Ann. Inst. Statist. Math.
Vol. 41, No. 4, 699–716 (1989)

ADMISSIBLE ESTIMATORS OF BINOMIAL PROBABILITY AND THE INVERSE BAYES RULE MAP

MORRIS SKIBINSKY AND ANDREW L. RUKHIN*

Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, U.S.A.

(Received December 16, 1987; revised September 1, 1988)

Abstract. Explicit formulae for prior distribution moments through values of the Bayes estimator of binomial probability are obtained. These are used to derive a new admissibility criterion.

Key words and phrases: Binomial parameter, admissibility, complete class, Bayes rule map.

1. Introduction

Consider the classical problem of estimating an unknown binomial probability θ. Let X be a binomial random variable with parameters n and θ, and assume that θ is estimated by $\delta(X)$ under quadratic loss.

Johnson (1971) (see also Brown (1981)) has shown that every admissible estimator δ must be of the form

$$(1.1)\qquad \delta(x)=\begin{cases} 0 & x\le r\,,\\ \dfrac{\int_0^1 \theta^{x-r}(1-\theta)^{s-x-1}d\mu(\theta)}{\int_0^1 \theta^{x-r-1}(1-\theta)^{s-x-1}d\mu(\theta)} & r+1\le x\le s-1\,,\\ 1 & x\ge s\,. \end{cases}$$

Here $-1\le r<s\le n+1$, and μ is a probability measure such that

$$\mu(\{0\}\cup\{1\})<1\,.$$

(This condition implies that the corresponding Bayes rule is well-defined not only for $x=0$ and $x=n$ (cf. Lehmann (1983), p. 246).)

In other words, every admissible estimator has the form of a Bayes

*Research supported by National Science Foundation Grants #DMS-8702620 and #DMS-8802359.

procedure on the "middle" part of the range of X. It is easy to show that estimator (1.1) depends only on the first $s - r - 1$ moments of μ.

Kozek ((1982), Theorem 6) obtained the "analytic counterpart" to Johnson's Bayesian characterization of admissibility. Let δ be an estimator of θ based on X with values in (0, 1). Define

$$\tau_{n+1,i}(\delta) = \prod_{k=0}^{i-1} (\delta(k)/(1 - \delta(k))), \qquad i = 1, 2, \ldots, n + 1 . \tag{1.2}$$

In establishing the equivalence of his criterion to that of Johnson, Kozek demonstrated that δ is Bayes (hence admissible)—i.e., has the form given by (2.1)—if and only if there exists a probability measure η on $[0, \infty)$ and a non-negative constant K such that

$$\tau_{n+1,i}(\delta) = \int_0^\infty t^i d\eta(t), \qquad i = 1, 2, \ldots, n ,$$

and

$$\tau_{n+1,n+1}(\delta) = \int_0^\infty t^{n+1} d\eta(t) + K .$$

Thus, if δ has the form (2.1), then straightforward calculation (Kozek (1982), Remark 4) shows that

$$\eta([0, t]) = \int_0^{t/(t+1)} (1 - \theta)^{n+1} d\mu(\theta) \Big/ \int_0^1 (1 - \theta)^{n+1} d\mu(\theta), \qquad t \geq 0 ,$$

while

$$K = \mu(\{1\}) \Big/ \int_0^1 (1 - \theta)^{n+1} d\mu(\theta) .$$

The converse follows from a moment space theorem of Karlin and Studden (1966).

In this paper (Section 2), we derive a new admissibility criterion expressed in terms of the inverse Bayes rule map: Bayes estimators $\to$ moment $(n + 1)$-tuples. We define the inverse explicitly (see (2.12)). Benefits which are derived thereby include the ability to distinguish those admissible estimators that determine the prior, relative to which they are Bayes (these priors of necessity have finite support) from those that do not. (There are then uncountably many priors relative to which the admissible estimator is Bayes.)

In Section 3, the structure of the class of Bayes estimators is given in complete detail for $n = 3$, together with illustrative examples. Explicit criteria for an estimator to be Bayes are also given for $n = 4$. Proofs of technical lemmas and the main theorem are given in the Appendix.

We note that a characterization of prior distributions by linear Bayes estimators was obtained by Diaconis and Ylvisaker (1979) and complete class results for double sample estimation were given by Cohen and Sackrowitz (1984).

2. The inverse Bayes estimator map

In this section, we shall study the inverse mapping for Bayes estimators of θ. To be precise, we view such an estimator as an image under a mapping B_{n+1}

$$B_{n+1}: M_{n+1} \to \boldsymbol{R}^{n+1} .$$

Here n is an arbitrarily given integer ≥ 1; we interpret $\boldsymbol{R}^{n+1}$ as the space of functions $\delta: \{0, 1, \ldots, n\} \to \boldsymbol{R}$; and take M_{n+1} to be the space of the first $n+1$ moments for probability distributions on $[0, 1]$ with support not confined solely to the endpoints (the Bayes estimators of θ with respect to these prior distributions are uniquely defined). For such a distribution μ, we take

$$c_i = \int_0^1 \theta^i d\mu(\theta), \qquad i = 0, 1, \ldots, n+1 .$$

It is well known that the Bayes estimator depends on μ only through $c_1, c_2, \ldots, c_{n+1}$. Indeed, for each

$$\boldsymbol{c} = (c_1, c_2, \ldots, c_{n+1}) \in M_{n+1} ,$$

one has

$$\text{(2.1)} \qquad \begin{aligned} B_{n+1}(\boldsymbol{c})(i) &= \int_0^1 \theta^{i+1}(1-\theta)^{n-i} d\mu(\theta) \Big/ \int_0^1 \theta^i (1-\theta)^{n-i} d\mu(\theta) \\ &= c_{i+1}^{(n-i)} / c_i^{(n-i)}, \qquad i = 0, 1, \ldots, n , \end{aligned}$$

where

$$\text{(2.2)} \quad c_i^{(j)} = c_i^{(j-1)} - c_{i+1}^{(j-1)}, \qquad i = 0, 1, \ldots, n+1-j, \quad j = 1, 2, \ldots, n+1 ,$$

with

$$c_i^{(0)} = c_i, \qquad i = 0, 1, \ldots, n+1 .$$

(See Skibinsky (1968) for a more explicit rendering of (2.1) in terms of $\boldsymbol{c}$.)

We note in passing that (2.1) affirms the conjecture of Cheng (1982) to the effect that the Bayes estimator of θ relative to a Beta (α, β) prior is Bayes also relative to distributions of finite support on $[0, 1]$. Indeed, by (2.1) it is Bayes relative to each of the uncountable infinity of such distributions whose first $n+1$ moments coincide with those of Beta (α, β).

In keeping with our interpretation of $\boldsymbol{R}^{n+1}$, we shall write δ_i for $\delta(i)$, $i = 0, 1, \ldots, n$ so that $\delta = (\delta_0, \delta_1, \ldots, \delta_n)$.

LEMMA 2.1. *Let*

$$\mathscr{M}_{n+1} = \{\delta \in \boldsymbol{R}^{n+1}: 0 < \delta_0 \leq \delta_1 \leq \cdots \leq \delta_n < 1\} .$$

Then

$$B_{n+1}(M_{n+1}) \subset \mathscr{M}_{n+1} .$$

Let M_{n+1}^* denote the collection of all $(n+1)$-tuples $\boldsymbol{c}$ such that $c_i^{(n-i)} \neq 0$, $i = 0, 1, \ldots, n$, i.e., for which the right-hand side of (2.1) is well defined. Note that M_{n+1}^*, hence M_{n+1}, does not contain any sequence all of whose components are the same. We denote the extension of B_{n+1} to M_{n+1}^* by B_{n+1}^*. Thus (deleting subscripts)

$$M \subset M^* \quad \text{and} \quad B = B^*|_M .$$

LEMMA 2.2. *B_{n+1}^* is one-one on M_{n+1}^*.*

Let $\delta \in \mathscr{M}_{n+1}$ and put

$$\delta_i^{(0)} = \delta_i, \quad i = 0, \ldots, n; \quad \delta_{n+1}^{(0)} = 1 \quad \text{(by convention)} . \tag{2.3}$$

For $j = 1, 2, \ldots, n$, recursively define

$$\delta_i^{(j)} = \frac{\delta_i^{(j-1)}}{1 - \lambda_i^{(j-1)}}, \quad i = 0, 1, \ldots, n - j + 1 , \tag{2.4}$$

where for $j = 0, 1, \ldots, n$

$$\lambda_i^{(j)} = \delta_{i+1}^{(j)} - \delta_i^{(j)}, \quad i = 0, 1, \ldots, n - j . \tag{2.5}$$

Note that by our convention in (2.3)

$$\delta_{n-j+1}^{(j)} = 1, \quad j = 0, 1, \ldots, n . \tag{2.6}$$

Hereafter, we shall also write λ_i for $\lambda_i^{(0)}$, $i = 0, 1, \ldots, n$.

LEMMA 2.3. *For* $j = 1, 2, \ldots, n$, *we have for* $\delta \in \mathscr{M}_{n+1}$,

$$\lambda_i^{(j)} = \frac{\lambda_i^{(j-1)}(1 - \delta_{i+1}^{(j-1)}) + \lambda_{i+1}^{(j-1)}\delta_i^{(j-1)}}{(1 - \lambda_i^{(j-1)})(1 - \lambda_{i+1}^{(j-1)})}, \quad i = 0, 1, \ldots, n-j. \tag{2.7}$$

Note that by (2.5) *and* (2.6), *whenever* $i = n - j$, (2.7) *is reduced to*

$$\lambda_{n-j}^{(j)} = 1 - \lambda_{n-j}^{(j)} = \frac{\lambda_{n-j+1}^{(j-1)}}{1 - \lambda_{n-j}^{(j-1)}}, \quad j = 1, 2, \ldots, n.$$

LEMMA 2.4. *For* $\delta \in \mathscr{M}_{n+1}$ *and* $j = 0, 1, \ldots, n$, $\delta^{(j)}$ *is positive, non-decreasing, and less than* 1 *on* $\{0, 1, \ldots, n-j\}$; *equivalently*,

$$\delta_0^{(j)} > 0; \quad \lambda_i^{(j)} \geq 0, \quad i = 0, 1, \ldots, n-j-1; \quad \lambda_{n-j}^{(j)} > 0.$$

Also,

$$\delta_0^{(j)} + \sum_{i=0}^{n-j} \lambda_i^{(j)} = 1.$$

Moreover,

$$0 < \delta_0 \leq \delta_0^{(1)} \leq \delta_0^{(2)} \leq \cdots \leq \delta_0^{(n)} < 1,$$
$$0 < \lambda_n \leq \lambda_{n-1}^{(1)} \leq \lambda_{n-2}^{(2)} \leq \cdots \leq \lambda_0^{(n)} < 1.$$

Some insight into computations deriving from these recursions (e.g., see Lemmas 3.1 and 3.2) is gained by observing that for $i = 0, 1, \ldots, n-j-1$, $j = 1, 2, \ldots, n$,

$$\lambda_i^{(j)} = 0 \Rightarrow \delta_i = \delta_{i+1} = \cdots = \delta_{i+j+1};$$

a consequence of (2.7) and the above lemma.

For each $\delta \in \mathscr{M}_{n+1}$, define

$$\nu_{n+1,i}(\delta) = \prod_{k=0}^{i-1} \delta_k^{(n-k)}, \quad i = 1, 2, \ldots, n+1, \tag{2.8}$$

and put

$$\boldsymbol{\nu}_{n+1} = (\nu_{n+1,1}, \nu_{n+1,2}, \ldots, \nu_{n+1,n+1}). \tag{2.9}$$

Now we apply the analogues of the definitions (2.2) to the $n+1$ components of $\boldsymbol{\nu}_{n+1}$ given by (2.8). Thus, for $i = 0, 1, \ldots, n+1-j$, $j =$

$1, 2, \ldots, n+1$, we define functions on $\mathscr{M}_{n+1}$:

$$v_{n+1,i}^{(j)} = v_{n+1,i}^{(j-1)} - v_{n+1,i+1}^{(j-1)}$$

with

$$v_{n+1,i}^{(0)} = v_{n+1,i}, \qquad i = 0, 1, \ldots, n+1, \quad v_{n+1,0} = 1 .$$

LEMMA 2.5. *For each $\delta \in \mathscr{M}_{n+1}$, and for each pair of non-negative integers i, j such that $i + j \leq n$,*

$$v_{n+1,i+1}^{(j)}(\delta) / v_{n+1,i}^{(j)}(\delta) = \delta_i^{(n-j-i)} .$$

COROLLARY 2.1. *For each $\delta \in \mathscr{M}_{n+1}$, $B_{n+1}^*(\boldsymbol{v}_{n+1}(\delta)) = \delta$.*

PROOF. By Lemma 2.5, setting $i + j = n$, we have, using (2.1) and the definition of B_{n+1}^*, that

$$\delta_i = \delta_i^{(0)} = v_{n+1,i+1}^{(n-i)}(\delta) / v_{n+1,i}^{(n-i)}(\delta) = B_{n+1}^*(\boldsymbol{v}_{n+1}(\delta))(i) ,$$

for $i = 0, 1, \ldots, n$ and $\delta \in \mathscr{M}_{n+1}$. But this is the desired result.

The inverse of the mapping $\mathscr{M}_{n+1} \to \boldsymbol{R}^{n+1}$ given by (2.9) and the inverse of the Bayes estimator mapping $M_{n+1} \to \mathscr{M}_{n+1}$ defined by (2.1) are specified in the following theorem. Note that the first inclusion statement of this theorem is just Lemma 2.1.

THEOREM 2.1.

$$B_{n+1}(M_{n+1}) \subset \mathscr{M}_{n+1}; \qquad M_{n+1} \subset \boldsymbol{v}_{n+1}(\mathscr{M}_{n+1}) \subset M_{n+1}^* ;$$

the inverse mapping $\boldsymbol{v}_{n+1}^{-1}$ is the restriction of B_{n+1}^* to $\boldsymbol{v}_{n+1}(\mathscr{M}_{n+1})$; the inverse mapping B_{n+1}^{-1} is the restriction of $\boldsymbol{v}_{n+1}$ to $B_{n+1}(M_{n+1})$.

We proceed by defining $n + 1$ disjoint subclasses of functions in $\mathscr{M}_{n+1}$ as follows. Let $\mathscr{D}_{n,0}$ denote the collection of all $\delta \in \boldsymbol{R}^{n+1}$ such that

$$\delta_i = 0, \quad a, \quad \text{or} \quad 1 ,$$

according as $0 \leq i < r$, $i = r$ or $r < i \leq n$, for some integer r, $0 \leq r \leq n$ and some a, $0 \leq a \leq 1$. For $m = 1, 2, \ldots, n$, denote by $\mathscr{D}_{n,m}$ the collection of all $\delta \in \boldsymbol{R}^{n+1}$ such that

$$\delta_i = 0, \quad B_{n+1}(\boldsymbol{c})(i - r), \quad \text{or} \quad 1 ,$$

according as $0 \le i < r$, $r \le i \le m + r$ or $m + r < i \le n$, for some integer r, $0 \le r \le n - m$ and some $\boldsymbol{c} \in M_{n+1}$.

Thus, a typical function in $\mathscr{D}_{n,1}$ is of the form

$$\delta_i = 0, \quad (c_1 - c_2)/(1 - c_1), \quad c_2/c_1, \quad \text{or} \quad 1 ,$$

according as $0 \le i < r$, $i = r$, $i = r + 1$, or $r + 1 < i \le n$, for some integer r, $0 \le r \le n - 1$, and some $(c_1, c_2) \in M_2$ (i.e., for some c_1, c_2 such that $c_1^2 \le c_2 < c_1$, $0 < c_1 < 1$). At the other extreme,

$$\mathscr{D}_{n,n} = B_{n+1}(M_{n+1}) ,$$

the class of all Bayes rules relative to prior distributions whose support is not confined to the endpoints of [0, 1].

If we now put

$$\mathscr{D}_n = \bigcup_{m=0}^{n} \mathscr{D}_{n,m} ,$$

we may reformulate Johnson's theorem referred to in the introduction.

JOHNSON'S (1971) THEOREM. *$\mathscr{D}_n$ is precisely the class of all admissible estimators of θ.*

Observe now that by Theorem 2.1, $\mathscr{D}_{n,m}$ may for $m = 1,2,\ldots,n$, be described as the collection of all functions $\delta \in \boldsymbol{R}^{n+1}$ such that

$$\delta_i = 0, \quad \hat{\delta}_{i-r}, \quad \text{or} \quad 1$$

according as $0 \le i < r$, $r \le i \le m + r$, or $m + r < i \le n$, for some integer r, $0 \le r \le n - m$, and some $\hat{\delta} \in v_{m+1}^{-1}(M_{m+1})$.

Thus in particular, $\mathscr{D}_{n,1}$ is partitioned in a natural way by the elements of

$$v_2^{-1}(M_2) = \{\hat{\delta} \in \mathscr{M}_2 \colon v_{2,1}^2(\hat{\delta}) \le v_{2,2}(\hat{\delta}) < v_{2,1}(\hat{\delta}), 0 < v_{2,1}(\hat{\delta}) < 1\} .$$

In fact, this class is $\mathscr{M}_2$ itself (see Lemma 3.1). At the other extreme

$$\mathscr{D}_{n,n} = v_{n+1}^{-1}(M_{n+1}) ,$$

which for $n \ge 3$ is a proper subclass of $\mathscr{M}_{n+1}$ (see Theorem 3.1).

Johnson's criterion for admissibility may now be given in terms of estimator-dependent determinants of Hankel matrices

(2.10) $\quad \underline{\Delta}_{s,t}, \quad \bar{\Delta}_{s,t}, \quad t = 1,2,\ldots,s; \quad s = 2,3,\ldots,n+1$.

Here, for positive integers k, we define $\underline{\Delta}_{s,2k}$, for $2k \leq s$ ($\underline{\Delta}_{s,2k+1}$, for $2k+1 \leq s$) to be the determinant of the $(k+1)$-square matrix with entry $v_{s,i+j-2}$ ($v_{s,i+j-1}$) in its i-th row and j-th column; $i,j = 1,2,\ldots,k+1$. Similarly, $\bar{\Delta}_{s,2k-1}$, $1 \leq 2k-1 \leq s$ ($\bar{\Delta}_{s,2k}$, $2k \leq s$) denotes the determinant of the k-square matrix with entry $v^{(1)}_{s,i+j-2}$ ($v^{(1)}_{s,i+j-1}$) in its i-th row and j-th column; $i,j = 1,2,\ldots,k$.

Direct application of Theorems 17.2 and 17.3 (Karlin and Shapley (1953)) yields the following useful result.

THEOREM 2.2. *Let $\delta \in \mathscr{M}_{n+1}$ then $\delta \in \mathscr{D}_{n,n}$ (equivalently $\delta \in B_{n+1}(M_{n+1})$; equivalently $\mathbf{v}_{n+1}(\delta) \in M_{n+1}$), if and only if one of the following two conditions holds.*

1. $\underline{\Delta}_{n+1,k}(\delta) > 0$, $\bar{\Delta}_{n+1,k}(\delta) > 0$, $k = 1,2,\ldots,n+1$.

2. *For some integer k, $2 \leq k \leq n+1$, either $\underline{\Delta}_{n+1,k}(\delta) = 0$, $\bar{\Delta}_{n+1,k}(\delta) > 0$ or $\underline{\Delta}_{n+1,k}(\delta) > 0$, $\bar{\Delta}_{n+1,k}(\delta) = 0$ and*

$$\textit{all } \Delta(\delta)\textit{'s of index} \quad \leqq k \textit{ are} \quad \geq 0 \,.$$

Note that each $\delta \in \mathscr{D}_{n,n}$ which satisfies condition 1 of the above theorem is the Bayes estimator with respect to each of the uncountably many distributions μ on $[0,1]$ whose first $n+1$ moments coincide with $\mathbf{v}_{n+1}(\delta)$. On the other hand, each $\delta \in \mathscr{D}_{n,n}$ which satisfies Condition 2 is the Bayes estimator with respect to precisely one distribution μ. This distribution has finite support, being the lower or the upper principal representation of $(v_{n+1,1}(\delta),\ldots,v_{n+1,k-1}(\delta))$, according as the first or the second pair of inequalities obtains (see Propositions 1–6 of Skibinsky (1986)).

To see the connection between Kozek's criterion for admissibility and our own, observe first that if $\delta \in \mathscr{M}_{n+1}$ (hence takes all its values in $(0,1)$), we may, by an easy inversion of (1.2), express its values in the form

$$\delta_i = \tau_{n+1,i+1}(\delta)/(\tau_{n+1,i}(\delta) + \tau_{n+1,i+1}(\delta)), \qquad i = 0,1,\ldots,n \,.$$

Substitution into (2.4) and induction yield

$$\delta_i^{(j)} = r_{i+1}^{(j)}(\delta)/r_i^{(j+1)}(\delta), \qquad i = 0,1,\ldots,n-j, \quad j = 0,1,\ldots,n \,,$$

where, viewed as functions on $\mathscr{M}_{n+1}$,

$$r_i^{(j)} = \sum_{k=0}^{j} \binom{j}{k} \tau_{n+1,i+k}, \qquad i = 0,1,\ldots,n+1-j, \quad j = 0,1,\ldots,n+1 \,.$$

By (2.8), therefore,

$$\begin{aligned}(2.11)\qquad v_{n+1,i} &= r_i^{(n+1-i)}/r_0^{(n+1)}\\ &= \sum_{k=0}^{n+1-i}\binom{n+1-i}{k}\tau_{n+1,i+k}\Big/\sum_{k=0}^{n+1}\binom{n+1}{k}\tau_{n+1,k}\,.\end{aligned}$$

Note that by (1.2), the inverse Bayes rule map (2.9) is now explicitly defined. Thus, for each $\delta \in \mathscr{M}_{n+1}$ and for $i = 1, 2, \ldots, n$,

$$(2.12)\qquad v_{n+1,i}(\delta) = \frac{\sum_{k=0}^{n+1-i}\binom{n+1-i}{k}\prod_{s=0}^{i+k-1}(\delta_s/(1-\delta_s))}{\sum_{k=0}^{n+1}\binom{n+1}{k}\prod_{s=0}^{k-1}(\delta_s/(1-\delta_s))}\,.$$

On the other hand, if we restrict δ to $B_{n+1}(M_{n+1})$ (by Lemma 2.1, a subclass of $\mathscr{M}_{n+1}$)—equivalently, if $v_{n+1}(\delta) \in M_{n+1}$—we have by (2.1), following Kozek ((1982), Remark 4), that

$$\tau_{n+1,i}(\delta) = \int_0^1 \theta^i(1-\theta)^{n+1-i}d\mu(\theta)\Big/\int_0^1(1-\theta)^{n+1}d\mu(\theta), \qquad i = 1, 2, \ldots, n+1\,.$$

It follows that everywhere on $B_{n+1}(M_{n+1})$ and for $i = 1, 2, \ldots, n+1$,

$$(2.13)\quad \tau_{n+1,i} = \sum_{k=0}^{n+1-i}(-1)^k\binom{n+1-i}{k}v_{n+1,i+k}\Big/\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}v_{n+1,k}\,.$$

One may also obtain this result directly via formula (4) in Skibinsky (1968). Clearly, the $(n+1)$-tuples with components (2.13) and (2.11) restricted to Bayes estimators are composite maps on $B_{n+1}(M_{n+1})$; say

$$\tau_{n+1} = \psi_{n+1}(v_{n+1}), \qquad v_{n+1} = \zeta_{n+1}(\tau_{n+1})\,.$$

Note first that

$$\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}v_{n+1}(k) = 1\Big/\sum_{k=0}^{n+1}\binom{n+1}{k}\tau_{n+1,k}\,.$$

It is now easily established that $\psi_{n+1} = \zeta_{n+1}^{-1}$.

3. Admissible estimators, $n = 3, 4$

We illustrate our method of development by making use of the notation and results of Section 2 to explicitly exhibit the class of all admissible estimators, $n = 3$. Thereby we show that not all estimators in

$\mathscr{M}_4$ are admissible, a fact already noted by Kozek ((1982), Remark 1). In this process, moreover, we are able to characterize the class of all prior distributions, relative to each of whose members any specific admissible estimator is Bayes (e.g., see Lemma 3.3). In particular, we can identify those admissible estimators for which the prior is unique, and if necessary exhibit this prior in terms of the estimator (see Examples 1 through 4).

Finally, for $n = 4$, we characterize by explicit inequalities the estimators in $\mathscr{M}_5$ which are Bayes and display (Example 5) $\mathscr{M}_5$ estimators which are not.

We are concerned, first of all, with the class

$$\mathscr{D}_3 = \bigcup_{m=0}^{3} \mathscr{D}_{3,m} .$$

By Johnson's theorem and Theorem 2.1, it will suffice to characterize

$$B_s(M_s) = \{\delta \in \mathscr{M}_s \colon v_s(\delta) \in M_s\}, \qquad s = 2, 3, 4 .$$

We begin with Kozek's observation that

LEMMA 3.1.

$$B_s(M_s) = \mathscr{M}_s, \qquad s = 2, 3 .$$

PROOF. Let $\delta = (\delta_0, \delta_1) \in \mathscr{M}_2$. (2.4), (2.5), Lemma 2.3 and (2.8), together with definitions for (2.10), *taking $n = 1$ throughout*, yield $\underline{\Delta}_{2,1}(\delta) = \delta_0^{(1)}$, $\bar{\Delta}_{2,1}(\delta) = \lambda_0^{(1)}$, $\underline{\Delta}_{2,2}(\delta) = \delta_0^{(1)}\lambda_0^{(1)}\lambda_0$, $\bar{\Delta}_{2,2}(\delta) = \delta_0^{(1)}\lambda_1$. The $\lambda_i^{(j)}$ are defined by (2.5). By Lemma 2.4 with $n = 1$, $\underline{\Delta}_{2,2}(\delta) \geq 0$ (equality holds if and only if $\lambda_0 = \delta_1 - \delta_0 = 0$), and the 3 remaining determinants are positive. By Theorem 2.2, our lemma holds when $s = 2$. Thus, if $n = 1$, every estimator in $\mathscr{M}_2$ is Bayes.

Now let $\delta = (\delta_0, \delta_1, \delta_2) \in \mathscr{M}_3$. Following the same route, *but now taking $n = 2$ throughout*, one finds that

$$\begin{aligned} \underline{\Delta}_{3,1}(\delta) &= \delta_0^{(2)}, & \bar{\Delta}_{3,1}(\delta) &= \lambda_0^{(2)} , \\ \underline{\Delta}_{3,2}(\delta) &= \delta_0^{(2)}\lambda_0^{(2)}\lambda_0^{(1)}, & \bar{\Delta}_{3,2}(\delta) &= \delta_0^{(2)}\lambda_1^{(1)} , \\ \underline{\Delta}_{3,3}(\delta) &= (\delta_0^{(2)})^2\delta_1^{(1)}\lambda_1^{(1)}\lambda_1, & \bar{\Delta}_{3,3}(\delta) &= \frac{\delta_0^{(2)}\lambda_0^{(2)}\lambda_1^{(1)}(1-\delta_1)}{1-\lambda_0}\lambda_0 . \end{aligned}$$

By Lemma 2.4 with $n = 2$, both order 1 determinants are positive. $\underline{\Delta}_{3,2}(\delta) \geq 0$ (equality holds here if and only if $\lambda_0^{(1)} = 0$, equivalently, if and only if $\lambda_0 = \lambda_1 = 0$. Note, therefore, that $\underline{\Delta}_{3,2}(\delta) = 0$ implies that $\underline{\Delta}_{3,3}(\delta) = \bar{\Delta}_{3,3}(\delta) = 0$). $\bar{\Delta}_{3,2}(\delta) > 0$, $\underline{\Delta}_{3,3}(\delta) \geq 0$ (equality if and only if $\lambda_1 = \delta_2 - \delta_1 = 0$),

$\bar{\Delta}_{3,3}(\delta) \geq 0$ (equality if and only if $\lambda_0 = \delta_1 - \delta_0 = 0$). Thus, by Theorem 2.2 our lemma holds when $s = 3$; i.e., if $n = 2$, every estimator in $\mathscr{M}_3$ is Bayes.

By the discussion following Theorem 2.2, it is apparent that the formulae above enable us to determine precisely which estimators in $\mathscr{M}_2$ and $\mathscr{M}_3$ are Bayes only with respect to a distribution of finite support, and which are Bayes with respect to uncountably many prior distributions.

LEMMA 3.2. *Let $\delta = (\delta_0, \delta_1, \delta_2, \delta_3) \in \mathscr{M}_4$. Then*

$$\begin{aligned}
\underline{\Delta}_{4,1}(\delta) &= \underline{K}_1(\delta), & \bar{\Delta}_{4,1}(\delta) &= \bar{K}_1(\delta)\,,\\
\underline{\Delta}_{4,2}(\delta) &= \underline{K}_2(\delta) \cdot \lambda_0^{(2)}, & \bar{\Delta}_{4,2}(\delta) &= \bar{K}_2(\delta)\,,\\
\underline{\Delta}_{4,3}(\delta) &= \underline{K}_3(\delta) \cdot \lambda_1^{(1)}, & \bar{\Delta}_{4,3}(\delta) &= \bar{K}_3(\delta) \cdot \lambda_0^{(1)}\,,\\
\underline{\Delta}_{4,4}(\delta) &= \underline{K}_4(\delta) \cdot (\lambda_0\lambda_2\delta_2(1-\delta_1) - \lambda_1^2\lambda_3\delta_0), & \bar{\Delta}_{4,4}(\delta) &= \bar{K}_4(\delta) \cdot \lambda_1\,,
\end{aligned}$$

where $\underline{K}_i$, $\bar{K}_i$, $i = 1, 2, 3, 4$ are the strictly positive functions on $\mathscr{M}_4$ which are defined below.

$$\begin{aligned}
\underline{K}_1(\delta) &= \delta_0^{(3)}, & \bar{K}_1(\delta) &= \lambda_0^{(3)}\,,\\
\underline{K}_2(\delta) &= \delta_0^{(3)}\lambda_0^{(3)}, & \bar{K}_2(\delta) &= \delta_0^{(3)}\lambda_1^{(2)}\,,\\
\underline{K}_3(\delta) &= (\delta_0^{(3)})^2\delta_1^{(2)}\lambda_1^{(2)}, & \bar{K}_3(\delta) &= \frac{\delta_0^{(3)}\lambda_0^{(3)}\lambda_1^{(2)}(1-\delta_2)}{(1-\lambda_0^{(1)})(1-\lambda_1)}\,,
\end{aligned}$$

$$\underline{K}_4(\delta) = \frac{(\delta_0^{(3)})^2\lambda_0^{(3)}\delta_1^{(2)}\lambda_1^{(2)}(1-\delta_1^{(1)})}{(1-\lambda^{(1)})(1-\lambda_0)(1-\lambda_1)(1-\lambda_2)}\,,$$

$$\bar{K}_4(\delta) = \frac{(\delta_0^{(3)})^2\delta_1^{(2)}\lambda_1^{(2)}\lambda_2^{(1)}(1-\delta_2)}{(1-\lambda_0^{(1)})(1-\lambda_1)}\,.$$

PROOF. Factorization of the Hankel determinants proceeds in a straightforward manner as in Lemma 3.1, but *now taking $n = 3$, throughout*. The strictly positive nature of the coefficients $\underline{K}_i$, $\bar{K}_i$ follows directly from Lemma 2.4.

LEMMA 3.3. *Let $\delta \in \mathscr{M}_4$. If $\lambda_1 = 0$ (equivalently, if $\delta_1 = \delta_2$) or if*

$$\lambda_1 > 0 \quad \textit{and} \quad \lambda_0\lambda_2\delta_2(1-\delta_1) = \delta_0\lambda_1^2\lambda_3\,, \tag{3.1}$$

then δ is the Bayes estimator of θ with respect to exactly one prior distribution for θ and this prior distribution has finite support. Specifically, *if* $\lambda_1 = 0$ *and* $\begin{cases} \lambda_0 = 0, \quad \lambda_2 = 0, & \textit{it is degenerate at the constant value of } \delta\,; \\ \lambda_0 > 0, \quad \lambda_2 = 0, & \textit{it is the lower principal representation of } (\nu_{4,1}(\delta), \nu_{4,2}(\delta))\,; \\ \lambda_0 = 0, \quad \lambda_2 > 0, & \textit{it is the upper principal representation of } (\nu_{4,1}(\delta), \nu_{4,2}(\delta))\,; \\ \lambda_0 > 0, \quad \lambda_2 > 0, & \textit{it is the upper principal representation of } (\nu_{4,1}(\delta), \nu_{4,2}(\delta), \nu_{4,3}(\delta))\,; \end{cases}$

if (3.1) *holds, it is the lower principal representation of* $(\nu_{4,1}(\delta), \nu_{4,2}(\delta), \nu_{4,3}(\delta))$.

If on the other hand,

$$\lambda_1 > 0 \quad \textit{and} \quad \lambda_0\lambda_2\delta_2(1-\delta_1) > \delta_0\lambda_1^2\lambda_3\,, \tag{3.2}$$

then δ *is the Bayes estimator of* θ *with respect to each of the uncountably many prior distributions for* θ *whose first* 4 *moments coincide with* $\mathbf{v}_4(\delta)$.

Finally, if δ *does not satisfy any of the above conditions, then it is not a Bayes estimator for* θ *and must of necessity belong to the class specified by the following theorem.*

THEOREM 3.1.

$$\mathscr{M}_4 \backslash B_4(M_4) = \{\delta \in \mathscr{M}_4\colon \lambda_0\lambda_2\delta_2(1-\delta_1) < \lambda_1^2\lambda_3\delta_0\}\,.$$

PROOF. By Lemmas 3.2 and 2.4,

$$\underline{\Delta}_i \geq 0, \quad \bar{\Delta}_i \geq 0, \quad i = 1, 2, 3; \quad \bar{\Delta}_4 \geq 0; \quad \text{everywhere on } \mathscr{M}_4\,.$$

Moreover, as the classification in Lemma 3.3 indicates, one or the other condition in Theorem 2.2 is satisfied for each $\delta \in \mathscr{M}_4$ such that $\underline{\Delta}_4(\delta) \geq 0$. Of course, neither condition is satisfied if $\underline{\Delta}_4(\delta) < 0$. Therefore,

$$\mathscr{M}_4 \backslash B_4(M_4) = \{\delta \in \mathscr{M}_4\colon \underline{\Delta}_4(\delta) < 0\}\,.$$

The result now follows from the expression for $\underline{\Delta}_4(\delta)$ given in Lemma 3.2.

The minimal complete class for $n = 3$ may now be characterized explicitly as follows.

THEOREM 3.2. *$\delta \in \boldsymbol{R}^4$ is admissible if and only if it takes the form* $(0, a, b, c)$ *or* $(a, b, c, 1)$ *for some triple* a, b, c *such that*

$$0 \le a \le b \le c \le 1\,,$$

or else belongs to

$$B_4(M_4) = \{\delta \in \mathscr{M}_4 \colon (\delta_1 - \delta_0)(1 - \delta_1)(\delta_3 - \delta_2)\delta_2 \ge \delta_0(\delta_2 - \delta_1)^2(1 - \delta_3)\}\,. \tag{3.3}$$

The inequality in (3.3) *is of course the inequality of Kozek* ((1982), *Remark* 1).

Example 1.

$$\delta = (a, b, b, b), \qquad 0 < a < b < 1\,.$$

Here, $\lambda_1 = 0$, $\lambda_0 = b - a$, $\lambda_2 = 0$. Therefore, δ is admissible. It is Bayes (only) with respect to the lower principal representation of $(\nu_{4,1}(\delta), \nu_{4,2}(\delta))$. Specifically, δ is Bayes only with respect to the prior distribution whose support is at 0 and b with mass $a/[1 - (b - a)(b^2 + 3(1 - b))]$ at the latter point.

Example 2.

$$\delta = (\varepsilon, \varepsilon, 1 - \varepsilon, 1 - \varepsilon), \qquad 0 < \varepsilon < 1/2\,.$$

Here, $\lambda_0\lambda_2\delta_2(1 - \delta_1) = 0$. But $\delta_0\lambda_1^2\lambda_3 = (\varepsilon(1 - 2\varepsilon))^2 > 0$. Hence, by Theorem 3.1, $\delta \in \mathscr{M}_4 \backslash B_4(M_4)$, i.e., δ is inadmissible.

Example 3.

$$\delta = (\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon), \qquad 0 < \varepsilon < 1/4$$

satisfies (3.2), so is admissible and Bayes with respect to each of the uncountably many distributions whose first 4 moments coincide with $\nu_4(\delta)$.

Example 4.

$$\delta = (a, b, 1 - b, 1 - a), \qquad 0 < a \le b < 1/2\,.$$

Here, $\lambda_1 = 1 - 2b > 0$, and

$$\lambda_0\lambda_2\delta_2(1 - \delta_1) = [(1 - b)(b - a)]^2 \gtreqless [a(1 - 2b)]^2 = \delta_0\lambda_1^2\lambda_3\,,$$

according as

$$(1-b)(b-a) \gtreqless a(1-2b);$$

i.e., according as

$$a \lesseqgtr \frac{b(1-b)}{2-3b}.$$

Note that

$$0 < \frac{b(1-b)}{2-3b} < b, \quad \text{all} \quad b < \frac{1}{2}.$$

Thus, by Lemma 3.3 if

$$0 < a < \frac{b(1-b)}{2-3b}, \tag{3.4}$$

then δ is the Bayes estimator with respect to the infinitely many prior distributions whose first 4 moments coincide with

$$\nu_4(\delta) = \left(\frac{1}{2}, \frac{1-b+a}{2(1-b+3a)}, \frac{1-b}{2(1-b+3a)}, \frac{(1-a)(1-b)}{2(1-b+3a)} \right).$$

Note that the condition (3.4) implies $\nu_4(\delta) \in \operatorname{int} M_4$.

If

$$a = \frac{b(1-b)}{2-3b},$$

then δ is the Bayes estimator *only* with respect to the prior distribution given by the lower principal representation of

$$(\nu_{4,1}(\delta), \nu_{4,2}(\delta), \nu_{4,3}(\delta)) = \left(\frac{1}{2}, \frac{1-b+a}{2(1-b+3a)}, \frac{1-b}{2(1-b+3a)} \right).$$

This prior distribution is uniform with two-point support at

$$\frac{1}{2}\left[1 \pm \left(\frac{1-b-a}{1-b+3a} \right)^{1/2} \right].$$

If the third possibility,

$$\frac{b(1-b)}{2-3b} < a < b$$

holds, δ is not a Bayes estimator, hence is not admissible. In particular, $(1/4,\ 1/3,\ 2/3,\ 3/4)$ is neither Bayes nor admissible.

Finally, as an additional application of Theorem 2.2, we give an explicit characterization for $B_5(M_5)$.

When $n = 4$, an estimator

$$\delta = (\delta_0, \ldots, \delta_4) \in \mathscr{M}_5$$

is a Bayes estimator of θ if and only if it satisfies both of the following inequalities.

$$(\delta_2 - \delta_1)(1 - \delta_2)(\delta_4 - \delta_3)\delta_3 \geq \delta_1(\delta_3 - \delta_2)^2(1 - \delta_4)\,,$$

$$(\delta_1 - \delta_0)(1 - \delta_1)(\delta_3 - \delta_2)\delta_2 \geq \delta_0(\delta_2 - \delta_1)^2(1 - \delta_3)\,.$$

Classification of these estimators as in Lemma 3.3 is easily carried out.

Example 5.

$$\theta = (a, b, b, c, c) \in \mathscr{M}_5$$

is not a Bayes estimator of θ.

Acknowledgements

The authors wish to express appreciation for helpful remarks made by the referees; in particular, for reference to Kozek's (1982) paper, which we had overlooked.

Appendix

We give proofs here for Lemmas 2.1, 2.2, 2.5 and Theorem 2.1. The proofs of Lemmas 2.3, 2.4 and 3.3 are straightforward and so are not included.

PROOF OF LEMMA 2.1. Let $\boldsymbol{c} \in M_{n+1}$. By definition of M_{n+1}, there is a probability measure on $[0, 1]$ which assigns positive probability to the open interval $(0, 1)$, such that

$$c_i = \int_0^1 \theta^i d\mu(\theta), \qquad i = 1, 2, \ldots, n+1\,.$$

It follows that of necessity

$$0 < c_{n+1} < c_n < \cdots < c_1 < 1 .$$

But then

$$B_{n+1}(\boldsymbol{c})(n) = c_{n+1}/c_n < 1 .$$

For $i = 1, 2, \ldots, n-1$, the following assertions are easily seen to be equivalent.

$$B_{n+1}(\boldsymbol{c})(i) \leq B_{n+1}(\boldsymbol{c})(i+1) .$$
$$c_{i+1}^{(n-i)} c_{i+1}^{(n-i-1)} \leq c_i^{(n-i)} c_{i+2}^{(n-i-1)} .$$
$$[c_{i+1}^{(n-i-1)} - c_{i+2}^{(n-i-1)}] c_{i+1}^{(n-i-1)} \leq c_i^{(n-i)} c_{i+2}^{(n-i-1)} .$$
$$(c_{i+1}^{(n-i-1)})^2 \leq c_{i+2}^{(n-i-1)} c_i^{(n-i-1)} .$$

The last inequality follows from the Cauchy-Schwarz inequality. Finally,

$$B_{n+1}(\boldsymbol{c})(0) = \int_0^1 \theta(1-\theta)^n d\mu(\theta) \bigg/ \int_0^1 (1-\theta)^n d\mu(\theta) > 0 ,$$

because μ is not solely supported by the endpoints of [0, 1].

PROOF OF LEMMA 2.2. We must show that the equality

$$B_{n+1}^*(\boldsymbol{c}) = B_{n+1}^*(\boldsymbol{d}) \tag{A.1}$$

for points $\boldsymbol{c}, \boldsymbol{d}$ in M_{n+1}^* implies that $\boldsymbol{c} = \boldsymbol{d}$. Let p, q be arbitrarily given integers such that

$$p \geq 1, \quad q \geq 0, \quad p + q \leq n + 1 ,$$

and denote by $\kappa_p^{(q)}$ the condition that

$$c_p^{(q)} / c_{p-1}^{(q)} = d_p^{(q)} / d_{p-1}^{(q)} . \tag{A.2}$$

A little algebra shows that conditions $\kappa_{p+1}^{(q)}$ and $\kappa_p^{(q+1)}$ together imply $\kappa_p^{(q)}$. Also, if $\kappa_i^{(k)}$ is true for all integers i, k such that $i \geq 1$, $k \geq 0$, $i + k = n + 1 - j$, for some integer $j \leq n - 1$, then the same condition holds if j is replaced by $j + 1$.

By (2.1), the hypothesis (A.1) is equivalent to the above condition with $j = 0$. Thus, induction shows that (A.1) implies the truth of $\kappa_i^{(k)}$ for all integers i, k such that $i \geq 1$, $k \geq 0$, $i + k \leq n + 1$. In particular, it implies that $\kappa_i^{(0)}$ holds for $i = 1, 2, \ldots, n + 1$. By (A.2), this is equivalent to the equalities

$$c_{n+1}/d_{n+1} = c_n/d_n = \cdots = c_1/d_1 = c_0/d_0 = 1 \ ,$$

which is the desired conclusion.

PROOF OF LEMMA 2.5. By (2.13) and Lemma 2.4, the statement of Lemma 2.5 is true when $j = 0$. Let J be a non-negative integer, $J \leq n$, and suppose this statement holds when $j = J - 1$. We show that it must then also hold for $j = J$. Indeed, for $i = 0, 1, \ldots, n - J$ (deleting the first subscript for v, which is always $n + 1$, as a notational convenience),

$$\begin{aligned} &v_{i+1}^{(J)}(\delta)/v_i^{(j)}(\delta) \\ &\quad = [v_{i+1}^{(J-1)}(\delta) - v_{i+2}^{(J-1)}(\delta)]/[v_i^{(J-1)}(\delta) - v_{i+1}^{(J-1)}(\delta)] \\ &\quad = (v_{i+1}^{(J-1)}(\delta)/v_i^{(J-1)}(\delta))[1 - (v_{i+2}^{(J-1)}(\delta)/v_{i+1}^{(J-1)}(\delta))]/[1 - (v_{i+1}^{(J-1)}(\delta)/v_i^{(J-1)}(\delta))] \\ &\quad = \delta_i^{(n-J-i+1)}[1 - \delta_{i+1}^{(n-J-i)}]/[1 - \delta_i^{(n-J-i+1)}] \qquad \text{by induction hypothesis} \\ &\quad = \delta_i^{(n-J-i)} \ . \end{aligned}$$

PROOF OF THEOREM 2.1. As a notational convenience, the subscript $n + 1$ which should be understood as attached to M, M^*, B, B^*, $\mathscr{M}$ and v will be deleted throughout this proof.

By Lemma 2.2, $B(M) \subset \mathscr{M}$. Hence

$$v(B(M)) \subset v(\mathscr{M}) \ . \tag{A.3}$$

By Corollary 2.1, B^* is well defined at $v(\delta)$ for each $\delta \in \mathscr{M}$ so that $v(\delta) \in M^*$ for each $\delta \in \mathscr{M}$. Equivalently,

$$v(\mathscr{M}) \subset M^* \ . \tag{A.4}$$

To show that $v^{-1} = B^*|_{v(\mathscr{M})}$, it will suffice in view of Corollary 2.1 to show that

$$v(B^*(c)) = c, \qquad \text{for each} \qquad c \in v(\mathscr{M}) \ .$$

Thus, let $c \in v(\mathscr{M})$. There must then exist $\delta \in \mathscr{M}$ such that $c = v(\delta)$. By Corollary 2.1, it then follows that $B^*(c) = B^*(v(\delta)) = \delta$. But then $v(B^*(c)) = v(\delta) = c$.

To show that $B^{-1} = v|_{B(M)}$, we must show that

$$\boldsymbol{v}(B(\boldsymbol{c})) = \boldsymbol{c}, \qquad \forall\, \boldsymbol{c} \in M \quad \text{and} \quad B(v(\delta)) = \delta, \qquad \forall\, \delta \in B(M)\,.$$

Thus, let $\boldsymbol{c} \in M$. By Lemma 2.2, $B(\boldsymbol{c}) \in \mathscr{M}$. Hence, by Corollary 2.1

$$B^*(v(B(\boldsymbol{c}))) = B(\boldsymbol{c}) = B^*(\boldsymbol{c})\,.$$

But then by Lemma 2.1,

$$\boldsymbol{v}(B(\boldsymbol{c})) = \boldsymbol{c}\,.$$

Now, let $\delta \in B(M)$; then there exists $\boldsymbol{c} \in M$ such that $\delta = B(\boldsymbol{c})$. But then $\boldsymbol{v}(\delta) = \boldsymbol{v}(B(\boldsymbol{c})) = \boldsymbol{c}$. Hence, $B(\boldsymbol{v}(\delta)) = B(\boldsymbol{c}) = \delta$.

We now know that

$$\boldsymbol{v}(B(M)) = M\,.$$

Thus, combining (A.3) and (A.4), we have

$$M \subset \boldsymbol{v}(\mathscr{M}) \subset M^*\,.$$

REFERENCES

Brown, L. (1981). A complete class theorem for statistical problems with finite sample spaces, *Ann. Statist.*, **9**, 1289–1303.

Cheng, P. (1982). Admissibility of simultaneous estimation of several parameters, *J. Systems Sci. Math. Sci.*, **2**, 176–195.

Cohen, A. and Sackrowitz, H. (1984). Results on double sample estimation for the binomial distribution, *Ann. Statist.*, **12**, 1109–1116.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families, *Ann. Statist.*, **7**, 269–281.

Johnson, B. McK. (1971). On the admissible estimators for certain fixed sample binomial problems, *Ann. Math. Statist.*, **42**, 1579–1587.

Karlin, S. and Shapley, L. S. (1953). Geometry of moment spaces, *Mem. Amer. Math. Soc.*, No. 12, Amer. Math. Soc. Providence, Rhode Island.

Karlin, S. and Studden, W. J. (1966). *Tchebycheff Systems, with Applications in Analysis and Statistics*, Interscience, New York.

Kozek, A. (1982). Towards a calculus of admissibility, *Ann. Statist.*, **10**, 825–837.

Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.

Skibinsky, M. (1968). Minimax estimation of a random probability whose first N moments are known, *Ann. Math. Statist.*, **39**, 492–501.

Skibinsky, M. (1986). Principal representations and canonical moment sequences for distributions on an interval, *J. Math. Anal. Appl.*, **120**, 95–118.