Gamma-funnels in the domain of a probability, with statistical implications.

P.W. Millar[1]
Statistics Department
University of California
Berkeley, California 94720
U.S.A.

Technical Report No. 177
October 1988

[1]Research supported by National Science Foundation Grant DMS87-01426

Department of Statistics
University of California
Berkeley, California

Gamma-funnels in the domain of a probability, with statistical implications.

P.W. Millar[1]
Statistics Department
University of California
Berkeley, California 94720
U.S.A.

[1]Research supported by National Science Foundation Grant DMS87-01426

1. Introduction

This paper establishes a probabilistic result which has implications for the numerical implementation of certain non-parametric statistical procedures. To describe the probabilistic result, let λ_n, $n \geq 1$ be an increasing sequence of integers, totally arbitrary except for the condition that $\lambda_n \uparrow \infty$. Let $Y_1, Y_2, \ldots$ be i.i.d. random variables with values in an infinite dimensional Banach space B, and whose common distribution μ has as its support the unit ball of B. Then no matter what $\varepsilon > 0$ is selected, and no matter which θ_0 in the unit ball of B is chosen, and no matter what the rate at which $\lambda_n \uparrow \infty$, one clearly has

(1.1) $$\lim_{n\to\infty} P\{|Y_i - \theta_0| \leq \varepsilon \text{ for some } i \leq \lambda_n\} = 1.$$

Indeed, if $\delta_\varepsilon \equiv P\{|Y_1 - \theta_0| < \varepsilon\}$, then the probability in (1.1) is equal to $1 - [1 - \delta_\varepsilon]^{\lambda_n}$, which converges to 1 as $\lambda_n \uparrow \infty$, since $\delta_\varepsilon > 0$.

On the other hand, we show below that if ε in (1.1) is allowed to shrink with n, as $n \to \infty$, a result such as (1.1) cannot possibly hold, no matter how slowly ε shrinks with n. More precisely, pick $\gamma > 0$. Then no matter how small γ is, and *no matter how fast* $\lambda_n \uparrow \infty$, there will exist an infinite number of points θ_0 within the unit ball such that

(1.2) $$\lim_{n\to\infty} P\{|Y_i - \theta_0| \leq n^{-\gamma} \text{ for some } i \leq \lambda_n\} = 0.$$

The actual result proved below is a stronger almost sure version of (1.2). Statement (1.2) and its stronger version (Theorem 1, section 3) should be compared with (1.1) and its stronger version: for almost every w, the sequence $Y_1(w), Y_2(w), \ldots$ is dense in the unit ball of B.

There can be no uniform distribution on the unit ball of an infinite dimensional Banach space; the main result therefore quantifies how this fact is reflected in an i.i.d. sequence. The n-independent neighborhoods of θ_0 involved in (1.2) are given a formal description as "γ-funnels", in section 2, where they have a natural statistical interpretation.

If B is assumed *finite* dimensional, then results like (1.2) are often false, depending on μ. See Remark (3.1b) for elaboration.

The result (1.2) was motivated by considerations involving stochastic search for extrema as a means of calculating, approximately, non-parametric goodness of fit tests and maximum likelihood estimators for an infinite dimensional parameter. Such motivations are discussed in section 2; the main theorem is stated and proved in section 3. These two sections can be read independently of each other; section 2 is statistical, while section 3 is totally probabilistic.

2. Gamma funnels and some statistical motivations

Probabilities on infinite dimensional Banach spaces B arise as natural tools for numerical implementation of several non-parametric statistical procedures. An important example is the computation of extrema by Monte Carlo methods. For example, let Θ be a parameter set (in the usual sense of statistics) which is a subset of a Banach space B. For n = 1,2,... and for each $\theta \in \Theta$ let $\hat{D}_n(\theta)$ be a real random variable. Nonparametric goodness of fit statistics, minimum distance estimators, and maximum likelihood estimates involve computation of quantities like

$$\inf_{\theta \in \Theta} \hat{D}_n(\theta) \tag{2.1}$$

Example 2.1. Let $x_1, \ldots, x_n$ be iid random variables with values in the Euclidean space R^d. Let $\hat{P}_n$ be the empirical measure associated with this sample. Let Θ be the collection of all elliptically symmetric distributions on R^d. Then Θ can be identified as a subset of the unit ball of some Banach space in many ways. For example, if **H** is the collection of half-spaces on R^d, and if $\theta \in \Theta$, one may regard θ as an element of $L_\infty(\mathbf{H})$ by identifying θ with the mapping $H \to \theta(H)$, $H \in \mathbf{H}$. A plausible goodness of fit statistic for assessing the hypothesis that the data came from some elliptically symmetric distribution is

$$\inf_{\theta \in \Theta} \sup_{H \in \mathbf{H}} n^{1/2} |\hat{P}_n(H) - \theta(H)|. \tag{2.2}$$

This is obviously of the form (2.1), and the set Θ is clearly infinite dimensional. See Beran and Millar, 1988, for more information in this particular problem.

Actual computation of $\inf_{\Theta}$ in (2.1) in non-parametric situations, is usually impossible, as a glance at Example 2.1 makes clear. Standard methods of numerical analysis involving "derivatives" will fail, because $\theta \to \hat{D}_n(\theta)$ is often not even differentiable; in any case Θ is infinite dimensional so finding paths of "quickest descent" involves looking in an infinite number of directions (a hard thing to do!). Asymptotic methods, wherein one computes $\lim_{n \to \infty} \inf_{\theta \in \Theta} \hat{D}_n(\theta)$ instead of (2.1), fail because of the intractability of the limit distribution, and because Θ is still infinite dimensional. Since, in a number of applications, Θ is not compact, one cannot attempt to evaluate (2.1) by taking an ε-grid over Θ.

Given this difficult situation, one is tempted to try a simple Monte Carlo technique. To describe this, let μ be a probability on $\Theta \subset B$, and let $Y_1, \ldots, Y_{\lambda_n}$ be a sequence of i.i.d. Θ-valued random variables, with distribution μ; here, as in section 1, $\{\lambda_n\}$ is a sequence of integers increasing to infinity. One then replaces the computationally infeasible (2.1) by

$$\min_{1 \le i \le \lambda_n} \hat{D}_n(Y_i). \tag{2.3}$$

This is a possible improvement since the minimum over a finite set replaces an infimum over an infinite dimensional one. If μ has full support on Θ, and if θ_0 is the actual minimizing point for (2.3), then the sequence $Y_1, \ldots, Y_{\lambda_n}$ will eventually come within ε of θ_0, for any $\varepsilon > 0$. (cf. section 1, display (1.1).) Assuming reasonable continuity of $\theta \to \hat{D}_n(\theta)$, one could then surmise that, by taking $\lambda_n \uparrow \infty$ at a fast rate, then (2.3) would be an effective substitute for (2.1).

Unfortunately, this sanguine view ignores the facts that (a) $\hat{D}_n$ changes with n and (b) typically the size of the set about θ_0 which determines the infimum gets rapidly smaller with n.

To describe this phenomenon more precisely, define for $c > 0$, $\theta_0 \in B$, $n = 1,2,\ldots$

$$V_n(c) \equiv V_n(c, \theta_0) = \{y \in B : |y - \theta_0| \le cn^{-\gamma}\} \tag{2.4}$$

where γ is a fixed positive number (in most applications, $0 < \gamma \le 1/2$, with $\gamma = 1/2$ being the most common). The collection $\{V_n(c, \theta_0) : n \ge 1\}$ is called a Γ-funnel at θ_0 of width c. For many applications, the infimum of $\hat{D}_n(\theta)$ is achieved within an γ-funnel in the sense that for some fixed γ; and unknown θ_0:

(2.5) (i) if $\theta_n \notin V_n(c, \theta_0)$ $\forall$ large n and every c then $\hat{D}_n(\theta_n) \to +\infty$

(ii) if $\theta_n \in V_n(c, \theta_0)$ $\forall$ large n and c fixed then $\hat{D}_n(\theta_n)$ remains bounded.

Thus, in order that the search set $Y_1, \ldots, Y_{\lambda_n}$ be effective for all large n it must not only hit the set $\{y : |y - \theta_0| < \varepsilon\}$ repeatedly, (as mentioned above) but it must visit the γ-funnel repeatedly: at least one $Y_1, \ldots, Y_{\lambda_n}$ should land in $V_n(c, \theta_0)$ with probability 1, as $n \to \infty$. That is, an effective search $Y_1, \ldots, Y_{j_n}$ needs to satisfy

$$\underline{\lim}\, P\{|Y_i - \theta_0| \le cn^{-\gamma} \text{ for at least one } i \le \lambda_n\} = 1. \tag{2.6}$$

The main result of this paper asserts that one cannot expect results like (2.6), no matter how fast one lets $\lambda_n \uparrow \infty$, and no matter how you choose c or γ. This result thus gives a negative view on certain simple Monte Carlo techniques.

On the other hand, in certain special problems, Monte Carlo searches of an infinite dimensional set Θ can be effective. In such cases, one again replaces (2.1) by an expression of the form (2.3), but the random sequence $Y_1, Y_2, \ldots$ is *not* i.i.d. Two examples of such a method are Beran and Millar 1988a, 1988b; in these papers the

Y_i's are constructed by a bootstrap method. (cf. Efron, 1979). In other special infinite dimensional problems, grid-type searches (called "sieve methods") are possible; see, e.g., Geman and Huang 1982. General background on Monte Carlo techniques can be found in Rubenstein (1981).

3. Covering a set by a random sample

Let B be a separable Banach space, and let μ be a probability on the Borel sets of B. Let $Y_1, \ldots, Y_{\lambda_n}$ be i.i.d. B-valued random variables, with common distribution μ. Here λ_n is an arbitrary sequence of positive integers subject only to the condition $\lambda_1 < \lambda_2 < \cdots$, $\lambda_n \uparrow \infty$. Fix $\gamma > 0$. The following theorem in the main result of this paper.

Theorem 1. *There exists a countably infinite collection of points* $\{\theta_{oj}, j \geq 1\}$, $|\theta_{oj}| \leq 1$, $\theta_{oj} \in B$, *such that*

$$\underline{\lim}_{n} \min_{j} \min_{i \leq \lambda_n} |Y_i - \theta_{oj}| n^{\gamma} \geq 1 \quad \text{a.e. } (\mu).$$

Proof. Let $z \in B$, $r > 0$ and define the ball of radius r centred at z by

$$S(z,r) = \{y \in B : |y - z| < r\} \tag{2.1}$$

Lemma 3.1. *There exists an infinite collection of distinct points* $z_1, z_2, \ldots,$ *and balls* $\{S(z_i, 1/4), i \geq 1\}$ *such that* $S(z_i, 1/4) \subset S(0,1)$ *and the* $S(z_i, 1/4)$ *are all disjoint.*

This lemma is immediate if B is a Hilbert space; see Kuo, 1975, p.5. It may also be known in the generality given here, but I could not find a reference. In any case, a simple proof is provided at the end of this section.

Corollary 3.1. Let m be a positive integer, and z an arbitrary point of B. Then within $S(z, 2^{-m})$ are a countably infinite number of disjoint balls of radius 2^{-m-2}.

The proof of theorem 1 will now be completed several steps. First let

$$a_1 > a_2 > \cdots \tag{3.1}$$

be a sequence of real numbers such that $\Sigma a_i < \infty$. Define the number c and the functions $\xi(x)$, $\psi(x)$, $x \geq 0$, by $c = \gamma / \log 4$

$$\xi(x) = \exp\{-\exp(c^{-1}(x+2))\}$$

$$\psi(x) = \exp\{c^{-1}(x+2)\}$$

Step 1: construction of θ_{01}.

By lemma 3.1, and the assumption that μ is a probability there exist z_i, $|z_i| \leq 1$ such that

$$\sum_i \mu(S(z_i, 2^{-2})) \leq 1;$$

in fact the z_i can be chosen so that $|z_i| = 1/2$. Pick $z_1^* \in \{z_i\}$ such that

$$\mu(S(z_1^*, 2^{-z})) \leq a_1 \xi(1)/\lambda(\psi(1))$$

where we have written $\lambda_n \equiv \lambda(n)$ for typographical convenience.

Within the ball $S(z_1^*, 2^{-2})$ there are, by Corollary 3.1, disjoint balls $S(z_{2i}, 2^{-4})$, centred at a countable number of points $\{z_{2i}\}$ (satisfying $|z_{2i} - z_1^*| = 2^{-3}$.) Therefore, again $\Sigma\mu(S(z_{2i}, 2^{-4})) \leq 1$. Pick $z_2^* \in \{z_{2i}, i \geq 1\}$ such that

$$\mu(S(z_2^*, 2^{-4})) \leq a_1 \xi(2)/\lambda(\psi(2)).$$

Within $S(z_2^*, 2^{-4})$ one may again pick disjoint balls of radius 2^{-6}, leading as above to a z_3^* inside $S(z_2^*, 2^{-4})$ with

$$\mu(S(z_3^*, 2^{-6})) \leq a_1 \xi(3)/\lambda(\psi(3)).$$

Continuing through i steps of this instruction we obtain a sequence z_i^*, $i \geq 1$, such that

(3.2) $$\mu(S(z_i^*, 2^{-2i})) \leq a_1 \xi(i)/\lambda(\psi(i))$$

(3.3) $$|z_{i+1}^* - z_i^*| \leq 2 \cdot 2^{-i},$$

implying that, for $m > n$

(3.4) $$|z_m^* - z_n^*| \leq 2 \sum_{i=n}^{m} 2^{-2i}.$$

Thus we may define θ_{01} by

(3.5) $$\theta_{01} = \lim_{n \to \infty} z_n^*$$

and by (3.4) we see that

(3.6) $$|\theta_{01} - z_n^*| \leq 2^{-2(n-1)}.$$

Step 2: the chance of hitting a γ-funnel at θ_{01}.

The calculations of step 1 imply that

(3.7) $$\begin{aligned} P\{|Y_1 - \theta_{01}| < 2^{-2(i+2)}\} &= \mu\{S(\theta_{01}; 2^{-2(i+2)})\} \\ &\leq \mu\{S(z_i^*; 2^{-2i})\} \\ &\leq a_1 \xi(i)/\lambda(\psi(i)). \end{aligned}$$

Let $i_n = c \log n - 2$. Then, from (3.7),

$$\begin{aligned} P\{|Y_i - \theta_{01}| &\le n^{-\gamma} \text{ for some } i \le \lambda_n\} \\ &\le \lambda_n P\{|Y_1 - \theta_{01}| \le n^{-\gamma}\} \\ &= \lambda_n P\{|Y_1 - \theta_{01}| \le 2^{-2(i_n+2)}\} \\ &\le \lambda_n a_1 \xi(i_n)/\lambda(\psi(i_n)) \\ &= a_1 e^{-n}. \end{aligned}$$

Step 3: construction for $\theta_{02}, \theta_{03}, \ldots$

Next, working with a_2 in steps 1,2 instead of a_1, pick a $y_1^* \in S(0,1)$, $y_1^* \in \{z_i\}$, $y_1^* \ne z_1^*$ and such that

$$\mu(S(y_1^*, 2^{-2})) \le a_2 \xi(1)/\lambda(\psi(1)).$$

Continuing in the manner used to produce θ_{02}, produce θ_{02} with

$$P\{|Y_i - \theta_{02}| \le n^{-\gamma} \text{ for some } i \le \lambda_n\} \le a_2 e^{-n}.$$

Similarly produce θ_{0k} where

$$P\{|Y_i - \theta_{0k}| \le n^{-\gamma} \text{ for some } i \le \lambda_n\} \le a_K e^{-n}. \tag{3.8}$$

Let A_{nk} be the event that $\min_{i \le \lambda_n} |Y_i - \theta_{0k}| n^{\gamma} \le 1$. By (3.8) and the Borel Cantelli lemma, the probability that A_{nk} occurs for infinitely many n, k is 0. This completes the proof.

Proof of lemma 3.1. Let B_1 be a finite dimensional subspace of B, having dimension d. Let us first show that $B_1 \cap S(0,1)$ contains d disjoint balls of radius 1/4 centred at some $z_1, \ldots, z_d$, with $|z_i| = 1/2$. Let $b_1, \ldots, b_d$ be a linearly independent subset from B_1, and let $e_1, \ldots, e_d$ be the usual basis for R^d. The mapping of B_1 to R^d given by $b_i \to e_i$ gives an isometric isomorphism of B^1 and R^d, when the norm on R^d is the one inherited from B_1. Thus to establish the first claim it suffices to work on R^d, with arbitrary norm $|\cdot|$ there. If x, $h \in (R^d)^*$, $x = \Sigma x_i e_i$, $h = \Sigma h_i e_i$, write $<x, h>$ for $\Sigma h_i x_i$. Let $|\ |_*$ denote the norm on the dual space. Then for $x \in R^d$, $|x| = \sup_{|h|_*=1} |<x,h>|$, $h = \Sigma h_i e_i \in (R^d)^*$. Let x have the form $x = \sum_{i=1}^{d-1} a_i e_i$, $y = a_d e_d$. Then

$$|x - y| = \sup_{|h|_*=1} |<x - y, h>|$$

$$\geq \sup_{\substack{h:|h|_*=1 \\ h_d=0}} |\langle x-y, h\rangle| \vee \sup_{\substack{h:|h|_*=1 \\ h_1=\cdots=h_{d-1}=0}} |\langle x-y, h\rangle|$$

$$= |x| \vee |y|.$$

Take x, y further restricted to $|x| = |y| = 1/2$, so that for such x, y

$$|x - y| \geq 1/2.$$

Let $D^{d-1} = \{x \in B^1 \doteq R^d : x = \sum_1^{d-1} a_i e_i, |x| = 1/2\}$. Then by the above inequality the distance between $1/2\, e_d$ and D^{d-1} is at least 1/2. Thus the ball $S(1/2\, e_d, 1/4)$ is contained in the unit ball of $B_1 \cong R^d$, and its distance from D^{d-1} is at least 1/4. In particular, it will not intersect a ball of radius 1/4 centred at any $x \in D^{d-1}$. An inductive argument completes the proof of our first claim.

As a consequence, we see that $S(0, 1)$ contains, for any d, at least d disjoint balls $\{S(z_i, 1/4): 1 \leq i \leq d\}$ where $|z_i| = 1/2$. Let P denote the set whose points are given by $\bigcup_{i=1}^{K} S(z_i, 1/4)$, where $S(z_i, 1/4)$ $1 \leq i \leq K$ are disjoint balls centred at z_i with $|z_i| = 1/2$; the z_i and K are allowed to vary here. Then P is partially ordered (by set inclusion) and each chain in P has an upper bound (given by the set theoretic union of all the sets in the chain). By Zorn's lemma, there is a maximal element, and this completes the proof of lemma 2.1, since said maximal element would necessarily have to involve more than d disjoint balls, for any finite d.

Remarks 3.1. (a) The hypothesis that B be separable was adopted only for convenience of exposition If B were not separable, and if μ were a *Borel* measure, then its support must be a separable subspace B_1 of B. Hence the arguments above can be carried through on B_1 instead of B, establishing theorem 3.1 in this context. If μ were not Borel measurable on B, but measurable with respect to the sigma field generated by, say, the open balls of B, as specified in some of the recent literature on empirical processes (cf., Gaenssler, 1983), then the main theorem continues to hold, by exactly the same argument.

(b) If B is assumed finite dimensional, then results like that of theorem 3.1 are false, in general, depending on μ. For example, let μ be a probability supported by the unit ball S^d of the Euclidean space R^d, such that μ has a continuous density bounded away from zero on S^d. Then no matter how $\gamma > 0$ is chosen, there *always* exists a sequence λ_n such that for each θ_0:

$$\lim_{n\to\infty} P\{|Y_i - \theta_0| \leq n^{-\gamma} \text{ for some } i \leq \lambda_n\} = 1$$

where here $Y_1, \ldots, Y_{\lambda_n}$ are iid μ. Indeed it suffices to take λ_n subject only to $\lim_{n\to\infty} \lambda_n n^{-d\gamma} = +\infty$. See Beran and Millar, 1986, section 5, where such a result was indicated for $\gamma = 1/2$.

(c) The concept of γ-funnel was introduced to ease the exposition and also because such a rate appears in many statistical applications. In theorem 3.1, one can easily replace the rate $n^{-\gamma}$ by slower rates, such as $[\log \log n]^{-\gamma}$, and the conclusion will still remain.

References

Beran, R.J. and Millar, P.W. (1987). Stochastic estimation and testing. *Ann. Statist.* **14**, 431-53.

—— (1988a). Multivariate symmetry models. Tech. Report No. 159, Statistics Dept., Univ. of Calif., Berkeley.

—— (1988b). Tests of fit for logistic models. Tech. Report No. 160, Statistics Dept., Univ. Dept., Berkeley.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Gaenssler, P. (1983). Empirical Processes. IMS Lecture Notes Series, vol 3.

Geman, S. and Huang, C.-R. (1982). Non-parametric maximum likelihood estimation by the method of sieves. *Ann. Statist.* **10**, 401-414.

Kuo, H.-H. (1975). Gaussian Measures in Banach Spaces. Lecture Notes in Math, vol 463. Springer. New York.

Rubenstein, Reuven Y. (1981). Simulation and the Monte Carlo Method. Wiley, N.Y.

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. 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. (Aug 1987, revised Oct 1988). 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 $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: 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.

177. MILLAR, P.W. (October 1988). Gamma-funnels in the domain of a probability, with statistical implications.

178. DONOHO, D.L. and LIU, R.C. (October 1988). Hardest one-dimensional subfamilies.

179. DONOHO, D.L. and STARK, P.B. (October 1988). Recovery of sparse signals from data missing low frequencies.

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.