The Annals of Statistics
1974, Vol. 2, No. 3, 528-534

ON THE MAXIMUM LIKELIHOOD ESTIMATION OF STOCHASTICALLY ORDERED RANDOM VARIATES[1]

BY TIM ROBERTSON AND F. T. WRIGHT

The University of Iowa

Brunk, Franck, Hanson and Hogg (1966) ("Maximum likelihood estimation of the distributions of two stochastically ordered random variables," *J. Amer. Statist. Assoc.* **61** 1067–1080) found and studied maximum likelihood estimates of a pair of stochastically ordered distribution functions. In this paper we discuss a generalization of this problem in which we do not require the domain of these "distribution functions" to be the real line. We think of the order restriction we impose on these "distribution functions" as an analogue of stochastic ordering on the line. Maximum likelihood estimates are found and strong uniform consistency properties are discussed.

1. Introduction. In this paper we consider the following generalization of a problem explored by Brunk, Franck, Hanson and Hogg (1966). Suppose $(\Omega, \mathscr{A})$ is a measurable space, $\mathscr{A}$ contains all of the one point subsets of Ω, F and G are probability measures on $\mathscr{A}$ and $\mathscr{L}$ is a σ-lattice of measurable subsets of Ω. Suppose $P_1, P_2, \cdots, P_m$ are observed values of random variates, each having range space Ω and induced probability measure F, and $Q_1, Q_2, \cdots, Q_n$ are observed values of random variates, each having range Ω and induced measure G. Suppose we have prior information that $F(L) \geqq G(L)$ for each $L \in \mathscr{L}$ and we wish to find estimates of F and G (based on $P_1, P_2, \cdots, P_m, Q_1, Q_2, \cdots, Q_n$) which satisfy this restriction. The case when $\Omega = (-\infty, \infty)$, $\mathscr{A}$ is the collection of Borel subsets of Ω and $\mathscr{L}$ is the collection of semi-infinite intervals of the form $(-\infty, x)$ or $(-\infty, x]$ was considered by Brunk, *et al.* (1966). They found maximum likelihood estimates and argued that these estimates are strongly consistent.

Suppose $\Omega = E_2 = (-\infty, \infty) \times (-\infty, \infty)$, $\mathscr{A}$ is the collection of Lebesgue subsets of E_2 and $\mathscr{L}$ is the collection of all lower layers (i.e., $L \in \mathscr{L}$ if and only if $x_1 \leqq x_2$, $y_1 \leqq y_2$ and $(x_2, y_2) \in L$ imply that $(x_1, y_1) \in L$). (The members of $\mathscr{L}$ are Lebesgue but not necessarily Borel measurable.) Suppose Q_1 and Q_2 are two dimensional random vectors which induce the measures F and G on $\mathscr{A}$, respectively. Note that $F(L) \geqq G(L)$ for all lower layers L if and only if $\alpha(Q_1)$ is stochastically less than or equal to $\alpha(Q_2)$ for all functions $\alpha(\cdot)$ of two variables which are non-decreasing in each variable. On the other hand, it is rather easy to construct examples of random vectors $Q_1 = (Q_{11}, Q_{12})$ and $Q_2 = (Q_{21}, Q_{22})$ with $P[Q_{11} \leqq x, Q_{12} \leqq y] \geqq P[Q_{21} \leqq x, Q_{22} \leqq y]$ for all (x, y), and yet $Q_{11} + Q_{12}$ is not stochastically less than or equal to $Q_{21} + Q_{22}$. It might, therefore, be reasonable

Received August 1972; revised July 1973.

[1] This research was sponsored by the National Science Foundation under Grant GP-30966X.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 60F15.

Key words and phrases. Stochastic ordering, maximum likelihood estimation, consistency.

to say in E_2 that finding maximum likelihood estimates subject to the restriction $F(L) \geqq G(L)$ for all lower layers L is the natural two dimensional analogue of the problem considered by Brunk, *et al.* (1966).

We return to the general case. Because we are motivated by the problem described in the preceding paragraph we shall continue to refer to members of $\mathscr{L}$ as lower layers. Let $\mathscr{U}$ be the σ-lattice of complements of members of $\mathscr{L}$. We refer to members of $\mathscr{U}$ as upper layers. In order to simplify some of the notation used in the remainder of this paper we will henceforth only use the symbol $L(U)$, with or without subscripts, primes, etc., to denote lower (upper) layers. Define the relation $\ll$ on Ω by:

$$\omega' \ll \omega \quad \text{if and only if} \quad \omega \in L \in \mathscr{L} \quad \text{imply that} \quad \omega' \in L .$$

This relation is a preorder on Ω in the sense that it is reflexive and transitive but not necessarily antisymmetric (cf. Robertson (1967)). Throughout this paper we make the following assumptions about $\mathscr{L}$.

(1.1) $\mathscr{L}$ is complete (i.e., $\mathscr{L}$ is closed under arbitrary unions and intersections).

(1.2) Ω together with $\ll$ forms a lattice. (Note that this implies $\ll$ is a partial ordering.)

Define the counting measures $N_F(\cdot)$ and $N_G(\cdot)$ on $\mathscr{A}$ by: $N_F(A)(N_G(A))$ is the number of observations from $F(G)$ in A. Let $N(A) = N_F(A) + N_G(A)$ and let $R(A) = N_G(A)/N_F(A)$ where this ratio is understood to be ∞ if $N_F(A) = 0$. The following remark is easily demonstrated.

REMARK 1.1. If $N(A \cap B) = 0$, then $R(A \cup B)$ is between $R(A)$ and $R(B)$. (We shall refer to this as the averaging property of $R(\cdot)$.)

If $\omega \ll \omega'$ then we say ω is to the lower left of ω'. Define the discrete probability measures $\bar{F}$ and $\bar{G}$ on $\mathscr{A}$ by the following construction: Let U_1 be the smallest upper layer containing all the observation points. (Recall, we are assuming $\mathscr{L}$, and thus $\mathscr{U}$, is complete.) If there are both F-observations and G-observations at a point then these F-observations are considered to be to the lower left of these G-observations. Choose U_2 a proper subset of U_1 so that $R(U_1 - U_2) \geqq R(U_1 - U')$ for all proper subsets U' of U_1. (This is possible since $R(\cdot)$ has only a finite number of values on such sets.) The next remark guarantees that U_2 may be chosen as small as possible.

REMARK 1.2. If U_2' is another proper subset of U_1 with the property that $R(U_1 - U_2') \geqq R(U_1 - U')$ for all proper subsets U' of U_1, then $U_2 \cap U_2'$ also has this property.

PROOF. Recalling the way U_2 was chosen and applying the averaging property to the two disjoint sets $U_1 - (U_2 \cup U_2')$ and $U_2' - U_2$, we obtain

$$(1.3) \qquad R(U_1 - (U_2 \cup U_2')) \leqq R(U_1 - U_2) \leqq R(U_2' - U_2) .$$

Similarly, we obtain

$$R(U_2' - U_2) \leqq R(U_1 - (U_2 \cap U_2')) \leqq R(U_1 - U_2') . \tag{1.4}$$

It follows from (1.3) and (1.4) that $R(U_1 - (U_2 \cap U_2')) \geqq R(U_1 - U_2)$ and the proof is completed in this case.

Thus we may choose U_2 "as small as possible." If $U_2 = \emptyset$ we are finished; otherwise choose the "smallest" proper subset U_3 of U_2 so that $R(U_2 - U_3) \geqq R(U_2 - U')$ for all proper subsets U' of U_2. Continuing in this fashion the process terminates since there are only a finite number of observation points and since $N(U_i - U_{i+1}) > 0$ for each i. (Note that $N(U_1 - U_2) > 0$ by the definition of U_1 and U_2 and that if $N(U_i - U_{i+1}) = 0$ for some $i > 1$ then $R(U_{i-1} - U_i) = R(U_{i-1} - U_{i+1})$ which contradicts the way U_i was chosen.) We thus obtain a sequence $U_1 \supset U_2 \supset \cdots \supset U_k$ $(U_{k+1} = \emptyset)$ of upper layers such that $R(U_i - U_{i+1}) \geqq R(U_i - U')$ for all U' a proper subset of U_i, and in fact $R(U_i - U_{i+1}) > R(U_i - U'')$ for U'' a proper subset of U_{i+1}. Let $S_i = U_i - U_{i+1}$ and $L_i = U_{i+1}^c$ for $i = 1, 2, \cdots, k$. We shall abuse the notation by writing $N_F(P)$ $(N_G(P))$ instead of $N_F(\{P\})$ $(N_G(\{P\}))$ for the number of observations from F (G) at P. Define the discrete measures $\bar{F}$ and $\bar{G}$ on $\mathscr{A}$ by:

$$\bar{F}(P) = \frac{N_F(P)}{N_F(S_i)} \cdot \frac{N(S_i)}{m+n}$$

if $P \in S_i$ and $N_F(S_i) > 0$;

$$\bar{G}(P) = \frac{N_G(P)}{N_G(S_i)} \cdot \frac{N(S_i)}{m+n}$$

if $P \in S_i$ and $N_G(S_i) > 0$. If $P \in S_i$ and $N_F(S_i) = 0$ (which implies that $i = 1$) let $\bar{F}(P) = \bar{G}(P)$ and if $P \in S_i$ and $N_G(S_i) = 0$ (which implies that $i = k$) let $\bar{G}(P) = \bar{F}(P)$.

It is easy to see that $\bar{F}$ and $\bar{G}$ are probability measures. In Section 2 we argue that $\bar{F}(L) \geqq \bar{G}(L)$ for every lower layer L; $(\bar{F}, \bar{G})$ provide "maximum likelihood estimates" subject to these restrictions and give a representation theorem for $(\bar{F}, \bar{G})$. In Section 3 we discuss the strong, uniform consistency of $\bar{F}$ and $\bar{G}$.

2. Properties of $(\bar{F}, \bar{G})$.

THEOREM 2.1. *$\bar{F}(L) \geqq \bar{G}(L)$ for all lower layers L.*

PROOF. Using the definitions of $\bar{F}$, $\bar{G}$, and U_{i+1}, one can argue that $\bar{F}(L \cap S_i) \geqq \bar{G}(L \cap S_i)$ for each i and the desired result follows.

For each upper layer U such that $N_F(U) > 0$ let $\mathscr{F}(U)$ be the collection of all upper layers U' such that $N_F(U - U') \neq 0$. Similarly, if $N_G(U) \neq 0$ let $\mathscr{G}(U)$ be the collection of all upper layers U' such that $N_G(U - U') \neq 0$.

THEOREM 2.2. *If $N_F(P) > 0$ then*

$$\bar{F}(P) = \frac{N_F(P)}{m+n} \min_{U \ni P} \max_{U' \in \mathscr{F}(U)} \frac{N(U-U')}{N_F(U-U')}$$

and if $N_G(P) > 0$ *then*

$$\bar{G}(P) = \frac{N_G(P)}{m+n} \max_{U \ni P} \min_{U' \in \mathscr{S}(U)} \frac{N(U - U')}{N_G(U - U')} .$$

PROOF. An argument similar to one given for Theorem 2.4 of Robertson and Wright (1973) may be employed here. The proofs and many of the details that have been omitted in this paper are contained in Robertson and Wright (1972).

Let $\hat{F}$ and $\hat{G}$ be the empirical probability measures (i.e., $\hat{F}(A) = N_F(A)/m$, $\hat{G}(A) = N_G(A)/n$) and for any pair (F, G) of probability measures on $\mathscr{A}$ let $\Lambda(F, G)$ denote the likelihood product for (F, G) (i.e., $\Lambda(F, G) = \prod_{i=1}^m F(P_i) \prod_{j=1}^n G(Q_j)$ with $F(P) = F(\{P\})$). The next result follows directly from the definitions of $\bar{F}$ and $\bar{G}$.

LEMMA 2.3. *If* $k = 1$ *then* $\bar{F} = \hat{F}$ *and* $\bar{G} = \hat{G}$.

Let $\mathscr{O} = \{O_1, O_2, \cdots, O_\alpha\}$ be the smallest lattice of points in Ω which contains all of the observation points and suppose these points are labeled so that $O_1 \ll O_i \ll O_\alpha$ for all i. Let $\mathscr{S}_1$ be the collection of all pairs (F^*, G^*) of discrete probability measures on $\mathscr{A}$ with the following three properties:

(2.1) $$F^*(L) \geqq G^*(L) \quad \text{for all} \quad L \in \mathscr{L},$$

(2.2) $$F^*(P) = 0 \quad \text{unless} \quad P = P_i \quad \text{for some } i \quad \text{or} \quad N_F(S_1) = 0 \quad \text{and} \quad P = O_1,$$

(2.3) $$G^*(P) = 0 \quad \text{unless} \quad P = Q_i \quad \text{for some } i \quad \text{or} \quad N_G(S_k) = 0 \quad \text{and} \quad P = O_\alpha .$$

LEMMA 2.4. *If* (F_1^*, G_1^*) *is any pair of probability measures on* $\mathscr{A}$ *such that* $F_1^*(L) \geqq G_1^*(L)$ *for all* $L \in \mathscr{L}$ *then there exists a pair* (F^*, G^*) *in* $\mathscr{S}_1$ *such that* $\Lambda(F^*, G^*) \geqq \Lambda(F_1^*, G_1^*)$.

PROOF. Construct a sequence $(F_1^*, G_1^*), (F_2^*, G_2^*), \cdots, (F_4^*, G_4^*) = (F^*, G^*)$ of pairs of probability measures on $\mathscr{A}$ such that $F_i^*(L) \geqq G_i^*(L)$ for all $L \in \mathscr{L}$ and $\Lambda(F_i^*, G_i^*) \leqq \Lambda(F_{i+1}^*, G_{i+1}^*)$ for $i = 1, 2, 3$. For each point P in Ω let $L(P)$ be the smallest lower layer containing P. Define $F_2^*(G_2^*)$ by concentrating all the $F_1^*(G_1^*)$ mass in the set $\{O_\alpha\} \cup L(O_\alpha)^c$ at O_α. Next, for each $O_i \in \mathscr{O}$ let $A(O_i)$ be the collection of all points P in Ω such that $P \ll O_i$ and it is not true that $P \ll O_j$ for any $O_j \ll O_i$ with $O_j \neq O_i$. Observe that $A(O_1), A(O_2), \cdots, A(O_\alpha)$ partition $L(O_\alpha)$ and then define the discrete probability measures F_3^* and G_3^* on $\mathscr{A}$ by concentrating all of the $F_2^*(G_2^*)$ mass on $A(O_i)$ at O_i. Clearly, $\Lambda(F_3^*, G_3^*) \geqq \Lambda(F_2^*, G_2^*)$. Suppose $L \in \mathscr{L}$ and let $A = \sum_{O_i \in L} A(O_i)$. Since $\mathscr{L}$ is complete, $\mathscr{L}$ can be characterized as the collection of all sets B in $\mathscr{A}$ with the property that $\omega \ll \omega' \in B$ imply that $\omega \in B$. We want to argue that $A \in \mathscr{L}$ so suppose $P \ll Q \in A$. Then $Q \in A(O_j)$ where $O_j \in L$, $P \in L(O_\alpha)$ so that it can be argued that $P \in A(\inf\{O_i; P \ll O_i\}) = A(O_l)$. Thus, $O_l \ll O_j$ and $O_l \in L$. It follows that $P \in A$ and $A \in \mathscr{L}$. Hence

$$F_3^*(L) = F_2^*(A) \geqq G_2^*(A) = G_3^*(L) .$$

Finally, define $F_4{}^*(G_4{}^*)$ by concentrating all the $F_3{}^*(G_3{}^*)$ mass at point O_ι such that $N_F(O_\iota) = 0$ $(N_G(O_\iota) = 0)$ at O_1 (O_α). This completes the argument since $(F_4{}^*, G_4{}^*) \in \mathscr{S}_1$.

Thus, in our search for a maximum likelihood estimate we may restrict our attention to $\mathscr{S}_1$. However, as in Brunk *et al.* (1966), this becomes a problem of finding a maximum of a continuous function over a closed and bounded subset of some Euclidean space whose dimension is no larger than $m + n + 2$. It is well known that such a maximum exists.

A maximum likelihood estimate in $\mathscr{S}_1$ can be constructed from $(\bar{F}, \bar{G})$ by concentrating all of the $\bar{F}$ $(\bar{G})$ mass on S_1 (S_k) at O_1 (O_α). However, this might result in an estimate of, say, F which assigns mass to a point, O_1, which was not an observation from either F or G. This seems very arbitrary, so let $\mathscr{S}$ be the class of all pairs (F^*, G^*) of discrete probability measures on $\mathscr{A}$ such that (2.1) holds together with:

(2.4) $$F^*(\mathscr{O}) = G^*(\mathscr{O}) = 1$$

(2.5) $$F^*(O_\iota) = 0 \quad \text{unless} \quad N_F(O_\iota) > 0 \quad \text{or} \quad O_\iota \in S_1, \quad N_F(S_1) = 0 \quad \text{and} \quad N_G(O_\iota) > 0,$$

(2.6) $$G^*(O_\iota) = 0 \quad \text{unless} \quad N_G(O_\iota) > 0 \quad \text{or} \quad O_\iota \in S_k, \quad N_G(S_k) = 0 \quad \text{and} \quad N_F(O_\iota) > 0.$$

LEMMA 2.5. *If (F^*, G^*) is a maximum likelihood pair in $\mathscr{S}_1$ then there exists a pair (F', G') in $\mathscr{S}$ such that $\Lambda(F', G') \geqq \Lambda(F^*, G^*)$.*

PROOF. We construct F'. The construction of G' is similar. If $N(O_1) > 0$ or $F^*(O_1) = 0$ then F^* has the desired properties. Thus, we assume $N(O_1) = 0$ and $F^*(O_1) > 0$. If $N_F(S_1) > 0$ then every observation from G has an F-observation to its lower left and it follows that there exists a lower layer L_0 such that $L_0 \cap U_1 \neq \emptyset$ and $F^*(L_0 - L') < G^*(L_0 - L')$ for all proper subsets L' of L_0 which contain O_1. Now if O_2 is any one of the points in $L_0 \cap U_1$ such that $N_F(O_2) > 0$, then one can shift some of the F^* probability from O_1 to O_2 without destroying the stochastic ordering and thereby increasing the likelihood product. Thus $N_F(S_1) = 0$. If $F^*(O_1) > G^*(S_1)$ then, as above, one can find a point O_2 such that $N_F(O_2) > 0$ and shifting some of the F^* probability from O_1 to O_2 does not destroy the stochastic ordering. Thus $F^*(O_1) = G^*(S_1)$ and letting $F' = G^*$ on S_1 and $F' = F^*$ otherwise gives the desired function.

THEOREM 2.6. *$(\bar{F}, \bar{G})$ is a maximum likelihood estimate.*

PROOF. Suppose $(F' G') \in \mathscr{S}$ is a maximum likelihood estimate. Define (F'', G'') as follows:

$$F''(O_j) = \frac{N_F(O_j)}{N_F(S_\iota)} F'(S_\iota); \qquad O_j \in S_\iota,\ N_F(S_\iota) > 0$$

$$= \frac{N_G(O_j)}{N_G(S_1)} F'(S_1); \qquad O_j \in S_1,\ N_F(S_1) = 0,$$

$$G''(O_j) = \frac{N_G(O_j)}{N_G(S_i)} G^*(S_i)\,; \qquad O_j \in S_i,\ N_G(S_i) > 0$$
$$= \frac{N_F(O_j)}{N_F(S_k)} G^*(S_k)\,; \qquad O_j \in S_k,\ N_G(S_k) = 0\,.$$

Note that

(2.7) $$\Lambda(F'', G'') \geqq \Lambda(F', G')$$ and

(2.8) $$F''(S_1 + S_2 + \cdots + S_j) \geqq G''(S_1 + S_2 + \cdots + S_j)\,; \qquad j = 1, 2, \cdots, k\,.$$

(It may not be the case that $F''(L) \geqq G''(L)$ for all L.) Now consider the result in Brunk *et al.* (1966) and apply it to a sampling situation on the line where we observe $N_F(S_1)x^s$ followed by $N_G(S_1)y^s$ followed by $N_F(S_2)x^s$, etc. It follows that

$$\frac{\Lambda(\bar{F}, \bar{G})}{\Lambda(F'', G'')} = \frac{\prod_{i=1}^k (N(S_i)/(m+n))^{N_F(S_i)} \prod_{i=1}^k (N(S_i)/(m+n))^{N_G(S_i)}}{\prod_{i=1}^k (F''(S_i))^{N_F(S_i)} \prod_{i=1}^k (G''(S_i))^{N_G(S_i)}} \geqq 1\,.$$

This, together with (2.7) and Theorem 2.1, completes the argument.

3. Consistency. Let

$$\delta_m = \sup_{L \in \mathscr{L}} |\hat{F}(L) - F(L)| \qquad \text{and} \qquad \varepsilon_n = \sup_{L \in \mathscr{L}} |\hat{G}(L) - G(L)|\,.$$

THEOREM.

$$\sup_{A \in \mathscr{A}} |\bar{F}(A) - \hat{F}(A)| \leqq \frac{n}{m+n} [\delta_m + \varepsilon_n] \qquad \textit{and}$$

$$\sup_{A \in \mathscr{A}} |\bar{G}(A) - \hat{G}(A)| \leqq \frac{m}{m+n} [\delta_m + \varepsilon_n]\,.$$

PROOF. It can be shown that $N(S_i)/N_F(S_i) > N(S_{i+1})/N_F(S_{i+1})$ for $i = 1, 2, \cdots, k-1$ and choosing $l = \max\{j;\ N(S_j)/N_F(S_j) \geqq (m+n)/m\}$, we see that for any $A \in \mathscr{A}$

$$\hat{F}(U_{l+1}^c) - \bar{F}(U_{l+1}^c) \leqq \bar{F}(A) - \hat{F}(A) \leqq \bar{F}(U_{l+1}^c) - \hat{F}(U_{l+1}^c)\,.$$

Next, one shows that for $i = 1, 2, \cdots, k$,

$$0 \leqq \bar{F}(U_{i+1}^c) - \hat{F}(U_{i+1}^c) \leqq n \cdot (m+n)^{-1}[\delta_m + \varepsilon_n]$$

and the first conclusion of the theorem follows immediately. The second conclusion is obtained similarly.

Combining this result with well-known properties of the empirical distribution function yields consistency properties for $(\bar{F}, \bar{G})$. For example, it follows from the strong law of large numbers and Scheffé's Theorem that if F and G are discrete then

$$P[\lim_{m,n\to\infty} \sup_{A \in \mathscr{A}} |\bar{F}(A) - F(A)| = 0,$$
$$\lim_{m,n\to\infty} \sup_{A \in \mathscr{A}} |\bar{G}(A) - G(A)| = 0] = 1\,.$$

If $\Omega = E_\beta$, $\mathscr{A}$ is the collection of Lebesgue subsets of E_β, $\mathscr{L}$ is the σ-lattice induced by the partial ordering $(x_1, x_2, \cdots, x_\beta) \ll (y_1, y_2, \cdots, y_\beta)$ if and only if $x_i \leqq y_i$; $i = 1, 2, \cdots, \beta$ and both F and G are absolutely continuous with respect to Lebesgue measure, then Blum (1955) proved that $\delta_m \to_{\text{a.s.}} 0$ and $\varepsilon_n \to_{\text{a.s.}} 0$. It follows that

$$(3.1) \qquad P[\lim_{m,n\to\infty} \sup_{L\in\mathscr{L}} |\bar{F}(L) - F(L)| = 0, \quad \lim_{m,n\to\infty} \sup_{L\in\mathscr{L}} |\bar{G}(L) - G(L)| = 0] = 1 .$$

On the other hand, if $\beta = 2$ and both F and G are singular and assign their mass uniformly to $\{(x, y); x \geqq 0, y \geqq 0, x+y = 1\}$ and $\{(x, y); x \geqq 1, y \geqq 1, x+y = 3\}$, respectively, then $F(L) \geqq G(L)$ for all lower layers L. Furthermore, with probability one, any observation from F is to the lower left of all the G-observations. It follows that $\bar{F} = \hat{F}$, $\bar{G} = \hat{G}$, and DeHardt's (1970) theorem gives

$$P[\lim_{m\to\infty} \sup_{L\in\mathscr{L}} |\bar{F}(L) - F(L)| = 1] = P[\lim_{m\to\infty} \sup_{L\in\mathscr{L}} |\hat{F}(L) - F(L)| = 1] = 1 .$$

A similar result holds for $\bar{G}$. Theorem 2 of DeHardt (1971) could be used to obtain more general conditions which ensure that (3.1) holds.

Kiefer (1961) proves that if F is absolutely continuous with respect to Lebesgue measure, then

$$P\left[\lim_{m\to\infty} \left(\frac{m}{\log\log m}\right)^{\frac{1}{2}} \sup_{P\in E_\beta} |\hat{F}(L(P)) - F(L(P))| = 2^{\frac{1}{2}}\right] = 1 .$$

It would be interesting to know if such an iterated logarithm result holds for δ_m or if $\sup_P |\bar{F}(L(P)) - \hat{F}(L(P))|$ can be related to $\sup_P |\hat{F}(L(P)) - F(L(P))|$ and $\sup_P |\hat{G}(L(P)) - G(L(P))|$ to obtain an iterated logarithm result for $(\bar{F}, \bar{G})$.

REFERENCES

[1] BLUM, J. R. (1955). On the convergence of empiric distribution functions. *Ann. Math. Statist.* **26** 527–529.

[2] BRUNK, H. D., FRANCK, W. E., HANSON, D. L. and HOGG, R. V. (1966). Maximum likelihood estimation of the distributions of two stochastically ordered random variables. *J. Amer. Statist. Assoc.* **61** 1067–1080.

[3] DEHARDT, JOHN (1970). A necessary condition for Glivenko-Cantelli convergence in E_n. *Ann. Math. Statist.* **41** 2177–2178.

[4] DEHARDT, JOHN (1971). Generalizations of the Glivenko-Cantelli theorem. *Ann. Math. Statist.* **42** 2050–2055.

[5] KIEFER, J. (1961). On large deviations of the empiric distribution function of vector random variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649–660.

[6] ROBERTSON, TIM (1967). On estimating a density which is measurable with respect to a σ-lattice. *Ann. Math. Statist.* **38** 482–493.

[7] ROBERTSON, TIM and WRIGHT, F. T. (1972). On the maximum likelihood estimation of stochastically ordered random variates. Technical Report Number 16, Statistics Department, Univ. of Iowa.

[8] ROBERTSON, TIM and WRIGHT, F. T. (1973). Multiple isotonic median regression. *Ann. Statist.* **1** 422–432.

DEPARTMENT OF STATISTICS
101 MACLEAN HALL
THE UNIVERSITY OF IOWA
IOWA CITY, IOWA 52242