Asymptotics for a Multivariate Location Estimator

Deborah Nolan
University of California at Berkeley

Technical Report No. 149
March 1988

Department of Statistics
University of California
Berkeley, California

Asymptotics for a Multivariate Location Estimator

Deborah Nolan

University of California at Berkeley

ABSTRACT

Asymptotics for a location estimator are obtained in this paper. The location estimator examined is that point in $\mathbf{R}^2$ with greatest depth, where the depth of a point is defined to be the minimum empirical measure of the half planes that contain the point. For a spherically symmetric distribution this estimator is consistent and it behaves, in the limit, as the best fit to a Gaussian process. That is, it behaves as the maximizer of a functional of the Gaussian process.

Introduction

In this paper, we determine the asymptotic behavior of a multivariate location estimator via an empirical process treatment. The location estimator is a set statistic; it can be represented as a functional of an empirical process indexed by the collection of half spaces in $\mathbf{R}^2$. Such a representation is useful because the collection of half spaces meets the metric entropy condition imposed by empirical process methodology to obtain central limit theorems. Other set statistics can not be so represented, and great difficulty may be encountered in achieving limit results. The estimator is shown to behave as the maximizer of a functional of a Gaussian process on the surface of the unit sphere in $\mathbf{R}^2$.

Donoho (1982) suggested the set statistic as an extension of Tukey's (1977, page 30) notion of depth from the real line to the plane. According to Tukey, given a sample of size n on the line, the depth of a point t is either the number of observations to the left of t or the number of observations to the right of t, whichever is smaller. Normalize with respect to n; then t's depth is either the empirical measure of the half closed interval $(-\infty,t]$ or that of the interval $[t,\infty)$, again, whichever is smaller. This normalization provides a general definition of depth for probability measures other than the empirical and for dimensions greater than one: the depth of a point t in $\mathbf{R}^d$, with respect to a probability measure P on $\mathbf{R}^d$, is the smallest P-measure of all closed halfspaces that contain t. It follows that a natural multivariate extension of the sample median to the plane is that t in $\mathbf{R}^2$ with greatest depth with respect to the empirical measure constructed from a sample of size n on the plane. This multivariate location estimator was first studied by Donoho (1982) for its affine equivariance and excellent robustness properties. Also, there is a projection pursuit flavor to this estimator; the assignment of depth to a point t is equivalent to pursuing the one-dimensional projection that has the largest value for its one-dimensional distribution function at t's projection.

To formalize the definition, let $\xi_1, \cdots, \xi_n$ be a sample from a distribution P on $\mathbf{R}^2$. In this paper, we consider only distributions with densities that are spherically symmetric about 0; that is, the density $p:\mathbf{R}^2 \to \mathbf{R}$ must have the property that $p(x)=p(y)$, if $|x|=|y|$. The symbol $|\cdot|$ denotes the Euclidean norm. Let $H_{u,t}$ be the closed halfspace with boundary that passes through t in the direction normal to the unit vector u. Call the collection of these half spaces $\mathbf{H}$. Then the location estimator T_n maximizes:

$$(1) \quad \inf_u P_n H_{u,t}$$

where P_n stands for the empirical measure that places mass $1/n$ on each of $\xi_1, \cdots, \xi_n$. Nonuniqueness can be handled by any arbitrary, fixed rule. We assume such a rule is employed in the definition of T_n, but ignore the slight complication that would result in its representation.

For a sample in three dimensions, T_n is also defined as in (1), except $H_{u,t}$ represents, instead, the closed half space in $\mathbf{R}^3$ through $t \in \mathbf{R}^3$ in the direction normal to the unit vector $u \in \mathbf{R}^3$, and P_n is the appropriate empirical measure. In this case, we have been unable to overcome difficulties involving the limiting distribution, but the author conjectures that it is the three-dimensional analogue of the result stated below.

Main Result

(2) Theorem: Suppose T_n is defined as in (1) for a sample of size n from a distribution P on $\mathbf{R}^2$. Also, suppose P is spherically symmetric about 0 and its one-dimensional marginal density has value $p_0>0$ at 0. Then

$$\operatorname{argsup}_t \inf_u P_n H_{u,t} \to \operatorname{argsup}_t \inf_u [Z(u) - u'tp_0]$$

where Z is a Gaussian process on the unit sphere with zero means and covariance structure:

$$(3) \quad PH_{u,0}H_{v,0} - 1/4 = 1/4 - \arccos(u'v)/2\pi .$$

The symbol "–>" denotes convergence in distribution. Notice the periodicity of the limit process. If $u'v = -1$ then the correlation is -1, and if $u'v = 0$ then the correlation is 0.

We explain the form the limit takes before we prove the theorem. Rewrite P_n as the sum of a deterministic expectation and a stochastic contribution:

$$P_n H_{u,t} = PH_{u,t} + \frac{1}{\sqrt{n}} \nu_n H_{u,t} \ .$$

Here ν_n is $\sqrt{n}\,(P_n - P)$. Follow the example of the one-dimensional sample median, and proceed with a one-term Taylor expansion of $PH_{u,t}$, for t in a neighborhood of 0.

$$(4) \qquad P_n H_{u,t} = \tfrac{1}{2} - u't\, p_0 + o(u't) + \frac{1}{\sqrt{n}} \nu_n H_{u,t} \ .$$

Spherical symmetry allows ½ to replace $PH_{u,0}$ and forces the one-dimensional marginals of P to be the same. The form of the limit as stated in the theorem is now revealed.

Proof of (2): The proof has three main steps. The first establishes consistency of T_n via an adaption of the Glivenko-Cantelli lemma. Next, an argument that sharpens the Glivenko-Cantelli result leads to $\sqrt{n}$-consistency of T_n. The argument uses a functional limit theorem for the empirical process and the Taylor expansion in (4). Finally, we show $argsup_t \, inf_u [x(H_{u,0}) - p_0 u't]$ is an almost surely continuous function of the real-valued functionals $\{x\}$ on H; the continuous mapping theorem gives the asymptotic behavior of $\sqrt{n}\,T_n$.

Consistency of T_n is a consequence of the Glivenko-Cantelli lemma for half planes in $\mathbf{R}^2$ which says that the empirical measure of a half plane converges to its expected measure uniformly over H the collection of all half planes. For $\varepsilon > 0$, let $c_\varepsilon = \tfrac{1}{2}(\inf_u PH_{u,0} + \sup_{|t|>\varepsilon} \inf_u PH_{u,t})$. Then

$$\begin{aligned} &\mathbf{P}\{\ |T_n| < \varepsilon\ \} \\ &\geq \mathbf{P}\{\inf_u P_n H_{u,0} > \sup_{|t|>\varepsilon} \inf_u P_n H_{u,t}\ \} \\ &\geq \mathbf{P}\{\inf_u P_n H_{u,0} > c_\varepsilon > \sup_{|t|>\varepsilon} \inf_u P_n H_{u,t}\} \end{aligned}$$

$$\geq \mathbf{P}\{ \sup_{\mathbf{H}} |(P_n - P)H_{u,t}| > \tfrac{1}{2} - c_\varepsilon \}$$

The last lower bound converges to 1 by the Glivenko-Cantelli lemma for **H**.

Next, strengthen this result to: $T_n = O_p(1/\sqrt{n})$. Express $P_n H_{u,T_n}$ as in (4),

$$\tfrac{1}{2} - u'T_n p_0 + o(u'T_n) + \frac{1}{\sqrt{n}} \nu_n H_{u,T_n} .$$

Empirical process methodology provides the limit theory for the stochastic process ν_n, a random element of the space **X** of real functions on **H**. Equip **X** with the supremum norm. The collection of half planes is a Vapnik-Cervonenkis class. It easily meets the regularity conditions for convergence of ν_n to a Gaussian process with uniformly continuous sample paths in **X** (see Theorem VII.21, Pollard 1984); uniform continuity is with respect to the $L^2(P)$ semi-norm on **H**. In addition, ν_n is stochastically equicontinuous (Lemma VII.15, Pollard 1984). Stochastic equicontinuity of ν_n says that if the random sequence $\{t_n\}$ converges to t in probability then

$$\nu_n(H_{u,t_n} - H_{u,t}) \to 0 \qquad \textit{in probability.}$$

In fact, the convergence is uniform in u.

Weak convergence of ν_n and the continuous mapping theorem imply $\sup_{u,t} |\nu_n H_{u,t}|$ is $O_p(1)$. Thus

$$\begin{aligned} \inf_u P_n H_{u,T_n} &= \tfrac{1}{2} + \inf_u [-u'T_n p_0 + o(u'T_n) + \frac{1}{\sqrt{n}} \nu_n H_{u,T_n}] \\ &\leq \tfrac{1}{2} + \inf_u [-u'T_n p_0 + o(u'T_n)] + \sup_{u,t} [\frac{1}{\sqrt{n}} \nu_n H_{u,t}] \\ &= \tfrac{1}{2} + \inf_u [-u'T_n p_0 + o(u'T_n)] + O_p(\frac{1}{\sqrt{n}}) \end{aligned}$$

Compare $\inf_u P_n H_{u,T_n}$ with $\inf_u P_n H_{u,0}$. The defining property of T_n requires $\inf_u P_n H_{u,T_n}$ to be greater than $\inf_u P_n H_{u,0}$, and the later is $O_p(\frac{1}{\sqrt{n}})$ because:

$$\inf_u P_n H_{u,0} = \tfrac{1}{2} + \inf_u \frac{1}{\sqrt{n}} \nu_n H_{u,0} .$$

It follows from the upper bound on $\inf_u P_n H_{u,T_n}$ that T_n must be $O_p(\frac{1}{\sqrt{n}})$.

Prepare $\sqrt{n}T_n$ for an application of the continuous mapping theorem. The estimator is stochastically bounded: given $\varepsilon>0$, there is an $M(\varepsilon)$ for which $\sqrt{n}\,|T_n| < M(\varepsilon)$ with probability at least $1-\varepsilon$, eventually.

$$\sqrt{n}T_n = \sqrt{n}\, \underset{|t|}{argsup}\ \underset{u}{inf}\,[\tfrac{1}{2} - p_0 u't + o(u't) + \frac{1}{\sqrt{n}}\nu_n H_{u,t}]$$

$$= \sqrt{n} \underset{|t|\le M(\varepsilon)/\sqrt{n}}{argsup}\ \underset{u}{inf}\,[\tfrac{1}{2} - p_0 u't + o(u't) + \frac{1}{\sqrt{n}}\nu_n H_{u,0} + o_p(\frac{1}{\sqrt{n}})]$$

$$= \underset{|t|\le M(\varepsilon)}{argsup}\ \underset{u}{inf}\,[-p_0 u't + o(1) + \nu_n H_{u,0} + o_p(1)]$$

Stochastic equicontinuity implies the o_p term is uniform in u. The process ν_n converges in distribution to a Gaussian process Z on **H** with zero means and covariance structure as in (3). The addition of the o terms to ν_n does not change its asymptotic distribution.

Finally, check continuity of the function $g(x)=argsup_t\, inf_u [x(H_{u,0})-p_0 u't]$ at those functionals x in **X** on which the limit process Z concentrates. The function g is continuous provided $inf_u [x(H_{u,0})-p_0 u't]$ has a unique maximum in t. The behavior of $\sqrt{n}T_n$ is determined by Z on those half planes that pass through the origin. There is no harm in the simplification that treats Z as a process indexed by u on the unit sphere, rather than by $H_{u,0}$ in **H**.

Consider one particular sample path of Z; keep the dependence on ω suppressed. Without loss of generality, set p_0 to 1. Define C as the constant: $sup_t\, inf_u [Z(u)-u't]$; define t^* as some t for which: $C = inf_u [Z(u)-u't^*]$. We show that there is only one such t^*, for almost all ω.

Call those u for which $Z(u)-u't^* = C$ the zeroes of t^*. Then

(5) *The zeroes of t^* are not contained in any open half sphere of the unit sphere.*

To prove this assertion, suppose the contrary that all the zeroes belong to some open half sphere called $S_{1/2}$. Let $s=t^* - \tfrac{1}{2}\varepsilon u_{1/2}$ where $u_{1/2}$ is the unit vector with direction opposite the direction of the center of $S_{1/4}$. The ε is chosen so that $Z(u)-u't^* \geq C+\varepsilon$ for u in $S_{1/4}^c$, the closed half sphere complement to $S_{1/4}$. Such a positive ε exists because $Z(u)-u't^* > C$ on $S_{1/4}^c$. Then, for all u:

$$Z(u) - u's \geq C + \tfrac{1}{2}\varepsilon .$$

This contradicts the definition of C. The zeroes of t^* satisfy (5).

Implicit in statement (5) is that t^* must have at least two zeroes. Suppose there are only two zeroes; call them u_1, u_2. Then $u_1 = -u_2$, and an unnatural constraint is placed on the sample path of Z, i.e. sample paths with exactly two zeroes live on a set of probability 0. From the negative correlation of $Z(u)$ and $Z(-u)$ we arrive at:

$$Z(-u_1) = C - u_1't^*$$
$$Z(-u_1) = -C - u_1't^* ,$$

which imply $C=0$. In turn, this implies $Z(v_1)+Z(v_2)+Z(v_3) = 0$ for $v_1=(0,1)$, $v_2=(\sqrt{3}/2,-1/2)$, and $v_3=(-\sqrt{3}/2,-1/2)$. Choice of v_1, v_2, v_3 does not depend on ω. Certainly, this relationship occurs with probability 0.

Now t^* has at least three zeroes and these zeroes do not all belong to some open half sphere. In this case, multiple maxima are impossible, almost surely. If s^* locates another maximum then:

$$Z(u)-u'\tfrac{1}{2}(s^*+t^*) \geq Z(u)-max(u's^*,u't^*) \geq C .$$

The first inequality is strict unless $u's^*=u't^*$. The second inequality is strict unless u is a zero for s^* or t^*. We must have equality between the left and righthand terms for at least three unit vectors that satisfy (5). (Equality at only two unit vectors can be ignored, as shown earlier). So, s^* and t^* share at least three zeroes. These shared zeroes can not all belong to some open half sphere. This forces s^* to equal t^*, a contradiction.

Uniqueness of the suprema of the functional $inf_u[x(u)-u't]$ is established. In turn, continuity of the functional $argsup_t inf_u[x(u)-u't]$ is established. The asymptotic behavior of $\sqrt{n}T_n$ is that of the $argsup_t inf_u [Z(u) - u't]$.

Remarks

The unusual form of the limit process results from the response of the function $argsup_t inf_u \, x(u,t)$ to rescaling and from the nondifferentiability of the indicator function $H_{u,0}$. A more typical maximization problem takes a two term Taylor expansion of the expected value $PH_{u,t}$ about $t=0$, and balances it against the linear contribution made by ν_n evaluated at a one-term Taylor expansion of $H_{u,t}$. Here, the linear term in the expansion of $PH_{u,t}$ does not disappear because T_n maximizes $inf_u P_n H_{u,t}$ rather than $P_n H_{u,t}$. Also, the indicator function for the halfplane $H_{u,t}$ is not differentiable in t, so matching quadratic terms in t with linear terms in $t\nu_n$ in order to produce the more familiar limit result is not possible.

Acknowledgements

The author thanks D. Donoho and M. Gasko for their generosity in introducing her to the multivariate location estimator. The author also thanks D. Pollard for helpful comments and criticisms.

References

Donoho, D.L. (1982), "Breakdown properties of multivariate location estimators". Ph.D. Qualifying paper, Dept of Statistics, Harvard University.

Pollard, D. (1984), *Convergence of Stochastic Processes*, Springer-Verlag, New York.

Tukey, J.W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, Massachusets.

TECHNICAL REPORTS

Statistics Department

University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist. 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (November 1986). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, 23, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in *Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.*

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (August 1987). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \leq k \leq n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Achieving information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (October 1987, revised March 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. To appear in *Statistics a Guide to the Unknown.*

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. DUBINS, L.E. and SCHWARZ, G. (December 1987). A sharp inequality for martingales and stopping-times.

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (January 1988). Two reports on trend analysis: a) An Elementary Trend Analysis of Rio Negro Levels at Manaus, 1903-1985 b) Consistent Detection of a Monotonic Trend Superposed on a Stationary Time Series

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (February 1988). Randomization in the two-armed bandit problem.

144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.

145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.

146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.

147. STARK, P.B. (March 1988). Strict bounds and applications.

148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.

149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.

150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.

151. NOLAN, D. (March 1988). Limit theorems for a random convex set.

152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.

153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.

154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.

155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.

156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.

157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.

158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.

159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.

160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.

161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).

162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.

163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform.

164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.

165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.

166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.

167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.

168. LE CAM, L. (August 1988). Maximum likelihood an introduction.

169. BREIMAN, L. (August 1988). Submodel selection and evaluation in regression-The conditional case and little bootstrap.

170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.

171. STONE, C.J. (September 1988). Large-sample inference for logspline models.

172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

173. MILLAR, P.W. (October 1988). Optimal estimation in the non-parametric multiplicative intensity model.

174. YOR, M. (October 1988). Interwinings of Bessel processes.

175. ROJO, J. (October 1988). On the concept of tail-heaviness.

176. ABRAHAMS, D.M. and RIZZARDI, F. (September 1988). BLSS - The Berkeley interactive statistical system: An overview.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.