Alphabetical Listing of Abstracts: Conference on Extreme Value Theory and Its Applications

RECORD VALUES FROM UNIVARIATE DISTRIBUTIONS

Mohammad Ahsanullah
Rider College Lawrenceville, NJ

The paper aims to present recent developments about the inference of univariate distributions based on record values. Estimation of parameters, test of hypothesis, characterization problems and recurrence relations of moment and entropy of univariate distributions are discussed.

THE LAW OF ITERATED LOGARITHM FOR THE SQUARE INTEGRAL OF BROWNIAN MOTION

J. M. P. Albin
Chalmers University of Technology
Gothenburg, SWEDEN

Let $X(t) \equiv \int_0^t W(s)^2\,ds$ where W is Brownian motion. Cameron and Martin (1944) found the distribution of $X(t)$. Strassen (1964) proved that $\overline{\lim}_{t\to\infty} X(t)/(t^2 \ln\ln t) = 8\pi^{-2}$ (w. p. 1). (Li (1992) gave a probabilistic proof.) Csáki (1979) showed that $\overline{\lim}_{t\to\infty} \left[X(t) - 8\pi^{-2}t^2 \ln\ln t\right] / [t^2 \ln\ln\ln t] \leq 12\pi^{-2}$. We shall prove that $\overline{\lim}_{t\to\infty} \left[X(t) - 8\pi^{-2}t^2 \ln\ln t\right] / \left[t^2 \ln\ln\ln t\right] = 8\pi^{-2}$.

THE SEVERITY OF EXTREMES

Clive W. Anderson
University of Sheffield, U.K.

The severity of a storm or a flood is often a function not only of the peak value of whichever environmental variable is concerned, but also of other aspects of the extreme event, such as its duration and temporal shape. A run of days with temperatures just below freezing, for example, can be more disruptive to everyday human activity and to animal and plant life than a single day with a much sharper frost. Similarly, sustained moderately high water levels in a river or the sea can lead to greater flooding than a more extreme level lasting for only a short time. The statistical analysis of such examples of severity demands new methods for extremes, going beyond the traditional modelling of maxima or storm peaks.

This paper suggests that for an important class of applications the severity of an extreme event can usefully be modelled by the sum of the excesses of the environmental variable over a high threshold during the event. Under mild conditions the general form of the asymptotic distribution of such a sum is derived, and special cases are used to motivate statistical procedures for severity.

Applications of the methods to floods on the River Thames near Oxford and to ozone levels north of London are described.

ENVIRONMENTAL SAMPLING, OBSERVATIONAL ECONOMY, AND HIGHER RANK ORDER STATISTICS

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

Issues in environmental sampling and ecological monitoring often involve extreme values as the main inferential target (for instance, as indicators of hazardous waste "hot spots") or as design tools for cost-effective selection and acquisition of sampling units while maintaining desired levels of precision. Although conventional sampling methods address the problem of estimating the population mean or total with a desired precision or at a desired cost, these classical procedures may not always be resource-effective for studies involving extreme values or order statistics. In this paper, we review some procedures that allow :

1. inference on sample extreme values based on sample means;
2. inference on the population mean or total based on sample extreme values; and
3. inference on both the population mean and extreme values,

while maintaining observational economy at the same time. The methods are illustrated with examples in compliance monitoring and enforcement, stock assessment based on fisheries data, and hazardous waste site characterization.

BOOTSTRAPPING EXTREMES OF I.I.D. RANDOM VARIABLES

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

Let $X_1, X_2, \ldots$ be i.i.d. random variables. Define $M_n \equiv max(X_1, X_2, \ldots, X_n)$. Assume that there exist $\{a_n\}$ and $\{b_n\}$ such the $G_n(x) \equiv P\{a_n(m_n - b_n) \leq x\}$ converges to one of Gnedenko's extreme value distributions. We consider the problem of estimating $G_n(x)$ by the bootstrap technique. Angus (*Communication in Statistics* (1993) 15-30) showed that the naive bootstrap estimator of $G_n(x)$ has a random limit if the resample size $m = n$ and thus bootstrap fails in this case. In this paper we show that the boostrap estimator is weakly consistent if $n = o(n)$ and it is strongly consistent if $m = o(\frac{n}{logn})$ both when a_n and b_n are known and when a_n and b_n are estimated from the data. Some applications to the problem of obtaining confidence intervals for population parameters are included.

EXTREME ANALYSIS OF WAVE PRESSURE AND CORROSION FOR STRUCTURAL LIFE PREDICTION

Bilal M. Ayyub
University of Maryland
College Park, MD

The factors that affect the life of a structure include design parameters, design safety factors, design methods, type of structure 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 occurence, (2) statistical uncertainties due to using limited amount of information in estimating the characteristics of the population, (3) model uncertainies 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 (1989), Ayyub and White (1988), Harris et al (1981), Tallin et al (1988), and Yazdani and Albrecht (1987).

RECORD VALUES FROM RAYLEIGH AND WEIBULL DISTRIBUTIONS AND ASSOCIATED INFERENCE

N. Balakrishnan
McMaster University
Hamilton, Ontario, CANADA
and
P. S. Chan
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.

EXTREMES AND MEAN RESIDUAL PLOTS

J. Beirlant, J. L. Teugels and P. Vynckier
Katholieke University
Leuven, BELGIUM

The tail of a mean residual life plot has been used extensively in tail estimation techniques. For example, the well-known Hill (1975) - estimate provides an estimator of the mean residual life function of the log-transformed data at an upper quantile. The moment estimator of Dekkers, Einmahl and de Haan (1989) for the extreme value index incorporates an estimate of the variance function corresponding to the mean residual life function.

In this contribution it is shown that the domains of attraction of extreme value distributions can be characterized using the composition of the mean residual life function and the generalized quantile function. This result is then used to motivate a new estimator for the extreme value index. Simulation results together with the first asymptotics are derived. Comparisons with the corresponding results for the Pickands (1975) estimator and the estimator presented in Dekkers et al (1989) are presented.

Special attention will be paid to the so-called Gumbel domain which contains a very broad spectrum of tails. Here the class of Weibull type models plays an important role. Estimators for the so-called Weibull coefficient are given which again are motivated by proper mean residual life modeling. This approach also leads to new high quantile estimator.

References

1. Dekkers, A. L. M., Einmahl, J. H. J. and de Haan, L. (1989) A moment estimator for the index of an extreme value distribution. *Ann. Statist.* **17** 1833-1855.
2. Hill, B. M. (1975) A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163-1174.
3. Pickands III, J. (1975) Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119-131.

A NEW CLASS OF RANDOM FIELDS AND THEIR EXTREME VALUES

Simeon M. Berman
Courant Institute of Mathematical Sciences
New York University, New York, NY

Let X be a random vector R^n, and let $f(t), t\epsilon R^N$, be a continuous function into R^n, for $n \geq 1, N \geq 1$. Let (x, y) be the usual inner product in R^n, and let $\| x \|$ be the corresponding norm, for x, y in R^n. We define the real random field $X(t)$ as $X(t) = (X, f(t)), t\epsilon R^N$. The basic hypothesis on the random vector is that it has a joint distribution which is invariant under all orthogonal transformations of R^n, that is, X and UX have the same distribution for any real $n \times n$ orthogonal matrix U. The function f can be of a very general form, requiring continuous partial derivatives up to order 2. It is assumed that the marginal distributions of $X(t)$ are identical. A necessary and sufficient condition for this is that $\| f(t) \| =$ constant in t.

This stochastic process was introduced by the author in *Communications in Statistics - Stochastic Models*, 4 (1988) 1-43 in the special case $N = 1$. There the main results were limit theorems for $Z = max(X(t) : 0 \leq t \leq 1)$ and the extreme sojourn time $L_u = mes(t : 0 \leq t \leq 1, X(t) > u)$. In particular there was given i) the asymptotic formula for $P(Z) > u$, for $u \to x_0$, where x_0 is the least upper bound of the support of the marginal distribution, and ii) the asymptotic formula for $P(v(u)L_u > x)$ for fixed $x > 0$ and $u \to x_0$, where $v(u)$ is a positive nondecreasing function u. The current work extends the results to the more general case $N \geq 1$. The higher dimensionality of the time parameter t introduces novel questions, and so the analysis requires more general tools than those in the case $N = 1$. The practical aim of this generalization to $N \geq 1$ is to introduce a new class of stochastic models of random fields for possible applications to extreme value analysis in areas such as environmental modeling and image processing. For example, the moving height of the water surface in a region of the ocean can be modeled as a random field $X(t) = X(t_1, t_2, t_3)$, where t_1 is a time coordinate, and t_2 and t_3 are plane coordinates.

To indicate the novelty of the mathematical questions arising in the case $N \geq 2$, recall that the limit of the conditioned "high level processes" for $N = 1$ is of the form

$$\eta + \xi s \| f'(t) \| - \frac{1}{2} s^2 \| f'(t) \|^2, s \epsilon R^1,$$

where η and ξ are independent, η has a standard exponential distribution, and ξ has a standard exponential distribution. This accounts for the limiting behavior of Z and L_u. In the case $N \geq 2$, the limiting random field is of the more general form,

$$\eta + \xi \prime As - \frac{1}{2} \| As \|^2,$$

where η and ξ are independent, η is exponential, and ξ is a normal random vector with i.i.d. standard normal components; $s \epsilon R^N$, and A is the $n \times N$ matrix with (i, k) entry equal to the partial derivative of the kth component of f with respect to t_i. The positive sojourn time of this random field is represented as the volume of an ellipsoid in R^N associated with the (positive definite) matrix $A'A$.

ON THE ESTIMATION OF THE PARETO TAIL-INDEX USING K-RECORD VALUES

Alexandre M. Berred
Universités Paris-VI & VII
Le Havre, 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. Let $k = \{k_n, n \geq 1\}$ be a sequence of integers such that $1 \leq k_n < n$ for $n \geq 1$. Denote by $\{X(n), n \geq 1\}$ the corresponding sequences of k-record values. In this paper we consider k-record versions of two statistics introduced in Berred (1992) and sstudy their asymptotic behavior. Such statistics can be used as alterative estimates for the exponent of regular variation of the tail $1 - F$.

References

Berred, M. (1991), On record values and the estimation of the estimation of the Weibull tail-coefficient, *C. R. Acad. Paris, Série I*, **312**, 943-946.
Berred, M. (1992), On recor values and the exponent of a distribution with a regularly varying tail, *J. Appl. Probab.*, **29**, 575-586.
Berred, M. !992), K-record values and the extreme-value index, 1992, *presented at $XIII^{es}$ Rencontres Franco-Belges de Statisticiens, submitted to JSPI.*

PROBING THE NATURE OF THE BRIGHTEST GALAXIES USING EXTREME VALUE THEORY

Suketu P. Bhavsar
University of Kentucky University of Virginia
Lexington, KY Charlottesville, VA

An application of extreme value theory to a problem in Astronomy regarding the nature of the brightest galaxies in rich clusters is described. The extremely small dispersion in the luminosities of these galaxies has been cause for much debate regarding their status as "special" objects or "normal" extremes of a statistical distribution. A comparison of the Gumbel distribution with the observations shows that these galaxies are NOT the tail end of a statistical population. A model proposing that these first-ranked galaxies are drawn from two distinct populations describes the observed probability of their luminosities very well. The theory and model make predictions regarding relations between mean physical properties of the two populations, which are borne out by observations. It is a triumph for extreme value theory that astronomers have learned more about these objects by treating them as a class of extremes, rather than as unique members of individual clusters.

HIGH BOUNDARY EXCURSIONS OF LOCALLY STATIONARY GAUSSIAN PROCESSES

H. U. Bräker
IMS
Bern, SWITZERLAND

Let $\{X(t), t\epsilon T\}$ be a locally stationary Gaussian process and $f(t), t\epsilon T$ a continuous function, where T is a infinite or infinte interval. An asymptotic estimate for small probabilities $P(Z(t) \geq f(t)$, some $t\epsilon T)$ is derived by approximating the density of the first passage time and integrating over T. This work extends a result proven by J. Cuzick (1981) for stationary Gaussian processes.

ASYMPTOTIC APPROXIMATIONS FOR THE CROSSING RATES OF FUNCTIONS OF POISSON SQUARE WAVE PROCESSES

K. Breitung
München, GERMANY

For the extreme value distribution of functions of random vector processes it is difficult to derive exact expressions; therefore approximations are needed. An often used model for loads in structural reliability is a Poisson square wave process $x(t)$. Such a process is defined by a Poisson point process and an additional sequence of i.i.d. random variables $Y_0, Y_1, Y_2, \ldots$ where the value of $x(t)$ is defined by $x(t) = Y_n$ between two points t_n and t_{n+1} of the Poisson process. In this paper the asymptotic Poissonian behavior of the point process of level crossings of functions of independent Poisson wave processes is shown. This can be used to approximate the extreme value distribution of such functions.

AN ASYMPTOTIC EXPANSION FOR LARGE DEVIATION PROBABILITIES OF GAUSSIAN RANDOM VECTORS

K. Breitung
München, GERMANY
and
W. D. Richter
Universität Rostock
Rostock, GERMANY

RANDOM RECORD PROCESSES AND STATE DEPENDENT THINNING

Sid Browne
Columbia University
New York, NY
and
John Bunge
Cornell University
Ithaca, NY

Suppose that a point process $\bar{N}_t = T_1, T_2, \ldots$ on $[0, \infty)$ is thinned by independently retaining T_n with probability p_n. Our main examples are the classical *p-thinning*$(p_n \equiv p)$ and the *random record process* $(p_n = 1/n)$. When $\bar{N}_t$ is a mixed, nonhomogeneous Poisson process, we find conditions such that the thinned process is Poisson, and we discuss several examples. When $\bar{N}_t$ is a linear birth process (gamma-mixed Poisson with exponential rate), we show that the record process is Markov renewal, with an interesting structure, and we compare this with related asymptotic results. When $\bar{N}_t$ is a Mittag-Leffler renewal process (the homogeneous Poisson is a special case),

we give a "Deheuvels-type" representation of the record process (Deheuvels, 1982) and related characterization results.

EXTREMES IN ENGINEERING APPLICATIONS

E. Castillo
University of Cantabria
Santander, Spain

In this paper some engineering applications of extremes are described. An analysis of the main problems is performed and a revision of the most common techniques of the statistical analysis of extremes, including models, methods of estimation, domains of attraction and families of distributions is made. Finally, the problems of dependence and extrapolation are analyzed.

EXTREME VALUE ANALYSIS OF WAVE HEIGHTS

E. Castillo and J. M. Sarabia
University of Cantabria
Santander, Spain

This paper discusses the most common problems associated with the determination of design wave heights. It analyses two common methods used in fitting wave data and shows some of the stability or inconsistency problems associated with commonly used distributions. Methods to obtain confidence intervals, detecting of outliers and treatment of missing data are given.

AN EXPERT SYSTEM PROTOTYPE FOR THE ANALYSIS OF EXTREME VALUE PROBLEMS

E. Castillo, E. Alvarez, A. Cobo and M. T. Herrero
University of Cantabria
Santander, Spain

This paper presents an expert prototype for the analysis of extreme value problems. The system includes a package of computer-aided instructions 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.

SHORT-RECORD-BASED EXTREME WIND SIMULATION

Edmond D. H. Cheng and Arthur N. L. Chiu
University of Hawaii at Manoa
Honolulu, Hawaii

In the current state-of-the-art, there are at least three approaches for estimating extreme wind speeds from short-term records. The first approach is based on the analysis of the largest monthly wind speeds of at least three years' data. This method has been well tested at 36 weather stations situated in various locations in the United States, and it has proven to be practical and adequate. The second approach is the determination of extreme wind speeds from a parent distribution. Using this method, extreme wind speeds can be estimated even from records shorter than three years. The third approach is simulation models based on short-term continuous hourly wind speed records. Our contribution to the third approach is a short-record-based stochastic simulation model, which is capable of generating long-term hourly wind data.

The basic strategy of our simulation model is to generate the time series of hourly wind speeds in parts: for those winds associated with well-behaved climates in which extraordinary strong winds are not expected to occur; and, for those extreme winds belonging to tropical cyclones or other extratropical winds. In other words, the time series of hourly wind speeds is a time-dependent system, and this system is decomposed into a set of components for which operating rules of these components may be designed. Under this concept, a master computer program is designed to generate hourly wind data on a daily cycle basis.

Applications of this model to generate long-term extreme winds, on the basis of short records at Dallas/Fort Worth Airport and Corpus Christi NAS of Texas, will be demonstrated.

SOME ASPECTS OF SPATIAL EXTREMES

Stuart G. Coles
The University of Nottingham
Nottingham, U.K.

Most meteorological or environmental processes have spatial aspects as well as temporal. Thus, in the context of extreme value modelling, it is often required that the spatial variation in extremal behaviour is assessed or modelled. This gives rise to two specific modelling issues: models for *marginal* extremal behaviour, and models for *spatial dependence* of extremes, as well as procedures of inference for each. This paper reviews current approaches to these aspects in the context of the specific problem of modelling spatial aspects of the extreme wind climate over the United Kingdom.

APPLICATION OF EXTREME VALUE THEORY TO STRUCTURAL SYSTEM RELIABILITY ANALYSIS

Ross B. Corotis, Wei Wang and Martin R. Ramirez
The Johns Hopkins University
Baltimore, MD

Engineering structures are designed to withstand the uncertain process of environmental and man-made forces as well as deterioration over an economic lifetime, often assumed as 50 years. It is understood that certain information is lost when the extreme value of a process is used instead of the full process itself. In the cases of the structural system reliability, this loss includes the information needed to determine modal failure probabilities, defined as the probability of occurrence of each mode. It is shown in this paper that modal failure probabilities cannot be uniquely defined under the framework of time-invariant analysis.

Several definitions used by structural reliability engineers, implicitly or explicitly, are discussed. Both the magnitude and the relative ranking of the modal failure probabilities are shown to be quite different using these various definitions and the true process-related definition. Therefore, more information than the extreme value distributions must be known to assess the consequence of failure of a structural system by its various collapse modes.

A MULTIVARIATE MODEL FOR REPEATED FAILURE TIME MEASUREMENTS

Martin Crowder
Surrey University
Guildford, U.K.

A multivariate survival distribution is proposed. It includes as special cases certain distributions which have been used for multivariate failure time in recent years. Some properties are derived and some tests for particular aspects are proposed.

CONCOMITANTS OF EXTREME ORDER STATISTICS

H. A. David
Iowa State University
Ames, IO

Suppose the independent pairs of variates $(X_i, Y_i), i = 1, \ldots, n$, are ordered by the X_i. Then the Y-variate paired with the r th order statistic $X_{r:n}$ is called the concomitant of $X_{r:n}$ and denoted by $Y[r:n]$. This paper treats the asymptotic distribution theory of concomitants when r or $n-r$ remains fixed as $n \to \infty$. The special case of a linear regression linking X and Y is examined in detail and a theroem of Galambos (1978) is generalized. Results on the joint distribution of extreme concomitants are reviewed and some applications are indicated.

THE JOINT LONG-TERM DISTRIBUTION OF DESIGN ENVIRONMENTAL FACTORS FOR OFFSHORE STRUCTURES

Liu Defu, Shi Jiangang, and Wang Zheng
Tianjin University
Tianjin, CHINA

In the development of ocean engineering, one of the most important factors which determines the structural safety, cost and suitable form of engineering is the selection of the ocean environmental design criteria. Due to the complexity,variation and randomness of ocean environmental conditions, the commonly used methods for determining design criteria do not consider the jointly occurring probabilities of several environmental factors. This problem leads to overestimating the design criteria and often results in unnecessary overspending in engineering. Based on the measured and hindcasting data and multi-dimensional joint probability theory, the joint effect of wind, wave and current on offshore structures and is studied by the application of random simulation techniques. The paper presents design criteria of environmental loads for the realistic engineering design.

PROBABILISTIC MODELS FOR SUMS OF CUBES AND HIGHER POWERS

Jean-Marc Deshouillers
Bordeaux, FRANCE

The structure of the divisors of integers which are sums of 2 squares, known since the work of Fermat in the XVII th centrury, easily implies that sums of 2 squares have a zero (upper) asymptotic density, i.e.

$$\limsup {}^{\#}\{n < x, n = a^2 + b^2\}/x = 0, \quad \text{when } x \to \infty.$$

To determine the density of sums of 3 cubes, 4 biquadrates ... seems presently out of reach. Three conjectual approaches have been proposed:

- a probabilistic one, by Erdös and Rényl in 1060, who suggest a Poissonian model. Their analysis however has two limitations: it does not take into account arithmetical properties of the involved sequences, and thus does not lead to a correct modelling of sums of 2 squares; on the other hand, it is not complete and a gap in the treatment of quasi-independent has to filled.
- a purely numerical one, by Barrucand in 1968, which leads to a conjectural zero asymptotic density for sums of 3 cubes, or 4 biquadrates.
- an arithmetical one, by Hooley in 1986, which is based on the study of the square of the number of representations of an integer as a sum of s powers, and leads to a conjectural positive lower asymptotic density of such numbers, when s is at least equal to 3.

The lecture will present an approach due to B. Landreau and J-M. Deshouillers which mingles the probabalistic, arithmetic and numerical aspects: our analysis supports Hooley's conjectures and conjectually implies the existence of a positive density for the sums of s integral s powers when s is a least equal to 3. This relies, among other things, on correlation inequalities proved by B. Landreau.

ESTIMATING THE EXTREMAL INDEX UNDER A LOCAL DEPENDENCE CONDITION, BY THE RECIPROCAL OF THE AVERAGE LENGTH OF SUCCESSIVE RUNS

Luisa C. C. Duarte
University of Lisbon, PORTUGAL

Whenever a strictly stationary sequence that is strongly mixing satisfies the local dependence condition $D''(u_n)$ of Leadbetter and Nandagopalan (1989), the point process $\tilde{N}_n$ defined by the upcrossing 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 us denote by $N(n, u_n) = \sum_{i=1}^{n} I(X_i > u_n)$ the r.v. that represents the number of exceedances of u_n in a sample of size n, and by $\tilde{N}(n, u_n) = \sum_{i=1}^{n} I(X_{i-1} \leq u_n < X_i)$ 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 $n \to \infty$, it is known that the extremal index Θ verifies

$$\lim_{n\to\infty} \frac{nP[X_1 \geq u_n(\tau)]}{nP[X_0 \leq u_n(\tau) < X_1]} = \lim_{n\to\infty} \frac{E[N(n, u_n(\tau))]}{E[\tilde{N}(n, u_n(\tau))]} = \frac{1}{\Theta}.$$

Now, if the sequence of levels is such that $E[\tilde{N}(n, u_n)] \to 1$, when $n \to \infty$, 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 of dividing the sample in k_n blocks of size r_n and taking the average number of execeedances 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

$$\tilde{\tilde{\theta}}_n = \left(\frac{\sum_{i=1}^{k_n} N_i(r_n, u_{ni})}{k_n} \right)^{-1}$$

References:

Leadbetter, M. R., Nandagopalan, S. (1989). On Exceedance Point Processes for Stationary Sequences under Mild Oscillation Restrictions. *Extreme Value Theory.* (eds. J. Hüsler and R.-D. Reiss). Springer-Verlag, p.69-80.

A TRIVARIATE EXTREME VALUE DISTRIBUTION APPLIED TO FLOOD FREQUENCY ANALYSIS

Carlos A. Escalante-Sandoval
Universidad Nacional Autonoma de Mexico
Mexico, DF, MEXICO
and
José A. Raynal-Villaseñor
Universidad Autonoma de Chihuahua
Chihuahua, MEXICO

A trivariate extreme value distribution has been derived from the logistic model for multivariate extreme value distribution. The construction of its corresponding probabiliity distribution and density function is depicted. In order to obtain the parameters of such a trivariate distribution, a generalized naxumum likelihood estimation procedure allows for samples with different record lengths. Furthermore the reliability of estimated parameters of the trivariate extreme value distribution is measured through the use of relative information ratios.

The trivariate model has been applied to the data from a region in Northern Mexico with 6 gauging stations. Results produced by the proposed model have been compared with those obtained by General Extreme Value (GEV) distribution functions.

A RANDOM FIELD EXCURSION MODEL OF SALT INDUCED CONCRETE DELAMINATION

Gordon D. Fenton
Technical University of Nova Scotia
Halifax, Nova Scotia, CANADA

A random field model of salt induced concrete delamination, a problem often encountered in parking garage slabs in northern climates, is investigated and compared against some field data. The chloride ion concentration at the reinforcement depth is viewed as a random field over the two-dimensional slab area, and delamination is assumed to occur when the concentration exceeds some threshold and the growth of these regions with time is represented by a falling threshold. Simulation based excursion statistics are used to obtain the mean statistics of the delamination process using this model.

EXTREME VALUE THEORY FOR APPLICATIONS

Janos Galambos
Temple University
Philadelphia, PA

The paper is a nontechnical introduction to the conference with the aim of pointing out what extreme value theory can or cannot do in applications. It stresses that the theory is more than an approximation by the classical model, and that applications are not limited to data analysis. This is demonstrated through examples, which can be split into three groups:

i Beautiful conclusions can be drawn from extreme value theory without actual calculations (If a cancer patient, still alive, was told three years ago that his life expectancy had been one year, was the doctor wrong?);

ii Different models are needed, and are available, for data analysis (beside the widely accepted classical model, dependent models are needed for modeling life distributions with certain characteristics of failure rate; and exchangeability in Bayesian statistics);

iii Extreme value theory can be applied to further theoretical studies (large prime divisors via a dependent model, and extremes in models of physics).

ANALYTICAL AND EMPIRICAL STUDY OF THE TAILS OF PROBABILITY DISTRIBUTIONS

Janos Galambos and Nicholas Macri
Temple University
Philadelphia, PA

In an introduction the reader is reminded that a number of basic concepts of probability theory are related to the concept of tails of the underlying probability distributions, not merely extreme value theory. In the main body of the paper we discuss, on an empirical basis, the accuracy of the approximation of practical extreme value problems by the classical model. Assume that the random quanity to be investigated is the maximum Z of n random variables, where n cannot be determined. Data $X_1, X_2, \ldots, X_N$ are gathered on Z and it is assumed that n is large enough for using an asymptotic extreme value distribution II for statistical inference. The conclusion is that such modeling might lead to errors in sensitive areas. Since the unknown n can be considerably larger if the classical assumptions are replaced by some dependent model, the study enforces our long standing belief that dependent models should be used in many applied areas. The paper ends with some tail estimates for multivariate distributions based on lower dimensional marginals.

THE RATE OF CONVERGENCE OR DIVERGENCE FOR PERCENTILES OF GAMMA DISTRIBUTION AND ITS APPLICATION TO SAMPLE EXTREMES

Gaxiong Gan and Lee J. Bain
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 100 pth percentile of the gamma distribution with mean $k\theta$ and variance $k\theta^2$, and if $\lim_{n\uparrow\infty} \frac{\alpha_n}{\beta_n} = a \geq 0$, and $\lim_{n\uparrow\infty} \frac{\eta_{\alpha_n}(\theta_1,k_1)}{\eta_{\beta_n}(\theta_2,k_2)} = 0$ if $k_1 < k_2, = \infty$ if $k_2 < K_1$, and $= a^{\frac{1}{k}}\frac{\theta_1}{\theta_2}$ if $k_1 - k_2$. An example of its application to sample extremes is given.

THE SELECTION OF EXTREME EVENTS

P. W. Gerstengarbe and P. C. Werner
Potsdam-Institut für Klimafolgenforschung e. v.
Potsdam, GERMANY

The complexity of various events demands appropriate complex methods for their analysis. This means a general withdrawal from the usual linear investigation of single values and a change towards investigation and description of the interaction between a great number of quantities. The use of multivariate statistical methods seeems to be a practicable way to describe such events. In this way events can be described for a special time period and/or a limited area by a fixed number of parameters. A combined method has been developed to define extremes and to describe their temporal and spatial behaviour. This procedure is demonstrated by the analysis of the occurrence of hot and cold summers in Central Europe.

CONICAL EXTREMES OF A MULTIVARIATE SAMPLE

Alexander V. Gnedin
University of Goettingen
Goettingen, GERMANY

We introduce multivariate extremes in the direction of a given cone. Convergence results for the number of the k-th extremes are obtained for sampling from a distribution having asymptotically independent radial and sperical components and regularly varying tail of radial component.

PENULTIMATE BEHAVIOR OF THE EXTREMES

M. Ivette Gomes
University of Lisbon
Lisbon, PORTUGAL

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables, and let $\{M_n = \max_{1\leq i\leq n} X_i\}_{n\geq 1}$, be the associated sequence of maximum values. Let $F_n(x) = P[M_n \leq x], x\epsilon I\!R$, be the distribution function of $M_n, n \geq 1$. In the independent identically distributed set-up, $F_n = F^n$, and the class of possible non-degenerate limiting laws of M_n, suitably mormalized, is the class $\{G = G_\theta, \theta\epsilon I\!R\}$ of *max-stable* distributions where G_θ is called the *generalized extreme value* distribution. Looking for a unified*theory of extreme values* on θ, the characterization of the *domain of attraction* of G_θ depends on the *rate of convergence* of $F^n(a_nx + b_n)$ to $G_\theta(x)$ and on the *penultimate* behavior of $F^n(a_nx + b_n)$, in the class of *max-stable* distributions. We also study other possible penultimate approximations to $F^n(a_nx + b_n)$. These approximations will surely produce a great impact on the statistical theory of extremes, if the applied statisticians in the area feel that the class of max-stable distributions is too small, while Mejzler's class (associated to a suitable non identically distributed scheme) is too broad for practical applications. The penultimate behavior of the joint distribution of top order statistics is also considered.

References:

Anderson, C. W. (1971). *Contributions to the Asymptotic Theory of Extreme Values.* Ph.D. Thesis. University of London.

Canto e Castro, L. (1989). *Sobre a Teoria Assintótica de Extremos.* Tese de Doutoramento, DEIO, Faculdade de Ciências da Universidade deLisboa.

Cohen, J. P. (1928b). Convergence rates for the ultimate and penultimate approximations in extreme value theory. *Adv. Appl. Probab.* **14**, 833-854.

Fisher, R. A. & Tippett, L. H. C. (1928). Limiting forms of the frequency distributions of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**, 180-190.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics.* Krieger.

Gomes, M. I. (1978). *Some Probabilistic and Statistical Problems in Extreme Value Theory.* Ph.D. Thesis. Univ. Sheffield.

Gomes, M. I. (1984). Penultimate limiting forms in extreme value theory. *Ann. Inst. Statist. Math.***36**, 71-85.

Gomes, M. I. & Pestana, D. D. (1987). Nonstandard domains of attraction and rates of convergence. In L. Puri et al (eds.) *New Perspectives in Theoretical and Applied Statistics.*Wiley, 467-477.

Smith, R. L. (1987). *Approximations in extreme value theory.* Preprint, Univ. North-Carolina.

APPLICATION OF EXTREME-VALUE THEORY TO RELIABILITY PHYSICS OF ELECTRONIC PARTS AND TO ON-ORBIT SINGLE EVENT PHENOMENA

Tateo Goka
National Space Development Agency of Japan
Tokyo, JAPAN

Some models, for example weakest link model and bundle of fiber models have been applied to reliability physics in the fields of reliability and destructive engineering. Extreme-value theory (particularly doubly exponential distribution) can be applied 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. The purpose of this paper is to examine the application of extreme-value theory to accelerated life test data of tantalum capacitors for space use and to on-orbit data of single event phenomenon of memory integrated circuits in the space radiation environment. Actually, the application of extreme-value theory is compared with conventional analytical methods for Weibull and Poisson distributions to verify the effectiveness of extreme-value theory (doubly exponential distribution).

CERTAIN IDENTITIES IN EXPECTATIONS OF FUNCTIONS OF ORDER STATISTICS AND CHARACTERIZATION OF DISTRIBUTIONS

Z. Govindarajulu
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 exponential distributions. The specialized identities are used to characterize the exponential, the generalized truncated normal, the folded logistic and the uniform distributions.

A HIGHER ORDER EXPANSION FOR THE JOINT DENSITY OF THE MEAN AND MAXIMUM OF A SERIES

Amy Marie Grady
University of North Carolina
Chapel Hill, NC

In this talk, we will develop a higher order expansion for the joint density of the mean, $\overline{X}$, and the maximum, M, of a series of *iid* random variables. The results of Chow and Teugels [*Proc. Second Prague Symp. Asymptotic Statistics*; (1978): 81-92] proved that when the underlying distribution has a finite variance as well as an infinite endpoint and both $\overline{X}$ and M have a limiting distribution, then $\overline{X}$ and M are asymptotically independent. Therefore, to characterize the dependence between

$\overline{X}$ and M, we need a higher order expansion. Such an expansion is developed here from the Edgeworth expansion for the conditional distribution of $\overline{X}$ given M, along with the appropriate higher order expansions for the distribution of M. Since this result shows the dependence structure between mean and the maximum, it should be beneficial in studying climate change. Specifically, this higher order expansion will eventually be used to model the extreme temperatures across the United States in the presence of a possible changing mean temperature.

INVESTIGATING THRESHOLD SELECTION FOR TAIL ESTIMATORS BASED ON OBSERVATIONS EXCEEDING A THRESHOLD FOR THE CAUCHY DISTRIBUTION

Scott D. Grimshaw
University of Maryland
College Park, MD

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. The selection of this threshold can greatly affect the tail estimates, but only ad hoc procedures and asymptotic criteria have been presented to date. This investigation on tail estimates for the Cauchy probability model will indicate how the threshold selection affects the tail estimator properties for long tailed distributions.

COMPARISON OF EXTREME VALUE ESTIMATES BY CONDITIONAL MEAN EXCEEDANCE, PICKANDS, DE HAAN-EINMAHL-DEKKERS, AND PROBABILITY PLOT CORRELATION COEFFICIENT

John Gross, Alan Heckert, James Lechner, Stefan Leigh and **Emil Simiu**
National Institute of Standards and Technology
Gaithersburg, MD

Whether extreme wind speed distributions have finite tails is a matter of great interest in structural engineering since the answer to this question may affect significantly the estimation of safety factors for wind-sensitive structures. Failure probability estimates based on the assumption that wind speeds are described by the Gumbel distribution have been shown to be unrealistically high. This may be the case because the probability of occurrence of very high speeds, rather than being finite as in the Gumbel model, is in fact zero, since it is obvious on physical grounds that extratropical wind speeds must be bounded.

In previous work, the authors have examined the application of estimators due to Pickands and de Haan-Einmahl-Dekker to extratropical wind speed annual maxima at 120 conterminous US locations. They have proposed, following N. Bingham, a *graphical* assay of Generalized-Pareto-Distribution goodness-of-fit and domain of attraction discrimination based on serial nonlinear least-squares fitting of the GPD model to cumulative exceedance distribution function estimates for varying order statistic thresholds. It has been noted that all of the information incorporated in such plots is embodied in the cumulative mean exceedance (CME, or mean residual life (MRL) of reliability theory) function. Application of these graphical techniques to the assay of domain of attraction for wind speeds led to a tentative conclusion of compatibility of the annual wind speed maxima with a (tail-limited) Weibull model.

This work has been extended by simulations designed to test the relative efficiencies of four estimators: the CME, modified Pickands, de Haan-Einmahl-Dekkers, and probability plot correlation. For sets of simulated extreme data with statistics mimicking those of actual extreme wind speeds, each of the four estimation procedures was applied, for the epochal approach and the peaks over threshold approach. According to the simulation results for the peaks over threshold approach, the CME procedure is optimal in terms of minimizing mean square error. In the epochal approach, the optimal procedure appears to depend upon the ratio of mean recurrence interval to length of record.

EXTREME VALUE ESTIMATION - SEMIPARAMETRIC APPROACH

Laurens de Haan
Erasmus University
Rotterdam, THE NETHERLANDS

Estimation problems are considered connected with extreme values when observations are taken from a probability distribution in the domain of attraction of an extreme-value distribution. First, a semi-parametric estimator for gamma, the main parameter of the extreme-value distribution, is developed. Next an estimator for extreme quantiles will be discussed. This involves semiparametric estimators of scale and location. A closely related problem is that of estimating the endpoint of the distributioon when gamma < 0. Applications of well-known results will be given to water level data and to maximal age of humans.

MULTIVARIATE TAIL AND QUANTILE ESTIMATION

Laurens de Haan
Erasmus University
Rotterdam, THE NETHERLANDS

We consider the problem of estimating exceedance probabilities of a high level and also that of estimating high quantiles, when observations are taken from a probability distribution in the domain of attraction of some multivariate extreme-value distribution. First the structure of the limit distribution is estimated (joint work with S. Resnick). This estimate is then used as a stepping stone for the estimation of high quantiles (joint work with X. Huang) and exceedance probabilities of a high level. Confidence strips are provided.

A PROBABILITY MODEL FOR PREDICTING CORROSION AND CORROSION FATIGUE LIFE OF ALLUMINUM ALLOYS*

D. G. Harlow and **Robert P. Wei**
Lehigh University
Bethlehem, PA

Localized corrosion (in the form of pitting and exfoliation in aluminum alloys) and corrosion fatigue crack initiation and growth are recognized as degradation mechanisms that affect the durability and integrity of commerical transport aircraft. A quantitative methodology for defining suitable inspection intervals and mandating repairs is needed for effective management of the aging fleet of commericial aircraft. More importantly, a comparable methodology is needed for estimating the life of new airframe components and structures.

In this effort, a mechanistically based probability approach of life prediction for pitting corrosion and corrosion fatigue crack growth in aluminum alloys in aqueous environments is described. The differences between the statistical and mechanistic approaches to life prediction are addressed. The need for adopting a mechanistically based approach is demonstrated, and the process for developing this type of model is described. The model is developed from the worst flaw from which pitting corrosion leads to corrosion fatigue. The probability considerations are part of the initial stages of the model construction, and they are integrated throughout the process. Along with being able to make predictions well beyond the range of available data, the model can account for the contributions of each random variable, and thus, provide an assessment of the significance of each. It should be noted that the mechanistic model and the probability functions were assumed for this work, that is, they were not based on detailed experiments.

*This research was supported by the U.S. Department of Transportation under Contract DTRS57-92-Q-80189.

ESTIMATING QUANTILES FOR A TYPE III DOMAIN OF ATTRACTION BASED ON THE k LARGEST OBSERVATIONS

A. M. Hasofer
The University of New South Wales, AUSTRALIA
and
Z. Wang
University of Western Sydney (Nepean), AUSTRALIA

A method of estimating the high quantiles of a distribution belonging to the domain of attraction of type III extreme value distributions (reversel 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.

BAYESIAN FORECASTING OF RECORD VALUES IN AN EXCHANGEABLE SEQUENCE

Bruce M. Hill
University of Michigan
Ann Arbor, MI

This article develops new theory and methodology for the forecasting of record values in an exchangeable sequence of random variables. The Hill (1975) tail index estimator for long-tailed distributions is modified so as to be appropriate for prediction of future variables. Some basic and subtle issues regarding the use of finite, versus infinite, idealized models are discussed. It is shown that the standard idealized long-tailed model with tail index $\alpha \leq 2$ can lead to unrealistic predictions if the observable data are assumed to be unbounded. However, if the model is instead viewed as valid only for some appropriate finite domain, then it is compatible with, and leads to a sharper version of sensible methods for prediction. In particular, the prediction of the next record value is then at most a few multiples of the current record. It is argued that there is no more reason to eschew posterior expectations for forecasting in the context of long-tailed distributions than to do so in any other context, such as in the many applications where expectations are routinely used for scientific inference and decision-making.

EXTREME VALUE ESTIMATION APPLIED TO AERSOL SIZE DISTRIBUTIONS AND RELATED ENVIRONMENTAL PROBLEMS

Philip K. Hopke
Clarkson University
Potsdam, NY
and
Pentti Paatero
University of Helsinki
Helsinki, FINLAND

A significant problem in studies of airborne particles is the estimation of the size distribution of the aersol from a set of measurements of the ability of particles to penetrate through some type of separation device (diffusion battery, cascade impactor, etc.). In these systems, the problem is to estimate more points in the distribution than measurements. Thus, in this underdetermined system, a unique solution is not possible. In addition to being underdefined, the problem also contains considerable collinearity so that the problem is also ill-conditioned. A variety of approaches have been developed to obtain a solution including non-linear, iterative least-squares fitting, expectation-maximization, and regularization procedures. In these approaches, a solution within the feasible solution space is obtained, but there is no certainly that the correct solution has been obtained. In the case of the regularization approaches, it is certain that the solution obtained is biased from the

true solution, but the degree of bias cannot be assessed. Therefore, an alternative approach has been adopted in which *extreme value estimation* is used to obtain the set of bounds on the feasible region. In this way, a particular solution is not found, but the limits within which the solution must lie are determined. It is often of interest to obtain the value of a function of the size distribution such as the fraction of particles deposited in the respiratory tract or the dose to particular cells at risk in the bronchial epithellum arising from the deposition of radon decay products on the walls of the bronchus. Particle deposition is particle size dependent, and thus, a bound on the deposited fraction or dose can be obtained by integrating the bounds of the function of the particle size distribution. Examples of this approach to assessment of exposure and dose to radon progeny in indoor atmospheres will be presented.

EXTREMES OF NONSTATIONARY SEQUENCES

Juerg Huesler
IMS
University of Bern
Bern, SWITZERLAND

We review the properties of extreme values and of exceedances in nonstationary sequences. In particular we focus on the questions of dependence and influence of nonstationarity by analyzing limiting laws and limiting point processes. Comparing these results with the related results in the stationary case, we deduce some facts on the robustness of the stationary theory.

LEVERAGE, SAFETY FIRST AND THE SYMMETRY OF EXTREMAL STOCK RETURNS

Dennis W. Jansen
Texas A& M University
College Station, TX
and
Kees Koedijk
University of Limburg
BELGIUM
and
Casper G. de Vries
Tinbergen Institute
BELGIUM

Roy's safety first criterion for portfolio selection focuses on minimizing the probability of a disastrously low return. We investigate the theoretical and empirical implications of leverage on this portfolio selection rule. The (a)symmetry of extremal returns is investigated by means of extreme value theory.

APPLICATION OF AN EMPIRICAL EXTREME VALUE DISTRIBUTION TO LOAD MODELS

Jun Kanda
University of Tokyo
Tokyo, JAPAN

An empirical extreme value distribution with lower and upper bounds proposed by the author is applied to represent probability distribution models for maximum load intensities of the earthquake ground motion, the wind speed and the live load in supermarkets. The characteristics of the distribution are reviewed with contrasts to three well-known types of extreme value distributions, i.e. the Gumbel, Fréchet and Weibull distributions.
One of the difficulties in estimating parameters can be pointed out, that is to determine the upper bound value. Nevertheless, the annual maximum earthquake ground motion model is considerably improved by applying the proposed distribution. Possible improvements to the annual maximum wind speed model are also pointed out. Extraordinary live load models are simulated by assuming crowding situations. The proposed distribution is also a fair candidate for the live load extremes.

The applicability is discussed in terms of statistical values representing extreme load models. The significance of the existence of an upper bound is stressed for load models with a large variability or in a range of very low probability of exceedance.

THE REFLECTED WEIBULL DISTRIBUTION FOR DATA ANALYSIS OF FLEXURAL STRENGTH OF HIGH PERFORMANCE CERAMICS

Shigeo Kase
The University of the Air
Chiba City, JAPAN
and
Tateo Goka
High-Reliability Component Corporation
Tokyo, JAPAN

The application of the Weibull distribution to analysis of flexural strengths of ceramics prevails all over the world with the understanding that its shape parameter m determines the variability of the material. The present study shows another interpretation of the shape parameter as a measure of how well the Weibull distribution approximates the doubly expomential distribution (or the third asymptotic distribution of extreme-value).

Apart from the shape parameter, ceramics researchers are not necessarily satisfied with the prevailing opinion on the statistical analysis of flexural strengths because the Weibull distribution gives a poor fit to data in the region of low strength which is the most important region for structural materials. The goodness of fit in the tail is also reviewed in this paper for the Weibull and the doubly exponential distributions.

The conclusion is that the doubly exponential distribution is preferable to the Weibull distribution for analysis of flexural strengths of high performance ceramics, e.g. silicon nitride.

ON Q-ADDITIVE FUNCTIONS

Imre Kátai
Loránd Eötvös University
Budapest, HUNGARY

Let $\mathcal{A}_q, \mathcal{M}_q$ be the classes of q-additive and q-multiplicative functions, where $q \geq 2, q \epsilon \mathcal{N}$. For $f \epsilon \mathcal{A}_q$ let $\xi_0^*, \xi_0, \xi_1, \xi_2, \ldots$ be independent random variables, $P(\xi_j = f(aq^j)) = 1/q$ if $a \epsilon \boldsymbol{A} = \{0, 1, \ldots, q-1\}, j = 0, 1, 2, \ldots$. $P(\xi_0^* = f(a)) = \frac{1}{\varphi(q)}$ if $a \epsilon \boldsymbol{A}, (a, q) = 1, N = N(x) = [\frac{\log x}{\log q}], \eta_x = \sum_{\nu=0}^N \xi_\nu, \eta_x^* = \xi_0^* + \sum_{\nu=1}^N \xi_\nu$. By using the method of trigonometric sums, we prove the following theorem: Let $k \geq 2, x \geq 2$. Then

$$\frac{1}{x} \sum_{n \leq x} (f(n^k) - M(\eta_{x^k}))^2 < c(q,k) D^2(\eta_{x_k}) \tag{1}$$

$$\frac{1}{\pi(x)} \sum_{q < p \leq x} (f(p^k) - M(\eta_{x^k}^*))^2 < c(q,k) D^2(\eta_{x^k}^*) \tag{2}$$

where p runs over the primes. $c(q,k)$ is a suitable constant. Some consequences of this theorem will be given.

POISSON APPROXIMATION OF POINT PROCESSES OF EXCEEDANCES UNDER VON MISES CONDITIONS

Edgar Kaufmann
Universitaet Siegen
Siegen, GERMANY

Bounds on the Hellinger distance between certain truncated empirical processes and certain Poisson processes are constructed. These bounds depend, roughly speaking, on the rate at which a fairly general von Mises condition holds. Applying these results, approximations of the k largest order statistics w.r.t. the variational distances are also established.

ON THE REQUIREMENTS FOR A REASONABLE EXTREME VALUE PREDICTION OF MAXIMUM PITS ON HOT-WATER-SUPPLY COPPER TUBING

Shigeru Komukai and Komei Kasahara
Tokyo Gas Company, Ltd.
Fundamental Technology Research Laboratory
Tokyo, JAPAN

Application of extreme value statistics to the problem of Type-II pits growth prediction on hot-water-supply copper tubing is described. A recommendation is suggested for optimum combinations of the number and the size of unit samples required for reasonable extreme value predictions.

ON THE LARGEST PRIME DIVISORS OF AN INTEGER

Jean-Marie De Koninck
Laval University
Quebec, CANADA

Let $\Omega(n) = \sum_{p^\alpha \| n} \alpha$ and, for each integer n such that $\Omega(n) \geq k$, let $P_k(n)$ denote its k^{th} largest prime factor. Further, given a set of primes Q of positive density $\delta < 1$ satisfying a certain regularity condition, define $P(n, Q)$ as the largest divisor of n belonging to Q, assuming that $P(n, Q) = +\infty$ if no such prime factor exists. We provide estimates of $\sum_{2^k \leq n \leq x, \Omega(n) \geq k} 1/P_k(n)$, for $k \geq 2$, and of $\sum_{n \leq x} 1/P(n, Q)$. We also study the median value of the function $P(n, Q)$ and that of the function $P_k(n)$ for each $k \geq 1$.

RATE OF CONVERGENCE OF THE NUMBER OF EXCEEDANCES OF NONSTATIONARY NORMAL SEQUENCES

Marie-France Kratz
University of Paris VI
Paris, FRANCE

It is known that the number of exceedances is asymptotically a Poisson random variable, under certain restrictions. We analyze the rate of convergence to a Poisson limit and extend the result known in the stationary case to nonstationary normal sequences by using the Stein-Chen method. In addition, we consider the cases of exceedances of a constant level as well as of a nonconstant level.

EXTREME VALUES OF MONOTONIC FUNCTIONS AND EVALUATION OF CATASTROPHIC FLOOD LOSS

James H. Lambert, **Duan Li** and Yacov Y. Haimes
University of Virginia
Charlottesville, VA

The model of a univariate monotonic function is often useful in engineering risk assessment, for example in relating the magnitude of flood discharge to the consequent economic losses. An engineering approach to study the transformation of the domain of attraction of an underlying random variable by a monotonic function is developed in this paper. Using von Mises criteria, sufficient conditions are derived to find the domain of attraction of the transformed variable for situations where only uncertain knowledge concerning the monotonic function and the underlying variable 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.

EXACT SOLUTION TO AN INTERACTING EXTREME VALUE PROBLEM: THE PURE FLAW MODEL

P. L. Leath
Rutgers University
Piscataway, NJ
and
P. M. Duxbury
Michigan State University
East Lansing, MI

Simple models play a key role in the microstructural analysis of mechanical failure in composites and other materials with complex and often disordered microstructures. Although equal load sharing models are amenable to rigorous statistical analysis, problems with local load enhancements near failed regions of the material have so far resisted exact analysis. Here we show for the first time, that some of the simpler of these *local load sharing models* can be solved exactly using a Markov matrix method. For diluted fiber bundles with local load sharing, it is possible to find a compact expression for the characteristic equation of the Markov matrix, and from it obtain an asymptotic expansion for the largest eigenvalue of the Markov matrix. This in turn gives the asymptotic behavior of the size effect and statistics of the fiber bundle models. We summarize these results, and show that the important features of the exact result can be obtained from a simple scaling analysis we had developed previously. We also compare the statistics of fracture with the appropriate limiting extreme failure distribution (a Gumbel distribution), and, as previously noted by Harlow and Phenonix, note that the Gumbel distribution performs quite poorly in this problem. We comment on the physical origin of this discrepancy. Finally, we have made significant progress in analysing the failure of fiber bundles with members whose strengths are drawn from continuous distributions, and we give a preliminary report of these results.

PERFORMANCE COMPARISON BETWEEN AN EXTREME VALUE STATISTICAL MODEL, A DETERMINISTIC MODEL, AND AN ARTIFICAL NEURAL NETWORK MODEL FOR PREDICTING DAMAGE FUNCTIONS FOR PITTING CORROSION

M. Urquidi-Macdonald and D. D. Macdonald
The Pennsylvania State University
University Park, PA

Various attemps have been made to develop models for predicting the development of damage in metals and alloys in aqueous environments to pitting corrosion. Generally speaking, these models may be divided into two classes: empirical and deterministic. By far the most prominent empirical method employs extreme value statistics using a Weibull distribution function, which is strictly valid only for static (time-independent) systems. Because no *a priori* method exists for predicting the time-dependencies of the shape and location parameters of the distribution, they must be determined experimentally. In the few cases, where sufficient data exist, no agreement exists as to the functional form of these dependencies, so that, strictly reliable predictions can only be made within the boundaries of the calibrating database. The apparent success of statistical models of this type, in representing the (time) development of damage, can be traced to the extreme mathematical flexibility of the distribution function.

At the other extreme of the empirical/deterministic spectrum, are models based on perceived mechanisms for the nucleation and growth of damage due to pitting corrosion. Perhaps the most highly developed algorithm of this type is the combination of the Point Defect Model (PDM) for describing the nucleation of damage and the Coupled Envirnment Model (CEM) for modeling the growth. These models are deterministic, in the sense that they conform with known physical laws (e.g. the conservation of charge) and they yield definite (and sometimes analytical) relationships between the damage function and the model inputs. However, they are based on preconceived notions of the damaging mechanisms which may or may not be correct.

More recently, Artificial Neutral Networks (ANNs), a non-deterministic type of model, have been developed to describe the progression of damage due to pitting corrosion. ANN contains no proconceived mechanishm for the development of damage, but it may be "trained" by exposure to reliable data and knowledge. No preconceived distribution function is employed, although a choice is made of 1) the addition functionality ("knowledge") based on the data, 2) the form of the transfer function, and 3) the number of independent variables.

In this paper we compare and contrast the three approaches outlined above for describing the progression of damage due to pitting corrosion in metal and alloy systems in contact with aqueous systems. Our goal is to illustrate the advantages and disadvantages of each approach with reference to real experimental data in order that the most reliable methods may be employed in future algorithms for predicting pitting damage functions for engineering structures.

TECHNIQUES USED TO DETERMINE EXTREME WAVE HEIGHTS FROM THE NESS DATA SET

Marc A. Maes and George Z. Gu
Mobil Research & Development
Dallas, TX

The objective of the present paper is to explore the application of different extreme value procedures to a selected grid point in the NESS (North European Storm Study) data set. The NESS project (Grant et al., 1993) was funded by eleven participants and conducted by a consortium of five research institutes in Europe. for the purpose of the present benchmark study, each of the participants was invited to submit estimates of a significant wave height with a high return period at a given grid point in the North Sea, together with a detailed description and a rationale of the approach used in deriving extreme wave heights. Submissions from five industry groups were received, compared, and benchmarked against one another.

The NESS users group built up considerable experience in the field of extermal value (E) applications. Valuable input was received from recognized experts in EV theory. It should therefore be made clear that the objective of the present comparison is not to review/criticize the use of particular extreme value techniques. Rather, the objective is (1) to reflect upon the diversity of the modeling assumptions and the procedures used to determine extreme wave heights, (2) to report on how the different groups set out to deal with difficult issues such as data reduction, statistical and parameter uncertainty, hindcast model uncertainty, and the consideration of measured data, and (3) to seek constructive guidance in this area from the extreme value specialists present at the NIST/Temple University EV Conference.

EXTREME VALUE DISTRIBUTIONS FOR LINEAR AND NON-LINEAR SYSTEMS AND APPLICATIONS TO MARINE STRUCTURES

A. F. Mansour
University of California
Berkeley, CA

Prediction of extreme wave loads acting on a marine structure is an essential step in its design. The structure may respond linearly or non-linearly to extreme wave conditions depending on its flexibility and type. This paper presents extreme value distributions that can be used for linear systems such as tankers and non-linear systems such as jack-up platforms. The results of application to these structures are compared and discussed.

ON TESTING THE EXTREME VALUE INDEX

Frank Marohn
Katholische Universitat Eichstatt
Eichstatt, Germany

Consider the statistical experiment $E = \{H_\beta : \beta \in I\!R\}$, where H_β denotes the generalized Pareto distribution given by the von Mises parametrization, and H_0 is the standard exponential distribution. We consider the two–sided testing problem H_0 against $\{H_\beta : \beta \neq 0\}$. For that testing problem an asymptotic optimal test sequence is proposed. As a main tool we show that SES is differentiable in quadratic mean (DQM) at $\beta = 0$, which is a crucial condition in the asymptotic setting. Moreover, we consider the extreme value distributions G_β, $\beta \in I\!R$, which are the most important ones in the neighborhood of the generalized Pareto distributions. We treat the problem of testing the Gumbel (G_0) against the Fréchet ($G_\beta : \beta > 0$) and the Weibull ($G_\beta : \beta < 0$). It turns out that the differentiabilty in quadratic mean carries over to the extreme value distributions. Monte-Carlo simulations visualize the results.

INCLUSION RATING BY STATISTICS OF EXTREMES AND ITS APPLICATION TO FATIGUE STRENGTH PREDICTION AND QUALITY CONTROL OF MATERIALS

Y. Murakami
Faculty of Engineering Kyushu University
Fukuoka, JAPAN

The inclusion rating method by statistics of extremes (IRMSE) using $\sqrt{area}$ of inclusions as the size parameter enables one to discriminate distinctly current super-clean steels. Moreover, IRMSE enables one to predict the size ($\sqrt{area_{max}}$) of maximum inclusions to be contained in larger domains than the inspection domain. The statistical distribution of $\sqrt{area_{max}}$ can be used for the quality control of materials and for the prediction of a scatter band of fatigue strength. Practical procedures for inclusion rating and prediction of a scatter band of fatigue strength are shown.

RECORD STATISTICS FROM POINT PROCESS MODELS

H. N. Nagaraja
The Ohio State University
Columbus, OH

Let $X_1, X_2, ...$ be a sequence of independent random variables with continuous distribution functions, being observed at successive time points determined by an independent point process P defined on positive reals. We say X_j is an upper record value if $X_j = max(X_1, ..., X_j)$. Let $\{R_j, j \geq 1\}$ be the sequence of upper records and let $\{T_j, j \geq 1\}$ be the times at which these records are observed. We describe the distribution of these and related record statistics under

various assumptions on P and the distribution functions of X_j. We discuss some applications of the results to situations such as the shock models and the best choice problem.

THE MULTIVARIATE EXTREMAL INDEX

S. Nandagopalan Nandu
Colorado State University
Fort Collins, CO

We consider the limiting distribution of the vector of componentwise maxima associated with a multivariate stationary sequence. This limiting distribution is influenced by the two dependence structures present: (a) the dependence between the components of each observation vector, and (b) the dependence between successive observation vectors over time. The interplay between the two can be quantified by the "multivariate extremal index". Some simple properties of the multivariate extremal index are explored and some examples are presented.

IS THERE A KHINCHINE-TYPE DECOMPOSITION THEOREM FOR EXTREME VALUES?

E. Pancheva
Institute of Mathematics
Sofia, BULGARIA

Traditionally, extreme value theory has been treated in the multiplicative semigroup $\mathcal{P}$ of d.f's on $\mathbb{R}$ endowed with the Levy metric L (which metrizes the weak convergence in $\mathcal{P}$). Unfortunately, in $(\mathcal{P}, L, .)$ there is no Khinchine-type decomposition theorem, as is shown in Zempleni (1991). We choose another approach to extreme values, namely, we consider the multiplicative semigroup $\mathcal{F}$ of distributions on $\bar{\mathbb{R}}^d = [-\infty, \infty)^d$, introduce in it a metric $\mathcal{L}$, corresponding to the weak convergence in $\mathcal{F}$, and show that in the structure $(\mathcal{F}, \mathcal{L}, .)$ there are analogues of the well known first and second Khinchine's theorems.

SECOND ORDER BEHAVIOUR FOR DOMAINS OF ATTRACTION AND THE BIAS OF THE GENERALIZED PICKANDS ESTIMATOR

M. Teresa Themido Pereira
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 results are presented for measurable domains of attraction, and the particular important case of Pickands' estimator for the main parameter of an extreme value distribution is proved. Moreover, under quite general conditions

that include second order behaviour and differentiability, the asymptotic normality of the estimator is proved and the 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.

References

Aczél, J. (1966) - Lectures on Functional Equations and their Applications. Academis Press, N.Y.

Cooil, B. (1985) - *Limiting multivariate distributions of intermediate order statistics.* Ann. Probab. **13** 469-477.

Cooil, B. (1988) - *When are intermediate processes of the same stochstic order?* Statist. & Probab. Letters **6** 159-162.

De Haan, L. (1970) - On Regular Variation and its Application to the Weak Convergence of Sample Extremes. Math. Centre Tracts 32. Mathematical Centre, Amsterdam.

De Haan, L. (1984) - *Slow variation and charactization of domains of attraction.* In Statistical Extremes and Applications (J. Tiago de Oliveira, ed¿) 31-38. Reidel, Dordrecht.

De Haan, L. and Dekkers, A. L. M. (1989) *On the estimation of the extreme-value index and large quantile estimation.* Ann. Statist.**17** 1795-1832.

Dekker, A. L. M. (1991) - On Extreme-Value Estimation. Ph.D. thesis, Erasmus University, Rotterdam.

Galambos, J. (1987) - The Asymptotic Theory of Extreme Order Statistics. Robert E. Krieger Publ. Comp., Malabar, Florida.

Geluk, J. L. and De Haan, L. (1987) - Regular Variation, Extensions and Tauberian Theorems. Math. Centre Tracts 40. Centre for Mathematics and Computer Science, Amsterdam.

Gnedenko, B. V. (1943) - *Sur la distribution limite du terme maximum d'une séries aléatoire.* Ann. of Math. **44** 423-453.

Hall, P. (1982) - On some simple estimates of an exponent of regular variation. J. R. Statist. Soc. B **44** 37-42.

Omey, E. and Willekens, E. (1987) - Π-*variation with remainder*, J. London Math. Soc. (2) **37** 105-118.

Pickands III, J. (1975) - *Statistical inference using extreme order statistics.* Ann. Statist. **3** 119-131.

Pickands III, J. (1986) - *The continuous and differentiable domains of attraction of the extreme value distributions.* Ann. Probab. **14** 996-1004.

Resnick, S. I. (1987) - Extreme Values, Regular Variation, and Point Processes. Springer-Verlag, N.Y.

Tiago de Oliveira, J. (1991) - *Extreme values and tail estimation.* Proc. Order Statist. Nonparametrics; Theory Appl., Alexandria.

Von Mises, R. (1936) - *La distribution de la plus grande de n valeurs.* Reprinted in Selected Papers II. Amer. Math. Soc., Providence , R.I. (1954) 271-294.

EXTREME VALUE THEORY FOR A CLASS OF MARKOV CHAINS ON $\mathcal{R}^d$

Roland Perfekt
University of Lund, SWEDEN

Let $\{\boldsymbol{\xi}_n = (\xi_{n,1}, \ldots, \xi_{n,d})\}_{n\geq 0}$ be a stationary Markov chain on $(E, \mathcal{E})$ with transition kernel K. Here $E \subseteq \mathcal{R}^d$ for some $d \geq 1$ and extremal properties of $\{\boldsymbol{\xi}_n\}$ and specifically study the limiting behaviour of $\boldsymbol{M}_n$, the vector of componentwise maxima:

$$\boldsymbol{M}_n = (M_{n,1}, \ldots, M_{n,d}) = \left(\bigvee_{i=1}^{n} \xi_{i,1}, \ldots, \bigvee_{i=1}^{n} \xi_{i,d} \right).$$

Suppose that $\{\boldsymbol{\xi}_n\}$ satisfies a weak long range dependence restriction and that there exists a family of normalizing sequences in $\mathcal{R}^d$,

$$\{\boldsymbol{u}_n(\boldsymbol{x}) = (u_{n,1}(x_1), \ldots, u_{n,d}(x_d))\}_{n\geq 1}, \quad \boldsymbol{x} = (x_1, \ldots, x_d) \in (0, \infty)^d,$$

such that

(i) $n(1 - P\{\xi_{0,i} \leq u_{n,i}(x)\}) \to x^{-1}$ for each

(ii) $n(1 - P\{\boldsymbol{\xi}_0 \leq \boldsymbol{u}_n(\boldsymbol{x})\}) \to -\log G(\boldsymbol{x})$, $\boldsymbol{x} \in (0,\infty)^d$,
for some function $G : (0,\infty)^d \mapsto [0,1]$.

If (i) and (ii) hold then G must necessarily be a

$$G(\boldsymbol{x}) = \exp\left\{ -\mu([0,\infty]^d \setminus [\boldsymbol{0}, \boldsymbol{x}]) \right\}, \quad \boldsymbol{x} \in (0,\infty)^d,$$

for some exponentmeasure μ on $[0,\infty]^d \setminus \{\boldsymbol{0}\}$ with $\mu((x,\infty] \times [0,\infty] \times \cdots \times [0,\infty]) = \ldots = \mu([0,\infty] \times \cdots \times [0,\infty] \times (x,\infty]) = x^{-1}$, $x \in (0,\infty)$; see Resnick (1987).

For such a chain we can show, under suitable assumptions on the transition probabilities K, that the extremal properties are determined by another Markov chain, $\{\boldsymbol{\xi}_n^*\}$ say, with state space $E^* \subseteq [0,\infty)^d$ such that $\boldsymbol{x} \in E^*$ implies Here K^* is given by the limiting behaviour of $K(\boldsymbol{u}, \cdot\,)$ for, in a suitable sense, large $\boldsymbol{u}$ and satisfies $\cdot\,)$, $t \in (0,\infty)$, $\boldsymbol{x} \in E^*$. In this case

$$P\{\boldsymbol{M}_n \leq \boldsymbol{u}_n(\boldsymbol{x})\} \to \exp\left\{ - \int_{E^* \setminus [\boldsymbol{0}, \boldsymbol{x}]} P\{\boldsymbol{\xi}_i^* \leq \boldsymbol{x},\ i \geq 1 | \boldsymbol{\xi}_0^* = \boldsymbol{y}\}\, \mu(d\boldsymbol{y}) \right\},$$

where μ is the exponentmeasure associated with G. Also, the limit distribution of $\boldsymbol{M}_n$ is max-stable although the marginals are not necessarily non-degenerate. This extends results in Smith (1992) and Perfekt (1991) and the multivariate setting of course allows treatment of extremes of real-valued Markov chains of higher order then one as well.

Main references:

Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *J. Multivariate Anal.* **29** 274-291.

Hüsler, J. (1990). Multivariate extreme values in stationary random sequences. *Stochastic Proc. Appl.* **35** 99-108.
Nandagopalan, S. (1990). Multivariate extremes and estimation of the extremal index. *Technical Report 315, Center for Stochastic Processes, Univ. North Carolina.*
Perfekt, R. (1991). Extremal behaviour of stationary Markov chains. *Technical Report 353, Center for Stochastic Processes, Univ. North Carolina.*
Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York.
Smith, R.L. (1992). The extremal index for a Markov chain. *J. Appl. Probab.* **29** 37-45.

BAYES QUANTILE PARAMETER ESTIMATION AND THRESHOLD SELECTION FOR THE GENERALIZED PARETO FAMILY

James Pickands III
University of Pennsylvania
Philadelphia, PA

There is always a tradeoff between computational and statistical efficiency. The phenomenal recent advances in computing have shifted the advantage towards Bayesian methods of estimation. We propose an interactive program for analyzing observations above a high threshold. The first phase of the analysis is the "identification" phase. Using parameter estimators which are not too efficient but easy to compute, we perform many iterations to find the empirical optimal combination choice of threshold and choice of monotone increasing function transformation. We review existing methods of estimation. The second phase is the "estimation" phase. We examine the use of Bayes estimators for this phase. A few close choices may still be examined. We also suggest a method of finding confidence regions.

ESTIMATION OF THE SIZE DISTRIBUTIONS OF THE FATIGUE CRACK INITIATING PORES IN ALUMINUM ALLOYS

Marek A. Przystupa
University of California
Los Angeles, CA

Most of the aluminum alloys have porosities and their pore size distributions can be characterized, for instance, by measuring the pores on the alloy plane sections. Distributions obtained in this way are however not the same as distributions of fatigue cracks initiating pores, as the latter are the extreme values of the former. To estimate the size distribution of fatigue cracks from the distribution of pore sizes measured on the plane sections one needs to make several assumptions about the pore spatial distribution, stress distribution, etc. Such predictions can then be verified by comparing estimates with the true distributions obtained by measuring the crack initiating pores

on fractured surfaces of fatigued samples. The analysis outlined above has recently been performed for the aluminum 7050 alloy: the results and questions raised will be discussed during this talk.

EXTREME VALUE THEORY AND ITS APPLICATIONS IN MICROSTRUCTURAL SCIENCES

Krishna Rajan
Rensselaer Polytechnic Institute
Troy, NY

The quantitative characterization of microstructure usually involves the determination of an "average" behavior of a microstructural parameter. These parameters may measure dimensional scales such as grain size or the frequency of a particular defect density . Most models of materials behavior are associated with this value. As such, most statistically oriented measurements in microstructural characterization of materials are only concerned with this mean. In this presentation we provide experimental examples where significant deviations from the average microstructure exist in pertinent materials physics. Some of the examples include: the deviations from classical Mullins-von Neumann law for two-dimensional grain growth; the changes occurring in the extreme values of grain size distributions associated with significant changes in materials properties; the evolution in the extreme value of grain boundary misorientation distributions reflecting variations in an otherwise "constant" microstructure; and the role of extreme values of pore size distributions in synthetic membranes. It is pointed out that the importance of extreme value theory arises in these examples due to the unique nature of the experiments as well as the increased sensitivity of data analysis. Details of the experimental aspects of our studies which help to emphasize the importance of extreme value theory and its applications are outlined.

A BIVARIATE EXTREME VALUE DISTRIBUTION APPLIED TO DROUGHT FREQUENCY ANALYSIS

José A. Raynal-Villaseñor
Universidad Autonoma de Chihuahua
Chihuahua, MEXICO

Drought frequency analysis has been carried out using widely-known univariate probability distributions for the minima. A few examples are the Three-Parameter Log-Normal, Pearson Type III and the Extreme Value Type I (Gumbel) distributions.

The methods for estimating the parameters of such distributions vary from the method of moments to newer methods such as the probability weighted moments and maximum entropy methods.

Multivariate extreme value models for flood and drought frequency analysis have been explored with application to hydrological problems. The bivariate extreme value model of the logistic type for the minima is explored in this paper. The characteristics, properties and construction of the model are reported in the paper with application in drought frequency analysis. The method of estimation of parameters is the method of maximum likelihood.

Several examples are shown in the paper, which is illustrated with several graphs to show the capabilities of modeling with the proposed model.

USER FRIENDLY PERSONAL COMPUTER PACKAGES IN FLOOD AND DROUGHT FREQUENCY ANALYSES: THE FLODRO PACKAGE

José A. Raynal-Villaseñor
Universidad Autonoma de Chihuahua
Chihuahua, MEXICO
and
Carlos A. Escalante-Sandoval
Universidad Nacional Autonoma de Mexico
Mexico, D.F., MEXICO

Personal computers have dramatically changed studies related to hydological problems. The increased capabilities of the current generation of personal computers allow the hydrological engineer to explore options that a few years ago were unthinkable without a mainframe computer. The user-friendly personal computer package FLODRO was developed to the advanage of these advances. It has two components:

a) Program FLOOD which performs flood frequency analysis by using eight typical distribution functions: Normal, Two and Three Parameter Log-Normal, Two and Three Parameter Gamma, Log-Pearson Type III, Extreme Value Type I and General Extreme Value.

b) Program DROUGHT which performs drought frequency analysis by using four typical distribution functions: Three Parameter Log-Normal, Pearson Type III, Extreme Value Types I and III and General Extreme Value.

The characteristics, properties and construction of the program FLOOD and DROUGHT of computer package FLODRO are reported with application to flood and drought frequency analyses. The methods of estimation of parameters are moments and maximum likelihood for flood analysis, and moments for drought analysis. Several applications are shown in the paper, which is illustrated with several graphs to show the capabilities of modeling with the computer model FLODRO.

XTREMES: EXTREME VALUE ANALYSIS AND ROBUSTNESS

R.-D. Reiss, S. Haßmann and M. Thomas
Universität Siegen
Siegen, GERMANY

The software package XTREMES provides a graphical representation of curves such as generalized Pareto, extreme value or Cauchy-normal densities and q.f.'s. Secondly, data sets may be generated according to these densities and in non-i.i.d. models. Finally, the data may be inspected by means of parametric and nonparametric methods. Implemented estimators are kernal densities, empirical q.f. and mean residual life function, Hill-, M-, mixed Pickands- and moment-estimator. An external implementation of new estimators is possible. The performance of estimators may be inspected by means of diagrams, plots and Monte Carlo simulations of the MSE and the distribution. XTREMES is menu-driven, runs on IBM-compatible PC's and comes with a user's guide manual.

EXTREME VALUE THEORY AND CHOICE MODELS

Sidney Resnick
Cornell University
Ithaca, NY
and
Rishin Roy
University of Toronto
Toronto, CANADA

A theory of choice models is presented where the choice space E is either finite or a continuum. Choice is made according to the principle of maximum utility so that at time t, choice $e \in E$ is chosen if its utility $Y_t(e)$ achieves $\vee_{s\in E} Y_t(s)$. When E is a finite set, modelling utility as an extremal process leads to a tractable theory. When E is a continuous set, say the unit square for the case that location is being chosen, we model utility as a super–extremal process $Y_t(s), t > 0, s \in E$ where for fixed t we suppose $Y_t(s), s \in E$ is an upper semi–continuous function of s. If more than one choice achieves maximum utility at time t, then we set

$$M_t = \{e \in E : Y_t(e) = \vee_{s\in E} Y_t(s)\}$$

and then the choice set at time t is a random closed set. Properties of the set valued process $\{M_t, t > 0\}$ may be studied in the case that for each t, $Y_t(\cdot)$ is max–stable.

ON THE WEAK CONVERGENCE OF CERTAIN ESTIMATORS OF STOCHASTICALLY ORDERED SURVIVAL FUNCTIONS

Javier Rojo
University of Texas
El Paso, TX

Let F and G be distribution functions with corresponding survival functions $\bar{F}$ and $\bar{G}$. When $\bar{F}$ is known and $\bar{G}(x) \geq \bar{F}(x)$ for all x, the estimator $\hat{\bar{G}}_n = \max(\bar{F}, \bar{G}_n)$, where $\bar{G}_n$ represents the empirical survival function, has been considered by Rojo and Ma (1992). In the two-sample problem, when both $\bar{F}$and $\bar{G}$ are unknown with $\bar{G}(x) \geq \bar{F}(x)$ for all x, Lo (1987)and Rojo and Ma (1992) have considered the estimators $\hat{\bar{G}}_{m,n} - \max(\bar{F}_n, \bar{G}_m)$ and $\hat{\bar{F}}_{m,n} = \min(\bar{F}_n, \bar{G}_m)$ where $\bar{F}_n$ and $\bar{G}_m$ represent the empirical survival functions based on random samples from F and G

of sizes n and m respectively. Here, the weak convergence of $\hat{\bar{G}}_n, \hat{\bar{G}}_{m,n}$and $\hat{\bar{F}}_{m,n}$ is considered. It is shown, under mild conditions on F and G, that $\hat{\bar{G}}_n, \hat{\bar{G}}_{m,n}$, and $\hat{\bar{F}}_{m,n}$ converge weakly when $\bar{G}(x) > \bar{F}(x)$ for all x. However, when $\bar{G}(x_o) = \bar{F}(x_0)$ for some x_0, weak convergence does not hold. The results yield asymptotic confidence bands for the survival functions of interest. An example in the estimation of pitting times under two different corrosive environments is used to illustrate the results.

TAIL EMPIRICAL PROCESSES FOR DEPENDENT SEQUENCES

Holger Rootzén
University of Lund
Lund, SWEDEN

In this talk we give conditions for weak convergence of the tail empirical process for stationary sequences. The main assumptions needed are absolute regularity with a specified mixing rate and assumptions on correlations of indicators. The result may be used for goodness of fit tests and to find limiting distributions for tail parameter and extremal index estimators.

NORMAL SAMPLE RANGE: ASYMPTOTIC DISTRIBUTION, APPROXIMATIONS AND POWER COMPARISONS

Andrew L. Rukhin
University of Maryland
Baltimore, MD

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.

OPTIMAL STOPPING AND EXTREME VALUE RESULTS FOR MINIMIZING THE SUM OF THE RANKS IN A MULTIPLE CRITERIA PROBLEM

Stephen M. Samuels
Purdue University
Lafayette, IN

Suppose $\{U_{i,j} : i = 1, \ldots, n;\ j = 1, \ldots, m\}$ are mn independent random variables, each uniformly distributed on the interval $(0, n)$.

Let $X_{i,j}$ be the rank of $U_{i,j}$ among $U_{1,j} \ldots U_{n,j}$ (rank 1 is the smallest), so the m vectors,

$$\{< U_{1,j} \ldots U_{n,j} >: j = 1, \ldots, m\},$$

are m independent random permutations of the integers 1 to n.

Let $Y_{i,j}$ be the *relative* rank of $U_{i,j}$ among $U_{1,j} \dots U_{i,j}$, $\mathcal{F}_k$ be the σ-algebra generated by $\{Y_{i,j} : i \le k\}$, and $\mathcal{T}_n$ be the class of stopping rules adapted to $\{\mathcal{F}_k : 1 \le k \le n\}$.

For $m \ge 2$, the limits, as $n \to \infty$, of both

$$n^{-(1-1/m)} E[\min_{\tau \in \mathcal{T}_n} \sum_{j=1}^{m} U_{\tau,j}] \quad \text{and} \quad n^{-(1-1/m)} E[\min_{\tau \in \mathcal{T}_n} \sum_{j=1}^{m} X_{\tau,j}]$$

are shown to be equal and are evaluated. Likewise, the limits of both

$$n^{-(1-1/m)} E[\min_{1 \le i \le n} \sum_{j=1}^{m} U_{i,j}] \quad \text{and} \quad n^{-(1-1/m)} E[\min_{1 \le i \le n} \sum_{j=1}^{m} X_{i,j}]$$

are shown to be equal and are evaluated. Equality of the pairs of limits fails when $m = 1$.

APPLICATIONS OF EXTREME VALUE THEORY IN CORROSION ENGINEERING

P. A. Scarf
University of Salford
Salford, U. K.
and
P. J. Laycock
U. M. I. S. T. Manchester, U. K.

In the context of corrosion engineering it is often natural to be concerned with "extreme events". This is because, firstly, it is these extreme events that often lead to failure and, secondly, it may only be possible to measure the extremes, with much of the underlying measurements by their very nature unobservable. A variety of statistical techniques, under the general heading of "extreme values", can be used to model and hence predict the statistical behaviour of extremes such as the largest pit, thinnest wall, maximum penetration or similar assessment of a corrosion phenonmenon. These techniques can be applied to measurements over a regular grid or to measurements of selected extremes. They can be adapted to accommodate all values over a selected threshold, or a selected number of the largest values, or just the single largest value. Data can come from just one coupon or several coupons, and be modelled to allow for dependence on environmental conditions, surface area examined, and the duration of exposure or of experimentation. A selection of these techniques are demonstrated on data collected in an industrial context and on data from laboratory experiments. Recent developments in the multivariate modelling of extreme values is also considered in the context of extreme pit depths and diameters.

APPLICATION OF EXTREME VALUE STATISTICS TO CORROSION

T. Shibata
Osaka University
Suita, JAPAN

SIMULATIONS FOR THE EXTREME STATISTICS

Masaaki Sibuya
Keio University
Yokohama, JAPAN

The speed of the present-day work stations and personal computers allows us to make a larger number of repetitions of simulations (Monte Carlo computations) to reduce sampling errors. To take advantage of the reduced sampling errors we now need random numbers of higher quality to reduce the numerical errors in simulations. We assume that computations with single-word precision are good enough in comparison with the sampling error.

First, in this report, recent results on uniform quasi-random number generation are reported. The Generalized Feedback Shift Register developed by Fushimi and Tezuka is highly recommended. Second, the possible errors in using the quantile function to transform uniform quasi-random numbers to random numbers of a non-uniform distribution are examined. Third, efficient and accurate methods for generating Gumbel random numbers are discussed. The exact-approximation method by Marsaglia is shown to be attractive.

MULTIVARIATE THRESHOLD METHODS

Richard L. Smith
University of North Carolina
Chapel Hill, NC

Statistical methods based on modeling the exceedances over a high threshold have emerged as one of the best ways to apply extreme value theory in practice. However, work on multivariate extensions - when there are several processes contributing to the extremes and we want some way of jointly modeling them - is still in a more primitive state. Nevertheless, the expansion of research on multivariate extreme value theory over the past several years (due especially to the work of J. Tawn and S. Coles) has now created a number of possibilities for developing multivariate threshold methods. In this talk, I shall briefly review both univariate threshold methods and multivariate extreme value theory, and will then discuss ways of putting them together. A novel point of the theory is that it suggests new ways to implement univariate threshold methods: it is possible to create models explicitly incorporating dependence by using multivariate techniques to model the joint distribution of exceedances in a time series. The talk will be illustrated with some examples taken from ongoing work being carried out at the National Institute of Statistical Sciences, on modeling high-level ozone exceedances as a function of time and meteorology.

SEISMIC RISK ANALYSIS BASED ON STRAIN ENERGY ACCUMULATION IN FOCAL REGION

Motoyuki Suzuki and Yoshio Ozaka
Tohoku University
Sendai, JAPAN

The object of this paper is to propose a stochastic method for evaluating the magnitude of future earthquakes taking account of nonstationariness in earthquake occurrence. For this purpose, the strain energy accumulation in focal region was estimated by means of the earthquake data of the past 100 years in Japan. Furthermore, the distributions of maximum ground acceleration were derived by means of the attenuation law. As a result, we found that the distributions of maximum ground acceleration fit the type III extreme value distributions and that the expected values of those distributions depend on the strain energy accumulation significantly. Finally, it is pointed out that the nonstationariness in earthquake occurrence should be taken into consideration in order to evaluate the earthquake load in design.

NOTE ON DOMAINS OF ATTRACTION OF MULTIVARIATE EXTREME VALUE DISTRIBUTIONS

Rinya Takahashi
Kobe University of Mercantile Marine
Kobe, JAPAN

By using dependence functions some necessary and sufficient conditions for the domains of attraction of multivariate extreme value distributions are shown. Joint asymptotic distributions of the multivariate extreme statistics are also shown.

APPLICATIONS OF MULTIVARIATE EXTREMES

Jonathan Tawn
University of Lancaster
Lancaster, UK

Many problems requiring application of extreme value methodology are multivariate in nature. Traditionally however, these problems have been simplified so that univariate methods can be applied. This often leads to analyses which neither incorporate the existing scientific knowledge about the processes under study nor enable the questions of practical relevance to be fully addressed. The recent advances in the statistical modelling of multivariate extremes (reviewed by Richard Smith at this meeting) provide a basis to address these practical problems more realistically.

In this talk I will review a number of problems arising from environmental applications and illustrate in each how multivariate extreme value methods are applied. Problems considered will be from the fields of coastal engineering (involving spatial dependence, spatial coherence of trends and

tides, wave - sea level dependence), temperature studies (involving modelling temporal dependence of the the process within extreme events) and reservoir safety (involving rainfall aggregations over different durations, concomitants, wind - rain dependence and use of hydrological models).

THE POISSON-WEIBULL FLAW MODEL FOR MODELING BRITTLE MATERIAL STRENGTH

Howard M. Taylor
University of Delaware
Newark, DE

The Weibull law is the distribution of choice for modeling the random strengths of many brittle materials such as modern high strength-high modulus fibrous composites, or their constituent fibers. The theoretical justification for the Weibull is compelling. Under quite general and plausible conditions, the Weibull arises as the asymptotic distribution of the minimum of identically distributed random variables, and the convergence is usually quite fast. Many reported sets of strength data on fibers and fibrous composites fit the Weibull distribution quite well.

Unfortunately, there are a significant number of sets of strength and fatigue data that clearly do not agree with the Weibull law. Discrepancies are of the three types. First, several authors have reported bimodal distritutions of strength. Of course, bimodality often signifies that the data arise from two or more populations. In the case of fibers, it might well indicate a significant spool effect, an effect that has been verified in several instances, sometimes as a diameter variation. Or it may signify distinct types of flaws, or distinct fracture morphologies. In long term static fatigue studies, clearly the properties of a material at the end of an experiment lasting months or years may differ distinctly from those at the start of the test. A plastic matrix may flow over time, for example. The most common approach to accomodating bimodality in data has been to use a mixture of distributions, sometime Weibull and sometimes not.

A second observed departure from Weibull behavior has been a slight, but unmistakable, curvature of the plot of data on Weibull coordinates. Such data can usually be adequately modeled in a descriptive sense by adding parameters. If mere data description is not the objective, but explanation or extrapolation is desired, the added parameters should have physical meaning and justification. Competing risks models are often used here, and while adding risks is not the same as mixing distributions, competing risks models can also model distinct fracture modes or flow types. As an example, surface flaws competing with internal flaws leads to complex size effects, since the surface flaws vary with surface area, while the internal flaws vary with volume.

An inconsistency between the Weibull parameter as estimated from the slope of the Weibull plot, and the same slope as estimated by the proportional change in mean strength with length is a third way in which the Weibull model has been found lacking. Under the Weibull, both phenomena should be driven by the same Weibull shape parameter.

We will show how a Poisson-Weibull flaw model provides a plausible framework for explaining all of the above discrepancies. Flaws, possibly of distinct types, are distributed throughout the material according to a nonhomogeneous Poisson process whose intensity may vary stochastically. Each flow contributes to a local weakness, the weakest of which may lead to failure. By analyzing the model in a variety of circumstances, we show that a variety of behaviors may be accomodated. Furthermore, the Poisson-Weibull flaw model exactly satisfies the usual, and desirable, weakest-link size effect scaling.

INDENTIFICATION OF FAILURE ORIGIN THROUGH TESTING AND THE WEIBULL RISK-OF-RUPTURE

John A. Tesk, Martin Y. M. Chiang, and Spurgeon M. Keeny, III*
National Institute of Standards and Technology
Gaithersburg, MD 20899
and
Jun Tang
Shanghai Pharmaceutical Industry Design Institute
Shanghai, CHINA

The stress distribution in bond layers of two different thicknesses (50 and 200 μm) was calculated by finite element analysis for rectangular, bonded metal specimens subjected to four point bending. These stresses were used in Weibull risk-of rupture (RR) calculations to aid in identification of the failure origin as interfacial. The characteristic strength from one test mode and bond thickness was correlated with that for another test mode and configuration. The finite element meshes were refined and the ratios of characteristic strengths were recalculated and virtually unchanged by the mesh refinement. The identification of the interface as the failure origin remained consistent. Further, the use of stresses extrapolated to zero mesh size also producted the same ratios. Hence, the RR calculation are not sensitive to the mesh sizes used for stress calculations.

*This work was partially supported by Interagency Agreement 2Y01 DE 30001 with the National Institute of Dental Research, Bethesda, MD 20892.

EXTREMES IN INSURANCE

J. L. Teugels and J. Beirlant
Katholieke Universitat
Leuven, BELGIUM

If an insurance portfolio is contaminated by large claims then the insurance company can consider taking reinsurance. Most of the currently employed treaties (as for example quota-share, excess of loss, stop-loss) are characterized by the fact that they use the entire portfolio on which a certain retention is applied. If reinsurance is geared to protect against large claims then it seems more natural to use treaties based on extreme claims (i.e. largest claims reinsurance, ECOMOR).

Since the mathematics involved in dealing with large claims is often forbidding we will suggest specific models for the main ingredients of the portfolio. For $N(t)$, the number of claims up to time t, we mention renewal processes and mixed Poisson processes given by $P\{N(t) = n\} = \sum_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dU(\lambda)$ where U is the so-called *structure distribution.* For the claim size distribution we will use a domain of attraction model.

Based on these assumptions we provide a survey of results that seem relevant to the practicing actuary (approximations, expectations, limits). However much still needs to be done and only an open discussion between practitioners and theoreticians guarantees that the really important issues are tackled.

LIMIT PROPERTIES OF MAXIMA OF WEIGHTED I.I.D. RANDOM VARIABLES

R. J. Tomkins
University of Regina
Regina, CANADA

Let $Z_n = max\{a_1 X_1, \ldots, a_n X_n\}, n \geq 1$, where $\{a_n\}$ is a positive real sequence and $X_1, X_2, \ldots$ 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 $\{Z_n\}$.

FRACTAL THEORY AND THE ESTIMATION OF EXTREME FLOODS

Donald L. Turcotte
Cornell University
Ithaca, NY

Floods and draughts constitute extreme values of great consequence to modern civilization. A wide variety of statistical techniques have been applied to the evaluation of the flood hazard. A primary difficulty is the relatively short time span over which historical data is available, quantitative estimates for paleofloods are generally suspect. It was in the context of floods that Hurst introduced the concept of the rescaled range. This was subsequently extended by Mandelbrot and his colleagues to concepts of fractional Gaussian noises and fractional Brownian walks. These studies introduced the controversial possibility that the extremes of floods and droughts could be fractal. An extensisve study of flood gauge records at 1200 stations in the United States indicates a good correlation with fractal statistics. It is convenient to introduce the parameter β which is the ratio of the ten-year flood to the one-year flood; for fractal statistics β is also the ratio of the 100-yearflood. It is found that β has strong regional variations associated with climate. The acceptance of power-law rather than exponentially based statistics would lead to far more conservative estimates of future flood hazards.

THE MEASUREMENT OF AVERAGES AND EXTREMES OF ENVIRONMENTAL VARIABLES

K. F. Turkman and C. W. Anderson
University of Lisbon, SPAIN and University of Sheffield, U.K.

Measurements of many environmental variables such as wind speeds, rainfall and the concentrations of atmospheric and aquatic pollutants are generally duration-specific: the actual quantities meaasured are averages over a specific time interval rather than instantaneous values. In practice however the scientists wishing to understand the environmenetal processes which lie behind such measurements, and the regulatory bodies which monitor pollution, often wish to deal with characteristics measured over some other time interval: peak concentrations over a day, for example, or high wind gust values, which in practice correspond to peak 3.5 second averages. There is therefore a need to understand the relationship between the statistics of environmental measurements averaged over different time scales, and to relate these statistics to extreme levels of the variables.

On the basis of theoretical probability models obtained for the relationship between averages of a stationary series and instantaneous extreme values of the series (Anderson and Turkman [1], [2]), statistical models have been developed for the dependence between hourly wind speeds and hourly maximum gust speeds (Anderson and Turkman [3]), and these models have been applied to extensive records of wind-speed data by Walshaw in [4].

In this paper, relationship between extreme averages over different time-periods will be investigated. A potential application is the linking of long historical data series containing information about extremes of daily rainfall (sometimes extending back to the early years of the century) to the shorter series of extreme hourly rainfall which have been recorded only in the past twenty or thirty years.

References

1. C. W. Anderson and K. F. Turkman. The joint limiting distributions of sums and maxima in stationary sequences. *J. Appl. Probab.*, 28:33-44, 1991.
2. C. W. Anderson and K. F. Turkman. Limit distributions for sums and maxima. *J. Appl. Probab.*, 28:715-716, 1991.
3. C. W. Anderson and K. F. Turkman. Limiting joint distributions of sums and maxima in a statistical context. *Theoriya Veroyatnostei ee Prim.*, 37:352-354, 1992.
4. D. Walshaw. *Statistical analysis of extreme wind speeds.* PhD thesis, University of Sheffield, 1991.

A MODEL FOR CLUSTER MAXIMA OF EXCEEDANCES OVER A THRESHOLD FOR OZONE DATA

José A. Villaseñor Alva
Colegio de Postgraduados
Chapingo, MEXICO

The study of the probability behaviour of the concentration of air pollutants in the atmosphere is of great interest in order to take actions in favor of human health protection in large cities. This work proposes a parametric distributional model for the maximum concentration of a pollutant coresponding to a group of daily maximum concentrations exceeding a threshold taking into consideration their dependence. This model also provides a joint probability distribution for the maximum and the group size. The parameter estimation problem is discussed and a set of ozone data from Mexico City is analyzed.

ZERO-ONE LAWS FOR LARGE ORDER STATISTICS WITH VARIABLE RANKS

Hong Wang
Memorial University of Newfoundland
St. John's, Newfoundland, CANADA

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common distribution function $F(x)$. Let $\{u_n\}$ be a real sequence and $\{r(n)\}$ be a non-decreasing integer sequence such that $r(n) \leq n$ for all positive integers n. Denote by $Z_{(r(n);n)}$ the $r(n)$-th largest of $\{X_1, X_2, \ldots, X_n\}$. We prove that, under certain conditions, the probability $P\{Z_{(r(n);n)} \leq u_n \text{ i.o.}\} = 0$ or 1 according as $\sum_{k=1}^{\infty} P\{Z_{(r(n_k);n_k)} \leq u_{n_k}\} < \infty$ or $= \infty$, where $\{n_k\}$ is a specific subsequence. Based on this result, one of Deheuvels's (1986) results will be proved under weaker assumptions. In particular, Klass's (1984, 1985) results will be extended to the case when $r \geq 1$.

GETTING THE MOST FROM YOUR EXTREME WIND DATA: A STEP-BY-STEP GUIDE

David Walshaw
University of Newcastle upon Tyne
Newcastle, U.K.

Models for extremes of environmental processes have been studied extensively in recent years. The particular problems arising when attempting to estimate return levels from sequences of measurements on the appropriate variables have been considered in some detail. In particular, the aspects of seasonal variation and short–range dependence have received a great deal of attention.

In this paper we present a case study based on ten years of hourly wind speed measurements collected at a single U.K. site, elucidating the most successful procedure emerging from an extensive

study of this data. The basic model (in which an extreme value distribution is fitted to cluster peak excesses over a high threshold) is standard. However the emphasis is on a number of practical problems which will always arise when such models are fitted to wind speeds, but which have received little consideration. These include:

- model selection and assessment of model adequacy when the threshold, and some or all of the parameters, are allowed to vary seasonally;
- the choice of the best combination of threshold and cluster identification procedure;
- the choice of precision measurement on return level estimates.

The aim is to suggest an algorithm which can be generally applied to the problem of gust return level estimation at a single site.

ORDER STATISTICS AND PROOFS OF COMBINATORIAL IDENTITIES

Roberta S. Wenocur
University of Pennsylania
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, the multivariate inverse Pólya distribution, and extreme 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 extremes, in particular the *maximum*, of a random sample.

CONVERGENCE AND DIVERGENCE OF FRACTIONAL PARTS OF EXTREME ORDER STATISTICS

R. J. G. Wilms
Eindhoven University of Technology
Eindhoven, THE NETHERLANDS

Let $x \stackrel{d}{=} x_1, x_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables (rv's) with distribution function (df) F, and suppose that the right endpoint of F is infinite. We write $F\epsilon\mathcal{D}(G)$ if there exist sequences of real numbers $(a_n)(a_n > 0 \text{ for } n\epsilon I\!N)$ and (b_n) such that $\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \quad (x\epsilon I\!R)$. Furthermore, $\Phi_\alpha(\alpha > 0)$ and Λ denote the Fréchet type and Gumbel type extreme value distribution, respectively.

Let further $\{X\}$ denote the fractional part of a rv X, defined by $\{X\} : = X - [X]$, where [X] is the integer part of X. In addition, we use the following notation:

$Z_n = \max(X_1, \ldots, X_n)$;

U is a rv with a uniform distribution on $[0,1]$;

V is a rv with an exponentional distribution and $\mathbb{E}V = 1$;

$x_+ = \max(x, 0) \quad (x \epsilon \mathbb{R})$.

We study the limit distribution of $\{Z_n\}$ and give sufficient conditions on F for $(Z_n\}$ to converge or to diverge in distribution as n tends to infinity. To this end we consider the following three cases:

1. $\{Z_n\} \xrightarrow{d} U$.

Representing F in terms of hazard functions, i.e. writing X as $X \stackrel{d}{=} h(V)$ for some non-decreasing function $h : \mathbb{R}_+ \to \mathbb{R}$, we find our main result:

$$h((Y + \log n)_+) \xrightarrow{d} U \ \text{ iff } \ \{Z_n \xrightarrow{d}\}U \quad (n \to \infty),$$

where Y has df Λ. We then give sufficient conditions for $h((Y + \log n)_+) \xrightarrow{d} U \quad (n \to \infty)$ to hold. A sufficient condition is $h(V + t) \xrightarrow{d} U \quad (t \to \infty)$, and a rather simple one is the requirement that the hazard rate of F tends to zero monotonically. These results have been proved in [1].

The following result shows the connection with extreme value theory. Suppose $F \epsilon \mathcal{D}(G), G = \Phi_\alpha$ or Λ. If $a_n \to \infty$, and if $\frac{d}{dx}(F^n(a_n x + b_n)) \to G'(x) \quad (x \epsilon \mathbb{R})$ locally uniformly on $(0, \infty)$ and $(-\infty, \infty)$, respectively, then $(Z_n) \xrightarrow{d} U \quad (n \to \infty)$.

2. $\{Z_n\} \xrightarrow{d} W \neq U$.

For any rv $W \neq U$ with $P(0 \leq W < 1) = 1$ we can construct a sequence of i.i.d. rv's $(X_m)_1^\infty$ such that $\{Z_n\} \xrightarrow{d} W \quad (n \to \infty)$. For the proof we need results from [2].

3. $\{Z_n\} \not\xrightarrow{d}$.

In this case, divergence of the distribution of $\{Z_n\}$, we succeeded in giving sufficient conditons on F by assuming $F \epsilon \mathcal{D}(\Lambda)$. In that case, if a $a_n \leq M$ for some $M > 0 \quad (n \epsilon \mathbb{N})$, and $\lim_{n\to\infty} b_{n+1} - b_n = 0$, then $\{Z_n\}$ diverges in distribution $(n \to \infty)$.

Representing F in terms of hazard functions we can prove the following result:

Suppose that $F \epsilon \mathcal{D}(\Lambda)$ has a non-increasing density if $x \geq x_0$, for some $x_0 \epsilon \mathbb{R}$. Assume further that the hazard rate g of F is monotone. Then we have

[i] if $a_n \leq M$ for some $M > 0 \quad (n \epsilon \mathbb{N})$, then $\{Z_n\}$ diverges in distribution $(n \to \infty)$.

[ii] if $g(t) \to \infty$ or $g(t) \to c > 0 \quad (t \to \infty)$, the $\{Z_n\}$ diverges in distribution $(n \to \infty)$.

References.

1. Wilms, R. J. G. and J. J. A. M. Brands (1991), On the asymptotically uniform distribution modulo 1 of extreme order statistics, Memorandum COSOR 91-38, Eindhoven University of Technology, Eindhoven, The Netherlands (submitted for publication).
2. Brands, J. J. A. M., F. W. Steutel and R. J. G. Wilms (1992), On the number of maxima in a discrete sample, Memorandum COSOR 92-16, Eindhoven University of Technology, Eindhoven, The Netherlands (submitted for publication).

DISTRIBUTIONS OF EXTREMES WHEN A TREND IS PRESENT WITH APPLICATIONS TO N.Z. RAINFALL AND RIVERFLOW SERIES

Kit Withers
New Zealand institute of Industrial Research
Lower Hutt, NEW ZEALAND

Fisher and Tippett showed that if the distribution of the maximum of a random sample, suitably scaled, has a limit as the sample size increases, then the limit has one of three forms. These are commonly called the EV1, EV2 and EV3 distributions. They may be combined as a parametric family called the GEV (generalized extreme value) distribution.

Hosking and colleagues proposed using the L-moments to fit the parameters of the GEV, and showed that this could be more efficient than the maximum likelihood estimates.

We consider the problem of fitting a trend to extreme data. One approach is to adapt Hosking's L-moment method. We report on applying this to selected N.Z. annual extreme rainfall and riverflow series.

We also report on how this model can be expanded in line with some theoretical results for the distribution of extremes when a trend is present.

THE EXTREMAL INDEX OF A HIGHER ORDER STATIONARY MARKOV CHAIN

Seokhoon Yun
University of North Carolina
Chapel Hill, NC

The extremal index turns out ot be a key parameter for studying extremal behavior of a stationary sequence of random variables. However, it is not known how to calculate the extremal index analytically except for a few mathematical examples. In this talk, a method of calculating the extremal index of a real-valued, higher order (kth order, $k \geq 1$) stationary Markov chain is introduced. The method is based on the assumption that the joint distribution of $k+1$ consecutive variables belongs to the domain of attraction of some multivariate extreme value distribution. In most cases, this assumption is equivalent to the existence of limits of rescaled stationary transition densities. These limiting transition densities define a new $(k-1)$th order Markov chain from which a nice representation for the extremal index is obtained. Since most multivariate distribution models are specified by densities, not by distribution functions, this approach in terms of transition densities is particularly helpful for practical purposes. The representation is then well suited for simulation to calculate the extremal index. Some examples with simulation results are presented.

EXTREME VALUES IN BUSINESS INTERRUPTION INSURANCE

Daniel Zajdenweber
PARIS X Nanterre University
Paris, FRANCE

In case of an accident stopping the economic activity of a firm covered by a business interruption insurance policy, the insurance company reimburses the damaged firm for the sales lost, minus the costs spared because of the interruption of the production. Those accidents must not be confused with fire-damages, which are the most frequent causes of business interruption (they are covered separately by an "ordinary" fire-damage policy), and the value of a business interruption claim is not a simple proportion of the size of the physical damages due to a fire. Business interruption claims are characterized by extreme variability of the yearly claims recorded in a country. In France, for instance, the total amount of the business interruption claims paid in 1988 by the insurance industry (1,198 million Francs) was more than two times larger than in 1987 (520 million Francs). And for the first time in the history of that type of insurance in France, the total amount of claims paid to the firms was larger than the premiums collected (1,158 million Francs). That variability is easily explained by the occurrence of huge claims larger than one hundred million or even sometimes larger than three hundred million Francs. They are not numerous but they make the tail of the yearly size-distribution extremely long. Hence, the analysis of the size-distribution and the valuation of the fair premium relevant to that kind of risk are the main purposes of the paper.