Comparing Location Experiments*

By E. L. Lehmann

University of California, Berkeley

Technical Report No. 75
August 1986

*Research supported by National Science Foundation DMS84-01388.

Department of Statistics
University of California
Berkeley, California

Comparing Location Experiments*

By E. L. Lehmann

University of California, Berkeley

*Research supported by National Science Foundation DMS84-01388.

Department of Statistics
University of California
Berkeley, California

American Math. Soc. 1980 Classifications: Primary — 62C05

Key words and phrases: Comparison of experiments, Location families, Monotone decision procedures, Monotone likelihood ratio, Spread ordering, Tail ordering

Comparing Location Experiments*

By E.L. Lehmann

University of California, Berkeley

In Sections 1-3, the classical theory of the comparison of two experiments is reviewed with particular reference to the comparison of two location experiments. It is shown that the requirement of domination of one experiment by another for *all* decision problems is too strong to provide a reasonable basis for comparison. For one-parameter problems with monotone likelihood ratio it is therefore proposed to restrict the comparison to decision problems which are monotone in the sense of Karlin and Rubin (1956). Application of this weaker definition to the location problem is shown to give satisfactory results. A scale-free comparison of this type leads to a new tail-ordering of distributions, and this is explored in Section 6.

1. Introduction. An experiment $\mathbb{E}$ is a random quantity X and a family $\mathbb{P} = \{P_\theta, \theta \in \Omega\}$ of possible distributions of X. Let $\mathbb{F} = (Y, \mathbb{Q} = \{Q_\theta, \theta \in \Omega\})$ be another experiment, with the distributions P_θ and Q_θ corresponding to the same state of nature θ. The idea of patterning the definition of one experiment being more informative than another on the concept of sufficiency was initiated in an unpublished memorandum by Bohnenblust, Shapley and Sherman and developed into a theory by Blackwell (1951, 1953).

DEFINITION (1.1). The experiment $\mathbb{F}$ is more informative than (or sufficient for) $\mathbb{E}$ if there exists a random quantity Z with known distribution and a function h such that for all $\theta \in \Omega$

$$Y \textit{ is distributed as } Q_\theta \Longrightarrow h(Y,Z) \textit{ is distributed as } P_\theta.$$

An immediate consequence of (1.1) is:

(1.2)For any decision procedure δ based on X and any loss function $L(\theta,d)$ there exists a (possibly randomized) procedure δ' based on Y such that $R(\theta,\delta') = R(\theta,\delta)$ for all θ.

It was shown by Blackwell, and under more general conditions by Le Cam (1964) and Feldman and Ramamoorthi (1986) that typically not only does (1.1) imply (1.2) but the inverse implication also holds. In fact, in the same papers it is shown that (1.1) is implied by the following apparently even weaker statement

* Research supported by National Science Foundation DMS84-01388.

(1.3)Statement (1.2) with the conclusion $R(\theta,\delta') = R(\theta,\delta)$ for all θ replaced by $R(\theta,\delta') \leq R(\theta,\delta)$ for all θ.

A fourth condition which typically is equivalent to (1.1)-(1.3) is the Bayes condition that given any prior distribution Λ for θ, the Bayes risk is no larger when the experiment is based on Y than when it is based on X.

If Y is more informative in the sense of these definitions, which are equivalent in the situations to be considered in this paper, we shall write $Y \geq X$. If $Y \geq X$ and $X \geq Y$, the experiments X and Y will be said to be *equivalent* $(Y \sim X)$. The experiment Y is strictly more informative than $X (Y > X)$ if it is more informative than X but not equivalent to it.

The various possibilities are illustrated by the following example.

EXAMPLE 1.1 (NORMAL). Let $X = (X_1, \cdots, X_n)$, $Y = (Y_1, \cdots, Y_n)$ where the X_i and Y_i are independently normally distributed as $N(\xi,\sigma^2)$ and $N(\xi,\rho^2\sigma^2)$ respectively, with ρ known and $0 < \rho < 1$.

(i) $\sigma = \sigma_0$ KNOWN. Here $Y \geq X$ since $Y_i + Z_i$ has the same distribution as X_i when Z_i is $N(0,(1-\rho^2)\sigma_0^2)$. That Y is strictly more informative than X is seen by noting that the UMV unbiased estimators of ξ based on X and Y are respectively $\overline{X}$ with variance σ_0^2/n and $\overline{Y}$ with variance $\rho^2\sigma_0^2/n$. The latter variance cannot be matched by an unbiased estimator based on X.

(ii) $\xi = \xi_0$ KNOWN. Assuming without loss of generality that $\xi_0 = 0$, one sees that $Y \sim X$ since the variables Y_i/ρ have the same distribution as the X_i, and the variables ρX_i the same distribution as the Y_i.

(iii) ξ AND σ BOTH UNKNOWN. The surprising fact (see Hansen and Torgersen (1974)) is that in this case X and Y are not comparable.

This example illustrates the three possibilities: strict comparability, equivalence, and noncomparability, and the two principal methods used to determine comparability. If Y is more informative than X it is typically easy to determine the function $h(Y,Z)$ required in (1.1). To prove that Y is not more informative than X, one attempts to construct a statistical task which can be performed on the basis of X, but which either cannot be performed, or at least not as well, on the basis of Y. For this latter purpose it is often most convenient to find a function $a(\theta)$ which has an unbiased estimator based on X but not on Y. One reason for looking at this particular kind of task is that unbiasedness of an estimator requires only the calculation of first moments. Another reason is a result of DeGroot (1966, Theorem 4.1), which states essentially that if $\mathbb{P}$ is complete, then Y is more informative than X if and only if for each set B for which probability is defined $a(\theta) = P_\theta(X \in B)$ has a non-negative unbiased estimator based on Y.

2. Comparing two location experiments.

PROBLEM 1: In most of this paper, we shall be concerned with the case that $\mathbb{P}$ and $\mathbb{Q}$ are location families, i.e. are given by

(2.1) $\qquad P_\theta(X \leq x) = F(x - \theta)$ and $Q_\theta(Y \leq y) = G(y - \theta)$.

For this case it was shown by Boll (1955) and independently by Stone (1961) that condition (1.1) greatly simplifies in that the function $h(Y,Z)$ can always be taken to be of the form $h(Y,Z) = Y + Z$, i.e. $Y \geq X$ if and only if there exists a random variable Z independent of Y such that

(2.2) $\qquad Y$ has distribution $Q_\theta \Longrightarrow Y + Z$ has distribution P_θ.

As an example, suppose that F is the uniform distribution on $\left(-\frac{1}{2},\frac{1}{2}\right)$ and G the triangular distribution on $(-1,1)$. Then $Y \geq X$ since (2.2) holds with Z uniformly distributed on $\left(-\frac{1}{2},\frac{1}{2}\right)$.

If ϕ_X and ϕ_Y denote the characteristic functions of the distributions F and G respectively, (2.2) is equivalent to the condition that

(2.3) $\qquad \psi(t) = \phi_X(t)/\phi_Y(t)$ is a characteristic function.

As an example, suppose that F is the double exponential distribution with density $\frac{1}{2}e^{-|x|}$ and G the exponential distribution with density e^{-x}, $x > 0$. Then

$$\phi_X(t) = \frac{1}{1+t^2} \text{ and } \phi_Y(t) = \frac{1}{1-it}$$

so that $\psi(t) = 1/(1 + it)$ which is the characteristic function of $-Y$. Thus Y is more informative than X.

An immediate consequence of (2.3) is the surprising fact that if X is normally distributed then Y can not be more informative than X unless Y is also normal. This follows immediately from Cramér's theorem that if X is normal and X is the sum of two independent random variables Y and Z, then Y and Z must be normal. It is however disconcerting to learn that a normal location family $F(x - \theta) = \frac{1}{\sigma}\Phi\left(\frac{x-\theta}{\sigma}\right)$ even for very large σ is never less informative than a nonnormal $G(y - \theta)$ even if the latter distribution is very tightly concentrated about θ. Of course, if $G(y - \theta) = \frac{1}{\tau}\Phi\left(\frac{y-\theta}{\tau}\right)$ with $\tau < \sigma$, then Y is more informative than X by Example 1.1(i).

Given any Y with distribution $G(y - \theta)$, it is trivial to construct a less informative X simply by taking $F(x - \theta)$ to be the distribution of $Y + Z$ for any independent Z. On the other hand, any particular given F and G will be comparable only in very exceptional cases since this would require one of the characteristic functions ϕ_X, ϕ_Y to be a factor of the other.

The normal example suggests the possibility that such comparisons may be more readily available when G differs from F only by a scale factor, say

PROBLEM 2: $\mathbb{P}$ and $\mathbb{Q}$ are given by (2.1) with

$$G(y - \theta) = F\left(\frac{y - \theta}{\sigma}\right), \quad 0 < \rho < 1. \tag{2.4}$$

This problem has been considered by Stone (1961) and Goel and DeGroot (1979). A necessary and sufficient condition for Y given by (2.4) to be more informative than X is that

$$\psi(t) = \frac{\phi_X(t)}{\phi_X(\rho t)} \text{ is acharacteristic function for } all \ 0 < \rho < 1. \tag{2.5}$$

As was pointed out by Goel and DeGroot, the distributions F whose characteristic functions ϕ_X satisfy (2.5) were investigated by Lévy in a quite different context, and are called self-decomposable or belonging to class L. In particular, it follows from Lévy's work that all stable laws are self-decomposable and that on the other hand all self-decomposable distributions are infinitely divisible.

EXAMPLE 2.1 (DOUBLE EXPONENTIAL). As a simple example of a self-decomposable distribution which is not stable, consider the double exponential distribution with density $f(x) = \frac{1}{2} e^{-|x|}$ and characteristic function $\phi(t) = \frac{1}{1 + t^2}$. To see that this is self-decomposable note that

$$\psi(t) = \frac{1 + \rho^2 t^2}{1 + t^2} = \rho^2 \cdot 1 + (1 - \rho^2) \cdot \frac{1}{1 + t^2}.$$

Thus, $\psi(t)$ is the characteristic function of a variable which is equal to 0 with probability ρ^2 and has density $f(x)$ with probability $1 - \rho^2$.

On the other hand, let $F(x)$ be any distribution whose support is a finite interval such as the uniform or triangular distribution. Then F cannot be self-decomposable since it is not infinitely divisible.

3. The uniform case. It follows from the discussion of the preceding section that if X and Y are uniformly distributed as

$$X : U(\theta - \frac{1}{2}, \theta + \frac{1}{2}); \quad Y : U(\theta - \frac{\rho}{2}, \theta + \frac{\rho}{2}) \tag{3.1}$$

then very surprisingly Y is not more informative than X for all $0 < \rho < 1$. This does not rule out the possibility that it may be more informative for some ρ, and this is in fact the case.

THEOREM 3.1. Under (3.1), Y is more informative than X if and only if $\rho = 1/k$ for some positive integer k.

PROOF. (i) If $\rho = 1/k$ then (2.2) holds with Z being uniformly distributed over the points

$$-\frac{(k-1)}{2k}, -\frac{(k-3)}{2k}, \ldots, \frac{k-3}{2k}, \frac{k-1}{2k}.$$

This is easily checked and well known.

(ii) To prove that Y is not more informative than X for $\rho \neq \frac{1}{k}$, we shall exhibit a statistical task which X can perform in this case but Y can not. For this purpose, consider the estimation of

$$a(\theta) = P_\theta(X < 0) = \begin{cases} 1 \text{ if } \theta < -1/2 \\ \frac{1}{2} - \theta \text{ if } -\frac{1}{2} \leq \theta \leq \frac{1}{2} \\ 0 \text{ if } \theta > \frac{1}{2}. \end{cases} \tag{3.2}$$

The estimator

$$\delta_0(X) = \begin{matrix} 1 \text{ if } X < 0 \\ 0 \text{ otherwise} \end{matrix}$$

has the following two properties:

(a) $\delta_0(X)$ is unbiased;

(b) $Var_\theta[\delta_0(X)] = 0$ when $\theta < -1$ and when $\theta > 1$.

We shall now show that an estimator based on Y which shares these properties can exist only if $\rho = 1/k$. Here attention can be restricted to non randomized estimators $\delta(Y)$ since if $\delta'(Y,Z)$ has properties (a) and (b) where Z has a known distribution, so does $\delta(Y) = E[\delta'(Y,Z) \mid Y]$.

Suppose now that $\delta(Y)$ satisfies (a). Then $\delta(Y)$ must be constant (a.e.) for y sufficiently close to either $+\infty$ or $-\infty$, and these constants must be 0 and 1 respectively if $\delta(Y)$ is to be unbiased. By differentiating the unbiasedness condition

$$\int_{\theta-\frac{\rho}{2}}^{\theta+\frac{\rho}{2}} \delta(y)dy = a(\theta) \text{ for all } \theta,$$

one finds further that $\delta(y)$ must satisfy (a.e.)

$$\delta(y - \rho_2) - \delta(y + \frac{\rho}{2}) = \begin{array}{l} \text{if } -1/2 < y < 1/2 \\ 0 \text{ otherwise} \end{array} . \tag{3.3}$$

For almost all sequences $\delta(y \pm j\rho)$, $(j = 0,1,2,...)$, δ must therefore decrease from its value 1 near $j = -\infty$ to its value 0 near $j = +\infty$ by steps the sizes of which are restricted to 0 and ρ. This is possible only when $\rho = 1/k$ for some positive integer k.

It seems plausible that the difficulty in this example is caused by the insistence on δ being unbiased, and that an estimator $\delta(Y)$ with risk uniformly smaller than that of $\delta_0(X)$ does exist for most reasonable loss functions. It will be seen at the end of Section 4 that this is indeed the case.

4. Monotone decision problems. The examples discussed in the preceding sections make it clear that condition (1.1) for comparability is too strong to hold in many situations in which intuition suggests that one experiment is more informative than another.

Three approaches to weakening the requirements for comparability have been proposed.

(i) Le Cam (1964) replaces condition (1.2) by the approximate condition that each risk function based on X can be matched within ϵ by one based on Y.

(ii) Several authors suggest that comparisons should be made in terms of some measure(s) of information such as Fisher, Shannon, or Kullback-Leibler information.

(iii) Throughout the literature on the comparison of experiments, the suggestion occurs of comparing two experiments not for all decision problems but only for some family $\mathcal{C}$ of problems. For such restricted comparisons, (1.3) is the relevant determining condition. So as to distinguish this approach from the classical one, we shall say that Y is ***more effective*** than X with respect to the class $\mathcal{C}$ of decision problems concerning θ if for any problem in $\mathcal{C}$ (specified by a set of possible decisions and a loss function) and any decision procedure for this problem based on X, there exists a procedure δ' based on Y such that $R(\theta,\delta') \leq R(\theta,\delta)$ for all θ. It is this last approach with which we shall be concerned in the remainder of this paper. More particularly we shall be interested in defining a class $\mathcal{C}$ which is large enough to include most of the statistical problems of interest, but not so large that comparability becomes practically impossible.

A suitable such class was introduced by Karlin and Rubin (1956) for the case that θ is a real-valued parameter. This class, the class of monotone procedures, which we shall denote by $\mathcal{M}$ is defined in terms of the decision space and the permissible loss functions. It may be loosely described as follows. (For more detail, see Karlin and Rubin (1956).

CASE 1. (Infinite decision space). Here the problems being considered essentially constitute the class of point estimation problems in which the estimand is a real-valued increasing function $a(\theta)$. The loss function is assumed to satisfy $L(\theta, a(\theta)) = 0$ for all θ, with $L(\theta, d)$ increasing as a function of d as d moves away from $a(\theta)$ on either side.

CASE 2. (Finite decision space). A corresponding definition is given for the case of a finite number of decisions, for example the case of two decisions corresponding to testing a hypothesis $H : \theta \leq \theta_0$ against the alternatives $\theta > \theta_0$; or the three-decision problem, in which the hypothesis $H : \theta_0 \leq \theta \leq \theta_1$ $(\theta_0 \leq \theta_1)$ is to be accepted (decision d_0) or rejected in favor either of the alternatives

$\theta < \theta_0$ (decision d_1) or $\theta > \theta_1$ (decision d_2), with suitable losses for wrong decisions. (This class is treated in some detail in Ferguson (1967), Section 6.1).

The basic result of Karlin and Rubin (KR) is concerned not with a comparison problem but with a single family of densities $p_\theta(x)$ with monotone likelihood ratio in $T(x)$ which without loss of generality we shall take to be x. It then states that for any monotone decision problem for such a family p_θ, the class $\boldsymbol{M}$ of monotone procedures is essentially complete. Here a nonrandomized procedure $\delta(x)$ is monotone if

$$x < x' \Longrightarrow \delta(x) \leq \delta(x').$$

For the case of a finite number of possible decisions $d_1, \cdots, d_k$, a decision procedure can be described by k functions $\phi = (\phi_1, \cdots, \phi_k)$ with

$$0 \leq \phi_i \leq 1 \text{ and } \sum_{i=1}^{k} \phi_i(x) = 1.$$

A procedure ϕ is monotone provided there exist $k+1$ points $x_1 = -\infty \leq x_2 \leq \cdots \leq x_{k+1} = +\infty$ such that $\phi_i(x) = 1$ or 0 as x lies in or outside the interval (x_i, x_{i+1}) with possible randomization on the end points. [Under slightly stronger assumptions, Brown, Cohen and Strawderman (1976) show $\boldsymbol{M}$ to be complete rather than only essentially complete.]

The KR-class of monotone problems does not cover the estimation of the function $a(\theta)$ defined by (3.2) since $a(\theta)$ is not strictly increasing. This problem can be viewed as a combination of the finite case (decisions that $\theta < -\frac{1}{2}$ or $> \frac{1}{2}$) and the infinite case (estimation of $a(\theta)$ for $-1/2 \leq \theta \leq 1/2$), and their proof is easily seen to extend to such a mixed case.

5. Relative effectiveness of location families. The aim of the present section is to provide necessary and sufficient conditions for Y to be more effective than X in problems 1 and 2 with respect to the class $\mathbb{M}$. However, we begin by presenting the result in its natural, somewhat more general setting.

Let $\mathbb{P}$ and $\mathbb{Q}$ consist of the families

(5.1) $$P_\theta(X \leq x) = F_\theta(x) \text{ and } Q_\theta(Y \leq y) = G_\theta(y).$$

THEOREM 5.1. Let F_θ and G_θ have densities f_θ and g_θ with respect to a common σ-finite measure μ, and suppose that the families $f_\theta(x)$ and $g_\theta(x)$ have monotone likelihood ratio in x. Then a necessary and sufficient condition for Y to be more effective than X with respect to $\mathbb{M}$ is that the function

(5.2) $$h_\theta(x) = G_\theta^{-1}[F_\theta(x)] \text{ is a nondecreasing function of } \theta \text{ for each } x.$$

PROOF. By the KR-theorem, attention can be restricted to the ability of Y to dominate any monotone procedure δ based on X, i.e. to prove the existence of a procedure δ' based on Y with risk uniformly no greater than that of δ. Although this is not required for the proof, the construction will produce a monotone δ'.

Assuming (5.2), we begin by showing that for any θ_0 and any $0 < \alpha < 1$, given any level α test of $H : \theta \leq \theta_0$ against $\theta > \theta_0$ based on X, there exists a test for the same problem based on Y which is uniformly at least as powerful for $\theta > \theta_0$ and uniformly at most as powerful for $\theta < \theta_0$.

Since there exist tests based on X and Y respectively that are simultaneously uniformly most powerful against $\theta > \theta_0$ and uniformly least powerful against $\theta < \theta_0$, it is enough to show that the claimed relationship holds for these two tests. The optimal test based on X is given by the rejection region $X > a$ and that based on Y by $Y > b$, where for the sake of simplicity we assume for the moment that no randomization is required and that the points a and b are unique. Then a and b satisfy $F_{\theta_0}(a) = G_{\theta_0}(b) = 1 - \alpha$ and are therefore related by

(5.3) $$b = G_{\theta_0}^{-1}[F_{\theta_0}(a)].$$

The power of the two tests against any $\theta > \theta_0$ is $1 - F_\theta(a)$ and $1 - G_\theta(b)$ respectively, so that domination of X by Y requires

$$G_\theta(b) \leq F_\theta(a) \quad \text{for all } \theta > \theta_0$$

and analogously

$$G_\theta(b) \geq F_\theta(a) \quad \text{for all } \theta < \theta_0.$$

substitution from (5.3) shows these inequality to be equivalent to

$$G_{\theta_0}^{-1}[F_{\theta_0}(a)] \leq G_\theta^{-1}[F_\theta(a)] \quad \text{for all } \theta > \theta_0$$

with the inequality being reversed for $\theta < \theta_0$. Since this must hold for all a, it is seen that (5.2) is a necessary condition for Y to be more effective than X for all one-sided testing problems.

To complete the proof of sufficiency, one can now plug into the proof of the KR-theorem. The result established so far corresponds exactly to Lemma 3 of Karlin and Rubin. To construct a procedure based on Y to dominate any given monotone procedure with finite decision space based on X one now simply applies the arguments of their Lemma 4 and Theorem 1. Having established Y's dominance for all finite problems in $\boldsymbol{M}$, one then passes to the limit as in Section 7 of Karlin and Rubin to establish the result for the infinite case.

We next specialize Theorem 5.1 to Problem 1.

THEOREM 5.2. Let the distributions of X and Y be given by (2.1) and suppose that $F(x-\theta)$ and $G(x-\theta)$ have densities (with respect to Lebesgue measure) which are strongly unimodal. Then a necessary and sufficient condition for Y to be more effective than X with respect to $\boldsymbol{M}$ is that

$$\frac{G^{-1}[F(b)] - G^{-1}[F(a)]}{b-a} \le 1 \text{ for all } a < b. \tag{5.4}$$

PROOF. If f and g are strongly unimodal, the families $f(x-\theta)$ and $g(x-\theta)$ have monotone likelihood ratio in x. The function h defined by (5.2) in the present case reduces to

$$h_\theta(x) = G^{-1}[F(x-\theta)] + \theta$$

and hence the inequality $h_\theta(x) \le h_{\theta'}(x)$ to (5.4) with $a = x - \theta'$ and $b = x - \theta$.

Putting $F(a) = u$, $F(b) = v$, we can rewrite (5.4) as

$$G^{-1}(v) - G^{-1}(u) \le F^{-1}(v) - F^{-1}(u) \text{ for all } 0 < u < v < 1. \tag{5.5}$$

Condition (5.5) states that F is more spread out than G in the sense that any two quantities are at least as far apart under F than they are under G. These are just the circumstances under which one would expect inferences about the location of F to be more difficult than those about the location of G. Restriction to monotone problems has thus replaced the original rather strange and - as it turned out - not very useful condition (2.2) with one that nicely quantifies our intuition.

Condition (5.5) was discussed in Bickel and Lehmann (1976) as the definition of F being *more spread out* than G. It was also pointed out there that if F^{-1} and G^{-1} are differentiable, (5.5) is equivalent to

$$\frac{f[F^{-1}(y)]}{g[G^{-1}(y)]} \le 1 \text{ for all } 0 < y < 1. \tag{5.6}$$

Still another equivalent condition is given in Theorem 1 of Bickel and Lehmann.

Let us finally specialize the above results to Problem 2, defined by (2.4). Since in that case $G(y) = F(y/\rho)$, $0 < \rho < 1$, we have $G^{-1}(y) = \rho F^{-1}(y)$, and it is seen that (5.5) holds for all F. We have thus proved

THEOREM 5.3. Let the distributions of X and Y be given by (2.4) and suppose that F has a density f which is strongly unimodal. Then Y is more effective than X relative to $\boldsymbol{M}$ for all $0 < \rho < 1$.

Since the uniform distribution is strongly unimodal, this establishes in particular the conjecture expressed at the end of Section 3 for any loss function meeting the KR conditions of case 1 of Section 4.

6. Scale-free comparisons and a tail-ordering. Conditions (5.4)-(5.6) make it possible for any particular F and G to decide on the comparability of the two families (2.1) with respect to $\boldsymbol{M}$. We shall in the present section consider a scale-free version of this comparison.

PROBLEM 3. Given F and G, does there exist ρ sufficiently small so that $G\left(\frac{y-\theta}{\rho}\right)$ is more effective than $F(x-\theta)$ with respect to $\boldsymbol{M}$?

When such a ρ exists, we shall say that G has *a more effective shape* than F with respect to $\boldsymbol{M}$.

It follows from (5.5) that G has a more effective shape than F if and only if there exists ρ such that

$$\frac{G^{-1}(v)-G^{-1}(u)}{F^{-1}(v)-F^{-1}(u)} \leq \frac{1}{\rho} \quad \text{for all } 0<u<v<1$$

i.e. if and only if

$$\frac{G^{-1}(v)-G^{-1}(u)}{F^{-1}(v)-F^{-1}(u)} \text{ is bounded.} \tag{6.1}$$

If F^{-1} and G^{-1} are differentiable, this reduces to

$$\frac{f\,[F^{-1}(y)]}{g\,[G^{-1}(y)]} \text{ is bounded.} \tag{6.2}$$

The function $f\,[F^{-1}(y)]$ is studied in a different context by Parzen (1979) who calls it the density-quantile function, and considers its limiting behavior as $y \to 0$ or 1 a measure of tail weight. He also evaluates this function for a number of important distributions. Condition (6.2) is also closely related to the s-ordering of symmetric distributions introduced by van Zwet (1964), who requires that

$$\frac{f\,[F^{-1}(y)]}{g\,[G^{-1}(y)]} \text{ be } \begin{matrix} \text{decreasing for } y>1/2 \\ \text{increasing for } y<1/2 \end{matrix}. \tag{6.3}$$

The ratio therefore attains its maximum at $y=1/2$ and if f is bounded, (6.3) implies (6.2).

As discussed by Loh (1984b), Van Zwet's s-ordering and some related orderings (including Loh's t-ordering) take into account not only the heaviness of the tail but also the behavior of f (its "peakedness") at the center. In contrast, condition (6.2) provides a definition of pure tail-ordering. As an example, if G is double exponential and F Cauchy, then F is not tail-heavier according to the s-ordering, but it is strictly tail-heavier according to the ordering (6.2) since (see Parzen (1979))

$$f\,[F^{-1}(u)] = \begin{matrix} 1-u & \text{for } u > 1/2 \\ u & \text{for } u < 1/2 \end{matrix} \quad \text{when } F \text{ is standard double exponential}$$

and

$$f\,[F^{-1}(u)] = \frac{1}{\pi}\sin^2(\pi u) \sim (1-u)^2 \quad \text{when } F \text{ is standard Cauchy}$$

where here and below $\sim$ means that the ratio tends to a positive finite construct as $u \to 1$. As another example, note that according to the s-ordering, the double exponential is heavier-tailed than the logistic, while the two are equivalent according to the ordering (6.3). This is seen from the fact that for the logistic distribution

$$f\,[F^{-1}(u)] = u(1-u).$$

Let us now return to Problem 3 and provide examples of some situations in which G has a more effective shape than F by being lighter-tailed according to the definition (6.2). Suppose for example that G is uniform (or has any other distribution whose density g is bounded away from 0 and ∞ on its support). It then follows from (6.2) that the shape of G is more effective than that of any distribution F with bounded density, and that it is strictly more effective if the bounded density of F is not bounded away from 0, e.g. if F is triangular, normal, etc.

As another example, suppose that G is the triangular distribution with density

$$g(x) = 1 - |x|, \quad -1 < x < 1.$$

Then $g\,[G^{-1}(y)] = \sqrt{2(1-y)}$, and G is lighter-tailed than the logistic and double exponential distributions but heavier-tailed than the extrema value distribution for which $f\,[F^{-1}(y)] = -(1-y)\log(1-y)$ (see Parzen, l.c.)

7. The case of n observations. So far, attention has been restricted to a single observation from model (2.1) or (2.4). We shall now generalize Theorem 5.1 to the case that $X_1, \cdots, X_n$; $Y_1, \cdots, Y_n$ are i.i.d. with distributions F_θ and G_θ respectively, where θ continuous to be real-valued. The definition of the class $\boldsymbol{M}$ does not require any modification since it concerns only the parameter and decision space but not the sample space.

THEOREM 7.1. Under the assumptions of Theorem 5.1, condition (5.2) is necessary and sufficient for $Y = (Y_1, .., Y_n)$ to be more effective than $X = (X_1, .., X_n)$ with respect to $\boldsymbol{M}$.

PROOF. A decision procedure $\delta(x_1, \cdots, x_n)$ is said to be monotone if

(7.1) $$x_i \le x'_i \text{ for } all\ i = 1, \cdots, n \text{ impliesthat } \delta(x) \le \delta(x').$$

The KR-theorem discussed in Section 4 was generalized to n i.i.d. variables from a MLR family by Oosterhoff (1969), Brown, Cohen and Strawderman (1976), and Van Houwelingen and Verbeek (1985), who state that as in the one-dimensional case the class of monotone procedures is essentially complete. We can therefore as in the proof of Theorem 5.1 restrict attention to the problem of dominating any monotone procedure based on X.

In analogy to the proof of Theorem 5.1, we begin by showing that for any θ_0 and $0 < \alpha < 1$, given any monotone level α test ϕ of $H : \theta \le \theta_0$ against $\theta > \theta_0$ based on $X = (X_1, \cdots, X_n)$ there exists a test for the same problem based on $Y = (Y_1, \cdots, Y_n)$ which is uniformly at least as powerful for $\theta > \theta_0$ and uniformly at most as powerful for $\theta < \theta_0$. Note, however, that there is now no longer a unique monotone level α test (which is uniformly most powerful) but a large class of such test.

To establish a test dominating ϕ, let us denote the distributions F_{θ_0} and G_{θ_0} by F and G respectively and replace the experiment Y by the equivalent experiment Z with $Z_i = h(Y_i)$ where $h = F^{-1}G$. Then X_i and Z_i have the same distribution $F = F_{\theta_0}$ when $\theta = \theta_0$. We shall now show that the distribution H_θ of Z_i satisfies

(7.2) $$\begin{aligned} H_\theta(z) \le F_\theta(z) &\quad \text{when } \theta > \theta_0 \\ \ge \qquad &\quad \text{when } \theta < \theta_0 . \end{aligned}$$

To see this, note that

$$H_\theta(z) = P_\theta[F^{-1}G(Y_i) \le z] = G_\theta\{G^{-1}[F(z)]\}.$$

Therefore $H_\theta(z) \le F_\theta(z)$ provided

$$G_{\theta_0}^{-1}[F_{\theta_0}(z)] \le G_\theta^{-1}[F_\theta(z)]$$

which by (5.2) is the case when $\theta_0 < \theta$ with the opposite inequality holding when

$\theta < \theta_0$.

Condition (7.2) states that Z_i is stochastically larger than X_i.

Let $\phi(X)$ be any monotone test based on $X = (X_1, \cdots, X_n)$ and let $\phi^*(Y) = \phi(Z)$ be the same test based on $Z = (Z_1, \cdots, Z_n)$. Then it follows from the basic property of stochastically ordered random variables, given for example in chapter 3, Lemma 1 of Lehmann (1986), that

$$\begin{aligned} E_\theta \phi(X) &\leq E_\theta \phi(Z) \quad \text{for } \theta > \theta_0 \\ &\geq \qquad\qquad \text{for } \theta < \theta_0 \end{aligned}$$

as was to be proved.

For more general decision procedures, we can proceed as in Brown et al (1976) or Van Houwelingen and Verbeek (1985).

It is an immediate consequence of Theorem 7.1 that not only Theorem 5.1 but also Theorems 5.2 and 5.3 and the results of Section 6 remain valid when the single variable X is replaced by a sample $X_1, \cdots, X_n$.

References

Bickel P.J. and Lehmann, E.L. (1976). Descriptive Statistics for nonparametric models. IV in Contributions to Statistics. Jaroslav Hajek Memorial Volume Prague.

Blackwell, D. (1951). Comparison of experiments. Proc. Second Berk. Symp. Math. Statist. Probab. Univ. of Calif. Press, 93-102.

Blackwell, D. (1953). Equivalent comparison of experiments. Ann. Math. Statist. 24, 265-272.

Boll, C.H. (1955).

Comparison of experiments in the infinite case and the use of invariance in establishing sufficiency. Unpublished Ph.D. thesis, Stanford Univ.

Brown, L.D.; Cohen, A.; and Strawderman, W.E. (1976). A complete class theorem for strict monotone likelihood ratio with applications. Ann. Statist. 4, 712-722.

Feldman, D. (1972). Some properties of Bayesian ordering of experiments. Ann. Math. Statist. 43, 1428-1440.

Feldman, D. and Ramamoorthi, R.V. (1986). Equivalent comparison of experiments - another proof of Blackwell's theorem. To be published.

Ferguson, T.S. (1976). Mathematical Statistics. Academic Press. New York.

Goel, P.K. and DeGroot, M.H. (1979). Comparison of experiments and information measures. Ann. Statist. 7, 1066-1077.

Hansen, O.H. and Torgersen, E.N. (1974). Comparison of linear normal experiments. Ann. Statist. 2, 367-373.

Karlin, S. and Rubin, H. (1956). The theory of decision procedures for distributions with monotone likelihood ratio. Ann. of Math. Statist. 27, 272-299.

Le Cam L. (1964). Sufficiency and approximate sufficiency. Ann. Math. Statist. 35, 1419-1455.

Lehmann, E.L. (1986). Testing Statistical Hypotheses. 2nd Ed. John Wiley. New York.

Loh, W.-Y. (1984a). Strong unimodality and scale mixtures. Ann. Inst. Statist. Math. 36, 441-449.

Loh, W.-Y. (1984b). Bounds on ARE's for restricted classes of distributions defined via tail-orderings. Ann. Statist. 12, 685-701.

Oosterhoff, J. (1969). Combination of one-sided statistical tests. Mathem. Centrum Tracts 28. Amsterdam.

Stone, M. (1961). Non-equivalent comparisons of experiments and their use for experiments involving location parameters. Ann. Statist. 32, 326-332.

Van Houwelingen, H. and Verbeek, A. (1985). On the construction of monotone symmetric decision rules for distributions with monotone likelihood ratio. Scand. J. Statist. 12, 73-81.

Van Zwet, W.R. (1964). Convex transformations of random variables. Math. Centrum. Amsterdam.

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. DkAPER, 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. RIZZARDI, F. (Aug 1987). Two-Sample t-tests where one population SD is known.

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.

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.

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 $E \max_{1\le k\le n} S_k^+ / ES_n^+$: 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.

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. (September 1987). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (September 1987). Unachievable 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. and MACGIBBON, B. (November 1987). 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. (November 1987). 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. (February 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.

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.