A NOTE ON MUTUAL SINGULARITY OF PRIORS[1]

By David A. Freedman

University of California, Berkeley

1. Introduction. The main result of this note is an improvement of (8.1) of [3], and is a by-product of an extensive collaboration with Dubins on that paper. I am also indebted, as usual, to David Blackwell for many helpful suggestions. Section 3, on branching processes, is almost self-contained, and may be of general interest.

2. Summary. Let $0 < r < 1$. Let μ be a probability on S, the closed unit square, which assigns probability 1 to the vertical line with abscissa r. Let Δ be the set of distribution functions on the closed unit interval I. Endow Δ with its weak $*$ Borel σ-field. For any closed sub-interval H of I, let $\langle H\rangle$ be the linear function which maps I onto H and sends 0 to the left endpoint of H; and H_0 (respectively, H_1) be the image of $[0, r]$ (respectively, $[r, 1]$) under $\langle H\rangle$. Write B for the set of all finite sequences of 0's and 1's (including the empty sequence $\emptyset$). For $b \varepsilon B$ and $\epsilon = 0$ or 1, $I_{b\epsilon} = (I_b)_\epsilon$, where $I_\emptyset = I$. For $G \varepsilon \Delta$, let $|G|$ be the unique probability on I with $G(x) = |G|[0, x]$ for all $x \varepsilon I$. Let $Y_b : b \varepsilon B$ be independent random variables, such that (r, Y_b) has distribution μ. Let P_μ, a probability on Δ, be the distribution of the (random) $F \varepsilon \Delta$ satisfying: $|F|(I_\emptyset) = 1$ and $|F|(I_{b0}) = Y_b\,|F|(I_b)$ for all $b \varepsilon B$. For a more detailed description, see [2], or Sections 1 and 2 of [3].

Say $F \varepsilon \Delta$ is *strictly singular* with respect to $G \varepsilon \Delta$ if there is no x for which the ratio of $F(x + h) - F(x)$ to $G(x + h) - G(x)$ converges to a finite, positive limit as h tends to 0. The object of this note is to prove the

THEOREM. *Let* $0 < r < 1$. *Let* μ *and* ν *be distinct probabilities on* S, *both assigning measure* 1 *to the vertical line with abscissa* r. *Then there are weak* $*$ *Borel subsets* C *and* D *of* Δ, *with* $P_\mu(C) = P_\nu(D) = 1$, *and each distribution function in* C *strictly singular with respect to each distribution function in* D.

Let $\mu^* = \mu\{(r, 0)\} + \mu\{(r, 1)\}$, and $\nu^* = \nu\{(r, 0)\} + \nu\{(r, 1)\}$. The easy case (either $\mu^* = 1$ or $\nu^* = 1$) of the Theorem is proved in Section 5. The harder case ($\mu^* < 1$ and $\nu^* < 1$) is proved in Section 4. Section 3 contains preliminary material on branching processes, which may be of general interest.

3. Branching processes. For this section, j is a positive integer, and n is a non-negative integer. J_n is the set of n-tuples formed with $0, \cdots, j - 1$; the only element of J_0 is the (empty) 0-tuple $\emptyset$. If $b \varepsilon J_n$, and $i = 0, \cdots, j - 1$, then b followed by i, namely bi, is in J_{n+1}, and is a *child* of b. The *j-tree* is $J = \bigcup_{n=0}^\infty J_n$, and $b \varepsilon J$ is a *node*. Moreover, p is a real number with $0 \leqq p \leqq 1$, and $(p_1, \cdots, p_j)$ is a probability distribution on $(1, \cdots, j)$, with probability

Received 6 August 1965.

[1] Prepared with the partial support of the Sloan Foundation and the National Science Foundation GP 2593.

generating function Ψ; that is, $\Psi(x) = p_1x + \cdots + p_jx^j$. The stochastic process $\{Z_b, L_b : b \varepsilon J\}$ is a *(p, Ψ)-process over the j-tree J*; that is, its joint distribution satisfies the following conditions. Z_b takes only three values: 0, 1, undefined. L_b takes only two values: live, dead. Z_b is undefined if and only if L_b is dead. If L_b is dead, so is L_{bi} for $i = 0, \cdots, j - 1$. Let N_b be the number of L_{bi} which are live, for $i = 0, \cdots, j - 1$. Then $L_\varnothing$ is live, $Z_\varnothing$ is 1 with probability p, 0 with probability $1 - p$, and $N_\varnothing$ has probability generating function Ψ. Of course, $Z_\varnothing$ and $N_\varnothing$ need not be independent. For $m = 0, 1, \cdots$, given Z_b and L_b for $b \varepsilon \bigcup_{n=0}^{m-1} J_n$, and L_b for $b \varepsilon J_m$: provided $c \varepsilon J_m$ and L_c is live, the pairs (Z_c, N_c) are conditionally independent, with common conditional distribution equal to the unconditional distribution of $(Z_\varnothing, N_\varnothing)$. Finally, for $m = 0, 1, \cdots$, given Z_b, L_b and N_b for $b \varepsilon \bigcup_{n=0}^{m} J_n$: provided $c \varepsilon J_m$, the j-tuples $(L_{c0}, \cdots, L_{c(j-1)})$ are conditionally independent; and for each $c \varepsilon J_m$, all j-tuples formed with N_c values live and $j - N_c$ values dead are conditionally equally likely for $(L_{c0}, \cdots, L_{c(j-1)})$.

"L_b is live" may be abbreviated to "b is live." If $b \varepsilon J_n$ and $0 \leqq i \leqq n$, then $b(i) \varepsilon J_i$ is the first i components of b, and is an *ancestor* of b.

LEMMA 1. (i) *$b \varepsilon J_n$ is live with probability $[j^{-1}\Psi'(1)]^n$.*

(ii) *Given that $b \varepsilon J_n$ is live, $b(0), \cdots, b(n - 1)$ are live; and $N_{b(0)}, \cdots, N_{b(n-1)}$ are conditionally independent with common conditional probability generating function*

$$x \to x\Psi'(x)/\Psi'(1).$$

PROOF. Verification is needed only for $n = 1$. Given $N_\varnothing = i$, b is live with conditional probability $\binom{j-1}{i-1}/\binom{j}{i} = i/j$. Thus, b is live and $N_\varnothing = i$ with unconditional probability ip_i/j. ●

LEMMA 2. *Given that $b \varepsilon J_n$ is live, and given which children of $b(0), \cdots, b(n - 1)$ are live: as c ranges over the live children other than $b(i + 1)$ of $b(i)$, for $i = 0, \cdots, n - 1$, Z_c are conditionally independent,* 1 *with conditional probability* p, 0 *with conditional probability* $1 - p$.

PROOF. Clear. ●

If $0 < \alpha < 1$ and $0 \leqq \theta \leqq 1$, let

$$m(\theta, \alpha) = [\theta/\alpha]^\alpha[(1 - \theta)/(1 - \alpha)]^{1-\alpha}.$$

An interesting fact in Chernoff (1952) is recorded here as:

LEMMA 3. *Let $X_1, X_2, \cdots$ be independent random variables, each assuming the value* 1 *with probability θ, and* 0 *with probability $1 - \theta$. Let $\theta < \alpha < 1$, and let n be a positive integer. Then $X_1 + \cdots + X_n \geqq n\alpha$ with probability no more than $[m(\theta, \alpha)]^n$.*

If b and c in J_{n+1} are different children of the same node in J_n, they are *brothers*. If $0 \leqq \alpha, \gamma \leqq 1$ are real numbers, $b \varepsilon J_n$ is *α-exceptional* if it is live and of the live brothers c of b, a fraction at least α have $Z_c = 1$. Moreover, b is *(α, γ)-exceptional* if it is live and there are at least $n\gamma$ integers $i = 1, \cdots, n$ for which $b(i)$ is α-exceptional.

LEMMA 4. *Let α be a real number, with $p < \alpha < 1$, and let $b \varepsilon J_n$. Given that b is live: for $i = 1, \cdots, n$, the events "$b(i)$ is α-exceptional" are conditionally independent, with common conditional probability no more than $\Psi'[m(p, \alpha)]/\Psi'(1)$.*

PROOF. As usual, only $n = 1$ needs verification. By Lemma 1, given b is live, the number $N_\varnothing - 1$ of live brothers of b has conditional probability generating function $x \rightarrow \Psi'(x)/\Psi'(1)$. By Lemmas 2 and 3, given b is live and given $N_\varnothing$, the conditional probability that b is α-exceptional is no more than $m(p, \alpha)^{N_\varnothing - 1}$. ●

LEMMA 5. *Let α and γ be real numbers, with $p < \alpha < 1$ and $\beta = \Psi'[m(p, \alpha)]/\Psi'(1) < \gamma < 1$. The probability that there is an (α, γ)-exceptional node in J_n is at most $[\Psi'(1)m(\beta, \gamma)]^n$.*

PROOF. $m(\theta, \gamma)$ increases with θ for $\theta < \gamma$. So, using Lemmas 3 and 4, given that $b \varepsilon J_n$ is live: b is (α, γ)-exceptional with conditional probability at most $[m(\beta, \gamma)]^n$. But there are j^n nodes in J_n; apply Lemma 1. ●

A *path* through J is a sequence $b_0, b_1, \cdots$ such that $b_0 = \varnothing$, and for all n, there is an $i = 0, \cdots, j - 1$ with $b_{n+1} = b_n i$. The probability P on the two-point set $\{0, 1\}$ assigns mass p to 1, and P^J is the probability on the set of functions f from J to $\{0, 1\}$ for which: as b ranges over J, the functions $f \rightarrow f(b) = 0$ or 1 are independent with common distribution P.

The next Lemma, and its proof, are taken from [3], with the permission of Dubins.

LEMMA 6. *If $p < \alpha < 1$ and $m(p, \alpha) < j^{-1}$, then for P^J-almost all functions f from J to $\{0, 1\}$, there is an $n(f) < \infty$ such that: for each $n \geqq n(f)$ and path $b_0, b_1, \cdots$ through J, $f(b_0) + \cdots + f(b_{n-1}) < n\alpha$.*

PROOF. Let E_n be the set of all functions g from J to $\{0, 1\}$ such that, for some path $b_0, b_1, \cdots$ through J, $g(b_0) + \cdots + g(b_{n-1}) \geqq n\alpha$. By Lemma 3, $P^J(E_n) \leqq j^{n-1}[m(p, \alpha)]^n$, which is summable in n. ●

4. Proof of the Theorem in case $\mu^* < 1$ and $\nu^* < 1$. It is necessary to give a more formal definition of P_μ; unfortunately, this involves further notation. B is the 2-tree, V is the vertical line segment $\{(x, y): x = r, 0 \leqq y \leqq 1\}$, and V^B is the set of functions τ from B to V. For $b \varepsilon B$ and $\tau \varepsilon V^B$, $\tau(b) = (r, \tau_2(b))$, with $0 \leqq \tau_2(b) \leqq 1$. If μ is a probability on (the Borel subsets of) V, then μ^B is the probability on (the Borel subsets of) V_B for which: as b ranges over B, the functions $\tau \rightarrow \tau(b)$ are independent with common distribution μ. A function M will now be defined from V^B to Δ so that $P_\mu = \mu^B M^{-1}$. Introduce the set 2^Z of all functions from the positive integers Z to the two-point set $\{0, 1\}$. For $b \varepsilon B$, $\bar{b}$ is the set of $\xi \varepsilon 2^Z$ which extend b; thus $\overline{\varnothing} = 2^Z$ and $\overline{00} = \{\xi: \xi \varepsilon 2^Z, \xi(1) = \xi(2) = 0\}$. For $\tau \varepsilon V^B$, $P(\tau)$ is the probability on 2^Z for which

$$P(\tau)(\overline{b0}) = \tau_2(b)P(\tau)(\bar{b}), \qquad \text{all} \quad b \varepsilon B.$$

For $\xi \varepsilon 2^Z$, let $f(\xi) = \bigcap_{n=0}^\infty I_{\xi(1)\cdots\xi(n)} \varepsilon I$. Then $M(\tau) \varepsilon \Delta$ satisfies: $|M(\tau)| = P(\tau)f^{-1}$.

The object of this section is to construct Borel subsets C^* and D^* of V^B, such that:

$$\mu^B(C^*) = \nu^B(D^*) = 1; \tag{1}$$

$$\sigma \varepsilon C^* \text{ and } \tau \varepsilon D^* \text{ implies } M(\sigma) \text{ is strictly singular with respect to } M(\tau). \tag{2}$$

If $\tau \varepsilon V^B$, then $b \varepsilon B$ is *τ-live* unless it has an ancestor either of the form $c1$ with $\tau_2(c) = 1$, or of the form $c0$ with $\tau_2(c) = 0$. If n is a non-negative integer, B_n is the set of n-tuples of 0's and 1's. If k is a positive integer and $b \neq c \varepsilon B_n$ have the same ancestor in B_{n-k} (that is, $b(n-k) = c(n-k)$), then b and c are *k-cousins*. If A is a Borel subset of V, n and k are positive integers, $0 \leqq \alpha \leqq 1$, then $b \varepsilon B_{nk}$ is *τ-(A, k, α)-exceptional* if it is τ-live, and of its τ-live k-cousins c, a fraction at least α have $\tau(c) \varepsilon A$. For $0 \leqq \gamma \leqq 1$, $(A, k, \alpha; n, \gamma)$ is the set of $\tau \varepsilon V^B$ for which: there is a τ-live $b \varepsilon B_{nk}$, with a fraction at least γ of $b(k), \cdots, b(nk)$ being τ-(A, k, α)-exceptional.

Recall $\mu^* = \mu\{(r, 0)\} + \mu\{(r, 1)\}$. Let $f_1(x) = \mu^* x + (1 - \mu^*)x^2$, $f_{n+1}(x) = f_1(f_n(x))$. From XII.5 of [4], f_n is the μ^B-probability generating function of the random variable whose value at τ is the number of τ-live $b \varepsilon B_n$.

Abbreviate

$$M = m(\mu(A), \alpha) \quad \text{and} \quad \beta_k = f_k'(M)/f_k'(1). \tag{3}$$

The main step in proving the Theorem is establishing this inequality: if $\mu(A) < \alpha < 1$, and $\beta_k < \gamma < 1$, then

$$\mu^B(A, k, \alpha; n, \gamma) \leqq [(2 - \mu^*)^k m(\beta_k, \gamma)]^n. \tag{4}$$

Proof of (4). The inequality will be verified by applying Lemma 5 to a suitable $(\mu(A), f_k)$-process on the 2^k-tree, as follows. In Section 3, put $j = 2^k$. Order B_{nk} and J_n lexicographically, and let 0_n be the order-preserving map of J_n onto B_{nk}. If $\tau \varepsilon V^B$, then τ^* is this function from J to $\{0, 1, \text{undefined}\} \times \{\text{live, dead}\}$. If $b \varepsilon J_n$, the second coordinate of $\tau^*(b)$ is live or dead according as $0_n b \varepsilon B_{nk}$ is τ-live or τ-dead; the first coordinate is undefined if and only if the second is dead; if the second is live, the first is 1 or 0 according as $\tau(0_n b) \varepsilon A$ or $\not\varepsilon A$. Unfortunately, the process $\{\tau \to \tau^*(b) : b \varepsilon J\}$, on the probability space (V^B, μ^B), is not a $(\mu(A), f_k)$-process.

Let π be a mapping of J into $(0, \cdots, j-1)!$; that is, for each $b \varepsilon J$, $\pi(b)$ is a permutation of $(0, \cdots, j-1)$. Then π^*, a permutation of J, is defined by this induction: $\pi^*(\varnothing) = \varnothing$; if π^* has been defined on J_n, $b \varepsilon J_n$, and $i = 0, \cdots, j-1$, then $\pi^*(bi)$ is $\pi^*(b)$ followed by the $\pi(b)$-image of i. Of course, if $b \varepsilon J_n$ is an ancestor of c, then $\pi^*(b) \varepsilon J_n$ is an ancestor of $\pi^*(c)$. If g is a function from J to some set, then πg is this function from J to the same set: $(\pi g)(b) = g(\pi^*(b))$, $b \varepsilon J$.

Let (Ω, Q) be a probability space. Let $\boldsymbol{\pi}$ be a mapping from $\Omega \times J$ to $(0, \cdots, j-1)!$, with these properties:

(i) for each $b \varepsilon J$, $\boldsymbol{\pi}(\cdot, b)$ is measurable, and takes each of its $j!$ possible values with probability $1/j!$;

(ii) as b ranges over J, the $\boldsymbol{\pi}(\cdot, b)$ are independent.

If $\omega \varepsilon \Omega$ and $\tau \varepsilon V^B$, then $[\pi(\omega, \cdot)]\tau^*$ is a function from J to {0, 1, undefined} $\times$ {live, dead}. If also $b \varepsilon B$, let $Z_b(\omega, \tau)$ be the first coordinate of $\{[\pi(\omega, \cdot)]\tau^*\}(b)$, and $L_b(\omega, \tau)$ the second. Plainly, the stochastic process $\{Z_b, L_b : b \varepsilon J\}$, defined on the probability triple $(\Omega, Q) \times (V^B, \mu^B)$, is a $(\mu(A), f_k)$-process on the 2^k-tree, in the sense of Section 3. There is an (α, γ)-exceptional node in J_n for this process evaluated at (ω, τ) if and only if $\tau \varepsilon (A, k, \alpha; n, \gamma)$. Apply Lemma 5, noting that $f_k'(1) = (2 - \mu^*)^k$, to complete the proof of (4).

Let $(A, k, \alpha; f.o., \gamma)$ be the set of $\tau \varepsilon V^B$ with $\tau \varepsilon (A, k, \alpha; n, \gamma)$ for finitely many n only. If $\mu(A) < \alpha < 1$, and $\mu^* < 1$, there is a positive integer k and a positive $\gamma < \frac{1}{2}$ with

$$\mu^B(A, k, \alpha; f.o., \gamma) = 1. \tag{5}$$

Proof of (5). By (4) and the Borel-Cantelli lemma, it is enough to choose k and γ so that

$$(2 - \mu^*)^k m(\beta_k, \gamma) < 1 \tag{6}$$

and

$$\beta_k < \gamma, \tag{7}$$

where β_k and M are defined in (3). Since $\mu(A) < \alpha < 1$, $M < 1$, $f_k(M) \to 0$ as $k \to \infty$. Since $f_1'(0) = \mu^*$, for each $\epsilon > 0$ there is an $E(\epsilon) < \infty$ with

$$\beta_k \leqq E(\epsilon)(\mu^* + \epsilon)^k/(2 - \mu^*)^k, \qquad k = 1, 2, \cdots. \tag{8}$$

Since $m(\theta, \gamma) \leqq 2\theta^\gamma$, the left side of (6) is no more than

$$2E(\epsilon)^\gamma[(2 - \mu^*)^{1-\gamma}(\mu^* + \epsilon)^\gamma]^k. \tag{9}$$

Since $\mu^* < 1$, $(2 - \mu^*)\mu^* < 1$ and $\mu^*/(2 - \mu^*) < 1$. Choose $\epsilon > 0$ so small that $(2 - \mu^*)(\mu^* + \epsilon) < 1$, and $(\mu^* + \epsilon)/(2 - \mu^*) < 1$. Choose $\gamma < \frac{1}{2}$ so large that $(2 - \mu^*)^{1-\gamma}(\mu^* + \epsilon)^\gamma < 1$. Then choose k so large that $\beta_k < \gamma$, using (8), and so large that (9) is less than 1, which completes the proof of (5).

Proof of the Theorem in case $\mu^* < 1$ and $\nu^* < 1$. Find a Borel subset A of V and a real number α with $\mu(A) < \alpha < \nu(A)$. Let A' be the complement of A in V. Use (5) to find a positive integer k and a positive $\gamma < \frac{1}{2}$ with

$$\mu^B(A, k, \alpha; f.o., \gamma) = 1$$

and

$$\nu^B(A', k, 1 - \alpha; f.o., \gamma) = 1.$$

A special case. Some nuisances in the rest of the proof disappear if, for example, $r = \frac{1}{2}$, μ and ν both concentrate on the two-point set $\{(\frac{1}{2}, 0), (\frac{1}{2}, w)\}$, and $0 < \mu\{(\frac{1}{2}, w)\} < \nu\{(\frac{1}{2}, w)\} < 1$, where $0 < w < 1$. This case will now be argued. For A, use the one-point set $\{(\frac{1}{2}, w)\}$. For C^*, take the set of $\sigma \varepsilon (A, k, \alpha; f.o., \gamma)$ with $\sigma(b) = (\frac{1}{2}, 0)$ or $(\frac{1}{2}, w)$ for all $b \varepsilon B$. For D^*, take the set of $\tau \varepsilon (A', k, 1 - \alpha; f.o., \gamma)$ with $\tau(b) = (\frac{1}{2}, 0)$ or $(\frac{1}{2}, w)$ for all $b \varepsilon B$. Property (1) is clear. For (2), let $\sigma \varepsilon C^*$, $\tau \varepsilon D^*$, $x \varepsilon I$. It must be seen that

(10) the ratio of $M(\sigma)(x + h) - M(\sigma)(x)$ to $M(\tau)(x + h) - M(\tau)(x)$ does not converge to a finite, positive limit as $h \to 0$.

Write $b(n, x)$ for the first n digits of the non-terminating binary expansion of x. Clearly, (10) holds if for some n, $b(n, x)$ is either σ-dead or τ-dead. So, suppose that for all n, $b(n, x)$ is both σ-live and τ-live. If $b \varepsilon B_n$ is an ancestor of $c \varepsilon B_{n+k}$, then c is a *k-child* of b. Let N be the set of n for which: at least one k-child c of $b(nk, x)$ is σ-live and τ-live and has $\sigma(c) = (\frac{1}{2}, 0)$ and has $\tau(c) = (\frac{1}{2}, w)$. Plainly, N is infinite, even having density $\geqq 1 - 2\gamma > 0$.

For $b \varepsilon B_n$, $b^* \varepsilon I$ has binary expansion b followed by 0's, and $b^{**} = b^* + 2^{-n}$. If $\rho \varepsilon V^B$ and $u, v \varepsilon I$, then $\rho[u, v]$ is the slope of the chord joining the points in the graph of $M(\rho)$ whose abscissas are u and v. There is a $\delta > 0$, depending only on k and w, with the following property. If $n \varepsilon N$, there is an $i \leqq k + 1$ and an i-child d of $b(nk, x)$ such that:

$$\frac{\sigma[b(nk, x)^*, d^{**}]}{\tau[b(nk, x)^*, d^{**}]} \Big/ \frac{\sigma[b(nk, x)^*, b(nk, x)^{**}]}{\tau[b(nk, x)^*, b(nk, x)^{**}]}$$

and

$$\frac{\sigma[d^{**}, b(nk, x)^{**}]}{\tau[d^{**}, b(nk, x)^{**}]} \Big/ \frac{\sigma[b(nk, x)^*, b(nk, x)^{**}]}{\tau[b(nk, x)^*, b(nk, x)^{**}]}$$

both differ from 1 in absolute value by δ or more. Since the closed interval $[b(nk, x)^*, b(nk, x)^{**}]$ shrinks to x, and x is in either $[b(nk, x)^*, d^{**}]$ or $[d^{**}, b(nk, x)^{**}]$, (10) holds. This completes the proof in the special case.

To return to the general case, if $b \varepsilon B_n$ is an ancestor of $c \varepsilon B_{n+i}$ for some $i = 0, \cdots, k$, then c is a *k-descendant* of b. If K is a subset of V, $\tau \varepsilon V^B$, then $b \varepsilon B$ is (τ, K, k)*-good* if $\tau(c) \varepsilon K$ for all k-descendants c of b. If $0 \leqq g \leqq 1$, then $b \varepsilon B_n$ is (τ, K, k, g)*-good* if a fraction at least g of $b(0), \cdots, b(n)$ are (τ, K, k)-good. Let $\langle K, k, g\rangle$ be the set of all $\tau \varepsilon V^B$ for which there is an $n(\tau) < \infty$ such that $n \geqq n(\tau)$ and $b \varepsilon B_n$ imply b is (τ, K, k, g)-good.

Use Lemma 6 to find compact subsets K and K' of V, with $K \subset A$, $K' \subset A'$, $(r, 0)$ and $(r, 1)$ not points of accumulation of $K \cup K'$, and for $g = 1 - [(1 - 2\gamma)/k]$,

$$\mu^B\langle K \cup K', k, g\rangle = \nu^B\langle K \cup K', k, g\rangle = 1.$$

Let

$$C^* = \langle K \cup K', k, g\rangle \cap (A, k, \alpha; f.o., \gamma)$$

and

$$D^* = \langle K \cup K', k, g\rangle \cap (A', k, 1 - \alpha; f.o., \gamma).$$

Property (1) is clear. The proof of (2) is a routine generalization of the special case, especially in view of the similar material in Sections 1 and 5–8 of [3]. This completes the discussion of the Theorem in case $\mu^* < 1$ and $\nu^* < 1$.

5. Proof of the Theorem in case $\mu^* = 1$ or $\nu^* = 1$. If A is a Borel subset of the closed unit interval I, then $\hat{A}$ is the weak* Borel subset of $F \varepsilon \Delta$ which assign measure 1 to some point in A. If $\mu^* = \nu^* = 1$, there are disjoint Borel subsets U and V of I, with $P_\mu(\hat{U}) = P_\nu(\hat{V}) = 1$. For example, use (9.17) of [3] and the Strong Law of Large Numbers. Clearly, $F \varepsilon \hat{U}$ is strictly singular with respect to $G \varepsilon \hat{V}$. If $\mu^* = 1$ but $\nu^* < 1$, take $\hat{I}$ for C, and take for D the set of all continuous $F \varepsilon \Delta$ with $F(0) = 0$. By (4.4) of [3], $P_\nu(D) = 1$. This completes the proof of the Theorem.

REFERENCES

[1] CHERNOFF, HERMAN (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[2] DUBINS, LESTER E. and FREEDMAN, DAVID A. (1963). Random distribution functions. *Bull. Amer. Math. Soc.* **69** 548–551.

[3] DUBINS, LESTER E. and FREEDMAN, DAVID A. (1965). Random distribution functions. To appear in Fifth Berkeley Symposium.

[4] FELLER, WILLIAM (1958). *An Introduction to Probability Theory and Its Applications.* **1** (2nd ed.). Wiley, New York.