MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 14/1984

Order Statistics, Quantile Processes and Extreme Value Theory

25.3. bis 31.3.1984

Die Tagung fand unter der Leitung von Herrn R.-D. Reiß (Siegen) und Herrn W.R. van Zwet (Leiden) statt. In den insgesamt 34 Vorträgen wurde fast die gesamte Breite der drei verschiedenen, sich zum Teil überschneidenden Themenkreise angesprochen.

Als thematische Schwerpunkte der Tagung sind (a) Grenzwertsätze für Differenzen von Ordnungsstatistiken, (b) Schätzverfahren in den Schwänzen einer Verteilung, (c) Regressions-Quantile in linearen Modellen, (d) starke Approximationen von empirischen Prozessen und Quantil-Prozessen und (e) Grenzwertsätze für Maxima von Zufallsvariablen zu nennen. Von verschiedenen Seiten wurden Problemstellungen im Zusammenhang mit Funktionen von Ordnungsstatistiken angesprochen, wobei das zentrale Gebiet der L-Statistiken mehr sporadisch als systematisch behandelt wurde. Neben neueren Ansätzen wie in (a) und (b) waren in den Vorträgen als wesentliche Tendenzen die Entwicklung von verfeinerten Approximationsmethoden und die Behandlung von klassischen Fragestellungen unter abgeschwächten Modellannahmen zu erkennen.

Die Tagung hat es insbesondere ermöglicht, zwischen Wissenschaftlern, die auf benachbarten Gebieten arbeiten, einen engeren Kontakt zu knüpfen, wobei die einzigartige Atmosphäre des Mathematischen Forschungsinstituts Oberwolfach ganz entscheidend dazu beigetragen hat.

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Program for monday, march 26

Morning session

Chairman: V. Mammitzsch

09.00: Order statistics in finite sets of non-iid variables
H.A. David

09.40: Strong limiting bounds for maximal spacings
P. Deheuvels

11.00: An improved Erdös-Rényi strong law for moving quantiles
J. Steinebach

11.30: Local time of the empirical and quantile process
P. Révész

Afternoon session

Chairman: H.A. David

16.00: Estimating tails of probability distributions
R.L. Smith

16.50: On estimation of large quantiles
D.D. Boos

17.40: Berry-Esseen theorems for the kernel quantile estimator
M. Falk

Program for tuesday, march 27

Morning session

Chairman: R. L. Smith

09.00: Point processes of exceedances with clustering
M.R. Leadbetter

09.50: Extremes of moving averages
H. Rootzén

11.00: Optimal prediction of extremes and level crossings in Gaussian processes
G. Lindgren

11.40: Extreme values of non-stationary sequences
J. Hüsler

Afternoon session

Chairman: W. Stute

16.00: Properties of order statistics under outlier - generating models
U. Gather

16.40: On asymptotic normality of Hill's estimator for the index of regular variation
E. Häusler

17.15: Order statistics in insurance mathematics
J.L. Teugels

17.45: Order statistics and symmetric functions
L. Rüschendorf

Program for wednesday, march 28

Morning session

Chairman: J.L. Teugels

09.00: Asymptotic theory for generalized L-statistics
R. Helmers

09.50: Asymptotics for uniform spacings
C.A.J. Klaassen

10.20: Higher order asymptotics for statistics based on uniform spacings
R.J.M.M. Does

11.15: A multivariate two sample test based on the number of nearest neighbor-type coincidences
N. Henze

11.45: Good news and bad news on nonparametric density and regression function estimate
L. Györfi

Program for thursday, march 29

Morning session

Chairman: H. E. Daniels

09.00: A new approximation for empirical and quantile processes I.
(weighted processes)
M. Csörgö

09.50: A new approximation for empirical and quantile processes II.
(processes indexed by classes of functions)
D.M.Mason

11.00: A new approximation for empirical and quantile processes III.
(applications to fine structure of partial sums)
S. Csörgö

11.55: The limit behaviour of the maximum of random variables with applications to outlier resistance
R. Mathar

Afternoon session

Chairman: M. Csörgö

16.00: Some remarks on multivariate empirical processes
F.H. Ruymgaart

16.40: Some open questions in quantile processes for linear models I
R. Koenker

17.40: Some open questions in quantile processes for linear models II
G.W. Bassett

Program for friday, march 30

Morning session

Chairman: R. Helmers

09.00: Records from an improving population
L.de Haan

09.40: Rates of convergences to the extreme value distributions
A.A. Balkema

10.40: Bivariate extremes: Basics and model-building
J. Tiago de Oliveira

11.20: On the distribution of M-estimators
H. Dinges

Afternoon session

Chairman: J. Tiago de Oliveira

16.00: A test for a change-point
Y. Mittal

16.40: Limit distributions for extreme order statistics in the non-identically distributed case
D. Mejzler

17.40: The maximum of a random walk whose mean path has a maximum
H.E. Daniels

Abstracts

A.A. BALKEMA

Rates of convergences to the extreme value distributions

Assume that F lies in the domain of attraction of $\Lambda(x) = e^{-e^{-x}}$, and set $\delta_n = \sup |F^n(a_n x + b_n) - \Lambda(x)|$. Then $\delta_n \to o$ for some choice of norming constants a_n and b_n. Choosing a_n and b_n optimal, the difference $F^n(a_n x + b_n) - \Lambda(x)$ will have three extremes of alternating sign of the same magnitude.

By the standard representation $X_n = f(U_n)$ with f nondecreasing, $U_1, U_2, \ldots$ iid from Λ, one can derive rate of convergence results by introducing the curvature $\alpha(u) = f''(u)/f'(u)$.

This research has been done in collaboration with L. de Haan und S. Resnick.

G.W. BASSETT, R. KOENKER

Some open questions in quantile processes for linear models

Noting that the sample quantiles may be obtained as M-estimates with the p function $p_\theta(t) = \theta|t|^+ + (1-\theta)|t|^-$ we have suggested p-dimensional analogues of the sample quantiles for the linear model which solve $\min \sum_{i=1}^{n} p_\theta(y_i - x_i b)$ for $b \in \mathbb{R}^p$. Letting solutions to this problem be $B_{\theta n}$, an associated quantile function may be defined as $\hat{Q}_n(\theta) = \inf\{\bar{x}b \mid b \in B_{\theta n}\}$. We show that properly normalized versions of this empirical quantile function have finite dimensional distributions of the Brownian Bridge for linear models with iid errors. A Glivenko-Cantelli result is also proved. Some questions regarding conditions for and rates of convergence are raised and some problems regarding applications are also discussed.

D.D. BOOS

On estimation of large quantiles

Weissman (1978, JASA) suggested quantile estimators $\hat{\eta}_p$ of $F^{-1}(p)$ based on the joint limiting distribution of the k largest order statistics. The talk focuses on the asymptotic distribution of $\hat{\eta}_p$ under three different limiting schemes. For practical choice of k, the CLT approach based on $k/n \to \alpha > 0$ gives the most insight.

M. CSÖRGÖ, S. CSÖRGÖ, D.M. MASON

A new approximation for empirical and quantile processes

Let $U_1, U_2, \ldots$ be a sequence of independent uniform (0,1) random variables and for each $n \geq 1$ let $U_{1,n} < \ldots < U_{n,n}$ denote the order statistics based on the first n uniform (0,1) random variables. Also let G_n denote the empirical distribution function based on $U_1, \ldots, U_n$ and let $U_{n(s)} = U_{i,n}$ when $(i-1)/n < s \leq i/n$ for $i=1,\ldots,n$ denote the corresponding empirical quantile function based on these random variables. Finally, let $\alpha_n(s) = \sqrt{n}(G_n(s)-s)$ and $\beta_n(s) = \sqrt{n}(s-U_n(s))$ for $0 < s < 1$ denote the uniform empirical and quantile processes. A probability space is constructed carrying a sequence of independent uniform (0,1) random variables and a sequence of Brownian Bridges $B_n, n=1,2,\ldots,$ such that the following inequality holds:

$$P\{\sup_{0 \leq s < d/n} |\beta_n(s) - B_n(s)| \geq n^{-1/2}(a \log d + x)\} \leq be^{-cx}$$

and

$$P\{\sup_{1-d/n \leq s \leq 1} |\beta_n(s) - B_n(s)| \geq n^{-1/2}(a \log d + x)\} \leq be^{-cx}$$

whenever $n_o \leq d \leq n$ and $0 \leq x \leq d^{1/2}$, where n_o, a, b and c are suitably chosen positive constants.

These two inequalities lead to two important facts. On the probability space that we have constructed for all $0 \leq \nu < 1/4$

$$\sup_{0 \leq s \leq 1} |\alpha_n(s) - \bar{B}_n(s)| / (s(1-s))^{1/2-\nu} = O_p(n^{-\nu}),$$

where $\bar{B}_n(s) = B_n(s)$ if $1/n < s \leq 1-1/n$ and zero otherwise; whereas for all $0 \leq \nu < 1/2$

$$\sup_{1/(n+1) \leq s \leq n/(n+1)} |\beta_n(s) - B_n(s)| / (s(1-s))^{1/2-\nu} = O_p(n^{-\nu}).$$

These last two parts appear to have wide ranging application in probability and statistics. We mention just a few of these here. The important Chibisov-O'Reilly and the Jaeschke-Eicker theorems together with generalizations of Rényi-type results are easy consequences. On addition, very general approximations for the empirical and quantile processes indexed by classes of functions are obtained, which in turn give weak convergence results for the empirical moment generating function, empirical Hall, moment and generalized mean function processes together with a simple probabilistic proof of the sufficiency of the normal convergence criterion. Also these two facts are important tools in the study of the fine structure of the limiting distributional behavior of partial sums of independent identically distributed random variables. We would like to mention that this was a three part talk given by the undersigned three authors describing work done jointly with Lajos Horvâth (Szeged), who was invited to participate in this conference, but was unable to attend.

H.E. DANIELS

The maximum of a random walk whose mean path has a maximum

The joint distribution of the maximum and the time at which it is attained is discussed. In a commom type of problem - e.g. the maximum number of infectives present during the course of an epidemic - the problem reduces to one of diffusion in the presence of a parabolic boundary $\frac{1}{2}b(t-t_o)^2$, $b > 0$. This leads to a certain integral equation

$$\int_0^\infty F(u+v)e^{-(v/2)(u+v/2)^2}\frac{dv}{\sqrt{2\pi\sigma}} = 1$$

which was solved by T. Skyrme in the form

$$F(\kappa) = 2^{-1/3} e^{\kappa^3/6} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{s\kappa}ds}{Ai(-2^{1/3}s)} .$$

Applying this to the case of a Brownian Bridge on (O,T), this leads to the following results with $\kappa = N^{1/3} b^{2/3}(t_o - t)$. Marginal density of $m: f(m) = \sqrt{\frac{T}{2\pi t_o(T-t_o)}} \exp\left(\frac{T}{2\, t_o(T-t_o)}\right.$ $\{m-\lambda N^{-1/6} b^{-1/3}\}^2\{1+O(N^{-1/3})\})$ where $\lambda = 2\int_{-\infty}^{\infty}F'(\kappa)F(-\kappa)d\kappa - \int_{-\infty}^{\infty}\kappa^2F(\kappa)F(-\kappa)d\kappa = .995$. Marginal density of κ: $f(\kappa) = 2F(\kappa)F(-\kappa)\{1+O(N^{-1/3})\}$. Conditional density of $m|\kappa$:

$$f(m|\kappa) = \sqrt{\frac{T}{2\pi t_o(T-t_o)}} \exp\left(-\frac{T}{2\, t_o(T-t_o)}\{m - N^{-1/6} b^{-1/3}\mu(\kappa)\}^2\{1+O(N^{-1/3})\}\right)$$

where

$$\mu(\kappa) = \left(1-\frac{t_o}{T}\right)\nu(\kappa) + \frac{t_o}{T}\nu(-\kappa),$$

$$\nu(\kappa) = F'(\kappa)/F(\kappa) - 1/2\,\kappa^2$$

(N is a "typical" time on the original scale - e.g. the time at the maximum of an epidemic and $m = N^{-1/2}$ max on the original scale).

H.A. DAVID

Order statistics in infinite sets of non-iid variates

An attempt is made to systematize the treatment of order statistics arising from non-iid variates $X_1,\ldots,X_n$. After a review of the main approaches available, special emphasis is given to the case where the X's are independent non-identically distributed (inid). This situation is motivated by (a) k-out-of-n systems of unlike independently failing components (Proschan) and (b) robustness against an outlier. Some new results are obtained for (b). Attention is also drawn to some 15 articles concerned with finite sets of non-iid variates.

P. DEHEUVELS

Limiting results for maximal spacings

Let $0 = U_{0,n} < U_{1,n} < \ldots < U_{n,n} < 1 = U_{n+1,n}$ be the order statistics of iid random variables, uniformly distributed on $(0,1)$. Define for $1 \leq K \leq n$ the k-th maximal K-spacing as the k-th maximum $M_{k,K}^{(n)}$ of the sequence $\{U_{i,n} - U_{i-K,n}, i=K,\ldots,n+1\}$.

The following results have been obtained for $M_{k,K}^{(n)}$; $\log_p n$ is the p-th iterated log.

a) Weak laws. Let $K \geq 1$ be fixed, and also $k \geq 1$. Then, for any $u \in \mathbb{R}$,

$$\lim_{n\to\infty} P\{n M_{k,K}^{(n)} - \log n - (K-1)\log_2 n + \log\{(K-1)!\}<u\} = e^{-e^{-u}}\{\sum_{l=0}^{k-1} \frac{e^{-lu}}{l!}\}.$$

b) Strong laws. Let $K \geq 1$ be fixed, then:

$$\liminf_{n\to\infty} \frac{nM_{1,K}^{(n)} - \log n - (K-1)\log_2 n}{\log_2 n} = 0, \quad \limsup_{n\to\infty} (\cdot) = 2.\text{a.s.}$$

c) Strong laws. Let $K = K_n \geq 1$ be non-decreasing and such that $K_n = o(\log n)$, $K_n \to \infty$. Then

$$\lim_{n\to\infty} \frac{n\, M_{1,K}^{(n)} - \log n}{(K-1)\log(\frac{e \log n}{K})} = 1 \text{ a.s.}$$

d) An evaluation is also given for $K_n = [c \log n]$. Here b,c,d corresponds to a joint work with Luc Devroye to appear in the Z. Wahrscheinlichkeitstheor. Verw. Geb., a will appear in the proceedings of the Pannonian Symposium, Bad Tatzmannsdorf, 1983.

Consider now the non-uniform case. Assume that $X_1, X_2, \ldots$ is an iid sequence with density $f > 0$ on (A,B) and 0 otherwise. Let then the order statistics be $X_{1,n} < X_{2,n} < \ldots < X_{n,n}$ and put $\Delta_n = \max_{1\leq i\leq n-1} (X_{i+1,n} - X_{i,n})$. The following strong bounds have been obtained.

e) Assume that $-\infty < A < B < +\infty$ and that, for $x_o \in (A,B)$, $f(x_o) > 0$ and $f(x_o) < f(x)$ for any $x \in (A,B)$. Assume also that $f(x) - f(x_o) \sim C|x-x_o|^r$ as $x \to x_o$. Then we have

$$-\frac{1}{r} = \liminf_{n\to\infty} \frac{n\Delta_n f(x_o) - \log n}{\log_2 n}, \quad 2-\frac{1}{r} = \limsup_{n\to\infty} \frac{n\Delta_n f(x_o) - \log n}{\log_2 n} \text{ a.s.}$$

f) If we assume $-\infty < A < B \leq +\infty$, then the limiting behavior of $X_{n,n}$ characterises the limiting behavior of Δ_n if $X_{n,n} \in \mathcal{D}(G)$, where $G = \Lambda$, $G = \Phi_\alpha$ or $G = \Psi_\alpha$, $\alpha > 1$. The corresponding limiting laws and strong bounds are given in precise form, with examples such as the Gaussian law. e is due to appear in the Annals of Probability.

H. DINGES

<u>On the distribution of M-estimators</u>

If Q_n has a distribution close to $N(\mu, \frac{1}{n}\sigma^2)$ for large n, then for any smooth $T(\cdot)$ with positive derivative also $P_n := T(Q_n)$ is approximately normally distributed. In many classical cases there exists a (unique!) normalizing transformation $A(\cdot)$, i.e. a transformation which makes $A(Q_n)$ very similar to $N(0,\frac{1}{n})$. This was made precise and exemplified in the lecture. In particular the concepts of skewness and excess got a specification which applies to M-estimators of location (being "strongly normal" in the sense of the talk). Furthermore large deviation results extending Cramer's result (1938) were mentioned. The technique behind the investigation is the saddle point method exploited by H. Daniels in quite similar situations.

R. DOES

Higher order asymptotics for statistics based on uniform spacings

Let $U_1, U_2, \ldots$ be a sequence of independent uniform (0,1) random variables. For $n = 1,2,\ldots, U_{1:n} \leq U_{2:n} \leq \cdots \leq U_{n:n}$ denote the ordered $U_1, U_2, \ldots, U_n$. Let $U_{0:n} = 0$ and $U_{n+1:n} = 1$. Uniform spacings are defined by $D_{jn} = U_{j:n} - U_{j-1:n}$ for $j = 1,2,\ldots,n+1$. Let $g: [0,\infty) \to \mathbb{R}$ be a fixed nonlinear measurable function and define statistics T_n by

$$(*) \quad T_n = \sum_{j=1}^{n+1} g((n+1)D_{jn})$$

Statistics of this form can be used for testing uniformity. It is well-known that suitably normalized T_n are asymptotically normally distributed under quite general conditions.

The aim of this talk is to discuss the problems of obtaining Berry-Esseen bounds of the order $n^{-1/2}$ and establishing Edgeworth expansions with remainder $O(n^{-1})$ for statistics of the form (*). This work is joint with R. Helmers and C.A.J. Klaassen.

M. FALK

Berry-Esseen theorems for the kernel quantile estimator

Several Berry-Esseen type theorems are established for the kernel estimator of the underlying q-quantile. These results make it possible to compare the kernel estimator with the sample q-quantile on the basis of their covering probabilities of symmetric intervals. Lower bounds for the relative deficiencies of the empirical q-quantile with respect to the kernel estimator are established.

U. GATHER

Properties of order statistics under outlier-generating models

The most usual outlier-generating distribution models for real-valued r.v.'s $X_1,...,X_n$ are presented. These are the Identified-Outliers-Model (H_I), the Ferguson-type model (H_F), the Exchangeable model (H_E), the Mixture model (H_M) and the Labelled model (H_L) (cf. Barnett, Lewis (1978)).

It is shown that the distributions of the order statistics $(X_{1,n},...,X_{n,n})$ under the models H_I, H_F and H_E are equal. Moreover, the above distribution-models are studied with respect to their degree of 'outlier-proneness' which is defined in terms of the limit behaviour of $P\{X_{n,n} - X_{n-1,n} > \varepsilon\}$ or $P\{X_{n,n}/X_{n-1,n} > 1+\varepsilon\}$ for $\varepsilon > 0$, $n \to \infty$, generalizing notions of Green (1976).

L. GYÖRFI

Good news and bad news on nonparametric density and regression function estimate

For a density f and sample size n let I_n denote the L_1 error of a density estimate.

(i) If $a_n > 0$, $a_n \to 0$ and estimates f_n are given then there exists f such that

$$\limsup_{n\to\infty} \frac{E\, I_n}{a_n} = \infty,$$

thus an arbitrary bad rate of convergence can be achieved.

(ii) Consider the histogram f_n for parameters h_n such that $h_n \to 0$, $nh_n^d \to \infty$ (d is the dimension). For each f and for each $\varepsilon > 0$ there exists n_o such that

$$P\{I_n > \varepsilon\} \leq \exp(-c\varepsilon^2 n) \text{ for each } n \geq n_o,$$

where c is a universal constant.

Similar properties can be shown for the L_1 error of regression function estimation.

L. DE HAAN

Records from an improving population

A parametric and a non-parametric model are proposed for the limiting behaviour of records from a sequence $(X_i+b_i)_{i\geq 1}$ where $X_1, X_2 \ldots$ are iid random variables and $(b_i)_{i\geq 1}$ an increasing sequence of constants.

E. HÄUSLER

On asymptotic normality of Hill's estimator for the index of regular variation

Let F be a distribution function with regularly varying upper tail, i.e. assume $1 - F(x) = x^{-\alpha}L(x)$ with $\alpha > 0$ and a function L which is slowly varying at infinity. Consider iid observations $\xi_1, \xi_2, \ldots$

distributed according to F, and let $\xi_{1:n},\ldots,\xi_{n:n}$ denote the order statistics based on $\xi_1,\ldots,\xi_n$. It is known that averages of the extreme data $\xi_{k:n},\ldots,\xi_{n:n}$ of the form $H_k^{(n)} = \frac{1}{k}\sum_{i=1}^{k}\log\xi_{n+1-i:n} - \log\xi_{n-k:n}$ are consistent estimators of $\frac{1}{\alpha}$ provided that $k = k(n)\to\infty$ and $k = o(n)$. Asymptotic normality, however, is present for $\sqrt{k}(H_k^{(n)} - \frac{1}{\alpha})$ only for sequences $k = k(n)$ which converge to infinity slowly enough. We give a condition which can be used to determine these sequences if some prior knowlege about L is available.

R. HELMERS

Asymptotic theory for generalized L-statistics

Recently generalized L-statistics were introduced by R.J. Serfling. The class of these statistics is quite large. It includes U-statistics, linear functions of order statistics and many other statistics of interest. Generalized L-statistics can be viewed as L-functionals of the empirical df of U-statistic structure. This representation was used by Silverman (1983) and Serfling (1984) to prove asymptotic normality for generalized L-statistics. We supplement their results by deriving an exponential bound and some Glivenko-Cantelli type theorems for the empirical df of U-statistic structure and a strong law as well as a Berry-Esseen bound for generalized L-statistics. As an application the asymptotic behaviour of the bootstrap approximation for generalized L-statistics is examined. This work is joint with P. Janssen and R.J. Serfling.

N. HENZE

A multivariate two sample test based on the number of nearest neighbor-type coincidences

For independent s-variate samples $X_1,\ldots,X_n$ iid $\sim f(x)$, $Y_1,\ldots,Y_n$

iid$\sim g(x)$, where the densities $f(\cdot)$ and $g(\cdot)$ are assumed to be continuous on their respective sets of positivity, consider the number $T_{m,n}$ of points Z of the pooled sample (which are either of type 'X' or of type 'Y') such that the nearest neighbor of Z is of the same type as Z. We show that, as $m,n \to \infty$, $\frac{m}{m+n} \to \tau$, $o < \tau < 1$, $f(\cdot) \equiv g(\cdot)$, $\lim \mathrm{Var} \left(\frac{1}{\sqrt{m+n}} T_{m,n}\right) = \sigma^2(\tau,s)$, independently of $f(\cdot)$. An omnibus test for the two sample problem '$f(\cdot) \equiv g(\cdot)$ or $f(\cdot) \not\equiv g(\cdot)$?' may be obtained by rejecting the hypothesis $f(\circ) \equiv g(\cdot)$ for large values of $T_{m,n}$.

J. HÜSLER

Extreme values of non-stationary sequences

The conditions used to generalize the extreme value theory for stationary random sequences to non-stationary sequences are discussed with respect to sufficiency and necessity. We mention some limit results as the convergence to non-degenerate limit distributions for maxima and the Poisson-like-behaviour of the exceedances. Related to the local dependence we discuss the extremal index with respect to non-stationary random sequences or moving boundaries. Finally we state some results on the normalization of maxima for a random sequence which consists of a stationary random sequence and a deterministic trend sequence.

C.A.J. KLAASSEN

Asymptotics for uniform spacings

For statistics based on uniform k-spacings a limit theorem will be derived. The method relies on a characterization of these statistics as conditional statistics. The technique used is applicable more generally; namely to those conditional statistics, for which a mild regularity condition is satisfied on the density of the quantity one conditions on.

M.R. LEADBETTER

Point processes of exceedances with clustering

It is known that under appropriate "long range" and "local" dependence conditions, the central results for classical extreme value theory remain true for dependent (stationary) sequences. In particular under the conditions "$D(u_n)$" (a distributional mixing condition) and "$D'(u_n)$" (a local dependence restricition) the asymptotic distributions of the maximum of extreme order statistics are the same as if the sequence were iid with the same marginal df. Here we investigate the effect of relaxing the condition D', noting that the distribution of the maximum is essentially unchanged (at least as far as its distributional type) whereas the distributions of extreme order statistics are altered in a specific way. This is shown through a study of the convergence properties of the point processes of high level exceedances, which become Poisson when $D'(u_n)$ holds, but involve clustering (compounding) when D' is relaxed.

G. LINDGREN

Optimal prediction of extremes and level crossings in Gaussian processes

One sometimes wants to predict (in advance!) upcrossings of a critical level u by a stationary Gaussian process X(t), using data $Y(t) \in \mathbb{R}^p$ available at time t, jointly stationary and Gaussian with X(t). The "optimal" predictor studied in the talk, uses the mean square predictor $\hat{X}_t(t+h) = E(X(t+h)|Y(t))$ and the expected (conditional) growth rate $\hat{Z}_t(t+h) = E(X'(t+h)|Y(t), X(t+h) = u)$ (here $h > o$ is any fixed quantity). Let $\sigma^2_{x \cdot y} = V(X(t+h)|Y(t))$, $\sigma^2_{x' \cdot xy} = V(X'(t+h)|X(t+h), Y(t))$ be the residual variances, and define $\psi(x) = \varphi(x) + x\Phi(x)$ from the standard normal density and distribution function.

The optimal alarm policy gives alarm at time t for an upcrossing at time t+h if

$$\left(\frac{u-\hat{X}_t(t+h)}{\sigma_{x \circ y}}\right)^2 \leq 2 \log \psi\left(\frac{\hat{Z}_t(t+h)}{\sigma_{x' \cdot xy}}\right) + 2 \log \frac{\sigma_{x' \cdot xy}}{\sigma_{x \cdot y}} + K_h,$$

where K_h is a constant "of your choice". It is shown that this alarm has the best possible detection probability under the restriction of a fixed total alarm time.

The optimal alarm was compared to the naive alarm which gives alarm for an upcrossing if

$$\left(\frac{u-\hat{X}_T(t+h)}{\sigma_{x \cdot y}}\right)^2 \leq K_h$$

and it was shown that the optimal alarm locates the correct time of the upcrossings more precisely and at an earlier stage than the naive alarm, which has a tendency to give late alarms.

R. MATHAR

<u>The limit behaviour of the maximum of random variables with applications to outlier resistance</u>

We consider degenerate limit laws for the sequence $\{X_{n,n}\}_{n \in \mathbb{N}}$ of successive maxima of identically distributed random variables.

It turns out that the concentration of $X_{n,n}$ for large n can be determined in terms of a tail ratio of the underlying distribution function F. Applications to the outlier-behaviour of probability distributions are given.

D. MEJZLER

Limit distributions for extreme order statistics in the non-identically distributed case

Let $X_1,\ldots,X_n$ be independent random variables and let $X_{1n}\leq\ldots\leq X_{nn}$ be the corresponding order statistics. Let F_i be the distribution function (df) of X_i and let F_{kn} be the df of X_{kn}, where $k = k(n)$ is a function of n $(1 \leq k \leq n)$, which satisfies the conditions

(*) $k \to \infty$, $p = n - k \to \infty$, $p/n \to 0$.

Let S* be the class of all df's H(x), that have the following property:

There exists a sequence of random variables X_n with corresponding df's F_n, a numerical sequence b_n and positive integers k(n), that satisfy the following conditions:

a) $F_{kn}(x+b_n) \to H(x)$

b) The sequence k(n) satisfies conditions (*), and

$$\sqrt{p(n+1)} - \sqrt{p(n)} \to 0.$$

c) For every $x > \underline{H} = \inf\{x:H(x) > 0\}$ we have

$$\min_{1\leq i\leq n} F_i(x+b_n) \to 1.$$

d) The sequence $\{b_n\},\{(p(n)\}$ are related in the following way:

If for some integer-valued function $m = m(n)$

$b_{n+m} - b_n \to \beta$, $(0 < \beta < \infty)$,

then $p(n+m)/p(n) \to 1$.

Theorem: The proper df H belongs to the class S* if and only if it is of the form

$$H(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u(x)} e^{-t^2/2}\,dt,$$

where $u(x)$ is a concave function on $(\underline{H},\infty)$.

Y. MITTAL

A test for a change point

Let $X_1,\ldots,X_n$ be iid with distribution F and with mean $\mu = 0$. It is possible that $\exists\ \nu$, $1<\nu<n$ such that $\mu_1 = \ldots = \mu_\nu = 0$ but $\mu_{\nu+1} = \ldots = \mu_n = \mu > 0$, say. If so then ν is called a change point. Based on n observations $X_1,\ldots,X_n$ (n fixed) we first want to determine if there is a change point and if yes, to estimate it.

Under the null hypothesis that there is no change point, the locations $L_1,\ldots,L_k$ of the top K ($K<<n$) maxima and $l_1,\ldots,l_k$ of the lowest k minima will be asymptotically independent uniform on $1,2,\ldots,n$. We test $H_o: F_n = G_n$ by the Kolmogorov-Smirnov test where F_n and G_n are the empirical distribution functions of $\{L_i\}$ and $\{l_i\}$ respectively. If the null hypothesis is rejected then we estimate ν and study the properties of this estimator as well as the power of the test against various other alternatives.

P. RÉVÉSZ

Local time of the empirical and quantile process

Let $U_1,U_2,\ldots$ be a sequence of independent uniform (0,1) r.v.'s. Let $F_n(x)$ be the empirical distribution and $0 < U_{n:1} < U_{n:2} < \ldots < U_{n:n} < 1$ be the ordered sample based on the sample $U_1,U_2,\ldots,U_n$.

Let $\alpha_n(x) = n^{1/2}(F_n(x)-x)$ be the empirical process. The r.v. $\gamma_n = \#\{k: 1 \leq k \leq n,\ \alpha_n(U_{n:k})\alpha_n(U_{n:k+1}) < 0\}$ is called the local time of α_n. Let $\{K(x,t),\ 0 \leq x \leq 1,\ t \geq 0\}$ be a Kiefer process. The process $\eta(t) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon}\lambda\{x: 0 \leq x \leq 1,\ t^{-1/2}|K(x,t)| \leq \varepsilon\}$ is called the local time of K. (Here λ is the Lebesgue measure and the existence of the involved a.s. limit is proved.) The following invariance principle is presented.

$$|\eta(n) - n^{-1/2}\gamma_n| = O(n^{-1/4+\varepsilon}) \text{ a.s.}$$

for any $\varepsilon > 0$. As an application of this theorem one can prove, for example,

$$\limsup_{n\to\infty} \frac{\gamma_n}{\sqrt{n \log \log n}} = \frac{1}{\sqrt{2}} \text{ a.s.}$$

The same theorems are true if we consider instead of the uniform law an arbitrary continuous distribution or instead of the empirical process the quantile process.

H. ROOTZÉN

Extremes of moving averages

Let $X_t = \Sigma_\lambda c_\lambda Z_{t+\lambda}$ be an (infinite) moving average of iid r.v.'s $\{Z_\lambda\}$. A rather complete account of the behaviour of extremes of $\{X_t\}$ is given for the case when the tails of the innovations are of the form $P\{Z_\lambda > z\} \sim k\, z^\alpha e^{-z^p}$, for some $\alpha,\ p > 0$.

L. RÜSCHENDORF

Order statistics and symmetric functions

We extend the markov property of order statistics to general underlying distributions and show that the distribution of order statistics has the MTP_2 property. In the second part of

Gefördert durch DFG Deutsche Forschungsgemeinschaft

the talk we prove a representation of symmetric functions by typical symmetric functions. This result generalizes the well known Schmidt-Mercer result for functions of two variables. As application we derive a sufficient condition for the symmetric completeness of a class of product measures.

F.H. RUYMGAART

Some remarks on multivariate empirical processes

Consider the empirical process based on a sample from the uniform distribution over the d-dimensional unit cube. In joint work with Einmahl (Nijmegen) and Wellner (Seattle) it has been shown that the O'Reilly (1974) characterization of weak convergence of the weighted empirical process holds true in the multivariate case, both when the process is indexed by points and when it is indexed by rectangles. Similar results were obtained by Alexander (1982). A result due to Csâki (1974 - 1980) on the a.s. convergence of the empirical process weighted by its variance (for d = 1) has been generalized by Einmahl to arbitrary dimensions. Eventually some results on the modulus of continuity of the multivariate empirical process were mentioned.

R.L. SMITH

Estimating tails of probability distributions

This talk is concerned with estimating a distribution function F in its extreme upper (or lower) tail, based on a sample of independent observations from F.
Many authors have proposed estimators based on the largest (or smallest) k order statistics from a sample of size n. In most cases $k = k(n) \to \infty$ as $n \to \infty$. However, even now not much is known about the "optimal" choice of k(n).

In my talk I concentrate on the case $1-F(y) = y^{-\alpha}L(y)$ for $\alpha > 0$ and L slowly varying at ∞. Results of Hall (1982), based on an estimator of Hill (1975), are extended by the case of a general L satisfying a "slow variation with remainder" condition. My viewpoint differs from Hall's (and Häusler's at this meeting) in that it is not k(n) which is fixed but the threshold level x_n: thus k(n) is the random number of exceedances of x_n. A central limit theorem is stated, from which one can determine the optimal choice of x_n.

This is still not a practical result, as the optimal x_n depends on higher-order terms which are unknown. In the final part of my talk I suggest an empirical choice of x_n based on an idea of Pickands (1975). I conjecture that this achieves the optimal rate of convergence of $\hat{\alpha}$ to α.

J. STEINEBACH

An improved Erdös-Rênyi strong law for moving quantiles

Consider a sequence $U_1, U_2, \ldots$ of independent, uniformly (0,1)-distributed random variables. For fixed $\alpha \in (0,1)$, let $U_\alpha(n,K)$ denote the $[K\alpha]$-th order statistic of the subsample $U_{n+1}, \ldots, U_{n+K}$, and set $M_\alpha(N,K) = \max_{0 \le n \le N-K} U_\alpha(n,K)$, $1 \le K \le N$. Book and Truax (1976) proved the following analogue of the Erdös-Rênyi (1970) law of large numbers. For $\alpha<u<1$ and $C_\alpha = C_\alpha(u)$ such that $\exp(-1/C_\alpha) = (u/\alpha)^\alpha ((1-u)/(1-\alpha))^{1-\alpha}$, it holds

$$(*) \quad \lim_{N\to\infty} M_\alpha(N,[C_\alpha \log N]) = u \quad \text{a.s.}$$

In view of the Deheuvels-Devroye (1983) improvements of the original Erdös-Rênyi strong law, we show that the best possible convergence rate in (*) is of order $O((\log N)^{-1} \log \log N)$. Using the quantile transformation, this result can be extended to a general sequence $X_1, X_2, \ldots$ of iid r.v.'s with df. F satisfying a strict monotonicity condition.

J. TEUGELS

Order statistics in insurance mathematics

More and more the actuary is facing situations where large claims upset the statistical procedures on which he used to base his decisions concerning premium, type of reinsurance, retention, etc. We survey the acturial literature for information on large claims. We postulate a specific form for the claim size distribution and investigate its consequences. Special attention is given to reinsurance.

J. TIAGO DE OLIVEIRA

Bivariate extremes: Basics and model-building

The limiting distributions of pairs of maxima of iid samples $\{(X_i, Y_i)\}$ are described. As the margins must be Ψ_α, Λ or Φ_α only the case is considered where they are $\Lambda(x)$ and $\Lambda(y)$, denoted by $\Lambda(x,y)$. From the stability equation $\Lambda(x,y) = \Lambda^k(x+\log k, y+\log k)$, $k>0$, it follows that $\Lambda(x,y) = \exp\{-(e^{-x}+e^{-y})k(y-x)\}$, $k(\cdot)$ being called the dependence function. The Boole-Fréchet inequality together with the stability equation imply $\Lambda(x)\Lambda(y) \leq \Lambda(x,y) \leq \min\{\Lambda(x), \Lambda(y)\}$. Necessary and sufficient conditions for $k(\cdot)$ such that $\Lambda(x,y)$ is a proper and non-degenerate df are given. Necessary and sufficient conditions for attraction to the independence case ($k(w) = 1$) and the diagonal case ($\mathrm{Prob}\{Y=X\} = 1$, $k(w) = \frac{\max(1,e^w)}{1+e^w}$) are given. References to regression and correlation coefficients are given.

The mix-technique and the max-technique of generating bivariate models of extremes are described.

Finally the question of a non-parametric estimation of k(•), within the set of dependence functions is raised.

Berichterstatter: R.-D. Reiss

Tagungsteilnehmer

Prof. Dr. A.A. Balkema
Prinsengracht 844
1017 JM AMSTERDAM/NETHERLANDS

Prof. Dr. H.A. David
Dept. of Statistics
Iowa State University
Ames IA 50011/USA

Prof. Dr. Gilbert Basset, jr.
University of Illinois at
Chicao Circle
Chicago Il 606680 / USA

Prof. Dr. Paul Deheuvels
Institute de Statistique
7. Avenue du Château
F 92340 Bourg-la-Reine/France

Prof. Dr. Dennis Boos
Statistics Department North
Carolina
State University Raleigh
North Carolina 267650 / USA

Prof. Dr. H. Dinges
Angewandte Mathematik
der Universität Frankfurt
Robert-Mayer-Str. 10
D-6000 Frankfurt

Prof. Dr. M. Csörgö
Dept. of Math. & Statistics
Carleton University
Colonel-by-Drive
Ottawa Ont. Canada KIS 586

Dr. R.J.M.M. Does
Romulushof 43
6215 EJ Maastricht/Netherlands

Prof. Dr. S. Csörgö
Szeged Univ. Bolyai Inst.
Aradi Vertanuk Tere 1
H 6720 Szeged / Hungary

Prof. Dr. F. Eicker
Universität Dortmund
Abteilung Statistik
Vogelpothsweg
D - 4600 Dortmund

Prof. Dr. H.E. Daniels
Stat. Lab. Univ. of Cambridge
16 Mill Lane
Cambridge CB 21SB/Great Britain

Prof. Dr. Gunnar Englund
Department of Mathematics
Royal Institute of Technology
S-10044 Stockholm /Sweden

Dr. Michael Falk
Universität-Gesamthochschule
Siegen
Hölderlinstr. 3

D-5900 Siegen

Dr. U. Gather
RWTH Institut für Statistik
und Wirtschaftsmathematik
Wüllnerstr. 3

D-5100 Aachen

Prof. Dr. L. Györfi
Technical University of Buda-
pest
Stoczek u. 2

H-111 Budapest / Hungary

Prof. Dr. L. de Haan
Erasmus Universiteit
Faculteit der Economische
Wetenschappen
Postfach 1738

Rotterdam / Netherlands

Dr. E. Häusler
Math. Institut der
Universität München
Theresienstr. 39

D-8000 München 2

Prof. Dr. R. Helmers
Stichting Mathematisch Centrum
Tweede Boerhaavstraat 49

1091 AL Amsterdam/Netherlands

Dr. N. Henze
Universität Hannover
Mathematische Stochastik
Welfengarten 1

D-3000 Hannover 1

Prof. Dr. J. Hüsler
Universität Bern
Institut für Mathem. Statistik
und Versicherungslehre
Sidlerstr. 5

CH 3012 Bern/Switzerland

Dr. C.A.J. Klaassen
Dept. Mathematics
University of Leiden
P.O.Box 9512

2300 Leiden/Netherlands

Prof. Dr. R. Koenker
Bell Laboratories Murray Hill

New Jersey 07974 / USA

Dr. H.D. Keller
Universität Trier
FB IV
Postfach 3825

D - 5500 Trier

Dr. Wolfgang Kohne
Universität-Gesamthochschule
Siegen
Hölderlinstr. 3

D - 5900 Siegen

Prof. Dr. M. Leadbetter
University of North Carolina
Stat. Dept.

Chapel Hill NC 27514 / USA

Prof. Dr. G. Lindgren
Dept. of Math. Stat.
University of Lund
Box 725
S -220 07 Lund/Sweden

Prof. Dr. D.M. Mason
Dept. of Math.
501 Kirkbride
Univ. of Delaware

Newark DE 19711 / USA

Prof. Dr. V. Mammitzsch
Fachbereich Mathematik
der Universität Marburg
Lahnberge

D - 3550 Marburg

Dr. R. Mathar
Universität Aachen
Institut für Statistik und
Wirtschaftsmathematik
Wüllnerstr. 3

D - 5100 Aachen

Prof. Dr. D. Mejzler
Hebrew University of Jerusalem

Jerusalem / Israel

Prof. Dr. D.Y. Mittal
Department of Statistics
Virginia Polytechnic Inst.

Blacksburg VA 24061 / USA

Prof. Dr. L. Rüschendorf
Mathematisches Institut der
Universität Freiburg
Albertstr. 23 b

D - 7800 Freiburg

Prof. Dr. F.H. Ruymgaart
Mathematical Institute
Katholieke University
Toernooiveld

Nymegen / Netherlands

Prof. Dr. R.-D. Reiss
Universität-Gesamthochschule
Siegen
Hölderlinstr. 3

D - 5900 Siegen

Prof. Dr. P. Révész
Mathematical Inst. of the Hungaria
Academy of Science
Realtanod 13/15

Budapest V / Hungary

Prof. Dr. H. Rootzén
Institute of Mathematical
Statistics
5, Universitetsparken
University of Copenhagen

DK 2100 Copenhagen / Denmark

Prof. Dr. W. Sendler
Universität Trier
FB IV
Postfach 3825

D - 5000 Trier

Prof. Dr. J.L. Teugels
University of Louvain
200-B, Celestijnenlaan

B-3030 Heverlee / Belgium

Prof. Dr. R.L. Smith
Math. Dept. Huxley Bldg.
Imperial College
Queen's Gate

London SW 7 28Z/ Great Britain

Prof. Dr. J. Tiago de Oliveira
Faculdade de Ciencies da Univ
sidade de Lisboa
Dept. de Estatistica
Investigacao Opercial e Computacao
58, Rue de Escola Politecnica
1294 Lisboa Codex / Portugal

Prof. Dr. J. Steinebach
Fachbereich Mathematik
der Universität Marburg
Lahnberge

D - 3550 Marburg

Prof. Dr. W.R. van Zwet
University of Leiden
Department of Mathematics
Wassenaarseweg 80
Postbus 9512

2300 RA Leiden / Netherlands

Prof. Dr. W. Stute
Fachbereich Mathematik
der Universität Gießen
Arndtstr. 2

D - 6300 Gießen