MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Tagungsbericht 10/96

Mathematische Stochastik

10.03. – 16.03.1996

The meeting was organized by Peter Bickel (UC Berkeley), Ursula Gather (Dortmund), and Friedrich Liese (Rostock).

The major areas of emphasis were

- Semiparametric models, censoring models
- Robust methods
- Inference and probabilistic analysis of complex stochastic structures.

New results were presented e.g. on constructing stable but still effective methods of estimation in semiparametric models and on the analysis of efficient methods in special models of importance in biostatistics.

Further highlights were discussions of application and implementation of high breakdown methods, clustering methods and jackknife methods. Stimulating presentations concerned graphical models for vector processes, the contact process, network analysis and modelling polymer growth.

Many novel topics were addressed such as quantum statistics, probabilistic algorithms for the travelling salesman and other problems. New insight was also gained into the normal approximation of the t-statistics, goodness of fit tests and martingale - inequalities.

All talks were well attended, discussion was lively and collaborations progressed and were initiated.

The following abstracts presented by the authors show the full range of the interesting topics which have been treated at the conference.

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

VORTRAGSAUSZÜGE:

P. BICKEL:

What is a linear process?

We ask what kinds of processes (stationary) can be distinguished from MA(p) processes:

(1) $x_t = \sum_{j=0}^{p} \psi_j \varepsilon_{t-j}$

where ε_t are i.i.d. F, $E\varepsilon_1 = 0$, $E\varepsilon_1^2 < \infty$ and $(\psi_0, \ldots, \psi_p)$, p, F are unknown, using realizations of arbitrary length n. That is, we characterize the closure of the set of probabilities on R^∞ induced by $\{ x_t \}$ as in (1). This closure is in terms of Mallows (2) convergence of finite dimensional distributions. Our analysis leads to the conclusion that to meaningfully speak of the hypothesis of linearity for a stationary stochastic process (ergodic) we must specify restrictions on the order of processes that are admissible in terms of the length n of realizations that is available. We make similar analyses for AR(p) processes,

(2) $x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + \varepsilon_t$, ε_t i.i.d. as above.

This is joint work with P. Bühlmann.

L. BIRGÉ:

An adaptive compression algorithm for Besov spaces

We present a new algorithm for the compression of a function f belonging to an unknown Besov space $B^{\alpha}_{p,\infty}[0,1]$. Assuming that the function f has a proper expansion of the form $f = \bar{f} + \sum_{j,k} a_{j,k}\varphi_{j,k}$ (wavelet or spline expansion) we want to get a compressed version of f of the form $\tilde{f} = \bar{f} + \sum_{(j,k)\in\Lambda} a_{j,k}\varphi_{j,k}$ with $|\Lambda| = D$ is given. One wishes that $\tilde{f}$ provides approximately the best approximation with respect to the $\mathbb{L}_q$-norm, for a given value of D. It is known that provided that $\alpha > \frac{1}{p} - \frac{1}{q}$ the best rate of approximation is of order $D^{-\alpha}$. One provides a new algorithm which achieves such a rate and is such that $\tilde{f}$ depends only on D, but not on any knowledge of α, p or q.

This is joint work with P. Massart.

E. BOLTHAUSEN:

Random perturbations of random walks

For a random walk $S_0 = 0, S_1, \ldots, S_n$ on $\mathbf{Z}^d$ with law P and a function $W: \mathbb{N} \times \mathbf{Z}^d \to \mathbb{R}$, a perturbation of P is defined by

$$\hat{P}_{n,W}(S) = \exp\left\{\lambda \sum_{j=1}^{n} W(j, S_j)\right\} P(S)/Z_{n,W},$$

where $Z_{n,W}$ is the appropriate norming. A random perturbation arises when W itself is random, with law $\mathbb{P}$.

The best known case is when the $W(j,x)$ are all i.i.d., the directed polymer in random environment. Most of the questions one is interested in, e.g. the scaling behavior of the endpoint S_n for large n, are still unsolved. An easier model is $W(j,x) = W(j)f(x)$ where $W(j)$ are i.i.d. and f is a deterministic function, which is monotone and bounded. We take as a special case $f(x) = \mathrm{sign}(x)$, $W(j) = \pm 1$, $\mathbb{P}(W(j) = 1) = h$. Sinai proved recently in the symmetric case $h = 0$ that the path localizes in a strong sense for all $\lambda > 0$. In a recent joint paper with Frank den Hollander it is proved that there is a transition from a localized phase for small h to a delocalized for large h at a critical point $h_c(\lambda)$. We also discuss the behavior of the function $\lambda \to h_c(\lambda)$.

R.J. CARROLL:

Nonparametric regression via local estimating equations

We show that essentially all previously defined nonparametric function estimation problems can be phrased as solutions to locally weighted estimating equations based on an unbiased estimating function. We derive asymptotic theory in this general context. These methods generally rely on a tuning constant (span in loess, local bandwidth for kernels), and we show how to estimate these tuning constants without direct reference to the asymptotics. We illustrate the power of the approach by consideration of a number of problems for which nonparametric function estimates have not been addressed previously.

This is a joint work with D. Ruppert and A.H. Welsh.

R. DAHLHAUS:

Graphical interaction models for multivariate time series

The concept of graphical interaction models for multivariate data is extended to multivariate time series and point processes. We define an edge in the graph by using the partial spectral coherence between two components of the time series given all remaining components, i.e. by looking at the correlation structure of the two series after the linear effects of all other series have been removed. It is proved that these partial coherences can be obtained by inversion and rescaling of the spectral matrix of the process. By using this result a separation theorem for time series graphs is proved. The method is applied to air pollution data, neural nets and the Australian stock exchange.

L.T. FERNHOLZ:

Target estimates for bias and variance reduction

Given a parametric family and a statistical functional T with finite expectation we define the corresponding target estimate. We show that under mild conditions the target estimate will have smaller bias and variance. The von Mises expansion of T is used to explore further these issues.

S. VAN DE GEER:

A maximal inequality for martingales

Let $(\Omega, F, \mathbb{P})$ be a probability space, and $F_0 \subset F_1 \subset \ldots$ an increasing sequence of sub-σ-algebras of F. We consider F_i-measurable random variables $Y_{i,\theta}$, $i=1,\ldots,n$, indexed by a parameter $\theta \in \Theta$, with $E(Y_{i,\theta} \mid F_{i-1}) = 0$, $i=1,\ldots,n$, $\theta \in \Theta$. If for some constant K, $|Y_{i,\theta}| \le K$, $i=1,\ldots,n$, $\theta \in \Theta$, then by Bernstein's inequality, for $a \le b^2/K$,

$$P\left(\left|\sum_{i=1}^{n} Y_{i,\theta}\right| \ge a \wedge V_{n,\theta} \le b^2\right) \le 2 \exp\left\{\frac{-a^2}{4b^2}\right\},$$

where $V_{n,\theta}$ is the predictable variation of $\sum_{i=1}^{n} Y_{i,\theta}$. We present conditions on a, b and Θ, such that

$$P\left(\left|\sum_{i=1}^{n} Y_{i,\theta}\right| \ge a \wedge V_{n,\theta} \le b^2 \text{ for some } \theta \in \Theta\right) \le C \exp\left\{\frac{-a^2}{C^2 b^2}\right\},$$

for some universal constant C. The assumption $|Y_{i,\theta}| \le K$, $i = 1,\ldots,n$, $\theta \in \Theta$, can be weakened, using higher order variational processes.

As application, we consider maximum likelihood and auto-regression. For example, suppose we have real-valued observations $X_0, X_1, \ldots, X_n$. Let $F_i = \sigma(X_0, \ldots, X_i)$, and suppose that the conditional density of X_i given X_{i-1}, with respect to Lebesgue measure on $\mathbb{R}$, is equal to $\theta_0(X_i - X_{i-1})$, with

$$\theta_0 \in \Theta = \left\{\theta: \mathbb{R} \to [0,\infty), \int \theta(x)dx = 1, \theta(x) \uparrow \text{ for } x < 0, \theta(x) \downarrow \text{ for } x > 0\right\}.$$

Let $\hat{\theta}_n$ be the maximum likelihood estimator of θ_0. We show that for all T sufficiently large

$$\mathbb{P}\left(\int \left(\sqrt{\hat{\theta}_n(x)} - \sqrt{\theta_0(x)}\right)^2 dx > T^2 n^{-2/3}\right) \le C \exp\left\{-T^2 n^{1/3}/C^2\right\},$$

provided that for some $m>0$ $\overline{\lim}_{|x|\to\infty} |x|^m \theta_0(x) < \infty$, $\overline{\lim}_{|x|\to 0} |x|^{1/m} \theta_0(x) < \infty$.

R.D. GILL:

Towards quantum statistics

Technological advances are giving great impetus to quantum computing and quantum cryptography. Small quantum systems are becoming directly observable. Quantum phenomena are intrinsically random but many experts, both from physics and from probability theory, have claimed that 'a different kind of probability' is involved.

I give an introduction to quantum probability with a view toward (coming) statistical applications. In particular I discuss old and new results on hidden variables models; these are models which attempt to provide a classical deterministic explanation for the random phenomena observed at the quantum level. Finally I draw attention to striking phenomena in quantum statistics, such as: quantum randomisation allows one to extract more information, not less.

Going back to hidden variables models, a new contribution is a short geometric proof (joint work with M. Keane) of the fundamental theorem of Kochen and Specker (1967). This theorem shows that any hidden-variables theory for quantum measurement (on an at least three-dimensional system) must be contextual: i.e., in a deterministic theory, randomness is explained not just by hidden states in the quantum system under study but also from hidden states in the measurement devices.

F. GÖTZE:

When is Student's statistic asymptotically standard normal?

The asymptotic normal approximation is investigated for Student's statistic and selfnormalizing statistics like $t_N := (X_1+\ldots+X_N)/\sqrt{X_1^2+\ldots+X_N^2}$ for sequences of independent r.v.'s $X_j, j \in \mathbb{N}$.

In joint work with V. Bentkus, SFB Bielefeld, it is proved that the error in the normal approximation admits the same type of bound as in the case of sums with nonrandom norming. The result holds in the i.i.d. case and extends to the non identically distributed case. (The latter is joint work with V. Bentkus and Bloznelis, Vilnius).

In particular it is shown that Lindeberg's condition still implies normal convergence but the Lindeberg-Feller result does not hold any more.

Furthermore, a conjecture made by (Logan et al. 73, AP) that in the i.i.d. case $t_N \Rightarrow N(0,1)$ is equivalent to X being in the domain of normal attraction and $EX=0$ has been verified. Moreover, it is shown that t_N, $N \in \mathbb{N}$ is stochastically bounded is equivalent to sub Gaussian behaviour of the tails of t_N, i.e. $\sup_N E\exp[\alpha t_N^2]$ for some $\alpha>0$. The last two results are joint work together with E. Giné and D. Mason.

R. GRÜBEL:

Hoare's selection algorithm

Hoare's selection algorithm finds the ℓ^{th} smallest of a set of n numbers, requiring a random number $C_{n,\ell}$ of comparisons. We consider convergence in distribution of the processes $\left(\frac{1}{n} C_{n,\lfloor n t \rfloor}\right)_{0<t\leq \ell}$ as $n \to \infty$. We also obtain stochastic upper bounds for $\frac{1}{n} C_{n,\ell}$ that hold uniformly in n and ℓ. These bounds arise as perpetuities in insurance mathematics and have a tail behaviour similar to that of Poisson distributions.

N. HENZE:

Do components of smooth tests of fit have diagnostic properties?

Smooth goodness of fit tests were introduced by Neyman (1937). They can be regarded as a compromise between globally consistent tests of fit and procedures having high power in the direction of a specific alternative. It is commonly believed that components of smooth tests like, e.g., skewness and kurtosis measures in the context of testing for normality, have special diagnostic properties in case of rejection of a hypothesis H_0 in the sense that they constitute direct measures of the kind of departure from H_0. Recent years, however, have witnessed a complete change of attitude towards the diagnostic capabilities of skewness and kurtosis measures in connection with normality testing. We argue that any component of any smooth test of fit is strictly non-diagnostic when used conventionally. However, a proper rescaling of components does indeed achieve the desired „directed diagnosis".

I. IBRAGIMOV:

Some statistical problems arising in the theory of SPDE

In this work (joint with R. Khasminskii) we investigate some nonparametric estimation problem occuring in parabolic SPDE of the form

$$d\,U(t,x) = L\,U(t,x)\,dt + \varepsilon\,d\,W(t,x)$$

where L is a partial differential operator with partially unknown coefficients and W is an infinite dimensional Wiener process. We derive asymptotic properties of estimators when $\varepsilon \to 0$.

A. JANSSEN:

Testing statistical functionals

The talk offers a method how to deduce tests for nonparametric statistical functionals. The approach is based on differentiable functionals and the framework of asymptotic statistics. For one- and two-sample testing problems asymptotic maximin most powerful tests are obtained. In various cases these tests turn out to be asymptotically efficient with respect to implicitly defined local alternatives given by the functional. It is surprising that their asymptotic power function does not depend on the special (implicit) direction of the alternatives. It only depends on the functional.
In connection with the full nonparametric two-sample problem the efficiency of the Wilcoxon, the log-rank, and the median test is obtained for special functionals.
The present work continues the research of H. Strasser (1985) which was done for one-sample problems.

U. KAMPS:

The inspection paradox with random time and applications

When considering a delayed renewal process one may be interested in both, the renewal function and the expected length of the interarrival time that contains some fixed time t. In general, it is difficult to obtain explicit expressions for specific underlying distributions. Replacing t by a random variable T, representations and bounds of the quantities are derived. The results lead to simple identities if T is exponentially distributed, and to an extension as well as to a refinement of the classical inspection paradox. Moreover, related characterizations of exponential distributions are shown. The results can be applied to the analysis of control charts with variable sampling intervals in quality control and to the premium principles in insurance mathematics.

C.A.J. KLAASSEN:

Efficient estimation of maximum correlation

Consider a two-dimensional random vector and the correlation between its two components after application of possibly different transformations of these components. Maximizing over all possible transformations we obtain the maximum correlation coefficient.
Our random vector fits into the normal copula model, if the above transformations may be chosen monotone and such that the resulting random vector is normal. In this case the maximum correlation coefficient equals the absolute value of the ordinary correlation coefficient of the corresponding normal vector. We show that the Van der Waerden normal scores rank correlation coefficient is an asymptotically efficient estimator of this correlation coefficient in the semiparametric normal copula model.

This is joint work with J.A. Wellner.

Y.A. KUTOYANTS:

Some problems of nonparametric estimation by observations of ergodic diffusion processes

We consider the problems of density and distribution function estimation by the observations of an ergodic diffusion process.
Two lower bounds on the risks of any estimators are proposed and the properties of the nonparametric estimators of the density and distribution function are studied. Particularly, the $\sqrt{T}$-consistency of the Kernel-type and a new (unbiased) estimator as well as their asymptotic normality are established. Unfortunately, both estimators are not asymptotically efficient. Then the asymptotic efficiency of the empiric distribution function is shown.

E. MAMMEN:

Direct estimation of low dimensional components in additive models

Additive regression models have turned out to be a useful statistical tool in analyses of high dimensional data sets. Recently, an estimator of additive components has been introduced by Linton and Nielsen (1994) which is based on marginal integration. The explicit definition of this estimator makes possible a fast computation and allows an asymptotic distribution theory. In this paper a modification of this procedure is introduced. We propose to introduce a weight function and to use local linear fits instead of kernel smoothing. These modifications have the following advantages:
(i) We demonstrate that with an appropriate choice of the weight function, the additive components can be efficiently estimated: An additive component can be estimated with the same asymptotic bias and variance as if the other components were known.
(ii) Application of local linear fits reduces the design related bias.

The talk reports on joint work with J. Fan and W. Härdle.

V. MAMMITZSCH:

Remarks on optimal kernels

According to T. Gasser and H.G. Müller (1979) a function $K:[-1,+1]\to \mathbb{R}$ is said to be of order (ν, k) if there holds

$$(1)\quad \int_{-1}^{+1} K(x)x^j\,dx = \begin{cases} 0 & \text{if } 0\le j<k, j\neq \nu \\ (-1)^\nu \nu! & \text{if } j=\nu \\ \pm 0 & \text{if } j=k, \nu\le k-2; \nu,k\in \mathbb{N}. \end{cases}$$

A kernel of order (ν , k) is defined to be optimal iff

$$\begin{cases} \# \text{ changes of sign of } K = k-2 \text{ and} \\ A(K) := \left| \int_{-1}^{+1} K(x)x^k dx \right|^{k-\nu} \cdot \left(\int_{-1}^{+1} K^2(x)dx \right)^{2\nu+1} = \min. \end{cases}$$

B. Granovsky and H.G. Müller (1989) showed that optimal kernels are of the form

(2) $K(x) = p(x)$ for all $x \in \{x \in [-1,1] : K(x) \neq 0\}$,

K continuous, p polynom of degree k. If $k - \nu > 3$, p is not uniquely defined by (1) and (2), as C. Pfeifer (1991) showed that there exist $\alpha_i \in [0,1]$ with $0 \le \alpha_1 < \alpha_2 < \cdots < \alpha_m \le 1$, m = $\left[\frac{k-\nu}{2}\right]$, and polynomials p_i of degree k with $p_i(-1) = p_i(\alpha_i) = 0$, i=1,...,m, such that for

$$K_i(x) = \begin{cases} p_i(x) & \text{if } -1 \le x \le \alpha_i \\ 0 & \text{if } \alpha_i < x \le 1 \end{cases}$$

equality (1) holds, cf. also the paper by B. Granovsky, H.G. Müller, and C. Pfeifer (1995). Using Gegenbauer polynomials as the main tool, in this paper we show $A(K_1) > A(K_2) > \ldots > A(K_m)$.

R. MARONNA:

Computationally cheap robust regression estimators

Robust regression estimators with high breakdown point involve the minimization of functions with many local minima, and this implies a computational cost that rapidly increases with the number of predictors. It is well-known that multiple least-squares regression can be performed by the repeated application of univariate regressions. We investigate the behaviour of procedures based on this idea, but using robust univariate regressions. Simulations show that the resulting estimators perform much better than regression S-estimators, with a much lower computing time.

G. NEUHAUS:

On testing tumor onset times

In a typical bio-assay a substance is studied with respect to its cancerogenicity by building a treatment group of n_1 rats and a control group without treatment of n_2 rats. When rat i dies at the time C_i, say, it is checked whether it died with tumor ($\Delta_i = 1$) or without tumor ($\Delta_i = 0$), while the tumor onset time T_i, say, is not observable. Hence the observations are

$Z_i = (C_i, \Delta_i)$, $1 \le i \le n_1 + n_2 =: n$. All rv's are assumed to be independent and identically distributed in the treatment groups $1 \le i \le n_1$ resp. in the control groups $n_1 + 1 \le i \le n$ with distribution functions F_1, F_2 of the tumor onset times resp. G_1, G_2 of the times of deaths. One wants to test the null hypothesis of no cancerogenicity, i.e., H_0: $F_1 = F_2$ versus the alternative of cancerogenicity, i.e. K: $F_1 \ge F_2$, $F_1 \ne F_2$, whereby G_1 and G_2 are nuisance parameters.
In the talk asymptotically optimal tests are constructed by LAN-theory for the above situation. Moreover, counting process theory is used for justifying replacement of the basic point $F := F_1 = F_2$ by some estimator $\hat{F}$ and for removing the dependence on the unknown G_1, G_2. As examples analogues of the log-rank test and the Wilcoxon test for the above „interval censoring" problem are presented.

I. PIGEOT:

A survey on jackknife estimators in contingency table analysis

We discuss the estimation of a common relative risk and a common odds ratio in stratified cohort studies and in stratified case-control studies, respectively. Since these estimators are typically biased, the jackknife principle is applied to achieve a bias reduction. The asymptotic properties of the resulting jackknife estimators are discussed mainly for the case that the number of strata remains fixed while the sample sizes within each stratum tend to infinity. Here, a general result is presented which under certain regularity conditions yields consistency and asymptotic normality of every jackknife estimator of a variety of functions of binomial probabilities. Furthermore, it can be shown that jackknifing indeed leads to a bias reduction.
In addition, another jackknife approach is presented and the asymptotic properties are discussed under the so-called sparse data model.
Finally, we mention a computer system we developed which makes these techniques not only available for practitioners, but also gives recommendations on request concerning the choice of an estimator in a given data situation. These recommendations are based on rules gained from simulation studies using artificial intelligence.

D. RASCH:

Replicationfree optimal designs in regression analysis

Exact locally D- and C-optimal experimental designs for non-linear regression problems based on model functions with p parameters often contain p support points $x_1, x_2, \ldots, x_p$ and need $n_i > 1$ measurements for the point x_i. For some practical applications we construct optimal designs with at most one measurement at any of N points $x_j \in [x_l, x_u]$. The paper presents a search algorithm for the construction of so-called replication-free exact locally optimum experimental designs for certain optimality criteria for non-linear regression functions with p parameters.

W.-D. RICHTER:

Geometric approach to quantile approximation for noncentral generalized chi-square distributions

Asymptotic expansions for large deviation probabilities can be used to derive quantile approximation procedures in many statistical situations. Tail probabilities of noncentral generalized chi-square distributions with k d.f. can be viewed as tail probabilities of respective k dimensional shifted spherically distributed random vectors. For this purpose, multivariate asymptotic expansions for a special type of Laplace integrals are needed. Namely, one has to study the case that there exists a uniquely determined maximum point of the function to be integrated over the large deviation domain and that this maximum point belongs to the boundary and degenerates asymptotically when the large deviation paramter approaches infinity. The latter case will be dealt with using global and local geometric properties as well of the spherical distributions as of the large deviation domains.

H. RIEDER:

Estimation of mortality rates

The estimation of true mortalities at ages x on the basis of observed mortalities q_x has been treated by many authors; for a textbook account see e.g. Benjamin and Pollard (1982). More recently, Miller and Olbricht (1996) have fit a regression line $f(x) = b \cdot \min\{31, x\} + c$ to $\log q_x$ by minimum $-L_1$ followed by the backtransformation $\hat{q}_x = \exp\{\hat{f}(x) + \hat{\sigma}^2/2\}$, where the scale estimate is the (standardized) interquartile range of the min $-L_1$ residuals.

Using the Laplace transform we prove that such estimates are systematically biased unless the error distribution in the log-linear model is normal and the scale estimate is gauged to normal variance. We determine this bias under the contaminated normal $F = (1-r)N(0,1) + rM$ for $M = \delta_z$ (Dirac), $M = N(z,1)$, and $M = \frac{1}{2}(\delta_{-z} + \delta_z)$, $M = N(0, z^2)$ - both for min $-L_1$, IQR, and min $-L_2$, VAR. Already for $z \approx 2.5$ the true mortalities may be underestimated by 50%! Instead, we therefore propose the square-root model: $\sqrt{q_x} = f(x) + (\sigma/\sqrt{L_x})\varepsilon_x$ (L_x: size of group of age x). As opposed to log, the $\sqrt{\ }$ is variance-stabilizing, can cope with $q_x = 0$, and has the "nonparametric" backtransformation formula $\hat{q}_x = \hat{f}(x)^2 + \hat{\sigma}^2/L_x$ (based on $EY^2 = \mu^2 + \sigma^2$ vs. $E\,e^Y = \exp\{\mu + \sigma^2/2\}$).

For the 1994 DAV-smooth, one real data set, and simulated binomial mortalities based on the DAV-smooth we fit a 2nd and 3rd degree polynomial f(x) and an exponential $f(x)=\exp\{a+b\cdot \min(31,x)\}$ by min $-L_2$, min $-L_1$, and by robust regression rg_Huber (as implemented in the ISP-program), and obtain fits, which are excellent and practically indistinguishable.

J. M. ROBINS:

Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semiparametric models

We argue, due to the curse of dimensionality, that there are major difficulties with any pure or smoothed likelihood-based method of inference in studies with randomly missing observations when missingness depends on a high-dimensional vector of variables. We study in detail a semiparametric superpopulation version of continuously stratified random sampling. We show that all estimators of the population mean that achieve any prespecified rate of convergence, no matter how slow, require the use of the selection (randomization) probabilities. We argue that, in contrast to likelihood methods which ignore these probabilities, inverse selection probability-weighted estimators continue to perform well. We develop a curse of dimensionality appropriate (CODA) asymptotic theory for inference in non- and semiparametric models in an attempt to formalize our arguments. We discuss whether our results constitute a fatal blow to the likelihood principle and study the attitude towards these that a committed Bayesian would adopt. Finally, we apply our CODA theory to analyze the effect of the "curse of dimensionality" in several interesting semiparametric models.

D. M. ROCKE:

Robust scale-free cluster analysis

Many cluster analysis techniques assume that the appropriate distance between any two points is known; yet in many cases this assumption is not reasonable. We investigate a class of methods that are affine equivariant and can cope with clusters each of whose shape is unknown and in which the shape may or may not be the same between clusters. Normal maximum likelihood would satisfy these requirements, but there are two large problems with this method. First, in dimensions higher than two or three, the global maximum of the likelihood is very difficult to find. Second, even a small percentage of contamination can entirely ruin the process. The method we describe is a robustification of normal maximum likelihood which, with associated search techniques, allow for broad application.

P. J. ROUSSEEUW:

Recent applications of robust statistics

The purpose of this talk is to discuss some recent applications of robust methods. We will focus on "positive breakdown" techniques, which can resist a substantial fraction of contamination in the data without breaking down.
The first part of the talk is concerned with robust regression techniques, such as the least median of squares (LMS) estimator. Contrary to the classical least squares approach, which tends to mask outliers and other substructures, the LMS method attempts to fit the majority of the data and thereby detects the outliers. Apart from illustrative examples, we will consider some substantive applications. One is to electric power systems, where one needs to estimate

the system's state variables robustly. Other applications are to computer vision, where LMS has been used for image recovery, surface reconstruction, and to detect moving objects in video from a mobile camera.
The second part of the talk is about detecting outliers in a multivariate point cloud, using the minimum volume ellipsoid (MVE) estimator of location and scatter. In a geochemistry application, MVE-based robust distances were used to detect mineralizations hidden beneath the surface. The MVE can also be used to robustify multivariate techniques such as principal component analysis and discriminant analysis. Applications in computer vision include image segmentation and determining shapes from color images.

L. RÜSCHENDORF:

Stochastic analysis of algorithms - the contraction method

The contraction method for the asymptotic analysis of (stochastic) recursive algorithms is based on essentially two ingredients. Firstly, to find a suitable probability metric allowing to obtain asymptotic contraction properties of the normalized algorithm and suitable estimates from above in terms of moments. Secondly, one has to prove convergence to some limiting equation, typically by some decomposition technique.
This method has applications to algorithms from different fields. In particular we will discuss sorting type algorithms, random trees and permutations, Arch- and branching-type algorithms as well as applications of the contraction method to some probabilistic problems.

N. SCHMITZ:

Permutation tests - a revival?

It is shown that permutation tests have convincing optimum properties for interesting classes of continuous distributions as well as discrete distributions with fixed support. Conditions sufficient for uniformly maximal power on subclasses are given. Moreover, a variety of examples is presented. To make these results applicable an efficient algorithm for computing the critical region of the permutation test is derived. The performance of this algorithm, which is based on an idea of Pagano & Tritchler, is demonstrated by simulation results. Moreover, this algorithm is compared with an algorithm due to Green.

J. STEINEBACH:

Variance estimation based on invariance principles

Consistent variance estimators for certain stochastic processes are suggested using the fact that (weak or strong) invariance principles may be available. Convergence rates are also derived, the latter being essentially determined by the approximation rates in the corresponding invariance principles. Some multivariate extensions are briefly discussed, too, together with possible applications in the changepoint analysis of renewal processes.

H. STRASSER:

Invariance properties of estimators and incidental nuisance parameters

It is shown that the asymptotic information bound which is valid for the estimation of a parameter in the mixture model remains valid in the model with incidental nuisance parameters if only perturbation invariant estimators are considered. Perturbation invariance is a property which is closely related to permutation invariance. In particular, equicontinuous functions of empirical processes are perturbation invariant. Thus, the result settles a problem posed by Pfanzagl of how permutation invariance is to be modified in order to exclude superefficiency in models with incidental nuisance parameters.

A. VAN DER VAART:

Likelihood inference in the proportional odds model

The proportional odds model specifies that the odds ratios $F(t|z)/(1-F(t|z))$ of the distribution $F(t|z)$ of a survival time T given a covariate Z are equal to the product $H(t)\exp(z^T b)$ of an unknown function H and an exponential regression factor. The function H is necessarily monotonely increasing, cadlag and 0 at the origin. We give a definition of a likelihood function for this model and discuss the behavior of the maximum likelihood estimator for the parameter (H,b), the behavior of the likelihood ratio statistic for testing hypotheses on b, and a discretized derivative of the profile likelihood function. The maximum likelihood estimator is asymptotically normal and efficient, the likelihood ratio statistic asymptotically chisquared, and the derivative of the profile likelihood a consistent estimator of the asymptotic variance of the maximum likelihood estimator of b. This is shown by deriving the likelihood equations, and inverting them, after showing that the information operator is continuously invertible, and that certain classes of score functions form Donsker classes of functions.

I. VAJDA:

Statistical inference based on convex distances of probability distributions

Let $\mathcal{P}_\mu$ be the set of all probability densities on a measurable space $(\Xi,\mathcal{A})$ with respect to a σ-finite measure μ. We consider a sample $X_1,\dots,X_n$ with i.i.d. components $\sim p_0 \in \mathcal{P}_\mu$. We discuss estimators p_n of p_0 considered in [1] which are consistent in the sense

$$n^{2/3} \int \frac{(p_n - p_0)^2}{p_0} d\mu = O_p(1).$$

Further, if $\mathcal{P} = \{p_\theta : \theta \in \Theta\} \subset \mathcal{P}_\mu$ for some $\Theta \subset \mathbb{R}^d$ and $p_0 = p_{\theta_0}$, we consider the estimator θ_n of θ_0 which minimizes the total variation $V(p_n, p_\theta) = \int |p_n - p_\theta| d\mu$ on Θ. We show that if for some $c(\theta_0) > 0$

$$V(p_\theta, p_{\theta_0}) = c(\theta_0)\|\theta - \theta_0\| + o(\|\theta - \theta_0\|) \text{ as } \theta \to \theta_0$$

then θ_n is consistent in the sense

$$n^{1/3}\|\theta_n - \theta_0\| \leq O_p(1).$$

This is compared with the results of [2].

[1] L. Györfi, F. Liese, I. Vajda, E. van der Meulen: Distribution estimates consistent in χ^2-divergence (preprint).
[2] P.V. Rao, E.F. Schuster, R.C. Littel: Estimation of a shift and center of symmetry based on Kolmogorov-Smirnov statistics. Ann. Statist. 3, 862-873.

Y. VARDI:

Network Tomography: Estimating source-destination traffic intensities from link data

We consider the problem of estimating the traffic intensity between all (directed) pairs of nodes of a (strongly connected) directed network, based on data measured on the links of the network repeatedly. We assume that the traffic originating at nodes i and i′ is Poisson with parameter $\lambda_i, i = (i, i')$, and is independent of traffic between other pairs. Thus there are n(n-1) Poisson parameters to be estimated. The number of links is of order o(n), and we repeat the link measurement K times. We consider two different types of traffic routing (1) deterministic (2) Markovian. We show that an exact MLE is computationally too complicated and we discuss an alternative based on the normal approximation. We then derive in detail an estimate based on the method of moments. The moment equations give rise to a LININPOS problem (Vardi & Lee, JRSS B, 1993) which is then solved using an EM/ML method. A small simulation study based on a network with 4 nodes (and hence 12 directed pairs) and 7 directed links is carried out, and the results indicate good statistical characteristics for moderate size data sets.

W. H. WONG:

A Monte Carlo method and its application to the salesman problem and nonlinear PDE

We propose a weighted Markov sampling scheme for simulation and global optimization. The original sampling problem is embedded into a sequence of problems with decreasing levels of complexity. To make transitions between levels possible, it is necessary to introduce an "importance weight" variable into the Markov system. Traditional Metropolis-Hasting transition rules are not applicable because the target density for the augmented system (original variables + weight) is not known explicitly. We present a method for the construction of transition rules that will yield valid weighted samples. We also discuss ways to control the behaviour of the weights. The method is tested on the travelling salesman problem and on multigrid Monte Carlo from Gibbs distributions whose minimum energy states correspond to solutions of nonlinear elliptic PDEs.

C.-H. ZHANG:

Linear regression with doubly censored data

Linear regression with doubly censored responses is considered. Buckley-James-Ritov type estimators are proposed. An expansion of the estimating equations is obtained under fairly general assumptions. Sufficient conditions are given for the asymptotic consistency and normality of the estimators. Semiparametric information and projective scores are also discussed.

This is a joint work with Xin Li.

W.R. VAN ZWET:

Statistical estimation for the contact process

At every time $t \geq 0$, each site in the lattice $\mathbf{Z}^d$ is either healthy or infected. The dynamics of the contact process are as follows. Healthy sites are infected by each of their infected immediate neighbors with rate λ, and infected sites become healthy with rate 1. All processes involved are independent. The process starts e.g. with a single infected site at the origin at time $t = 0$.

The discussion centers on the estimation of λ on the basis of an observation of the process at a single time t. An estimator is constructed which is consistent and asymptotically normal as $t \to \infty$.

Berichterstatterin: Ursula Gather

E-Mail-Adressen der Tagungsteilnehmer

Bickel, P.J.	bickel@stat.berkeley.edu
Birge, L.	lb@ccr.jussieu.fr
Bolthausen, E.	eb@amath.unizh.ch
Carroll, R.J.	rayc@helix.nih.gov
Dahlhaus, R.	dahlhaus@statlab.uni-heidelberg.de
Fernholz, L.	fernholz@vm.temple.edu
Gaenssler, P.	gaensler@rz.mathematik.uni-muenchen.de
Gather, U.	gather@omega.statistik.uni-dortmund.de
van de Geer, S.	geer@wi.leidenuniv.nl
Gill, R.D.	gill@math.ruu.nl
Goetze, F.	goetze@math200.mathematik.uni-bielefeld.de
Gruebel, R.	rgrubel@stochastik.uni-hannover.de
Henze, N.	henze@mspcdip.mathematik.uni-karlsruhe.de
Hilker, T.	hilker@omega.statistik.uni-dortmund.de
Ibragimov, I.A.	ibr32@pdmi.ras.ru
Janssen, A.	janssena@clio.rz.uni-duesseldorf.de
Kamps, U.	u.kamps@stochastik.rwth-aachen.de
Keiding, N.	nk@kubism.ku.dk
Klaassen, C.A.J.	chrisk@fwi.uva.nl
Klueppelberg, C.	cklu@ikaros.Mathematik.Uni-Mainz.DE
Kutoyants, Y.	kutoyants@univ-lemans.fr
Lehn, J.	lehn@mathematik.th-darmstadt.de
Liese, F.	liese@mathematik.uni-rostock.d400.de
Mammen, E.	enno@statlab.uni-heidelberg.de
Mammitzsch, V.	mammit@mathematik.uni-marburg.de
Maronna, R.	maron@elrond.uc3m.es
Neuhaus, G.	fm3a007@TILL.math.uni-hamburg.de
Pigeot-Kuebler, I.	pigeot@stat.uni-muenchen.de
Rasch, D.	dieter.rasch@wts.wk.wau.nl
Reiss, R.-D.	reiss@hrz.uni-siegen.d400.de
Richter, W.-D.	richter@likeli.math.uni-rostock.de
Rieder, H.	helmut.rieder@uni-bayreuth.de
Ritov, J.	yaacov@olive.mscc.huji.ac.il
Robins, J.	robins@episun1.harvard.edu

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Rocke, D.	dmrocke@ucdavis.edu
Rousseeuw, P.J.	rousse@uia.ua.ac.be
Rueschendorf, L.	ruschen@buffon.mathematik.uni-freiburg.de
Schmitz, N.J.	schmnor@math.uni-muenster.de
Steinebach, J.	jost@mathematik.uni-marburg.de
Strasser, H.	helmut.strasser@isis.wu-wien.ac.at
Stute, W.	Winfried.Stute@math.uni-giessen.de
van der Vaart, A.W.	aad@cs.vu.nl
Vajda, I.	vajda@utia.cas.cz
Vardi, Y.	vardi@stat.rutgers.edu
Wong, W.-H.	whwong@sparc2.sta.cuhk.hk
Zhang, C.-H.	chzhang@rci.rutgers.edu
van Zwet, W.R.	vanzwet@wi.leidenuniv.nl

Tagungsteilnehmer

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

Berkeley , CA 94720-3860
USA

Prof.Dr. Luisa Fernholz
Statistics Department
Temple University

Philadelphia , PA 19122
USA

Prof.Dr. Lucien Birge
L.S.T.A.
45/55 - 3e - Boite 158
Universite Paris VI

F-75252 Paris Cedex 05

Prof.Dr. Peter Gänssler
Mathematisches Institut
Universität München
Theresienstr. 39

80333 München

Prof.Dr. Erwin Bolthausen
Institut für Angewandte Mathematik
Universität Zürich
Winterthurerstr. 190

CH-8057 Zürich

Prof.Dr. Ursula Gather
Fachbereich Statistik
Universität Dortmund

44221 Dortmund

Prof.Dr. Raymond Carroll
Department of Statistics
Texas A & M University

College Station , TX 77843-3143
USA

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

NL-2300 RA Leiden

Prof.Dr. Rainer Dahlhaus
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294

69120 Heidelberg

Prof.Dr. Richard D. Gill
Mathematisch Instituut
Rijksuniversiteit te Utrecht
P. O. Box 80.010

NL-3508 TA Utrecht

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Prof.Dr. Friedrich Götze
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131

33501 Bielefeld

Prof.Dr. Rudolf Grübel
Institut für Mathematische Stochastik
Universität Hannover
Postfach 6009

30060 Hannover

Prof.Dr. Norbert Henze
Institut für Mathematische Stochastik
Universität Karlsruhe

76128 Karlsruhe

Dipl.-Stat. Torsten Hilker
Fachbereich Statistik
Universität Dortmund

44221 Dortmund

Prof.Dr. Ildar A. Ibragimov
St. Petersburg Branch of Steklov Mathematical Institute - POMI
Russian Academy of Science
Fontanka 27

191011 St. Petersburg
RUSSIA

Prof.Dr. Arnold Janssen
Mathematisches Institut
Angewandte Mathematik
Universität Düsseldorf
Universitätsstr. 1

40225 Düsseldorf

Dr. Udo Kamps
Institut für Statistik und Wirtschaftsmathematik
RWTH Aachen

52056 Aachen

Prof.Dr. Niels Keiding
Department of Biostatistics
University of Copenhagen
Blegdamsvej 3

DK-2200 Kobenhavn N

Prof.Dr. Chris A. J. Klaassen
Fakulteit Wiskunde en Informatica
Universiteit van Amsterdam
Plantage Muidergracht 24

NL-1018 TV Amsterdam

Prof.Dr. Claudia Klüppelberg
Fachbereich Mathematik
Universität Mainz
~~Postfach 3980~~

55099 Mainz

Prof.Dr. Yurii A. Kutoyants
Faculte des Sciences
Universite du Maine
B.P. 535
Avenue Olivier Messiaen

F-72017 Le Mans Cedex

Prof.Dr. Ricardo Maronna
Departamento de Estadistica y
Econometria
Universidad Carlos III
Calle Madrid, 126

E-28903 Getafe

Prof.Dr. Jürgen Lehn
Fachbereich Mathematik
TH Darmstadt
Schloßgartenstr. 7

64289 Darmstadt

Prof.Dr. Georg Neuhaus
Institut für Mathematische
Stochastik
Universität Hamburg
Bundesstr. 55

20146 Hamburg

Prof.Dr. Friedrich Liese
Fachbereich Mathematik
Universität Rostock

18051 Rostock

Prof.Dr. Iris Pigeot-Kübler
Institut für Statistik
Universität München
Ludwigstr. 33

80539 München

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

69120 Heidelberg

Prof.Dr. Dieter Rasch
Department of Mathematics
Agricultural University Wageningen
Dreijenlaan 4

NL-6703 HA Wageningen

Prof.Dr. Volker Mammitzsch
Fachbereich Mathematik
Universität Marburg

35032 Marburg

Prof.Dr. Rolf-Dieter Reiß
Fachbereich 6 Mathematik
Universität/Gesamthochschule Siegen

57068 Siegen

Prof.Dr. Wolf-Dieter Richter
Fachbereich Mathematik
Universität Rostock

18051 Rostock

Prof.Dr. Peter J. Rousseeuw
Dept. of Mathematics
Universitaire Instelling Antwerpen
Universiteitsplein 1

B-2610 Wilrijk

Prof.Dr. Helmut Rieder
Fakultät für Mathematik und Physik
Universität Bayreuth

95440 Bayreuth

Prof.Dr. Ludger Rüschendorf
Institut für Mathematische
Stochastik
Universität Freiburg
Eckerstr. 1

79104 Freiburg

Prof.Dr. Jacob Ritov
Department of Statistics
The Hebrew University of Jerusalem
Mount Scopus

Jerusalem 91905
ISRAEL

Prof.Dr. Norbert J. Schmitz
Institut für Mathematische
Statistik
Universität Münster
Einsteinstr. 62

48149 Münster

Dr. James Robins
Dept. of Biostatistics
Harvard School of Public Health
677 Huntington Ave.

Boston , MA 02115
USA

Prof.Dr. Josef Steinebach
Fachbereich Mathematik
Universität Marburg

35032 Marburg

Prof.Dr. David M. Rocke
Graduate School of Management
University of California

Davis , CA 95615-8609
USA

Prof.Dr. Helmut Strasser
Institut für Statistik
Wirtschaftsuniversität Wien
Augasse 2 - 6

A-1090 Wien

Prof.Dr. Winfried Stute
Mathematisches Institut
Universität Gießen
Arndtstr. 2

35392 Gießen

Prof.Dr. Cun-Hui Zhang
Dept. of Statistics
Rutgers University
Hill Center, Busch Campus

New Brunswick , NJ 08903
USA

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

NL-1081 HV Amsterdam

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

NL-2300 RA Leiden

Prof.Dr. Igor Vajda
Institute of Information Theory and
Automation
Pod vodarenskou vezi 4

182 08 Praha 8
CZECH REPUBLIC

Prof.Dr. Yehuda Y. Vardi
Dept. of Statistics
Rutgers University
Hill Center, Busch Campus

New Brunswick , NJ 08903
USA

Prof.Dr. Wing-Hung Wong
Department of Statistics
The University of Chicago
5734 University Avenue

Chicago , IL 60637
USA

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Gefördert durch
DFG Deutsche Forschungsgemeinschaft