ADMISSIBLE AND MINIMAX INTEGER-VALUED ESTIMATORS OF AN INTEGER-VALUED PARAMETER[1]

By D. S. Robson

Cornell University

1. Summary. The decision problem considered here is that of deciding which element of a finite parametric family of probability distributions $p(x, \mu)$ represents the true distribution of the statistic X. It is assumed that $p(x, \mu)$ satisfies certain regularity conditions which essentially require that the parameter μ be integer-valued with known bounds and that $p(x, \mu_1)/p(x, \mu_0)$ be an increasing function of x whenever $\mu_0 < \mu_1$. Complete classes are characterized for various loss functions $W(\mu, \alpha)$ which are convex functions of the decision α for each fixed value of μ. Minimax procedures are considered for the case $W(\mu, \alpha) = |\alpha - \mu|^k$.

2. Introduction. The problem of estimating an integer-valued parameter is viewed as a special case of Wald's general statistical decision problem. The chance variable X is known to be distributed over the sample space M according to a probability distribution $p(x, \mu)$ depending upon a single unknown integer-valued parameter μ with the known bounds $0 \leqq \mu \leqq N$. The statistician is required to make one of $N + 1$ decisions, corresponding to the $N + 1$ different possible values of μ, on the basis of a single observed value of X. A decision function δ therefore has the form

$$\delta(x) = (\delta_0(x), \delta_1(x), \cdots, \delta_N(x)) \tag{1}$$

where $\delta_\alpha(x) \geqq 0$ for $\alpha = 0, 1, \cdots, N$ and $\sum_{\alpha=0}^{N} \delta_\alpha(x) = 1$ for all x in M, with the interpretation that when the procedure δ is used and the observed value of X is x_0 then the decision that the true distribution of X is $p(x, \alpha)$ is to be made with probability $\delta_\alpha(x_0)$, $\alpha = 0, 1, \cdots, N$. The loss associated with the decision α when the true value of the parameter is μ is expressed by a loss function $W(\mu, \alpha)$ which, for each fixed value of μ, is a convex function of α with $W(\mu, \mu) = 0$ and, for α between μ and β, $W(\mu, \alpha) < W(\mu, \beta)$.

The following regularity conditions are imposed upon the function $p(x, \mu)$.

Condition 1. $p(y, \mu)p(x, v) < p(x, \mu)p(y, v)$ *if and only if* $p(y, \mu)p(x, v)$ *and* $p(x, \mu)p(y, v)$ *are not both zero and* $x < y$, $\mu < v$.

Condition 2. *If* $p(x, v) = 0$ *for all* x *in* M *then* $p(x, \mu) = 0$ *for all* x *in* M *either for every* $\mu \leqq v$ *or for every* $\mu \geqq v$.

Condition 3. *If* $M = (x_0, x_1, \cdots, x_n)$, $x_{i-1} < x_i$, *then for every* i, $0 < i \leqq n$, *there exists an integer* μ_i *such that* $p(x_{i-1}, \mu_i) > 0$ *and* $p(x_i, \mu_i) > 0$.

Conditions 1 and 2 are essentially a more precise way of saying that the likelihood ratio $p(x, v) / p(x, \mu)$ is a strictly increasing function of x whenever $\mu < v$.

Received October 5, 1955; revised January 10, 1958.

[1] Part of a doctoral thesis presented to the Faculty of the Graduate School of Cornell University.

A simple but useful consequence of Condition 1 is the following

LEMMA 1. *If the distribution* $p(x, \mu)$ *satisfies Condition* 1 *and if* $p(y, \alpha) = 0$ *and if there exists a pair* z, β *such that* $p(z, \alpha) > 0$ *and* $p(y, \beta) > 0$ *then either*

(i) $p(y, \mu) = 0$ *for all* $\mu \leqq \alpha$ *and* $p(x, \alpha) = 0$ *for all* $x \geqq y$, *or*

(ii) $p(y, \mu) = 0$ *for all* $\mu \geqq \alpha$ *and* $p(x, \alpha) = 0$ *for all* $x \leqq y$.

3. A Karlin-Rubin Complete Class Theorem. A general approach to decision problems involving distributions with a monotone likelihood ratio has been developed by H. Rubin [1] and S. Karlin and H. Rubin [2]. Since the finite action problem posed here represents a special case of the Karlin-Rubin problem, a direct application of their results concerning completeness of the class of monotone decision procedures gives the following

THEOREM 1. *Let C be the class of decision functions such that*

(i) *for every x in M there exists an integer* α_x *such that* $\delta_{\alpha_x}(x) + \delta_{\alpha_x+1}(x) = 1$

(ii) $\delta_\alpha(x) > 0$ *only if* $p(x, \alpha) > 0$

(iii) if $x < y$ then $\bar{\alpha}_x \equiv \alpha_x \delta_{\alpha_x}(x) + (\alpha_x + 1)\delta_{\alpha_x+1}(x) \leqq \alpha_y$

If $p(x, \mu)$ *satisfies Conditions* 1 *and* 2 *and if, for each fixed* μ, *the loss function* $W(\mu, \alpha)$ *is a convex function of* α *with* $W(\mu, \mu) = 0$ *and, for* α *between* μ *and* β, $W(\mu, \alpha) < W(\mu, \beta)$ *then the class C is complete.*

The theorem remains valid under weaker conditions on the loss function[2]; however, in what follows only convex loss functions are considered.

4. Admissible procedures when $W(\mu, \alpha) = | \alpha - \mu |^k$ **for large** k. The class C may, under the hypotheses of Theorem 1, contain inadmissible procedures. This is effectively demonstrated by the special case where $W(\mu, \alpha)$ is a convex function of $| \alpha - \mu |$ and increases very rapidly with $| \alpha - \mu |$. $W(\mu, \alpha) = | \alpha - \mu |^k$ is one example of such a loss function and, clearly, any convex function $W(| \alpha - \mu |)$ with $W(0) = 0$ can be dominated by $K | \alpha - \mu |^k$ by choosing the constants K and k sufficiently large. The most stringent requirements for admissibility are then encountered when the range of x for which $p(x, \mu) > 0$ is independent of μ; in particular,

THEOREM 2. *If* $p(x, \mu)$ *satisfies Condition* 1 *and* $p(x, \mu) > 0$ *for all integer pairs* (x, μ) *such that* $0 \leqq x \leqq n$, $0 \leqq \mu \leqq N$, *and* $p(x, \mu) = 0$ *otherwise, then there exists an integer* $k_p > 0$ *such that if* $W(\mu, \alpha) = | \alpha - \mu |^k$ *and* $k \geqq k_p$ *then every admissible procedure is of the form*

$$\begin{aligned} \delta_\alpha(x) &= 1 \text{ for } x < y \\ \delta_\alpha(y) + \delta_{\alpha+1}(y) &= 1 \\ \delta_{\alpha+1}(x) &= 1 \text{ for } x > y \end{aligned}$$

where $0 \leqq y \leqq n, 0 \leqq \alpha \leqq N$.

[2] The author proved the theorem as it is stated and a referee pointed out that the Karlin-Rubin theorem for the finite action problem includes this result.

PROOF OF THEOREM 2. The conclusion is obtained by showing that, under the hypotheses specified, every Bayes solution has this form when k is sufficiently large. Let $\xi = (\xi_0, \xi_1, \cdots, \xi_N)$ be an a priori distribution on the parameter space and let $r(\xi, \delta)$ be the integrated risk of the procedure δ. Then δ^ξ is said to be a Bayes solution relative to ξ if $\inf_\delta r(\xi, \delta) = r(\xi, \delta^\xi)$. For every ξ, however, there exists a non-randomized Bayes solution; consequently, if

$$r_x(\xi, \alpha) = \sum_\mu | \alpha - \mu |^k p(x, \mu)\xi_\mu$$

then

$$r(\xi, \delta^\xi) = \sum_x \inf_\alpha r_x(\xi, \alpha).$$

The function $r_x(\xi, \alpha)$ is seen to have the following properties:

I: If $r_x(\xi, \alpha) \leqq r_x(\xi, \alpha + 1)$ then $r_x(\xi, \alpha + \beta) < r_x(\xi, \alpha + \beta + 1)$ for all $\beta > 0$

II: If $r_y(\xi, \alpha) \leqq r_y(\xi, \alpha + 1)$ then $r_x(\xi, \alpha) < r_x(\xi, \alpha + 1)$ for all $x \leqq y$

III: If $r_0(\xi, \alpha - 1) \leqq r_0(\xi, \alpha)$ then $r_n(\xi, \alpha) < r_n(\xi, \alpha + 1)$ for all $k \geqq k_p$.

Let $\Delta_k(\alpha) = (\alpha + 1)^k - \alpha^k$ then $r_x(\xi, \alpha) \leqq r_x(\xi, \alpha + 1)$ is equivalent to

$$\sum_{\mu=\alpha+1}^{N} \Delta_k(\mu - \alpha - 1)p(x, \mu)\xi_\mu \leqq \sum_{\mu=0}^{\alpha} \Delta_k(\alpha - \mu)p(x, \mu)\xi_\mu .$$

Then property I follows from

$$\begin{aligned}\sum_{\mu=\alpha+\beta+1}^{N} \Delta_k(\mu - \alpha - \beta - 1)p(x, \mu)\xi_\mu &\leqq \sum_{\mu=\alpha+1}^{N} \Delta_k(\mu - \alpha - 1)p(x, \mu)\xi_\mu \\ &\leqq \sum_{\mu=0}^{\alpha} \Delta_k(\alpha - \mu)p(x, \mu)\xi_\mu \leqq \sum_{\mu=0}^{\alpha+\beta} \Delta_k(\alpha + \beta - \mu)p(x, \mu)\xi_\mu\end{aligned} \tag{2}$$

for all β, $0 < \beta \leqq N - \alpha - 1$. Since $p(x, \mu) > 0$, $0 \leqq x \leqq n$, $0 \leqq \mu \leqq N$, then either the first or last inequality (or both) of (2) is strict for $k > 1$. Property II is a direct result of the restrictions upon $p(x, \mu)$, for if $r_y(\xi, \alpha) \leqq r_y(\xi, \alpha + 1)$ then, since $p(x, \mu)$ satisfies Condition 1 and $p(x, \mu) > 0, 0 \leqq x \leqq n, 0 \leqq \mu \leqq N$,

$$\begin{aligned}\sum_{\mu=\alpha+1}^{N} \Delta_k(\mu - \alpha - 1) \frac{p(x, \mu)}{p(x, \alpha)} \xi_\mu &\leqq \sum_{\mu=\alpha+1}^{N} \Delta_k(\mu - \alpha - 1) \frac{p(y, \mu)}{p(y, \alpha)} \xi_\mu \\ &\leqq \sum_{\mu=0}^{\alpha} \Delta_k(\alpha - \mu)\frac{p(y, \mu)}{p(y, \alpha)} \xi_\mu \leqq \sum_{\mu=0}^{\alpha} \Delta_k(\alpha - \mu) \frac{p(x, \mu)}{p(x, \alpha)} \xi_\mu\end{aligned} \tag{3}$$

for all $x \leqq y$, with either the first or last inequality (or both) of (3) being strict for all k. Property III is derived by noting that if $r_0(\xi, \alpha - 1) \leqq r_0(\xi, \alpha)$ then

$$\Delta_k(\alpha - 1)\xi_0 \geqq - \sum_{\mu=1}^{\alpha-1} \Delta_k(\alpha - 1 - \mu) \frac{p(0, \mu)}{p(0, 0)} \xi_\mu + \sum_{\mu=\alpha}^{N} \Delta_k(\mu - \alpha) \frac{p(0, \mu)}{p(0, 0)} \xi_\mu$$

and

$$
\begin{aligned}
&\frac{\Delta_k(\alpha-1)}{p(n,0)}(r_n(\xi,\alpha+1)-r_n(\xi,\alpha))\\
&\geqq \Delta_k(\alpha)\left(-\sum_{\mu=1}^{\alpha-1}\Delta_k(\alpha-1-\mu)\frac{p(0,\mu)}{p(0,0)}\xi_\mu+\sum_{\mu=\alpha}^{N}\Delta_k(\mu-\alpha)\frac{p(0,\mu)}{p(0,0)}\xi_\mu\right)\\
&+\Delta_k(\alpha-1)\left(\sum_{\mu=1}^{\alpha}\Delta_k(\alpha-\mu)\frac{p(n,\mu)}{p(n,0)}\xi_\mu-\sum_{\mu=\alpha+1}^{N}\Delta_k(\mu-\alpha-1)\frac{p(n,\mu)}{p(n,0)}\xi_\mu\right)\\
&=\sum_{\mu=1}^{\alpha-1}\left(\Delta_k(\alpha-1)\Delta_k(\alpha-\mu)\frac{p(n,\mu)}{p(n,0)}-\Delta_k(\alpha)\Delta_k(\alpha-\mu-1)\frac{p(0,\mu)}{p(0,0)}\right)\xi_\mu\\
&\qquad+\left(\Delta_k(\alpha-1)\frac{p(n,\alpha)}{p(n,0)}+\Delta_k(\alpha)\frac{p(0,\alpha)}{p(0,0)}\right)\xi_\alpha\\
&+\sum_{\mu=\alpha+1}^{N}\left(\Delta_k(\alpha)\Delta_k(\mu-\alpha)\frac{p(0,\mu)}{p(0,0)}-\Delta_k(\alpha-1)\Delta_k(\mu-\alpha-1)\frac{p(n,\mu)}{p(n,0)}\right)\xi_\mu .
\end{aligned}
\tag{4}
$$

For $0 \leqq \mu \leqq \alpha - 1$, $\Delta_k(\alpha-1)\ \Delta_k(\alpha-\mu) \geqq \Delta_k(\alpha)\Delta_k(\alpha-\mu-1)$ and, by Condition 1, $(p(n,\mu)/p(n,0)) > p(0,\mu)/p(0,0))$ so the coefficient of ξ_μ in (4) is positive for all $\mu \leqq \alpha$. And since, for $\mu > \alpha$, $(\Delta_k(\alpha)\Delta_k(\mu-\alpha)/\Delta_k(\alpha-1)$ $\Delta_k(\mu-\alpha-1))$ can be made arbitrarily large by choosing k sufficiently large then $k_p(\mu,\alpha)$ may be defined as the smallest integer k such that

$$\frac{\Delta_k(\alpha)\Delta_k(\mu-\alpha)}{\Delta_k(\alpha-1)\Delta_k(\mu-\alpha-1)} > \frac{p(n,\mu)p(0,0)}{p(n,0)p(0,\mu)}.$$

Hence, for $k \geqq \max_\mu k_p(\mu,\alpha)$ the coefficient of ξ_μ in (4) is positive for all $\mu > \alpha$, and property III is thus established by taking

$$k_p = \max_{\substack{0\leqq\alpha\leqq N\\ \alpha\leqq\mu\leqq N}} k_p(\mu,\alpha)$$

Now let α_x^ξ be an integer such that $\min_{0\leqq\alpha\leqq N} r_x(\xi,\alpha) = r_x(\xi,\alpha_x^\xi)$. Then for $x < y$, $\alpha_x^\xi \leqq \alpha_y^\xi$. For suppose $x < y$ and $\alpha_y^\xi < \alpha_x^\xi$; since $r_y(\xi,\alpha_y^\xi) \leqq r_y(\xi,\alpha_y^\xi+1)$ then, by II, $r_x(\xi,\alpha_y^\xi) < r_x(\xi,\alpha_y^\xi+1)$ and then I implies the contradiction $r_x(\xi,\alpha_x^\xi-1) < r_x(\xi,\alpha_x^\xi)$. For $k \geqq k_p$, III gives $r_n(\xi,\alpha_0^\xi+1) < r_n(\xi,\alpha_0^\xi+2)$ which implies by II, that $r_x(\xi,\alpha_0^\xi+1) < r_x(\xi,\alpha_0^\xi+2)$ for all $x \leqq n$ and this implies, by I, that $r_x(\xi,\alpha_0^\xi+\beta) < r_x(\xi,\alpha_0^\xi+\beta+1)$ for all $\beta > 1$. Hence, $\alpha_0^\xi \leqq \alpha_x^\xi \leqq \alpha_0^\xi + 1$ for all x. If there exists a value of x such that $\alpha_x^\xi = \alpha_0^\xi + 1$ let y be the least such x, then $0 \leqq y \leqq n$ and

$$\alpha_x^\xi = \begin{cases}\alpha_0^\xi \text{ for } x < y\\ \alpha_0^\xi + 1 \text{ for } x \geqq y;\end{cases}$$

in this case, randomized Bayes solutions exist and are of the form

$$
\begin{aligned}
\delta_{\alpha_0^\xi}(x) &= 1 \text{ for } x < y\\
\delta_{\alpha_0^\xi}(y) + \delta_{\alpha_0^\xi+1}(y) &= 1\\
\delta_{\alpha_0^\xi+1}(x) &= 1 \text{ for } x > y.
\end{aligned}
$$

Since every admissible procedure is a Bayes solution this completes the proof of Theorem 2.

The distribution

$$(5) \qquad p(x,\mu) = \binom{n}{x}\left(\frac{\mu}{N}\right)^{x}\left(1-\frac{\mu}{N}\right)^{n-x}, \qquad 0 \leqq x \leqq n, 0 \leqq \mu \leqq N$$

satisfies the hypotheses of Theorem 2 for $0 < x < n$, and since Theorem 1 applies for $x = 0, n$ then

COROLLARY. *If $p(x, \mu)$ is the distribution* (5) *then when $k \geqq k_p$ a procedure δ is admissible only if*

$$\begin{aligned} \delta_\alpha(0) + \delta_{\alpha+1}(0) &= 1 \\ \delta_\beta(x) &= 1 \textit{ for } 0 < x < y \\ \delta_\beta(y) + \delta_{\beta+1}(y) &= 1 \\ \delta_{\beta+1}(x) &= 1 \textit{ for } y < x < n \\ \delta_\gamma(n) + \delta_{\gamma+1}(n) &= 1 \end{aligned}$$

where the integers α, β, γ satisfy $0 \leqq \alpha < \beta < \gamma \leqq N$.

5. Admissible procedures when $W(\mu, \alpha) = | \alpha - \mu |$. If C_k denotes the class of procedures which are admissible when $W(\mu, \alpha) = | \alpha - \mu |^k$ then C_k is contained in the class C of Theorem 1. As demonstrated by Theorem 2, however, when k is sufficiently large the class C_k may reduce to a collection of procedures which virtually designate the same decision for all values of x, so in this case little significance could be attached to the mere fact that a procedure belonged to the class C. Since $W(\mu, \alpha) = | \alpha - \mu |^k$ is a conventional type of loss function for estimation problems Theorem 2 therefore raises a question of the practical importance of the class C; hence, it is of special interest that

THEOREM 3. *If $p(x, \mu)$ satisfies Conditions* 1 *and* 2 *and if the sample space M is finite then the class C_1 of procedures which are admissible relative to $W(\mu, \alpha) = | \alpha - \mu |$ is the class C itself.*

PROOF OF THEOREM 3. If a member δ of C is inadmissible then, since C is a complete class, there exists a member δ' of C which is better than δ. Then for all possible μ

$$(6) \qquad r(\mu, \delta) - r(\mu, \delta') = \sum_x p(x, \mu) \left(| \bar{\alpha}_x - \mu | - | \bar{\alpha}'_x - \mu |\right) \geqq 0.$$

Theorem 3 is proved by showing that (6) cannot hold for all possible μ; and, in particular, that there exists x in M such that either

$$r(\alpha_x, \delta) > r(\alpha_x, \delta') \quad \text{or} \quad r(\alpha_x + 1, \delta) > r(\alpha_x + 1, \delta').$$

Only the ordering of the sample space is pertinent so, without loss of generality, let $M = (0, 1, \cdots, n)$ and let $| \bar{\alpha}_x - \bar{\alpha}'_x | = A_x > 0$ for all x in M. Let

$$
\begin{aligned}
d &= \min_x(x \mid \bar{\alpha}_x < \bar{\alpha}'_x) \\
e &= \max_x(x \mid \bar{\alpha}_x > \bar{\alpha}'_x) \\
Y &= (x \mid d < x < e \text{ and } \bar{\alpha}'_{x-1} < \bar{\alpha}_{x-1} \leqq \bar{\alpha}_x < \bar{\alpha}'_x) \\
&= (y_1, y_2, \cdots, y_{m-1}),\ y_i < y_j \text{ for } i < j \\
Z &= (x \mid d < x < e \text{ and } \bar{\alpha}_{x-1} < \bar{\alpha}'_{x-1} \leqq \bar{\alpha}'_x < \bar{\alpha}_x) \\
&= (z_1, z_2, \cdots, z_m),\ z_i < z_j \text{ for } i < j
\end{aligned}
$$

then let $y_0 = d$, $y_m = e$, and

$$
\begin{aligned}
u_{2i} &= v_{2i-1} + 1 = y_i \text{ for } i = 0, 1, \cdots, m \\
u_{2i+1} &= v_{2i} + 1 = z_{i+1} \text{ for } i = 0, 1, \cdots, m-1 \\
\mu_{2i} &= \alpha_{u_{2i}} \text{ for } i = 0, 1, \cdots, m-1 \\
\mu_{2i+1} &= \alpha_{v_{2i+1}} + 1 \text{ (or } \alpha_{v_{2i+1}} \text{ if } \alpha_{v_{2i+1}} = \bar{\alpha}_{v_{2i+1}}) \\
&\qquad\qquad \text{for } i = 0, 1, \cdots, m-1
\end{aligned}
$$

Since δ and δ' are in C then $\bar{\alpha}_x \leqq \bar{\alpha}_y$ and $\bar{\alpha}'_x \leqq \bar{\alpha}'_y$ for all $x < y$ so

$$
\begin{aligned}
&u_{2i} \leqq v_{2i} < u_{2i+1} \leqq v_{2i+1} && \text{for } i = 0, 1, \cdots, m-1 \\
&\mu_0 < \mu_{2i-1} \leqq \mu_{2i} < \mu_{2i+1} && \text{for } i = 1, 2, \cdots, m-1
\end{aligned}
$$

and for $2k-1 \leqq q \leqq 2k$

$$
\begin{aligned}
r(\mu_q, \delta) - r(\mu_q, \delta') = &-\sum_{x=0}^{u_0-1} p(x, \mu_q)A_x + \sum_{i=0}^{2k-1} (-1)^i \sum_{x=u_i}^{v_i} p(x, \mu_q)A_x \\
&- \sum_{i=2k}^{2m-1} (-1)^i \sum_{x=u_i}^{v_i} p(x, \mu_q)A_x - \sum_{x=u_{2m}}^{n} p(x, \mu_q)A_x .
\end{aligned}
$$

Let

$$
\begin{aligned}
b_i(\mu_q) &= \sum_{x=u_i}^{v_i} p(x, \mu_q)A_x \\
B_{i,j}(\mu_q) &= \sum_{t=2i}^{2(i+j)-1} (-1)^t b_t(\mu_q)
\end{aligned}
$$

then, since $A_x > 0$ for all x,

$$r(\mu_q, \delta) - r(\mu_q, \delta') \leqq B_{0,k}(\mu_q) - B_{k,m-k}(\mu_q).$$

A contradiction to (6) is then obtained by showing that there exists a pair (k, q) such that $2k - 1 \leqq q \leqq 2k$ and $B_{0,k}(\mu_q) < B_{k,m-k}(\mu_q)$.

Let $S_i(k_j)$ be the following statement, defined for all integer pairs (i, j) such that $0 \leqq i \leqq m - j, 1 \leqq j \leqq m$.

$S_i(k_j)$: There exists an integer pair $(k_j, q(i, j))$ such that $0 \leqq k_j \leqq j$,

$$2(i + k_j) - 1 \leqq q(i, j) \leqq 2(i + k_j), \quad \text{and} \quad B_{i,k_j}(\mu_{q(i,j)}) < B_{i+k_j, j-k_j}(\mu_{q(i,j)}).$$

The negation of $S_i(k_j)$, written *not* $S_i(k_j)$, is then

not $S_i(k_j)$: For every integer pair (k, q) such that $0 \leqq k \leqq j$ and $2(i + k) - 1 \leqq q \leqq 2(i + k)$

$$B_{i,k}(\mu_q) \geqq B_{i+k,j-k}(\mu_q).$$

The desired contradiction to (6) may then be written $S_0(k_m)$.

The statement $S_i(k_1)$ is easily proved by contradiction. Note first that since δ belongs to C then

$$\begin{aligned} p(u_{2i}, \mu_{2i}) &> 0 \qquad \text{for } i = 0, 1, \cdots, m - 1 \\ p(v_{2i-1}, \mu_{2i-1}) &> 0 \qquad \text{for } i = 1, 2, \cdots, m \end{aligned} \tag{7}$$

so that

$$b_i(\mu_j) \begin{pmatrix} \geqq \\ > \end{pmatrix} 0 \text{ for } i \begin{pmatrix} \neq \\ = \end{pmatrix} j. \tag{8}$$

If $b_{2i+j}(\mu_{2i}) > 0$ then there exists $x_1 \geqq u_{2i+j}$ such that $p(x_1, \mu_{2i}) > 0$, and since $p(u_{2i}, \mu_{2i}) > 0$ then, by Lemma 1,

$$\text{if } b_{2i+j}(\mu_{2i}) > 0 \text{ then } p(x, \mu_{2i}) > 0 \text{ for } u_{2i} \leqq x \leqq u_{2i+j}. \tag{9a}$$

Similarly, if $b_{2i}(\mu_{2i+j}) > 0$ then there exists $x_0 \leqq v_{2i}$ such that $p(x_0, \mu_{2i+j}) > 0$. By (7), however, there exists $x_1 \geqq u_{2i+j}$ such that $p(x_1, \mu_{2i+j}) > 0$; hence, by Lemma 1,

$$\text{if } b_{2i}(\mu_{2i+j}) > 0 \text{ then } p(x, \mu_{2i+j}) > 0 \text{ for } v_{2i} \leqq x \leqq u_{2i+j}. \tag{9b}$$

It then follows that

$$\text{if } B_{i,1}(\mu_{2i}) \leqq 0 \text{ then } B_{i,1}(\mu_{2i+j}) \begin{pmatrix} < \\ \leqq \end{pmatrix} 0 \text{ for } j \begin{pmatrix} = \\ > \end{pmatrix} 1 \tag{10}$$

The statement in (10) is easily seen to hold for all $j \geqq 1$ such that either $b_{2i}(\mu_{2i+j}) = 0$ or $b_{2i+1}(\mu_{2i+j}) = 0$, for if $b_{2i}(\mu_{2i+j}) = 0$ then (8) implies (10) and if $b_{2i+1}(\mu_{2i+j}) = 0$ then $p(u_{2i+1}, \mu_{2i+j}) = 0$ but, by (7), there exists $x_1 > u_{2i+1}$ such that $p(x_1, \mu_{2i+j}) > 0$ so, by Lemma 1, $p(x, \mu_{2i+j}) = 0$ for all $x \leqq u_{2i+1}$ and, in particular, $b_{2i}(\mu_{2i+j}) = 0$. Now suppose that both $b_{2i}(\mu_{2i+j}) > 0$ and $b_{2i+1}(\mu_{2i+j}) > 0$ but that (10) does not hold; in particular, suppose $b_{2i+1}(\mu_{2i}) \geqq b_{2i}(\mu_{2i})$ and $b_{2i}(\mu_{2i+j}) \geqq b_{2i+1}(\mu_{2i+j})$. Since, by (8), $b_{2i}(\mu_{2i}) > 0$ then $b_{2i+1}(\mu_{2i}) > 0$ and, by (9a), $p(x, \mu_{2i}) > 0$ for $u_{2i} \leqq x \leqq u_{2i+1}$; and since $b_{2i}(\mu_{2i+j}) > 0$ then, by (9b), $p(x, \mu_{2i+j}) > 0$ for $v_{2i} \leqq x \leqq u_{2i+j}$. Then, by Condition 1,

$$\frac{b_{2i}(\mu_{2i})}{p(v_{2i}, \mu_{2i})} > \frac{b_{2i}(\mu_{2i+j})}{p(v_{2i}, \mu_{2i+j})};$$

but then the assumption that $b_{2i}(\mu_{2i+j}) \geqq b_{2i+1}(\mu_{2i+j})$ implies

$$\frac{a_{2i+1}(\mu_{2i})}{p(v_{2i}, \mu_{2i})} > \frac{b_{2i+1}(\mu_{2i+j})}{p(v_{2i}, \mu_{2i+j})},$$

which contradicts Condition 1. Hence (10) holds for all $j \geqq 1$. The statement *not* $S_i(k_1)$ implies that $B_{i,1}(\mu_{2i}) \leqq 0$ and $B_{i,1}(\mu_{2i+1}) \geqq 0$ and is therefore a contradiction of (10); hence, $S_i(k_1)$ for all i such that $0 \leqq i \leqq m - 1$.

Now suppose $S_i(k_j)$ for all (i, j) such that $0 \leqq i \leqq m - j$, $1 \leqq j < s \leqq m$ but *not* $S_h(k_s)$, where $0 \leqq h \leqq m - s$. Then $(k_1, q(h, 1))$ can be chosen as $(0, 2h)$; otherwise $B_{h,1}(\mu_{2h}) \leqq 0$ and, since $2h < 2(h + 1 + k_{s-1}) - 1 \leqq q(h + 1, s - 1)$, then, by (10), $B_{h,1}(\mu_{q(h+1,s-1)}) \leqq 0$ which, together with the assumption $S_{h+1}(k_{s-1})$ implies the contradiction $S_h(k_s)$.

If for all $j < g < s$, $(k_j, q(h, j))$ can be chosen as $(0, 2h)$ then $(k_g, q(h, g))$ can be chosen as $(0, 2h)$. Otherwise, $B_{h,g}(\mu_{2h}) \leqq 0$ and, since $S_{h+g}(k_{s-g})$ but *not* $S_h(k_s)$, $B_{h,g}(\mu_{q(h+g,s-g)}) > 0$. And since $p(u_{2h}, \mu_{2h}) > 0$ then $p(x, \mu_{2h}) > 0$ for $u_{2h} \leqq x \leqq v_{2(h+g)-1}$; otherwise, by Lemma 1, $p(x, \mu_{2h}) = 0$ for all $x \geqq v_{2(h+g)-1}$, and then $(k_{g-1}, q(h, g - 1)) = (0, 2h)$ implies that $(k_g, q(h, g))$ can be chosen as $(0, 2h)$. Also, $p(v_{2(h+g-1)}, \mu_{q(h+g,s-g)}) > 0$; otherwise, by Lemma 1, $p(x, \mu_{q(h+g,s-g)}) = 0$ for all $x \leqq v_{2(h+g-1)}$ since $v_{2(h+g-1)} < v_{2(h+g+k_{s-u})-1} < u_{2(h+g+k_{s-u})}$, and then $B_{h,g}(\mu_{q(h+g,s-g)}) \leqq 0$ to contradict *not* $S_h(k_s)$. Hence,

$$-\frac{B_{h+g-1,1}(\mu_{2h})}{p(v_{2(h+g-1)}, \mu_{2h})} \geqq \frac{B_{h,g-1}(\mu_{2h})}{p(v_{2(h+g-1)}, \mu_{2h})} \tag{11}$$

and

$$\frac{B_{h,g-1}(\mu_{q(h+g,s-g)})}{p(v_{2(h+g-1)}, \mu_{q(h+g,s-g)})} \geqq -\frac{B_{h+g-1,1}(\mu_{q(h+g,s-g)})}{p(v_{2(h+g-1)}, \mu_{q(h+g,s-g)})}. \tag{12}$$

Observe, however, that if

$$\frac{B_{h,j}(\mu_{2h})}{p(v_{2(h+j)}, \mu_{2h})} > \frac{B_{h,j}(\mu_q)}{p(v_{2(h+j)}, \mu_q)} \tag{13}$$

where $1 \leqq j \leqq g - 1$, $2h < q$, and $p(x, \mu_q) > 0$ for $v_{2(h+j)} \leqq x \leqq v_{2(h+g)}$ then, by Condition 1,

$$\frac{B_{h+j,1}(\mu_{2h})}{p(v_{2(h+j)}, \mu_{2h})} > \frac{B_{h+j,1}(\mu_q)}{p(v_{2(h+j)}, \mu_q)} \tag{14}$$

and, since $B_{h,j}(\mu) + B_{h+j,1}(\mu) = B_{h,j+1}(\mu)$,

$$\frac{B_{h,j+1}(\mu_{2h})}{p(v_{2(h+j)}, \mu_{2h})} > \frac{B_{h,j+1}(\mu_q)}{p(v_{2(h+j)}, \mu_q)}. \tag{15}$$

Since $(k_{j+1}, q(h, j + 1))$ can be chosen as $(0, 2h)$ then $B_{h,j+1}(\mu_{2h}) > 0$, and since, by Condition 1, $p(v_{2(h+j)}, \mu_{2h})p(v_{2(h+j+1)}, \mu_q) > p(v_{2(h+j)}, \mu_q)p(v_{2(h+j+1)}, \mu_{2h})$, then

$$\begin{aligned} \frac{B_{h,j+1}(\mu_{2h})}{p(v_{2(h+j+1)}, \mu_{2h})} &= \frac{p(v_{2(h+j)}, \mu_{2h})}{p(v_{2(h+j+1)}, \mu_{2h})} \cdot \frac{B_{h,j+1}(\mu_{2h})}{p(v_{2(h+j)}, \mu_{2h})} \\ &> \frac{p(v_{2(h+j)}, \mu_q)}{p(v_{2(h+j+1)}, \mu_q)} \cdot \frac{B_{h,j+1}(\mu_q)}{p(v_{2(h+j)}, \mu_q)} = \frac{B_{h,j+1}(\mu_q)}{p(v_{2(h+j+1)}, \mu_q)}. \end{aligned} \tag{16}$$

Thus, if (13) then (16). Now let j' be the least j, $1 \leqq j \leqq g - 1$, such that $p(v_{2(h+j)}, \mu_{q(h+g,s-g)}) > 0$. Then (13) holds for $j = j'$ and $q = q(h + g, s - g)$, for if $j' = 1$ and $p(v_{2h}, \mu_{q(h+g,s-g)}) > 0$, then let $j = 0$ in (14), (15), (16) to get the desired result; otherwise, if $j' \geqq 1$ and $p(v_{2h}, \mu_{q(h+g,s-g)}) = 0$ then the right side of (13) is nonpositive while the left side is positive since $(k_{j'}, q(h, j'))$ can be chosen as $(0, 2h)$. Hence, by finite induction, (13) holds for $j = g - 1$, $q = q(h + g, s - g)$; i.e.,

$$\frac{B_{h,g-1}(\mu_{2h})}{p(v_{2(h+g-1)}, \mu_{2h})} > \frac{B_{h,g-1}(\mu_{q(h+g,s-g)})}{p(v_{2(h+g-1)}, \mu_{q(h+g,s-g)})}.$$

Hence, by (11) and (12),

$$-\frac{B_{h+g-1,1}(\mu_{2h})}{p(v_{2(h+g-1)}, \mu_{2h})} > -\frac{B_{h+g-1,1}(\mu_{q(h+g,s-g)})}{p(v_{2(h+g-1)}, \mu_{q(h+g,s-g)})}$$

in contradiction to Condition 1. This proves that if $S_i(k_j)$ for all (i, j) such that $0 \leqq i \leqq m - j$, $1 \leqq j < s \leqq m$ but *not* $S_h(k_s)$ then $(k_j, q(h, j))$ can be chosen as $(0, 2h)$ for all j such that $1 \leqq j < s < m$.

With this result, however, simply take $j = s - 1$, $q = 2(h + s) - 1$ in (13) to get

$$\frac{B_{h,s-1}(\mu_{2h})}{p(v_{2(h+s-1)}, \mu_{2h})} > \frac{B_{h,s-1}(\mu_{2(h+s)-1})}{p(v_{2(h+s-1)}, \mu_{2(h+s)-1})}.$$

The denominator $p(v_{2(h+s-1)}, \mu_{2(h+s)-1})$ must be positive; otherwise, since $p(v_{2(h+s)-1}, \mu_{2(h+s)-1}) > 0$ then $p(x, \mu_{2(h+s)-1}) = 0$ for all $x \leqq v_{2(h+s-1)}$ so that $B_{h,s}(\mu_{2(h+s)-1}) = -b_{2(h+s)-1}(\mu_{2(h+s)-1}) < 0$ to contradict the assumption *not* $S_h(k_s)$. Then *not* $S_h(k_s)$ gives, as before,

$$-\frac{B_{h+s-1,1}(\mu_{2h})}{p(v_{2(h+s-1)}, \mu_{2h})} \geqq \frac{B_{h,s-1}(\mu_{2h})}{p(v_{2(h+s-1)}, \mu_{2h})}$$
$$\geqq \frac{B_{h,s-1}(\mu_{2(h+s)-1})}{p(v_{2(h+s-1)}, \mu_{2(h+s)-1})} \geqq \frac{B_{h+s-1,1}(\mu_{2(h+s)-1})}{p(v_{2(h+s-1)}, \mu_{2(h+s)-1})}$$

to contradict Condition 1. Hence, $S_h(k_s)$. And since $S_i(k_1)$ for all i such that $0 \leqq i \leqq m - 1$ then, by finite induction, $S_i(k_j)$ for all (i, j) such that $0 \leqq i \leqq m - j$, $1 \leqq j \leqq m$. In particular, $S_0(k_m)$, which establishes Theorem 3.

6. Minimax procedures when $W(\mu, \alpha) = |\alpha - \mu|$. When the minimax estimator does not have constant risk, as is obviously the general case here, then the Bayes method of finding the minimax procedure by guessing a least favorable a priori distribution becomes extremely difficult, if not hopeless. For distributions of the type considered here, however, it is possible to reduce the problem of guessing a least favorable a priori distribution to one of guessing which points in the parameter space are assigned positive probability by a least favorable a priori distribution.

THEOREM 4. *If* $p(x, \mu)$ *satisfies Conditions* 1, 2, *and* 3 *and* $W(\mu, \alpha) = |\alpha - \mu|$ *then there exists a least favorable a priori distribution which assigns positive probability to at most* $n + 2$ *values of* μ, $\mu_0 \leqq \mu_1 \leqq \cdots \leqq \mu_{n+1}$, *and if* δ *is a Bayes* *solution with respect to a least favorable priori distribution then* $\mu_i \leqq \bar{\alpha}_{x_i} \leqq \mu_{i+1}$ *for* $i = 0, 1, \cdots, n$.

PROOF OF THEOREM 4. Assume, without loss of generality, that $M = (0, 1, \cdots, n)$. Let

$$r_x(\xi, \alpha) = \sum_{\mu=0}^{N} | \alpha - \mu | p(x, \mu)\xi_\mu$$

and let a_x^ξ be the collection of integers

$$a_x^\xi = (\alpha, 0 \leqq \alpha \leqq N \mid \inf_\beta r_x(\xi, \beta) = r_x(\xi, \alpha).$$

From the proof of Theorem 2 the function $r_x(\xi, \alpha)$ has the properties

I′: if $r_x(\xi, \alpha) \leqq r_x(\xi, \alpha + 1)$ then $r_x(\xi, \alpha + \beta) \leqq r_x(\xi, \alpha + \beta + 1)$ for all $\beta \geqq 0$

II: if $r_y(\xi, \alpha) \leqq r_y(\xi, \alpha + 1)$ then $r_x(\xi, \alpha) < r_x(\xi, \alpha + 1)$ for all $x \leqq y$

Hence, a_x^ξ has the form

$$a_x^\xi = (\alpha_x^\xi, \alpha_x^\xi + 1, \cdots, \alpha_x^\xi + \beta_x^\xi)$$

where $0 \leqq \alpha_x^\xi \leqq N$, $0 \leqq \beta_x^\xi \leqq N - \alpha_x^\xi$, and $\alpha_x^\xi + \beta_x^\xi \leqq \alpha_{x+1}^\xi$. Furthermore, since $r_x(\xi, \alpha_x^\xi) = r_x(\xi, \alpha_x^\xi + 1) = \cdots = r_x(\xi, \alpha_x^\xi + \beta_x^\xi)$ or, for $i = 1, \cdots, \beta_x^\xi - 1$,

$$p(x, \alpha_x^\xi + i)\xi_{\alpha_x^\xi+i} + \sum_{\mu=\alpha_x^\xi+i+1}^{N} p(x, \mu)\xi_\mu = \sum_{\mu=0}^{\alpha_x^\xi+i-1} p(x, \mu)\xi_\mu$$

and

$$\sum_{\mu=\alpha_x^\xi+i+1}^{N} p(x, \mu)\xi_\mu = \sum_{\mu=0}^{\alpha_x^\xi+i-1} p(x, \mu)\xi_\mu + p(x, \alpha_x^\xi + i)\xi_{\alpha_x^\xi+i}$$

then $p(x, \alpha_x^\xi + i)\xi_{\alpha_x^\xi+i} = 0$ for $0 < i < \beta_x$. Hence, since $r_x(\xi, \alpha_x^\xi - 1) > r_x(\xi, \alpha_x^\xi)$, or

$$p(x, \alpha_x^\xi)\xi_{\alpha_x^\xi} + p(x, \alpha_x^\xi + \beta_x^\xi)\xi_{\alpha_x^\xi+\beta_x^\xi} + \sum_{\mu=\alpha_x^\xi+\beta_x^\xi-1}^{N} p(x, \mu)\xi_\mu > \sum_{u=0}^{\alpha_x^\xi-1} p\,(x, \mu)\xi_\mu$$

and $r_x(\xi, \alpha_x^\xi + \beta_x^\xi) < r_x(\xi, \alpha_x^\xi + \beta_x^\xi + 1)$, or

$$\sum_{\mu=\alpha_x^\xi+\beta_x^\xi+1}^{N} p(x, \mu)\xi_\mu < \sum_{\mu=0}^{\alpha_x^\xi-1} p(x, \mu)\xi_\mu + p(x, \alpha_x^\xi)\xi_{\alpha_x^\xi} + p(x, \alpha_x^\xi + \beta_x^\xi)\xi_{\alpha_x^\xi+\beta_x^\xi},$$

then $p(x, \alpha_x^\xi)\xi_{\alpha_x^\xi} > 0$ and $p(x, \alpha_x^\xi + \beta_x^\xi)\xi_{\alpha_x^\xi+\beta_x^\xi} > 0$. Therefore, since $p(x, \alpha_x^\xi) > 0$ and $p(x, \alpha_x^\xi + \beta_x^\xi) > 0$ imply, by Lemma 1, that $p(x, \alpha_x^\xi + i) > 0$ for $0 \leqq i \leqq \beta_x^\xi$, then

$$\xi_{\alpha_x^\xi+i} \begin{cases} > 0 \text{ for } i = 0 \\ = 0 \text{ for } 0 < i < \beta_x^\xi \\ > 0 \text{ for } i = \beta_x^\xi. \end{cases} \tag{17}$$

If ξ^0 is a least favorable a priori distribution; i.e., if ξ^0 maximizes $\inf_\delta r(\xi, \delta)$, then since, for a fixed δ, $r(\xi, \delta)$ is linear in ξ, every ξ which satisfies

$$(18)\quad r_x(\xi, \alpha_x^{\xi^0} - 1) \geqq r_x(\xi, \alpha_x^{\xi^0}) = \cdots = r_x(\xi, \alpha_x^{\xi^0} + \beta_x^{\xi^0}) \leqq r_x(\xi, \alpha_x^{\xi^0} + \beta_x^{\xi^0} + 1)$$

for $x = 0, 1, \cdots, n$ is likewise a least favorable a priori distribution. But since $p(x, \mu)$ satisfies Conditions 1, 2, and 3, and, for every $x > 0$, $p(x - 1, \alpha_{x-1}^{\xi^0} + \beta_{x-1}^{\xi^0}) > 0$ and $p(x, \alpha_x^{\xi^0}) > 0$, where $\alpha_{x-1}^{\xi^0} + \beta_{x-1}^{\xi^0} \leqq \alpha_x^{\xi^0}$, then for every $x > 0$ there exists an integer μ_x such that $\alpha_{x-1}^{\xi^0} + \beta_{x-1}^{\xi^0} \leqq \mu_x \leqq \alpha_x^{\xi^0}$ and both $p(x - 1, \mu_x) > 0$ and $p(x, \mu_x) > 0$. Let (μ_x), $x = 1, 2, \cdots, n$, $\mu_x \leqq \mu_{x+1}$, be a sequence of such integers and define $\mu_0 = \alpha_0^{\xi^0}$ and $\mu_{n+1} = \alpha_n^{\xi^0} + \beta_n^{\xi^0}$. Then every ξ which satisfies

$$(19)\qquad r_x(\xi, \mu_x) = r_x(\xi, \mu_x + 1) = \cdots = r_x(\xi, \mu_{x+1})$$

for $x = 0, 1, \cdots, n$ also satisfies (18) and has $\xi_\mu = 0$ for $\mu_x < \mu < \mu_{x+1}$ for $x = 0, 1, \cdots, n$.

It remains, then, to show that a solution ξ' to (19) exists and may be chosen so that $\xi_\mu' = 0$ for $\mu < \mu_0$ and for $\mu > \mu_{n+1}$. This, however, follows directly from Theorem 3, for the problem of proving the existence of such a ξ' is easily seen to reduce to the problem of proving that a set of equations of the form

$$\sum_{\mu=0}^{x} p'(x, \mu)\xi_\mu = \sum_{\mu=x+1}^{m} p'(x, \mu)\xi_\mu, \; x = 0, 1, \cdots, m - 1$$

where $p'(x, \mu)$ satisfies Conditions 1 and 2 for $x = 0, 1, \cdots, n \geqq m - 1$, $\mu = 0, 1, \cdots, m$, and $p'(x, x) > 0$ and $p'(x, x + 1) > 0$, has a solution $\xi = (\xi_\mu)$ such that $\xi_\mu > 0$, $\mu = 0, 1, \cdots, m$, and $\sum_{\mu=0}^{m} \xi_\mu = 1$, and this may be viewed as a special case of Theorem 3 with $N = m$ and $n \geqq m - 1$. Theorem 3 then asserts that a procedure δ with $\delta_x(x) + \delta_{x+1}(x) = 1$, $\delta_z(x) < 1$, for $x = 0, 1, \cdots, m - 1$ and $\delta_m(x) = 1$ for $x \geqq m$ is admissible, and therefore δ is a Bayes solution relative to some a priori distribution ξ and, by (17), $\xi_\mu > 0$ for $\mu = 0, 1, \cdots, m$. Hence, a ξ' of the desired form exists and the theorem is established.

The construction of the minimax procedure δ^0 is easily accomplished once the integer μ_x is known for every x. δ^0 is defined by $(\bar{\alpha}_0^0, \bar{\alpha}_1^0, \cdots, \bar{\alpha}_n^0)$ which is uniquely determined by the equations

$$r(\mu_x, \delta^0) = \sum_{y=0}^{x-1} p(y, \mu_x)(\mu_x - \bar{\alpha}_y^0) + \sum_{y=x}^{n} p(y, \mu_x)(\bar{\alpha}_y^0 - \mu) = r(\mu_0, \delta^0)$$

7. Discussion. The requirement that an estimator of an integer-valued parameter must itself be integer-valued is almost a logical necessity in any rigorous approach to the estimation problem. For practical purposes, of course, such a requirement has been regarded as an unnecessary refinement, and statisticians conventionally estimate an integer-valued parameter by means of a real-valued statistic, presenting as their estimate either the real number itself or the nearest integer. The problem is frequently encountered, for example, in such a form that the statistician wishes to present an estimate of the fraction μ/N. Certainly, division by the known constant N is a trivial alteration of the estimation problem; it would be unheard of, however, to require in this case that the estimate assume one of the values $0/N$, $1/N$, $\cdots$, N/N.

If real-valued procedures are allowed then when loss is absolute error the randomized, integer-valued procedure δ is equivalent to the non-randomized procedure which estimates the real number $\bar{a}_x$ when x is observed. Any optimum property ascribed to an integer-valued procedure therefore applies to its real-valued counterpart so, as a corollary to Theorem 3, when real-valued procedures are allowed then the class of non-randomized real-valued procedures derived from the class C in the above manner is a minimal essentially complete class. Likewise, if δ^0 is the minimax integer-valued procedure then the non-randomized real-valued procedure $\bar{a}_x^0$ is also minimax. Theorems 3 and 4 thus remain essentially unaffected by the introduction of real-valued procedures.

REFERENCES

[1] H. RUBIN, "A complete class of decision procedures for distributions with monotone likelihood ratio," Technical Report 9, Dept. of Stat., Stanford University, October 31, 1951.

[2] S. KARLIN AND H. RUBIN, "The theory of decision procedure for distributions with monotone likelihood ratio." *Ann. Math. Stat.*, Vol. 27 (1956), pp. 272–299.