The Annals of Statistics
1974, Vol. 2, No. 4, 822-828

ON THE ADMISSIBILITY OF THE M.L.E. FOR ORDERED BINOMIAL PARAMETERS

BY H. SACKROWITZ AND W. STRAWDERMAN

Rutgers—The State University

Admissibility properties of the M.L.E. for the parameters of m independent binomial distributions (when these parameters are known to be ordered) are determined for certain convex loss functions. It is shown that, except in two special cases, the M.L.E. is inadmissible whenever the total sample size is 7 or more.

1. Introduction. Let $X_1, \cdots, X_m$ be independent random variables with X_i having a binomial distribution with parameters θ_i and n_i, $i = 1, \cdots, m$. The problem is to estimate $\theta_1, \cdots, \theta_m$ where $n_1, \cdots, n_m$ is known, and in addition, it is assumed that $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_m$. Such situations occur in reliability ([4]), in the biological sciences ([6] and [8]), and in other disciplines as well.

In reliability studies, for example, testing is often carried out in stages, and improvements are made after each testing stage to correct problems found during testing. Presumably, therefore, the reliability (or probability of successful operation) should increase from stage to stage. This and related problems (i.e. estimating ordered parameters for other distributions) have received considerable attention in the literature (see for example [1], [3] and [5] and references therein as well as the recent book by Barlow, *et al.* [2]).

We are primarily concerned in this work with the question of admissibility of the Maximum Likelihood Estimator (M.L.E.) relative to loss functions of the form $L(\theta_1, \cdots, \theta_m; a_1, \cdots, a_m) = \sum_{i=1}^m W(|\theta_i - a_i|)$ where $W(\cdot)$ is strictly convex. Our main result (Theorem 4.1) states that, except in two special cases, the M.L.E. is inadmissible if $n_1 + \cdots + n_m \geqq 7$. We also show for the case of sum of squared error loss that the M.L.E. is admissible in all cases not covered by the above. This is in marked contrast to the unordered case where the M.L.E. is admissible for all $(n_1, \cdots, n_m)$ (see, for example, Johnson [9]). The method of proof is to show that the M.L.E. is not admissible in a suitably chosen sub-problem by showing that it cannot be Bayes in that sub-problem. This method suffers from the shortcoming that no new estimator is exhibited which beats the M.L.E. The situation is somewhat analogous to the problem of estimating ordered normal means with sum of squared error loss. There, the M.L.E. is inadmissible if there are two or more means to be estimated since the estimator is not smooth and hence cannot be generalized Bayes (see Sacks [10] for a proof that the generalized Bayes procedures are a complete class).

Received December 1972; revised September 1973.

AMS 1970 *subject classifications.* Primary 62C15; Secondary 62C10.

Key words and phrases. Maximum likelihood estimate, admissibility, binomial distribution, Bayes, ordered parameters, convex loss.

However, to the author's knowledge there is no estimator which is known to beat the M.L.E.

In Section 2 we formally state the problem, as well as introduce some notation and definitions. Section 3 contains general conditions for admissibility or inadmissibility. These results are applied in Section 4 to obtain the main results.

2. Notation and definitions. We will introduce some notation and put the problem in the framework of statistical decision theory.

Any letter, either Roman or Greek, capital or lower case, which is boldface denotes a row vector. The coordinates of the vector are denoted by the same letter but using subscripts. For example, $\mathbf{X} = (X_1, \cdots, X_m)$ and $\mathbf{x} = (x_1, \cdots, x_m)$. All vectors are m-dimensional.

Let $X_1, \cdots, X_m$ be mutually independent random variables. We construct the following statistical decision games about these random variables (the notation for sample space, action space, etc. is essentially that of Ferguson [7]).

$$\begin{aligned}
\mathscr{G}\colon \mathscr{X} &= \{\mathbf{x} = (x_1, \cdots, x_m) : x_i = 0, \cdots, n_i, i = 1, \cdots, m\} \\
\Theta &= \{\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m) : 0 \leqq \theta_1 \leqq \cdots \theta_m \leqq 1\} \\
\mathscr{A} &= \Theta \\
L(\boldsymbol{\theta}, \mathbf{a}) &= \textstyle\sum_{i=1}^m W(|\theta_i - a_i|) \qquad \text{all} \quad \boldsymbol{\theta} \in \Theta, \quad \mathbf{a} \in \mathscr{A} \\
P(\mathbf{X} = \mathbf{x} \mid \boldsymbol{\theta}) &= \textstyle\prod_{i=1}^m \binom{n_i}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i}, \qquad \mathbf{x} \in \mathscr{X}, \boldsymbol{\theta} \in \Theta .
\end{aligned}$$

$$\begin{aligned}
\mathscr{G}_* \colon \mathscr{X}_* &= \{\mathbf{x} : \mathbf{x} \in \mathscr{X}, x_m = n_m\} \\
\Theta_* &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \Theta, \theta_m = 1\} \\
\mathscr{A}_* &= \{\mathbf{a} : \mathbf{a} \in \mathscr{A}, a_m = 1\} \\
L_*(\boldsymbol{\theta}, \mathbf{a}) &= L(\boldsymbol{\theta}, \mathbf{a}) \qquad \text{all} \quad \boldsymbol{\theta} \in \Theta_*, \quad \mathbf{a} \in \mathscr{A}_* \\
P_*(\mathbf{X} = \mathbf{x} \mid \boldsymbol{\theta}) &= \textstyle\prod_{i=1}^{m-1} \binom{n_i}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i}, \qquad \mathbf{x} \in \mathscr{X}_*, \theta \in \Theta_* .
\end{aligned}$$

$$\begin{aligned}
\mathscr{G}^* \colon \mathscr{X}^* &= \{\mathbf{x} : \mathbf{x} \in \mathscr{X}, x_1 = 0\} \\
\Theta^* &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \Theta, \theta_1 = 0\} \\
\mathscr{A}^* &= \{\mathbf{a} : \mathbf{a} \in \mathscr{A}, a_1 = 0\} \\
L^*(\boldsymbol{\theta}, \mathbf{a}) &= L(\boldsymbol{\theta}, \mathbf{a}) \qquad \text{all} \quad \theta \in \Theta^*, \quad \mathbf{a} \in \mathscr{A}^* \\
P^*(\mathbf{X} = \mathbf{x} \mid \boldsymbol{\theta}) &= \textstyle\prod_{i=2}^m \binom{n_i}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i}, \qquad \mathbf{x} \in \mathscr{X}^*, \boldsymbol{\theta} \in \Theta^* .
\end{aligned}$$

In addition we define the game $\mathscr{G}_0$ as follows.

$$\begin{aligned}
\mathscr{G}_0 \colon \mathscr{X}_0 &= \{\mathbf{x} : \mathbf{x} \in \mathscr{X}, x_1 \neq 0, x_m \neq n_m\} \\
\Theta_0 &= \Theta \\
\mathscr{A}_0 &= \mathscr{A} \\
L_0(\boldsymbol{\theta}, \mathbf{a}) &= L(\boldsymbol{\theta}, \mathbf{a}) \qquad \text{all} \quad \theta \in \Theta_0, \quad \mathbf{a} \in \mathscr{A}_0 .
\end{aligned}$$

$$\begin{aligned}
P_0(\mathbf{X} = \mathbf{x} \mid \boldsymbol{\theta}) = {} & \binom{n_1}{x_1} \theta_1^{x_1 - 1} (1 - \theta_1)^{n_1 - x_1} \binom{n_m}{x_m} \theta_m^{x_m} (1 - \theta_m)^{n_m - x_m - 1} \\
& \times \textstyle\prod_{i=2}^{m-1} \binom{n_i}{x_i} \theta_1^{x_i} (1 - \theta_i)^{n_i - x_i} \\
& \times [\textstyle\sum_{i=0}^{n_1 - 1} (1 - \theta_1)^i]^{-1} [\textstyle\sum_{i=0}^{n_m - 1} \theta_m^{\,i}]^{-1}, \qquad \mathbf{x} \in \mathscr{X}_0, \theta \in \Theta_0 .
\end{aligned}$$

Note that P_0 is just the conditional distribution of $\mathbf{X}$ given $(X_1 \neq 0, X_m \neq n_m)$.

In each of the above games a (non-randomized) decision rule or estimator is a vector-valued function of the observations and will be denoted, for example, as follows: $\mathbf{D}(\mathbf{x}) = (D_1(\mathbf{x}), \cdots, D_m(\mathbf{x}))$. Although the same notation applies to all decision rules, the domain and range of these rules will determine their appropriateness for the individual games (for example, a rule $\mathbf{D}$ for the game $\mathscr{G}$ may also be used as a rule for the game $\mathscr{G}_0$ as $\mathscr{X}_0 \subset \mathscr{X}$ and $\mathscr{A}_0 = \mathscr{A}$). We will use the notation $R(\boldsymbol{\theta}, \mathbf{D})$, $R_*(\boldsymbol{\theta}, \mathbf{D})$, $R^*(\boldsymbol{\theta}, \mathbf{D})$, $R_0(\boldsymbol{\theta}, \mathbf{D})$ to denote the risk function of a decision rule $\mathbf{D}$ when used for the respective games $\mathscr{G}$, $\mathscr{G}_*$, $\mathscr{G}^*$, and $\mathscr{G}_0$. Of course use of this notation will imply that the $\boldsymbol{\theta}$ value belongs to the appropriate parameter space.

We will always assume the function $W(\cdot)$ to have the following properties:

(2.1)
(i) $W(\cdot)$ is a strictly convex function.
(ii) $W(\cdot) \geqq 0$ with $W(0) = 0$.
(iii) The derivative of $W(y)$ with respect to y (denoted by $W'(y)$) exists for all $0 \leqq y \leqq 1$ and $W'(1) < \infty$, $W'(0) = 0$.

Let the decision rule $\mathbf{T}(\mathbf{x})$, defined on $\mathbf{x} \in \mathscr{X}$, denote the maximum likelihood estimator (as derived in [1]) for $\boldsymbol{\theta} \in \Theta$. Denote by $\mathbf{T}_*$, $\mathbf{T}^*$, and $\mathbf{T}_0$, respectively, the restrictions of $\mathbf{T}$ to $\mathscr{X}^*$, $\mathscr{X}_*$, and $\mathscr{X}_0$. We note the following properties of these estimators:

(2.2)
(i) $T_1(x) = 0$ all $\mathbf{x} \in \mathscr{X}^*$,
(ii) $\mathbf{T}_*(\mathbf{x})$ is the maximum likelihood estimator for $\boldsymbol{\theta} \in \Theta_*$,
(iii) $T_m(x) = 1$ all $\mathbf{x} \in \mathscr{X}_*$,
(iv) $\mathbf{T}^*(x)$ is the maximum likelihood estimator for $\boldsymbol{\theta} \in \Theta^*$.

3. Conditions for admissibility. We immediately state the following result (as Theorem 3.1) the proof of which is essentially given in [9], page 1579.

THEOREM 3.1. $\mathbf{T}(\mathbf{X})$ *is admissible for* $\mathscr{G}$ *if and only if*

(i) $\mathbf{T}_*$ *is admissible for* $\mathscr{G}_*$,
(ii) $\mathbf{T}^*$ *is admissible for* $\mathscr{G}^*$,
(iii) $\mathbf{T}_0$ *is admissible for* $\mathscr{G}_0$.

LEMMA 3.1. *If* $\mathbf{T}$, *restricted to* $\mathscr{X}_0$, *is not Bayes for* $\mathscr{G}_0$, *then* $\mathbf{T}$ *is inadmissible for* $\mathscr{G}$.

PROOF. By Wald ([11], Theorem 3.20) the Bayes rules comprise a complete class for $\mathscr{G}_0$. This with Theorem 3.1 establishes the result.

Our main tool for showing inadmissibility will be Theorem 3.2 stated below. Following the lead of Lemma 3.1, Theorem 3.2 will give a method for deciding that T, restricted to $\mathscr{X}_0$, is not Bayes for $\mathscr{G}_0$. Basically Theorem 3.2 reflects the notion that an admissible rule must exhibit a certain amount of consistency on $\mathscr{X}_0$ (i.e. there are points in $\mathscr{X}_0$ on which an admissible rule must behave in a similar fashion as to whether or not the estimate of θ_1 is equal to the estimate of θ_m). This consistency on $\mathscr{X}_0$ will be expressed through the functions $B(\cdot)$ and $U(\cdot)$ defined on $\mathscr{X}_0$ as follows:

$$\begin{aligned} B(\mathbf{x}) &= +\infty \quad \text{if} \quad x_1 = 1\,, \quad x_2 = \cdots = x_m = 0 \\ &= 1 \quad \text{if} \quad x_1 \geqq 2 \\ &= \text{the first} \quad i > 1 \quad \text{such that} \quad x_i \geqq 1 \quad \text{if otherwise.} \end{aligned}$$

$$\begin{aligned} U(\mathbf{x}) &= -\infty \quad \text{if} \quad x_m = n_m - 1\,, \quad x_j = n_j \quad \text{all } j = 1, \cdots, m-1 \\ &= m \quad \text{if} \quad x_m \leqq n_m - 2 \\ &= \text{the last} \quad i < m \quad \text{such that} \quad x_i \leqq n_i - 1 \quad \text{if otherwise.} \end{aligned}$$

Intuitively, $B(\mathbf{x})$ is the first index i such that x_i is greater than its minimum value on $\mathscr{X}_0$ and $U(\mathbf{x})$ is the last index i such that x_i is less than its maximum on $\mathscr{X}_0$.

THEOREM 3.2. *If there exist two points y and z in $\mathscr{X}_0$ such that*

$$\text{(a)} \quad T_1(\mathbf{y}) = T_m(\mathbf{y}) \quad \textit{and} \quad T_1(\mathbf{z}) \neq T_m(\mathbf{z})$$

and

$$\text{(b)} \quad B(\mathbf{y}) \geqq B(\mathbf{z}) \quad \textit{and} \quad U(\mathbf{y}) \leqq U(\mathbf{z})\,,$$

then $\mathbf{T}(\mathbf{X})$ *is inadmissible for* $\mathscr{G}$.

PROOF. We will simply sketch the proof, as the method of proof is straightforward (the details are tedious though not difficult). The idea is to show that conditions (a) and (b) of this theorem imply that $\mathbf{T}_0$ cannot be Bayes for $\mathscr{G}_0$. The arguments are basically as follows. For $\mathbf{T}_0$ to be Bayes with respect to the prior distribution τ means

$$\int_\Theta W(|\theta_i - T_i(\mathbf{x})|) f(\mathbf{x}, \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) = \inf_{0 \leq a \leq 1} \int_\Theta W(|\theta_i - a|) f(\mathbf{x}; \theta)\, d\tau(\boldsymbol{\theta})$$

for all $i = 1, \cdots, m$ and all $\mathbf{x} \in \mathscr{X}_0$ and where

$$f(\mathbf{x}; \boldsymbol{\theta}) = \theta_1^{x_1 - 1}(1 - \theta_1)^{n_1 - x_1}\theta_m^{x_m}(1 - \theta_m)^{n_m - x_m - 1} \prod_{i=2}^{m-1} \theta_i^{x_i}(1 - \theta_i)^{n_i - x_i} \times [\sum_{i=0}^{n_1 - 1} (1 - \theta_1)^i]^{-1}[\sum_{i=0}^{n_m - 1} \theta_m^{\,i}]^{-1}\,.$$

This implies (with conditions (2.1) allowing the change in order of integration and differentiation)

$$\int_\Theta V(\theta_i, T_i(\mathbf{x})) f(\mathbf{x}; \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) = 0 \tag{3.1}$$

for all $i = 1, \cdots, m$ and all $\mathbf{x} \in \mathscr{X}_0$ and where $V(\theta, a)$ is the derivative of $W(|\theta - a|)$ with respect to a. Two properties of $V(\theta, a)$ which are due to the strict convexity of $W(\cdot)$ are

(i) $V(\theta, a)$ is a strictly increasing function in a;
(ii) $V(\theta, a)$ is a strictly decreasing function in θ.

Since each integral of the form (3.1) is equal zero, the difference of any two must equal zero. Thus

$$\int_\Theta [V(\theta_i, T_i(\mathbf{x})) - V(\theta_j, T_j(\mathbf{x}))] f(\mathbf{x}; \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) = 0 \tag{3.2}$$

all $i, j = 1, \cdots, m$ and all $\mathbf{x} \in \mathscr{X}_0$. In particular, we take $i = m, j = 1, \mathbf{x} = \mathbf{y}$. By assumption (a) of the theorem, (3.2) becomes

$$\int_\Theta [V(\theta_m, T_1(\mathbf{y})) - V(\theta_1, T_1(\mathbf{y}))] f(\mathbf{y}, \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) = 0 . \tag{3.3}$$

Since $P_\tau(\theta_1 \leqq \cdots \leqq \theta_m) = 1$, by the monotonicity properties of $V(\theta, a)$, if $\theta_m > \theta_1$ then

$$V(\theta_m, T_1(y)) - V(\theta_1, T_1(y)) < 0 .$$

It now follows from (3.3) that τ must be such that

$$[V(\theta_m, T_1(y)) - V(\theta_1, T_1(y))] f(\mathbf{y}, \boldsymbol{\theta}) = 0$$

with τ probability one. Equivalently we have

$$P(\theta_1 = \theta_m \text{ or } f(\mathbf{y}, \boldsymbol{\theta}) = 0) = 1 . \tag{3.4}$$

By the hypotheses of this theorem the point $\mathbf{z}$ is such that $B(\mathbf{y}) \geqq B(\mathbf{z})$ and $U(\mathbf{y}) \leqq U(\mathbf{z})$. Thus (for $\theta_1 \leqq \cdots \leqq \theta_m$)

$$f(\mathbf{y}, \boldsymbol{\theta}) = 0 \Rightarrow f(\mathbf{z}, \boldsymbol{\theta}) = 0$$

for any $\boldsymbol{\theta} \in \Theta$. We now see that equation (3.4) implies $P(\theta_1 = \theta_m \text{ or } f(\mathbf{z}, \boldsymbol{\theta}) = 0) = 1$ which in turn implies

$$\int_\Theta [V(\theta_m, T_1(\mathbf{z})) - V(\theta_1, T_1(\mathbf{z}))] f(\mathbf{z}, \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) = 0$$

which implies (by (3.1) with $\mathbf{x} = \mathbf{z}$ and $i = 1$ and then $i = m$)

$$\begin{aligned} 0 &= \int_\Theta V(\theta_m, T_1(\mathbf{z})) f(\mathbf{z}, \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) \\ &= \int_\Theta [V(\theta_m, T_1(\mathbf{z})) - V(\theta_m, T_m(\mathbf{z}))] f(\mathbf{z}, \boldsymbol{\theta})\, d\tau(\boldsymbol{\theta}) . \end{aligned} \tag{3.5}$$

By assumption $T_m(\mathbf{z}) > T_1(\mathbf{z})$. By the monotonicity of $V(\theta, a)$, $V(\theta_m, T_1(\mathbf{z})) - V(\theta_m, T_m(\mathbf{z})) < 0$ for all θ_m. Thus the last integral in (3.5) being zero is a contradiction.

4. Main results.

In this section we establish the admissibility and inadmissibility of $\mathbf{T}$ for various choices of $n_1, \cdots, n_m$.

THEOREM 4.1. *Let $m \geqq 2$. If $n_1 + \cdots + n_m \geqq 7$ and neither*

$$\text{(i)} \quad m = 2 , \quad \textit{and} \quad n_1 = 1 \quad \textit{or} \quad n_2 = 1$$

nor

$$\text{(ii)} \quad m = 3 , \quad n_1 = n_3 = 1$$

holds, then $\mathbf{T}$ *is inadmissible.*

PROOF. The method of proof will be to find points $\mathbf{y}$ and $\mathbf{z}$ satisfying the conditions of Theorem 2.

Case I. $n_1 \geqq 3$. Choose

$$\mathbf{y} = (n_1, n_2, \cdots, n_{m-1}, n_m - 1) , \qquad \mathbf{z} = (2, n_2, \cdots, n_{m-1}, n_m - 1) .$$

Here $B(\mathbf{y}) = 1$, $U(\mathbf{y}) = -\infty$, $B(\mathbf{z}) = 1$, $U(\mathbf{z}) = 1$ and

$$T_1(\mathbf{y}) = 1 - (n_1 + \cdots + n_m)^{-1} = T_m(\mathbf{y})$$

$$T_1(\mathbf{z}) = \frac{2}{n_1} < 1 - (n_2 + \cdots + n_m)^{-1} = T_m(\mathbf{z})$$

under the assumptions on $n_1, \cdots, n_m$. By Theorem 3.2, T is inadmissible in this case.

The remaining cases can also be handled in this way, that is, by specifying an appropriate **x** and **y**. This is done in the following table.

TABLE 1

Conditions in addition to $n_1 + \cdots + n_m \geqq 7$	**x** **y**	$B(\mathbf{x})$ $B(\mathbf{y})$	$U(\mathbf{x})$ $U(\mathbf{y})$
$n_m \geqq 3$	$(1, 0, \cdots, 0, 0)$ $(1, 0, \cdots, 0, n_m - 2)$	$+\infty$ m	m m
$n_1 = 2, n_m = 2$	$(2, n_2, \cdots, n_{m-1}, 1)$ $(2, 0, \cdots, 0, 1)$	1 1	$-\infty$ $m - 1$
$m = 3$ $n_1 = 2, n_m = 1$	$(1, 0, 0)$ $(1, n_2 - 1, 0)$	$+\infty$ 2	2 2
$m = 3$ $n_1 = 1, n_m = 2$	$(1, n_2, 1)$ $(1, 1, 1,)$	2 2	$-\infty$ 2
$m \geqq 4$ $n_1 = 2, n_m = 1$	$(2, n_2, \cdots, n_{m-1}, 0)$ $(2, 0, \cdots, 0, n_{m-1}, 0)$	1 1	$-\infty$ $m - 2$
$m \geqq 4$ $n_1 = 1, n_m = 2$	$(1, 0, \cdots, 0)$ $(1, 0, n_3, \cdots, n_{m-1}, 0)$	$+\infty$ 3	m m
$m = 4$ $n_1 = n_2 = n_4 = 1$	$(1, 0, 1, 0)$ $(1, 0, n_3 - 1, 0)$	3 3	3 3
$m = 4$ $n_1 = n_3 = n_4 = 1$	$(1, n_2 - 1, 1, 0)$ $(1, 1, 1, 0)$	2 2	2 2
$m = 4$ $n_1 = n_4 = 1$ $n_2 > 1, n_3 > 1$	$(1, n_2, n_3, 0)$ $(1, 1, n_3, 0)$	2 2	$-\infty$ 2
$m > 4$ $n_1 = n_m = 1$	$(1, n_2, \cdots, n_{m-1}, 0)$ $(1, 1, 0, \cdots, 0, n_{m-1}, 0)$	2 2	$-\infty$ $m - 3$

In the important case of squared error loss (i.e. $W(u) = u^2$) the converse of Theorem 4.1 is also true.

THEOREM 4.2. *Let* $W(u) = u^2$. *Then* T(X) *is admissible if and only if either*

$$\text{(i)}\quad n_1 + \cdots + n_m < 7 \qquad \textit{or}$$

$$\text{(ii)}\quad m = 2\,,\quad n_i = 1 \quad \textit{some} \quad i = 1, 2 \qquad \textit{or}$$

$$\text{(iii)}\quad m = 3\,,\quad n_1 = n_3 = 1\,.$$

PROOF. If none of the conditions (i), (ii), or (iii) hold, then by Theorem 4.1, T(X) is inadmissible.

To prove admissibility we first note that if T(X) is admissible for $\mathscr{G}$ when $m = k$ ($k = 1, 2, \cdots$) then $\mathbf{T}_*(\mathbf{X})$ and $\mathbf{T}^*(\mathbf{X})$ are admissible for $\mathscr{G}_*$ and $\mathscr{G}^*$, respectively, when $m = k + 1$. We also note that for $m = 1$ T(X) is scalar-valued and is just the usual M.L.E. for a single binomial distribution. This is known to be admissible ([7]). Thus the proof reduces to (by Theorem 3.1) showing that $\mathbf{T}(\mathbf{X})$, restricted to $\mathscr{X}_0$, is admissible for $\mathscr{G}_0$ when $m = 2, \cdots, 6$. This can be done by showing that $\mathbf{T}(\mathbf{X})$ is unique Bayes for $\mathscr{G}_0$ in each of these situations. We will not do the computations but just state that (when the assumptions of Theorem 4.2 are satisfied) this can be shown by considering prior distributions of the following form. Let τ be a mixture of two distributions; τ_1 which is concentrated on the line $\theta_1 = \cdots = \theta_m$ with density function $K \sum_{i=0}^{n_1-1} (1-\theta)^i \sum_{i=0}^{n_m-1} \theta^i$ and τ_2 is a discrete distribution concentrated on the vertices of Θ. Thus we use a prior distribution of the form $\tau = q\tau_1 + (1-q)\tau_2$ for some $0 \leqq q \leqq 1$.

5. Remarks. We suspect that Theorem 4.1 and Theorem 4.2 hold for many loss functions other than those considered here. However, there does not appear to be a method which is applicable to all situations. For example, it can be shown that Theorems 4.1 and 4.2 hold if $L(\boldsymbol{\theta}, \mathbf{a}) = \sum_{i=1}^{m} (\theta_i - a_i)^2/\theta_i(1-\theta_i)$. Unfortunately some of the techniques used to show this are pertinent only for this loss function.

REFERENCES

[1] AYER, M., BRUNK, H. D., EWING, G. M., REID, W. J. and SILVERMAN, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26** 641–647.

[2] BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference Under Order Restrictions.* Wiley, New York.

[3] BARLOW, R. and BRUNK, H. D. (1972). The isotonic regression problem and its dual. *J. Amer. Statist. Assoc.* **67** 140–147.

[4] BARLOW, R. E. and SCHEUER, E. M. (1966). Reliability growth during a development testing program. *Technometrics* **8** 53–60.

[5] BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454.

[6] EISENHART, E., HASTAY, M. W. and WALLIS, W. A. (eds). (1947). *Techniques of Statistical Analysis.* McGraw-Hill, New York.

[7] FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

[8] GOULDEN, C. H. (1952). *Methods of Statistical Analysis*, (2nd ed.). Wiley, New York.

[9] JOHNSON, B. M. (1971). On the admissible estimators for certain fixed sample binomial problems. *Ann. Math. Statist.* **42** 1579–1587.

[10] SACKS, J. (1963). Generalized Bayes solutions in estimation problems. *Ann. Math. Statist.* **34** 751–768.

[11] WALD, A. (1971). *Statistical Decision Functions*, (2nd ed.). Chelsea, New York.

STATISTICS CENTER
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903