The Annals of Mathematical Statistics
1972, Vol. 43, No. 1, 308-319

A STABILITY OF SYMMETRIZATION OF PRODUCT MEASURES WITH FEW DISTINCT FACTORS

BY JAMES HANNAN[1] AND J. S. HUANG[2]

Michigan State University and University of Waterloo

0. Summary. Let $\mathscr{P} = \{F_0, \cdots, F_m\}$ be a class of probability measures on $(\mathscr{X}, \mathscr{B})$. For any signed measure τ on $\mathscr{B}^N$, let τ^* be the average of τg over all $N!$ permutations g and let $||\tau|| = \vee\{|\tau(C)| : C \in \mathscr{B}^N\}$. Let $d_{ij} = ||F_i - F_j||$ and $K(x) = .5012 \cdots x(1-x)^{-\frac{3}{2}}$. For any nonnegative integral partitions $\mathbf{N} = (N_0, \cdots, N_m)$ and $\mathbf{N}' = (N_0', \cdots, N_m')$ of N, let $\delta_i = N_i' - N_i$ and $\Lambda_i = (N_i' \wedge N_i) + 1$. With $\tau = \times F_i^{N_i} - \times F_i^{N_i'}$ and $n = \#\{i \mid \delta_i \neq 0\} - 1$, we bound $||\tau^*||^2$ by

$$\text{(T 3)} \qquad nK(d) \sum \delta_i^2 \Lambda_i^{-1} \quad \text{with} \quad d = \vee\{d_{ij} \mid \delta_i \neq 0, \delta_j \neq 0\}$$

and, if $\mathscr{P}$ is internally connected by chains with non-orthogonal successive elements, by

$$\text{(T 4)} \qquad \tfrac{1}{2} m K(\check{d})(\sum |\delta_i|)^2 (\sum \Lambda_i^{-1}) \quad \text{with} \quad \check{d} = \vee\{d_{ij} \mid F_i \not\perp F_j\}.$$

The bound (T 3) is finite iff the F_i with $\delta_i \neq 0$ are pairwise non-orthogonal and (T 4) is designed to replace it otherwise.

1. Introduction. Section 2 investigates some general properties of signed measures and their symmetrization w.r.t. general groups. Section 3 specializes to the permutation group notes a contraction effect of probability factors in product signed measures. The properties developed in Sections 2 and 3 will be used throughout the paper and, in particular, in Lemma 2, in the completion of the proof of Theorem 1, and in the proofs of Theorems 2, 3 and 4.

Section 4 proves Theorem 1, which is the special case of (T 3) for $m = 1$ and $\delta_1 = 1$. This is the main result of the paper. Its proof contains a detailed outline of itself including Lemmas 2, 3 and 4. An example, consisting of a simple special case, shows that the bound of Theorem 1 is sometimes asymptotically sharp to within a factor of $3.149 \cdots$.

Section 5 proves Theorems 2, 3, and 4, all as corollaries to Theorem 1. Theorem 2 is the special case of (T 3) for $m = 1$, Theorem 3 is (T 3) and Theorem 4 is (T 4).

Our main results are a strengthened generalization of Theorem II.1 of Hannan (1953). The latter is easily characterized in terms of the $m = 1$ case of (T 3),

$$\text{(T 2)} \qquad ||\tau^*||^2 \leq K(d_{01})\delta_1^2(1/\Lambda_0 + 1/\Lambda_1),$$

Received February 6, 1970.

[1] This research was partially supported by NSF Grant GP-14591.

[2] Now at the University of Guelph.

amounting to the assertion that, for $m = 1$ and fixed $F_0 \not\perp F_1$,

(T II. 1) $$||\tau^*||^2 \to 0 \quad \text{as} \quad \text{RHS}(\text{T}\,2) \to 0\,.$$

As partially indicated in Section 3 and Section 5, the derivation in Section 5 of (T 3) and (T 4) as corollaries of Theorem 2 equally well would yield weakened (as (T II. 1) weakens (T 2)) forms of (T 3) and (T 4) as corollaries of (T II. 1).

A corollary of the $\delta_1 = 1$ case of (T II. 1), the Lemma of Hannan and Robbins (1955), was there used to show (Theorem 5) that the difference between the simple and equivariant envelopes converges to zero. A generalization of the Lemma, Theorem 2 of Horn (1968), is shown in Section 3 to be improved by the $\delta_1 = 1$ case of a rather immediate corollary to (T II. 1). The special case of (T 3) with two non-zero δ_i is used in Hannan and Huang (1972) (Theorem 1) to bound a more general case of the difference. A similar application, Theorem 1 of Horn (1968), inherits the deficiencies of her Theorem 2.

2. Properties of symmetrization and of an $\mathscr{L}_1$-norm. Although in this paper we will only be concerned with the difference between two symmetrized probability measures, some of the properties used in our proofs hold true and are easier to prove in a more general context. This section investigates some of these properties of symmetrized signed measures and their $\mathscr{L}_1$-norms.

Let $(\mathscr{Y}, \mathscr{C})$ be a measurable space and $\mathscr{G}$ be a finite group of measurable transformations g on $(\mathscr{Y}, \mathscr{C})$. For a signed measure τ on $(\mathscr{Y}, \mathscr{C})$, we define τg^{-1} as the induced signed measure and τ^* as the symmetrization of τ by

$$(\tau g^{-1})C = \tau(g^{-1}(C)) \qquad C \in \mathscr{C}, \quad \tau^* = \text{AV}(\tau g)\,,$$

where AV denotes the average over $g \in \mathscr{G}$. Thus symmetrization (*, hereafter) is a linear operator.

For any real valued function f on $\mathscr{Y}$, define $f \circ g$ and f^* by

$$(f \circ g)y = f(gy)\,, \qquad f^* = \text{AV}(f \circ g)\,.$$

τ and f are said to be symmetric if $\tau = \tau^*$ and $f = f^*$, respectively. Since $\tau^* g \equiv \tau^*$, τ is symmetric iff $\tau \equiv \tau g$.

Throughout this paper we shall denote supremum and infimum by $\vee$ and $\wedge$, and express integrals of left operators by

$$\tau(f) = \int f(y)\, d\tau(y)\,.$$

For any signed measure τ, define an $\mathscr{L}_1$-norm of τ by

(1) $$||\tau|| = \vee\{|\tau(C)| : C \in \mathscr{C}\}\,.$$

It follows from the Jordan decomposition that

(2) $$||\tau|| = ||\tau^+|| \vee ||\tau^-||$$

and hence, if $\tau(1) = 0$,

(3) $$||\tau|| = ||\tau^+|| = ||\tau^-||\,.$$

In particular, if P and Q are probability measures, then

$$0 \leqq ||P - Q|| \leqq 1 , \tag{4}$$

with equality at 0 iff $P = Q$, and equality at 1 iff $P \perp Q$.

For use in the proof of Lemma 1, let μ be a measure such that $d\tau/d\mu$ exists. Since $d\tau^+/d\mu = (d\tau/d\mu)^+$ and $d\tau^-/d\mu = (d\tau/d\mu)^-$,

$$||\tau|| = \mu(d\tau/d\mu)^+ \vee \mu(d\tau/d\mu)^- .$$

Hence, if $\tau(1) = 0$,

$$2||\tau|| = \mu(|d\tau/d\mu|) . \tag{5}$$

It follows from the transformation theorem (Theorem 39. C, Halmos (1950)) that

$$(d\tau/d\mu) \circ \check{g} = d\tau g/d\mu g , \qquad g \in \mathscr{G} .$$

In particular, if $\mu = \mu^*$, AV of the above equality yields

$$(d\tau/d\mu)^* = d\tau^*/d\mu . \tag{6}$$

Let μ be a measure and let h and f be such that the products h^*f and hf^* are μ^*-integrable. By the transformation theorem and symmetry, $\mu^*(h^*f \circ g) = \mu^*(h^*f)$ for all $g \in \mathscr{G}$. Averaging over $\mathscr{G}$ and interchanging h and f yields

$$\mu^*(h^*f) = \mu^*(h^*f^*) = \mu^*(hf^*) . \tag{7}$$

For the special case of (7) obtained by letting $h = d\mu/d\mu^*$, we have $h^* = 1$ by (6) and therefore, if f (and hence also f^*) is μ^*-integrable, then

$$\mu^*(f) = \mu^*(f^*) = \mu(f^*) . \tag{8}$$

For use in the completion of proof of Theorem 1, we note that, by subadditivity of norm and by $||\tau g|| \equiv ||\tau||$ for a signed measure τ,

$$||\tau^*|| \leqq \text{AV}\,||\tau g|| = ||\tau|| . \tag{9}$$

It follows from norm subadditivity and the Schwarz inequality that if τ_i are signed measures then

$$||\sum \tau_i||^2 \leqq (\sum 1) \sum ||\tau_i||^2 . \tag{10}$$

If $\tau = \tau_1 \times \tau_2$ is a product signed measure, $||\tau||$ is simply related to the corresponding norms of its factors. We abbreviate by omission subscripts on the norms. Since $\tau^+ = (\tau_1^+ \times \tau_2^+) + (\tau_1^- \times \tau_2^-)$, $\tau^- = (\tau_1^+ \times \tau_2^-) + (\tau_1^- \times \tau_2^+)$, it follows easily that

$$||\tau_1|| \cdot ||\tau_2|| \leqq ||\tau_1 \times \tau_2|| \leqq 2||\tau_1|| \cdot ||\tau_2|| , \tag{11}$$

with the first equality iff either τ_1 or τ_2 is a measure or the negative of a measure, the second equality iff $\tau_1(1) = \tau_2(1) = 0$. In particular, if τ_2 is a probability measure, the first equality of (11) yields

$$||\tau_1 \times \tau_2|| = ||\tau_1|| . \tag{12}$$

To assist in comparing our results with Theorem II. 1 of Hannan (1953), consider a signed measure τ with $\tau(1) = 0$. By (3),

$$(13) \qquad ||\tau|| = ||\tau^+|| = \bigvee\{\tau(f) \mid 0 \leqq f \leqq 1\}.$$

Since $\tau^*(1) = 0$ by linearity of * operator, upon applying (13) to τ^*, and applying (8) to $\tau^*(f)$, it follows that

$$(14) \qquad ||\tau^*|| = \bigvee\{\tau(f^*) \mid 0 \leqq f \leqq 1\} = \bigvee\{\tau(f) \mid 0 \leqq f^* = f \leqq 1\},$$

with the second equality following from $\{f^* \mid 0 \leqq f \leqq 1\} = \{f \mid 0 \leqq f = f^* \leqq 1\}$.

3. Permutations; contraction effect of probability factors. Henceforth we specialize $\mathscr{G}$ to be the group of transformations on $(\mathscr{X}, \mathscr{B})^N$ induced by the group of permutations on N objects, where $(\mathscr{X}, \mathscr{B})$ is a measurable space. We also let $\mathscr{G}$ denote the permutation group itself. Thus a generic element $g \in \mathscr{G}$ will be used both as a permutation and the transformation $g\mathbf{x} = (x_{g1}, \cdots, x_{gN})$.

The following lemma will be used, in Section 5, for the extension, via appropriate triangulation, from Theorem 2 with 2 factors to Theorem 3 with $m + 1$ factors. Starting from Theorem II. 1 of Hannan (1953) rather than Theorem 2, Lemma 1 together with (10) would yield an extension paralleling our Theorem 3 extension from Theorem 2. It will be shown that Lemma 1 alone would yield an extension of Theorem II. 1 improving Theorem 2 of Horn (1968), where there is a stronger restriction of the F_i than mutual absolute continuity, and where the non-zero δ_i are 1 and -1 respectively.

LEMMA 1. *If $\tau = \check{\tau} \times P$ for a signed measure $\check{\tau}$ with $\check{\tau}(1) = 0$ and a probability measure P, then, abbreviating affixes on * and on $||\ ||$ by omission,*

$$||\tau^*|| \leqq ||(\check{\tau})^*||.$$

PROOF. Since $\tau(1) = 0$, $||\tau^*|| = \bigvee\{\check{\tau}(P(f)) \mid 0 \leqq f = f^* \leqq 1\}$ by (14). Since $\check{f} = P(f)$ is symmetric in the remaining variables and since $0 \leqq \check{f} \leqq 1$, one more application of (14) completes the proof.

For our extension of Theorem II. 1, let $\tau = \times F_i^{N_i} - \times F_i^{N_i'}$ where the F_i are pairwise non-orthogonal and fixed, and the $\delta_i = N_i' - N_i$ are zero except for two i's. By judicious choice of g, we have $\tau g = \check{\tau} \times P$ with $P = \times\{F_i^{N_i} \mid \delta_i = 0\}$. By $(\tau g)^* \equiv \tau^*$ and Lemma 1, $||\tau^*|| \leqq ||(\check{\tau})^*||$ which, by (14) and Theorem II. 1, converges to zero as this case of the bound in (T 2) does.

4. Two distinct factors with unit differences in multiplicities.

THEOREM 1. *Let F_1 and F_0 be p-measures and let $\mathbf{N} = (N_1, N_0)$ be an integral partition of N with $N_1 \geqq 0 < N_0$. With $d = ||F_1 - F_0||$, and $K(d) = .5012 \cdots \times d(1 - d)^{-\frac{3}{2}}$,*

$$(15) \qquad ||(F_1^{N_1} \times F_0^{N_0})^* - (F_1^{N_1+1} \times F_0^{N_0-1})^*||^2 \leqq K(d)\left(\frac{1}{N_1 + 1} + \frac{1}{N_0}\right).$$

PROOF. Since both sides of (15) vanish if $d = 0$, henceforth assume $F_1 \neq F_0$.

The proof proceeds according to the following outline: As in Hannan (1953), a parametric family of densities of the F_i is introduced and some of their moment properties are related to d. Starting as in Hannan (1953) but then weakening by the Schwarz inequality, Lemma 2 obtains a family of upper bounds for a slight generalization of LHS (15). Lemma 3 develops lower bounds for modal binomial probabilities for application to the denominators of a bound of Lemma 2. Lemma 4 bounds the second difference of generalized binomial probabilities for application to the rest of the bound of Lemma 2. The bound of (15) is then obtained as the minimum of d^2, from Lemma 1, and a bound resulting from the application of the other lemmas.

For $0 < p = 1 - q < 1$ and $i = 1, 0$, let

$$F_p = pF_1 + qF_0 , \qquad f_i = dF_i/dF_p , \tag{16}$$

$$\theta = F_1(f_0), P_i = pF_1(f_i) = 1 - Q_i(= 1 - qF_0(f_i)) , \tag{17}$$

and note

$$F_i(f_i) = F_p(f_i^2) > (F_p(f_i))^2 = 1 \tag{18}$$

by the Schwarz inequality, with equality eliminated by $F_p \neq F_i$. Thus

$$\theta = p^{-1}(1 - qF_0(f_0)) < 1 . \tag{19}$$

Let μ be any σ-finite measure dominating the F_i, let $h_i = dF_i/d\mu$ and let $h_p = ph_1 + qh_0$. Since $f_1 f_0 h_p = h_1 h_0 h_p^{-1} \geqq h_1 \wedge h_0$, it follows that

$$\theta \geqq \mu(h_1 \wedge h_0) = 1 - d . \tag{20}$$

Finally, note that from (17) and (18)

$$P_1 - P_0 = 1 - \theta , \qquad P_i Q_i = pqF_1(f_i)F_0(f_i) > pq\theta . \tag{21}$$

For integer N and $\mathbf{p} \in [0, 1]^N$, let $b(k; \mathbf{p})$ denote the generalized binomial probability of k successes in N independent trials with success probabilities $\mathbf{p}$.

At the cost of notational complications, the following lemma could be stated and proved for $m + 1$ instead of 2 probability measures.

LEMMA 2. *For nonnegative integral partitions of* N, $\mathbf{N} = (N_0, N_1)$ *and* $\mathbf{N}' = (N_0', N_1')$, *define* $\mathbf{F} = F_1^{N_1} \times F_0^{N_0}$, $\mathbf{F}' = F_1^{N_1'} \times F_0^{N_0'}$ *and* $\tau = \mathbf{F} - \mathbf{F}'$. *For* $0 < p < 1$

$$\mathbf{F}^* \left(\frac{d\mathbf{F}'^*}{dF_p^N}\right) = \frac{b(N_1; P_1^{N_1'} P_0^{N_0'})}{b(N_1; p^N)} = \frac{b(N_1'; P_1^{N_1} P_0^{N_0})}{b(N_1'; p^N)} , \tag{22}$$

$$4\|\tau^*\|^2 < \frac{b(k; P_1^r P_0^{N-r})}{b(k; p^N)} \Bigg]_{k=N_1}^{k=N_1'} \Bigg]_{r=N_1}^{r=N_1'} . \tag{23}$$

PROOF. Since $\mathbf{F}^*(d\mathbf{F}'^*/dF_p^N) = \mathbf{F}'^*(d\mathbf{F}^*/dF_p^N)$, the second equality in (22) will follow from the first.

By (6) the integrand in LHS (22) is $(d\mathbf{F}'/dF_p^N)^*$, whence, by (8), LHS (22) is AV of

$$\mathbf{F}\left(\frac{d\mathbf{F}'}{dF_p^N}\circ g\right). \tag{24}$$

Since the integrand in (24) is the product $\prod_1^{N_1'} f_1(x_{g\alpha}) \prod_{N_1'+1}^{N} f_0(x_{g\alpha})$ of $\mathbf{F}$-independent variables, the integral is expressible, in terms of $K = \#[\alpha > N_1' \mid g\alpha \leqq N_1]$, as the product

$$\begin{aligned}&[F_0(f_0)]^{N_0'-k}[F_1(f_0)]^{K}[F_0(f_1)]^{N_0-N_0'+K}[F_1(f_1)]^{N_1-K}\\ &\qquad = p^{-N_1}q^{-N_0}P_0^{\,K}Q_0^{\,N_0'-K}P_1^{\,N_1-K}Q_1^{\,N_1'-N_1+K} = H(K)\,.\end{aligned} \tag{25}$$

Then (22) follows since

$$\mathrm{AV}\,H(K(g)) = \sum_k \binom{N_0'}{k}\binom{N_1'}{N_1-k}\binom{N}{N_1}^{-1} H(k) = \frac{b(N_1;\, P_1^{N_1'}P_0^{N_0'})}{b(N_1;\, p^N)} \tag{26}$$

with the first equality following from the transformation theorem.

From (5) and the Schwarz inequality,

$$4\|\tau^*\|^2 = \left[F_p^N\left(\left|\frac{d\tau^*}{dF_p^N}\right|\right)\right]^2 < F_p^N\left(\left|\frac{d\tau^*}{dF_p^N}\right|^2\right) = \tau^*\left(\frac{d\tau^*}{dF_p^N}\right). \tag{27}$$

Four-fold application of (22) to this bound results in (23), completing the proof.

For integers $0 \leqq k < N$, let

$$a_{Nk} = \{(N-1)pq\}^{\frac{1}{2}} b(k;\, p^N) \quad \text{with} \quad p = (k+1)/(N+1)\,. \tag{28}$$

LEMMA 3. *$a_{NN-1-k} = a_{Nk} \geqq a_{N0} \overset{\uparrow}{\searrow} e^{-1}$ as $N \uparrow \infty$.*

PROOF. Note that $f(x) = (1 + x^{-1})^{x+\frac{1}{2}} \underset{\downarrow}{\searrow} e$ as $0 < x \uparrow \infty$ since, with $t = (2x+1)^{-1}$, $\log f(x) = 1 + t^2/3 + t^4/5 + \cdots$. The lemma then follows from the representations,

$$a_{Nk} = a_{Nk-1}\frac{f(k)}{f(N-k)} = a_{N0}\prod_1^k \frac{f(j)}{f(N-j)}\,, \qquad a_{N0} = \left\{\frac{N-1}{N+1}\right\}^{\frac{1}{2}}\frac{1}{f(N)}\,. \tag{29}$$

The following lemma uses characteristic function inversions to bound the second difference of generalized binomial probabilities. In our application to certain numerators of (23), the generalized binomials involve only two distinct probabilities but the added generality simplifies the proof and may serve to motivate other applications.

LEMMA 4. *Let $B(x) = (4/\pi)\int_0^1 y^{\frac{1}{2}}(1-y)^{-\frac{1}{2}}e^{-2xy}\,dy$ and $\sigma^2 = \sum_1^M P_\alpha(1-P_\alpha)$. For $\mathbf{P} \in [0,1]^M$ and integral k*

$$|\Delta^2 b(k;\, \mathbf{P})| \leqq B(\sigma^2)\,, \tag{30}$$

$$x^{\frac{3}{2}}B(x) \to (2\pi)^{-\frac{1}{2}} \qquad \text{as} \quad x \uparrow \infty\,, \tag{31}$$

$$s = \vee\{x^{\frac{3}{2}}B(x) \mid 0 \leqq x < \infty\} = .545447\cdots. \tag{32}$$

PROOF. Note for later use that, with $\phi(u) = Pe^{iu} + Q$ the ch.f. of a Bernoulli variable with parameter $P = 1 - Q$,

$$(33)\qquad |\phi(u)|^2 = 1 - 4PQ\sin^2 u/2 \leqq \exp - 4PQ\sin^2 u/2 .$$

Since $-\Delta^2 e^{-iu(k-1)} = e^{-iuk}4\sin^2 u/2$, it follows that $-\Delta^2 b(k-1; \mathbf{P})$ is given by the Fourier inversion

$$(34)\qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iuk}4\sin^2\frac{u}{2}\prod_1^M \phi_\alpha(u)\,du .$$

By application of (33) to the ϕ_α, the modulus of (34) is bounded by

$$(35)\qquad \frac{4}{\pi}\int_0^\pi \sin^2\frac{u}{2}\, e^{-2\sigma_*^2\sin^2 u/2}\,du = B(\sigma^2) .$$

(Since $v^{-1}\sin v \downarrow$ on $0 < v \leqq \pi/2$, RHS (33) $\leqq \exp - 4PQ(u/\pi)^2$ on $|u| \leqq \pi$. The corresponding weakening of (35) implies $s < \pi^{5/2}2^{-7/2} = 1.546\cdots$:

$$B(x) \leqq (4/\pi)\int_0^\pi \sin^2(u/2)e^{-2x(u/\pi)^2}\,du < (1/\pi)\int_0^\infty u^2 e^{-2x(u/\pi)^2}\,du = (\pi^{5/2}/2^{7/2})x^{-\frac{3}{2}} .)$$

With I_0, I_1 denoting the modified Bessel functions,

$$(36)\qquad B(x) = \left[-\frac{2}{\pi}\int_0^1 y^{-\frac{1}{2}}(1-y)^{-\frac{1}{2}}e^{-2xy}\,dy\right]'$$
$$= [-2e^{-x}I_0(x)]' = 2e^{-x}[I_0(x) - I_1(x)] ,$$

where the second equality follows by differentiation from the corresponding Laplace transform (cf. 29. 3. 124 of NBS-AMA55 (1964)), and the last from $I_0' = I_1$.

In view of (36), (31) is an immediate consequence of the usual asymptotic expansions of I_0 and I_1 (cf. 9. 7. 1 *ibid.*).

To verify (32), first note that $[x^{\frac{3}{2}}B(x)]' = x^{\frac{1}{2}}e^{-x}F(x)$ with

$$F = (3 - 4x)I_0 + (4x - 1)I_1 .$$

Since $F(1) > 0 > F(2)$, the behavior of F' will imply that $\exists x_0 \in (1, 2)$ with $F \underset{\downarrow}{\leq} 0$ on $[0, x_0]$ and $F < 0$ on (x_0, ∞) as follows. Since $I_1 \leqq xI_0$, $xF' = x(4x-5)I_0 + (1-x)(4x+1)I_1 \leqq -(15/16)xI_0$ on $[0, 1]$. Since $x(4x-1)F' = (4x+1)(1-x)F - 3I_0$, $1 \leqq x$ and $F(x) \geqq 0 \Rightarrow F'(x) < 0$ and hence $1 \leqq x_0$ and $F(x_0) = 0 \Rightarrow F < 0$ on (x_0, ∞).

Since, in addition, $.00012 > F(1.452) > 0 > F(1.453)$ (page 228 BAAS v VI (1937)),

$$0 \leqq s - (1.452)^{\frac{3}{2}}B(1.452) \leqq (1.452)^{\frac{1}{2}}e^{-1.452}(.00012)(.001) < 4(10^{-8})$$

which results in (32) and thus completes the proof.

COMPLETION OF PROOF OF THEOREM 1. From Lemma 2 with $N_1' = N_1 + 1$ we have this case of the bounds (23). Let $p = (N_1 + 1)/(N + 1)$ henceforth.

This choice insures both denominators in (this case of) (23) are equal and, by Lemma 3, are bounded below by $\{(N-1)pq\}^{-\frac{1}{2}}a_{N0}$. Use of the recursion relation resulting from the convolution interpretation of the generalized binomials shows that the mixed second difference remaining in the numerator is $-(P_1-P_0)\Delta^2 b(N_1-1;\, P_1^{N_1}P_0^{N_0-1})$. Application of Lemma 4 results in the upper bound $(P_1-P_0)B(\sigma^2)$, with $P_1-P_0=1-\theta\leqq d$ and $\sigma^2=N_1P_1Q_1+(N_0-1)P_0Q_0>(N-1)pq(1-d)$ by (20) and (21), so that, with $C(d)=d(1-d)^{-\frac{3}{2}}$,

$$(37)\qquad ||\tau^*||^2\leqq\frac{\{(N-1)pq\}^{\frac{1}{2}}}{4a_{N0}}\,dB(\sigma^2)\leqq\frac{N+1}{N-1}\,\frac{\sigma^3B(\sigma^2)}{4a_{N0}}\,C(d)\left(\frac{1}{N_1+1}+\frac{1}{N_0}\right).$$

With the s of (32) and $c_N=(N+1)/(N-1)(s/4a_{N0})$, (37) and Lemma 1 yield

$$(38)\qquad ||\tau^*||^2\leqq d^2\wedge c_NC(d)\left(\frac{1}{N_1+1}+\frac{1}{N_0}\right).$$

Since $c_N\downarrow$ w.r.t. N while

$$v_N=\mathrm{V}_{d,\mathbf{N}|N}\frac{d^2}{C(d)\left(\frac{1}{N_1+1}+\frac{1}{N_0}\right)}=\frac{.4(.6)^{\frac{3}{2}}}{N+1}\left[\frac{N+1}{2}\right]\left[\frac{N+2}{2}\right]\uparrow\text{ w.r.t. }N$$

and $v_9<c_{10}=.5012\cdots<v_{10}$, the bound of (15) is the best of its kind following from (38), and the proof of Theorem 1 is complete.

REMARK. As $\sigma^2\to\infty$ (which occurs if $(N_1\wedge N_0)(1-d)$ does), so does N so that, from (31), the RHS of (37) is then asymptotic to the bound of (15) with c_{10} improved to $e(32\pi)^{-\frac{1}{2}}=.2711\cdots$.

That the bound of Theorem 1 is sometimes relatively sharp will follow from examination of one of the simplest special cases.

EXAMPLE. Let F_i put mass ξ on i and $1-\xi$ on $\frac{1}{2}$ for $i=0,1$. For every triple of nonnegative integers $M=(M_0,M_{\frac{1}{2}},M_1)$ with sum N, $F_1^{N_1'}\times F_0^{N_0'}$ (and hence also its symmetrization) assigns mass $b(M_1;\,\xi^{N_1'})b(M_0;\,\xi^{N_0'})$ to the symmetric set of $\mathbf{x}$ in $\{0,\frac{1}{2},1\}^N$ with the frequencies M. For the τ of Theorem 1, a little algebra shows

$$(39)\qquad (1-\xi)\tau[M]=\left(\frac{M_0}{N_0}-\frac{M_1}{N_1+1}\right)b(M_0;\,\xi^{N_0})b(M_1;\,\xi^{N_1+1}).$$

Thus $2(1-\xi)||\tau^*||$ is seen to be $E|\hat{\xi}_0-\hat{\xi}_1|$ where the $\hat{\xi}_i$ are the independent relative binomials indicated in (39). As both $N_i\xi(1-\xi)\to\infty$,

$$X_{\mathbf{N}}=\{[N_0^{-1}+(N_1+1)^{-1}]\xi(1-\xi)\}^{-\frac{1}{2}}(\hat{\xi}_0-\hat{\xi}_1)$$

converges in distribution to the standard normal since both $\{N_i^{-1}\xi(1-\xi)\}^{-\frac{1}{2}}\times(\hat{\xi}_i-\xi)$ do. Since $EX_{\mathbf{N}}^2\equiv1$, it follows from the Corollary to the Moment

Convergence Theorem (page 184 Loève (1963)) that

$$E|X_{\mathbf{N}}| \to \int |z|(2\pi)^{-\frac{1}{2}}e^{-z^2/2}\,dz = (2/\pi)^{\frac{1}{2}}$$

and therefore

$$||\tau^*||^2 \sim \frac{1}{2\pi}\frac{\xi}{1-\xi}\left(\frac{1}{N_1+1}+\frac{1}{N_0}\right). \tag{40}$$

When $\xi \to 0$, the comparison is most favorable and the bound of Theorem 1 exceeds RHS (40) only by the factor $3.149\cdots$.

5. $m+1$ **Distinct factors.** We now obtain various extensions of Theorem 1 as corollaries to that theorem. Theorems 1, 2 and 3 represent successive extensions each subsuming, yet corollary to, its predecessor. Theorem 3 is useless unless the F_i are pairwise non-orthogonal, and Theorem 4 is designed to replace it in this case. Thus, as implied in the summary, our final results are merely Theorems 3 and 4. Let

$$\begin{aligned} &\tau = \times_i F_i^{N_i} - \times_i F_i^{N_i'}, \qquad d_{ij} = ||F_i - F_j|| \\ &\delta_i = N_i' - N_i, \qquad \wedge_i = (N_i' \wedge N_i) + 1, \qquad i, j = 0, \cdots, m. \end{aligned} \tag{41}$$

THEOREM 2. *For* $m = 1$,

$$||\tau^*||^2 \leqq K(d_{01})\delta_1^2\left(\frac{1}{\wedge_0}+\frac{1}{\wedge_1}\right). \tag{42}$$

PROOF. Assume without loss of generality $\delta_1 \geqq 1$ (we may rename $\mathbf{N}$ and $\mathbf{N}'$ otherwise), and for $j = 1, \cdots, \delta_1$, let

$$\tau_j = F_0^{N_0-j+1} \times F_1^{N_1+j-1} - F_0^{N_0-j} \times F_1^{N_1+j}. \tag{43}$$

Since $\tau = \sum \tau_j$, it follows from linearity of * and (10) that

$$||\tau^*||^2 \leqq \delta_1 \sum ||\tau_j^*||^2. \tag{44}$$

Applying Theorem 1 to each summand in RHS (44) completes the proof.

THEOREM 3. *Let* $n = \#\{i \mid \delta_i \neq 0\} - 1$ *be positive* (*otherwise* $\tau^* = 0$), *and let* $d = \vee\{d_{ij} \mid \delta_i \neq 0, \delta_j \neq 0\}$. *Then*

$$||\tau^*||^2 \leqq nK(d) \sum \delta_i^2 \wedge_i^{-1}.$$

PROOF. Given $\mathbf{N}$ and $\mathbf{N}'$ we construct a sequence of partitions $\mathbf{N} = \mathbf{N}_0, \mathbf{N}_1 \cdots$, $\mathbf{N}_r = \mathbf{N}'$, for some $r \leqq n$, as follows. To construct $\mathbf{N}_1 = (N_{10}, \cdots, N_{1m})$, let s be such that $|\delta_s| = \wedge\{|\delta_j| \mid \delta_j \neq 0\}$, and let t be such that δ_t has opposite sign from δ_s. Let $N_{1s} = N_s + \delta_s$, $N_{1t} = N_t - \delta_s$ and $N_{1j} = N_j$ for the other j's. Thus $\mathbf{N}_1$ stays between $\mathbf{N}$ and $\mathbf{N}'$ coordinatewise and differs from $\mathbf{N}$ in two coordinates. Repeating this construction, the process terminates in $r \leqq n$ steps since each successive $\mathbf{N}_i$ identifies at least one more coordinate with $\mathbf{N}'$. Define

$$\tau_i = \times_{j=0}^m F_j^{N_{i-1j}} - \times_{j=0}^m F_j^{N_{ij}}, \qquad i = 1, \cdots, r.$$

Since $\mathbf{N}_i$ differs from $\mathbf{N}_{i-1}$ in two coordinates, say s and t, the fact that $(\tau_i g)^* = (\tau_i)^*$ enables us to use Lemma 1 to obtain

(45) $$||\tau_i^*|| \leqq ||(F_s^{N_{i-1s}} \times F_t^{N_{i-1t}})^* - (F_s^{N_{is}} \times F_t^{N_{it}})^*|| \,.$$

Applying Theorem 2 to RHS (45), we note that, since each $\mathbf{N}_i$ stays between $\mathbf{N}_{i-1}$ and $\mathbf{N}_{i+1}$ coordinatewise, the denominators in this application of RHS (42) are bounded below by Λ_s and Λ_t respectively. Since K is increasing, we furthei weaken this application of the bound (42) by replacing d_{st} by d. Thus

(46) $$[\text{RHS}(45)]^2 \leqq K(d)(N_{i-1s} - N_{is})^2\left(\frac{1}{\Lambda_s} + \frac{1}{\Lambda_t}\right).$$

Since $N_{i-1j} = N_{ij}$ except for $j = s, t$, RHS (46) is

$$K(d) \textstyle\sum_{j=0}^{m} \Lambda_j^{-1}(N_{i-1j} - N_{ij})^2 \,.$$

Since $\tau = \sum_{i=1}^{r} \tau_i$, by (10) and the above representation of RHS (46),

(47) $$||\tau^*||^2 \leqq r \textstyle\sum ||\tau_i^*||^2 \leqq rK(d) \sum_i \Lambda_j^{-1} \sum_i (N_{i-1j} - N_{ij})^2 \,.$$

Since $\sum_i (N_{i-1j} - N_{ij}) = \delta_j$ and since $N_{i-1j} - N_{ij}$ are of the same sign for fixed j, the summation w.r.t. i in the last term of (47) is bounded by δ_j^2. Since $r \leqq n$ the proof is complete.

THEOREM 4. *Let $F_0, F_1, \cdots, F_m$ be internally connected by chains with successive elements non-orthogonal and let $\check{d} = \vee\{d_{ij} \mid F_i \not\perp F_j\}$. Then*

$$||\tau^*||^2 \leqq \tfrac{1}{2}mK(\check{d})(\textstyle\sum_i |\delta_i|)^2 \sum_i \Lambda_i^{-1} \,.$$

PROOF. For any connected graph of finitely many vertices there exists a vertex whose removal leaves the remaining graph connected. We shall rename $F_0, \cdots, F_m$ in such a way that successive removal of $F_0, F_1, \cdots$ leaves the remaining connected. For each $i \leqq m - 1$, let $t(i)$ be such that $t(i) > i$ and $F_{t(i)} \not\perp F_i$.

Given $\mathbf{N}$ and $\mathbf{N}'$ we consider the partition which differs from either $\mathbf{N}$ (if $\delta_0 \leqq 0$) or $\mathbf{N}'$ (if $\delta_0 > 0$) only on the 0th and the $t(0)$th coordinates, where the 0th coordinate is $N_0 \wedge N_0'$ and the $t(0)$th coordinate, compared to that of $\mathbf{N}$ or $\mathbf{N}'$, is increased by $|\delta_0|$. By weakening Theorem 3 on both K and the second denominator (1 + the $t(0)$th coordinate of the new partition $\geqq \Lambda_{t(0)}$), we see that the square norm of * of the difference between the product measures associated with the two partitions is bounded by

(48) $$K(\check{d})\delta_0^2\left(\frac{1}{\Lambda_0} + \frac{1}{\Lambda_{t(0)}}\right).$$

Iterate the process. Letting $\delta_j^{(i)}$ be the difference in the jth coordinate at the ith iteration of this process, we see that $\delta_j^{(i)} = 0$ for $j < i$ and

(49) $$\delta_j^{(i)} = \delta_j + \textstyle\sum \{\delta_r^{(r)} \mid 0 \leqq r < i,\ t(r) = j\}$$

for $j \geqq i$. We also note that the $\delta_r^{(r)}$ above are disjoint sums of δ's from $\{\delta_0, \cdots, \delta_{i-1}\}$.

Since $\Lambda_j^{(i)} - 1$, the minimum of the jth coordinates for the two partitions at ith iteration, is increasing w.r.t. i, the bound corresponding to (48), further weakened by $\Lambda_i^{(i)} \geqq \Lambda_i$, is

$$K(\check{d})[\delta_i^{(i)}]^2\left(\frac{1}{\Lambda_i} + \frac{1}{\Lambda_{t(i)}}\right). \tag{50}$$

Since each iteration results in reducing one coordinate difference to zero, the process terminates in m steps. By (10) as in the proof of Theorem 3, we see that, by (50),

$$||\tau^*||^2 \leqq mK(\check{d}) \sum_{i=0}^{m-1} \delta_i^{(i)2}\left(\frac{1}{\Lambda_i} + \frac{1}{\Lambda_{t(i)}}\right). \tag{51}$$

The coefficient of $1/\Lambda_i$ in the summation above is, with $\delta_m^{(m)} = 0$,

$$\delta_i^{(i)2} + \sum_r \{\delta_r^{(r)2} \mid t(r) = i\}. \tag{52}$$

By (49) and the note following it, complementing in $\delta_i^{(i)}$ by $\sum \delta_i = 0$, we see that $\delta_i^{(i)}$ and the $\delta_r^{(r)}$ in (52) are disjoint sums of $\{\delta_0, \cdots, \delta_{i-1}, \delta_{i+1}, \cdots, \delta_m\}$. Thus the maximum of (52) over all $\boldsymbol{\delta}$ is $\leqq (\sum \delta^+ - \delta_i^+)^2 + (\sum \delta^- - \delta_i^-)^2 \leqq \frac{1}{2}(\sum |\delta|)^2$, and the proof is complete.

The hypothesis of Theorem 4 fails to hold if and only if $\mathscr{P}$ is disconnected. Then either (i) there exists a component, i.e., a maximal connected set of factors, whose **N**-multiplicity differs from its **N**$'$-multiplicity, in which case it follows easily that $||\tau^*|| = 1$, or (ii) every component has identical **N**- and **N**$'$-multiplicity, in which case $||\tau^*||$ is simply related to the $||(\tau_c)^*||$ corresponding to the separate components:

$$\bigvee_c ||(\tau_c)^*|| \leqq ||\tau^*|| \leqq \sum_c ||(\tau_c)^*||, \tag{53}$$

where the second inequality follows from Lemma 1 and the triangle inequality.

Addendum. After this paper was submitted, Horn and Schach (1970) obtained qualitative results of a more general character as a corollary to their 0–1 law (a product probability measure with each factor recurring i.o. takes only the values 0 or 1 on sets invariant under all finite permutations): if λ and ν are probability measures such that $\tau = \lambda - \nu$ is dominated by a recurring $\mu = \times_1^\infty \mu_i$, then, reinterpreting our choice of $\mathscr{G}$ in Section 3 as transformations on $(\mathscr{X}, \mathscr{B})^\infty$,

$$||\tau^*|| \downarrow 0 \qquad \text{as } N \uparrow \infty. \tag{54}$$

REMARK 1. This result is self-strengthening in that the conclusion (54) continues to hold if only there is a finite permutation of ν, say ν', such that $\lambda - \nu'$ is dominated by a recurring μ.

REMARK 2. With the same reinterpretation of $\mathscr{G}$, our proof of Lemma 1 continues to apply provided that the domain of $\check{\tau} \subseteq \mathscr{B}^N$ and shows that equality here implies equality in the lemma. Thus, recognizing the "τ" in Theorems 3 and 4 as present $\check{\tau}$, the bounds of these theorems apply to $||\tau^*||$ with $\tau = \check{\tau} \times P$, whatever be the probability measure P on $\mathscr{B}^\infty$.

REMARK 3. When λ and ν are product probability measures with factors in a fixed set $\mathscr{F} = \{F_0, F_1, \cdots, F_m\}$, then it follows easily by a Kakutani (1948) criterion that $||\tau^*|| \equiv 1$ unless

(55) $$\lambda_i = \nu_i \quad \text{except} \quad \text{f.o.}$$

When (55) obtains, the hypothesis of the (strengthened) Horn-Schach result holds iff

(56) there exists ν' such that each $\lambda_i + \nu_i'$ is dominated by a recurrent factor of λ.

If, for example, $F_i \ll F_j$ only if $i = j$, then (55) and (56) will hold only if $\tau^* = 0$ except f.o. although the bounds of Theorems 3 and 4 will converge to 0 much more generally.

If, on the other hand, $F_0 \gg \{F_1, \cdots, F_m\}$, then (55) and (56) may hold without any of the $\Lambda_1 \cdots \Lambda_m$ going to ∞ with N, so that all these terms of the bounds of Theorems 3 and 4 fail to converge to 0.

In the presence of special domination assumptions, some improvements in our bounds are readily obtainable. For example, the proof of Theorem 1 used $P_1 - P_0 \leqq d$. Since $P_1 = F_1(ph_1h_p^{-1}) \downarrow F_1[h_0 = 0]$ as $p \downarrow 0$ the bound of Theorem 1, strengthened by the insertion of $(P_1 - P_0)/d$ in RHS(15), is $o(1)$ if $F_0 \gg F_1$ and only $N_0 \uparrow \infty$.

REFERENCES

BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (1937). *Mathematical Tables* **6** *Bessel Functions, Part 1, Functions of Order Zero and Unity*. University Press. Cambridge.

HALMOS, PAUL R. (1950). *Measure Theory*. Van Nostrand, Princeton.

HANNAN, JAMES F. (1953). Asymptotic solutions of compound decision problems. UNC Institute of Statistics Mimeo Series No. 68.

HANNAN, JAMES and HUANG, J. S. (1972). Equivariant procedures in the compound decision problem with finite state component problem. *Ann. Math. Statist.* **43** 102–112.

HANNAN, JAMES F. and ROBBINS, HERBERT (1955). Asymptotic solutions of the compound decision problem for two completely specified distributions. *Ann. Math. Statist.* **26** 37–51.

HORN, SUSAN DADAKIS (1968). The optimality criterion for compound decision problems. Technical Report No. 10, Department of Statistics, Stanford Univ.

HORN, SUSAN and SCHACH, SIEGFRIED (1970). An extension of the Hewitt-Savage zero-one law. *Ann. Math. Statist.* **41** 2130–2131.

KAKUTANI, S. (1948). Equivalence of infinite product measures. *Ann. Math.* **49** 214–224.

LOÈVE, MICHEL (1963) *Probability Theory*, (3rd ed.). Van Nostrand, Princeton.

NATIONAL BUREAU OF STANDARDS APPLIED MATHEMATICS SERIES 55 (1964). *Handbook of Mathematical Functions*, U.S. Government Printing Office, Washington, D.C.