Probability and Statistical Decision Theory

Volume A

Edited by

F. Konecny,

J. Mogyoródi, and

W. Wertz

PROBABILITY AND STATISTICAL DECISION THEORY

Proceedings of the 4th Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Austria, 4–10 September, 1983

Volume A

Edited by

F. KONECNY

University of Agriculture, Vienna, Austria

J. MOGYORÓDI

Eötvös Loránd University, Budapest, Hungary

W. WERTZ

Technical University of Vienna, Austria

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER / TOKYO

Library of Congress Cataloging-in-Publication Data

CIP

Pannonian Symposium on Mathematical Statistics (4th:
1983: Bad Tatzmannsdorf, Austria)
Probability and statistical decision theory.

(Proceedings of the 4th Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Austria, 4–10 September 1983; v. A)
Bibliography: p.
Includes index.
1. Probabilities—Congresses. 2. Statistical decision—Congresses. I. Mogyoródi, J. II. Konecny, F. III. Wertz, Wolfgang. IV. Series: Pannonian Symposium on Mathematical Statistics (4th: 1983: Bad Tatzmannsdorf, Austria). Proceedings of the 4th Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Austria, 4–10 September 1983; v. A.
QA276.A1P36 1983 vol. A 519.5 s [519.2] 85-14465 [QA273.A1]
ISBN 90-277-2089-4

Distributors for the U.S.A. and Canada
Kluwer Academic Publishers,
190 Old Derby Street, Hingham, MA 02043, U.S.A.

Distributors for Albania, Bulgaria, Chinese People's Republic,
Cuba, Czechoslovakia, German Democratic Republic, Hungary,
Korean People's Republic, Mongolia, Poland, Romania,
the U.S.S.R., Vietnam and Yugoslavia
Kultura Hungarian Foreign Trading Company
P.O.B. 149
H-1389 Budapest 62, Hungary

Distributors for all remaining countries
Kluwer Academic Publishers Group,
P.O.Box 322, 3300 AH Dordrecht, Holland

Joint edition published by D. Reidel Publishing Company
Dordrecht, Holland
and
Akadémiai Kiadó, Budapest, Hungary

Printed in Hungary

PREFACE

The Fourth Pannonian Symposium on Mathematical Statistics was held in Bad Tatzmannsdorf, Austria, 4-10 September, 1983. The first two Symposia were held there in 1979 and 1981, whereas the Third Symposium was staged in Visegrád, Hungary in 1982. The proceedings volumes of these conferences, published by Springer, D. Reidel, and D. Reidel & Akadémiai Kiadó, respectively, give information about the objective of the Pannonian Symposia and the topics covered.

About 130 participants from 17 countries took part in this Fourth Symposium, and 92 lectures were presented. This volume contains 24 reviewed contributions which are mainly mathematically oriented. A second group of papers dealing with problems of applied statistics, probability theory and related topics is published in a separate volume entitled "Mathematical Statistics and Applications".

The contributions dealing with probability theory concentrate on two (intersecting) main topics: on the one hand, stochastic processes, and, on the other hand, limit theorems and invariance principles: Gaussian Processes, approximation of the Wiener Process, distribution of spacings and of order statistics, limit theorems in triangular arrays. Besides, adjacent topics like erdogic theory, maximal inequalities, approximation of convolutions of distribution functions, stochastic programming etc. are dealt with. In many of these contributions stress is laid upon the weakening of the assumption of independence.

The subjects of the statistical contributions are even more homogeneous: there are decision-theoretic papers (treating admissibility, sufficiency, limits of experiments, etc.) and papers on nonparametric estimation (nonparametric estimators of regression curves, further rank-, L- and shrinkage estimators). One paper gives a comprehensive survey of two-sided parametric tests.

The reader will observe at once the close connections between many of the questions under consideration; in particular, many results of papers which can be related to probability theory have immediate applications in statistics, and some belong to the common boundary of these subjects.

We wish to express our thanks to many persons who gave us indispensable aid in the edition of this volume: S. Csörgő, P. Deheuvels, M. Denker, L. Devroye, W. Eberl jr., B. Gyires, L. Györfi, J. Hurt, M. Hušková, I. Kátai, A. Kozek, W. Philipp, D. Plachky, P. Révész, P.K. Sen, F. Schipp and W. Sendler for their help in refereeing the papers and in other editorial matters, the secretarial staff of Professor I. Kátai at the University of Budapest, who did the laborious typing of the manuscript under the supervision of Mrs. Z. Andrásné Králik and, last but not least, Akadémiai Kiadó and D. Reidel Publishers for their good cooperation. (After the refereeing process and retyping, all papers were returned to the authors for correction.)

The organization of the Symposium was made possible by the help of many individuals and institutions. The organizers gratefully acknowledge the generous support given by the State Government of Burgenland, the Federal Ministry of Science and Research, the Austrian Statistical Society, the Control Data Co., Hypobank of Burgenland Co., the Volksbank Oberwart Co., the Raiffeisenbank Oberwart Co., the Kurbad Tatzmannsdorf Co. and the Local Autority of Bad Tatzmannsdorf; special thanks go to Th. Kery and Dr. R. Grohotolsky (Head and Vice-Head of Burgenland resp.), Dr. J. Karall (Member of the State Government of Burgenland). DDr. Schranz (Member of State Parliament of Burgenland) and Mag. R. Luipersbeck (Director of the Kurbad Tatzmannsdorf AG) for their help in many respects. Finally, we express

our thanks to the ladies for their splendid work in the preparation and local organization of the meeting, in particular to Ms. Ingrid Danzinger who patiently undertook most of the typing for the organization.

Wolfgang WE[illegible]

CONTENTS

Preface ix

INVITED PAPERS

P. DEHEUVELS
Spacing and Applications 1

U. MÜLLER-FUNK, F. PUKELSHEIM and H. WITTING
Locally Most Powerful Tests for Two-Sided Hypotheses 31

P. RÉVÉSZ
Approximation of the Wiener Process and its Local Time 57

CONTRIBUTED PAPERS

L. DEVROYE and L. GYÖRFI
Distribution-Free Exponential Bound for the L_1-Error of Partitioning-Estimates of a Regression Function 67

R. DUNCAN and D. SZYNAL
Pointwise Convergence Theorems for Some Classes of Operators on $L^p(\mu)$, $1\leq p\leq\infty$ 77

W. EBERL, Jr. and O. MOESCHLIN
On Measurability in Stochastic Programming 97

L. GAJEK
Sufficient Conditions for Admissibility 107

Z. GRUDZIEŃ and D. SZYNAL
On the Expected Values of k-th Record Values and Associated Characterizations of Distributions 119

G. HAIMAN
Almost Sure Asymptotic Bounds for the Extremes of a Stationary Sequence 129

N. HERRNDORF
A Note on the Invariance Principle for Strongly Mixing Sequences of Random Variables 139

M. HUŠKOVÁ and J. JUREČKOVÁ
Asymptotic Representation of R-Estimators of Location 145

M. IOSIFESCU
Ergodic Piecewise Monotonic Transformations and Dependence with Complete Connections 167

Z.A. IVKOVIĆ
On Non-Linear Time-Domain Analysis of Gaussian Processes 175

A. JANSSEN
The Convergence of Almost Regular Statistical Experiments to Gaussian Shifts 181

J. JUREČKOVÁ
Tail-Behaviour of L-Estimators and M-Estimators 205

A. KOVÁTS
On the Deviation of Distributions of Sums of Independent Integer Valued Random Variables 219

K.S. KUBACKI
On a Random-Sums Limit Problem 231

J.P. LECOUTRE
Almost Complete Convergence of the Statistically Equivalent Blocks Estimator of the Regression Function 265

D. MUSSMANN
Generating Systems for L-Spaces with Applications to Sufficiency 277

D. PLACHKY
A Converse of Lehmann-Scheffé's Theorem About MVU-Estimators 287

R.-D. REISS
Asymptotic Expansions of Moments of Central Order Statistics 293

Z. RYCHLIK and I. SZYSZKOWSKI
On the Rate of Convergence for Linear Functionals of Sums of Martingale Differences 301

P.K. SEN and A.K. EHSANES SALEH
Nonparametric Shrinkage Estimators of Location in a Multivariate Simple Regression Model 311

Á. SOMOGYI
Some Results on Dependent R.V.'s under Moment Type Restrictions 325

Subject Index 339

Proc. of the 4th Pannonian Symp. on Math. Stat., Bad Tatzmannsdorf, Austria 1983
F. Konecny, J. Mogyoródi, W. Wertz, eds

DISTRIBUTION FREE EXPONENTIAL BOUND FOR THE L_1 ERROR OF PARTITIONING-ESTIMATES OF A REGRESSION FUNCTION

LUC DEVROYE AND LÁSZLÓ GYÖRFI

Montreal and Budapest

ABSTRACT

Let $m_n(x)$ be a partitioning estimate of a regression function $m(x) = \mathbb{E}(Y/X=x)$ for the sample $(X_1,Y_1) \ldots (X_n,Y_n)$. If $|Y| \le K < \infty$ a.s. and μ denotes the probability measure of X then, under some mild conditions on the sequence of partitions, for each $\varepsilon>0$ there exists an n_0 such that

$$\mathbb{P}(\int |m_n(x) - m(x)|\mu(dx)>\varepsilon) \le \exp(-c(\varepsilon/K)^2 n)$$

for every $n \ge n_0$ where c is a universal constant.

INTRODUCTION

Let (X,Y) $(X_1,Y_1), \ldots (X_n,Y_n)$ be independent identically distributed $\mathbb{R}^d x[-K,K]$ valued random vectors, and let $m(x) = E(Y/X=x)$ be the regression function of Y on X, to be estimated from the data $D_n = \{(X_1,Y_1), \ldots (X_n,Y_n)\}$.

Let μ, $\tilde{\mu}$ be the probability measures of X and (X,Y), μ_n and $\tilde{\mu}_n$ be the empirical probability measure for $X_1, X_2, \ldots X_n$ and $(X_1,Y_1), \ldots (X_n,Y_n)$, respectively.

The signed measures $\nu_n(A)$ and $\nu(A)$ are defined by

$$\nu_n(A) = \frac{1}{n}\sum_{i=1}^{n} I_{[X_i \in A]} Y_i$$

and (1)

$$\nu(A) = \mathbb{E}\nu_n(A).$$

for every Borel set A.

<u>A partitioning estimate</u> is based on a finite or countable infinite Borel measurable partition $\mathcal{P}_n$ of $\mathbb{R}^d$ ($\mathcal{P}_n = \{A_{n_1}, A_{n_2}, \ldots\}$). If $A_n(x)$ is the set from $\mathcal{P}_n$ to which x belongs, then the estimate is given by

$$m_n(x) = \begin{cases} \frac{\nu_n(A_n(x))}{\mu_n(A_n(x))}, & \text{if } \nu_n(A_n(x))>0 \\ \frac{1}{n}\sum_{i=1}^{n} Y_i, & \text{otherwise.} \end{cases} \quad (2)$$

It is called fixed, if $\mathcal{P}_n$ does not depend upon the data, and variable, if $\mathcal{P}_n$ depends upon the data (for measurability reasons, in latter case, we have to assume the membership function mapping x and D_n to {1,2, ...} to be Borel measurable). Fixed partitioning estimates are also called histogram estimates.

A sequence of estimates is said to be <u>cubic</u> if there exist positive numbers $a_1, \ldots a_d$ such that each $\mathcal{P}_n$ consists entirely of sets of the form $\prod_{i=1}^{n} [a_i j_i h_n, a_i(j_i+1)h_n)$ where $j_1, \ldots j_d$ are integers and $\{h_n\}$ is a sequence of positive numbers.

Cubic sequences are particularly attractive from the computational point of view, because the set $A_n(x)$ can be determined for each x in constant time, independent of the data size n. This property is not shared by other sequences: for example, if all the A_{ni}'s are rectangles of different proportions, the membership determination can be done in time bounded from below by logn provided that a binary tree is created for locating an interval. In most cases, partitioning estimates are computationally superior to the other nonparametric estimates, particularly if the search for $A_n(x)$ is organized using binary decision trees (see for example, Friedman (1977)).

In this note we study only the fixed partitioning. The weak universal consistency of variable partitioning estimate was established by Gordon and Olshen (1980).

MAIN RESULT

Theorem If for each sphere S centered at the origin

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i:A_{ni}\cap S\neq\emptyset} 1 = 0 \tag{3}$$

and

$$\lim_{n\to\infty} \max_{i;A_{ni}\cap S\neq\emptyset} \{\sup_{x,y\in A_{ni}} ||x-y||\} = 0 \tag{4}$$

then for the partitioning estimate (2) and for each $\varepsilon>0$ there exists n_o such that

$$\mathbb{P}(\int|m_n(x) - m(x)|\mu(dx)\geq\varepsilon)<\exp(-c(\varepsilon/K)^2 n) \qquad n\geq n_o$$

where c is universal (for example c=1/400).

Remark 1 If $\mathcal{P}_n$ is a cubic sequence of partitions then the conditions (3) and (4) are equivalent to

$$\lim_{n\to\infty} h_n = 0, \quad \lim_{n\to\infty} nh_n^d = \infty \tag{6}$$

Applying Stone's Theorem (Stone, 1977) it is easy to check that for cubic partitions (6) is necessary and sufficient condition for

$$\lim_{n\to\infty} \mathbb{E}(\int |m_n(x) - m(x)|\mu(dx)|) = 0$$

for all distributions of (X,Y) to hold.

Remark 2 One of the most important consequences of this theorem is the exponential rate of convergence of the error probability in discrimination. If $X, X_1, \ldots X_n$ are as before, $Y, Y_1, \ldots Y_n$ take values in $\{1,2, \ldots M\}$ and the task of the discrimination is to decide on Y given X and D_n, then partitioning rule is defined as follows:

$$g_n(X)=i \text{ if } \sum_{j=1}^{n} I_{[X_j\in A_n(X), Y_j=i]} = \max_{\ell} \sum_{j=1}^{n} I_{[X_j\in A_n(X), Y_j=\ell]}.$$

If L^* denotes the (optimal) Bayesian error probability and

$$L_n = \mathbb{P}(g_n(X)\neq Y|D_n),$$

then our Theorem implies that for each $\varepsilon>0$

$$\mathbb{P}(L_n-L^*>\varepsilon) \leq \exp(-c\varepsilon^2 n) \qquad n\geq n_0.$$

The proof of this bound is based on the well known relation of $L_n - L^*$ and the L_1 errors of aposteriori probability estimates (see, for example Devroye(1982)).

PROOFS

For a sphere S put

$$I_{Sn} = \{i; A_{ni} \cap S \neq \emptyset\}$$

Lemma 1 If $|I_{Sn}|/n < (\varepsilon/K)^2/20$ then

$$\mathbb{P}(K \sum_{i \in I_{Sn}} |\mu_n(A_{ni}) - \mu(A_{ni})| > \varepsilon) \leq 3 \exp(-\tfrac{1}{25}(\varepsilon/K)^2 n) \tag{7}$$

and if $|I_{Sn}|/n < (\varepsilon/K)^3/2160$ then

$$\mathbb{P}(\sum_{i \in I_{Sn}} |\nu_n(A_{ni}) - \nu(A_{ni})| > \varepsilon) \leq 3 \exp(-\tfrac{1}{26}(\varepsilon/K)^2 n) \tag{8}$$

Proof (7) was proved in Devroye (1983, Lemma 3). (8) is a consequence of (7). To proving this, for a fixed integer N put

$$D_i = [i \tfrac{\varepsilon}{N}, (i+1)\tfrac{\varepsilon}{N}],$$

and

$$c_i = \frac{2i+1}{2} \frac{\varepsilon}{N}, \qquad (-\frac{K}{\varepsilon/N} \leq i \leq \frac{K}{\varepsilon/N}),$$

and

$$\tilde{Y}_j = \sum_i c_i \; I_{D_i}(Y_j)$$

then $\tilde{Y}_j$ is a quantized version of Y_j with step size ε/N, therefore

$$|\tilde{Y}_j - Y_j| \le \frac{\varepsilon}{2N}$$

and if $\tilde{\nu}_n$, $\tilde{\nu}$ stand for ν_n, ν, resp., when $\{Y_j\}$ is replaced by $\{\tilde{Y}_j\}$ then

$$\sum_i |\tilde{\nu}_n(A_{ni}) - \nu_n(A_{ni})| \le \frac{\varepsilon}{2N}$$

and

$$\sum_i |\tilde{\nu}(A_{ni}) - \nu(A_{ni})| < \frac{\varepsilon}{2N}.$$

Thus

$$\mathbb{P}(\sum_{i \in I_{Sn}} |\nu_n(A_{ni}) - \nu(A_{ni})| > \varepsilon) <$$

$$< \mathbb{P}(\sum_{i \in I_{Sn}} | \tilde{\nu}_n(A_{ni}) - \tilde{\nu}(A_{ni})| > \varepsilon - \frac{\varepsilon}{N}) =$$

$$= \mathbb{P}(\sum_{i \in I_{Sn}} | \sum_{i'} \frac{1}{n} \sum_{j=1}^{n} c_{i'} (I_{[X_j \varepsilon A_{ni}, Y_j \varepsilon D_{i'}]} - \mathbb{E}(I_{[X_j \varepsilon A_{ni}, Y_j \varepsilon D_{i'}]}))|$$

$$> \varepsilon (1 - \frac{1}{N})) <$$

$$< \mathbb{P}(K \sum_{i \in I_{Sn}} \sum_{i'} | \frac{1}{n} \sum_{j=1}^{n} I_{[X_j \in A_{ni}, Y_j \in D_{i'}]} - \mathbb{E}(I_{[X_1 \in A_{ni}, Y_1 \in D_{i'}]}) | >$$

$$> \varepsilon(1 - \frac{1}{N})) =$$

$$= \mathbb{P}(K \sum_{i \in I_{Sn}} \sum_{i'} | \tilde{\mu}_n(A_{ni}, D_{i'}) - \tilde{\mu}(A_{ni}, D_{i'}) | > \varepsilon(1 - \frac{1}{N}))$$

Applying (7) for N=52 and for the measure $\tilde{\mu}$ we get (8).

<u>Lemma 2</u> Under the condition (4), for each sphere S

$$\lim_{n \to \infty} \sum_{i \in I_{Sn}} \int_{A_{ni}} \left| \frac{\nu(A_{ni})}{\mu(A_{ni})} - m(x) \right| \mu(dx) = 0 \qquad (9)$$

<u>Proof</u>

$$\frac{\nu(A)}{\mu(A)} = \frac{\int_A m(z)\mu(dz)}{\mu(A)},$$

therefore if m is uniformly continuous then (4) implies (9).

For a given $\varepsilon>0$ and for an arbitrary m choose a uniformly continuous functionuous function $\tilde{m}$ such that

$$\int | \tilde{m}(z) - m(z) | \mu(dz) < \varepsilon$$

then

$$\sum_{i \in I_{Sn}} \int_{A_{ni}} \left| \frac{\nu(A_{ni})}{\mu(A_{ni})} - m(x) \right| \mu(dx) <$$

$$\leq \sum_{i\in I_{Sn}} \int_{A_{ni}} \left| \frac{\int_{A_{ni}} m(z)\mu(dz)}{\mu(A_{ni})} - \frac{\int_{A_{ni}} \tilde{m}(z)\mu(dz)}{\mu(A_{ni})} \right| \mu(dx) +$$

$$+ \sum_{i\in I_{Sn}} \int_{A_{ni}} \left| \frac{\int_{A_{ni}} \tilde{m}(z)\mu(dz)}{\mu(A_{ni})} - \tilde{m}(x) \right| \mu(dx) +$$

$$+ \sum_{i\in I_{Sn}} \int_{A_{ni}} | \tilde{m}(x) - m(x) | \mu(dx) \leq$$

$$\leq 2\varepsilon + \sum_{i\in I_{Sn}} \int_{A_{ni}} \left| \frac{\int_{A_{ni}} \tilde{m}(z)\mu(dz)}{\mu(A_{ni})} - \tilde{m}(x) \right| \mu(dx);$$

by the first part of the ,roof, Lemma 2 follows.

<u>Proof of the Theorem</u> For an arbitrary sphere S

$$\int |m_n(x) - m(x)| \mu(dx) \leq$$

$$\leq 2K \mu(S^c) + \sum_{\substack{i \\ A_{ni}\cap S\neq\emptyset \\ \mu(A_{ni})>0}} \int_{A_{ni}} |m_n(x) - m(x)| \mu(dx) \tag{10}$$

For $\mu(A_{ni}) > 0$ we have

$$\int_{A_{ni}} |m_n(x) - m(x)| \mu(dx) \leq$$

$$\leq \int_{A_{ni}} \left|m_n(x) - \frac{\nu_n(A_{ni})}{\mu(A_{ni})}\right| \mu(dx) + \left|\frac{\nu_n(A_{ni})}{\mu(A_{ni})} - \frac{\nu(A_{ni})}{\mu(A_{ni})}\right| \mu(A_{ni}) +$$

$$+ \int_{A_{ni}} \left| \frac{\nu(A_{ni})}{\mu(A_{ni})} - m(x) \right| \mu(dx) \tag{11}$$

If $\mu(A_{ni})>0$ and $\mu_n(A_{ni})>0$ then

$$\int_{A_{ni}} \left| m_n(x) - \frac{\nu_n(A_{ni})}{\mu(A_{ni})} \right| \mu(dx) =$$

$$= \left| \frac{\nu_n(A_{ni})}{\mu_n(A_{ni})} - \frac{\nu_n(A_{ni})}{\mu(A_{ni})} \right| \mu(A_{ni}) \leq$$

$$\leq K\mu_n(A_{ni}) \left| \frac{1}{\mu_n(A_{ni})} - \frac{1}{\mu(A_{ni})} \right| \mu(A_{ni}) =$$

$$= K \left| \mu_n(A_{ni}) - \mu(A_{ni}) \right|. \quad (12)$$

If $\mu(A_{ni})>0$ and $\mu_n(A_{ni}) = 0$ then $\nu_n(A_{ni}) = 0$ therefore

$$\int_{A_{ni}} \left| m_n(x) - \frac{\nu_n(A_{ni})}{\mu(A_{ni})} \right| \mu(dx) =$$

$$= \left| \frac{1}{n} \sum_{j=1}^{n} Y_j \right| \mu(A_{ni}) \leq K \left| \mu(A_{ni}) - \mu_n(A_{ni}) \right| \quad (13)$$

By (10), (11), (12) and (13) we have

$$\int |m_n(x) - m(x)| \mu(dx) \leq 2K \mu(S^c) +$$

$$+ K \sum_{i \in I_{Sn}} |\mu_n(A_{ni}) - \mu(A_{ni})| + \sum_{i \in I_{Sn}} |\nu_n(A_{ni}) - \nu(A_{ni})| +$$

$$+ \sum_{i \in I_{Sn}} \int \left| \frac{\nu(A_{ni})}{\mu(A_{ni})} - m(x) \right| \mu(dx)$$

By appropriate choice of S the first term can be arbitrary small, the fourth term tends to zero by Lemma 2. The exponential bounds of the second and third terms result from condition (4) and from Lemma 1.

REFERENCES

L. DEVROYE (1982) "Necessary and sufficient conditions for the pointwise convergence of nearest neighbor regression function estimates", Z. Wahrscheinlichkeitstheorie verw. Gebiete 61 pp. 467-481

L. DEVROYE (1983) "The equivalence of weak, strong and complete convergence in L_1 for kernel density estimates", Ann. Statist. 11 pp. 896-904

F.H. FRIEDMAN (1977) "A recursive partitionong decision rule for nonparametric classification", IEEE Trans. Computers C-26 pp. 404-408

L. GORDON, R.A. OLSHEN (1980) "Consistent nonparametric regression from recursive partitioning schemes" J. Multivariate Analysis 10 pp. 611-627.

C. J. STONE (1977) "Consistent nonparametric regression" Ann. Statis. 5 pp. 595-645

Prof. Luc Devroye
McGill University
School of Computer Sciences
805 Sherbrooke Street West
Montreal, PQ
Canada H3A 2K6

Dr. László Györfi
Technical University of Budapest
Stoczek u. 2.
H-1111 Budapest