MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Tagungsbericht 53/1992

# Asymptotische Statistik

13.12. bis 19.12.1992

Die Tagung fand unter der Leitung von F. Götze (Bielefeld) und J. Pfanzagl (Köln) statt. Sie beschäftigte sich in 35 Vorträgen und zahlreichen Diskussionen insbesondere mit neuesten Entwicklungen auf den folgenden Gebieten: Schätzen von Regressionsfunktionen und Dichten, effiziente Verfahren für semiparametrische und nichtparametrische Modelle, Statistik für stochastische Prozesse, Bootstrap. Ein Abend war einer Diskussion über die Zukunft der Statistik gewidmet.

# Vortragsauszüge

M. Akahira

## Asymptotic expansions for confidence intervals

An asymptotically unbiased confidence interval is constructed from an unbiased test up to the third order, where the second and third order derivatives of the log-likelihood function are used. And also its application to the location parameter case is described. Further, from the viewpoint of a posterior risk, the upper and lower confidence limits are derived and, in practice, obtained up to the second order in case of the normal, uniform and truncated normal distributions. The relationship between the loss function and a confidence level is also discussed. It is noted that the level can be determined from a shape of the loss function and is corrected to the length of a confidence interval.

O.E. Barndorff–Nielsen

## Some aspects of random trees

A survey was given of some recent results concerning binary rooted trees with stochastic structures. More specifically, models for river networks and for disordered electrical networks were considered.

The stochastic and statistical analysis of river networks has, within the last five to ten years, grown into a very substantial and exciting subject area. This is however not so widely known among probabilists and statisticians, because most of the important papers have been published in Earth Sciences journals, particularly in *Water Resources Research*. Much of the work relates to *Horton's Laws*, which are empirically determined regularities pertaining to order, length and drainage area of stream links (or sections). Of some special interest are a series of recent papers that establish a connection between fractality and Horton ratios. Other results describe the limiting behaviour, in stochastic river networks, of the main channel length, the width of the network, etc., the statements being conditional on the number of sources or on the order of the network.

For electrical networks with random resistances one of the questions of interest is to characterize the distributional properties of the total resistance $R$ of the network. The distribution function of $R$ can generally not be found explicitly. However, certain specifications in terms of inverse Gaussian and reciprocal inverse Gaussian distributions do allow an explicit solution.

V. BENTKUS

## Asymptotic expansions for $L^p$–norms of empirical processes

Let $B$ be a real Banach space. Let $X, X_1, X_2, \ldots \in B$ denote a sequence of i.i.d. r.v.'s with values in $B$. Denote $S_n = n^{-1/2}(X_1 + \ldots + X_n)$. Let $\pi : B \to R$ be a polynomial. We consider Edgeworth expansions for the distribution function of $\pi(S_n)$, as well as for the derivatives of the distribution function. As an application we get asymptotic expansions in the integral and local limit theorems for the general $\omega$–statistic

$$\omega_n^p(q) = n^{p/2} \int_0^1 (F_n(x) - x)^p q(x) dx.$$

Here $p = 2, 4, \ldots$, $q : [0,1] \to [0, \infty]$ is a weight function, and $F_n$ is an empirical distribution function.

P.J. BICKEL

## A new approach to testing goodness of fit

We study the properties of a method of testing for goodness of fit. This method can be viewed as a way of selecting the best among a family of tests corresponding to test statistics $\{T_j\}$, $j \geq 1$, such that the tests based on $T_j$ are "best" against alternatives in a finite dimensional parametric family $\mathcal{F}_j$. We suppose $\mathcal{F}_j \subset \mathcal{F}_{j+1}$, $j \geq 1$. The tests we propose roughly,

(i) Are powerful against low dimensional alternatives ($\mathcal{F}_j$ with $j$ small).

(ii) Perform as well as the best of the $T_j$ based tests for alternatives not belonging to $\cup_k \mathcal{F}_k$.

(iii) Are consistent against all alternatives.

This kind of approach is, in part, proposed but not studied analytically in Rayner and Best (1989).

D.M. CHIBISOV

## Chi–squared tests with large number of degrees of freedom

Let $x_1, \ldots, x_N$ be independent normally $N(\mu_i, 1)$ distributed r.v.'s, and $H_0 : \mu = (\mu_1, \ldots, \mu_N) = 0$ against $H_1 : \mu \neq 0$, $\sum \mu_i = 0$ is to be tested. It is proved that the chi-squared test based on $\chi^2_N = \sum(x_i - \bar{x})^2$ is asymptotically most powerful within the class of symmetric (permutation invariant) tests, as $N \to \infty$.

M. DENKER

**Rank statistics under dependent observations and applications to experimental designs**

This is joint work with E. BRUNNER. In a general model of the form

$$\mathbf{X}_i(n) = \left(X_{i1}(n), \ldots, X_{im_i(n)}(n)\right)'$$

with independent random vectors of (non-constant) dimension $m_i(n)$ we prove asymptotic normality of simple linear rank statistics under overall ranking. When the score function is of class $C^2[0,1]$, and very likely only in this case, such a statistic can be used in applications, i.e. the unknown variance can be estimated and the assumptions in the theorem be checked. Applications include multivariate test for symmetry, Kruskal–Wallis test under repeated measurements, Friedman test with overall ranking (and repeated measurements) and many more. The proof of the result follows known schemes.

D. DONOHO

**Minimax estimation of regression, densities, inverse problems**

In the 1980's the theory of minimax estimation of nonparametric regression, densities, and in inverse problems was vigorously developed by researchers in the Soviet Union, Europe, and the U.S. This talk will discuss very recent developments (wavelets, renormalization, infinite–dimensional nonlinear shrinkage) which expose new phenomena in nonparametric estimation, and the implications for future research.

V. FABIAN

**An optimum design for estimating the first derivative**

This is a joint work with Roy E. ERICKSON and Jan MAŘÍK (both from East Lansing, MI). The result concerns an optimum choice of step lengths in the estimate of the first derivative, used in Fabian (*Ann. Math. Statistics*, 1967, **38**, 191–200; 1968, **39**, 457–466, 1327–1333). The optimum choice leads to the minimal expected squared error of a stochastic approximation procedure. The problem is equivalent to the minimization of $\Gamma$ treated in the Theorem below.

**Theorem.** *Let $m$ be an integer, $m \geq 2$,*

$$X = \{x;\ x \in R^m,\ 0 < x_1 < -x_2 < \ldots < (-1)^{m-1}x_m,\ \prod_{i=1}^{m} |x_i| = 1\}$$

*and let $\Gamma$ be defined on $X$ by*

$$\Gamma(x) = \frac{\det[1, x^3, \ldots, x^{2m-1}]}{\det[x, x^3, \ldots, x^{2m-1}]}$$

*$x^s = (x_1^s, x_2^s, \ldots, x_m^s)$). Then the minimal value of $\Gamma$ is $m$ and is attained at exactly one point $x$, given by*

$$x_i = (-1)^{i-1} 2 \cos\left(\frac{m+1-i}{2m+1}\pi\right) \qquad \text{for } i = 1, \ldots, m.$$

*($x_i/2$ are the roots of the 2nd type Chebyshev polynomial of degree $m$.)*

M. Falk

**On testing the extreme value index via the POT–method**

Consider an iid sample $Y_1, \ldots, Y_n$ of random variables with common distribution function $F$, whose upper tail belongs to a certain neighborhood of the upper tail of a generalized Pareto distribution $H_\beta$, $\beta \in \mathbb{R}$. We investigate the testing problem $\beta = \beta_0$ against a sequence $\beta = \beta_n$ of contiguous alternatives, based on the point processes $N_n$ of the exceedances among $Y_i$ over a sequence of thresholds $t_n$. It turns out that the (random) number of exceedances $\tau(n)$ over $t_n$ is the central sequence for the log–likelihood ratio $d\mathcal{L}_{\beta_n}(N_n)/d\mathcal{L}_{\beta_0}(N_n)$, yielding it's local asymptotic normality (LAN). This result implies in particular the surprising fact that $\tau(n)$ carries asymptotically all the information about the underlying parameter $\beta$, which is contained in $N_n$. We establish sharp bounds for the rate at which $\tau(n)$ becomes asymptotically sufficient, which show however that this is quite a poor rate.

S. van de Geer

**On the application of martingale inequalities to maximum likelihood estimation**

We consider uniform exponential probability inequalities for martingales, imposing entropy conditions which are analogous to those used in empirical process theory. These inequalities can be applied in several maximum likelihood problems.

For example, let $\{N_t\}$ be a counting process with continuous compensator $\{A_t\}$, and suppose that $dA/d\mu = a(\vartheta_0)$, $\vartheta_0 \in \Theta$. Let

$$L_T(\vartheta, \vartheta_0) = \int_0^T \log\Big(\frac{a(\vartheta)}{a(\vartheta_0)}\Big) dN - \int_0^T \big(a(\vartheta) - a(\vartheta_0)\big) d\mu$$

be the log–likelihood ratio and let

$$h_T^2(\vartheta, \tilde{\vartheta}) = \frac{1}{2} \int_0^T \big(a^{1/2}(\vartheta) - a^{1/2}(\tilde{\vartheta})\big)^2 d\mu$$

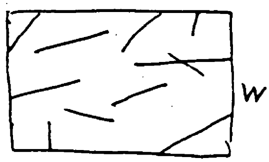
be the Hellinger process. Then one can define a function $\varphi_\sigma(b)$, expressed in terms of the entropy with bracketing of $\{a(\vartheta) : \vartheta \in \Theta\}$ endowed with metric $h_T(\vartheta, \tilde{\vartheta})$, such that for $b_* \geq \varphi_\sigma(b_*)$ we have

$$\begin{aligned} &P\big(L_T(\vartheta, \vartheta_0) \geq 0 \ \wedge \ h_T(\vartheta, \vartheta_0) > b_* \quad \text{for some } \vartheta \in \Theta\big) \\ &\leq C_1 \exp(-C_\alpha b_*^2) + P\big(A(T) > \sigma^2\big). \end{aligned}$$

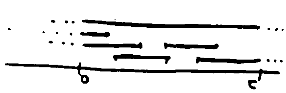
R.D. GILL

## Laslett's line segment problem

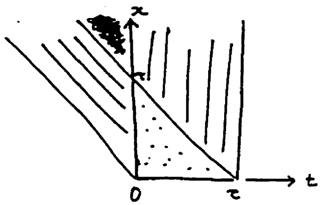
Joint with B. WIJERS.— G. LASLETT (*Biometrika* 1982) considered a number of related nonparametric estimation problems including the following: a Poisson line segment process is observed through a bounded window $W$, so that some line segments are completely observed, some are 'censored' on one or both sides:




Assuming independent line segment lengths and orientations the object is to estimate the length distribution $F$. Following LASLETT we show how the likelihood can be calculated. Computation of the NPMLE of $F$ is feasible though to derive its asymptotic properties is an open problem. We specialize to the one–dimensional case:



Here we show the NPMLE is consistent (WIJERS, 1992) using a general convexity argument and the fact that after reparametrization to a length biased version of $F$, we have a pure 'nonparametric missing data problem' in which underlying i.i.d. copies of $X,T$, where $X \sim V$ (unknown) and $T|X=x \sim \mathrm{Unif}(-x, \tau)$ are grouped onto lines or into a region or completely observed according to the following figure:




The sharp point at the top of the 'completely observed' region introduces a singularity which so far has prevented us from proving (root $n$ style) asymptotics. In fact by a result of van der Vaart root $n$ estimation of $V(\tau)$ is impossible. Some modifications are proposed which are almost efficient.

L. HEINRICH

**Normal approximation for mean–value estimates in absolutely regular Voronoi tessellations**

We consider a $d$–dimensional Voronoi tessellation $V(\Psi) = \{C_i,\ i \in \mathbb{N}\}$ generated by a stationary point process $\Psi = \{X_i,\ i \in \mathbb{N}\}$; $C_i = \{x \in \mathbb{R}^d : \|x - X_i\| \leq \|x - X_j\| \text{ for } j \neq i\}$. One of the best studied models for a random tessellation of the space $\mathbb{R}^d$ is the well-known Poisson–Voronoi tessellation which is generated by a homogeneous Poisson process $\Psi$. In order to measure the departure of an observed tessellation from a Poisson $VT$ (which often serves as a kind of gauge model) one has to find suitable characteristics and corresponding test statistics with known (approximate) distribution. The question arises how to find such a distribution. To tackle this problem we assume that the generating point process $\Psi$ satisfies a $\beta$–mixing condition (absolute regularity), e.g. in case of Poisson cluster and certain Gibbs processes, and show that the random closed set $\partial V(\Psi) = \bigcup_{i \in \mathbb{N}} \partial C_i(\Psi)$ —the skeleton of the $VT$ — and, hence, all associated point processes (nodes, midpoints of edges, circumcenters of facets) satisfy a similar absolute regularity condition.

Finally, applying some results and techniques from the limit theory of strongly mixing random fields, we obtain asymptotic normality of the proposed intensity estimators of the associated point processes. Using a suitable estimate of the asymptotic variance of these estimates (which is shown to be asymptotically unbiased and consistent) we can establish asymptotic $100(1-\alpha)\%$ confidence intervals for the intensities under consideration. Here asymptotics mean that the sampling region grows unboundedly in all directions.

R. HÖPFNER

## Estimating a parameter in a birth–and–death process model

We deal with estimation of an unknown one–dimensional parameter $\vartheta$ ranging over $\Theta = (-C, +\infty)$, in a particular birth-and-death process model where the observed process is either transient (case $\vartheta > 1$), positive recurrent (case $\vartheta < 0$), or recurrent null (case $0 \leq \vartheta \leq 1$). It is known that a certain random observation scheme establishes local asymptotic normality at all points $\vartheta \in \Theta$ of the model, everywhere with same local scale $1/\sqrt{n}$. We construct and discuss different families of estimator sequences for the unknown parameter. Some of these sequences, being regular in the sense of Hájek at all points $\vartheta \in \Theta$, fail to be efficient in the sense of the convolution theorem, on different subsets of $\vartheta$.

A. JANSSEN

## Recent results for Kolmogorov–Smirnov tests and related tests

The first part of the talk deals with two–sample goodness of fit tests of Kolmogorov–Smirnov, Cramér von Mises and Anderson–Darling type when ties are present. Two methods are presented in order to obtain valid (asymptotically) $\alpha$–similar tests. Also the power function is calculated in direction of non–parametric tangent vectors.

The second part deals with the local comparison of different tests. Each non–parametric unbiased test has a principal component decomposition of the curvature of the power function given by a Hilbert-Schmidt operator. Thus every non–parametric test has reasonable curvature only for a finite number of orthogonal directions of alternatives. As application one obtains results about the curvature of the two–sided Kolmogorov–Smirnov tests. It is shown that these tests prefer for small $\alpha$ approximately the same direction as the two–sample median rank test. The results are analogous to earlier results of Hájek and Šidák for one–sided Kolmogorov–Smirnov tests.

P. JEGANATHAN

**On the calculation of Bartlett correction in time series models**

Computing the Bartlett correction for the likelihood ratio statistic involves the computation of approximate bias of the log–likelihood ratio. This may be difficult in time series models. An alternative approach to the calculation is suggested. The approach is related to the one that uses a formula for the conditional distribution of MLE given the ancillaries in the case of transformation models.

J.L. JENSEN

**On asymptotic normality of pseudo likelihood estimates for pairwise interaction processes**

Due to the possibility of phase transitions it is not possible to prove asymptotic normality of maximum likelihood estimates in Gibbs point processes for all values of the parameters defining the process. The problem is to get a bound on the mixing coefficients of the process. Contrary to this, asymptotic normality can be proved for the maximum pseudo likelihood estimates without using the mixing properties. Instead one uses ergodicity and a property that resembles that of a martingale difference scheme, i.e. the score function is a sum of terms each of which has mean zero conditioned on the surroundings. The final result says that it is possible to calculate a stochastic norming of the pseudo likelihood estimate such that the normed estimates have asymptotically a standard normal distribution for a stationary Gibbs point process. The talk is based on joint work with Hans R. Künsch.

W.C.M. KALLENBERG

**Accurate test limits with estimated parameters**

Joint work with A. ALBERS and G.D. OTTEN (Enschede). — Due to measurement errors, procedures are typically forced to set test limits well within specification limits. The methods used in practice are rather informal and usually conservative with respect to consumer loss, thus leading to unnecessary loss of yield. We present approximations for test limits which are still relatively easy to evaluate and moreover very accurate. In addition, the analytical tractability of these approximations allows extension to the more realistic case where parameters are estimated.

C.A.J. KLAASSEN

**Transformation models**

A general class of semiparametric transformation models is considered. A second order differential equation is derived for estimation of the real parameters involved. A frailty model of Clayton and Cuzick for survival data is studied in some detail. It is shown that the information bounds are sharp by constructing an estimator attaining it.

H.R. KÜNSCH

**Linking blocks in the bootstrap for stationary observations**

Joint with E. CARLSTEIN, Chapel Hill. — We consider the problem of estimating the distribution of a statistic when the underlying observations are stationary. A truly model free procedure for this is the blockwise bootstrap which resamples independent blocks of consecutive observations. For consistency it only requires mixing and moment conditions. Still, for finite $n$ the bootstrap variance has a bias even in the case of the mean. In order to reduce this bias, we propose to link the blocks by choosing their starting points according to a Markov chain. The transition probabilities are determined by assuming a partially specified model for the observations, e.g. an ARMA– or a Markov model. The analysis of this method leads to the study of $U$–statistics with kernels depending on the sample size.

H. LÄUTER

**Bootstrapped nonlinear estimators**

We are looking for strong LLN for bootstrapped statistics. We give conditions for the uniform almost sure convergence for bootstrapped normalized weighted sums of variables where the coefficients depend on parameters. This result we use for the proof of the strong LLN for the bootstrapped least squares estimator in nonlinear distribution families. As a second problem we use the bootstrap distribution for the BIAS reduction of nonlinear estimators. If convex parameters are to be estimated then with this procedure we improve estimators in the sense of the mean squared error.

R.C. LIU

**Geometry in non–parametrics with censoring**

In this talk, we use a geometric approach to:

(1) identify optimal rates of convergence in estimating non–parametric functionals (such as density, hazard) with censoring

(2) construct optimal rate estimators through a geometric quantity and $F/R/W$ procedure

(3) construct nearly best estimators, which are, in general, within 25% to minimax with censoring

simply, for unified and general set up.

Our general tool is a geometric smoothness measure called "modulus of continuity". The issued (1) and (2) were not completely answered in the literature; and (3) was never studied before.

E. MAMMEN

**Bootstrap tests for multimodality**

Some results are presented for the expected number of modes of kernel density estimates. These results are applied for the study of test statistics for multimodality based on kernel density estimates. Asymptotic results are given which suggest that bootstrap can be used for achieving critical values and that these bootstrap tests are conservative.

R. NORVAIŠA

**Asymptotic behavior of distributions induced by the empirical processes on function spaces**

Consider a (centered) $L$–statistic $L_n$ given by the integral representation

$$L_n - \mu = \int_{\mathbb{R}} \int_{F(t)}^{F_n(t)} J(s) ds g(dt) =: \int_{\mathbb{R}} \Phi_J F_n dg,$$

where $g$ is an indefinite integral, $J$ is a score function and $F_n$ is an empirical d.f. based on a sample from the arbitrary d.f. $F$. One can consider $\Phi_J$ as a Nemytskij operator acting between Banach function spaces. The main statement is that the task of an asymptotic behavior of $L_n$ reduces to the asymptotic behavior of the empirical d.f. $F_n$ and smoothness properties of $\Phi_J$. We plan to illustrate this for the consistency and asymptotic normality questions. In this way we improve previously known results on this subject.

M. NUSSBAUM

**Asymptotic equivalence of density estimation and white noise**

We consider the question of asymptotic equivalence of the density estimation experiment, where $y_i$, $i = 1, \ldots, n$ are observed i.i.d. random variables with values in $[0, 1]$ having density $f$, with the white noise model. The latter one is given by observations

$$dy(t) = f(t)dt + n^{-1/2}r^{1/2}(t)dW(t), \qquad t \in [0,1]$$

where $dW(t)$ is standard Gaussian white noise, $r(t)$ is a variance function and $f$ the unknown drift function. Equivalence is construed in the sense of LeCam's deficiency distance between experiments. We first demonstrate that if $f$ varies in a certain shrinking neighborhood of some density $f_0$, then the i.i.d. experiment is asymptotically equivalent to a white noise model with variance function $f_0$. We then extend this local result to a global one by using a preliminary estimator $\hat{g}$ of $f_0$, based on a fraction of the sample, and by introducing a compound experiment where the first fraction of the i.i.d. sample is retained, and the second fraction is substituted by a white noise model with estimated variance function. That yields global asymptotic equivalence if $f$ varies in a set $\Sigma^*$, which can be specified to be genuinely global and nonparametric. This allows deducing asymptotic risk bounds for density estimation from the white noise model.

G. PFLUG

**Asymptotics of stochastic optimization problems**

We consider an optimization problem of the form

$$\text{(P)} \qquad E(H(x,\xi)) + \psi_C(x) = \min! \quad \text{where } \psi_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

and its "empirical" version

$$(\mathrm{P}_e) \qquad \frac{1}{n}\sum_{i=1}^{n} H(x,\xi_i) + \psi_C(x) = \min!$$



Gefördert durch DFG Deutsche Forschungsgemeinschaft

Let $x^*$ be the solution of $(P)$ and $\hat{X}_n$ be the solution of $(\mathrm{P}_\epsilon)$. Set

$$T_n = \Gamma_n^{-1}(\hat{X}_n - x^*).$$

We want to study the asymptotic distribution of $T_n$ for a suitably chosen sequence of regular matrices $\Gamma_n \to 0$.

By change of coordinates

$$T_n = \arg\min \varrho_n \sum_{i=1}^{n} \left[H(x^* + \Gamma_n t, \xi_i) - H(x^*, \xi_i)\right] + \psi_C(x^* + \Gamma_n t)$$

for $\varrho_n > 0$, arbitrary. For a variety of cases, the epi–limit in distribution of the process

$$Z_n = \varrho_n \sum_{i=1}^{n} \left[H(x^* + \Gamma_n t, \xi_i) - H(x^*, \xi_i)\right] + \psi_C(x^* + \Gamma_n t)$$

may be found. The limit of $\psi_C(x^* + \Gamma_n t)$ depends on the local curvature of the set $C$ near $x^*$. If $\Gamma_n$ is a multiple of the unit matrix, then the limiting constraint set is the tangent cone. If $\Gamma_n$ is different in different directions, other limiting constraint sets may appear. The stochastic process

$$\varrho_n \sum_{i=1}^{n} \left[H(x^* + \Gamma_n t, \xi_i) - H(x^*, \xi_i)\right]$$

converges typically to a process of the form $D(t) + S(t)$, where $D(t)$ is some deterministic function and $S(t)$ is a self–similar zero mean stochastic process. We discuss cases where $S(t) = t' \cdot Y$; $Y \sim N(0, \Sigma)$ as well where $S(t)$ is a generalized Wiener process or a generalized Poisson process.

Y. Ritov

**Estimating the expectation of a discounted reward process**

Let $\vartheta = E\Sigma\lambda^{t-1}R_t$ for an unknown random process $R_t$. We look for unbiased estimators efficient under white noise models and show that under some model, geometric stopping time and equal weights yield the efficient design. Strange things happen under other models.

L. RÜSCHENDORF

## Asymptotics of algorithms

We determine the asymptotic distribution for some examples of stochastic recursive algorithms. The proof based on a contraction technique consists of three steps. First determine the stable limiting equation of a normalized version of the recursion. Secondly choose a probability metric which leads to contraction properties of the operator describing the limiting equation. This metric has to reflect the structure of the recursion. Thirdly decompose the normalized recursion into one part which converges to zero and one part which approximates the limiting equation. The examples discussed include sorting algorithms, trie algorithms, search algorithms, the bootstrap estimator and iterated function systems. The talk is based on joint work with S.T. RACHEV.

A. SAMAROV

## Nonparametric functional estimation

In the context of nonlinear–nonparametric regression, we consider estimation of the Pearson correlation ratio defined as $\eta^2 = \mathrm{Var}(m(\mathbf{X}))/\mathrm{Var}(Y)$, where $m(\mathbf{X}) = E(Y|\mathbf{X})$. We show that, under certain conditions which include sufficient smoothness of $m(\mathbf{x})$, there exist $n^{1/2}$–consistent, asymptotically normal estimators of that functional, which are, in fact, asymptotically efficient in terms of nonparametric information bound of Koshevnik and Levit (1976). Together with asymptotic properties, we address the issue of data–dependent selection of smoothness parameter (bandwidth), and also use estimation of functionals closely related to $\eta^2$ for assessment of (non)linearity of regression and for measuring relative importance of subsets of predictor variables.

A. SCHICK

## On efficient estimation in semiparametric regression

The problem of constructing efficient estimates of the finite dimensional parameter in regular semiparametric regression models with unknown error and covariate distributions is discussed. The efficient influence function is derived and shown to depend on the projection of characteristics of the model onto some subspace. It is then shown how to construct an efficient estimate if appropriate estimates of the regression function and the projection are available. Special care has to be taken when estimating the projection. An example is presented in which the projection can be calculated explicitly and a plug–in estimate does not work under minimal conditions. It is shown that empirical projection estimates work in this case.

I.M. SKOVGAARD

**Saddlepoint and Laplace type expansions — an overview**

Saddlepoint approximations are known as a highly accurate form of asymptotic expansion of a density (DANIELS, 1954) or of a distribution function (ESSCHER, 1932, LUGANNANI & RICE, 1980). Basic results concern sums of i.i.d. random variables, but applications often go further.

An overview of the possibilities of deriving saddlepoint expansions for various statistics is given. This starts from the basic derivation of such an expansion, for example for sums of non–i.i.d. random variables and continues with the possible operations leading from one such expansion to another. These operations include conditioning, marginalization, non–linear transformations that are not necessarily one–to–one, and one–dimensional integration. This part of the talk is closely related to the theory of Wiener germs by H. DINGES.

Applications in statistics are discussed, in particular the difficulties related to expansions of distributions of test statistics.

H. STRASSER

**Random and incidental nuisance parameters**

Consider a model with structure parameter $\vartheta$ and nuisance parameter $\eta$. Suppose that $\vartheta \in \mathbb{R}$, $\eta \in (0,1)$, and that $\eta$ is governed by a uniform distribution. Then for a loss function $W$ there is a bound $\beta_W$ such that the following holds:

If the risks of a permutation invariant estimator sequence $(S_n)$ are asymptotically seldom worse than $\beta_W + \varepsilon$ for all $\varepsilon > 0$, then they are asymptotically seldom better than $\beta_W - \varepsilon$ for all $\varepsilon > 0$.

A. VAN DER VAART

**Bracketing smooth functions: new Donsker classes**

Let $X_1, X_2, \ldots$ be i.i.d. random elements with distribution $P$. A class $\mathcal{F}$ of measurable functions is called a Donsker class if the sequence of empirical processes $f \to n^{-1/2} \sum_{i=1}^{n} (f(X_i) - Pf)$ converges in distribution to a tight Gaussian process in the space of bounded functions from $\mathcal{F}$ to the real line. Presently there are two types of simple sufficient conditions for a class to be Donsker: in terms of random or uniform entropies (applicable to VC–classes) and in terms of $L_2$–bracketing entropy. For the latter one needs upper bounds on the number of brackets that are necessary to cover $\mathcal{F}$. Such bounds are classical for smooth functions on compact subsets of Euclidean space. We derived sharp upper bounds for classes of smooth functions on unbounded sets.

This study was motivated by problems of infinite dimensional maximum likelihood or $M$ estimation. $M$ estimators can be shown to be asymptotically normal by expanding and inverting a set of likelihood equations. This involves two main technical problems: to show that a certain derivative operator is continuously invertible and to control remainder terms of the expansion. The latter can be accomplished by showing that certain classes of functions are Donsker.

W. WEFELMEYER

## Quasi–likelihood models and efficient estimation

Consider an ergodic Markov chain on the real line, with parametric models for the conditional mean and variance of the transition distribution. Such a setting is an instance of a quasi–likelihood model. The customary estimator for the parameter is the maximum quasi–likelihood estimator. We show:

1. The maximum quasi–likelihood estimator is as good as the best estimator that ignores the model for the conditional variance.

2. There is an estimator which is as good as the maximum quasi–likelihood estimator if the conditional variance is correctly specified, and strictly better, and efficient, if it is not.

3. An efficient estimator in the quasi–likelihood model is given as a weighted nonlinear one–step least squares estimator, with weights involving predictors for the third and fourth contered conditional moments.

J.A. WELLNER

## Copula models

*A parametric copula model* $\{C_\vartheta : \vartheta \in \Theta\}$ is a parametric family of distribution functions on the unit square with uniform $(0,1)$ marginal distributions: $C_\vartheta(u,1) = u$, $C_\vartheta(1,v) = v$, $u, v \in [0,1]$. The Archimedean copulas are of the form $C(u,v) = \psi^{-1}\big(\psi(u)+\psi(v)\big)$, and have an interpretation in terms of frailty models when $\psi^{-1}(u) = Ee^{-uW}$ is completely monotone.

*Semiparametric copula models* can be obtained from any particular parametric copula models by composition with arbitrary marginal distributions:

*Model 1*

$\mathcal{P}_1 = \{P_{\vartheta,G} : P_{\vartheta,G}$ has d.f. $F_{\vartheta,G}(s,t) = C_\vartheta\big(G(s),t\big),\ \vartheta \in \Theta,\ G \in \mathcal{G}\}$.

*Model 2*

$\mathcal{P}_2 = \{P_{\vartheta,G,H} : P_{\vartheta,G,H}$ has d.f. $F_{\vartheta,G,H}(s,t) = C_\vartheta\big(G(s),H(t)\big),\ \vartheta \in \Theta,\ G \in \mathcal{G},\ H \in \mathcal{G}\}$.

In this talk I discussed information bound theory for semiparametric copula models with emphasis on "model 2", the case of two unknown marginal distributions. In this case the efficient score function $\ell_\vartheta^*$ for estimation of $\vartheta$ is $l_\vartheta^* = \dot\ell_\vartheta - \dot\ell_g a_* - \dot\ell_h b_*$ where $A_*' \equiv a_*$, $B_*' \equiv b_*$ are determined by the coupled differential equations

$$(*)\qquad \begin{aligned} A_*'' - \alpha A_*(u) &= -\gamma(u) - \int_0^1 B_*(v)K(u,v)dv \\ B_*'' - \beta B_*(v) &= -\delta(v) - \int_0^1 A_*(u)K(u,v)du \end{aligned}$$

where

$$\begin{aligned} \alpha(u) &= E(\dot\ell_u^2 \mid U = u), & \gamma(u) &= E(\dot\ell_\vartheta \dot\ell_u \mid U = u) \\ \beta(v) &= E(\dot\ell_v \mid V = v), & \delta(v) &= E(\dot\ell_\vartheta \dot\ell_v \mid V = v) \end{aligned}$$

$$K(u,v) = \ddot\ell_{uv}(u,v)C_\vartheta(u,v),$$
$$\dot\ell_\vartheta(u,v) = \frac{\partial}{\partial\vartheta}\log C_\vartheta(u,v).$$

For most families $C_\vartheta$ the equations $(*)$ do not have an explicit solution. However, in the special case of the bivariate normal copula family

$$\{C_\vartheta(u,v) = \Phi_\vartheta\big(\Phi^{-1}(u),\Phi^{-1}(v)\big) : -1 < \vartheta < 1\},$$

the equations $(*)$ have an explicit solution leading to the efficient influence function

$$\ell_\vartheta^*(s,t) = \frac{1}{(1-\vartheta^2)^2}\{xy - \frac{\vartheta}{2}(x^2+y^2)\}\big|_{x=\Phi^{-1}(G(s)),\ y=\Phi^{-1}(H(t))}$$

and the efficient information for $\vartheta$ is $I_\vartheta^* = (1-\vartheta^2)^{-2}$, which is exactly the same as for estimation of $\vartheta$ in the bivariate normal submodel. Thus the bivariate normal submodel is least favorable. The efficient score equation $0 = \mathbb{P}_n \hat{\ell}^*_\vartheta = n^{-1} \sum_1^n \hat{\ell}^*_\vartheta(S_i, T_i)$ leads to the normal scores rank correlation coefficient

$$\hat{\vartheta}_n = \sum_1^n \Phi^{-1}\left(\frac{i}{n+1}\right)\Phi^{-1}(R_{ni}/(n+1)) \Big/ \sum_1^n \Phi^{-1}\left(\frac{i}{n+1}\right)^2$$

as an estimator of $\vartheta$, and it follows from Ruymgaart, Shorack and van Zwet (1972) that $\hat{\vartheta}_n$ is asymptotically efficient:

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \rightarrow dN\big(0, (1-\vartheta^2)^2\big).$$

W.R. van Zwet

**Asymptotics for plug–in estimators with application to the bootstrap**

Let $X_1, X_2, \ldots$ be i.i.d. with (unknown) common distribution $P \in \mathcal{P}$. If $\hat{P}_N = p_N(X_1, \ldots, X_N)$ is a sequence of estimators of $P$ with values in $\mathcal{P}$, we can estimate a "parameter" sequence $\tau_N(P)$ by the plug–in estimators $\tau_N(\hat{P}_N)$. We show that under a variety of continuity conditions on $\tau_N$ and convergence assumptions on $\hat{P}_N$, the plug–in estimators are indeed consistent. The results are proved in sufficient generality to be directly applicable to the bootstrap.

Berichterstatter: W. Wefelmeyer

Tagungsteilnehmer

Prof.Dr. Masafumi Akahira
Institute of Mathematics
Tsukuba University

Ibaraki 305
JAPAN

Prof.Dr. Manfred Denker
Institut für Mathematische
Stochastik
Universität Göttingen
Lotzestr. 13

W-3400 Göttingen
GERMANY

Prof.Dr. Ole E. Barndorff-Nielsen
Afd. for Teoretisk Statistik
Matematisk Institut
Aarhus Universitet
Ny Munkegade

DK-8000 Aarhus C

Prof.Dr. David L. Donoho
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720
USA

Prof.Dr. Vidmantas J. Bentkus
Institute of Mathematics
and Mechanics
Lithuanian Academy of Sciences
Akademijos Str. 4

232600 Vilnius
LITHUANIA

Prof.Dr. Vaclav Fabian
Dept. of Statistics and Probability
Michigan State University

East Lansing , MI 48824-1027
USA

Prof.Dr. Peter J. Bickel
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720
USA

Prof.Dr. Michael Falk
Mathematisch-Geographische Fakultät
Universität Eichstätt
Ostenstr. 26-28

W-8078 Eichstätt
GERMANY

Prof.Dr. Dmitrii M. Chibisov
Steklov Mathematical Institute
MIAN
Academy of Sciences
42, Vavilova str.

Moscow 117 966 GSP-1
RUSSIA

Dr. Sara van de Geer
Mathematisch Instituut
Rijksuniversiteit Leiden
Postbus 9512

NL-2300 RA Leiden

Prof.Dr. Richard D. Gill
Mathematical Institute
University of Utrecht
6 Budapestlaan

NL-3584 DC Utrecht

Prof.Dr. Friedrich Götze
Fakultät für Mathematik
Universität Bielefeld
Postfach 10 01 31

W-4800 Bielefeld 1
GERMANY

Dr. Lothar Heinrich
Fachbereich Mathematik
Institut für Stochastik
Bergakademie Freiberg
Bernhard-von-Cotta-Str. 2

O-9200 Freiberg
GERMANY

Dr. Reinhard Höpfner
Laboratoire de Probabilités
Université P. et M. Curie
Tour 56, 3e étage
4, place Jussieu

F-75252 Paris Cedex 05

Prof.Dr. Arnold Janssen
Mathematisches Institut
Heinrich-Heine-Universität
Universitätsstraße 1

W-4000 Düsseldorf 1
GERMANY

Dr. Paramaraj Jeganathan
Department of Statistics
University of Michigan
1444 Mason Hall

Ann Arbor , MI 48109-1027
USA

Dr. Jens Ledet Jensen
Matematisk Institut
Aarhus Universitet
Ny Munkegade
Universitetsparken

DK-8000 Aarhus C

Prof.Dr. Wilbert C.M. Kallenberg
Department of Applied Mathematics
Twente University
P.O.Box 217

NL-7500 AE Enschede

Prof.Dr. Chris A. J. Klaassen
Department of Mathematics and
Computer Science
University of Amsterdam

NL-1018 TV Amsterdam

Prof.Dr. Hans Rudolf Künsch
Seminar f. Statistik
ETH-Zentrum
Sonneggstr. 33

CH-8092 Zürich

Prof.Dr. Henning Läuter
Fachbereich Mathematik
Universität Potsdam
Am Neuen Palais 10

O-1571 Potsdam
GERMANY

Prof.Dr. Richard C. Liu
Dept. of Mathematics
Cornell University
White Hall

Ithaca , NY 14853-7901
USA

Dr. Enno Mammen
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294

W-6900 Heidelberg 1
GERMANY

Dr. Hartmut Milbrodt
Fachbereich Mathematik
Universität Dortmund
Postfach 50 05 00

W-4600 Dortmund 50
GERMANY

Prof.Dr. Rimas Norvaisa
Institute of Mathematics and
Informatics
Lithuanian Academy of Sciences
Akademijos Str. 4

232600 Vilnius
LITHUANIA

Dr. Michael Nussbaum
Institut für Angewandte Analysis
und Stochastik
Hausvogteiplatz 5-7

O-1086 Berlin
GERMANY

Prof.Dr. Johann Pfanzagl
Mathematisches Institut
Universität zu Köln
Weyertal 86-90

W-5000 Köln 41
GERMANY

Prof.Dr. Georg Pflug
Institut für Statistik und
Informatik
Universität Wien
Universitätsstr. 5/3

A-1010 Wien

Prof.Dr. Ya'acov Ritov
Department of Statistics
The Hebrew University of Jerusalem
Mount Scopus

Jerusalem 91905
ISRAEL

Prof.Dr. Ludger Rüschendorf
Institut für Mathematische
Statistik
Universität Münster
Einsteinstr. 62

W-4400 Münster
GERMANY

Prof.Dr. Alexander Samarov
Sloan School of Management
MIT E40-131
1 Amherst Street

Cambridge , MA 02139
USA

Aad W. van der Vaart
Mathematisch Instituut
Vrije Universiteit
De Boelelaan 1081a

NL-1081 HV Amsterdam

Prof.Dr. Anton Schick
Dept. of Mathematical Sciences
State University of New York
at Binghamton

Binghamton , NY 13902-6000
USA

Dr. Wolfgang Wefelmeyer
Mathematisches Institut
Universität zu Köln
Weyertal 86-90

W-5000 Köln 41
GERMANY

Prof.Dr. Ib. Skovgaard
Department of Mathematics
The Royal Veterinary & Agricultural
University

DK-1871 Frederiksberg C

Prof.Dr. Jon A. Wellner
Department of Statistics
University of Washington

Seattle , WA 98195
USA

Prof.Dr. Helmut Strasser
Fakultät für Mathematik
Universität Wien
Dr.Karl Lueger-Ring 1

A-1010 Wien

Prof.Dr. Willem R. van Zwet
Mathematisch Instituut
Rijksuniversiteit Leiden
Postbus 9512

NL-2300 RA Leiden