MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 9/1993

Curves, Images, Massive Computation

21.-27.2.1993

Die Tagung fand unter der Leitung von Prof. Dr. L. Devroye (Montreal), Prof. Dr. W. Härdle (Berlin) und Prof. Dr. I. Johnstone (Stanford) statt. Im Mittelpunkt des Interesses standen Fragen zu aktuellen Entwicklungen im Bereich der Funktionenschätzung, deren Anwendung bei der Image-Analyse und der massive Einsatz von Computern, den diese Verfahren zum großen Teil benötigen.

Als Einführung wurden am Montagvormittag Vorträge zu Funktionenschätzung in angewandten Studien von B. Silvermann, S. Leurgans, A. Kneip, C.J. Stone und J.S. Marron gegeben. In diesem Feld, wie etwa Hazardstudien, treten Funktionen in natürlicher Weise als Schätzziel auf. Am Nachmittag standen Funktionenschätzungen in bestimmten Nichtstandard-Situationen (Abhängigkeit, zensierte Daten) im Mittelpunkt. Hierunter fallen die Vorträge von P. Doukhan, A. Tsybakov, J. Fan und I. Gijbels. Dabei wurde auch ein neueres Verfahren zur nicht-parametrischen Schätzung vorgestellt, das darauf beruht, lokal gewichtete Polynome anzupassen. Hier ist auf die Vorträge am Dienstag zu verweisen, die die Reduktion des numerischen Aufwands dieser Verfahren behandelten.

Der Dienstagvormittag war einer Reihe von Vorträgen über Wavelets gewidmet (I. Johnstone, D. Picard, D. Donoho, M. Neumann), einer neuen und vielversprechenden Methodik im Bereich der Signalverarbeitung und Funktionenschätzung. Aus dieser Sitzung entwickelte sich für Mittwochabend eine informelle Diskussion über (bzw. Einführung in die Theorie der Wavelets, die für alle Teilnehmer sehr fruchtbar war. Ein weiterer Vortrag (M. Nussbaum) beschäftigte sich mit der Äquivalenz von Dichteschätzung und Weißem Rauschen. Der Dienstagnachmittag behandelte die Implementation nicht-parametrischer Funktionenschätzer, so daß sie schnell berechnet werden und der intensive Rechenzeiteinsatz reduziert wird (Th. Gasser, K. Messer). Daneben wurden weitere Verfahren zur Funktionenschätzung und Imageanalyse vorgestellt und ihre Anwendung in Semiparametrischen Modellen gezeigt (E. Mammen, K.-C. Li, Ph. Vieu, B. Turlach).

Die Frage des massiven Computereinsatzes in der Statistik war Hauptthema des Mittwochs. Es wurden verschiedene Aspekte von Markov-Chain-Samplers diskutiert, sowie deren Anwendung in Monte-Carlo-Studien und Image-Analyse demonstriert (D.D. Cox, P. Green, F. Götze, L. Tierney). Die Sitzung endete mit dem Vortrag von O. Bunke über Bootstrap, ein rechenintensives Sichprobenwiederholungsverfahren.

Verschiedene Fragen zur Image-Analyse, wie z.B. Rekonstruktion und effiziente Speicherung, standen im Mittelpunkt der Donnerstagvormittag-Sitzung (R. Olshen, L. Schumaker, C. Jennison, L. Younes, F. Natterer, P. Clifford). Den Nachmittag eröffnete der Vortrag zu Sieves und Maximum-Likelihood von S. van de Geer, danach wurden nichtlineare und nicht-parametrische Kalibrationstechniken vorgestellt (E. Jolivet, M.A. Gruet). Weiterhin wurde die (für alle nicht-parametrischen Methoden wichtige) Frage der Wahl des Glättungsparameters diskutiert (D. Girard, B. Grund).

Der Freitagvormittag begann mit einem Vortrag von P. Hall über die Schätzung von fraktalen Dimensionen. Dies ist ein neues und hochinteressantes Gebiet in der Statistik mit vielen praktischen Anwendungen, die ebenfalls vorgestellt wurden. Weitere Themen waren die multivariaten Dichteschätzung in den Vorträgen von C. Huber, W. Polonik und Ph. Stark sowie die Regressorwahl in parametrischen und nicht-parametrischen Modellen (G. Eagleson, M. Müller). Die Tagung endete mit einer weiteren Sitzung über massiven Computereinsatz am Freitagnachmittag. In dieser letzten Sitzung wurden nochmals Aspekte der nichtparametrischen Funktionenschätzung behandelt (G. Golubev, M. Low) sowie Bootstrap-Anwendungen vorgestellt (R. Beran, W. Stute).

ABSTRACTS

Stein Confidence Sets

R. Beran
University of California, Berkeley

In the simplest general trend model, the observations $X = (X_1, \ldots, X_q)^T$ satisfy the relation $X_i = \xi_i + E_i$, the $\{E_i\}$ being i.i.d. standard normal random variables. Estimators of $\xi = (\xi_1, \ldots, \xi_q)^T$ include the classical estimator X, smoothed curve estimators, model selection estimators and the Stein estimators such as $\hat{\xi}_s = (1 - \frac{q-2}{|x|^2})X$. Our results concern the construction and asymptotic properties of Stein confidence sets for ξ of the form $C_s = \{f \in I\!R^q : |\hat{\xi}_s - f| \leq \hat{r}_q\}$. We discuss:

1. Asymptotic and bootstrap constructions of $\hat{r}_q$ such that $P_\xi[C_s \ni \xi] = \alpha + O(q^{-1/2})$.
2. Improved bootstrap constructions of $\hat{r}_q$ such that $P_\xi[C_s \ni \xi] = \alpha + O(q^{-1})$.
3. Risk superiority of C_s over the classical confidence sets centered at X.
4. Failure of natural bootstrap algorithm that resample from $N(X, I)$ or $N(\hat{\xi}_s, I)$.
5. Asymptotic geometry of Stein confidence sets.

Bootstrapping in Small Sample Regression Problems

O. Bunke
Humboldt University, Berlin

We investigate the small sample behavior of bootstrap procedures in parametric, semiparametric and nonparametric regression problems. The approach is based on exact calculations or a small variance approximation of the biases in the moments of the bootstrap distribution of the interesting estimator. We show that even if a parametric model for the regression model is known, there will be an essential bias in the 3rd– and 4th–order bootstrap moments. Moreover, we calculate additional bias terms in the case of bootstrap procedures based on residuals after fitting by an incorrect linear or nonlinear parametric regression model or by a nonparametric estimator.

A typical example exhibits the possibly disturbing magnitude of such terms in the case of the variance of the bootstrap distribution.

We discuss modification of bootstrap procedures that may diminish the above mentioned errors in approximating the true distribution of an estimator by bootstrapping. They are based on estimates of the first four central moments of the error distribution that have small biases or, in case of the error variance, have additionally mean square error optimality properties.

Polygonal Models for Spatial Pattern

P. Clifford
Oxford University

Let T be a 2-D region, and let χ be a surface defined on T. The values of χ on T constitute an image or pattern. The true value of χ at any point on T cannot be directly observed, but data can be recorded which provide information about χ. The aim is to reconstruct χ using the prior knowledge that χ varies smoothly over most of T, but may exhibit jump discontinuities over line segments. This information can be incorporated via Bayes' theorem, using a polygonal Markov random field as a prior distribution. In a similar fashion linear structure in 2-D images can be modeled as a random graph. In this case the statistical problem is that of efficiently estimating the defined parameter. A probabilistic basis for such discussions is outlined using base measures which are either spatial point or line processes.

Monte-Carlo Approximation of Posterior Distributions in Nonparametric Logistic Regression

Dennis D. Cox
Rice University, U.S.A.

A prior for nonparametric logistic regression is constructed by assuming the logit is a Gaussian process plus a standard regression model with an improper prior on the components. Many authors have proposed approximate inference using a Gaussian approximation of the posterior. We wished to assess the accuracy of this approximation. Analysis of the posterior tails leads to an importance sampling density which is a mixture of the Gaussian approximation and the tail dominating density. Our calculations show the Gaussian approximation works reasonably well for marginal inferences about the logit (or probability) function at a point. Also demonstrated were simultaneous credibility bands for the logit.

Massive Adaptation

David Donoho
Stanford University

Let $y_i = \theta_i + z_i, i = 1, \ldots, n, z_i$ i.i.d. $N(0,1)$. Let $\hat{\theta}^*_{n,i} = \eta_{t_n}(y_i)$, where $t_n = \sqrt{2\log(n)}$ and $\eta_t(g) = \mathrm{sgn}(y)(|y| - t)_+$. Now for almost any reasonable loss which is bowl-shaped and symmetric, and almost any a priori class Θ which is bowl-shaped and symmetric

$$\sup_{\theta\in\Theta} R_n(\hat{\theta}^*_n, \theta) \leq \text{" } \log n\text{-terms" } \inf_{\hat{\theta}} \sup_{\theta\in\Theta} R_n(\hat{\theta}, \theta).$$

Implications include the fact that wavelet estimators based on $\sqrt{2\log(n)}$ thresholding achieve within the log factors of the minimax risk over all Triebel and Besov classes.

Since this changes the picture of adaptive estimation - making it seem almost trivial if we are willing to accept log terms and Gaussian white noise assumptions, we discuss the problem of adaptively selecting a basis. Adopting the approach of Coifman, Meyer, Quake and Wickerhauser, we select a basis by adaptively tiling the time-frequency plane, and threshold in that basis. Denoising of Doppler and other mult-frequency chirps works well when the basis is selected using Stein's unbiased risk estimate.

The use of simple thresholding for estimating quadratic functionals is also briefly discussed, along with adaptivity results in inverse problems.

Minimax Rates for Weakly Dependent Regression Estimates

Paul Doukhan
Université d'Orsay, France

Regression estimation is a main tool for the general predictor problem. We estimate a regression function r in a mixing stationary model $(X_n, Y_n)_{n=1,2,...}$ $[r(x) = E(Y \mid X = x)]$ using δ-sequence estimates (Walter, Blum (1979)) which include the use of convolution kernels, wavelets (Doukhan (1988)) and other projections techniques. Some explicit models are explicited by their mixing properties (Doukhan L.N.S. (to appear)).

Extending optimal equivalents for variance of kernel density estimates (Tran (1990)) to our framework yields the MINIMAX bounds for the MISE of estimates of f and r. A simple by-product is the $\sqrt{n}$-consistent estimator of quadratic integral functionals of f and rf under restrictive assumptions. Berbee Reconstruction Lemma (1979) allows to prove uniform MINIMAX rates of this large class of estimates of a β-mixing process (in the a.s. and mean senses). In the weak dependent case the only result of this class was in Birgé (1986) concerning Dereblin recurrent Markov chains (this heavy condition is omitted here).

Variable Selection in Nonparametric Regression

Geoff Eagleson
University of New South Wales

As well as new methods of estimating response surfaces, attention needs to be paid to the consequent inferential problems. This point will be illustrated with a simple example. Analyzed by an automatic application of the additive nonparametric regression program ACE, the results obtained are spurious. To understand why this might be so we study the behavior of the selection criterion, R^2, under two extreme null models. Appropriate permutation tests provide one way of assessing the fits obtained. It may well be that smoothing should be used to assess goodness-of-fit, different to that which is used for estimating the response.

Localization: A Useful Principle

Jianqing Fan
University of North Carolina and Chapel Hill

The traditional statistical model selection is based on largely "trials-and-errors". In this talk, we introduce a simple and useful idea for model fitting. The smoothing parameter that governs the model complexity is estimated based on the idea of localization: one models locally the unknown function by a linear function and uses different order of approximation to assess the bias of the approximation. The variance of estimate can easily be determined by the usual least squares theory. In comparison with traditional approaches, our method is more efficient and has wider applicability. The idea can be used to determine the order of approximation at various points and can also be applied to all likelihood based model. These models can be in one case the partial likelihood from survival analysis data, and in another case the quasi likelihood function from generalized linear models. This new idea gives more strength to the statistical "maximum likelihood principle" via automatic model selection.

Fast Local Polynomial Fitting

Theo Gasser
Abteilung Biostatistik, Universität Zürich

Among the many methods proposed for non–parametric curve and surface fitting, local polynomial fitted by least squares turned out recently to have a number of optimality properties (in a minimax sense). There are a number of other advantages, such as automatic boundary correction. We have developed a fast algorithm for computing these estimates. It works in $O(n)$ operations, whereas the conventional algorithm needs $O(n^{\frac{9}{5}})$ operations. Typically half a second is needed on a SUN Sparc IPX to fit 1000 points, irrespective of the design. Due to an add-substract procedure, numerical stability is a problem, in particular for small bandwidths. This problem is dealt mainly by checking a stability factor, defined to be the smallest (standardized) Cholesky factor in the normal equation.

Nonparametric Regression Based on Censored Data

Irene Gijbels
Institut de Statistique, Université Catholique de Louvain

There are a variety of statistical tools available for modeling the relationship between response and covariate if the data are fully observable. In the situation of censored data however, those tools are no longer directly applicable. We provide an easily implemented methodology for modeling the regression relationship based on censored data, without making any assumptions about its form. Basic ideas behind the methodology are to transform the data in an appropriate simple way, and then to apply a locally weighted least squares regression. The proposed kernel type estimator involves an appealing variable bandwidth, and as a consequence the procedure automatically adapts to the design of the data points. The bandwidth depends on a tuning parameter which can be selected using cross-validation techniques. The methodology is illustrated via simulation studies and analysis of real data sets. Some basic asymptotic results are established.

On the Randomized Smoothing Parameters Choices

D. Girard
CNRS, University of Grenoble

We describe how and when we may use the fast randomized versions of the GCV and C_L-criterions (resp. $RGCV$ and RC_L). Even for finite sample sizes, a simple heuristic argument shows that, if we are confident that C_L can work for the problem at hand, then RC_L works too! We show that this can be extended to GCV and $RGCV$. Next we describe how a few runs of RC_L (or $RGCV$) can produce inferences about the validity of the final choice. This is illustrated by some simulations in a two-dimensional setting.

Sequential Design for Inhomogenous Sobolev Class

G. Golubev
Institute for Problems of Information Transmission, Moscow

We are given noisy data $Y_j = f(t_j) + \xi_j, 1 \leq j \leq n, t_j \in [0,1]$, where ξ_j are i.i.d. Gaussian variables and the regression function $f(\bullet)$ belongs to a certain inhomogeneous Sobolev class. Our aim is to recover the regression function minimum IMSE. It is assumed that knots may be sequentially chosen. It is proved that optimal rate of convergence of IMSE cannot be improved by means of sequential design. Nevertheless it is pointed out the sequential design when we don't know the parameters of the Sobolev class.

Stochastic Search in Image Reconstruction

Friedrich Götze
Universität Bielefeld

We investigate the behaviour of MAP–estimators in statistical models for noise degraded images. Starting from a white noise model on pixel errors and a Bayesian prior (Gibbs–measure) on the set of images which uses a nearest neighbor interaction we give an estimate for the accuracy of the MAP–estimate for uniformly colored original images.

Furthermore, stochastic search algorithms approximating the MAP–estimate are considered. It is shown that in case of a unique and stable MAP–estimate (in the sense of Peierls) a logarithmic cooling schedule yields a polynomial relaxation time until the probability of having met the optimal image is larger than $\frac{1}{2}$, say using simulated annealing. These bounds diverge when the amount of smoothing increases.

In case of an Ising–type of every function with symmetric ground states it is shown that an appropriate Cluster algorithm of FKSW type admits a polynomial running time too. This is a result of E. Weineck which is part of his PhD–thesis.

Markov Chain Monte Carlo: Posterior Risk, Over-Relaxation and Sensitivity Analysis

Peter Green
University of Bristol

Two variants of basic MCMC methods were first introduced: Hastings methods using randomized proposal distributions, and methods based on partial conditioning. To quantify the statistical performance of MCMC estimators, the concept of relative increase in posterior risk was advocated: this is the excess in risk of the MCMC estimator over the corresponding one based on the exact posterior distribution. To investigate the merits of over-relaxation and of grouping the variables, the Gaussian case, with deterministic sweeping, was discussed, where an exact representation of the autocorrelation time is available. This can be used to guide the design of efficient sampling methods, even in the non-Gaussian case. Among general conclusions are that autocorrelation time (i) is reduced by over-relaxation; (ii) is unaffected by the sweep schedule, given the grouping; and (iii) is reduced by grouping in the case of strong positive association and a non-negative linear objective functional. Finally, a key application of MCMC was discussed: sensitivity to the prior in Bayesian analysis. Importance sampling ideas are useful here, and the infinitesimal version is easy to apply: the change in an estimated posterior expectation can be approximated by the empirical covariance between the function of interest and the log prior ratio.

Nonparametric Calibration

M.A. Gruet
INRA, Biométrie, Jouy-en-Josas

Statistical calibration analysis provides a way to predict a quantity ξ which is not directly observable from the observation of another easily measured quantity Z related to the just one by some dose-relationship. In many situations, the knowledge of the experimental process can hardly be translated into some parametric model. A tempting alternative may be provided by a nonparametric modelisation. Let $Z = r(\xi) + e$ when e is a centered random variable. The unknown regression function r is supposed to be smooth and strictly monotone. The available information about r is given by experimental data $Y_i = r(\xi_i) + \epsilon_i$ obtained from a calibration experiment. The ϵ_i's are i.i.d. with zero mean and the training sample $\{X_i, Y_i\}$ is independent of Z. A nonparametric method is proposed for estimating directly the parameter of real interest ξ, r being here considered a nuisance parameter. We propose to estimate ξ by a solution of the estimating equation

$$H_n(\xi) = \frac{1}{n}\sum \frac{1}{h}K\left(\frac{\xi - X_i}{h}\right)\Psi(Y_i - Z) = 0$$

where K is a convolution kernel, h the smoothing parameter and Ψ an odd function. The asymptotic properties of the estimate are discussed. Calibration intervals of the type $\{\xi, |H_n(\xi)| \leq c\}$ are constructed. This nonparametric technique is applied to an important biological problem, that of estimating the relative potency of a new product relative to a known one.

Loss and Risk in Bandwidth Selection

Birgit Grund
University of Minnesota

Currently, there is considerable disagreement whether "loss" ISE or "risk" MISE should be used in kernel bandwidth selection. In this talk, we contribute two new arguments to the ISE vs. MISE discussion. First, we demonstrate that MISE measures "risk" in the classical decision theoretic framework; this contradicts the common argument that expected ISE is the only proper decision theoretic risk in bandwidth selection. Second, we show that bandwidth procedures may display opposite relative performance with respect to expected ISE and MISE. This defends the commonly adopted viewpoint that MISE optimal bandwidths are good estimates for the ISE optimal bandwidth.

On the Statistical Estimation of Fractal Dimension

Peter Hall
Australian National University/CSIRO

Fractal models for surfaces form a convenient and readily interpretable way of describing roughness. Practical examples include metal surfaces (e.g. prior to painting), plastic wrap (roughness influences the tendency of bacteria to adhere), bearing–and–shaft interface (lubrication qualities are affected by roughness), etc.. Methods for gathering data about surface roughness will be described, and new developments of data enhancements will be discussed. Stochastic process models for rough surfaces will be suggested, and statistical methods for estimating fractal dimension be outlined. The properties of those methods, when applied to stochastic process models for surfaces, will be noted.

Estimating Densities on the Plane: Optimal Balloon Estimates

C. Huber
University of Paris V, France
Joint work with Peter Hall, Art Owen, and Alex Covertey

Given a sample of n observations from a density f on $\mathbb{R}^d$, a natural estimator of $f(x)$ is formed by counting the number of points in some region $\mathcal{R}$ around x and dividing this count by the d–dimensional volume of $\mathcal{R}$. An asymptotically optimal choice for $\mathcal{R}$ is presented here, with respect to $\mathbb{L}^2$ norm risk. The optimal "balloons" turn to be ellipsoids with shape depending on x. An extension of the idea, that uses a kernel function to put weight greater on points nearer x is given. Among non negative kernels, the familiar Epanechnikov kernel used with an ellipsoidal region is optimal. When using higher order kernels, the optimal regions shapes are related to $\mathbb{L}^p$ balls for even positive p.

Stochastic Optimisation: The Genetic Algorithm, Simulated Annealing and Applications in Image Analysis

Chris Jennison
University of Bath, UK

The Genetic Algorithm (GA) has been proposed as a method for optimising functions $f(x)$ of many variables, $x = (x_1, \ldots, x_\ell)$, with multiple local optima. It operates by creating a population of vectors x and letting these evolve over time in a manner similar to natural evolution. In choosing parents for future generations, priority is given for those with high values of $f(x)$. A "crossover" operation is used in producing some offspring — the two parent vectors are broken at a randomly chosen point, the first part of parent A and the second part of parent B are combined to form one offspring, the remaining parts are combined to form the other. A small amount of random mutation is also introduced into each new generation. The vector x with largest $f(x)$ seen in a certain number of generations is returned as the solution.

The algorithm defines a complex stochastic process which is difficult to characterize. On close examination, the customary theoretical justifications of the GA are seen to lack rigour. In contrast to the successes claimed in many papers, our empirical studies have shown that the GA can perform poorly on very simple problems. In a more complex problem, searching for a MAP image estimate, the GA is seen to work as a very inefficient analogue of Simulated Annealing.

On a positive note, we believe that the problems for which the GA is designed to be important — even if the GA is not the right method for these problems. There is a need for general-purpose methods to find near–optima of functions with structure $f(x) = \sum_c V_c(x)$ where each $V_c(x)$ depends only on $\{x_i : i \in c\}$ for a small set c. The treatment of main effects and low–order interactions in classical experimental design is relevant here, but dealing with a very large set of variables and the presence of local optima poses new challenges.

Wavelet Shrinkage: Ideal Spatial Adaption and Denoising

I. Johnstone
Stanford University
Joint work with David Donoho, Gérard Kerkyacharian and Dominique Picard

We give first an overview of the method of soft thresholding of wavelet coefficients in the idealized setting of non-parametric regression on the interval with equally spaced data and Gaussian white noise errors. Use of a threshold at $\sqrt{2\log n}$ standard deviations of the noise yields estimates that adapt spatially to oscillation and discontinuities and have a high degree of visual smoothness. Some supporting theorems are presented in connection with spatial adaption: wavelet thresholding can be tuned to come within a logarithmic multiplicative factor of an ideal risk attained by an oracle who knows the optimal set of coefficients to estimate. This result, valid for L^2 global average risk for arbitrary functions, can be sharpened over estimation spaces of functions, can be estimated at a given rate by other spatially adaptive procedures such as piecewise polynomial fits. This exploits oracle inequalities over weak-L_p balls along with characterisations of approximation spaces developed by De Vore, Jawerth and Popov. Finally, we present some results concerning "denoising" and universal near-minimaxity (in L_2) properties of the estimates over certain Besov and Triebel scales of function classes.

Calibration in a Nonlinear Setting

E. Jolivet
INRA, Laboratoire de Biométrie, Jouy-en-Josas
Joint work with M.A. Gruet

Let observe the following data $(Y_i, x_i), i = 1, \ldots, n,\ (Z_j), j = 1, \ldots, m$ such that

$$Y_i = f(x_i, \theta)) + \epsilon_i, Z_j = f(z_0, \theta) + \eta_j, \theta \in \Theta, E(\epsilon_i) = E(\eta_j) = 0.$$

We want to estimate z_0 using the training set (Y_i, x_i) and the response Z_j. We are interested in set estimates. No consistent asymptotic framework is directly obtainable for this problem. We propose to consider $m = n^\delta, 0 < \delta < 1, n \to \infty$. Under these assumptions, various asymptotic pivots can be derived, and Edgeworth expansions can be computed. On the other hand simulations can be performed. Although no one of these two methods provide a definitive answer concerning the actual level of the prediction sets derived, we remark some convergences of the results obtained by both ways.

Nonparametric Estimation of Common Regressors for Similar Curve Data

Alois Kneip
Wirtschaftstheorie II, Universität Bonn

The talk is concerned with data from a collection of different, but related regression curves $(m_j)_{j=1,\ldots,N}, N >> 1$. In statistical practice analysis of such data is most frequently based on low dimensional linear models. It is then assumed that each regression curve m_j is a linear combination of a small number $L << N$ of common functions $g_1, \ldots, g_L$. For example, if all m_j's are straight lines this holds with $L = 2$, $g_1 \equiv 1$ and $g_2(x) = x$. In this paper the assumption of a prespecified model is dropped. A nonparametric method is presented which allows to estimate the smallest L and corresponding functions $g_1, \ldots, g_L$ from the data. The procedure combines smoothing techniques with ideas related to Principal Component Analysis. An asymptotic theory is presented which yields detailed insight into properties of the resulting estimators. An application to household expenditure data illustrates the approach.

Smooth Curves in Biomechanics

Sue Leurgans
Rush–Presbyterian–St. Luke's Medical Center, Chicago, IL, USA

Many biomechanical measurements are smooth curves recorded on a fine mesh. Examples include joint angles in gait analysis, foot pressures from pedabarographs, and lumbar sphere position from Lumbar Motion Monitors. Even though the curves are known to be smooth, some data sets contain sample curves that are not smooth, because the computations are not entirely wellposed. Some simple plots can assist in the identification of curves that cannot correspond to actual motions. The most important plot superposes all trials from one subject on one plot.

Although smooth principal components are a natural method for such data, the Rice & Silverman (1991) estimators are not robust to unusual (smooth) curves. Sample influence functions for smooth eigenvalues and eigenvectors can be used to locate unusual curves when multiple trials are available for each subject, smooth extension of random effects multivariate analysis of variance can be applied.

Helices in High Dimensional Regression

Ker–Chau Li
Math Department, UCLA, USA

The shape of 14–dimensional Boston housing data looks like a helix or slide when viewed from a three dimensional scatterplot of the output variable (median house price) against the two projected regressors found by SIR (sliced inverse regression). From the first viewing angle, a linear fit seems appropriate. But as we spin the data, a nonlinear trend begins to emerge. These patterns are further confounded with the curvilinear association between the two projections shown on the screen as the output axis is turned off the sight. The helix–like shape of a data cloud highlights a typical nonlinear confounding problem in regression, rarely addressed in the literature. These are four questions to ask: How often does a helix occur? How to expose a helix in a high dimension setting? How seriously does a helix impair regression analysis and modeling? How to repair it? Strategies based on SIR type analysis are suggested. Demonstration on Mac Powerbook is given.

A Constrained Risk Inequality with Applications to Adaptive Estimation

Mark Low
University of Pennsylvania

A general constrained minimum risk inequality is derived. Given two densities f_θ and f_0 we find a lower bound for the risk at the point θ given an upper bound for the risk at the point 0. The inequality sheds new light on adaptive estimation problems arising in nonparametric functional estimation.

Locally Adaptive Regression Splines

E. Mammen
Humboldt-Universität, Berlin
Joint work with S. van de Geer

A penalized least squares estimate is considered for nonparametric regression. The penalty is the total variation of a (higher order) derivative of the regression function. It turns out that the estimate is a spline with data adaptive chosen knot points. An algorithm is presented which shows that the estimate is based on stepwise addition and deletion of knot points and soft thresholding of empirical spline coefficients. Global and local asymptotic results are presented. In particular when the total variation of the function itself is used as penalty, the estimate is at monotone pieces asymptotically equivalent to the least squares monotone estimate.

Visual Error Criteria from Qualitative Smoothing

J. S. Marron
University of North Carolina

An important gap between the classical mathematical theory and the practice and implementation of nonparametric curve estimating is due to the fact that the usual norm on function spaces searches something different from what the eye can see visually in a graphical presentation. Mathematical error criteria, which more closely follow "visual impression" are developed, and analysed from both graphical and mathematical viewpoints. Examples from wavelet regression and kernel density are considered.

A Fast and Easy Smoothing Algorithm with an Application to Environmental Monitoring Data

Karen Messer
California State University

Soil sampling data from the Nevada Test Site for nuclear weapons are presented. Observations of CS–137 concentration were taken over a nearly regular grid, with variable grid size over an irregular, non–convex sampling region. It is desired to estimate the total inventory of CS–137 present, and to estimate a specific concentration contour ("isopleth") of the CS–137 concentration profile. Originally these data were analysed using Kriging with a parametric model for the covariance, with unsatisfactory results. A fast and easily implemented kernel smoothing algorithm was presented, and used to give a boundary corrected kernel smooth of the data. Error bounds were presented. The smoother is well behaved, and has to recommend it that it is easy to code and to analyse.

Asymtotic Properties of Model Selection Procedures

Marlene Müller
Humboldt University, Berlin

The talk concerns the asymptotic properties of a class of criteria for model selection in linear regression models, which covers the most well known criteria as e.g. MALLOWS' C_p, CV (cross-validation), GCV (generalized cross-validation), AKAIKE's AIC and FPE as well as SCHWARZ' BIC. These criteria are shown to be consistent in the sense of selecting the true or larger models, assuming i.i.d. errors and the possible inadequacy of the linear model. Additionally we prove that BIC-type criteria select the true model if the sample size is converging to infinity. These consistency properties are completed by convergence results for the risk and loss of the estimated regression functions.

On the Smoothness of Stochastic Image Models

Frank Natterer
Universität Münster

Let $Ef(x) = 0$ and $Ef(x)f(y) = K(x-y)$ describe an imge f. A 1-D image f is said to be scale invariant if

$$E\frac{1}{2}(f(x_i) + f(x_{i-1}))\frac{1}{2}(f(x_j) + f(x_{j-1})) = cEf(x_i)f(x_j)$$

for some constant c. One can show that f is scale invariant iff $K(x) = e^{-\lambda|x|}$ with some constant $\lambda > 0$. An image in $\mathbb{R}^2$ is scale invariant if $K(x) = e^{-\lambda|x|}$ with some norm $|x|$ in $\mathbb{R}^2$. It is shown that $E\|f\|_{H^\alpha} < \infty$ iff $\alpha < \frac{1}{2}$, where H^α is the Sobolev space of order α. By computer simulations it is shown that the covariance $K(x) = e^{-\lambda|x|}$ in fact leads to images which are more pleasing than images with other choices of K.

On the Asymptotic Efficiency of Wavelet Estimators under Arbitrary Error Distributions

Michael H. Neumann
Institute for Applied Analysis and Stochastics, Berlin

We assume a nonparametric regression model with fixed, not necessarily identically distributed errors and regular design. Further, we suppose that the unknown regression function f lies in a certain ball $\mathcal{F}$ of some Besov space $W^m_{p,q}, 1 \leq p < 2, p \leq q$.

We show that under the above restrictions the minimax risk can be bounded below by the minimax risk under Gaussian errors and equidistant design. Here the Fisher informations of the location families derived from the error distributions enter into the formula of the variance of the Gaussian errors. If, additionally, the error distributions vary in certain Hellinger neighborhoods of fixed, not necessarily identical distributions, the inverses of the Fisher informations will be replaced by the respective variances of the central measures.

We prove that the optimal rate is attained by wavelet estimators with nonrandom thresholds. We adapt these thresholds by some cross-validation technique and show that these (random) thresholds are asymptotically as good as the optimal ones. The corresponding estimator attains the lower bound within certain analytical constants, which depend on the particular shapes of the error distributions.

Asymptotic Equivalence of Density Estimation and White Noise

M. Nussbaum
Institute for Applied Analysis and Stochastics, Berlin

Consider the problem of estimating a density f, defined on the unit interval, from n i.i.d. observations $y_i, i = 1, \ldots, n$. Suppose f is a priori known to be in a set Σ. We show that the corresponding experiment is asymptotically equivalent in the sense of Le Cam's deficiency distance to an experiment

$$d_y(t) = f^{1/2}(t)dt + \frac{1}{2}n^{-1/2}dW(t),\ t \in [0,1],\ f \in \Sigma,$$

if $\Sigma = \Sigma_{\epsilon,M}$, where $\Sigma_{\epsilon,M}$ is the set of densities bounded below by ϵ and having a 4th derivative bounded in L_2-norm by M^2. Equivalence means that $\Delta(.,.) \longrightarrow 0$ as $n \to \infty$ for the deficiency distance Δ between the experiments. This generalizes a result of Le Cam (1985) to the infinite dimensional case (Σ nonparametric) and has various implications for asymptotic decision theory in the i.i.d. model.

Binary Tree–structured Methods for Image Compression

Richard Olshen
Division of Biostatistics, Stanford University

Applications of CART ideas to clustering concern what in the engineering literature is termed pruned, tree–structured vector quantization (PTSVQ); it is used for compressing images. The basic idea is that of "two–means" clustering algorithms applied successively, that is, in a binary tree–structured manner. The "predictor" at each terminal node is the centroid of the learning sample vectors in the node. The depth of the tree averaged over a suitable learning or test sample of vectors is the "bit rate." The talk was in part a discussion of algorithms and in part a report on medical experiments with the algorithms. Original and compressed CT and MR images were shown. The work is a collaboration among engineers, radiologists, and statisticians led by Robert Gray. There was some exposition of asymptotic properties of the algorithms, both their almost sure consistency in a described sense, and an empirical study of the rates of convergence of "distortions" of "codebooks" to their asymptotic values.

Wavelet Shrinkage: Density Estimation in Nonlinear Cases

D. Picard
Paris VII
Joint work with D. Donoho, I. Johnstone, G. Kerkyacharian

We present different problems in density estimation of nonlinear type where a wavelet shrinkage procedure (or other nonlinear wavelet treatment) achieve the minimax rate of the problem or some time at least nearly (up to a $\log n$-term), whereas it can be proved that all the linear procedures fail to have the proper rate:

1. Estimate f, $f \in B_{spq}$ with loss function $\|\hat{f} - f\|_p^{p\prime}, p\prime > p$.
2. Estimate f, $f \in \bigcup_{s,p} B_{spq}$ with loss function $\|\hat{f} - f\|_p^{p\prime}$.
3. Estimate $\int f^3$.

In the first two questions the proposed estimate is:

$$\sum_k \hat{\alpha}_{J_1(n),k}\phi_{J_1(n),k} + \sum_{J_1(n)}^{J_2(n)} \sum_k S_{\delta_n(j)}(\hat{\beta}_{J,k})\psi_{J,k}$$

where ϕ, ψ are father, mother wavelet

$$\hat{\alpha}_{J,k} = \frac{1}{n}\sum_i \phi_{J,k}(x_i), \quad \hat{\beta}_{J,k} = \frac{1}{n}\sum_i \psi_{J,k}(x_i), \quad S_{\delta_n(j)}(x) = x 1_{|x|>\delta_n(j)}.$$

It is worthy to observe that the last part of the sum is a nonlinear improvement of the first part which behaves just like a classical kernel or orthogonal series method.

Random Sets and Density Estimation

Wolfgang Polonik
Institut für angewandte Mathematik, Heidelberg

By using empirical process theory we study random sets in $\mathbb{R}^d$ which we call "empirical generalized λ-clusters". These sets can be used for estimating the density contour clusters of an underlying density and also for estimating the density itself. Let F be a distribution on $\mathbb{R}^d$ which has a Lebesgue density f and let $X_1, \ldots, X_n$ be i.i.d. observations drawn from F. Let F_n denote the empirical measure of the n observations. For any class $\mathcal{C}$ of measurable subsets of $\mathbb{R}^d$ the "empirical generalized λ-clusters in $\mathcal{C}$", denoted by $\Gamma_{n,\mathcal{C}}(\lambda)$, are defined as those sets in $\mathcal{C}$ where the supremum of $F_n(C) - \lambda Leb(C)$ over $\mathcal{C}$ is attained. Here Leb denotes the Lebesgue measure. We give consistency results and rates of convergence. Moreover, if "$\Gamma(\lambda) \in \mathcal{C}$ for all $\lambda \geq 0$", then the underlying density f can be estimated by means of the sets $\Gamma_{n,\mathcal{C}}(\lambda), \lambda \geq 0$. Consistency results and rates of convergence of the corresponding density estimator will be given. The assumption "$\Gamma(\lambda) \in \mathcal{C}$ for all λ" can be used for modeling qualitative aspects of the underlying distribution, as for example modality, symmetry, monotonicity. For a special choice of $\mathcal{C}$, our density estimator equals the Grenander estimator, which is the maximum likelihood estimator of a monotone density.

Data Dependent Triangulations and Applications to Surface Fitting and Image Compression

Larry Schumaker
Vanderbilt University, Nashville, Tenessie

We consider the space of bivariate Splines $\mathcal{S}_d^r(\Delta) = \{f \in C^r : f|_{T_i} \in \mathcal{P}_d\}$ where $\mathcal{P}_d$ is the set of polynomials of degree d and Δ is a triangulation in $\mathbb{R}^2$. The idea is to adjust Δ to the data being fit. This can be done by swapping edges in Δ to improve some criterion of interest. For example, using $\mathcal{S}_1^0(\Delta)$, we can interpolate at the vertices and adjust to minimize smoothness of the surface. Alternatively, we may seek to minimize ℓ_2-goodness of fit assuming we are given data at points other than the vertices. To apply these ideas to image compression, we simply observe that the image can be regarded as a large (say 512×512) set of data ($z_{ij} \in \{0, \ldots, 255\}$) on a regular grid. To fit the image we begin with an initial Δ, the adjust it. The fit can be further improved by adding new vertices, and further compressed by deleting vertices (on triangles). To reconstruct the image, one simply evaluates the spline on the grid.

What can we do when the Data are Curves?

Bernard Silverman
University of Bath, England

An increasing number of problems involve functional data analysis where the data points are "curves" or functions rather than numbers or vectors. Extensions of principal components analysis and canonical correlation analysis to the functional context (Rice and Silverman, 1991; Leurgans, Moyeed and Silverman, 1993; both JRSS-B) are described. In PCA, a method involving smoothing is desirable but not essential; in CCA smoothing, in the same form, is essential. Both PCA and CCA can be used, in different ways, to analyse "bivariate" data of the form $\{(X_i(s), Y_i(s)), i = 1, \ldots, n\}$. The methods were illustrated by reference to two sets of data, one on human gait and one on ozone levels. There are many other important potential applications.

Conservative Finite–Sample Confidence Envelopes in Density Estimation

P.B. Stark
Department of Statistics, UC, Berkeley, CA 94720
Joint work with N.W. Hengartner

If a density is known to be monotone or to have K or fewer modes, a conservative finite–sample confidence region for the density can be found by solving a finite set of finite–dimensional linear programs. The method requires no conventional smoothness assumptions, but if the density is Lipschitz (ρ) in a neighborhood of the point x containing no modes, the bounds at x converge at the rate $n^{-\frac{\rho}{2(1+\rho)}}$. The method also gives constructive nonparametric lower confidence intervals for the number of modes of the density.

Hazard Regression

Charles J. Stone
Department of Statistics, UC, Berkeley, CA 94720
Joint work with Charles Kooperberg and Young Truong

A version of hazard regression is discussed in which the log–hazard function is modeled as a sum of functions of at most two variables (covariates and/or time) with polynomial splines and their tensor products being used to fit the various functions in the model. Maximum likelihood, stepwise addition, stepwise deletion and BIC are combined to obtain the final fit. Rates of convergence for nonaptive versions of this methodology are also presented.

Improved Estimation under Random Censorship

W. Stute
University of Giessen, Germany

We briefly review the (very short) history of "Mean lifetime" estimation under random censorship. Bias considerations lead one to modify the Kaplan-Meier estimator. In a simulation study it is demonstrated that for small to moderate sample sizes the new estimator outperforms Kaplan-Meier (though the asymptotics are the same). Also confidence intervals for the mean lifetime are constructed which work pretty well.

Regeneration in Markov Chain Samplers

Luke Tierney
School of Statistics, University of Minnesota, Minneapolis, MN 55455

Markov chain sampling has received considerable attention in the recent literature, in particular in the context of Bayesian computation and maximum likelihood estimation. This talk discusses the use of Markov chain splitting, originally developed as a tool for the theoretical analysis of general state space Markov chains, to introduce regeneration times into Markov chain samplers. This allows the use of regenerative methods for analyzing the output of these samplers, and can also provide a useful diagnostic of the performance of the samplers. The general approach is applied to several different samplers and is illustrated in a number of examples.

Asymptotically Minimax Estimation of Sets and Boundaries

Alexander Tsybakov
Université Catholique de Louvain

The problem of estimating an unknown set G from the data driven from the distribution indexed by this set is considered. The two particular setups are treated: The "regression" setup and the "density" setup. For the regression setup the pairs (X_i, Y_i), $i = 1, \ldots, n$ are sampled such that, conditionally on $X_i \in G$, Y_i has the density $p_1(\bullet|X_i)$, and conditionally on $X_i \notin G$ it has the density $p_2(\bullet|X_i)$, such that the Hellinger distance between p_1 and p_2 is bounded away from zero uniformly in X_i. For the density setup it is assumed that G is a support of unknown probability density, and the sample $X_1, \ldots, X_n$ is taken from this density. It is shown that, under general conditions, there exist two types of estimators which achieve minimax rates of convergence to the true G in Hausdorff metric and in measure–of–symmetric–difference metric. The rates of convergence are related to the entropy properties of the classes of sets G.

Discretized Version of Average Derivative Estimators

Berwin A. Turlach
C.O.R.E. & Institut de Statistique, Université Catholique de Louvain

Average Derivative Estimation is a nonparametric method which may be used to estimate the unknown parameter β in the single index model $E[Y|X = x] = g(x^T\beta)$. Like most nonparametric methods it depends on the choice of a smoothing parameter. Analyzing samples will include a variation of this smoothing parameter which may lead to an enormous amount of calculations which make an interactive analysis impossible. In this talk we demonstrate how the computational burden can be reduced by discretization methods, i.e., binning the data. A simulation study shows that in the multivariate setting histogram binning is preferable over linear binning.

Sieves and Maximum Likelihood

Sara van de Geer
University of Leiden

We consider independent observations $X_1, \ldots, X_n$ with distribution depending on some unknown parameter $\theta_0 \in \Theta$, where Θ is a given parameter space. The sieved estimator $\hat{\theta} \in \Theta_N$ is obtained by minimizing a loss function $L(\theta) = \sum_{i=1}^{n} f_\theta(x_i)$ over a sieved parameter space Θ_N. Entropy methods are used to establish uniform probability inequalities for $L(\theta)$. As example, we show that a rate of convergence in Hellinger distance for the sieved maximum likelihood estimator follows from the entropy with bracketing of the class of densities endowed with Hellinger metric.

Additive Model for Repeated Measurements: A Nonparametric Point of View

Philippe Vieu
Université Paul Sabatier, 39062 Toulouse

The aim of this talk is to present a recent paper by Joel Baularan, Louis Ferre and myself (JSPI, 1993, in print). In this paper we develop a non-parametric approach, based on kernel estimation, to deal with a two-stage model for repeated measurements. One of the main interest of the proposed method is that it works well for unbalanced data as for balanced ones, while previous non-parametric two-stage approaches need crucially a balanced data set. The use of a two-stage model allows to get, in a first stage, an estimated "mean" curve which is obtained as if all the data where collected from a single individual. Then, in second stage, we can estimate individual curves by using not only the data collected for each individual but also the "mean" estimated curve. In fact, each individual is used to estimate the mean curve, and so information for each individual can be used for the other ones.

In this work we present asymptotics (including L_1, L_2 and L_∞ rates of convergence and optimal smoothing parameter selection). Then two examples are presented. The first one is a classical growth curve situation, which is a balanced data set. The second one is a geophysical data set and our method is particularly relevant on it because it is unbalanced.

Synchronous Random Fields for Image Reconstruction

L. Younes
Université Paris Sud

We investigate a class of random fields on a finite lattice which are invariant under the action of a synchronous probability kernel, i.e. a transition $P(x, y)$ which factorizes under the form

$$\prod_{s \in S} \pi^s(g, x_s),$$

where S is the given lattice. After reviewing the necessary and sufficient condition for P to be reversible w.r.t. its invariant distribution, we present an extension of the model by considering P-dependent Markov-chain with some properties generalizing reversibility in a proper way. It is then shown how this class of model may be used in practical applications, namely Bayesian reconstruction of images, even in the presence of blur results of experiments are finally given for this last context.

Berichterstatter:

Sigbert Klinke, Marlene Müller, Berwin Turlach

Tagungsteilnehmer

Prof.Dr. Rudolf J. Beran
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720
USA

Prof.Dr. David L. Donoho
Department of Statistics
Stanford University

Stanford, CA 94305
USA

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

F-75252 Paris Cedex 05

Prof.Dr. Paul Doukhan
Mathématiques
Université de Paris Sud (Paris XI)
Centre d'Orsay, Bâtiment 425

F-91405 Orsay Cedex

Prof.Dr. Olaf Bunke
Institut für Stochastik
Humboldt-Universtität Berlin
Unter den Linden 6

O-1086 Berlin
GERMANY

Prof.Dr. Geoff K. Eagleson
Australian Graduate School of
Management
The University of New South Wales
P.O. Box 1

Kensington N.S.W. 2033
AUSTRALIA

Prof.Dr. Peter Clifford
Mathematical Institute
Oxford University
24 - 29, St. Giles

GB- Oxford , OX1 3LB

Prof.Dr. Jianqing Fan
Statistics Department
University of North Carolina
Phillips Hall

Chapel Hill , NC 27599
USA

Prof.Dr. Dennis D. Cox
Dept. of Statistics
University of Illinois
725 S Wright St.

Champaign , IL 61820
USA

Prof.Dr. Jürgen Franke
Fachbereich Mathematik
Universität Kaiserslautern
Postfach 3049

W-6750 Kaiserslautern
GERMANY

Gefördert durch
DFG Deutsche Forschungsgemeinschaft

Prof.Dr. Theo Gasser
Department of Biostatistics ISPM
University of Zürich
Sumatrastr. 30

CH-8006 Zürich

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

W-4800 Bielefeld 1
GERMANY

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

NL-2300 RA Leiden

Prof.Dr. Peter J. Green
Department of Mathematics
University of Bristol
University Walk

GB- Bristol , BS8 1TW

Dr. Irene Gijbels
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 34

B-1348 Louvain-La-Neuve

Dr. Marie-Anne Gruet
Département de Biométrie
INRA

F-78352 Jouy-en-Josas Cedex

Prof.Dr. Didier A. Girard
Laboratoire LMC-IMAG
BP 53X
Université Joseph Fourier

F-38041 Grenoble Cedex

Dr. Birgit Grund
Department of Applied Statistics
School of Statistics
University of Minnesota
Classroom Office Building 352

St. Paul , MN 55108-6042
USA

Prof.Dr. Georgü K. Golubev
Institut für Angewandte Analysis
und Stochastik, Abt. 6
Hausvogteiplatz 5-7

O-1086 Berlin
GERMANY

Prof.Dr. Peter G. Hall
Department of Statistics
Australian National University
G.P.O. Box 4

Canberra ACT 26 01
AUSTRALIA

Prof.Dr. Wolfgang Härdle
Inst.für Statistik und Ökonometrie
Humboldt Universität zu Berlin
FB Wirtschaftswissenschaften
Spandauer Str. 1

O-1020 Berlin
GERMANY

Prof.Dr. Emmanuel Jolivet
Département de Biométrie
INRA

F-78352 Jouy-en-Josas Cedex

Prof.Dr. Nancy E. Heckman
Department of Statistics
The University of British Columbia
2021 West Mall

Vancouver , B.C. V6T 1Z2
CANADA

Sigbert Klinke
Center for Operations Research and
Econometrics
Université de Louvain
34 Voie du Roman Pays

B-1348 Louvain-la-Neuve

Prof.Dr. Catherine Huber
71 Boulevard des Etats Unis

F-78110 Le Vesinet

Dr. Alois R. Kneip
Wirtschaftstheorie II
Instit. f. Gesellschafts- u. Wirt-
schaftswiss. d. Universität Bonn
Adenauerallee 24 - 26

W-5300 Bonn 1
GERMANY

Dr. Chris Jennison
School of Mathematical Sciences
University of Bath
Claverton Down

GB- Bath Avon, BA2 7AY

Dr. Sue E. Leurgans
Dept.of Prevenetive Medicine
Rush-Presbyterian-St.Luke's Medical
Center, Suite 117
1725 West Harrison Str.

Chicago , IL 60612
USA

Prof.Dr. Iain M. Johnstone
Department of Statistics
Stanford University
Sequoia Hall

Stanford , CA 94305-4065
USA

Prof.Dr. Ker-Chau Li
Dept. of Mathematics
University of California
405 Hilgard Avenue

Los Angeles , CA 90024-1555
USA

Prof.Dr. Mark Low
724 Cornelia Place

Philadelphia , PA 19118
USA

Marlene Müller
Inst.für Statistik und Ökonometrie
Humboldt-Universität zu Berlin
FB Wirtschaftswissenschaften
Spandauer Str. 1

O-1020 Berlin
GERMANY

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

W-6900 Heidelberg 1
GERMANY

Prof.Dr. Frank Natterer
Institut für Numerische und
Instrumentelle Mathematik
Universität Münster
Einsteinstr. 62

W-4400 Münster
GERMANY

Prof.Dr. James Stephen Marron
Statistics Department
University of North Carolina
Phillips Hall

Chapel Hill , NC 27599-3260
USA

Michael Neumann
Institut für Angewandte Analysis
und Stochastik
im Forschungsverbund Berlin e.V.
Hausvogteiplatz 5-7

O-1086 Berlin
GERMANY

Prof.Dr. Karen Messer
Department of Mathematics
California State University

Fullerton , CA 92634
USA

Dr. Michael Nussbaum
Institut für Angewandte Analysis
und Stochastik
im Forschungsverbund Berlin e.V.
Hausvogteiplatz 5-7

O-1086 Berlin
GERMANY

Prof.Dr. Dietrich Werner Müller
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294

W-6900 Heidelberg 1
GERMANY

Prof.Dr. Richard A. Olshen
Department of Statistics
Stanford University
Sequoia Hall

Stanford , CA 94305-4065
USA

Prof.Dr. Dominique Picard
U. F. R. de Mathématiques
T. 45-55, 5ème étage
Université de Paris VII
2, Place Jussieu

F-75251 Paris Cedex 05

Prof.Dr. Charles J. Stone
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720
USA

Dr. Wolfgang Polonik
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294

W-6900 Heidelberg 1
GERMANY

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

W-6300 Gießen
GERMANY

Prof.Dr. Larry L. Schumaker
Dept. of Mathematics
Vanderbilt University
P.O. Box 6213 - B

Nashville , TN 37235
USA

Prof.Dr. Luke Jon Tierney
Dept. of Statistics
University of Minnesota
270 Vincent Hall
206 Church Street S. E.

Minneapolis , MN 55 455
USA

Prof.Dr. Bernard W. Silverman
School of Mathematical Sciences
University of Bath
Claverton Down

GB- Bath Avon, BA2 7AY

Prof.Dr. Alexander B. Tsybakov
CORE
Université Catholique de Louvain
34 voie du Roman Pays

B-1348 Louvain-la-Neuve

Prof.Dr. Phil Stark
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720
USA

Berwin Turlach
Center for Operations Research and
Econometrics
Universite de Louvain
34 Voie du Roman Pays

B-1348 Louvain-la-Neuve

Prof.Dr. Philippe Vieu
Laboratoire de Statistique et
Probabilités
Université Paul Sabatier
118, route de Narbonne

F-31062 Toulouse Cedex

Prof.Dr. Laurent Younes
Mathématiques
Université de Paris Sud (Paris XI)
Centre d'Orsay, Bâtiment 425

F-91405 Orsay Cedex

E-MAIL

Beran, R.	beran@stat.berkeley.edu
Birgé, L.	lb@ccr.jussieu.fr
Clifford, P.	clifford@stats.ox.ac.uk
Cox, D.	dcox@rice.edu
Donoho, D.	donoho@playfair.stanford.edu
Doukhan, P.	doukhan@stats.matups.fr
Eagleson, G.	G.Eagleson@unsw.edu.au
Fan, J.	jfan@vega.sph.unc.edu, jfan@biostat.sph.unc.edu
Gasser, Th.	k320259@czhrzu1a.bitnet
Gijbels, I.	gijbels@stat.ucl.ac.be
Girard, D.	girard@frgag51.bitnet
Götze, F.	goetze@mathematik.uni-bielefeld.de
Green, P.	P.J.Green@bristol.ac.uk
Gruet, M.A.	mag@jouy.inra.fr
Grund, B.	birgit@umnstat.stat.umn.edu
Hall, P.	halpstat@durras.anu.edu.au
Härdle, W.	haerdle@wiwi.hu-berlin.de
Heckman, N.	nancy@stat.ubc.ca
Huber, C.	huber@frciti51.bitnet
Jennison, C.	cj@maths.bath.ac.uk
Jolivet, E.	jolivet@jouy.inra.fr
Johnstone, I.	imj@playfair.stanford.edu
Kerkyacharian, G.	kerk@frmap711.mathp7.jussieu.fr
Klinke, S.	klinke@stat.ucl.ac.be
Kneip, A.	akneip@trick.or.uni-bonn.de
Leurgans, S.	leurgans@is.rpslmc.edu
Mammen, E.	enno@statlab.uni-heidelberg.de
Messer, K.	messer@osiris.ucsd.edu
Müller, M.	marlene@wiwi.hu-berlin.de
Neumann, M.H.	neumann@iaas-berlin.dbp.de
Nussbaum, M.	nussbaum@iaas-berlin.dbp.de
Olshen, R.	olshen@playfair.stanford.edu
Picard, D.	picard@frmap711.mathp7.jussieu.fr
Polonik, W.	polonik@statlab.uni-heidelberg.de
Schumaker, L.	s@mars.cas.vanderbilt.edu
Silverman, B.W.	bws@maths.bath.ac.uk
Stark, P.B.	stark@stat.berkeley.edu
Stone, C.J.	stone@stat.berkeley.edu
Tierney, L.	luke@umnstat.stat.umn.edu
Tsybakov, A.	mcore2@bucll n11.bitnet
Turlach, B.A.	turlach@core.ucl.ac.be
Younes, L.	younes@dmi.ens.fr

Gefördert durch
DFG Deutsche Forschungsgemeinschaft