GEV			High extreme sea level	Low extreme sea level	Max.(min.) observed
	LSM	CMM	PWM	LSM	CMM	PWM	LSM	CMM	PWM			
30	169.9	168.2	168.4	169.9	168.5	167.9	169.5	168.2	168.5	176.5	-43.9	174.0 (-41.0)
50	172.4	170.4	170.7	172.3	170.5	169.7	172.0	170.3	170.8	179.2	-45.4	
100	175.7	173.4	173.7	175.3	173.1	172.0	175.4	173.2	173.8	183.3	-47.2	
200	179.1	176.3	176.7	178.2	175.6	174.1	178.6	176.0	176.9	187.7	-48.9	
300	181.0	178.1	178.4	179.9	177.0	175.3	179.5	177.7	178.7	190.5	-49.8	

(e) Mukho

Return period (yr)	Annual maxima method									Joint probability method		Remark
	Gumbel			Weibull			GEV			High extreme sea level	Low extreme sea level	Max.(min.) observed
	LSM	CMM	PWM	LSM	CMM	PWM	LSM	CMM	PWM			
22	77.8	75.0	76.2	76.8	74.3	75.5	76.9	74.4	75.5	87.2	-33.8	
50	83.3	79.8	81.3	81.0	77.5	79.4	81.6	77.8	79.7	91.1	-35.5	77.0 (-29.0)
100	88.0	83.7	85.7	84.2	79.9	82.4	85.2	80.4	83.0	93.9	-36.8	
200	92.7	87.7	90.0	87.2	82.2	85.2	88.7	82.8	86.2	96.7	-37.8	
300	95.4	90.0	92.5	88.9	83.4	86.7	90.6	84.1	88.0	98.1	-38.4	

LSM : Least square method, CMM : Conventional moment method, PWM : Probability weighted moment method

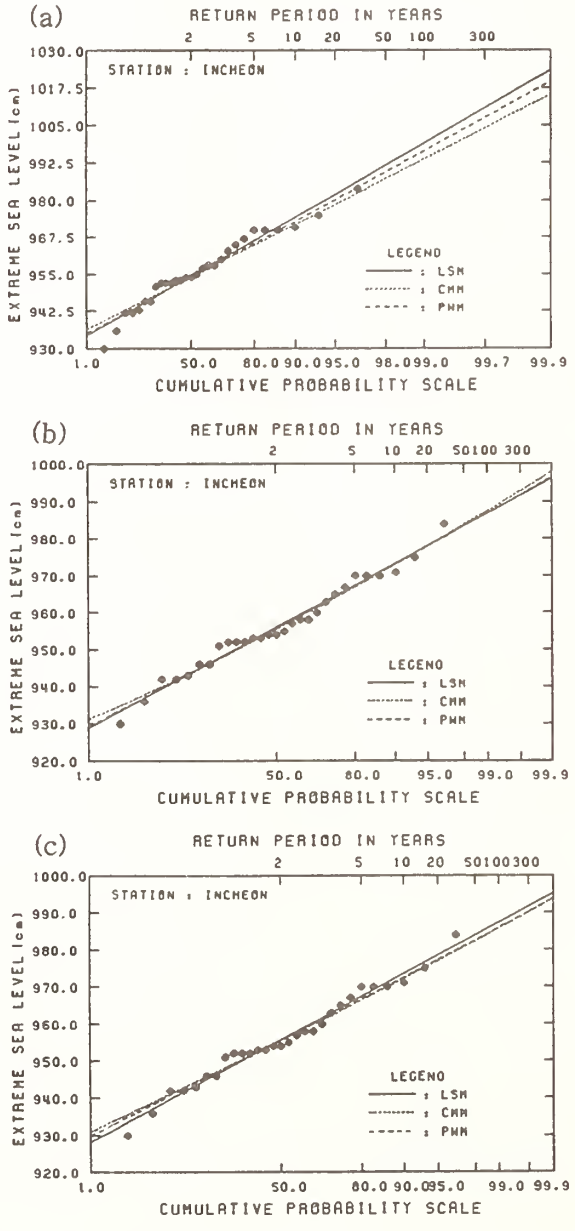


Fig. 2 Probability distributions of extreme sea levels at Incheon (a) Gumbel, (b) Weibull, (c) GEV

approximately 5-10cm higher than those by the annual maxima method at every port. This is clear from the fact that the surges are auto-correlated, i.e., successive hourly samples of the time series are not mutually independent.[8], [10] Surges persist for more than one hour. Another reason is that large surges tend to not occur with extreme tide levels and so the probability of an extreme total level due to a combination of extreme tide and extreme surge is lower than in that the case of their independence.

When using the same method for estimating the parameters as in the annual maxima method, the estimated extreme sea levels in Gumbel distribution tend to be higher than those of other distributions, and the estimated extreme sea levels by the least square method tend to be higher than those of the other methods of estimating parameters. In the case of annual maxima method, the extreme sea levels computed by the least square method in the Gumbel distribution are the highest of all, and the CMM and PWM methods of the Weibull distribution tend to get lower values. In evaluating of the extreme sea levels these relative deviations shown in Table 6 vary depending on not only what kinds of distributions are taken but also how the parameters are estimated.

The smallest extreme sea level for a certain return period is less than that from the method giving the highest value corresponding to one-third of the return period. For instance, the smallest value(989.5) for 300 years return period at Incheon is comparable to the value corresponding to 100 years or less of return period. Therefore to obtain the extreme sea levels which are essential for planning or

designing a coastal structure, it seems desirable to get the extreme sea levels, by several methods. One of the three statistical values (maximum, minimum and mean) of the estimated extreme sea levels can be taken as a design sea level for a return period considering the functional and economical aspects of the coastal structures.

5. Conclusions

Estimation methods of extreme sea level with data of relatively short term compared to structural life time were presented here to provide a design criteria needed in planning or designing coastal structures. The methods were applied to the tidal data recorded at several major ports in Korea, and results from each method were analyzed and compared. The major conclusions are summarized as follows;

1) The extreme sea levels by the joint probability method are approximately 5-10cm higher than those by the annual maxima method.

2) In the case of annual maxima method, the extreme sea levels are higher in the Gumbel distribution than those in any other distributions when the same method for estimating parameters is employed.

3) For the same distribution, the extreme sea levels evaluated by the least square method tend to be higher than those by any other methods.

4) The estimated smallest extreme sea level for a certain return period is less than the estimate based on the method giving the highest value corresponding to one-third of the return period.

The extreme sea levels are necessary for planing or designing the large scale coastal developments such as new airport construction, artificial island construction, and reclamation in the coastal zone. It is often the case that long-term sea level record may not be available for a specific site of possible coastal development. Even with short-term data the joint probability method can give a statistically acceptable extreme sea levels. Even if the method is generally found to give a slightly overestimated value, it is acceptable in a conservative sense from the engineering point of view.

6. References

[1] Lennon, G. W., A frequency investigation of abnormally high tide levels at certain west coast ports, Proc. Instn Civ. Engrs, 25(1963), 451-483.

[2] Suthons, C. T., Frequency of occurrence of abnormally high sea levels on the east and south coasts of England, Proc. Instn Civ. Engrs, 25(1963), 433-450.

[3] Blackman, D. L. and Graff, J., The analysis of annual extreme sea levels at certain ports in southern England, Proc. Instn Civ. Engrs., Part 2, 65(1978), 339-357.

[4] Graff, J., Concerning the recurrence of abnormal sea levels, Coastal Engng, 2(1979), 177-187.

[5] Graff, J., An investigation of the frequency distributions of annual sea level maxima at ports around Great Britain, Estuarine, Coastal and Shelf Science, 12(1981), 389-449.

[6] Graff, J. and Blackman, D. L., Analysis of maximum sea levels in southern England, Proc. 16th Coastal Engng Conf., Hamburg, ASCE, 931-947.

[7] Shim, J. S., Kim, S. I. and Kang, S. W., Comparison of computing methods for extreme sea levels, Ocean Research, KORDI, Korea, 11(1989), 43-50.

[8] Pugh, D. T. and Vassie, J. M., Extreme sea levels from tide and surge probability, Proc. 16th Coastal Engng Conf., Hamburg, ASCE, 911-930.

[9] Pugh, D. T. and Vassie, J. M., Applications of the joint probability method for extreme sea level computations, Proc. Instn Civ. Engrs, Part 2, 69(1980), 959-975.

[10] Walden, A. T., Prescott, P. and Webber, N. B., An alternative approach to the joint probability method for extreme high sea level computations, Coastal Eng., 6(1982), 71-82.

[11] Gumbel, E. J., Statistics of extremes, Columbia Univ. Press, New York, 1958.

[12] Weibull, W., A statistical distribution function of wide applicability, J. Applied Mech., 18(1951), 293.

[13] Jenkinson, A. F., The frequency distribution of the annual maximum (minimum) values of meteorological elements, Quart. J. Roy. Meteor. Soc., 81(1955), 158-171.

[14] Isaacson, M.de St. Q. and MacKenzie, N. G., Long-term distribution of ocean waves : A review, J. Waterways, Port, Coastal and Ocean

Div., ASCE, 107(1981), 93-109.
[15] Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R., Probability weighted moments, Water Resour. Res., 15(1979), 1049-1054.
[16] Takeuchi, K. and Tsuchiya, K., PWM solutions to normal, lognormal and Pearson-III distributions, JSCE, 393(1988), 95-101.
[17] Landwehr, J. M., Matalas, N. C. and Wallis, J. R., Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantities, Water Resour. Res., 15(1979), 1055-1064.

Limit Properties Of Maxima Of Weighted I.I.D. Random Variables

Tomkins, R.J.
University of Regina, Regina, Saskatchewan, Canada

Let $Z_n = \max\{a_1X_1, \cdots, a_nX_n\}, n \geq 1$, where $\{a_n\}$ is a positive real sequence and $X_1, X_2, \cdots$ form a sequence of independent, identically-distributed random variables. Define $Z = \lim_{n\to\infty} Z_n$. It will be shown that $P[Z < +\infty] = 0$ or 1. Necessary and sufficient conditions will be given for Z to be finite almost surely, or to be almost surely constant. This work is a preliminary step in the study of the stability of the sequence $\{Z_n\}$.

1. Introduction

Throughout this paper, $X_1, X_2, \cdots$ will be a sequence of independent, identically-distributed (i.i.d.) random variables (r.v.) with common distribution function (d.f.) $F(x)$. Define

$$x_0 = \sup\{x : F(x) < 1\}; \tag{1.1}$$

note that x_0 is well-defined, and $x_0 \leq +\infty$.

During the past half-century, a good deal of attention has been paid to the limiting behaviour of the sequence of maxima $\{M_n, n \geq 1\}$, where

$$M_n = \max\{X_1, \cdots, X_n\}. \tag{1.2}$$

The sequence $\{M_n\}$ is said to be *relatively stable* (respectively, *almost surely* (a.s.) *stable*) if a real sequence $\{b_n\}$ exists such that

$$\frac{M_n}{b_n} \to 1$$

in probability (resp., a.s.) as $n \to \infty$. Necessary and sufficient conditions for relative stability and for a.s. stability are well-known; see Gnedenko (Ref. [1]), Barndorff-Nielsen (Ref. [2]), and Resnick and Tomkins (Ref. [3]).

It is natural to wonder if analogous results exist for the case where the X_n's are independent, but not necessarily identically distributed. As a first step in achieving such a generalization of the i.i.d. results, it seems reasonable to focus on the maximum sequence $\{Z_n, n \geq 1\}$, where

$$Z_n = \max\{a_1X_1, \cdots, a_nX_n\} \tag{1.3}$$

for some positive real sequence $\{a_n, n \geq 1\}$. Obviously, M_n and Z_n are one and the same if $a_n = 1$ for all n. Since Z_n is non-decreasing in n, it makes sense to define

$$Z = \lim_{n\to\infty} Z_n. \tag{1.4}$$

The limiting behaviour of M_n is straightforward; it is easy to see that

$$\lim_{n\to\infty} M_n = x_0 \text{ a.s.} \tag{1.5}$$

In other words, the a.s. limit of the M_n-sequence is the right-hand end-point of the support of $F(x)$; thus the stability problem for $\{M_n\}$ is trivial and of little interest in the case where x_0 is

finite. But, not surprisingly, the behaviour of Z_n is somewhat more complex.

For example, if the X_n-sequence is uniformly-distributed on $(0,1)$, and if $a_n = n, n \geq 1$, then it follows from the Borel Zero-One Law that, for every $M > 0$,

$$P[nX_n > M \text{ infinitely often (i.o.)}] = 1,$$

and hence it follows that $Z = +\infty$ a.s. in this case, even though x_0 is finite ($x_0 = 1$). On the other hand, if each X_n is exponentially distributed with mean one and $a_n = 1/\log(n+1)$, then

$$P[a_n X_n > 2 \text{ i.o.}] = 0$$

by the Borel-Cantelli Lemma, from which it is clear that Z is a.s. finite in this case, even though $x_0 = +\infty$.

The goal of this paper is to determine completely the properties of Z. It will be shown that $P[|Z| < +\infty] = 0$ or 1, whatever the value of x_0 may be. Section 2 will deal with the case where $x_0 = +\infty$; a criterion for Z to be a.s. finite in this case will be given, and it will be shown and that Z cannot be a.s. constant in this case. Section 3 will present necessary and sufficient conditions for Z to be a.s. finite, and for Z to be a.s. constant, in the case where x_0 is finite. The paper will conclude with some results and remarks on stability questions in Section 4.

2. Properties of Z when $x_0 = +\infty$

Throughout this section, it will be assumed that $x_0 = +\infty$; that is, $F(x) < 1$ for all real x. Some fundamental properties of Z will now be presented.

Theorem 2.1. Let $X_1, X_2, \cdots$ be i.i.d. r.v. with d.f. $F(x)$ such that $F(x) < 1$ for all x. Let $\{a_n\}$ be a positive real sequence, and define Z by (1.4). Then

(i) $Z > \gamma$ a.s. for some real γ if

$$\sum_{n=1}^{\infty} \{1 - F(\gamma a_n^{-1})\} = +\infty; \tag{2.1}$$

(ii) $Z = +\infty$ a.s. if (2.1) holds for every real γ;

(iii) if

$$\sum_{n=1}^{\infty} \left\{1 - F(\gamma a_n^{-1})\right\} < \infty \tag{2.2}$$

for some γ, then Z is a.s. finite, Z is non-degenerate and $P[Z \leq \gamma] > 0$; and

(iv) If Z is a.s. finite then $a_n \to 0$ as $n \to \infty$.

Proof. If (2.1) holds for some γ, then $P[a_n X_n > \gamma \text{ i.o.}] = 1$ by the Borel Zero-One Law. It follows that

$$\begin{aligned} P[Z > \gamma] &= P[Z_n > \gamma \text{ i.o.}] \\ &\geq P[a_n X_n > \gamma \text{ i.o.}] = 1; \end{aligned}$$

i.e., $Z > \gamma$ a.s., establishing (i). Part (ii) follows easily from (i).

Now suppose that (2.2) holds for some γ. Recalling that, for $0 < c_n < 1$, $n \geq 1$, $\prod_{n=1}^{\infty} c_n$ converges (to a positive number) if and only if (iff) $\sum_{n=1}^{\infty} (1 - c_n) < \infty$, it follows that $\prod_{n=1}^{\infty} F(\gamma a_n^{-1}) > 0$, and $P[a_n X_n > \gamma \text{ i.o.}] = 0$ by the Borel-Cantelli Lemma. Hence, it makes sense to define the r.v. $N = N_\gamma$ as follows:

$$N_\gamma = \begin{cases} \max\{n : a_n X_n > \gamma\} \\ \quad \text{if } a_n X_n > \gamma \text{ for some } n \\ 0 \text{ if } a_n X_n \leq \gamma \text{ for all } n \geq 1 \end{cases}. \tag{2.3}$$

Consequently,

$$\begin{aligned} P[Z \leq \gamma] &= P[N = 0] \\ &= P[a_n X_n \leq \gamma \text{ for all } n \geq 1] \\ &= \prod_{n=1}^{\infty} F(\gamma a_n^{-1}) > 0. \end{aligned}$$

On the other hand,

$$P[Z > \gamma] = P[N \geq 1] \geq P[a_1 X_1 > \gamma]$$
$$= 1 - F(\gamma a_1^{-1}) > 0,$$

so Z is not degenerate. Moreover, $Z = Z_k$ on the event $[N = k], k \geq 1$, and $Z \leq \gamma$ on $[N = 0]$, so Z is a.s. finite. This proves (iii).

Finally, if Z is a.s. finite, it follows from (ii) that (2.2) holds for some real γ. But then, of necessity, $F(\gamma a_n^{-1}) \to 1$ as $n \to \infty$. Since $x_0 = +\infty$ by hypothesis, it follows that $a_n \to 0$, proving part (iv). □

Remarks. 1. It is an easy consequence of Theorem 2.1 that $P[Z = +\infty]$ is zero or one. This is no surprise, since $[Z = +\infty] = [\lim_{n\to\infty} \max(a_N X_N, \cdots, a_n X_n) = +\infty]$ for every $N \geq 1$, so that $[Z = +\infty]$ is a tail event and, hence, has probability zero or one by Kolmogorov's Zero-One Law.

2. It is evident from Theorem 2.1 that Z is a.s. finite iff (2.2) holds for some $\gamma > 0$.

3. The converse to part (iv) is not generally true. For instance, if $F(x) = 1 - e^{-x}, x > 0$, and $a_n = (\log\log(n+2))^{-1}$ for $n \geq 1$, then $a_n \to 0$ and $x_0 = +\infty$. But the series in (2.2) equals $\sum_{n=3}^{\infty} (\log\ n)^{-\gamma}$, which diverges for all γ. Hence $Z = +\infty$ a.s. by Theorem 2.1 (ii).

4. In the special case where $\{a_n\}$ is a monotone sequence, parts (i) and (ii) of Theorem 2.1 can be derived with the aid of a theorem of Mucci (Ref. [4], or see Theorem 4.4.1 of Galambos's book (Ref. [5])).

Corollary 2.2. Let a_n, X_n, Z and F be as given in Theorem 2.1. Define γ_0 to be the infimum of all γ, if any, such that (2.2) holds and let $\gamma_0 = +\infty$ if (2.1) holds for all γ. Then $Z \geq \gamma_0$ a.s. and $P[Z \leq \gamma_0 + \varepsilon] > 0$ for every $\varepsilon > 0$.

Proof. If $\gamma_0 = +\infty$ then $Z = +\infty = \gamma_0$ by Theorem 2.1 (ii); the second part is trivial in this case. If $\gamma_0 < \infty$ then $Z > \gamma_0 - \varepsilon$ a.s. for every $\varepsilon > 0$ by Theorem 2.1 (i); hence $Z \geq \gamma_0$ a.s. An application of part (iii) of the same theorem concludes the proof. □

Remark. 5. If $\gamma_0 < \infty$, it follows from Corollary 2.2 and the proof of Theorem 2.1 (iii) that, for any $\varepsilon > 0$,

$$\gamma_0 \leq Z \leq (\gamma_0 + \varepsilon) I(N = 0) + Z_N I(N \geq 1),$$

where $N = N_{\gamma_0+\varepsilon}$ is defined by (2.3).

3. Properties of Z when x_0 is finite

Throughout this section, it will be assumed that $x_0 < +\infty$; that is, that X_1 is a.s. bounded above. A simple criterion for Z to be a bona fide r.v. when $x_0 > 0$ will now be presented in Theorem 3.1, which also provides a necessary and sufficient condition for Z to be a.s. constant.

Theorem 3.1. Let $X_1, X_2, \cdots$ be a sequence of i.i.d. r.v. with d.f. F, and let $\{a_n\}$ be a positive sequence. Define x_0 and Z by (1.1) and (1.4) respectively, and assume $0 < x_0 < +\infty$. Then

(i) $Z = +\infty$ a.s. iff $\limsup_{n\to\infty} a_n = +\infty$;

(ii) $Z < +\infty$ a.s. iff $\limsup_{n\to\infty} a_n < \infty$; in which case $Z \leq x_0 \sup_{n\geq 1} a_n$; and

(iii) Z is a.s. constant iff $\sup_{n\geq 1} a_n < \infty$ and either X_1 is degenerate or $\limsup_{n\to\infty} a_n = \sup_{n\geq 1} a_n$. In either case, $Z = x_0 \sup_{n\geq 1} a_n$.

Proof. Since $x_0 > 0$ by hypothesis, $P[X_1 > 0] > 0$. Therefore, $P[X_{n_k} > 0$ i.o.$] = 1$ by the Borel Zero-One Law, for every subsequence $\{n_k\}$. Put another way, this says that $P[X_{n_k} = X_{n_k}^+$ i.o.$] = 1$, where $X^+ = \max(X, 0)$, as usual.

Suppose $\limsup_{n\to\infty} a_n = +\infty$. Then, for any given $M > 0$, a subsequence $\{n_k\}$ exists such

that $a_{n_k} \geq M$ for $k \geq 1$. Hence

$$\begin{aligned} Z &\geq \lim_{k\to\infty} \max_{i\leq k}(a_{n_i} X_{n_i}) = \lim_{k\to\infty} \max_{i\leq k}(a_{n_i} X_{n_i}^+) \\ &\geq \lim_{k\to\infty} M \max(X_{n_1}, \cdots, X_{n_k}) \\ &= M x_0 \text{ a.s.}, \end{aligned}$$

in view of (1.5). Clearly, then, $Z = +\infty$ a.s. in this case.

Now suppose $\limsup_{n\to\infty} a_n < \infty$ and define $\beta = \sup_{n\geq 1} a_n$. Then $\beta < \infty$ and

$$\begin{aligned} Z_n \leq Z_n^+ &= \max_{i\leq n}(a_i X_i^+) \\ &\leq x_0 \max(a_1, \cdots, a_n) \leq \beta x_0 \text{ a.s.} \end{aligned}$$

for each $n \geq 1$. Therefore, $Z \leq \beta x_0$ a.s. Thus, (i) and (ii) are established.

Now, suppose $Z = \lambda$ a.s. for some constant λ. Then, from (ii), $\beta < \infty$ and, as shown above, $\lambda \leq \beta x_0$. Moreover, $\lambda > 0$ since $P[Z > 0] \geq P[X_1 > 0] > 0$, in view of the assumption $x_0 > 0$. But if $\lambda < \beta x_0$, then $a_m > \lambda/x_0$ for some $m \geq 1$. Hence, if γ satisfies $\lambda < \gamma < a_m x_0$,

$$P[Z > \gamma] \geq P[a_m X_m > \gamma] = 1 - F(\gamma a_m^{-1}) > 0$$

by (1.1), since $\gamma a_m^{-1} < x_0$. This contradicts the assumption that $Z = \lambda$ a.s., so $\lambda = \beta x_0$.

Now assume that X_1 is not degenerate and $\limsup_{n\to\infty} a_n < \sup_{n\geq 1} a_n$. Then $b \equiv \sup_{n\geq N} a_n < \beta$ for some N; clearly $N \geq 2$. Define $Z_{N,n} = \max_{N\leq i\leq n}(a_i X_i)$ for $n \geq N$. Since $P[X_n > 0 \text{ i.o.}] = 1$,

$$\begin{aligned} Y_N \equiv \lim_{n\to\infty} Z_{N,n} &= \lim_{n\to\infty} \max_{N\leq i\leq n}(a_i X_i^+) \\ &\leq b \lim_{n\to\infty} M_n^+ \leq b x_0 < \beta x_0 \text{ a.s.} \end{aligned}$$

where $\{M_n\}$ is defined by (1.2). But $Z_n = \max(Z_{N-1}, Z_{N,n})$ for $n \geq N$, so taking $n \to \infty$ yields $Z = \max(Z_{N-1}, Y_N)$ a.s. But $Y_N < \beta x_0$ and $Z = \beta x_0$, so $Z_{N-1} = Z = \beta x_0$ a.s. Thus $Z_{N-1} > 0$ a.s., so

$$\beta x_0 = Z_{N-1} \leq \beta M_{N-1} \leq \beta x_0$$

in view of (1.1) and (1.2). Hence $M_{N-1} = x_0$ a.s. Consequently,

$$1 = P[M_{N-1} = x_0] = 1 - (P[X_1 < x_0])^{N-1}.$$

It follows that $P[X_1 < x_0] = 0$ and therefore, by (1.1), $P[X_1 = x_0] = 1$; that is, X_1 is degenerate, a contradiction. Hence, either $\limsup_{n\to\infty} a_n = \beta$ or X_1 is degenerate if $Z = \lambda$ a.s.

Finally, turn to the converse. Assume $\beta < \infty$. If X_1 is degenerate, then $P[X_1 = x_0] = 1$. Then $Z_n = x_0 \max(a_1, \cdots, a_n)$, so that $Z = \beta x_0$; i.e., Z is a.s. constant.

Now suppose $\sup_{n\geq N} a_n = \beta$ for all $N \geq 1$. Then, for every $\varepsilon > 0$, a sequence $\{n_k\}$ exists such that $a_{n_k} > \beta - \varepsilon, k \geq 1$. Note that $Z_n \leq \beta M_n$ if $Z_n > 0$. But $P[X_{n_k} > 0 \text{ i.o.}] = 1$, so $Z > 0$ and, by (1.4) and (1.5),

$$\begin{aligned} \beta x_0 = \lim_{n\to\infty} \beta M_n \geq Z &\geq \lim_{k\to\infty} \max_{i\leq k}(a_{n_i} X_{n_i}^+) \\ &\geq (\beta - \varepsilon) \lim_{k\to\infty} \max_{i\leq k} X_{n_i}^+ = (\beta - \varepsilon) x_0 \text{ a.s.} \end{aligned}$$

It follows that $Z = \beta x_0$ a.s. □

It remains to consider the behaviour of the Z_n-sequence when $x_0 \leq 0$.

Theorem 3.2. Let $\{a_n\}$ be a positive real sequence, and let $X_1, X_2, \cdots$ be i.i.d. r.v. Define x_0 and Z by (1.1) and (1.4) respectively, and assume $x_0 \leq 0$. Then $-\infty < Z \leq 0$ a.s. In fact,

(i) Z is not degenerate iff either (a) $x_0 = 0$ and

$$\sum_{n=1}^{\infty} \{1 - F(-\gamma a_n)\} < \infty \qquad (3.1)$$

for some $\gamma > 0$; or (b) $x_0 < 0$ and $\liminf_{n\to\infty} a_n > \inf_{n\geq 1} a_n$.

(ii) if neither (a) nor (b) holds, then

$$Z = x_0 \inf_{n\geq 1} a_n \text{ a.s.} \qquad (3.2)$$

Proof. The proof will be accomplished by considering three cases.

Case 1: $x_0 = 0$ and $P[X_1 = 0] > 0$. By the Borel Zero-One Law, $P[X_n = 0 \text{ i.o.}] = 1$. But $Z_n \leq 0$ since $x_0 = 0$, so $P[Z_n = 0 \text{ i.o.}] = 1$ inasmuch as $Z_n = 0$ if $X_n = 0$ in this case. It follows that $Z = 0$ a.s. and (3.2) holds.

For the remainder of the proof, it can now be assumed that $P[X_1 < 0] = 1$, so that $X_n^* \equiv -X_n^{-1}$ is a bona fide, positive r.v. for $n \geq 1$. Define $a_n^* = a_n^{-1}$ and $Z_n^* = \max(a_1^* X_1^*, \cdots, a_n^* X_N^*)$. It is easy to use (1.3) to check that

$$Z_n = -1/Z_n^* \tag{3.3}$$

and

$$Z = -1/Z^* \text{ if } Z^* < +\infty \text{ a.s.}, \tag{3.4}$$

where $Z^* = \lim_{n\to\infty} Z_n^*$.

Case 2. $x_0 = 0, P[X_1 = 0] = 0$. Note that X_n^* is unbounded on the right since $x_0 = 0$, so that $\{a_n^*\}$ and $\{X_n^*\}$ obey the conditions of Theorem 2.1. It is readily seen that (3.1) is equivalent to

$$\sum_{n=1}^{\infty} P[X_1^* > \alpha a_n^{-1}] < \infty, \tag{3.5}$$

where $\alpha = \gamma^{-1}$. By Theorem 2.1 (iii), if (3.1) holds, then Z^* is a.s. positive, non-degenerate and finite. In view of (3.4), Z is a.s. negative, non-degenerate and finite.

If (3.1) does not hold for any γ, then the series in (3.5) diverges for all α and, hence, $Z^* = +\infty$ a.s. by Theorem 2.1 (ii). It follows from (3.3) that $Z = 0$ a.s. in this case, and hence (3.2) holds.

Case 3: $x_0 < 0$. In this case, the a.s. least upper bound on X_n^* is $-x_0^{-1}$, so Theorem 3.1 is pertinent. Suppose

$$\liminf_{n\to\infty} a_n = \inf_{n\geq 1} a_n. \tag{3.6}$$

This is true if $\liminf_{n\to\infty} a_n = 0$, in which case $\limsup_{n\to\infty} a_n^* = +\infty$ so that $Z^* = +\infty$ a.s. by Theorem 3.2 (i). Hence $Z = 0$ a.s. by (3.3), and (3.2) clearly holds.

If $\liminf_{n\to\infty} a_n > 0$ then $\limsup_{n\to\infty} a_n^* < \infty$, and hence Z^* – and, by (3.4), Z – is a.s. finite and non-zero, in view of Theorem 3.1(ii). If (3.6) also holds, then $\limsup_{n\to\infty} a_n^* = \sup_{n\geq 1} a_n^*$, so that Z^* is degenerate by Theorem 3.1 (iii). In fact, $Z^* = -x_0^{-1} \sup_{n\geq 1} a_n^*$ a.s., so (3.2) holds, in view of (3.4). It remains only to note that, by Theorem 3.1, Z^* is a.s. finite, non-zero and non-degenerate if $\liminf_{n\to\infty} a_n > 0$ and (3.6) is false; consequently Z has the same properties, by (3.4). □

4. Connections with stability

The first result of this section links the limiting behaviour of Z_n and M_n.

Theorem 4.1. Let $X_1, X_2, \cdots$ be an i.i.d. sequence with d.f. $F(x)$ with $F(x) < 1$ for all x. Let $\{a_n\}$ be a non-increasing sequence such that $a_n \to 0$. Define M_n and Z as in (1.2) and (1.4) respectively, and γ_0 as in Corollary 2.2. Then

$$\limsup_{n\to\infty} a_n M_n = \lambda \text{ a.s.} \tag{4.1}$$

for some real λ iff

$$\begin{aligned} &P[Z \geq \lambda] = 1 \text{ and} \\ &P[Z \leq \lambda + \varepsilon] > 0 \text{ for every } \varepsilon > 0 \end{aligned} \tag{4.2}$$

and for some λ. Moreover, $\lambda = \gamma_0$ in either case.

Proof. It is well-known that

$$P[M_n > b_n \text{ i.o.}] = P[X_n > b_n \text{ i.o}]$$

if $\{b_n\}$ is non-decreasing and $b_n \to \infty$. By the Borel Zero-One Law, then, (4.1) holds iff

$$\sum_{n=1}^{\infty} \{1 - F(\gamma a_n^{-1})\} \tag{4.3}$$

converges or diverges according as $\gamma > \lambda$ or $\gamma < \lambda$. Hence $\lambda = \gamma_0$ by definition of the latter, and (4.2) holds with $\lambda = \gamma_0$ by Corollary 2.2.

Now assume (4.2) holds. Since $P[Z \leq \lambda+1] > 0, Z < +\infty$ a.s. by Theorem 2.1. Therefore, the series in (4.3) converges for some γ, so that γ_0 is finite. It follows that (4.3) converges if $\gamma > \gamma_0$ and diverges if $\gamma < \gamma_0$. As noted above, this is tantamount to $\limsup_{n\to\infty} a_n M_n = \gamma_0$ a.s. □

By considering some special values for $\{a_n\}$, a new necessary and sufficient condition for the a.s. stability of $\{M_n\}$ arises.

Corollary 4.2. Let $\{X_n\}, \{M_n\}$ and F be as in Theorem 4.1. Define, for $n \geq 1$, $\mu_n = F^{-1}(1-n^{-1})$ and $Z' = \lim_{n\to\infty} \max\{X_1/\mu_1, \cdots, X_n/\mu_n\}$, where $F^{-1}(x) = \inf\{y : F(y) > x\}$. Then $\{M_n\}$ is a.s. stable iff $P[Z' \leq 1+\varepsilon] > 0$ for every $\varepsilon > 0$.

Proof. By Theorem 1 of Resnick and Tomkins (Ref. [3]), the a.s. stability of $\{M_n\}$ is equivalent to $\limsup_{n\to\infty} M_n/\mu_n = 1$ a.s. Since it is readily apparent that $\mu_n \leq \mu_{n+1}$ for $n \geq 1$ and $\mu_n \to \infty$, taking $a_n = \mu_n^{-1}, n \geq 1$, in Theorem 4.1 reveals that $\{M_n\}$ is a.s. stable iff $P[Z' \geq 1] = 1$ and $P[Z' \leq 1+\varepsilon] > 0$ for every $\varepsilon > 0$. But, by definition of μ_n, $1 - F(\gamma\mu_n) \geq n^{-1}$ for all $n \geq 1$ and $\gamma < 1$, so $\gamma_0 \geq 1$. If $P[Z' \leq 1+\varepsilon] > 0$ for every $\varepsilon > 0$ then, by Theorem 2.1 and Corollary 2.2, γ_0 is finite and $Z' \geq \gamma_0$ a.s. If $\gamma_0 > 1$ then the contradiction $P[Z' \leq 1+\varepsilon] = 0$ arises for $\varepsilon < \gamma_0 - 1$. Hence $\gamma_0 = 1$ and $P[Z' \geq 1] = 1$ when $P[Z' \leq 1+\varepsilon] > 0$ for every $\varepsilon > 0$. The result is now apparent. □

An obvious open question relates to the stability of the sequence $\{Z_n\}$. Some work in this vein has been done by Mucci (Ref. [4], or see Theorems 4.4.1 and 4.4.2 of Ref. [5]). The author plans to explore this topic in a future paper.

References

[1] Gnedenko, B.V., Sur la distribution limite du terme maximum d'une séries aléatoire, *Ann. Math.*, **44**(1943), 423-453.

[2] Barndorff-Nielsen, O., On the limit behaviour of extreme order statistics, *Ann. Math. Statist.*, **34**(1963), 992-1002.

[3] Resnick, S.I. and Tomkins, R.J., Almost sure stability of maxima, *J. Appl. Probab.*, **10**(1973), 387-401.

[4] Mucci, R., Limit Theorems for Extremes, Ph.D. Thesis, Temple University, 1977.

[5] Galambos, J., The Asymptotic Theory of Extreme Order Statistics, 2nd ed., Robert E. Krieger, Malabar, 1987.

Large Deviations For Order Statistics

Vinogradov, V.
Concordia University, Montreal, Quebec, Canada

The presence of two polar types of the formation of large deviations (rare events) is reviewed from the point of view of the extreme value theory. Special consideration is given to the study of the asymptotic behavior of maxima for typical representatives of both polar types: normal samples and samples with regularly varying tails. The tail approximation/extreme value approximation alternative is suggested for the case of the normal sample. The results are compared with those obtained by P. Hall, R.L. Smith and J.P. Cohen. We also derive limit theorems on large deviations for trimmed sums and pose a number of open problems.

Let $\{X_n , n \geq 1\}$ be i.i.d. random variables with common distribution function $\mathbf{F}(\cdot)$; denote the corresponding order statistics by $X_n(n) \leq ... \leq X_1(n)$, and $X_1 + ... + X_n$ by S_n ; $S_0 := 0$. Set $a \wedge b := \min(a,b)$.

In this work, we study *probabilities of large deviations for order statistics and their sums,* i.e., the asymptotics of the probabilities such as

$\mathbf{P}\{X_1(n) > y\}$, $\mathbf{P}\{X_n(n) > y\}$,

$\mathbf{P}\{S_n - X_1(n) - ... - X_k(n) > y\}$, etc., where n, y and k vary such that these probabilities tend to zero. Note that our definition of large deviations (rare events) provides a more general approach than *the large deviation principle*, as well as the approach to large deviations as some *refinements of theorems on weak convergence or laws of large numbers.*

Let us first consider the case when the random sample $\{X_n , n \geq 1\}$ has the standard normal distribution. It is well known that in this case the distribution of the properly centered and normalized maximum $X_1(n)$ converges weakly to the Gumbel (double exponential) distribution

$\Lambda(x) := \exp\{ -e^{-x} \}$:

(1)

$\mathbf{P}\{X_1(n) \leq A_n + B_n \cdot x\} \Rightarrow \Lambda(x)$

as $n \to \infty$, where

$A_n := (2 \cdot \log n)^{1/2} - 1/2 \cdot (\log \log n + \log(4\pi)) \cdot (2 \cdot \log n)^{-1/2}$,

and $B_n := (2 \cdot \log n)^{-1/2}$. Note that (1) remains true if A_n and B_n are replaced by a_n and b_n respectively, such that $b_n / B_n \to 1$ and $(a_n - A_n)/B_n \to 0$ as $n \to \infty$ (see, e.g., Ref. [1] Lemma 2.2.2). It was known since the works Ref. [2] and Ref. [3] (cf. also Ref [4]) that the rate of convergence in (1) is very slow and worst on the tails. Let us now quote the rigorous result. It was obtained in Ref. [5] that the exact bounds in (1) are as follows:

There exists a positive constant C such that for any integer $n \geq 2$,

$$\frac{C}{\log n} \leq \sup_{x \in \mathbb{R}^1} \left| \mathbf{P}\{X_1(n) \leq A'_n + B'_n \cdot \Lambda(x)\right| \qquad (1')$$

$$\leq \frac{3}{\log n} ,$$

where

$B'_n := (A'_n)^{-1}$; $A'_n := (2 \cdot \log n)^{1/2} - (\log \log n + \log(4\pi)) \cdot (8 \cdot \log n)^{-1/2} - [(\log \log n + \log(4\pi))^2 - 4 \cdot (\log \log n + \log(4\pi)] \cdot (8 \cdot (2 \cdot \log n)^{3/2})^{-1}$.

The following representation for the distribution of the maximum from the normal sample containing the leading error term for (1') can be derived from formula (10) of Ref. [5] and Theorem 2 of Ref. [6]:

$$P\{X_1(n) \le A'_n + B'_n \cdot x\} - \Lambda(x)$$

(2)

$$= \Lambda(x) \cdot e^{-x} \cdot \frac{x^2/2+x+1}{2 \cdot \log n} + o\left(\frac{1}{\log n}\right)$$

as $n \to \infty$ uniformly in $x \in \mathbb{R}^1$.

Note that formula (10) of Ref. [5] is in fact an auxiliary result of that work used for the derivation of (1'), whereas slightly different centering and normalizing sequences were chosen in Ref. [6]. Let us also point out that Theorem 2 of Ref. [6] covers a wide class of distributions (which contains the normal distribution) known as *class* **N**.

In addition, in order to emphasize the fact that the rate of convergence in (1) is very slow and worst on the tails, we now quote the following remark given on p. 492 of Ref. [7]: "*... an approximation by an extreme value distribution is of little use in determining a critical point for the rejection of outliers. What is needed is a non-uniform estimate* $Q_n(x)$ *of* $P\{X_1(n) \le x\}$. *Ideally such an estimate should be simple to calculate, and the relative error*

$$|Q_n(x) - P\{X_1(n) \le x\}| \,/\, (1 - P\{X_1(n) \le x\})$$

should tend to zero as x and n tend to infinity." In this respect, the following result was derived in Ref. [7] Theorem 3, which can be viewed as a non-uniform estimate in (1) taking into account right-hand large deviations.

Let $a_n > 0$ be defined from the equation

$$2\pi \cdot a_n^2 \cdot \exp\{-a_n^2\} = n^2,$$

and let

$$z_n(x) := (2\pi)^{-1/2} \cdot n \cdot x^{-1} \cdot \exp\{-x^2/2\}.$$

Then for $x \ge a_n$,

$$\exp\{-z_n(x)\ [1 - x^{-2} + 3x^{-4} + z_n(x)/2(n-1)]\}$$

(3)

$$\le P\{X_1(n) \le x\} \le \exp\{-z_n(x) \cdot [1 - x^{-2}]\}.$$

Now, let us point out that Theorem 3 of Ref. [8] (see also Remark on p. 1195 therein) implies that the asymptotics of the probabilities of right-hand large deviations for the centered and normalized maximum from the normal sample is given by the tail of function Λ only in every narrow range of large deviations, namely as

$n \to \infty$ and $x \to \infty$, $x = o((\log n)^{1/2})$

(4) $P\{X_1(n) > A_n + B_n \cdot x\} \sim 1 - \Lambda(x)$.

In addition, it is easily shown that (3) implies the following more general result describing the exact asymptotics of the probabilities of right-hand large deviations in the full range:

$$P\{X_1(n) > a_n + b_n \cdot x\}$$

(4')

$$\sim (1 - \Lambda(x)) \cdot \frac{\exp\{-x^2/2a_n^2\}}{1 + x/a_n^2}$$

as $n \to \infty$, $x = x(n) \to \infty$, where $b_n := a_n^{-1}$, and a_n is

the same as in (3); $a_n \sim (2\log n)^{1/2}$ as $n \to \infty$.

Remark. It is obvious that in the range of deviations $x = o((\log n)^{1/2})$, the asymptotics of $P\{X_1(n) > a_n + b_n \cdot x\}$ is completely determined by the first factor (compare to (4)). On the other hand, for

$x \ge Const \cdot (\log n)^{1/2}$, $x = o(\log n)$,

the asymptotics of $P\{X_1(n) > a_n + b_n \cdot x\}$ is completely determined by the product

$(1 - \Lambda(x)) \cdot \exp\{-x^2/2a_n^2\}$,

whereas for $x \ge Const \cdot \log n$, all the three factors on the right-hand side of (4') should be taken into account.

In this work, we develop an alternative method (hereinafter referred to as the **tail approximation**), which provides asymptotic expansions for the probabilities of the right-hand large deviations along with accurate estimates of remainders. The just mentioned method is based on the following apparent representation, which is equally applicable when the distribution function **F** of i.i.d. random variables $\{X_n, n \ge 1\}$ is arbitrary (We do not even require the distribution function of properly centered and normalized maximum $X_1(n)$ to belong to the domain of attraction of anyone limiting distribution):

$P\{X_1(n) > y\}$

(5)

$$= \sum_{k=1}^{n} (-1)^{k+1}$$

$$\cdot P\left\{\exists i_1, \ldots, i_k : X_{i_1} > y, \ldots, X_{i_k} > y\right\}$$

$$= \sum_{k=1}^{n} (-1)^{k+1} \cdot \binom{n}{k} \cdot P\{X_i > y\}^k .$$

It is obvious that (5) easily follows from the apparent representation

$$P\{X_1(n) > y\} = P\left(\bigcup_{i=1}^{n} \{X_i > y\}\right)$$

and the fact that the latter probability can be easily rewritten by means of the well known formula for the probability of a union of non-disjoint events.

In particular, (5) implies that

$P\{X_1(n) > y\}$

(5')

$= n \cdot (1 - F(y)) + O((n \cdot (1- F(y))^2)$

as $n \to \infty$, $y \to \infty$, such that $n \cdot (1- F(y)) \to 0$.

An application of Representation (5) to the normal sample yields the following result:

Theorem 1. Let us assume that the i.i.d. sequence $\{X_n, n \geq 1\}$ is the standard normal, and denote their common Laplace distribution function by Φ. Let a_n be defined as in (3), and $b_n := a_n^{-1}$. Then for any positive x,

$P\{X_1(n) > a_n + b_n x\}$

(6)

$$= \sum_{k=1}^{n} (-1)^{k+1} \cdot \binom{n}{k} \cdot n^{-k}$$

$$\cdot \{e^{-x} \cdot \frac{\exp\{-x^2/2a_n^2\}}{1+x/a_n^2} \cdot \{1 - \frac{1}{a_n^2 \cdot (1+x/a_n^2)^2} +$$

$$\ldots + (-1)^l \cdot \frac{1 \cdot 3 \cdot \ldots \cdot (2l-1)}{a_n^{2l} \cdot (1+x/a_n^2)^{2l}} + \ldots\}\}^k .$$

Proof of Theorem 1 is straightforward. It involves an application of (5) with $y = a_n + b_n \cdot x$, an expansion of $1-\Phi(x)$ over powers of x as $x \to \infty$ (see, e.g., Ref. [9] (vol. I, Chapter 7, Section 7, Problem 1) and the fact that

$$\frac{n \cdot \exp\{-(a_n + b_n \cdot x)^2/2\}}{(2\pi)^{1/2}(a_n + b_n \cdot x)} = e^{-x} \cdot \frac{\exp\{-x^2/2a_n^2\}}{1+x/a_n^2} . \square$$

Remarks. (i) Note that both alternating sums on the right-hand side of (6) can be dropped at any term; the absolute value of the emerged error will be bounded by the absolute value of the first omitted term.

In particular, Theorem 1 implies that

$P\{X_1(n) > a_n + b_n \cdot x\} =$

(6')

$$(1-\Lambda(x)) \cdot \frac{\exp\{-x^2/2a_n^2\}}{1+x/a_n^2} \cdot \{1 - \frac{1}{a_n^2 \cdot (1+x/a_n^2)^2}\}$$

$$+ O\left((1-\Lambda(x))^2 \cdot \frac{\exp\{-x^2/2a_n^2\}}{1+x/a_n^2} \right)$$

$$+ O\left((1-\Lambda(x)) \cdot \frac{\exp\{-x^2/2a_n^2\}}{a_n^4 \cdot (1+x/a_n^2)^5} \right)$$

as $n \to \infty$, $x \to \infty$.

(ii) Note that our representations (6) and (6') are similar to Theorem 3 of Ref. [7]. However, our results seem to be more convenient for computations. In addition, the proposed *tail approximation* method is equally applicable to an arbitrary distribution function, whereas the range of applications of Theorem 3 of Ref. [7] is confined to the normal samples. In this respect, let us quote the following remark from Ref. [10] (see p. 329 therein): *"Hall (1980) suggested approximations to $\Phi^n(x)$ using some refined inequalities for the normal tail function. These approximations are much closer to $\Phi^n(x)$ than the penultimate approximations. Thus, if the X_i's are indeed independent and identically normally distributed and if n is known, then Hall (1980) gives better estimates of the distribution of*

$Y_n = max\{X_i\}$ than approximations based on extreme value theory. However, in practice we are often uncertain of the normality, the independence and perhaps the value of n. Since the three limit laws apply to a large class of initial distributions, and often in the dependent case (cf. Galambos (1978)), extreme value theory approximations are more robust than the alternatives suggested by Hall (1980)."

Now, in view of the above remark, let us suggest the following approach, which in our opinion provides a more appropriate approximation for distributions of

maxima. Hereinafter, we refer to this approach as the tail approximation / extreme value approximation alternative. Note that in this work, we apply this approach only to the standard normal sample, the classical test sample of the extreme value theory. Of course, the range of applications of this alternative is not confined by the normal sample.

It is natural to require the relative error $\delta(n,x)$ of the tail approximation be less than ε (given a priori). Then by Bonferroni's inequalities (cf., e.g., Ref. [9] (Vol. I, Chapter IV, (5.7)) if
$P_n := \mathbb{P}\{X_1 > a_n + b_n \cdot x\} = 1/n$ then

$$n \cdot P_n - n^2 \cdot P_n^2/2 \le \mathbb{P}\{\lambda_1(n) > a_n + b_n x\} \le n \cdot P_n.$$

Hence,

$$\delta(n, x) = \frac{n \cdot P_n}{2(1 - n \cdot P_n/2)} \quad (< \varepsilon).$$

Obviously, the product $n \cdot P_n$ is assumed to be small enough. Making simple computations we get that for fixed α and n, the above inequality is fulfilled (i.e. the relative error of the approximation of $\mathbb{P}\{X_1(n) > a_n + b_n x\}$ by the leftmost term on the right-hand side of (6') is less than α) if

$$x \ge X_1(n, \alpha) := \frac{a_n^2}{3} \cdot \{1 - \frac{5}{a_n^2} \cdot \log\frac{n}{2\sqrt{\pi}}\}^{1/2} - 1\}.$$

On the other hand, it is obvious in view of (4) and (4') that the extreme value approximation is accurate at least for the values of x being sufficiently small compared to $(\log n)^{1/2}$. Moreover, representation (6') implies that

$$\frac{|\mathbb{P}\{X_1(n) > a_n + b_n x\} - (1 - \Lambda(x))|}{1 - \Lambda(x)} \le 1 - \frac{\exp\{-x^2/2a_n^2\}}{1 + x/a_n^2} \cdot \{\frac{1 - \Lambda(x)}{2} + 1\} \quad (< \varepsilon).$$

Making simple transformations and neglecting the second order terms we get that the above inequalities are fulfilled for
$\varepsilon > x^2/(2a_n^2) + x/(a_n^2)$.
Hence, we obtain the result that for fixed α and, n the relative error of the extreme value approximation of $\mathbb{P}\{X_1(n) > a_n + b_n \cdot x\}$ is less than ε if

$$x \le X_2(n, \varepsilon) := (1 + 2 \cdot a_n^2 \cdot \varepsilon)^{1/2} - 1.$$

It is not surprising that for any positive (sufficiently small) ε, and for any integer
$n \ge N(\varepsilon) := [(\log(\varepsilon+1)/\varepsilon))^2 / 8\varepsilon] + 1$,
there is an overlap between a lower bound $X_1(n, \varepsilon)$ and an upper bound $X_2(n, \varepsilon)$. In other words, for $n \ge N(\varepsilon)$ we can always provide an approximation for the distribution of the maximum from the normal sample whose relative error is less than (a given a priori) ε. Moreover, there exists a certain range of deviations in which both the tail approximation and the extreme value approximation give comparable results.

Let us summarize the above discussion in the following algorithm: *assume that we should find an approximation of $\mathbb{P}\{X_1(n) > a_n + b_n x\}$ for given n and x, whose relative error would be less than ε. Than we can choose the tail approximation if $x \ge X_1(n, \varepsilon)$ or the extreme value approximation if $x \le X_2(n, \varepsilon)$. This is possible at least for $n \ge N(\varepsilon)$.*

Open Problems. (i) To extend the suggested alternative to a wider class of distributions (for example, class **N** introduced in Ref. [6]).
(ii) To derive sharper estimates for ranges of applications of both approximations. In particular, it would be interesting to find the border function $\bar{x}(n,\varepsilon)$ such that for $x \le \bar{x}(n,\varepsilon)$ the extreme value approximation should be chosen, whereas for $x \ge \bar{x}(n,\varepsilon)$ the tail approximation should be chosen.
(iii) To derive the asymptotics of the left-hand large deviations.

Now, let us proceed with the consideration of the asymptotics of the distribution of the maximum $X_1(n)$ from the sample of i.i.d. random variables whose common distribution function $\mathbf{F}(\cdot)$ has the right-hand tail of power type:
(7)
$1 - \mathbf{F}(x) = c_{\alpha_1} \cdot x^{-\alpha_1} + o(x^{-\alpha_1})$
as $x \to \infty$, where $c_{\alpha_1} > 0$ and $\alpha_1 > 0$. Note that it was proved in Ref. [11] Theorem 4 that under fulfilment of

(7), the properly normalized distribution function of the maximum $X_1(n)$,
$\mathbf{F}_n(x) := \mathbf{P}\{X_1(n) \le (c_{\alpha_1} \cdot n)^{1/\alpha_1} \cdot x\}$,
converges weakly as $n \to \infty$ to the limiting distribution $\Psi_{\alpha_1}(\cdot)$ defined as $\Psi_{\alpha_1}(x) := \exp\{-x^{\alpha_1}\}$ if $x > 0$, and
$\Psi_{\alpha_1} := 0$ otherwise. It is also well known (cf., e.g., Ref. [12]) that under fulfilment of (7) and certain supplementary conditions the following results on the asymptotic behavior of the probabilities of large deviations are valid:
$\mathbf{P}\{S_n > y\} \sim n \cdot c_{\alpha_1} \cdot y^{-\alpha_1} \sim \mathbf{P}\{X_1(n) > y\}$
(8)
$\sim 1 - \Psi_{\alpha_1}((y / (c_{\alpha_1} \cdot n)^{1/\alpha_1})^{-\alpha_1})$
as $n \to \infty$ where $\ge Y(n)$ is a positive monotone sequence that depends on certain parameters and increases to infinity as $n \to \infty$. Note that the rightmost relationship in (8) remains true even as $n \to \infty$,
$y/n^{1/\alpha_1} \to \infty$, and just under fulfilment of (7), without any supplementary restriction (cf., e.g., Section 2.3 of Ref. [13]).

Thus, it is clear in view of (8) and the above mentioned result on weak convergence to $\Psi_{\alpha_1}(\cdot)$ that the tail approximation and the extreme value approximation actually coincide in this case, and we do not have any alternative but the extreme value approximation.

Now, let us proceed with the derivation of refinements to (8). It seems reasonable to assume, that if more precise information on the tail behavior of function $\mathbf{F}$ is available (compare to (7)), then more precise representations for $\mathbf{P}\{S_n > y\}$ and $\mathbf{P}\{X_1(n) > y\}$ as $n \to \infty$ than those, valid up to equivalence can be derived. Various expansions for $\mathbf{P}\{S_n > y\}$ refining the first of relationships (8) have been constructed in a number of works by the author (cf., e.g., Ref. [14,15]) in the case, where the right hand tail of $\mathbf{F}$ admits the following expansion over negative powers of x:

$1 - \mathbf{F}(x)$
(9)

$$= \sum_{i=1}^{l} c_{\alpha_i} \cdot x^{-\alpha_i} + o(x^{-r})$$

as $x \to \infty$, where $c_{\alpha_i} > 0$, and
$0 < \alpha_1 < \alpha_2 < \ldots < \alpha_l \le r$.

On the other hand, the following result, Theorem 2, provides the asymptotic expansions for the distribution of $X_1(n)$ refining both the above mentioned result on weak convergence of $\mathbf{F}_n$ towards Φ_{α_1} obtained in Ref. [11] and the rightmost relationship in (8) in the case in which condition (9) is valid. The remainder term of the expansion of this theorem may be viewed as the unremovable error generated by the lack of the perfect information on the tail behavior of $\mathbf{F}$.

Theorem 2 (cf. Ref. [16] Theorem 1). Let Condition (9) be fulfilled. Then
$\mathbf{F}_n(x) = \mathbf{P}\{X_1(n) \le (c_{\alpha_1} \cdot n)^{1/\alpha_1} \cdot x\}$

$$= (1 - n^{-1} \cdot x^{-\alpha_1})^n$$

$$\cdot \left(1 + \sum_{m=1}^{[r/\alpha_2] \vee [(r-\alpha_1)/(\alpha_2-\alpha_1)]} \cdot \frac{(-1)^m}{m!}\right.$$

$$\cdot \left(n \cdot \sum_{s=1}^{[r/\alpha_2]} \cdot \frac{1}{s} \cdot \left(\sum_{i=2}^{l} \frac{C_{\alpha_i}}{C_{\alpha_1}^{\alpha_i/\alpha_1}} \cdot n^{-\alpha_i/\alpha_1} \cdot x^{-\alpha_i}\right)^s\right.$$

$$\left.\left. \cdot \sum_{k=0}^{[(r/m - s\cdot\alpha_2)/\alpha_1] \vee 0} \cdot (-1)^k \cdot \binom{-s}{k} \cdot n^{-k} \cdot x^{-k\cdot\alpha_1}\right)^m\right)$$

$$\cdot \left(1 + \delta(x,n) \cdot n^{-(r-\alpha_1)/\alpha_1} \cdot x^{-r/\alpha_1}\right),$$

where $\delta(x,n) \to 0$ as $n \to \infty$ uniformly on the rays $[C, +\infty)$;

$$\binom{-s}{0} := 1, \text{and} \binom{-s}{k} := (-1)^k \cdot s \cdot (s+1) \cdot \ldots \cdot (s+k-1)/k!$$

for integer $k \geq 1$ (here $C > 0$ being fixed).

Remarks. (i) Note that a uniform version of our Theorem 2 for a special case $\ell = 2$, $\alpha_2 < 2 \cdot \alpha_1$, $r = \alpha_2$ was obtained in Ref. [17] (see Example 1 in Section 6 therein).

(ii) Note that one can easily obtain the asymptotic expansions for distributions of the maximum $X_1(n)$ in the case, where

$$1 - F(x) = \sum_{j=1}^{\ell} c_{\alpha_j} \cdot (a-x)^{\alpha_j} + o((a-x)^r)$$

as $x \to a$, where $a < \infty$, $c_{\alpha 1} > 0$, and $0 < \alpha_1 < \alpha_2 < \ldots < \alpha_\ell \leq r$ by reformulating the result of Theorem 2. Note that distributions of such type often arise in various statistical estimation problems (cf., e.g., Ref. [18] for details).

Now, let us emphasize that the comparison of the above results related to maxima from normal samples with those related to maxima from samples of distributions having right-hand tails of the power type reveals the presence of two polar types of the limiting behavior of the probabilities of large deviations. Note that the presence of these two polar types has been first established within the framework of the classical scheme of summation of independent identically distributed random variables. Thereupon, it also became apparent during the study of the limiting behavior of the probabilities of large deviations of certain families of stochastic processes (cf., e.g., Ref. [19] Introduction). Following Ref. [19], we regard **the first type** of the limiting behavior of the probabilities of large deviations as one, being associated with the case of fulfilment of Cramér's condition of the finiteness of the exponential moment:

$\mathbf{E}\{z \cdot X_i\} < \infty$ for any $z \in \mathbb{R}^1$. In this case, the probability of a large deviation is generated mainly by approximately equal individual summands X_i. In contrast to that, **the second polar type** is characterized with the case in which the main part of the probability of a large deviation is generated by one large summand comparable with the whole sum S_n. The typical example is the case of power tails with index $\alpha_1 < 2$ (cf. condition (7) and relationships (8) above). Let us point out that only the polar types of the limiting behavior of the probabilities of large deviations do not cover all possible cases. Thus, few subtle results intermediate between the polar types are also known (see, e.g., Ref. [20-22]). The latter two papers contain a number of subtle results on the exact asymptotics of (conditional and unconditional) probablilites of large deviations for trimmed sums under fulfilment of condition (7) and its left-hand analog, obtained by transferring the problem to the càdlàg space $\mathbf{D}[0,1]$. Those results on the asymptotic behavior of the probabilities of large deviations of trimmed sums can be interesting from the point of view of possible applications. On the other hand, they also provide a better understanding of the nature of large deviations. To explain this, we introduce the random step-function

$S_{n,y}(t) := S_{[nt]} / y$ if $t \in [0,1]$.

We consider the realizations of $S_{n,y}(\cdot)$ in $\mathbf{D}[0,1]$ equipped with the uniform metric ρ. The question arises is what are the typical paths of $S_{n,y}$ like if an event of small probability (a large deviation) has occurred. It is well known that if the random step-function (r.s.f.) $S_{n,y}$ is constructed starting from random variables with finite exponential moments, then a large deviation is mainly contributed by close-to-continuous (or even close-to-smooth) paths (cf. Ref. [19] Introduction) - the result of the first polar type. On the other hand, under fulfilment of condition (7) and its left-hand analog with index $\alpha_1 < 2$, large deviations of $S_{n,y}$ occur mainly via almost piecewise constant paths, which perform one or several big jumps - the result of the second polar type. For the case of fulfilment of condition (7) and its left-hand analog with index $\alpha_1 > 2$ (note that in this case the weak convergence of $S_{n,\sqrt{n}}(\cdot)$ to the Wiener process $\mathbf{w}(\cdot)$ holds) the paths of both types can give comparable contributions to the probability of a large deviation - the result which is intermediate between the both polar types). Let us emphasize that this is possible for the case $\alpha_1 > 2$ and in a very narrow range of large deviations only. Outside this range of deviations, if $y \geq n^{1/2+k}$, events of small probability occur mainly due to the almost piecewise constant paths (here $\kappa > 0$ being any real). Moreover, for several sets of $\mathbf{D}[0,1]$ large deviations occur via the paths which perform one or several big jumps and close-to-continuous functions between them. Surprisingly, all the discovered sets with such properties are related to trimmed sums.

Now, let us state the following theorem (refining Theorem 2 of Ref. [21] for a special case), which contains the asymptotic expansion for the probabilities of the right-hand large deviations of

$S_n - X_1(n)$:

Theorem 3. Let condition (9) and the left-hand analog of condition (7) be fulfilled with $\alpha_1 \in (0,1) \cup (1,2)$.

Let $EX_i = 0$ if $\alpha_1 \in (1,2)$. Then

$$\mathbf{P}\{S_n - X_1(n) > y\}$$

$$= \binom{n}{2} \cdot \sum_{\substack{1 \le i < j \le l: \\ \alpha_i + \alpha_j < 3\alpha_1 \wedge r}} c_{\alpha_i} \cdot c_{\alpha_j} \cdot y^{-\alpha_i - \alpha_j}$$

$$+ O(n^3 \cdot y^{-3\alpha_1}) + O(n^2 \cdot y^{-r})$$

as $n \to \infty$, $y / n^{1/\alpha_1} \to \infty$.

Proof of Theorem 3 is straightforward and purely probabilistic. It follows along the same lines as that of Ref. [21] Theorem 2 and Ref. [22], Theorem 1, in which the results on the asymptotic behavior (up to equivalence) of the probabilities of large deviations of trimmed sums have been obtained. Applying slight modifications of auxiliary results of those works, we get that there exists $\kappa \in (0, 1/2)$ such that

$$\mathbf{P}\{S_n - X_1(n) > y\}$$

$$= \sum_{1 \le i < j \le n} \mathbf{P}\{S_n - X_1(n) > y, \ |X_i| > \kappa y, \ |X_j| > \kappa y\}$$

$$+ O(n^3 \cdot y^{-3\alpha_1})$$

as $n \to \infty$, $y / n^{1/\alpha_1} \to \infty$.

On the other hand, it is not difficult to show that the sum over i and j on the right-hand side of this representation is equal to

$$\binom{n}{2} \cdot \int_{-\infty}^{(1-2\kappa)y} \mathbf{P}\{X_{n-1} \wedge X_n > y - z\} \, dF_{S_{n-2}}(z)$$

(10)

$$+ O(n^3 \cdot y^{-3\alpha_1}).$$

It is obvious, in view of (9), that the right-hand tail of the distribution of $X_{n-1} \wedge X_n$ has the following asymptotics as $v \to \infty$:

$\mathbf{P}\{X_{n-1} \wedge X_n > v\}$

(11)

$$= \sum_{\substack{1 \le i < j \le l: \\ \alpha_i + \alpha_j < 3\alpha_1 \wedge r}} \cdot c_{\alpha_i} \cdot c_{\alpha_j} \cdot v^{-\alpha_i - \alpha_j}$$

$$+ O(v^{-3\alpha_1}) + O(v^{-r}).$$

Splitting the integral in (10) into three parts by analogy to the proof of Theorem 1 of Ref. [23] (see (7) - (11) therein) and replacing $\mathbf{P}\{X_{n-1} \wedge X_n > y-z\}$ by the sum over i and j on the right-hand side of (11) imply the result of Theorem 3. □

Remark. Note that from our perspective, the study of the asymptotic behavior of trimmed sums in the case in which the distributions of maxima belong to the domain of attraction of the Gumbel distribution is not of the same interest. Indeed, it is essentially similar to the study of the asymptotic behavior of S_n (cf., e.g., Ref. [24] for the results on weak convergence, and also the discussion below Theorem 2 of the present work given from the point of view of the presence of the two polar types of the limiting behavior of probabilities of large deviations).

Open Problems. (i) It seems possible to construct more accurate asymptotic expansions for $\mathbf{P}\{S_n - X_1(n) > y\}$ compare to those of Theorem 3.

(ii) To extend Theorem 3 to the case when an arbitrary fixed number of upper order statistics are deleted from S_n.

(iii) To extend Theorem 3 to the case of power tails (cf. condition (7)) with index $\alpha_1 > 2$.

Acknowledgement. The preparation of the preliminary version of this work was financially supported by an NSERC Canada International Research Award hosted at Carleton University, Ottawa. The author thanks Donald A. Dawson and Miklos Csörgő for their advice during the preparation of this work, as well as their hospitality. The technical assistance of Yaacov D.L. Shapiro, during the preparation of the final version is very much appreciated.

[1] Galambos, J., The Asymptotic Theory of Extreme Order Statistics, John Wiley & Sons, New York, 1978.

[2] Tippett, L.H.C., On the extreme individuals and the range of samples taken from a normal population, Biometrika, 17 (1925), 364-387.

[3] Fisher, R.A. and Tippett, L.H.C., Limiting forms of the frequency distribution of the largest or smallest number of a sample, Proc. Camb. Phil. Soc., 24 (1928), 180-190.

[4] Gumbel, E.J., Statistics of Extremes, Columbia University Press, New York, 1958.

[5] Hall, P., On the extreme terms of a sample from the parent extremes, J. Appl. Prob., 15 (1978), 420-426.

[6] Cohen, J.P., On the inflexion and penultimate approximation of extreme-value theory, Adv. Appl. Prob., 14 (1982), 833-854.

[7] Hall, P., Estimating probabilities for normal extremes, Adv. Appl. Prob., 12 (1980), 491-500.

[8] de Haan, L. and Resnick, S., On the rate of growth of sample maxima, Ann. Math. Statist., 43 (1972), 1185-1196.

[9] Feller, W., An Introduction to Probability Theory and its Applications, Vol. I, II, 2nd edition, John Wiley & Sons, New York, 1971.

[10] Cohen, J.P., The penultimate form of approximation to normal extremes, Adv. Appl. Prob. 14 (1982), 324-329.

[11] Gnedenko, B.V., Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math., 44 (1943), 423-453.

[12] Heyde, C.C., On large deviation probabilities in the case of attraction to a non-normal stable law, Sankhya, A30 (1968), 253-258.

[13] Resnick, S.I., Extreme Values, Regular Variation, and Point Processes, Springer, Berlin, 1987.

[14] Vinogradov, V., On asymptotic expansions in limit theorems on large deviations for sums of independent random variables in the case of power tails, C.R. Math. Rep. Acad. Sci. Canada, 14 (1992), 83-88.

[15] Vinogradov, V., A non-uniform estimate taking into account large deviations in the limit theorem on non-normal convergence to the normal law, C. R. Math. Rep. Acad. Sci. Canada, 14 (1992), 285-290.

[16] Vinogradov, V., Limit theorems for extreme order statistics: large deviations and asymptotic expansions with non-uniform estimates of remainders, C. R. Math. Rep. Acad. Sci. Canada, 14 (1992), 131-136.

[17] Smith, R.L., Uniform rates of convergence in extreme-value theory, Adv. Appl. Probab., 14 (1982), 600-622.

[18] Ibragimov, I.A. and Has'minskii, R.Z., Statistical Estimation - Asymptotic Theory, Springer, New York, 1981.

[19] Wentzell, A.D., Limit Theorems on Large Deviations for Markov Stochastic Processes, Kluwer, Dordredcht, 1990.

[20] Pinelis, I.F., A problem of large deviations in a space of trajectories, Theory Probab. Appl., 26 (1981), 69-84.

[21] Godovan'chuk, V.V. and Vinogradov, V., Large deviations of sums of independent random variables without several maximal summands, Theory Probab. Appl., 34 (1989), 512-515.

[22] Godovan'chuk, V.V. and Vinogradov, V., On large deviations for trimmed sums of independent random variables. In: Probability Theory and Mathematical Statistics. Proc. 5th Vilnius Conf., Vol. 1 (eds.: B. Grigelionis et al.), VSP/Mokslas, Utrecht/Vilnius, 1990, pp. 424-432.

[23] Vinogradov, V., Asymptotic expansions of the probability of large deviations for sums of independent random variables under violation of Cramér's condition, Theor. Probability and Math. Statist., 31 (1985), 21-27.

[24] Lo, G.S., A note on the asymptotic normality of sums of extreme values, J. Statist. Planning Inference, 22 (1989), 127-136.

Extremes For Independent Nonstationary Sequences

Weissman, I.
Technion-Israel Institute of Technology, Haifa, Israel

Extreme value theory for nonstationary sequences of independent random variables is discussed. We present limit distributions for extremes, point processes associated with extremes, extremal processes, record values and record times. The results shown are those which, we think, are useful for practitioners and are not found in textbooks. The last two sections include some new results.

1 Introduction

The literature of extreme value theory and its applications is huge and continues to grow. Articles on the subject appear frequently in journals of almost all sciences. A great deal of the literature is devoted to sequences of independent, identically distributed (iid) random variables (rv's). The interested practitioner can find the necessary material in texts such as Refs. [1–5].

A common generalization of the iid sequence is the stationary sequence. Reference [6] provides an extensive coverage of extreme value theory for stationary sequences. Another generalization is the independent nonstationary sequences. Except for two sections in Ref. [2], this case is not discussed in texts. Dependent nonstationary sequences are discussed in this volume in Ref. [7]. The purpose of the present paper is to bring to the attention of practitioners of all sciences the kinds of results available in the literature. The last two sections contain some new results.

2 Limiting Distributions for Extremes

Let $\{X_i\}$ be a sequence of independent rv's, where F_i is the distribution function (df) of X_i, and let

$$M_{nk} = k\text{-th largest of } \{X_1, X_2, \ldots, X_n\} .$$

The df of $M_n = M_{n1}$ is given by

$$P\{M_n \le x\} = \prod_{i=1}^{n} F_i(x) \equiv H_n(x) . \qquad (2.1)$$

What are the possible limits, as $n \to \infty$, of H_n? The answer is trivial, since every df G can appear as a limit of (2.1) (just take $F_i = G^{2^{-i}}$). In order to avoid trivialities, we normalize M_n with constants $a_n > 0$ and b_n such that as $n \to \infty$

$$\begin{aligned} P\{M_n \le a_n x + b_n\} = & \qquad (2.2) \\ = & \prod_{i=1}^{n} F_i(a_n x + b_n) \\ = & H_n(a_n x + b_n) \to G(x) \end{aligned}$$

at all continuity points of G (assumed to be nondegenerate) and such that

$$\lim_{n\to\infty} \min_{1\le i\le n} F_i(a_n x + b_n) = 1 \quad (x > x_L) . \quad (2.3)$$

Here $x_L = x_L(G) = \sup\{x : G(x) = 0\}$. Condition (2.3) is called the uniform right-negligibility (URN) condition. Under URN, no finite set of X_i can play a predominant role in determining the maximum M_n as $n \to \infty$.

For a df G let $\lambda(x) = \lambda_G(x) = -\log G(x)$. Define the following classes of df's:

$\mathcal{M}^o = \{G : \lambda(x)$ is convex$\}$
$\mathcal{M}^+ = \{G : x_L > -\infty$ and $\lambda(x_L + e^x)$ is convex$\}$
$\mathcal{M}^- = \{G : x_R < \infty$ and $\lambda(x_R - e^{-x})$ is convex$\}$.

Here $x_R = x_R(G) = \inf\{x : G(x) = 1\}$. The characterization of the possible limit laws of M_n is given by the following theorem.

Theorem 2.1 (Refs. [8, 9]) *A df G can be a limit in (2.2) under (2.3) if and only if it belongs to*

$$\mathcal{M} = \mathcal{M}^o \cup \mathcal{M}^+ \cup \mathcal{M}^- \ (= \mathcal{M}^o \cup \mathcal{M}^-) \ .$$

Note that if $\lambda(x_L + e^x)$ is convex so is $\lambda(x)$ $(-\infty < x < \infty)$, thus $\mathcal{M}^+ \subseteq \mathcal{M}^o$. The classical extreme value distributions (EVD) $\Lambda(x) = \exp\{-e^{-x}\}$, $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$ $(x > 0)$ and $\Psi_\alpha(x) = \exp\{-|x|^\alpha\}$ $(x < 0)$ belong respectively to $\mathcal{M}^o, \mathcal{M}^+$ and $\mathcal{M}^-$. The df $F(x) = x^\alpha$ $(\alpha > 0,\ 0 \le x \le 1)$ belongs to each one of the three subclasses. The normal distribution is not in $\mathcal{M}$. So, in situations where the iid assumption is not justified, the practitioner might fit a df for the extremes from a much larger class than the EVD.

Note that if $G \in \mathcal{M}$, then G is strictly increasing, continuous and differentiable inside (x_L, x_R). The right-end x_R can be a point of discontinuity only if $G \in \mathcal{M}^-$.

Note that under (2.3), $H_n(a_{n+1}x + b_{n+1})$ and $H_{n+1}(a_{n+1}x + b_{n+1})$ have the same limit (as $n \to \infty$), thus the Convergence of Types Theorem (Ref. [10], p. 253) implies

$$\frac{a_{n+1}}{a_n} \to 1 \ , \quad \frac{b_{n+1} - b_n}{a_n} \to 0 \quad (n \to \infty) \ . \tag{2.4}$$

If $\lambda(x)$ is convex, then $G(x)/G(x+\Delta)$ is a df for every $\Delta > 0$. Thus, for every increasing sequence $\{b_n\}$ with $b_o = 0, b_n \to \infty$ which satisfied (2.4), the df's $F_i(x) = G(x - b_i)/G(x - b_{i-1})$ satisfy (2.2) and (2.3) (with $a_n \equiv 1$). If $G \in \mathcal{M}^+$, then $G(x_L + x)/G(x_L + \alpha x)$ is a df for $\alpha > 1$ and thus for every increasing sequence $\{a_n\}$ $(a_n > 0,\ a_n \to \infty)$ which satisfies (2.4), the df's $F_i(x) = G(x_L + x/a_i)/G(x_L + x/a_{i-1})$ satisfy (2.2) and (2.3) (with $b_n = -a_n x_L$). Finally, if $G \in \mathcal{M}^-$, then $G(x_R + x)/G(x_R + \alpha x)$ is a df for $0 < \alpha < 1$ and thus for every decreasing sequence $\{a_n\}$ $(a_n > 0,\ a_n \to 0)$, which satisfies (2.4), the df's $F_i(x) = G(x_R + x/a_i)/G(x_R + x/a_{i-1})$ satisfy (2.2) and (2.3) (with $b_n = -a_n x_R$).

Let $X_{ni} = (X_i - b_n)/a_n$ and let J_n be the point process of the points $\{X_{ni} : i = 1, \ldots, n\}$; that is $J_n(x, \infty) = \sum_{i=1}^n I(X_{ni} > x)$ is the number of exceedances in the sample over the level $a_n x + b_n$. Since $P\{I(X_{ni} > x) = 1\} = 1 - F_i(a_n x + b_n) \equiv \bar{F}_{ni}(x)$, under the URN condition, (2.2) is equivalent to

$$\sum_{i=1}^n \bar{F}_{ni}(x) \longrightarrow -\log G(x) = \lambda(x) \quad (n \to \infty) \ . \tag{2.5}$$

Let $m_{ni} = (M_{ni} - b_n)/a_n$; then the points $\{m_{ni} : i = 1, \ldots, n\}$ are the points of J_n in descending order. We have the following Poisson convergence.

Theorem 2.2 (Ref. [11]) *Under (2.2) and (2.3), there exists a nonhomogeneous Poisson process J on (x_L, x_R) whose mean measure at (x, ∞) is $\lambda(x)$. Moreover, if $m_1 \ge m_2 \ge \cdots$ are the points of J in descending order, then for each k*

$$(m_{n1}, \ldots, m_{nk}) \xrightarrow{D} (m_1, \ldots, m_k) \quad (n \to \infty) \ . \tag{2.6}$$

Notice that $\{\lambda(m_i)\}$ are the points of a standard homogeneous Poisson process (SHPP) on $(0, \infty)$; thus a simple exercise shows that the joint density of $(m_1, \ldots, m_k)$ is

$$\psi(x_1, \ldots, x_k) = G(x_k) \prod_{i=1}^k (-\lambda'(x_i))$$
$$(x_1 \ge x_2 \ge \cdots \ge x_k) \, , \tag{2.7}$$

where $\lambda'(x)$ is the derivative of $\lambda(x)$. The marginal df of m_k is given by

$$\lim_{n\to\infty} P\{m_{nk} \le x\} = \tag{2.8}$$
$$= P\{m_k \le x\} = G(x)\sum_{i=0}^{k-1} \lambda^i(x)/i!$$
$$= \int_{\lambda(x)}^{\infty} e^{-u}u^{k-1}du/(k-1)! \,.$$

The results (2.4)–(2.8) are the same as in the iid case, except that G belongs to $\mathcal{M}$, a much larger class than the classical EVD.

Suppose a df G is a candidate for the limiting df of M_n, then for each fixed k and large n we should have

$$\{\lambda((M_{ni} - b_n)/a_n) : i = 1, \ldots, k\}$$
$$\overset{D}{\approx} \{T_i : i = 1, \ldots, k\}$$

where $\{T_i\}$ are the points of an SHPP.

Suppose n is large. Fix a $k \ll n$; thus if G is the right df then

$$\{M_{ni} : i = 1, \ldots, k\}$$
$$\overset{D}{\approx} \{a_n\lambda^{-1}(T_i) + b_n : i = 1, \ldots, k\} \,.$$

A graphical method to verify that G is indeed the right df is to plot M_{ni} vs. $\lambda^{-1}(i)$ (we replace T_i by its expectation $ET_i = i$) for $i = 1, \ldots, k$. If the points are scattered around a straight line, then we have statistical evidence in favor of this G. Moreover, the slope and intercept are estimates of a_n and b_n, respectively. Another possible approach is *maximum likelihood.* Suppose one wants to fit $G_\theta \in \mathcal{M}$ as the limiting df for M_n, where θ is a parameter to be estimated. Then the likelihood based on $M_{n1}, \ldots, M_{nk}$ is approximately

$$L(\theta, a_n, b_n) = a_n^{-k}G_\theta((M_{nk}$$
$$-b_n)/a_n)\prod_{i=1}^{k}(-\lambda'_\theta((M_{in} - b_n)/a_n)) \,.$$

The triple $(\hat\theta, \hat a_n, \hat b_n)$ which maximizes L will be used as our estimates. This approach is similar to Ref. [12] or Ref. [13], except that G_θ is not limited to the EVD class.

For the iid case, let $\xi_1(p)$ and $\xi_n(p)$ be the p-quantiles of X_1 and M_n, respectively. Then for all $0 \le p \le 1$

$$\xi_n(p) = \xi_1(p^{1/n}) \approx \xi_1(1 + (\log p)/n) \,. \tag{2.9}$$

The classical extreme value theory implies

$$(\xi_n(p) - b_n)/a_n \to \xi(p) = G^{-1}(p) \quad (n \to \infty) \,. \tag{2.10}$$

In the general case, where $X_i \sim F_i$, (2.9) is not necessarily true but (2.10) still holds.

3 Functional Limit Theorems

Suppose one is interested not just in the maximum M_n of the sample, but also in its evolution along time. The process $\{M_{[nt]} : t > 0\}$ ($M_0 = X_1$) is a pure-jump Markov process, whose distinct values consist of the set of upper records of the sequence $\{X_i\}$; $M_{[n\cdot]}$ jumps at t if and only if X_{nt} is an upper record. Let $m_{n1}(t) = (M_{[nt]} - b_n)/a_n$, then we have

$$P\{m_{n1}(t) \le x\} = \prod_{i=1}^{[nt]} F_i(a_nx + b_n) = \tag{3.1}$$
$$= H_{[nt]}\left(a_{[nt]} \cdot \frac{a_nx + (b_n - b_{[nt]})}{a_{[nt]}} + b_{[nt]}\right) \,.$$

Suppose $H_n(a_nx + b_n) \to G(x)$ (G being nondegenerate). Then (3.1) will have a nondegenerate limit if and only if

$$\frac{a_n}{a_{[nt]}} \to \alpha_t \,;\ \frac{b_n - b_{[nt]}}{a_{[nt]}} \to \beta_t \quad (n \to \infty) \tag{3.2}$$

for some constants $\alpha_t > 0$ and β_t. Moreover, in this case $\lim_{n\to\infty} P\{m_{n1}(t) \le x\} = G(\alpha_tx + \beta_t) \equiv G_t(x)$. In fact, we have the following result.

Theorem 3.1 *Suppose (2.2) and (2.3) hold. Then*

$$\lim_{n\to\infty} P\{m_{n1}(t) \le x\} = G_t(x) \tag{3.3}$$
$$= G(\alpha_tx + \beta_t) \qquad (t > 0)$$

if and only if (3.2) holds for all $t > 0$. Moreover, if (α_t, β_t) is not identically $(1, 0)$, (3.3) implies the URN condition (2.3).

Note that in the iid case $F^n(a_n x + b_n) \to G(x)$ implies both the URII condition $F(a_n x + b_n) \to 1$ and $F^{[nt]}(a_n x + b_n) \to G^t(x)$ for all $t > 0$ (i.e. (3.3)). Reference [14] shows that α_t, β_t must have the form

$$\alpha_t = t^\rho \quad ; \quad \beta_t = c(1 - t^\rho)/\rho \qquad (3.4)$$

for some constants ρ and c (interpret $(1 - t^\rho)/\rho$ as $-\log t$ if $\rho = 0$). The convergence of the finite-dimensional distributions of $m_{n1}(\cdot)$ implies a functional limit theorem:

Theorem 3.2 (Ref. [15]) *Under (3.3), there exists a pure-jump Markov process $m_1 = \{m_1(t) : t > 0\}$ such that for all $0 = t_0 < t_1 < \cdots < t_k$ and $x_1 \le x_2 \le \cdots \le x_k$*

$$P\{m_{n1}(t_i) \le x_i \cdot i = 1, \ldots, k\} \to P\{m_1(t_i) \le x_i\}$$

$$= \prod_{i=1}^{k} \{G_{t_i}(x_i)/G_{t_{i-1}}(x_i)\}^{t_i - t_{i-1}} \qquad (3.5)$$

Theorem 3.2 claims that all the finite dimensional laws (fdl) of m_{n1} converge to those of m_1 (write $m_{n1} \xrightarrow{\text{fdl}} m_1$). Note that $m_1(\cdot)$ is right continuous and we can always choose a right continuous version for $m_1(\cdot)$. Therefore both m_{n1} and m_1 are elements of the Skorokod space $D(0,\infty)$ with the J_1-topology. Reference [15] proves full weak convergence in $D(0,\infty)$:

The process m_1 is called an *extremal process* and its transition probabilities are

$$P\{m_1(t+s) \le x \,|\, m_1(t) = y\} = \begin{cases} \dfrac{G_{t+s}(y)}{G_t(y)} & x \ge y \\ 0 & x < y \end{cases}$$

For joint convergence of $m_{n1}(i) = (M_i - b_n)/a_n$ $(i \ge 1)$, it is useful to employ the point process K_n of the points $\{(i/n, \ldots)\}$. Let K be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_-$ whose mean measure at $(0,t] \times (x,\infty)$ is $\lambda_t(x) = -\log G_t(x)$. If $\{(T_i, Y_i)\}$ are the points of K, let $m(t)$ be the kth largest Y_i among points with $T_i \le t$. Then obviously we have

$$P\{m_k(t) \le x\} = P\{K((0,t] \times (x,\infty)) \le k-1\} = G_t(x) \sum_{i=0}^{k-1} \lambda_t^i(x)/i! \,.$$

The main result of this section is the following:

Theorem 3.3 (Refs. [16, 17]) *Under (3.3), the point process K_n converges to the nonhomogeneous Poisson process K. Moreover, for each k, as $n \to \infty$*

$$(m_{n1}, \ldots, m_{nk}) \xrightarrow{D} (m_1, \ldots, m_k) \quad \text{in} \quad D^k(0,\infty) \,.$$

4 Upper Records

Given a sequence of random variables $\{X_i : i \ge 1\}$, X_j is an (upper) record if $X_j > M_{j-1}$; when all F_i are continuous, X_j is a record if $X_j = M_j$. The indices $\{L(j) : j \ge 1\}$ $(L(1) = 1)$, where the Markov process $\{M_k : k \ge 1\}$ jumps are called *record times* and the values $\{M_{L(j)} : j \ge 1\} = \{X_{L(j)} : j \ge 1\}$ are the *record values*. Let N_n be the number of records among the first n observations.

Record values and record times for iid sequences and their relation to Poisson and the extremal process are treated extensively in Ref. [3]. A lovely review of this subject is provided by Ref. [18]. Records of independent nonstationary sequences are treated in Refs. [19–25]. Reference [21] is motivated by the unpredicted high sea levels in The Netherlands that caused the collapse of the sea dikes and the loss of 2000 human lives; all the other authors are motivated by sport-records. The models treated can be presented as follows. The sequence of independent rv's $\{X_j : j \ge 1\}$ is such that

$$Y_j = Z_j + cd_j \quad (j = 1, 2, \ldots) \,, \qquad (4.1)$$

where c is a constant, $\{d_j\}$ is a monotone sequence, $d_j \in \mathbb{R}$, and $\{Z_j\}$ are iid with a common

df F. Here the df of X_j is $F_j(x) = F(x - cd_j)$, i.e. all the F_j are of the same type. Climatologists who believe in global warming can use these models for their data. For a linear-growth model we have the following useful result.

Theorem 4.1 (Ref. [19]) *Suppose* $EZ_1^+ < \infty$, $c > 0$ *and* $d_j = j$.

(a) There exists $p \in [0,1]$ *such that as* $n \to \infty$ $N_n/n \to p$ *a.s. and in* L_2 *and* $L(n)/n \to p^{-1}$ *a.s.*

(b) If F *is continuous,* $E(X_1^+)^2 < \infty$ *and* $0 < p < 1$ *then as* $n \to \infty$

$$\sqrt{n}(n^{-1}N_n - p) \xrightarrow{D} \mathcal{N}(0, \sigma^2) ;$$

$$\sqrt{n}(n^{-1}L(n) - p^{-1}) \xrightarrow{D} \mathcal{N}(0, p^{-3}\sigma^2)$$

for some $\sigma^2 = \sigma^2(p)$.

The *record rate* p is obtained from

$$p = \lim_{k\to\infty} p_k = \lim_{k\to\infty} \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} F(x + cj)dF(x) , \tag{4.2}$$

where $p_k = P\{A_k\} = P\{X_k \text{ is a record}\}$. Note that $EN_n = \sum_1^n p_j$. Ballerini and Resnick show that A_k and A_{k+m} tend to be independent for k, m large. When $F(x) = \Lambda((x-b)/a)$ $(a > 0)$, the $\{A_j\}$ are mutually independent, $p_j = \delta_j/S_j$, where $\delta_j = \exp(cj/a)$ and $S_j = \sum_{i=1}^j \delta_i$ (Smith and Miller (1984)) and

$$p = 1 - e^{-c/a} . \tag{4.3}$$

Let $\Delta_n = L(n+1) - L(n)$ be the *inter-record times.*

Theorem 4.2 (Ref. [20]) *For* $F(x) = \Lambda((x-b)/a)$, $c > 0$ *and* $d_j = j$, *as* $n \to \infty$ *we have*

$$\{\Delta_{n+k} : k \geq 1\} \xrightarrow{\text{fdl}} \{\Gamma_k : k \geq 1\} , \tag{4.4}$$

where $\{\Gamma_k\}$ *are iid geometric with* $p = 1 - e^{-c/a}$.

The two results $N_n/n \to p$ a.s. and (4.4) are complementary; by Theorem 4.1 $p_j = P\{A_j\} \to p$. Thus, the independence of the A_j implies that for all $k = k(n) \to \infty$ and $j \geq 1$, the number of records among $\{X_{k+1}, \ldots, X_{k+j}\}$ tends to $Bin(j, p)$. Hence the inter-record times must converge to independent geometric rv's with parameter p.

[illegible] the X_j are iid) [illegible] are independent and [illegible] $E\Delta_k = \infty$ for all $k \geq 1$ and $(\log \Delta_k)/k \to 0$ a.s. $(k \to \infty)$.

Ballerini and Resnick analyze the mile-race data in Ref. [19]. The fastest time X_j of every year (1860–1982) is plotted vs. j $(j = 1, \ldots, 123)$ and indeed a linear trend is exhibited. There are 36 records altogether and the record-rates $\{N_j/j : j = 1, 2, \ldots, 123\}$ do stabilize around $\hat{p} = 36/123$ from about $j = 50$ to the end.

Let us assume now that $F \in \mathcal{D}(\Lambda)$, i.e. $F^n(a_n x + b_n) \to \Lambda(x)$ for some constants $a_n > 0$ and b_n. Define the extremal processes for $\{Z_j\}$ and $\{X_j = Z_j + cb_j\}$ by

$$m_n^Z(t) = (\max\{Z_1, \ldots, Z_{[nt]}\} - b_n)/a_n$$

and

$$m_n^X(t) = (M_{[nt]} - (1+c)b_n)/a_n .$$

Then clearly $m_n^Z \xrightarrow{D} m_o$ in $D(0,\infty)$, where m_o is an extremal process as in (3.5) with $G_t(x) = \Lambda^t(x) = \Lambda(x - \log t)$. We have the following result for m_n^X.

Theorem 4.3 (Ref. [21]) *For* $c > 0$, *as* $n \to \infty$

$$m_n^X \xrightarrow{D} m_c \quad \text{in} \quad D(0,\infty)$$

where

$$\{m_c(t) : t > 0\} \stackrel{D}{=} \{m_o(t^{c+1}/(c+1)) : t > 0\} . \tag{4.5}$$

The properties of m_c can be read off easily from the properties of m_o via (4.5). In particular, for large n and y, $P\{M_n \leq y\} \approx F^{n^{c+1}/(1+c)}(y)$. It means that M_n behaves like a maximum of n^{c+1} iid rv's (whose common df is $F^{1/(c+1)}$). This implies an analogous result to Theorem 4.1(a).

Theorem 4.4 (Ref. [21]) *For N_n, the number of records among $\{X_j = Z_j + cb_j : j = 1, \ldots, n\}$ with $c \geq 0$, we have*

$$\frac{N_n}{(c+1)\log n} \to 1 \quad a.s. \quad and \ in \quad L_2 \ . \tag{4.6}$$

Notice that when $F = \Lambda$, we have $b_n = \log n$, $a_n \equiv 1$ and $X_j = Z_j + c \log j$, i.e. — we have a logarithmic growth.

Another result of a similar nature is the following.

Theorem 4.5 (Ref. [21]) *Let $\{Z_j\}$ be iid with $F(x) = \Phi_1(x) = e^{-1/x}(x > 0)$. Let N_n be the number of records among $\{X_j = Z_j + j : j = 1, \ldots, n\}$. Then*

$$E\,\frac{N_n}{\log^2 n} \to \frac{1}{2} \ , \quad \mathrm{Var}\,\frac{N_n}{\log^2 n} \to \frac{1}{6} \ .$$

Note, here we have a linear-growth model, but since $Z_1 > 0, EZ_1 = \infty$, Theorem 4.1 does not apply.

5 Extreme Value Times

Let $R_{nk} = \min\{j \leq n : M_{jk} = M_{nk}\}$, so $R_{n1} = L(N_n)$ is the last record time in $\{1, \ldots, n\}$, R_{n2} is the time (or index) of the second largest observation, etc. Asymptotic results for R_{n1} are taken from Ref. [26]. Results for R_{nk} $(k > 1)$ are new.

Observe first that

$$\begin{aligned} & P\{R_{n1} \leq k,\ M_n \leq x\} \\ &= P\{M(k,n) < M_k \leq x\} \\ &= \int_{-\infty}^{x} \prod_{i=k+1}^{n} F_i(y) d \prod_{j=1}^{k} F_j(y) \ , \end{aligned} \tag{5.1}$$

where $M(k,n) = \max\{X_{k+1}, \ldots, X_n\}$. Thus, we have the following asymptotic result.

Theorem 5.1 (Ref. [26]) *Under (3.3) we have for $0 \leq t \leq 1$*

$$\begin{aligned} & \lim_{n\to\infty} P\{R_{n1} \leq nt,\ m_{n1} \leq x\} = \\ &= \int_{-\infty}^{x} \frac{G(y)}{G_t(y)} dG_t(y) \\ &= \int_{-\infty}^{x} G(y) d \log G_t(y) \equiv H_1(t,x). \end{aligned} \tag{5.2}$$

Let $H_1(t) = H_1(t, \infty)$ be the limiting df of R_{n1}, then (3.3) is only sufficient for the existence of a limiting distribution for R_{n1} (take $F_i \equiv F$, where F is not in any domain of attraction of an EVD, but $R_{n1}/n \xrightarrow{D} U(0,1)$ (uniform on $[0,1]$)).

Theorem 5.1 is generalized as follows.

Theorem 5.2 *Under (3.3) we have for $k \geq 1$ and $0 \leq t \leq 1$*

$$\begin{aligned} & \lim_{n\to\infty} P\{R_{nk} \leq nt,\ m_{nk} \leq x\} = \\ &= \int_{-\infty}^{x} \frac{G(y)}{G_t(y)} \lambda^{k-1}(y) dG_t(y) \equiv H_k(t,x) \end{aligned} \tag{5.3}$$

and for $0 \leq s,\ t \leq 1$

$$\begin{aligned} & \lim_{n\to\infty} P\{R_{n1} \leq nt,\ R_{n2} \leq ns\} = \\ &= \int_{-\infty}^{\infty} \frac{G(x)}{G_s(x)} \lambda_t(x) dG_s(x) \equiv Q(t,s). \end{aligned} \tag{5.4}$$

For the special case

$$G_t(x) = G^{\phi(t)}(x) \tag{5.5}$$

for some function ϕ we have the following results.

Theorem 5.3 *Under (3.3) and (5.5) we have*

(i) $\phi(t) = t^\gamma$ for some $\gamma > 0$ and G must be an EVD.

(ii) $H_k(t) \equiv H_k(t, \infty) = \phi(t)$ $(k \geq 1)$

(iii) For each k, R_{nk} and M_{nk} are (asymptotically) independent.

(iv) For each k, $R_{n1}, \ldots, R_{nk}$ are (asymptotically) independent.

We have a 0–1 law.

Theorem 5.4 *Under (3.3), if for some k (fixed) $R_{nk}/n \xrightarrow{P} c_k$, where c_k is a constant, then $c_k \equiv c$ and $c = 0$ or 1.*

6 Two Growth Models

Let $\{Z_j : j \geq 1\}$ be iid, $Z_j \sim \Lambda$ and let $X_j = Z_j + d_j$ ($d_j \in \mathbb{R}$). Define $\delta_j = \exp(d_j)$ and $S_n = \sum_1^n \delta_j$. Then $F_j(x) = \Lambda(x - d_j) = \Lambda^{\delta_j}(x)$ and $P\{M_n \leq x\} = \Lambda^{S_n}(x)$. Hence the right normalization is $a_n \equiv 1$, $b_n = \log S_n$ and

$$\begin{aligned} P\{m_n(t) \leq x\} &= P\{M_{[nt]} - b_n \leq x\} \\ &= \Lambda^{S_{[nt]}/S_n}(x) \quad (t > 0). \end{aligned}$$

Since $F_j(x + b_n) = \Lambda^{\delta_j/S_n}(x)$, URN holds when $\delta_j = o(S_n)$ $(1 \leq j \leq n)$.

(i) *Logarithmic models.* Suppose $d_j = c \log j$. For URN we need $c \geq -1$. If $c > -1$, we have $b_n = \log(n^{c+1}/(c+1))$ and $P\{m_n(t) \leq x\} \to G_t(x) = \Lambda^{t^{c+1}}(x) = \Lambda(x - (c+1)\log t)$. Thus (3.2) holds with $(\alpha_t, \beta_t) = (1, (c+1)\log t)$. By Theorem 5.3, $H_k(t) = t^{c+1}$ $(0 < t \leq 1,\ k \geq 1)$. Moreover, the extreme values and their times of occurrence are (asymptotically) independent. The de Haan-Verkade result $N_n \sim (c+1)\log n$ (a.s.) holds in fact for $c > -1$ and not just for $c \geq 0$.

For $c = -1$ we have $b_n = \log\log n$, (3.2) holds with $(\alpha_t, \beta_t) = (1, 0)$, which means that $m(t) \equiv m(1)$ a.s. $(t > 0)$ — the extremes occur very early. Indeed, here, since $G_t \equiv G$, $H_k(t) = 1$ $(t > 0)$ (i.e. $R_{nk}/n \xrightarrow{P} 0$) and $N_n \sim \log\log n$ (a.s.). For $c < -1$, we have $S_n \nearrow S < \infty$ and $P\{M_n \leq x\} = \Lambda^{S_n}(x) \to \Lambda^S(x)$. URN does not hold, but we can show that $P\{R_{n1} = k\} \to k^c/S$ $(k = 1, 2, \ldots)$, i.e. the last record occurs at a finite time w.p.1. Since $N_n \leq R_{n1}$, N_n remains finite as $n \to \infty$. This is not surprising since the sequence $\{X_j\}$ is stochastically decreasing, so the first few observations determine the sample maximum.

(ii) *Polynomial models.* Let $d_j = j^\alpha$. When $\alpha = 0$ the $\{X_j\}$ are iid. When $\alpha < 0$, $\delta_j \to 1$. Thus, the $\{X_j\}$ are "almost" iid and they yield the same asymptotic results as the iid sequence.

For $\alpha > 0$ we have $b_n = \log S_n \approx n^\alpha + (1-\alpha)\log n$ and $b_n - b_{[nt]} \to \infty$. Hence, $m_n(t) \to -\infty$ (a.s.) $(0 < t < 1)$ but $P\{m_n(1) \leq x\} = \Lambda(x)$. This means, $G_t(x) \equiv 1$ and $H_k(t) \equiv 0$ $(0 < t < 1)$, i.e. $R_{nk}/n \xrightarrow{P} 1$. As long as $0 < \alpha < 1$, URN holds, $p_j = \delta_j/S_j \to 0$, $\sum_1^n p_j \sim \log S_n$ and $N_n \sim \log S_n$ (a.s.), i.e. $N_n = O(n^\alpha)$. Since $P\{R_{n1} \leq k\} = S_k/S_n$, one can show that for $0 < \alpha < 1$

$$\lim_{n\to\infty} P\{(n - R_{n1})\alpha n^{\alpha-1} \geq x\} = e^{-x} \quad (x > 0). \tag{6.1}$$

For $\alpha = 1$ (the linear model), URN does not hold, $p_j = \delta_j/S_j = e^{j-1}(e-1)/(e^j - 1) \to p = 1 - e^{-1}$ — the case treated by Ballerini and Resnick. Here, $N_n \sim np$ (a.s.) and the inter-record times are (asymptotically) geometric (result (6.1) is actually a generalization of this fact).

When $\alpha > 1$, $p_j = \delta_j/S_j \to 1$ very fast, so not only $N_n \sim n$ but except for finitely many, all X_j are records.

Acknowledgement. This research was supported by the Fund for the Promotion of Research at the Technion.

References

[1] Gumbel, E.J., *Statistics of Extremes*, Columbia University Press, New York (1958).

[2] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics*, John Wiley, New York (1978). (Second edition, Robert E. Krieger, Malabar, Florida (1987)).

[3] Resnick, S.I., *Extreme Values, Regular Variation and Point Processes*, Springer, New York (1987).

[4] Castilo, E., *Extreme Value Theory in Engineering*, Academic Press, New York (1987).

[5] Reiss, R.-D., *Approximate Distribution of Order Statistics*, Springer-Verlag, N.Y. (1989).

[6] Leadbetter, M.R., Lindgren, G. and Rootzen, H., *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, Berlin (1983).

[7] Husler, J., Extremes: Limit results for (univariate and multivariate) nonstationary sequences, this volume (1994).

[8] Mejzler, D., On a Problem of B.V. Gredenko, *Ukrain Math. Ž* 2, 67–184 (1949) (Russian).

[9] Mejzler, D., On the Problem of the Limit Distributions for the Maximal Term of a Variational Series, Lvov. Politechn. Inst., *Nauč. Zap. Ser. Fig.Math.* 38 (1956), 90–109 (Russian).

[10] Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. II, Second edition, John Wiley, New York (1971).

[11] Mejzler, D. and Weissman, I., On Some Results of N.V. Smirnov Concerning Limit Distributions for Variational Series, *Annals of Mathematical Statistics* 40 (1969), 480–491.

[12] Weissman, I., Estimation of Parameters and Large Quantiles Based on the k Largest Observations, *Journal of the American Statistical Association* 73 (1978), 812–815.

[13] Smith, R.L. and Weissman, I., Maximum Likelihood Estimation of the Lower Tail of a Probability Distribution, *Journal of the Royal Statistical Society*, Ser. B 47 (1985), 285–298.

[14] Weissman, I., On Location and Scale Functions of a Class of Limiting Processes with Application to Extreme Value Theory, *Annals of Probability* 3 (1975b), 178–181.

[15] Weissman, I., Extremal Processes Generated by Independent Nonidentically Distributed Random Variables, *Annals of Probability* 3 (1975a), 172–177.

[16] Weissman, I., On Weak Convergence of Extremal Processes, *Annals of Probability* 4 (1976), 470–474.

[17] Weissman, I., Multivariate Extremal Processes Generated by Independent Nonidentically Distributed Random Variables, *Journal of Applied Probability* 12 (1975c), 477–487.

[18] Glick, N., Breaking Records and Breaking Boards, *The American Mathematical Monthly* 85 (1978), 2–26.

[19] Ballerini, R. and Resnick, S., Records From Improving Populations, *Journal of Applied Probability* 22 (1985), 487–502.

[20] Ballerini, R. and Resnick, S. Records in the Presence of a Linear Trend, *Advances in Applied Probability* 19 (1987), 801–828.

[21] de Haan, L. and Verkade, E., On Extreme Value Theory in the Presence of a Trend, *Journal of Applied Probability* 24 (1987), 62–76.

[22] Smith, R.L., Forecasting Records by Maximum Likelihood, *Journal of the American Statistical Association* 83 (1988), 331–338.

[23] Smith, R.L. and Miller, J.E., Predicting Records, Technical Report, Department of Mathematics, Imperial College, London (1984).

[24] Smith, R.L. and Miller, J.E., A Non-Gaussian State Space Model and Applications to the Prediction of Records, *Journal of the Royal Statistical Society*, Ser. B 48 (1986), 79–88.

[25] Yang, M.C.K., On the Distribution of the Inter-Record Times in an Increasing Population, *Journal of Applied Probability* 12 (1975), 148–154.

[26] de Haan, L. and Weissman, I., The Index of the Outstanding Observation Among n Independent Ones, *Stochastic Processes and Their Applications* 27 (1988), 317–329.

Order Statistics and Proofs Of Combinatorial Identities

Wenocur, R.S.
University of Pennsylvania, Philadelphia, PA

An urn model approach to exceedances of order statistics leads to new proofs of combinatorial identities, one of which is Gauss's ${}_2F_1$ summation formula. The relationships among Gauss's ${}_2F_1$ identity, inverse Pólya distributions, and order statistics emerge as a consequence. We discuss connections among combinatorial methods of proving hypergeometric identities and our probabilistic approach, with emphasis on exceeding the upper order statistics, in particular the *maximum*, of a random sample

1. INTRODUCTION

Proofs of combinatorial identities have a long history. For a classic proof of Gauss's ${}_2F_1$ summation theorem see for example Slater (Ref. [1], pp. 27-28).

Often a proof of an identity using order statistics — in particular *extreme* order statistics — appears in the literature; see for example Refs. [2, 3, 4]. The purposes of this paper are: to present proofs of combinatorial identities using order statistics, with emphasis on Gauss's ${}_2F_1$ summation theorem which generalizes the result in Ref. [3]; to show relationships with *WZ-pairs* which are explained in the following paragraph; and to illustrate how classic results can be proven by various methods which are the consequences of posing mathematical questions in different contexts.

Many combinatorial identities can now be verified by means of the Wilf-Zeilberger certification theorems (Ref. [5]) which provide a method for certifying couples of identities via *WZ-pairs*. If two functions $F(N,k)$ and $G(N,k)$, defined for integer k and nonnegative integer N, satisfy the condition

$$\Delta_N F = \Delta_k G \quad \text{(for integers } N \geq 0 \text{ and } k\text{)},$$

(F,G) is a *WZ-pair*. Under the conditions that for each integer k, the limit $f_k = \lim_{N\to\infty} F(N,k)$ exists and is finite; for each integer $N \geq 0$, $\lim_{k\to\pm\infty} G(N,k) = 0$; and $\lim_{L\to\infty} \sum_{N\geq 0} G(N,-L) = 0$, we have a couple of identities

$$\sum_k F(N,k) = \text{const} \quad (N = 0,1,2,\ldots), \qquad (1)$$

and

$$\sum_{N\geq 0} C(N,k) = \sum_{j\leq k-1} (f_j - F(0,j))$$

Moreover, computerized proofs of hypergeometric identities are now possible (see Ref. [6]).

2 ORDER STATISTICS

As a consequence of studying the works of Galambos, including Ref. [7], Galambos and Seneta (Ref. [8]), David (Ref. [9]), and Gumbel (Ref. [10]), we analyzed order statistics from the following point of view, using an urn model approach (see, for example, Refs. [11, 12, 13]).

For a very clear and concise introduction to *order statistics and extremes*, see Galambos's book (Ref. [7], pp.16-17).

Let X_1, X_2, ... X_N be a random sample from an arbitrary real-valued continuous distribution. Call X_1, X_2, ... X_N the *previous* sample, and consider *future* trials $X_{N+1}, X_{N+2}, \ldots X_{N+k}, \ldots$ from the same distribution. With probability 1, the order statistics $X_{(1)} < X_{(2)} < \ldots < X_{(N)}$ associated with the previous sample determine $N+1$ disjoint random intervals $I_1, I_2, \ldots I_{N+1}$ on the real line, into which any future observations must fall. For any I_α, $\alpha = 1, 2, \ldots, N+1$, the conditional probability $P(X_{N+k+1} \in I_\alpha \,|\, X_{N+1}, X_{N+2} \ldots X_{N+k})$ is equal to $\frac{r+1}{N+1+k}$, where r is the number of X_{N+j}'s, $j = 1, 2, \ldots k$, that fall into I_α. Since the I_α's are almost surely disjoint, determining $P(X_{N+k+1} \in \bigcup_{\alpha \in S} I_\alpha | X_{N+1}, X_{N+2} \ldots X_{N+k})$, where $S \subset \{1, 2, \ldots N+1\}$, poses no difficulty. See for example Refs. [11, 13].

3. EXCEEDANCES

Sample until exactly m future trials exceed the j^{th} order statistic $X_{(j)}$ where $1 \le j \le N$. If $W_{N,j,m}$ equals the number of future trials until $X_{(j)}$ is exceeded m times, then (adopting the convention that the empty product equals 1, of course)

$$P(W_{N,j,m} = m+k) =$$

$$\frac{(m+k-1)!\,(N+1-j)\cdots(N+m-j)}{k!\,(m-1)!\,(N+1)\cdots(N+m)} \times \frac{(j)\cdots(j+k-1)}{(N+m+1)\cdots(N+m+k)}$$

$$= \frac{(N+m-j)!\,N!}{(N+m)!\,(N-j)!} \times \frac{(\mathrm{m})(\mathrm{m}+1)\cdots(\mathrm{m}+\mathrm{k}-1)(\mathrm{j})\cdots(\mathrm{j}+\mathrm{k}-1)}{\mathrm{k!}\,(\mathrm{N}+\mathrm{m}+1)\cdots(\mathrm{N}+\mathrm{m}+\mathrm{k})} \quad (2)$$

Since $\sum_{k=0}^{\infty} P(W_{N,N,m} = m+k) = 1$ (see Johnson and Kotz, Ref. [14], sections 4.4 and 4.5; Wenocur, Refs. [11,13]), it must follow that

$$\sum_{k=0}^{\infty} P(W_{N,j,m} = m+k) = 1 \,. \quad (3)$$

4. PROBABILISTIC PROOF OF GAUSS'S ${}_2F_1$ IDENTITY

Let us now apply results of §2 and §3 to prove Gauss's ${}_2F_1$ identity and to show its relationship to an inverse Pólya distribution.

Gauss's ${}_2F_1$ summation theorem often appears in the form (see for example Ref. [1], pp. 27-28 or Ref. [4])

$${}_2F_1[a,b;c;1] = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \quad (4)$$

for $c > a+b$; where ${}_2F_1[a,b;c;1]$ is defined by

$${}_2F_1[a,b;c;1] = \sum_{k=0}^{\infty} \frac{a\cdots(a+k-1)\,b\cdots(b+k-1)}{c(c+1)\cdots(c+k-1)\,k!} \,;$$

see for example Slater (Ref. [1], p. 1) or Ref. [4]. The purpose of this section is to present a probabilistic proof of (4) where a, b, and c are positive integers.

Directly from (2) and (3),

$${}_2F_1[m,j;N+m+1;1] = \frac{(N+m)!\,(N-j)!}{(N+m-j)!\,N!} = \frac{\Gamma(c)\,\Gamma(c-a-b)}{\Gamma(c-a)\,\Gamma(c-b)} \,,$$

where $a = m$, $b = j$, $c = N+m+1$, and where $N+m+1 > j+m$ is equivalent to Gauss's condition $c > a+b$.

Notice that (2) describes an inverse Pólya distribution; see Ref. [13], Ref. [14], pp.194-200, and Ref. [15]. Hence, Gauss's sum is related to this particular type of waiting-time distribution. A combinatorial argument leads to a different proof of Gauss's ${}_2F_1$ identity in Ref. [4]; nevertheless, the relationship between Gauss's ${}_2F_1$ sum and the inverse Pólya distribution emerges.

Bose-Einstein statistics are related, too; see, for example the now classic texts of Feller, that of Galambos, and Refs. [11, 13].

5. PROBABILISTIC PROOF OF A CLASSIC IDENTITY WITH A TERMINATING SUM

The urn model approach to order statistics presented in §2 is useful for proving various hypergeometric identities. In Ref. [3], a less general form of Gauss's identity is proven by means of a similar but more restrictive method. For different techniques which are combinatorial in nature, see Refs. [4, 5, 6].

As another example of applying the approach of §2, let us derive a result which involves a finite sum. Certainly, this is a well-known identity that appears in Ref. [16] and in textbooks, but let us derive it again in the present context for at least two reasons: (a) to illustrate the utility of methods employed in this paper; (b) to lead to an identity (6) that has been of interest in industry (Ref. [17]) but proved by longer means, namely, an induction argument, that ignores order statistics. Suppose we consider n *future* trials and let q_k be the probability that exactly k of these n trials have values less than $X_{(j)}$, the j^{th} smallest of *past* N values. The fact that $\sum_{k=0}^{n} q_k = 1$ and direct application of the urn model presented in §2 lead to the identity

$$\sum_{k=0}^{n} \binom{j+k-1}{k} \binom{N-j+n-k}{n-k} = \binom{N+n}{n} \qquad (5)$$

If $j = N = p+1$, (5) reduces to a result from the ancient field of *summation calculus*:

$$\sum_{k=0}^{n} (k+1)\cdots(k+p) = \frac{(n+1)\cdots(n+p+1)}{(p+1)} . \qquad (6)$$

A probabilistic interpretation of (6) involves the N^{th} order statistic. (that is, the *maximum*). Divide both sides of Eq. (6) by its right-hand-side. Thus, we obtain the sum, from $k = 0$ to n, of the probabilities that exactly k of our n future values fall below the maximum $X_{(N)}$ of our previous sample.

6. RELATIONSHIP TO SOME WZ-PAIRS

Strong connections exist between the urn model approach to order statistics and the Wilf-Zeilberger certification theorems of Ref. [5] described in §1. Suppose we restrict our attention to the j^{th} upper order statistic.

Let $m = 1$ and consider exceeding the j^{th} *largest* (that is, $X_{(N+1-j)}$, the $(N+1-j)^{th}$ smallest) of the first N values observed. Letting $F(N,k) = P(W_{N,N+1-j,1} = k)$ for $k = 1, 2, \ldots$, we have the *WZ*-pair

$$F(N,k) = \frac{j\,N!\,(N+k-1-j)!}{(N-j)!\,(N+k)!} ,$$

$$G(N,k) = \frac{j\,(1-k)\,N!\,(N+k-1-j)!}{(N-j+1)!\,(N+k)!} .$$

Suppose $m = 1$ and consider exceeding the *maximum* $X_{(N)}$ of the first N values observed. Letting $F(N,k) = P(W_{N,N,1} = k)$ for $k = 1, 2, \ldots$, we obtain the *WZ*-pair

$$F(N,k) = \frac{N}{(N+k)\,(N+k-1)} ,$$

$$G(N,k) = \frac{1-k}{(N+k)\,(N+k-1)} ,$$

and potential function

$$\Phi(N,k) = \frac{-N}{(N+k-1)} ,$$

where, as defined in Ref. [5],

$$\Delta_k \Phi = F \quad \text{and} \quad \Delta_N \Phi = G .$$

Although the notation is different, $F(N,k)$ appears in Refs. [3, 11, 12, 13] as a probability depending on N and k, namely, previous sample size and stopping time respectively. Probabilistic analysis of $F(N,k)$ presented here, employed in Refs. [3, 11, 12, 13], and described from another perspective in Ref. [4], provides a counterpoint to combinatorial methods in Ref. [5] and Ref. [6].

ACKNOWLEDGMENTS

Many thanks to Janos Galambos, Herbert S. Wilf, and Doron Zeilberger for lively discussions that led to these results. Additional thanks to Professor Galambos and to the referee for suggestions that improved the presentation .

REFERENCES

[1] Slater, L. J., *Generalized Hypergeometric Functions*, Cambridge University Press, Cambridge, 1966.

[2] Kadell, K. W. J., A probabilistic proof of Ramanujan's ${}_1\psi_1$ sum, *SIAM J. Math. Anal.*, 18 (1987), 1539-1548.

[3] Wenocur, R. S., Rediscovery and alternate proof of Gauss's identity, *Ann. Discrete Math*, 9 (1980), 79-82.

[4] Wenocur, R. S., A probabilistic proof of Gauss's ${}_2F_1$ identity, to appear in *J. Combin. Th.* (1994).

[5] Wilf, H. S. and Zeilberger, D., Rational functions certify combinatorial identities, *J. Am. Math. Soc.*, **3** (1990), 147 - 158.

[6] Wilf, H. S. and Zeilberger, D., An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities. *Invent. mathem.*, **108** (1992), 575-633.

[7] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, New York, 1978.

[8] Galambos, J. and Seneta, E., Record times, *Proc. Amer. Math. Soc.*, **50** (1975), 383 - 387.

[9] David, H. A., *Order Statistics*, John Wiley and Sons, New York, 1970.

[10] Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, New York, 1958.

[11] Wenocur, R. S., Waiting times and return periods related to order statistics, PhD thesis, Temple University, Philadelphia, 1979.

[12] Wenocur, R. S., Waiting times and return periods to exceed the maximum of a previous sample, *Statis. Distrib. in Scientific Work*, **6** (1981), 411 - 418.

[13] Wenocur, R. S., Waiting times and return periods related to order statistics: an application of urn models, *Statis. Distrib. in Scientific Work*, 6 (1981), 419 - 434.

[14] Johnson, N. L. and Kotz, S., *Urn Models and their Application*, John Wiley and Sons, New York, 1977.

[15] Janardan, K. G. and Patil, G. P., On multivariate modified Pólya and inverse Polya distributions and their properties., *Annals of the Institute of Stat. Math.*, **26** (1974), 271-276.

[16] Galambos, J., Introductory article: *Statistical Extremes and Applications*, (ed.: J. Tiago de Oliveira), Reidel, 1984, 13.

[17] Newsletter, Daniel H. Wagner Associates, Paoli, 1987.

An Examination Of The Extremes Of Selected New Zealand Rainfall And RunOff Records For Evidence Of Trend

Withers, C.S. and Silby, W.W.
Institute for Industrial Research and Development, Lower Hutt, New Zealand

This report examines selected annual NZ rain and river flow maxima and minima for evidence of long term trend. The model assumes a linear time trend while the residuals are from the generalised extreme value (GEV) distribution. This is the most popular model for dealing with extremes, and gives a first-order approximation to the theory of extremes.

The iterative estimation technique combines the high-efficiency L-moment method of Hosking et al. (1985) for the GEV-parameters with the maximum likelihood equation for slope.

Regions chosen were prone to flood or drought. The results were surprising: the evidence for the presence of a trend in rainfall and runoff was very weak using the least squares estimate (LSE) of slope but extremely significant when using our mixed MLE/L-moment method.

It is recommended that the techniques used here be tried on NZ temperature series, as a greenhouse effect is much more likely to show up as a trend in temperature than a trend in rainfall.

To eliminate the possibility of programming error it is also recommended that the method be reprogrammed from scratch.

1 INTRODUCTION

This paper examines selected annual New Zealand (NZ) rain and river flow maxima and minima for evidence of long term trend.

The main results, given in §4, regress these series against time, assuming that the residuals are from the generalised extreme value (GEV) distribution. This is the most popular model for dealing with extremes, and gives a first-order approximation to the theory of extremes. (Ongoing work aims at improving this approximation by considering the two main families of distributions that arise in theory: distributions with power tails and distributions with exponential-power tails.)

§2 describes the data series chosen.

The theory developed for estimating the parameters is outlined in §3. It is based on an extension and modification of the method of L-moments (also called probability weighted moments) given by Hosking et al. (1985) that uses the maximum liklihood equation for the slope iteratively with the L-moment method for the 3 GEV parameters.

Regions chosen were prone to flood and drought. The evidence for the presence of a trend in rainfall and runoff was very weak using the least squares estimate (LSE) of slope but extremely significant when using our mixed MLE/L-moment method. At first sight this is

surprising — partly because examining extremes throws away a lot of information, partly because one might expect the data series the data series to be too short for confirmation of a trend — and partly because rainfall need not increase linearly as global CO2 increases according to global circulation models as presently developed. In fact it may decrease in some regions: see Mullan and Renwick (1990) for New Zealand climate change from increased CO2 as inferred from a global circulation model, and Salinger et al (1990) for a scenarios approach to changes in New Zealand climate. However since Hosking et al (1985) showed that the L-moment estimates for GEV parameters were more efficient than the MLEs, one does expect that the LSE of slope for a model with GEV residuals would be inefficient compared with our MLE/L-moment method.
It is intended to work out an analytic form for the asymptotic covariance, both to confirm that it is regular enough for its jackknife estimate to be consistent, and to cut down on the long amount of time jackknifing took — typically over one hour per run, using Splus.
We recommend that our iterative estimation method be started from the LSE for slope rather than from slope 0, as in seven out of the 18 runs, beginning iterations from slope 0 gave convergence to a wrong result.
Jump scenarios were also tested for in an ad hoc manner: see §4 for numerical results.
Detection of a trend in rainfall is a much more difficult problem than in temperature, and it is recommended that the methods developed as well as refinements in progress be applied to NZ temperature series. In addition more theoretical work needs to be done to settle the question of the efficiency of regression methods in detecting a trend over methods based purely on analysis of extremes.

2 THE DATA SERIES CHOSEN

RAINFALL SERIES — MAXIMUM DAILY PER ANNUM in mm

Gisborne	876902	1937–87
Masterton	59604	1926–87
Timaru	414201	1881–1985
Palmerston North	53603	1928–87
Arthur's Pass	219510	1957–89
Arthur's Pass	219501	1941–87

RIVER FLOW SERIES – ANNUAL 15 MINUTE FLOOD PEAKS (cubic metres/second)
– MINIMUM AVERAGE OVER 7 DAYS (litres/second)

Motu	16502	Flood 1960–1987	Minimum 1960–1990
Ruamahanga	29201	Flood 1955–1987	Minimum 1955–1987
Opihi	69618	Flood 1936–1987	Minimum 1965–1986
Opuha	69614	Flood 1936–1987	Minimum 1965–1986
Manawatu	1032560	Flood 1929–1988	Minimum 1972–1989
Waimakariri	66401	Flood 1930–1987	Minimum 1967–1989

The number following each series is the NZ Meteorological Service's code.

3 THEORY: ESTIMATES FOR THE GEV DISTRIBUTION WHEN A TREND IS PRESENT

Much has been written on the theory of extremes for the case of stationary observations. See for example the books of Gumbel (1958), Leadbetter et al (1980), Galambos (1987) and Resnick (1987). However these books give very little theory for the case when the observations are not stationary.

One of the few papers dealing with a search for a trend is Smith (1989): this paper analyses the

number of exceedances of a given (high) level based on 15 years of hourly measurements. He applies techniques to remove short-term dependency and seasonality in the data. Then he uses the 3 parameter generalised extreme value **(GEV)** model for exceedances, allowing the parameters to vary from period to period as well as from year to year. (Here a 'period' is one or two months.) Parameters are estimated by maximum liklihood. Fit is measured by plotting the ordered (estimated) residuals against their expected value for GEV exceedances (the **'generalised Pareto distribution'**) and also by plotting the ordered values transformed by their estimated distributions against the expected value of the order statistics from a uniform distribution. This paper also gives a number of useful references, such as Husler (1986) on extremes of nonstationary sequences.

Davison and Smith (1990) proceed further with this approach, allowing the parameters of this model to be given functions of regression covariates, with the parameters estimated by maximimum liklihood or least squares.

Tawn (1988) and Smith (1986) give a related approach but use the joint distribution of the r largest extremes rather than the threshold method.

McKerchar and Pearson (1989) looked at 13 New Zealand continuous water level recordings of more than 50 years each and found some evidence (p15–16) of long term trend. They fitted the GEV distribution to annual flood peaks to a great number of New Zealand flood peak series and noted (p33) that the EV 1 distribution (that is the GEV distribution with $k = 0$) gave an unsatisfactory fit to many series.

The approach taken here is to fit the model

$$Y_i = bt_i + X_i, 1 \leq i \leq n \qquad (3.1)$$

where Y_i = extreme value for ith observation, t_i = year of ith observation (typically i) and the ith residual X_i is assumed to come from the GEV distribution

$$F(x) = exp\{-\lambda(x)\}, \lambda(x) = (1-kz)^{1/k}, \quad (3.2)$$
$$z = (x-\xi)/\alpha \text{ where } \alpha > 0.$$

The GEV distribution is the limiting distribution of

$$(max_{j=1}^{N} Z_j - b_N)/c_N \text{ as } N \to \infty,$$

if a limit exists for some $c_N > 0$ and b_N when $Z_1, \ldots, Z_N$ is a random sample from some distribution on R. The following examples are from §4 of Withers (1992a).

EXAMPLE 3.1. If

$$P(Z_1 > x) \approx (a/x)^b \text{ as } x \to \infty \text{ where } a, b > 0,$$

then one can take $b_N = 0$, $c_N = aN^{1/b}$, $k = -b^{-1} < 0$, $\alpha = b^{-1}$, $\xi = 1$.□

EXAMPLE 3.2. If

$$P(Z_1 > x) \approx fz^d e^{-z} \text{ as } x \to \infty$$
$$\text{where } z = \{(x-b)/c\}^a \text{ and } a, c > 0,$$

then one can take $b_N = cN_1^{1/a}\{1 + a^{-1}N_1^{-1}(dN_2 + f_1)\}$ where $N_1 = logN$, $N_2 = log\, N_1$, $f_1 = log\, f$ and $c_N = ca^{-1}N_1^{1/a-1}$, $k = 0$, $\alpha = 1$, $\xi = 0$.□

EXAMPLE 3.3. If Z_1 is bounded above by b and

$$P(Z_1 > x) \approx c(b-x)^a \text{ as } x \uparrow b \text{ where } a, c > 0,$$

then one may take $b_N = a$, $c_N = (cN)^{-1/b}$, $k = \alpha = a^{-1} > 0$, $\xi = -1$.□

By the first two examples, if the density of the underlying variable Z has upper tail falling to zero as a power law (or exponentially) then for $max_{j=1}^{N} Z_j$, the corresponding GEV distribution has $k < 0$ (or $k = 0$ respectively).

By the third example (with $Y_j = -Z_j$) if Y_1 is bounded below (as for daily rainfall or riverflow) then for $min_{j=1}^{N} Y_j$ the corresponding GEV distribution has $k > 0$.

Without loss of generality we took

$$\bar{t} = n^{-1}\Sigma_{i=1}^{n} t_i$$

to be zero. (That is we replaced t_i by $t_i - \bar{t}$. This is the same as reparameterising ξ.)

We refer to the 4 parameters of the model as

$$\theta = (b, \xi, \alpha, k)$$

If $k = 0$, $\lambda(x) = exp(-z)$, the limit as $k \to \infty$, so F is the "$EV1$" distribution.

The range of x is given by $1 - kz \geq 0$, that is

$$x \leq \xi + \alpha/k \text{ if } k > 0, \text{ and } x \geq \xi + \alpha/k \text{ if } k < 0.$$

If b is known to be 0, Hosking et al. (1985) have shown by simulation that the L-estimates (also called PWMs) for (k, ξ, α) are more efficient than the maximum liklihood estimates (MLEs) for

$n = 15, 25, 50$ and $k = -.4, -.2, 0, .2, .4$ and are comparable for $n = 100$. For $n = \infty$, the relative efficiencies of these estimates at $k = 0$ are about .8 for $\hat{k}$, .95 for $\hat{\xi}$ and .85 for $\hat{\alpha}$, but fall to 0 at $k = -.5$. The relative efficiency of $\hat{k}$ falls to 0 at $k = .5$ while the other 2 remain high (Figure 4). The $L-$estimates do not exist if $k \leq -1$; the mean MLEs do not exist if $k \geq 1/3$: see p252,254. (See Phien (1987) for a comparison of estimates for the case when k is known to be 0.)
However no results are available for the case when b is unknown. Tawn and Dixon (1992) use the model (3.1) on extreme sea levels, estimating θ by the MLE. Our approach is to use the MLE equation for b iteratively with equations (14), (15) of Hosking et al for the $L-$estimates of (k, ξ, α) based on the estimated residuals, $\hat{X}_i = Y_i - \hat{b}t_i$. Beginning with $b = 0$ we found that in all cases except two, 5 iterations gave accuracy to at least 4 decimal places.
The derivative of the mean log likelihood ratio w.r.t. b is $-g(b)/\alpha$ where

$$g(b) = n^{-1}\Sigma_{i=1}^{n} t_i\{W_i^{(1/k-1)} + (k-1)W_i^{-1}\} \quad (3.3)$$

where $W_i = 1 - kZ_i$ for $Z_i = (X_i - \xi)/\alpha$; $g(b)$ was plotted against $b = 0$ for the first and last iterations, and in each case found to be monotonic with a single root. (See the figures at the end of §5 for some examples.) Thus the method is not complicated by the presence of multiple roots. (To evaluate the root we used Newton's method.)
Our sample sizes ranged from 17 to 95 years and our estimates $\hat{k}$ from $-.45$ to .23 with none significantly different from 0. Only one $\hat{b}$ was significantly different from 0.
We end this section with a refinement available to the theory of Hosking et al. when $k > 1$ or $\hat{k} > 1$. (It turned out that in our applications $\hat{k}$ was less than one in each case so that this refinement was not needed.)
Hosking et al. noted that MLEs are 'not always satisfactory' if $k > 1$ as the density approaches ∞ as max X_i approaches its upper bound $u = \xi + \alpha/k$. In situations where the range depends on the parameters — in this case $X \leq u$, the corresponding estimate — in this case $\hat{u} = max\ X_i$, is superefficient, that is has variance $O(n^{-2})$, not just $O(n^{-1})$, so that one can reduce the variances by replacing say the 3rd of Hosking et al's 3 L-moment equations, that is their (12), by

$$\hat{\xi} + \hat{\alpha}/\hat{k} = x_n$$

(their notation for $max\ X_i$).
By their (15) this is equivalent to replacing their (13) by

$$b_0 + \hat{\alpha}\Gamma(1+\hat{k})/\hat{k} = x_n.$$

This results in their implicit estimate $\hat{k}$ of (13) being replaced by the explicit estimate

$$\tilde{k} = -\log_2\{1 - (2b_1 - b_0)/(x_n - b_0)\} \quad (3.4)$$

where $b_0 = \bar{X}, b_1 = (n^2 - n)^{-1}\Sigma_{j=2}^{n}(j-1)x_j$ and $x_1 \leq ... \leq x_n$ are the ordered values of $X_1, ..., X_n$. The corresponding estimates of ξ and α are now given by their (15) with $\hat{k}$ replaced by $\tilde{k}$.
To summarise: if $\hat{k} > 1$, use $\tilde{k}$. This will decrease the variances of the estimates with high probability.
(With exponentially small probability as n increases one could find $\tilde{k} < 1 < \hat{k}$; in that case one cannot rely on either estimate of k.)

4 NUMERICAL RESULTS

For each series we give the minimum, quartiles, maximum, and number of years covered (n), followed by the number of iterations (5 except for two cases), parameter estimates, their covariance as estimated by the jackknife method, the estimate and variance of the quantile estimates for the GEV distribution fitted. By (8) of Hosking et al the $F-$quantile of the GEV is

$$x(F) = \xi + \alpha\{1 - (-logF)^k\}/k = t(\theta) \quad (4.1)$$

This is estimated by $t(\hat{\theta})$ and its variance by $v(\hat{\theta}, \hat{C})$ where

$$v\ (\theta, C) = \dot{t}'C\dot{t}, \dot{t} = \partial t(\theta)/\partial\theta$$

and
$\hat{C}$ is the jacknife estimate of the covariance C of $\hat{\theta}$.
We then estimate the years t for which $EY_t = bt + EX$ will equal 1.1 and .9 of $E\bar{Y}$, — that is, the mean will increase by 10% or drop by 10%, according to the model (3.1). The t for which the EY_t changes to a fraction H of $E\bar{Y}$ is given by

$$bt + EX = H(b\bar{t} + EX) \quad (4.2)$$

that is by $t(\theta) = H\bar{t} + b^{-1}(H-1)EX$ where $EX = \xi + \alpha EZ$ and $EZ = k^{-1} - \Gamma(k)$.

We then estimate the standard deviation of $t(\hat{\theta})$ as for the previous example of $t(\theta)$.
Next we compute the goodness-of-fit statistic

$$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2/\{(n-3)var(\hat{X})\} \qquad (4.3)$$

where $var(X) = (\alpha/k)^2\{\Gamma(1+2k) - \Gamma(k+1)^2\}$.

(This is only an ad hoc statistic as the usual regularity conditions for its asymptotic distribution are not satisfied.)
Finally we give the 2-sample t-statistic for testing that the first and second half of the data series have the same mean, without assuming equal variances. Under the null hypothesis of no change in mean, this is asymptotically normal (0,1).

RAINFALL SERIES — MAXIMUM DAILY PER ANNUM

			Mean Years	Years available
Gisborne	876902	1937–87	1962	51
Masterton	59604	1926–87	1956.5	62
Timaru	414201	1881–1985	1937.505	95
Palmerston North	53603	1928–87	1957.5	60
Arthur's Pass	219510	1957–89	1973	33
Arthur's Pass	219501	1941–87	1964	47

RIVER FLOW SERIES – ANNUAL 15 MINUTE FLOOD PEAKS

Motu	16502	1960–1987	1973.5	28
Ruamahanga	29201	1955–1987	1971	33
Opihi	69618	1936–1987	1961.5	52
Opuha	69614	1936–1987	1961.5	52
Manawatu	1032560	1929–1988	1958.5	60
Waimakariri	66401	1930–1987	1958.5	58

NOTE 1 *For Timaru, data for the years 1883,1887–1895 is not available.* □

RIVER FLOW SERIES – ANNUAL 15 MINUTE MINIMA

Motu	16502	1960–1991	1975.5	32
Ruamahanga	29201	1955–1987	1971	33
Opihi	69618	1965–1986	1975.5	22
Opuha	69614	1965–74,1978–84	1973.97	17
Manawatu	1032560	1972–1989	1980.5	18
Waimakariri	66401	1967–1989	1978	23

NOTE 2 *The minima are for the 7 day averages of the 15 minute observations. The data years for minima and maxima flow are different as the minima flow came from electronic files and the maxima flows from McKerchar and Pearson (1989) Appendices 1 and 2.* □

Before giving the results in detail, we give the following summary of the estimates of the slopes b and the extreme value parameters k, and the t-statistics for testing the difference in means between the first and second half of each series.

SUMMARY OF ESTIMATES OF b AND k AND THEIR SIGNIFICANCE AND THE T-TEST

		$\hat{b}$	$\hat{b}/s.d.$	$\hat{k}$	$\hat{k}/s.d.$	$t-stat'c$
RAINFALL MAXIMA						
Gisborne	876902	.451	7.77	−.197	−12.79	.67
Masterton	59604	−.085	−4.16	−.179	−16.27	−.04
Timaru	414201	−.096	−13.92	−.044	−3.69	.15
Palmerston North	53603	−.049	−3.29	−.036	3.50	.77

Arthur's Pass	219510	1.000	4.85	.118	4.37	1.12
Arthur's Pass	219501	.445	6.39	.082	4.29	1.36
RIVER FLOOD PEAKS						
Motu	16502	−.339	.69	.229	9.38	.00
Ruamahanga	29201	2.128	4.82	.102	3.81	1.81
Opihi	69618	1.775	5.98	−.400	−27.93	.62
Opuha	69614	−.198	−.90	−.244	−1.59	−.74
Manawatu	1032560	1.342	2.43	−.025	−2.04	−.34
Waimakariri	664014	.519	.74	−.254	−22.28	−1.01
RIVER FLOW MINIMA						
Motu	16502	−27.18	−10.68	−.039	−14.31	−2.02
Ruamahanga	29201	84.42	N/A	.877	N/A	2.04
Opihi	69618	−34.97	−12.00	.366	9.48	−1.56
Opuha	69614	−57.69	−5.63	.766	9.17	−1.10
Manawatu	1032560	169.72	N/A	.675	N/A	−.14
Waimakariri	66401	52.07	.86	.346	11.32	.08

NOTE 3 *The value of $\hat{b}/\hat{sd}(\hat{b})$ when residuals were GEV turned out to be larger than that for normal residuals by a factor of 2 to 5. This could possibly suggest that convergence to normality is far from being approached at the sample sizes considered here. Although a search for a programming error to explain this effect was fruitless, this possibility cannot be ruled out.* □

The most striking feature of the analyses was the reduction in variance of $\hat{b}$ when moving from the LSE to our estimate. The result is to make all the b's highly significant — ie non-zero, whereas using the LSE one would draw the opposite conclusion — ie that they were not significantly different from 0.

There are several possible explanations: (a) the LSE of slope is highly inefficient for the model trend plus GEV residuals; — recall that our estimate mixes the MLE method with the L-moment method , and that the latter is more efficient than the MLE method. The LSE is efficient for normal residuals but not for GEV residuals — but it seems surpising that it should be so inefficient. This feature will be checked by obtaining an analytic form for the asymptotic variance; (b) the jacknife estimate of variance is too small because the analytic form for the asymptotic variance is not regular (ie. not estimable by replacing the distribution of residuals by their estimated empirical distribution); — this is the case for quantiles such as the median, and will be checked by obtaining an analytic form for the asymptotic variance in the coming year using the method of stochastic expansion of Withers (1987); however if the asymptotic variance was not regular one would expect the jackknife estimate of variance to be too large, not too small; (c) program error; the program is listed in Appendix B and has been carefully checked.

The second feature that jumps out is the highly significant values of the shape parameters, due to their small standard deviations — as estimated by the jackknife. This is no doubt closely connected with the small s.d.s for the slope estimate.

In this regard we quote Pearson (1992) p 67: "Three-parameter distributions such as the GEV should not be used to analyse flood frequency at single sites, unless the annual series is at least 30 years long, since sampling errors are much larger than for two-parameter distributions". With our four-parameter model variability will be even greater.

A third surprising feature of the data is that the slope of $\hat{b}$ is positive (and highly significant) in six cases but negative (and highly significant) in six cases. (This allows for the fact that negative minima not minima should be fitted to GEV distributions: these estimated slopes have been reversed so that a positive slope indicates an increase.) This is not so surprising on further thought as different regions are expected to react differently to global warming — some will get more rain and some less — though globally the average rainfall will increase.

NOTE 4 *These results were run beginning our iterations with b=0 and also beginning with the*

LSW for b to check that convergence to the same results occurred. In fact this failed for seven cases:

1. *for Gisborne maximum rainfall starting at b=0 the final estimates were* $\hat{b} = .156\ \hat{k} = -.194\ \hat{\alpha} = 19.712\ \hat{\xi} = 63.419$

2. *for Palmerston North maximum rainfall starting at b=0 the final estimates were* $\hat{b} = -.012\ \hat{k} = -.049\ \hat{\alpha} = 12.583\ \hat{\xi} = 42.773$

3. *for Motu flood peak starting at b=0 the final estimates were* $\hat{b} = .950\ \hat{k} = .163\ \hat{\alpha} = 77.546\ \hat{\xi} = 218.559$

4. *for Opihi flood peak starting at b=0 the final estimates were* $\hat{b} = .161\ \hat{k} = -.447\ \hat{\alpha} = 55.551\ \hat{\xi} = 77.337$

5. *for Opuha flood peak starting at b=0 the final estimates were* $\hat{b} = 1.210\ \hat{k} = -.280\ \hat{\alpha} = 82.063\ \hat{\xi} = 124.867$

6. *for Manawatu flood peak starting at b=0 the final estimates were* $\hat{b} = -1.749\ \hat{k} = -.043\ \hat{\alpha} = 468.015\ \hat{\xi} = 1160.312$

7. *for Waimakariri flood peak starting at b=0 the final estimates were* $\hat{b} = 4.904\ \hat{k} = -0.2800\ \hat{\alpha} = 361.867\ \hat{\xi} = 1177.764$

This is another reason to begin iteration with the LSE rather than 0. The other was that starting with b=0 gave some residuals so large that in some cases the method fails unless those outliers are discarded. □

NOTE 5 *The fact that the t-tests did not pick up any differences in mean between the first and second half of each series is consistent with explanation (a) above that the LSE is much less efficient than our estimate of slope; but it could also be argued to be consistent with the explanations (b) and (c).* □

5 Acknowledgement

The computing was carried out by Sarah Harper and Howard Silby. This was done in S-plus. A copy of the program is available on request.

6 REFERENCES

Bradford, E (1991) *Investigation of some statistical methods for detecting global warming.* DSIR Physical Science Report, Box 1335 Wellington.

Bradford, E (1991) *Further methods used on lake flows.* — Appendix D of (1991) progress report by Withers and Pearson listed below.

Bradford, E (1992) *Some statistical methods for detecting changes in mean temperature.* DSIR Physical Science Report 45, Box 1335 Wellington.

Bradford, E (1992) *South Island lake inflow : possible trend changes and modelling.* DSIR Physical Sciences Report 51. Applied Mathematics Group, DSIR, Box 1335, Wellington.

Bradford, E (1992) *Multiple regressions of rainfall at various NZ locations on time, site, sunspot index and Southern Oscillation index.* DSIR Physical Sciences Report, Applied Mathematics Group, DSIR, Box 1335, Wellington.

Davison, AC and Smith, RL (1990) *Models for exceedances over high thresholds.* J.R. Statist. Soc. B, 393-442.

Galambos, J (1987) *The asymptotic theory of extreme order statistics*, 2nd edition, Krieger, Melbourne, Florida.

Gumbel, EJ (1958) *Statistics of extremes.* Columbia University Press, New York.

Hawkins, DM (1977) *Testing a sequence of observations for a shift in location.* JASA 72, 180-186.

Hosking, JRM, Wallis, JR, and Wood, EF (1985) *Estimation of the generalized extreme-value distribution by the method of probability-weighted moments.* Technometrics 27, 251-261.

Leadbetter, MR, Lindgren, G and Rootzen, H (1980) *Extremes and related properties of random sequences and processes.* Springer-Verlag, New York.

McKerchar, AI and Pearson, CP (1989) *Flood frequency in New Zealand.* Publication 20, Hydrology Centre, Christchurch. ISSN 0112-1197.

Mullan, AB and Renwick, JA (1990) *Climate change in the New Zealand region inferred from general circulation models.* NZ Meteorological Service report 30.11.90 for the NZ Ministry for the Environment.

Pearson, CP (1992) *Analysis of floods and low flows.* Chapter 6 of **Mosley, M.P.** (1992) (ed) *Waters of New Zealand* NZ Hydroligical Society, Wellington.

Phien, HN (1987) *A review of methods of parameter estimation for the extreme value type-1 distribution.* Jnl of Hydrology 90, 251-268.

Resnick, S (1987) *Extreme values, regular variation and point processes.* Springer, New York.

Salinger, MJ, Mullan, AB, Porteous, AS, Reid, SJ, Thompson, CS, Withers, CS, Coutts, LA, and Fouhy, E (1990) *New Zealand climate extremes: scenarios for 2050 AD.* Report prepared by NZ Meteorological Service and DSIR Applied Mathematics Group for the Ministry for the Environment.

Smith, RL (1986) *Extreme value theory based on the r largest annual events.* Journal of Hydrology 86, 27-43.

Smith, RL (1989) *Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone.* Statistical Science 4, 367-393.

Tawn, JA (1988) *An extreme-value theory model for dependent observations.* Journal of Hydrology 101, 227-250.

Tawn, JA and **Dixon, MJ** (1992) *Trends in extreme sea levels* p 313-318 of Proceedings 5th International Meeting on Statistical Climatology, Toronto, June 1992.

Withers,CS (1987) *The bias and skewness of M-estimates in regression.* DSIR Applied Mathematics Technical Report 139.

Withers, CS (1990) *Extreme rainfalls recognising mixed processes. - Jointly with K.J.A. Revfeim.* DSIR Applied Mathematics Division Unpublished Report 1990/013.

Withers, CS and Pearson, CP (1991) *Detecting Climate Change in NZ rainfall and runoff records.* DSIR Physical Sciences Report 22.

Withers, CS (1991) *Moment estimates for mixtures with common scale. Communications in Statistics, Theory and Methods 20, 1445-1461.*

Withers, CS (1991) *Tests and confidence intervals for the shape parameter of a gamma distribution. Communications in Statistics, Theory and Methods, 20 (7), 2663-2685.*

Withers, CS (1992a) *Asymptotic multivariate distributions and moments of extremes,* to be submitted.

Withers, CS (1992b) *Expansions for quantiles and multivariate moments of extremes for a class of Pareto-type distributions.* Submitted.

Withers, CS (1992c) *Expansions for quantiles and moments of extremes for distributions of exponential-power type.* Submitted.

Withers, CS (1992d) *Expansions for the distribution of the maximum from a Pareto-type distribution when a trend is present.* Submitted.

Withers, CS (1992e) *Saddlepoint expansions and Bell polynomials* Submitted.

Withers, CS (1992f) *Expansions for the beta function and its inverse for one parameter large.* Submitted.

Withers, CS (1992g) *Estimates for mixtures of compound Poisson-gamma distributions.* Submitted.

Withers, CS (1992h) *Multivariate Bell polynomials, series, chain rules, moments and inversion.* Submitted.

Withers, CS (1992i) *Moment estimates for mixtures of several distributions with different means or scales.* Submitted.

Withers, CS (1992j) *The moments and cumulants of a mixture.* Submitted.

Withers, CS (1992k) *Estimating trend: Is daily, monthly or annual data best?* Submitted.

Withers, CS (1992l) *The distribution of the range of a Wiener process.* Submitted.

Extreme Values In Business Interruption Insurance

Zajdenweber, D.
Universitè de Paris X-Nanterre, Nanterre Cedex, France

The size-distribution of yearly claims in the French business interruption insurance branch is a Pareto law with an extremely long tail. The behavior of that law reflects the fact that the total value of yearly claims is dominated by a small number of major claims. We estimate the characteristic exponent of the tail, which is very close to one. This value means that the theoretical probability distribution has no expectation, and that business interruption insurance may be a very hazardeous economic activity.

INTRODUCTION.

The insurance industry lies on a foundation stone : the subdivision of risks through the law of large numbers. More precisely, in case of a statistical distribution of claims with a finite expectation, the law of large numbers proves that the average amount of claim per head (or per policy) becomes closer and closer to the expectation as the number of policies becomes larger. In the business interruption branch, as in some other branches where the risks may be catastrophic, Ref.[1], the rate of convergence of the average claim towards the expectation can be very small. Even worse, it can be nil, because of a small number of extremely large claims. In that case, the main problems posed to managers of insurance companies are the estimation of the actuarial value of the risks and the estimation of the appropriate amount of reserves necessary to cope with the extreme values of some claims. In the present paper, we shall only analyze the yearly size-distribution of business interruption claims in France.

I.EMPIRICAL EVIDENCE.

Fire is the most frequent cause of business interruption. If a damaged firm is insured against business interruptions, the insurance company pays for the losses in sales, minus the costs spared because of the interruption of the production. Two facts motivate the analysis of business interruption.

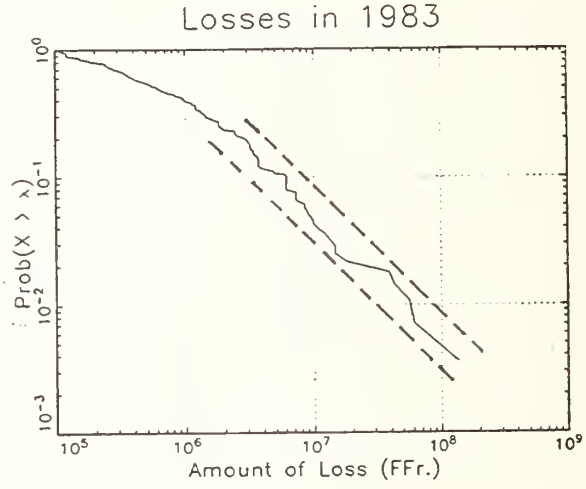
First, on a microeconomic level : the value of a business interruption claim is seldom a simple proportion of the size of the physical damges due to fire. Sometimes its value is insignificant compared to the fire-damages, and it may be much greater than the value of the equipment, machines, furniture or buildings burnt.

Second, on a macroeconomic level : the most striking feature of business interruption claims is the extreme variability of yearly claims recorded in a country. In France, for instance, the total amount of business interruption claims paid in 1988 by the insurance industry was nearly US $ 200 million, that was twice as much as 1987's (US $ 87 million). And for the first time in the history of that type of insurance in France, the total amount of the claims paid to the firms was greater than the premiums collected (US $193 million). The variability can be easily explained by the occurence of huge claims, the amount of which is greater than US $ 16 million and even sometimes greater than US $ 160 million. They are not numerous, but they make the tail of the yearly size-distribution of claims extremely long. The analysis of that tail is the main subject is the main subject of the paper.

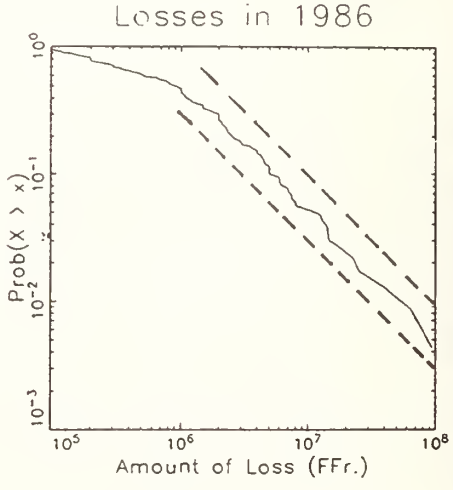
All the data used in our analysis originate from the statistics of the French insurance union : "ASSEMBLEE PLENIERE des SOCIETES D'ASSURANCES de DOMMAGES" (APSAD). it records all business interruption claims due to fire, the amount of which is greater than US $ 1600. All the claims are located in France. the available data span the years 1975 up to 1991. Yearly claims are ranked and valued in constant US $ (reference year 1988).

Fifteen out of the seventeen yearly size-distribution (all but 1979 and 1981) look like the typical distributions shown on the graphs #1 and #2. The abscissa is the Log of the sizes of the claims. The ordinate is the Log of the number of claims which are greater than the size on the abscissa. It is thus the Log of the complement of the cumulative distribution function of the random variable : "size of yearly claims". The common slope of the parallel dotted lines is -1. The small steps on the curves are by-products of the tendency to report rounded values of the damages.

N.B. Claims amounting to less than US $ 16000 (=100000 F.) are not shown on the graphs. That threshold may be approximatively considered as the median of the statistical distribution of claims : in 1988, for instance, 227 claims amounted to less than US $16000 and 223 amounted to more than this value. But the cumulated value of those small claims, below the median, is always insignificant. For instance, in 1988, 1989 and in 1990, their cumulated value amounted to a mere 0.3 percent of the total value of the yearly claims (in 1991 that proportion fell to 0.2 percent). Thus the loss of information is harmless.

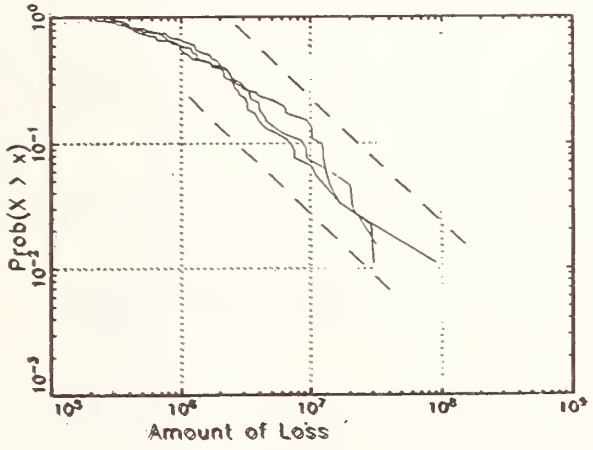


Graph #1



Graph #2

Except for the years 1979 and 1981 (fortunately, no very large claims occured these two years), the size distributions of yearly claims always display the same two features. Below a threshold of about US$ 330000 the curves are concave.But above that threshold, the tails fit a straight line the slope of which is close to -1. The behavior of the tail can be described and the value of its slope can be approximatively estimated by means of a graphical analysis. Graph #3 shows the superposition of three size-distributions of yearly claims corresponding to years 1975, 1976 and 1977. The common behavior is striking (in the next chapter we shall use a more rigorous estimation technique). The straight line is also known in economic literature as the Pareto line. Each year the cumulated value of the claims on the Pareto line amount to at least 80% of the total value of the claims. In 1988 and also in 1989 they even shared 92% of that total value. Thus it is easy to understand why the Pareto line is the heart of the matter for the insurance companies. Their managers speak about "lucky" years when no large claims occur, as in 1979 and in 1981. Nevertheless those happy years are exceptional, they give only a temporary relief. Most of the time the claims on the Pareto line strikes a severe blow to the profits of the insurance industry.



Graph #3

N.B. It is worth noting that the size-distribution of business interruption claims in Western Germany in 1989 also fits a Pareto line with a slope close to -1. But the data recorded by the COMITE EUROPEEN DES ASSURANCES in that country do not state the claims less than US$ 3,3 million. Hence the number of claims reported is small in absolute value (32 instead of 112 claims larger than US $ 330000 in France that year). Nevertheless the number of huge claims is relatively large in Western Germany, because the number of firms insured in that country is greater than in France (where there were only 12 claims greater than US $3,3 million the same year).

II.THEORETICAL PROBABILITY MODELS.

Since the data in our sample only record the yearly values of the claims greater than a (relatively) high threshold, and since the size-distribution is invariant (except for two "lucky" years), the theoretical model is to be found in the families of "max-stable" extreme distributions. One of the main theorem of the mathematical theory of extreme values, states that there are only three types of max-stable probability distribution functions,Ref.[2], Ref.[3].

- The **Pareto** d.f. : $\mathbf{1- x^{-\alpha}}$ $x > 1$ $\alpha > 0$
This d.f. has a "heavy" or "fat" upper tail. Its main mathematical feature is that it has no variance when $\alpha \leq 2$ and no expectation when $\alpha \leq 1$. The exponent α also gives the value the slope of the Pareto line on a double-log graph.

- The **Weibull** or "type II" d.f. : $\mathbf{1- (-x)^{-\alpha}}$ $1< x < 0$ $\alpha < 0$; This d.f. has a short upper tail. The Weibull d.f. is found in the context of a maximum value that cannot be passed. This is not the case in the data concerning business interruption : the maximum potential value of a claim is far greater than the record value already observed (between US $ 2,1 billion and US $ 18,7 billion!).

- The **exponential** d.f. : $\mathbf{1- e^{-x}}$ $x>0$
This d.f. has a medium upper tail. In practice, when dealing only with the large values exceeding high thresholds (US $ 3,3 million in our samples), this outstanding result means that only one of those three d.f. can be observed. However, since both the cumulated distribution function of the Weibull law and of the exponential law display a concave tail without a

straight line when drawn on a double-log graph, only the Pareto law may be relevant to the data on the business interruption claims. (With the exception of the two "lucky" years, 1979 and 1981, without major claims, that cannot fit the Pareto law, but may fit the exponential distribution or the Weibull law).

III. ESTIMATION OF THE SLOPES OF THE PARETO LINES.

Since each year we know the value of the empirical threshold (US$330000), the estimation of each yearly Pareto distribution is completely performed through the estimation of its characteristic exponent α.

Let N be the number of claims in the Pareto tail, x the size of a claim, x_0 the threshold value, i the rank of a claim with the value x_i (i=1,2,.., N); the maximum likelihood estimator of α (or Hill estimator, Ref[4]) is:

$$1/\hat{\alpha} = (1/N)\Sigma \text{Log}\, x_i - \text{Log} x_0$$

We have $E(1/\hat{\alpha}) = 1/\alpha$ and $\text{Var}(1/\hat{\alpha}) = 1/N\alpha^2$

Thus, the confidence interval defined by means of the central-limit theorem is, for N sufficiently large:

$$\sqrt{N}\alpha(1/\hat{\alpha} - 1/\alpha) \sim \Re(0,1)$$

This means that there is a probability 0.95 that the true value of $1/\alpha$ lies within the interval:

$$1/\hat{\alpha}\,(1 \pm 2/\sqrt{N})$$

Table #1 sums up all the estimations, except for the two "lucky" years 1979 and 1981 when no huge claims occurred, thus changing the Pareto line into a concave curve.The last column on the right shows the values of the largest claim (maxclaim) in thousands of US $, each year .

TABLE #1

ESTIMATED CHARACTERISTIC EXPONENTS

Year	N	α	conf.int.	maxcl.
1975	29	0.9523	0.69-1.51	6789
1976	46	0.9283	0.72-1.32	18844
1977	40	0.8937	0.68-1.31	6611
1978	46	1.0250	0.79-1.45	7434
1979	29	-	-	4685
1980	59	1.0130	0.80-1.37	11689
1981	48	-	-	4315
1982	71	1.1117	0.90-1.46	18667
1983	68	0.9775	0.79-1.29	24116
1984	52	0.9918	0.78-1.37	27191
1985	61	1.0549	0.84-1.42	6584
1986	71	1.0072	0.81-1.32	16776
1987	61	1.0152	0.81-1.36	5549
1988	91	0.8854	0.73-1.12	24547
1989	112	1.0235	0.86-1.26	13183
1990	114	0.9049	0.76-1.11	52542
1991	127	0.9779	0.83-1.19	16076

Average value of α =0.9842
Median value of α =0.9918

All these estimations are compatible with the theoretical value inferred from the graphical analysis of the data : one. The sample fluctuations are small (range of the estimated values : 0.8854-1.1117) and there is no trend in the estimated values of the characteristic exponents. But, a trend obviously appears in the number of yearly claims greater than US $3,3 million. It is due to the increasing number of firms insured. Nevertheless, no significant correlation is measured between the number of claims in the tail and the estimated characteristic exponents (R=-0.05).

A property of the Pareto law is its stability when the extreme values only are recorded, Ref.[5]. This means that the size-distribution of the n records of n identical and independent Pareto laws with a minimum value m and a characteristic exponent α=1 is the same Pareto law with a new minimum value n.m. Here the records are the 15 "maxclaims", hence the theoretical minimum value is US $ 5 million. We can verify that the Hill estimator gives the same characteristic exponent : α=0.9847. With other minimum values, greater than US $ 5 million, we have :

α=1.0001 with m= US $ 5078000.

α=1.0006 with m= US $ 6789000.

(but there are only 11 records larger than this very high value).

N.B. In Western Germany in 1989, α=1.0149 with the confidence interval 0.75-1.57. The largest claim amounted to US $ 240 million. This huge claim is not far from the greatest historical business interruption claim in France : US $ 330 million. It happened before 1975. It nearly equalled two years of business interruption premia paid in France. The second largest historical business interruption claim happened in 1992, in an oil refinery. It amounted to US $ 180 million.

CONCLUSION.

Large business interruption claims in France display a remarkable feature, their yearly size-distributions fit accurately a Pareto distribution with a constant characteristic exponent α around one. This means that without an objective ceiling, the theoretical probability distribution of the claims have no expectation. Claims can have a giant size, may be a size larger than the worst claims already experienced in the past.Those huge claims show that business interruption is of the same nature as natural hazards, for instance hurricanes or earthquakes, it can be an economic catastrophy.

REFERENCES.

[1] Zajdenweber D., Equité et Jeu de Saint-Petersbourg, Revue Econ.,1(1994)(in press).

[2] Galambos J., The Asymptotic Theory of Extreme Order Statistics, 2nd.Ed. Krieger, Malabar(Fla.), 1987.

[3] Gumbel E.J., Statistics of Extremes, Columbia University Press, New-York, 1958.

[4] Hill B.M., A Simple General Approach to Inference About the Tail of a Distribution, Ann.Stat.,3 (1975),1163-1174.

[5] Mandelbrot B.B., New Methods in Statistical Economics, J. of Pol. Econ., 71(1963), 421-440.

Acknowledgement. The author wishes to thank MM. Sylvain Tribouillois and Fabrice Genest for providing the data on business interruption claims in France, Mico Loretan for the graphs on the size-distribution of claims. Financial support from the "Fédération Française des Sociétés d'Assurances" (FFSA) is gratefully acknowledged.

* U.S. GOVERNMENT PRINTING OFFICE:1994-300-568/03050

NIST Technical Publications

Periodical

Journal of Research of the National Institute of Standards and Technology—Reports NIST research and development in those disciplines of the physical and engineering sciences in which the Institute is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Institute's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Institute's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NIST, NIST annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NIST under the authority of the National Standard Data Act (Public Law 90-396). NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published bimonthly for NIST by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Institute on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NIST under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NIST administers this program in support of the efforts of private-sector standardizing organizations.

Consumer Information Series—Practical information, based on NIST research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the **above** *NIST publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

Order the **following** *NIST publications—FIPS and NISTIRs—from the National Technical Information Service, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NIST pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NIST Interagency Reports (NISTIR)—A special series of interim or final reports on work performed by NIST for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.

U.S. Department of Commerce
National Institute of Standards and Technology
Gaithersburg, MD 20899