ASYMPTOTIC SOLUTIONS OF THE COMPOUND DECISION PROBLEM FOR TWO COMPLETELY SPECIFIED DISTRIBUTIONS[1]

By James F. Hannan and Herbert Robbins

Michigan State College and *Columbia University*

1. Summary. A compound decision problem consists of the simultaneous consideration of n decision problems having identical formal structure. Decision functions are allowed to depend on the data from all n components. The risk is taken to be the average of the resulting risks in the component problems. A heuristic argument for the existence of good asymptotic solutions was given by Robbins ([1] Sec. 6) and was preceded by an example (component decisions between $N(-1,1)$ and $N(1,1)$) exhibiting, for sufficiently large n, a decision function whose risk was uniformly close to the envelope risk function of "simple" decision functions.

The present paper considers the class of problems where the components involve decision between any two completely specified distributions, with the risk taken to be the weighted probability of wrong decision. For all sufficiently large n, decision functions are found whose risks are uniformly close to the envelope risk function of "invariant" decision functions.

2. Statement and reduction of the problem. The problem of testing a simple statistical hypothesis against a simple alternative can be formulated as follows. Let x be a random variable (of arbitrary dimensionality) which is known to have one of the two distinct distribution functions $F(x, \theta)$ for $\theta = 0$ or 1. On the basis of a single observation on x (we consider only the nonsequential case) it is require to decide whether the true value of the unknown parameter θ is 0 or 1.

Statistical decision problems of the same formal structure often occur, or can be considered, in large groups. We shall, therefore, take as a single entity the following compound decision problem. Let n be a fixed positive integer and let $x_1, \cdots, x_n$ be independent random variables, each of which has the distribution function $F(x, \theta)$ with respective parameter values $\theta_1, \cdots, \theta_n$, with $\theta_i = 0$ or 1. Let $\mathbf{x} = (x_1, \cdots, x_n)$ denote the vector of observations and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$ the unknown vector of parameters; $\boldsymbol{\theta}$ is known to belong to the set Ω consisting of all 2^n possible vectors of n components, each 0 or 1. On the basis of $\mathbf{x}$ it is required to decide the true value of $\boldsymbol{\theta}$, which amounts to deciding for every $i = 1, \cdots, n$ whether $\theta_i = 0$ or 1.

Any vector of n functions $\mathbf{t} = (t_1(\mathbf{x}), \cdots, t_n(\mathbf{x})$ is a (randomized) decision function for the compound decision problem if for $i = 1, \cdots, n$, $0 \leqq t_i(\mathbf{x}) \leqq 1$, and if the conditional probabilities, given $\mathbf{x}$, of deciding that $\theta_i = 0$

Received July 20, 1954.

[1] This research was completed at the University of North Carolina and was supported in part by the United States Air Force under Contract AF 18(600)-83 monitored by the Office of Scientific Research.

or 1 are respectively $1 - t_i(\mathbf{x})$ and $t_i(\mathbf{x})$. If for some function $t(x)$, $t_i(\mathbf{x}) = t(x_i)$ for $i = 1, \cdots, n$, then $\mathbf{t}$ will be called *simple* and will be denoted by t.

We assume the practical background of the problem provides two positive constants,

a, the loss incurred in deciding "$\theta_i = 0$" when the true value of $\theta_i = 1$,
b, the loss incurred in deciding "$\theta_i = 1$" when the true value of $\theta_i = 0$.

For any Borel set B we define

$$\mu_\theta(B) = \int_B dF(x, \theta), \qquad \theta = 0, 1; \tag{2.1}$$

$$\mu(B) = \mu_0(B) + \mu_1(B). \tag{2.2}$$

Since μ_0 and μ_1 are absolutely continuous with respect to the finite measure μ, generalized probability density functions $f(x, 0)$ and $f(x, 1)$ exist such that for any Borel set B

$$\mu_\theta(B) = \int_B f(x, \theta)\, d\mu, \qquad \text{for } \theta = 0, 1.$$

We note that the relation

$$f(x, 0) + f(x, 1) = 1, \qquad \text{a.e. } (\mu), \tag{2.3}$$

is obtainable from the identity in Borel B,

$$\int_B 1\, d\mu = \mu(B) = \mu_0(B) + \mu_1(B) = \int_B (f(x, 0) + f(x, 1))\, d\mu.$$

The joint generalized probability density function of $\mathbf{x}$ with respect to the product measure μ^n, when the parameter vector is $\boldsymbol{\theta} = (\theta_1, \cdots \theta_n)$, is $f(\mathbf{x}, \boldsymbol{\theta}) = \prod_1^n f(x_i, \theta_i)$. The expected loss on the ith decision in using a decision function $\mathbf{t} = (t_1(\mathbf{x}), \cdots, t_n(\mathbf{x}))$ is, for $i = 1, \cdots, n$,

$$R_i(\mathbf{t}, \boldsymbol{\theta}) = \int [a\theta_i(1 - t_i(\mathbf{x})) + b(1 - \theta_i)t_i(\mathbf{x})] f(\mathbf{x}, \boldsymbol{\theta})\, d\mu^n.$$

The average expected loss on all n decisions, which we define to be the risk of $\mathbf{t}$, is therefore

$$R(\mathbf{t}, \boldsymbol{\theta}) = \frac{1}{n}\sum_1^n R_i(\mathbf{t}, \boldsymbol{\theta}) = \int \frac{1}{n}\sum_1^n \{a\theta_i[1 - t_i(\mathbf{x})] + b(1 - \theta_i)t_i(\mathbf{x})\} f(\mathbf{x}, \boldsymbol{\theta})\, d\mu^n. \tag{2.4}$$

This is equivalent to defining the loss of the decision $\mathbf{d} = (d_1, \cdots, d_n)$ in Ω, given $\boldsymbol{\theta}$ in Ω, to be

$$W(\mathbf{d}, \boldsymbol{\theta}) = \frac{1}{n}\sum_1^n [a\theta_i(1 - d_i) + b(1 - \theta_i)\, d_i],$$

since this definition implies that the decision function $\mathbf{t}$ induces the conditional (for fixed $\mathbf{x}$) expected loss

$$W(\mathbf{t}(\mathbf{x}), \boldsymbol{\theta}) = \frac{1}{n}\sum_1^n \{a\theta_i[1 - t_i(\mathbf{x})] + b(1 - \theta_i)t_i(\mathbf{x})\}. \tag{2.5}$$

By the remark (2.3), $f(\mathbf{x}, \boldsymbol{\theta})$ is expressible as

$$f(\mathbf{x}, \boldsymbol{\theta}) = \prod_1^n f(x_i, \theta_i) = \prod_1^n [\theta_i f(x_i, 1) + (1 - \theta_i)(1 - f(x_i, 1))]. \tag{2.6}$$

The Halmos-Savage [2] form of the theorem linking sufficient statistics and density factorization shows that $(f(x_1, 1), \cdots, f(x_n, 1))$ is a sufficient statistic for $\boldsymbol{\theta}$ in Ω. Let

$$\begin{cases} \mathbf{z} = (z_1, \cdots, z_n), & z_i = f(x_i, 1); \\ \nu_\theta(I) = \mu_\theta(x \mid f(x, 1) \text{ in } I), \quad \text{for all Borel sets } I; & \theta = 0, 1; \\ \nu(I) = \nu_0(I) + \nu_1(I). \end{cases} \tag{2.7}$$

It is easily verified (see [3] Sec. 32) that the sufficient statistic $\mathbf{z}$ (and, consequently, the measures ν_0, ν_1, and ν) is independent of the choice (2.2) of an underlying measure μ relative to which μ_0 and μ_1 are absolutely continuous. We note that $\mathbf{z}$ has, with respect to the product measure ν^n, the generalized probability density

$$d(\mathbf{z}, \boldsymbol{\theta}) = \prod_1^n [\theta_i z_i + (1 - \theta_i)(1 - z_i)]. \tag{2.8}$$

Returning to (2.4) we see that the vector of conditional expectations

$$E[\mathbf{t}(\mathbf{x}) \mid \mathbf{z}] = (E[t_1(\mathbf{x}) \mid \mathbf{z}], \cdots, E[t_n(\mathbf{x}) \mid \mathbf{z}])$$

is a decision rule having the same risk as $\mathbf{t}$. We denote this rule by $\mathbf{t}(\mathbf{z}) = (t_1(\mathbf{z}), \cdots, t_n(\mathbf{z}))$ and, using (2.5) express its risk,

$$\begin{aligned} R(\mathbf{t}, \boldsymbol{\theta}) &= \int W(\mathbf{t}(\mathbf{z}), \boldsymbol{\theta})\, d(\mathbf{z}, \boldsymbol{\theta})\, d\nu^n \\ &= \frac{a}{n} \sum_1^n \theta_i \int (1 - t_i(\mathbf{z}))\, d(\mathbf{z}, \boldsymbol{\theta})\, d\nu^n + \frac{b}{n} \sum_1^n (1 - \theta_i) \\ &\qquad \cdot \int t_i(\mathbf{z})\, d(\mathbf{z}, \boldsymbol{\theta})\, d\nu^n. \end{aligned} \tag{2.9}$$

3. Simple decision functions. If $\mathbf{t} = t$ is a simple decision function, then (2.9) simplifies to

$$R(t, \boldsymbol{\theta}) = \frac{a}{n} \sum_1^n \theta_i \int (1 - t(z))z\, d\nu + \frac{b}{n} \sum_1^n (1 - \theta_i) \int t(z)(1 - z)\, d\nu. \tag{3.1}$$

Setting $\bar{\theta} = (1/n)\sum \theta_i$ = proportion of 1's among the n components of θ, we can write (3.1) in the form

$$R(t, \boldsymbol{\theta}) = \int a\bar{\theta}(1 - t(z))z + b(1 - \bar{\theta})t(z)(1 - z)\, d\nu. \tag{3.2}$$

The value $\bar{\theta}$ is necessarily rational, of the form k/n for $k = 0, \cdots, n$. We now define for any real number p such that $0 \leqq p \leqq 1$, and any ν-measurable function $t(z)$ such that $0 \leqq t(z) \leqq 1$, the expression

$$B(t, p) = \int ap(1 - t(z))z + b(1 - p)t(z)(1 - z)\, d\nu. \tag{3.3}$$

Thus (3.2) becomes

$$R(t, \boldsymbol{\theta}) = B(t, \bar{\theta}). \tag{3.4}$$

From (3.3) it is clear that for any fixed p, $B(t, p)$ is a minimum if and only if for almost every $(\nu)z$, $t(z)$ is of the form

$$t_p(z) = \begin{cases} 1, & \text{if } apz > b(1 - p)(1 - z); \\ 0, & \text{if } apz < b(1 - p)(1 - z); \\ \text{arbitrary}, & \text{if } apz = b(1 - p)(1 - z). \end{cases} \tag{3.5}$$

The minimum value of $B(t, p)$ is

$$\begin{aligned} \phi(p) = B(t_p, p) &= \int \min\, [apz, b(1 - p)(1 - z)]\, d\nu \\ &= ap \int_0^{C(p)} d\nu_1 + b(1 - p)\left(1 - \int_0^{C(p)} d\nu_0\right) \end{aligned} \tag{3.6}$$

where $C(p) = b(1 - p) / [ap + b(1 - p)]$ for $0 \leqq p \leqq 1$. It follows from (3.4) that, for any simple decision function t and any $\boldsymbol{\theta}$ in Ω,

$$R(t, \boldsymbol{\theta}) \geqq \phi(\bar{\theta}), \tag{3.7}$$

equality holding if and only if $t(z)$ is of the form (3.5) with $p = \bar{\theta}$.

We shall now establish some properties of the function $\phi(p)$ defined for $0 \leqq p \leqq 1$ by (3.6). For $0 \leqq p_1 < p_2 \leqq 1$, and $0 < s < 1$,

$$\begin{aligned} s \min[ap_1z,\ b(1 - p_1)(1 - z)] &+ (1 - s) \min[ap_2z,\ b(1 - p_2)(1 - z)] \\ &= \min[asp_1z,\ bs(1 - p_1)(1 - z)] \\ &\qquad + \min[a(1 - s)p_2z,\ b(1 - s)(1 - p_2)(1 - z)] \\ &\leqq \min[a\{sp_1 + (1 - s)p_2\}z,\ b\{1 - sp_1 - (1 - s)p_2\}(1 - z)], \end{aligned}$$

with strict inequality if and only if $p_1 < C(z) < p_2$ (or alternatively $C(p_2) < z < C(p_1)$). Integrating with respect to ν we obtain from (3.6) that ϕ is a concave function of p,

$$s\phi(p_1) + (1 - s)\, \phi(p_2) \leqq \phi(sp_1 + (1 - s)p_2), \tag{3.8}$$

with equality if and only if $\nu[z \mid C(p_2) < z < C(p_1)] = 0$. Since $\phi(0) = \phi(1) = 0$, this implies

$$\phi(p) > 0, \qquad 0 < p < 1, \qquad \text{unless } \nu[z \mid 0 < z < 1] = 0. \tag{3.9}$$

The exception here is the trivial case where ν_0 and ν_1 (and hence μ_0 and μ_1) are disjoint measures.

The continuity of $\phi(p)$ can be established in the following form. If $0 \leqq p < p' \leqq 1$, then $0 \leqq C(p') < C(p) \leqq 1$ and

$$\min[azp, \ \ b(1-z)(1-p)] - \min[azp', \ \ b(1-z)(1-p')]$$
$$= \begin{cases} -az(p'-p), & 0 \leqq z < C(p'); \\ b(1-z)(1-p) - azp', & C(p') \leqq z \leqq C(p); \\ b(1-z)(p'-p), & C(p) < z \leqq 1. \end{cases}$$

Hence it follows that, for $0 \leqq z \leqq 1$,

$$-az(p'-p) \leqq \min[azp, \ \ b(1-z)(1-p)] - \min[azp', \ \ b(1-z)(1-p')] \leqq b(1-z)(p'-p).$$

Integrating with respect to ν, we obtain

$$-a(p'-p) \leqq \phi(p) - \phi(p') \leqq b(p'-p), \qquad p' > p. \tag{3.10}$$

Interchanging p and p', multiplying by -1 achieves

$$-b(p-p') \leqq \phi(p) - \phi(p') \leqq a(p-p'), \qquad p' < p. \tag{3.11}$$

From (3.3) we have, for any $0 \leqq p, p' \leqq 1$,

$$\begin{aligned} B(t_p, p') - \phi(p) &= \int (p'-p)\{a(1-t_p(z))z - bt_p(z)(1-z)\}\, d\nu \\ &= (p'-p) \int \{a(1-t_p(z))z - bt_p(z)(1-z)\}\, d\nu, \\ &\leqq \begin{cases} (p'-p)a, & p' > p; \\ (p-p')b, & p' < p. \end{cases} \end{aligned} \tag{3.12}$$

The last three inequalities, (3.10) to (3.12), and the definition of ϕ imply

$$0 \leqq B(t_p, p') - \phi(p') \leqq (a+b) \mid p - p' \mid.$$

From (3.4) it follows that for any $0 \leqq p \leqq 1$ and any $\boldsymbol{\theta}$ in Ω,

$$0 \leqq R(t_p, \boldsymbol{\theta}) - \phi(\bar{\theta}) \leqq (a+b) \mid p - \bar{\theta} \mid.$$

Thus, we have proved

THEOREM 1. *Suppose that in some manner an approximate value p of the true proportion $\bar{\theta}$ is known. Then the statistician who uses the simple decision function $\mathbf{t} = t_p$ will achieve a risk $R(t_p, \boldsymbol{\theta})$ which is within $(a+b) \mid p - \bar{\theta} \mid$ of the minimum attainable risk $\phi(\bar{\theta})$ in the class of all simple decision functions.*

4. "Consistent" estimation of a proportion. We use (without loss of generality) the canonical form of the problem (see (2.6) to (2.9)), assuming only that the original measures, μ_0 and μ_1, are not identical. Thus we are concerned with the random vector $\mathbf{z} = (z_1, \cdots, z_n)$ having the density $d(\mathbf{z}, \boldsymbol{\theta})$ defined by (2.8). We consider the problem of finding an estimator $p_n(\mathbf{z})$ for the proportion $\bar{\theta}$ of 1's in the first n coordinates of $\boldsymbol{\theta}$.

Since our principal interest is in the asymptotic estimation problem, we will consider the sequence of problems defined for $n = 1, 2, \cdots$ as embedded in the probability space of infinite sequences $\mathbf{z} = (z_1, z_2, \cdots)$ with p-measure induced by the density $d(\mathbf{z}, \boldsymbol{\theta})$ defined on the first n coordinates of $\mathbf{z}$ and $\boldsymbol{\theta}$, respectively. We emphasize this aspect in this and the following section by referring to $\boldsymbol{\theta}$ in Ω_∞.

Avoiding a discussion of "optimum" estimation, we devote this section to the consideration of a subclass H of the class U of all unbiased estimators of $\bar{\theta}$. This subclass H is to be the class of all estimators of the form

$$\bar{h}(\mathbf{z}) = \frac{1}{n}\sum_{1}^{n} h(z_i), \tag{4.1}$$

where $h(z)$ is an unbiased estimator of θ.

As a measure of the risk of an estimator in H we use its variance,

$$n \operatorname{Var} \bar{h}(\mathbf{z}) = \bar{\theta}\, V_1(h) + (1 - \bar{\theta})\, V_0(h), \tag{4.2}$$

where $V_1(h) = \int (h(z) - 1)^2\, d\nu_1$ and $V_0(h) = \int (h(z))^2\, d\nu_0$. We single out an interesting subclass of H by investigating the existence and representation of elements h minimizing, for fixed p, with $0 < p < 1$,

$$pV_1(h) + (1 - p)V_0(h). \tag{4.3}$$

For any pair of real numbers $\lambda = (\lambda_0, \lambda_1)$ we define an extension of (4.3) to the set of all $g(z)$ such that $\int |g(z)|\, d\nu < \infty$:

$$\begin{aligned} F_\lambda(g) = p\left[\int g^2\, d\nu_1 - 1\right] + (1 - p)\int g^2\, d\nu_0 - 2\lambda_1\left[\int g\, d\nu_1 - 1\right] \\ -2\lambda_0 \int g\, d\nu_0 . \end{aligned} \tag{4.4}$$

Using the density representation of these integrals and the restriction on the domain of g, we have

$$\begin{aligned} F_\lambda(g) = \int g^2(z)[pz + (1 - p)(1 - z)] - 2g(z)[\lambda_1 z + \lambda_0(1 - z)] \\ \cdot d\nu - p + 2\lambda_1 . \end{aligned} \tag{4.5}$$

For fixed λ_0 and λ_1 the integrand here is a minimum, for each fixed z, if and only if

$$g(z) = g_p(z \mid \lambda) = \frac{\lambda_1 z + \lambda_0(1 - z)}{pz + (1 - p)(1 - z)}, \qquad \text{a. e. } (\nu). \tag{4.6}$$

Since $g_p(z \mid \lambda) \leqq \max [|\lambda_1|, |\lambda_0|] / \min [p, 1 - p]$, we find that $g_p(z \mid \lambda)$ is a unique (ν) minimum of $F_\lambda(g)$ over its domain. If there exists a determination of λ_0 and λ_1 such that $g_p(z \mid \lambda)$ is an unbiased estimator of θ, we will denote it by $h_p(z)$. Then we have for all unbiased h

$$(4.7) \quad pV_1(h) + (1 - p)V_0(h) = F_\lambda(h) \geqq F_\lambda(h_p) = pV_1(h_p) + (1 - p)V_0(h_p),$$

where equality holds only if $h(z) \equiv h_p(z) \quad (\nu)$.

The estimator $g_\lambda(z \mid p)$ will be unbiased if λ_0 and λ_1 satisfy

$$(4.8) \quad \lambda_1 \int \frac{z}{pz + (1 - p)(1 - z)}\, d\nu_i + \lambda_0 \int \frac{1 - z}{pz + (1 - p)(1 - z)}\, d\nu_i = i, \qquad i = 0, 1.$$

The determinant of these equations is

$$\left[\int \frac{z(1 - z)}{pz + (1 - p)(1 - z)}\, d\nu\right]^2 - \int \frac{(1 - z)^2}{pz + (1 - p)(1 - z)}\, d\nu \cdot \int \frac{z^2}{pz + (1 - p)(1 - z)}\, d\nu.$$

It is nonpositive by the Schwarz inequality. If it is equal to zero, z and $1 - z$ are linearly dependent a.e. (ν). Since z and $1 - z$ are densities, linear dependence would imply that ν_0 and ν_1 are identical; this possibility has been excluded here.

Thus, an explicit representation of estimators in H, minimizing (4.2) for $\boldsymbol{\theta}$ such that $\bar{\theta} = p$, is obtained as

$$\bar{h}_p(\mathbf{z}) = \frac{1}{n} \sum_1^n h_p(z_i), \qquad 0 < p < 1,$$

where $h_p(z) = g_p(z \mid \lambda)$, and λ is a solution of (4.8). This class of estimators merits our interest since each is admissible relative to H and each satisfies the following theorem.

THEOREM 2. *Let* $\bar{h}(\mathbf{z})$ *be any estimator in H such that* $|h(z)| + 1 < M < \infty$. *Define*

$$p_n(\mathbf{z}) = \begin{cases} 0, & \bar{h}(\mathbf{z}) < 0; \\ \bar{h}(\mathbf{z}), & 0 \leqq \bar{h}(\mathbf{z}) \leqq 1; \\ 1, & 1 < \bar{h}(\mathbf{z}). \end{cases}$$

Then, (a) $0 \leqq p_n(z) \leqq 1$, *and (b) for any* $\epsilon > 0$ *there exists an* $N(\epsilon)$ *such that, for any* $\boldsymbol{\theta}$ in Ω_∞,

$$\Pr[|p_n(z) - \bar{\theta}| > \epsilon \quad \textit{for some } n \geqq N(\epsilon)] \leqq \epsilon.$$

PROOF. It will clearly be sufficient to show that $\bar{h}(\mathbf{z})$ satisfies part (*b*). For this purpose we introduce $\bar{h}_i(k)$ for $i = 0$ or 1 and $k = 1, 2, \cdots$ to denote the arithmetic mean of k independent random variables, each having the probability measure $\nu_i h^{-1}$. We express

$$\bar{h}(\mathbf{z}) = (1/n)\sum[\theta_j h(z_j) + (1 - \theta_j)h(z_j)]$$

$$= \bar{\theta} \sum \theta_j h(z_j) / \sum \theta_j + (1 - \bar{\theta}) \sum (1 - \theta_j) h(z_j) / \sum (1 - \theta_j) \tag{4.9}$$

$$= \bar{\theta} \bar{h}_1(n\bar{\theta}) + (1 - \bar{\theta}) \bar{h}_0(n - n\bar{\theta}),$$

$$\bar{h}(\mathbf{z}) - \bar{\theta} = \bar{\theta}[\bar{h}_1(n\bar{\theta}) - 1] + (1 - \bar{\theta}) \bar{h}_0(n - n\bar{\theta}). \tag{4.10}$$

The strong law of large numbers yields the existence of functions $N_1(\eta)$ and $N_0(\eta)$, defined for $\eta > 0$ and such that

$$\Pr[|\bar{h}_i(k) - i| > \eta \quad \text{for some } k \geqq N_i(\eta)] \leqq \eta, \qquad i = 0, 1. \tag{4.11}$$

We fix $\epsilon > 0$ and consider the term $\bar{\theta}[\bar{h}_1(n\bar{\theta}) - 1]$ of (4.10). Since $\bar{\theta} \leqq 1$ and $|\bar{h}_1(n\bar{\theta}) - 1| \leqq M$, we have

$$|\bar{\theta}[\bar{h}_1(n\bar{\theta}) - 1]| \leqq \min \{\bar{\theta}M, \ |\bar{h}_1(n\bar{\theta}) - 1|\}. \tag{4.12}$$

Hence, for any fixed integer N and any fixed point $\boldsymbol{\theta}$ in Ω_∞, if $\mathbf{z}$ is such that there exist $n_{\mathbf{z}} \geqq N$ with $\bar{\theta}[\bar{h}_1(n_{\mathbf{z}}\bar{\theta}) - 1] > \epsilon/2$, then

$$n_{\mathbf{z}}\bar{\theta} > n_{\mathbf{z}} M^{-1}\epsilon/2, \qquad |\bar{h}_1(n_{\mathbf{z}}\bar{\theta}) - 1| > \epsilon/2.$$

Hence there exists $k > NM^{-1}\epsilon/2$ such that $|\bar{h}_1(k) - 1| > \epsilon/2$. Consequently

$$\begin{aligned} &\Pr\{\bar{\theta}[\bar{h}_1(n\bar{\theta}) - 1] > \epsilon/2 \text{ for some } n \geqq N\} \\ &\qquad \leqq \Pr\{|\bar{h}_1(k) - 1| > \epsilon/2 \text{ for some } k > NM^{-1}\epsilon/2\}. \end{aligned} \tag{4.13}$$

Thus, if $N \geqq 2M\epsilon^{-1}N_1(\epsilon/2)$, it follows from (4.11) for $i = 1$ that the right side of (4.13) is less than $\epsilon/2$ uniformly for all $\boldsymbol{\theta}$ in Ω_∞.

We can deal in a similar fashion with the term $(1 - \bar{\theta})\bar{h}_0(n - n\bar{\theta})$ of (4.10). Thus we obtain that part (b) is satisfied by

$$N(\epsilon) = 2M\epsilon^{-1} \max [N_1(\epsilon/2), N_0(\epsilon/2)]. \tag{4.14}$$

5. Nonsimple decision functions t. We have seen in Section 3 that if p is a good approximation to $\bar{\theta}$, then the simple decision function $\mathbf{t} = t_p$ (see (3.5)) does about as well as is possible for any simple decision function. Although a good approximation p to $\bar{\theta}$ is not generally available to the statistician, we have seen in Section 4 that for large n a good estimator $p_n(\mathbf{z})$ of $\bar{\theta}$ is always available.

It is natural to combine these two results by using the decision function t_p with the constant p replaced by the random variable $p_n(z)$. This amounts to using the nonsimple decision function $t^* = (t_1^*(\mathbf{z}), \cdots, t_n^*(\mathbf{z}))$ such that for $i = 1, \cdots, n$,

$$t_i^*(\mathbf{z}) = \begin{cases} 1, & \text{if } z_i > C(p_n(\mathbf{z})), \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

(We have chosen, arbitrarily, one way of resolving the ambiguity in the definition (3.5) when $z = C(p)$.)

In practice, $\mathbf{t}^*$ can be used only when all the values $z_1, \cdots, z_n$ are known before the individual decisions on the values of $\theta_1, \cdots, \theta_n$ have to be made.

We shall now investigate the behavior, for large n, of the risk function $R(\mathbf{t}^*, \boldsymbol{\theta})$.

We begin by considering the loss of the decision rule t_p (determined by (3.5) and by $t_p(C(p)) = 0$), where $\mathbf{z}$ is observed and $\boldsymbol{\theta}$ is the true parameter point. From (2.5) we obtain

$$
\begin{aligned}
W(t_p(\mathbf{z}), \boldsymbol{\theta}) &= a\bar{\theta}[\textstyle\sum\theta_i(1 - t_p(z_i)) / \sum\theta_i] \\
&\quad + b(1 - \bar{\theta})[\textstyle\sum(1 - \theta_i)t_p(z_i) / \sum(1 - \theta_i)] \qquad (5.2) \\
&= a\bar{\theta}S_1(C(p) \mid n\bar{\theta}) + b(1 - \bar{\theta})[1 - S_0(C(p) \mid n - n\bar{\theta})],
\end{aligned}
$$

where $S_i(v \mid k)$ is the sample distribution function of k independent random variables, each distributed with probability measure ν_i for $i = 0$ or 1 and $k = 1, 2, \cdots$. For future use we define

$$
(5.3) \quad D_1^+(k) = \sup_{0 \leq v \leq 1} \left[S_1(v \mid k) - \int_0^v d\nu_1 \right], \quad D_0^-(k) = \sup_{0 \leq v \leq 1} \left[\int_0^v d\nu_0 - S_0(v \mid k) \right].
$$

Using an alternative form of (3.3), it follows from (5.2) that

$$
\begin{aligned}
W(t_p(\mathbf{z}), \boldsymbol{\theta}) &= a\bar{\theta} \left\{ S_1(C(p) \mid n\bar{\theta}) - \int_0^{C(p)} d\nu_1 \right\} \\
&\quad + b(1 - \bar{\theta}) \left\{ \int_0^{C(p)} d\nu_0 - S_0(C(p) \mid n - n\bar{\theta}) \right\} + B(t_p, \boldsymbol{\theta}).
\end{aligned}
$$

Hence, from (5.3) and Theorem 1,

$$
W(t_p(\mathbf{z}), \boldsymbol{\theta}) \leq a\bar{\theta}D_1^+(n\bar{\theta}) + b(1 - \bar{\theta})D_0^-(n - n\bar{\theta}) + \phi(\bar{\theta}) + (a + b)\,|p - \bar{\theta}|\,.
$$

Since this holds for each $0 \leq p \leq 1$ we have, for any estimate $p_n(\mathbf{z})$ satisfying (a) of Theorem 1,

$$
\begin{aligned}
(5.4) \quad W(t_{p_n(\mathbf{z})}(\mathbf{z}), \boldsymbol{\theta}) &\leq a\bar{\theta}D_1^+(n\bar{\theta}) \\
&\quad + b(1 - \bar{\theta})D_0^-(n - n\bar{\theta}) + \phi(\bar{\theta}) + (a + b)\,|p_n(z) - \bar{\theta}|\,.
\end{aligned}
$$

By the Glivenko-Cantelli Theorem [4]

$$
\Pr\,[\lim_{k=\infty} D_1^+(k) = 0] = \Pr\,[\lim_{k=\infty} D_0^-(k) = 0] = 1.
$$

This is equivalent to the existence of functions $N_1^*(\eta)$ and $N_0^*(\eta)$, defined for $\eta > 0$, such that

$$
\Pr[D_1^+(k) > \eta \text{ for some } k \geq N_1^*(\eta)] \leq \eta,
$$
$$
\Pr[D_0^-(k) > \eta \text{ for some } k \geq N_0^*(\eta)] \leq \eta.
$$

As in (4.11) to (4.14), we have that if

$$
N^*(\epsilon) = 2\epsilon^{-1} \max[aN_1^*(\epsilon/2), bN_0^*(\epsilon/2)], \qquad \epsilon > 0,
$$

then, for any $\boldsymbol{\theta}$ in Ω_∞,

$$
(5.5) \quad \Pr[a\bar{\theta}D_1^+(n\bar{\theta}) + b(1 - \bar{\theta})D_0^-(n - n\bar{\theta}) > \epsilon \text{ for some } n \geq N^*(\epsilon)] \leq \epsilon.
$$

Returning to (5.4) we see that, for any $p_n(z)$ satisfying the conclusion of Theorem 2, (5.5) and Theorem 2 combine to furnish for any $\epsilon > 0$

$$n_0(\epsilon) = \max[N^*(\epsilon/2), \quad N(\epsilon/2(a + b))]$$

such that, for any $\boldsymbol{\theta}$ in Ω_∞,

$$\Pr\{W(t_{p_n(\mathbf{z})}(\mathbf{z}), \boldsymbol{\theta}) - \phi(\bar{\theta}) > \epsilon \text{ for some } n \geqq n_0(\epsilon)\} \leqq \epsilon. \tag{5.6}$$

The argument from (5.2) onward proves

THEOREM 3. *If $p_n(\mathbf{z})$ satifies the conclusions (a) and (b) of Theorem 2, then the decision function $\mathbf{t}^*$ defined by (5.1) is such that to any $\epsilon > 0$ there corresponds an $n_0(\epsilon)$ such that for any $\boldsymbol{\theta}$ in Ω_∞*

$$\Pr\{W(\mathbf{t}^*(\mathbf{z}), \boldsymbol{\theta}) \leqq \phi(\bar{\theta}) + \epsilon \text{ for all } n \geqq n_0(\epsilon)\} \geqq 1 - \epsilon.$$

Since $W(\mathbf{t}^*(\mathbf{z}), \boldsymbol{\theta}) \leqq \max\,[a, b]$, Theorem 3 implies

THEOREM 4. *Under the assumptions of Theorem 3, $\mathbf{t}^*$ is such that to any $\epsilon > 0$ there corresponds an $n_1(\epsilon)$ such that for any $n \geqq n_1(\epsilon)$ and any $\boldsymbol{\theta}$ in Ω,*

$$R(\mathbf{t}^*, \boldsymbol{\theta}) < \phi(\bar{\theta}) + \epsilon.$$

Thus, $\mathbf{t}^*$ is a nonsimple decision function which for large n does about as well as could be done by any simple decision function even if $\bar{\theta}$ were known exactly.

6. Invariant and R-invariant decision functions. Let $(P(1), \cdots, P(n))$ be an arbitrary permutation of $(1, \cdots, n)$. Define, for any real vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$,

$$P\boldsymbol{\xi} = (\xi_{P(1)}, \cdots, \xi_{P(n)}). \tag{6.1}$$

From (2.5) and (2.8) we note that $W(P\mathbf{t}, P\boldsymbol{\theta}) = W(\mathbf{t}, \boldsymbol{\theta})$ and $d(P\mathbf{z}, P\boldsymbol{\theta}) = d(\mathbf{z}, \boldsymbol{\theta})$. From (2.9) we obtain for any decision function $\mathbf{t}$, by a change of variable of integration,

$$\begin{aligned} R(\mathbf{t}, P\boldsymbol{\theta}) &= \int W(\mathbf{t}(\mathbf{z}), P\boldsymbol{\theta})\, d(\mathbf{z}, P\boldsymbol{\theta})\, d\nu^n \\ &= \int W(\mathbf{t}(P\mathbf{z}), P\boldsymbol{\theta})\, d(P\mathbf{z}, P\boldsymbol{\theta})\, d\nu^n = R(P^{-1}\mathbf{t}P, \boldsymbol{\theta}). \end{aligned} \tag{6.2}$$

We call a decision function $\mathbf{t}$ *invariant* if $P^{-1}\mathbf{t}P = \mathbf{t}$ for all P. We denote by *R-invariant* a decision function $\mathbf{t}$ such that

$$R(P^{-1}\mathbf{t}P, \boldsymbol{\theta}) = R(\mathbf{t}, \boldsymbol{\theta}), \qquad \text{for all } P \text{ and } \boldsymbol{\theta}. \tag{6.3}$$

The risk of an R-invariant decision function can be expressed as an explicit function of $\bar{\theta}$ by formally averaging the representation (2.9) over all P. From (2.9), (6.2), and (6.3),

$$R(\mathbf{t}, \boldsymbol{\theta}) = \frac{1}{n!}\sum_P R(\mathbf{t}, P\boldsymbol{\theta}) = \frac{1}{n!}\sum_P \int W(\mathbf{t}(\mathbf{z}), P\boldsymbol{\theta})\, d(\mathbf{z}, P\boldsymbol{\theta})\, d\nu^n. \tag{6.4}$$

Changing the order of summations and integration, we find that if $\mathbf{t}$ is R-invariant its risk, $R(\mathbf{t}, \boldsymbol{\theta})$, is expressible in the form

$$R(\mathbf{t}, \boldsymbol{\theta}) = \frac{1}{n} \sum_1^n \int \left\{ a \frac{1 - t_i(\mathbf{z})}{n!} \sum_P \theta_{P(i)}\, d(\mathbf{z}, P\boldsymbol{\theta}) + b \frac{t_i(\mathbf{z})}{n!} \sum_P (1 - \theta_{P(i)})\, d(\mathbf{z}, P\boldsymbol{\theta}) \right\} d\nu^n. \tag{6.5}$$

For any real vector $\mathbf{r} = (r_1, \cdots, r_n)$ we define for every integer k

$$L(k, n, \mathbf{r}) = \begin{cases} \dfrac{1}{n!} \sum_P \prod_1^k r_{P(j)} \prod_{k+1}^n (1 - r_{P(j)}), & k = 0, 1, \cdots, n; \\ 0, & \text{otherwise}; \end{cases}$$

$L(k, n - 1, \check{\mathbf{r}}_i = L[k, n - 1, (r_1, \cdots, r_{i-1}, r_{i+1}, \cdots, r_n)], \quad i = 1, 2, \cdots, n.$

Then

$$\begin{aligned} \sum_P \theta_{P(i)}\, d(\mathbf{z}, P\boldsymbol{\theta}) &= z_i \sum_{\theta_{P(i)}=1} \prod_{j \neq i} [\theta_{P(j)} z_j + (1 - \theta_{P(j)})(1 - z_j)] \\ &= z_i \left(\sum_1^n \theta_k \right) (n - 1)!\, L\left(\sum_1^n \theta_k - 1, \quad n - 1, \quad \check{\mathbf{z}}_i \right); \end{aligned} \tag{6.6}$$

$$\sum_P (1 - \theta_{P(i)})\, d(\mathbf{z}, P\boldsymbol{\theta}) = (1 - z_i) \sum_1^n (1 - \theta_k)(n - 1)! \cdot L\left(\sum_1^n \theta_k, \quad n - 1, \quad \check{\mathbf{z}}_i \right). \tag{6.7}$$

It follows from (6.5) that for R-invariant $\mathbf{t}$

$$R(\mathbf{t}, \boldsymbol{\theta}) = \frac{1}{n} \sum_1^n \int \{ a[1 - t_i(\mathbf{z})] z_i \bar{\theta} L(n\bar{\theta} - 1, \quad n - 1, \quad \check{\mathbf{z}}_i) + b t_i(\mathbf{z})(1 - z_i)(1 - \bar{\theta}) L(n\bar{\theta}, \quad n - 1, \check{\mathbf{z}}_i) \} \, d\nu^n. \tag{6.8}$$

Denoting the ith summand of (6.8) by $R_i(\mathbf{t} \mid \bar{\theta})$, we note for later use that, if $\mathbf{t}$ is actually invariant,

$$R_i(\mathbf{t} \mid \bar{\theta}) = R_1(\mathbf{t} \mid \bar{\theta}) = R(\mathbf{t}, \boldsymbol{\theta}), \qquad i = 1, \cdots, n. \tag{6.9}$$

This follows from the invariance and a permutation of the variables of integration in the representation of $R_i(\mathbf{t} \mid \bar{\theta})$ in (6.8). It further follows from this representation that, for any fixed $\boldsymbol{\theta}$, the integrand of $R_i(\mathbf{t} \mid \bar{\theta})$ is, for each fixed $\mathbf{z}$, at a minimum with respect to t_i if and only if t_i is of the form defined, for $p = 0$, $1/n, \cdots, (n - 1)/n, 1$, by

$$t_{ip}(\mathbf{z}) = \begin{Bmatrix} 1, \\ 0, \\ \text{arbitrary}, \end{Bmatrix} a z_i p L(np - 1, n - 1, \check{\mathbf{z}}_i) \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} b(1 - z_i)(1 - p) \cdot L(np, \quad n - 1, \quad \check{\mathbf{z}}_i), \tag{6.10}$$

with p equal to $\bar{\theta}$. We denote $(t_{1p}(\mathbf{z}), \cdots, t_{np}(\mathbf{z}))$ by $\mathbf{t}_p$. With a proper determination of the arbitrary part, any decision function of this form (6.10) is invariant. Hence, the representation of its risk in the form (6.8) is valid and $R(\mathbf{t}, \boldsymbol{\theta})$ for each fixed $\boldsymbol{\theta}$ is minimized over the class of R-invariant decision functions by an invariant decision function of the form (6.10).

Let

$$\phi_n(\bar{\theta}) = \inf_{R\text{-invariant } \mathbf{t}} R(\mathbf{t}, \boldsymbol{\theta}). \tag{6.11}$$

From (6.8), (6.9), and (6.10) we obtain

$$\begin{aligned} \phi_n(\bar{\theta}) &= R_1(t_{\bar{\theta}} \mid \bar{\theta}) \\ &= \int \min\,[az_1\bar{\theta}L(n\bar{\theta} - 1, \quad n - 1, \check{z}_1), \\ &\qquad\qquad b(1 - z_1)(1 - \bar{\theta})L(n\bar{\theta}, \quad n - 1, \quad \check{z}_1)]\, d\nu^n. \end{aligned} \tag{6.12}$$

Since $L(k, \quad n - 1, \quad \check{z}_1)$ is a symmetric function of $z_2, \cdots, z_n$, we have for any symmetric measurable set S

$$\int_S L(k, \quad n - 1, \quad \check{z}_1)\, d\nu^{n-1} = \nu_1^k \nu_0^{n-1-k}(S). \tag{6.13}$$

It follows that the integral with respect to $z_2, \cdots, z_n$ of a z_1-section of the integrand of (6.12) is

$$az_1\bar{\theta}[1 - \nu_1^{n\bar{\theta}-1}\nu_0^{n-n\bar{\theta}}(S_{z_1})] + b(1 - z_1)(1 - \bar{\theta})\, \nu_1^{n\bar{\theta}}\nu_0^{n-1-n\bar{\theta}}(S_{z_1})$$

where S_{z_1} is the symmetric set

$$\begin{aligned} S_{z_1} = [(z_2, \cdots, z_n) \mid az_1\bar{\theta}L(n\bar{\theta} - 1, \quad n - 1, \quad \check{z}_1) \\ > b(1 - z_1)(1 - \bar{\theta})L(n\bar{\theta}, \quad n - 1, \quad \check{z}_1)]. \end{aligned}$$

The developments from (6.4) onward set the stage for

THEOREM 5. *If $\phi(\bar{\theta})$ and $\phi_n(\bar{\theta})$ are defined by* (3.6) *and* (6.12), *respectively, then for any $\epsilon > 0$ there exists $N(\epsilon)$ such that for any $n \geqq N(\epsilon)$ and any $\boldsymbol{\theta}$*

$$\phi(\bar{\theta}) - \epsilon < \phi_n(\bar{\theta}) \leqq \phi(\bar{\theta}).$$

PROOF. That $\phi_n \leqq \phi$ for all $\boldsymbol{\theta}$ follows from the fact that every simple decision function is invariant.

For the nontrivial part of the proof we fix $\epsilon > 0$. From the continuity of $\phi(p)$ ((3.10) and (3.11)) we obtain a $\delta(\epsilon)$ such that

$$\phi(p) - \epsilon < 0 \qquad \text{if } p < \delta \quad \text{or } p > 1 - \delta. \tag{6.14}$$

Thus, it will be sufficient to show the existence of an $N(\epsilon)$ which suffices for the theorem when $0 < \delta \leqq \bar{\theta} \leqq 1 - \delta$. We will obtain this from the following measure theoretic lemma, a slight generalization of which is proved in [5].

LEMMA. *If m_0 and m_1 are nondisjoint p-measures on a σ-algebra of subsets χ of a set X, then for every $\epsilon, \delta > 0$ there exists $N(\epsilon, \delta)$ such that for any pair of positive integers r, s with $r + s \geqq N(\epsilon, \delta)$ and $\delta \leqq r/(r+s) \leqq 1 - \delta$,*

$$|m_0^r m_1^s(S) - m_0^{r-1} m_1^{s+1}(S)| < \epsilon \text{ uniformly for all symmetric } S \text{ in } \chi^{r+s}.$$

From this lemma we obtain that if $n - 1 \geqq N(\eta, \delta^*)$ and $\delta^* \leqq n\bar{\theta}/(n-1) \leqq 1 - \delta^*$,

$$[1 - \nu_1^{n\bar{\theta}-1}\nu_0^{n-n\bar{\theta}}(S_{z_1})] + \nu_1^{n\bar{\theta}}\nu_0^{n-1-n\bar{\theta}}(S_{z_1}) > 1 - \eta$$

uniformly for all z_1. From this it follows that $\phi_n(\bar{\theta}) > (1 - \eta)\phi(\bar{\theta}) = \phi(\bar{\theta}) - \eta\phi(\bar{\theta})$. Since the $\max_p \phi(p) \leqq c = ab\,/\,(a+b)$, we see that in view of (6.14) the choice $N(\epsilon) = N(c^{-1}\epsilon, 2^{-1}\delta(\epsilon))$ will complete the proof of Theorem 5.

Theorem 5 can be combined with Theorem 3 or Theorem 4 to give a strengthened endorsement of $\mathbf{t}^*$. For large n, $\mathbf{t}^*$ does about as well as could be done by any R-invariant decision function even if $\bar{\theta}$ were known exactly.

7. Bayes' and minimax solutions. Let $\mathbf{t} = (t_1(\mathbf{z}), \cdots, t_n(\mathbf{z}))$ be any decision function in the compound decision problem. For any $\boldsymbol{\theta}$ in Ω we can write the risk (2.9) in the form

$$R(\mathbf{t}, \boldsymbol{\theta}) = \int \frac{1}{n} \sum_1^n \{a\theta_i[1 - t_i(\mathbf{z})]\, d(\mathbf{z}, \boldsymbol{\theta}) + b(1 - \theta_i)t_i(\mathbf{z})\, d(\mathbf{z}, \boldsymbol{\theta})\}\, d\nu^n.$$

By a weight function we mean any function $\beta(\boldsymbol{\theta}) \geqq 0$ defined on Ω and not identically 0. For any weight function $\beta(\boldsymbol{\theta})$ and any decision function $\mathbf{t}$ we define the weighted risk of $\mathbf{t}$ relative to $\beta(\boldsymbol{\theta})$ as

$$\begin{aligned} B(\mathbf{t}, \beta) = \sum_{\boldsymbol{\omega}} \beta(\boldsymbol{\omega}) \cdot R(\mathbf{t}, \boldsymbol{\omega}) = \int \frac{1}{n} \sum_1^n \{(1 - t_i(\mathbf{z}))\, a \sum_{\boldsymbol{\omega}} \beta(\boldsymbol{\omega})\omega_i\, d(\mathbf{z}, \boldsymbol{\omega}) \\ + t_i(\mathbf{z}) b \sum_{\boldsymbol{\omega}} \beta(\boldsymbol{\omega})(1 - \omega_i)\, d(\mathbf{z}, \boldsymbol{\omega})\}\, d\nu^n. \end{aligned}$$

For fixed $\beta(\boldsymbol{\theta})$ this will be a minimum if and only if for almost every $(\nu^n)\mathbf{z}$ and for every $i = 1, \cdots, n$, $t_i(\mathbf{z})$ is of the form

$$t_i(\mathbf{z} \mid \beta) = \begin{Bmatrix} 1, \\ 0, \\ \text{arbitrary,} \end{Bmatrix} a \sum_{\boldsymbol{\omega}} \beta(\boldsymbol{\omega})\omega_i\, d(\mathbf{z}, \boldsymbol{\omega}) \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} b \sum_{\boldsymbol{\omega}} \beta(\boldsymbol{\omega})(1 - \omega_i)\, d(\mathbf{z}, \boldsymbol{\omega}). \tag{7.1}$$

Any decision function $\mathbf{t}$ of the form (7.1) is called a *Bayes' solution* relative to the weight function $\beta(\boldsymbol{\theta})$.

For the remainder of this section we restrict our attention to symmetric $\beta(\boldsymbol{\theta})$, that is, such that

$$\binom{n}{k} \beta(\boldsymbol{\theta}) = \beta_k \qquad \text{for all } \boldsymbol{\theta} \text{ with } \sum_1^n \theta_j = k, \quad k = 0, 1, \cdots, n. \tag{7.2}$$

For symmetric β we obtain the representation (see (6.8) to (6.11))

$$a \sum_{\omega} \beta(\omega)\omega_i\, d(z, \omega) = az_i \sum_{k=0}^{n} \beta_k \frac{k}{n} L(k - 1, \quad n - 1, \quad \check{z}_i),$$

$$b \sum_{\omega} \beta(\omega)(1 - \omega_i)\, d(z, \omega) = b(1 - z_i) \sum_{k=0}^{n} \beta_k(1 - k/n)L(k, \quad n - 1, \quad \check{z}_i).$$

Several particular cases of (7.2) hold special interest for us. If for some integer k_0 with $0 \leqq k_0 \leqq n$, (7.2) takes the form

$$\beta_{k_0} = 1, \qquad \beta_k = 0 \qquad\qquad k = 0, \cdots k_0-1, k_0 + 1, \cdots, n,$$

then the corresponding Bayes' solution (7.1) is a decision function of the form (6.10).

If for some constant p with $0 \leqq p \leqq 1$, expression (7.2) takes the form

$$\beta_k = \binom{n}{k} p^k(1 - p)^{n-k},$$

then the corresponding Bayes' solution (7.1) is a decision function of the form (3.5) provided the Bayes' solution is required to be simple on the arbitrary part of its definition.

Referring back to Section 3, we have by the previous paragraph that each of the simple decision functions $\mathbf{t} = t_p$ for $0 \leqq p \leqq 1$, defined by (3.5), is a Bayes' solution for the compound decision problem. It can be shown (we omit the simple proof) that there always exists a value $0 < p' < 1$ and a determination of $t_{p'}$ for which the coefficient of p in (3.3) vanishes. Letting r be the constant value of $B(t_{p'}, p)$ it follows that $R(t_{p'}, \boldsymbol{\theta}) \equiv r$ for every $\boldsymbol{\theta}$ in Ω.

Since $t_{p'}$ has constant risk and is a Bayes' solution relative to a weight function which is positive for every $\boldsymbol{\theta}$ in Ω, it follows that $t_{p'}$ is the admissible minimax decision function, unique in the sense of risk. That is

$$\sup_{\boldsymbol{\theta} \text{ in } \Omega} R(\mathbf{t}, \boldsymbol{\theta}) = \text{minimum for } \mathbf{t} = t_{p'},$$

and if $\mathbf{t}$ is any other decision function such that $\sup R(\mathbf{t}, \boldsymbol{\theta}) = r$, then $R(\mathbf{t}, \boldsymbol{\theta}) \equiv r$.

8. Admissibility. Since the minimax decison function $t_{p'}$ is simple and has constant risk r independent of n, it follows that

$$r = R(t_{p'}, \boldsymbol{\theta}) \geqq \inf_{t} R(t, \boldsymbol{\theta}) = \phi(\bar{\theta}) \qquad\qquad \text{for all } \boldsymbol{\theta} \text{ in } \Omega,$$

and hence $r = \phi(p') = \max_p \phi(p)$.

If μ_0 and μ_1 are nondisjoint, $\phi(0) = \phi(1) = 0 < r$ and we conclude from Theorem 4 that, for any $0 < \epsilon < r$, the minimax decision function is ϵ-inadmissible (See [6] for definitions) for all sufficiently large n, since decision functions of the type $\mathbf{t}^*$ are ϵ-better.

The present paper has failed to exhibit a $\mathbf{t}^*$ which is admissible (or even to show the existence of such). This deficiency is remedied, at least asymptotically, by the fact that, for any $\epsilon > 0$ Theorems 4 and 5 together imply that any decision function of the type t* is ϵ-uniformly-best (relative to the class of R-invariant decision functions) for all sufficiently large n.

REFERENCES

[1] HERBERT ROBBINS, "Asymptotically subminimax solutions of compound statistical decision problems," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951.

[2] P. R. HALMOS AND L. J. SAVAGE, "Application of the Radon-Nikodym theorem to the theory of sufficient statistics," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 225–242.

[3] P. R. HALMOS, *Measure Theory*, D. Van Nostrand Company, 1950.

[4] M. FRÉCHET, *Recherches Théoriques Modernes sur le Calcul des Probabilités*, Vol. 1, Gauthier-Villars, Paris, 1937.

[5] J. F. HANNAN, "Asymptotic solutions of compound decision problems," *Institute of Statistics Mimeograph Series*, University of North Carolina, No. 68 (1953), pp. 1–77.

[6] J. WOLFOWITZ, "On ϵ-complete classes of decision functions," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 461–465.