A Bayes Sequential Selection of the Least Probable Event

Hyung-Tae Hwang* & Woo-Chul Kim*

ABSTRACT

A problem of selecting the least probable cell in a multinomial distribution is studied in a Bayesian framework. We consider two loss components - the cost of sampling and the difference in cell probabilities between the selected and the least probable cells. A Bayes sequential selection rule is derived with respect to a Dirichlet prior, and it is compared with the best fixed sample size selection rule. The continuation sets with respect to the vague prior are tabulated for certain cases.

1. Introduction

The multinomial distribution often provides a statistical model for many problems in the real world. Also, we often ask the question : which events is the most probable or the least probable among the events under consideration? Bechhofer, Elmaghraby and Morse (1959), and Kesten and Morse (1959) were the first to study the problem of selecting the most probable cell in the so-called indifference-zone framework. Cacoulos and Sobel(1966) considered the so-called inverse-sampling procedure for the same problem. Recently, Ramey and Alam (1979, 1980) proposed sequential sampling procedures for selectin gthe most probable cell.

While various selection procedures have been proposed and studied for selecting the most probable cell, the problem of selecting the least probable cell has not been studied much. Following Bechhofer et al. (1959), Alam and Thompson (1972) considered the so-called preference zone

$$\Omega(\Delta) = \{(p_1, \cdots, p_k) \mid p_{[2]} - p_{[1]} \geq \Delta\}$$

* Department of Computer Science and Statistics, Seoul National University. This work was supported in part by the Ministry of Educations, Korean Government, through the Research Institue of Basic Sciences, Seoul National University.

where $0<\Delta<(k-1)^{-1}$ is given and $p_{[1]}\leq\cdots\leq p_{[k]}$ are the ordered values of $p_1, \cdots, p_k$ and they considered a fixed sample size procedure which guarantees a high probability of selecting the least probable cell as long as the cell probabilities $p_1, \cdots, p_k$ are in $\Omega(\Delta)$. We call such a method the indifference-zone approach. Although it has a certain statistical meaning, there is also a criticism on the pre-specification of the preference zone.

In this paper, we study the problem of selecting the least probable cell in a multinomial distibution from a Bayesian approach. The formulation of the problem and some notations are introduced in Section 2. In Section 3, the Bayes sequential selection rule is derived and the simplification of the rule is given. Section 4 consists of the comparison of the Bayes sequential selection rule with the best fixed sample size Bayes selection rule. Table I gives the continuation sets with respect to the vague prior for certain cases, and Table II shows how much savings can be obtained by the Bayes sequential selection rule. Also, the FORTRAN subroutine program has been prepared for the practical implementation of the Bayes sequential selection rule, and it is available upon request.

2. Statement of the problem.

This section gives a decision theoretic formulation of the problem along with the introduction of some notations to be used.

Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \cdots$ be a sequence of independent and identically distributed random vectors from the multinomial distribution $m(1, p_1, \cdots, p_k)$, $p_1+\cdots+p_k=1$, $p_i>0$, $1=1, \cdots, k$. we consider the problem of selecting the least probable cell $\pi_{[1]}$ associated with $p_{[1]}=\min_{1\leq i\leq k} p_i$ based on the observations which are taken one at a time. It is clear that for any fixed n the joint probability function of $\underset{\sim}{X}_1, \cdots, \underset{\sim}{X}_n$ is given by

$$f(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n|p_1, \cdots, p_k)=p_1^{t_{1n}}\cdots p_k^{t_{kn}} \tag{2.1}$$

where t_{in} denotes the cell frequency i.e.,

$$(t_{1n}, \cdots, t_{kn})=\sum_{j=1}^{n} \underset{\sim}{x}_j.$$

Let $c>0$ denote the relative cost of sampling per unit observation as regards to how much different the selected p_i is from $p_{[1]}$. Then the loss is given by

(2. 2) $$L((n,\ d_i),\ \underset{\sim}{p})=nc+(p_i-p_{[1]})$$

where (n, d_i) denotes the decision that the cell π_i associated with p_i is selected as the least probable cell after having observed n units.

We shall try to find a rule which minimizes the average loss given by (2. 2) from a Bayesian point of view. We assume a Dirichlet prior $D(\alpha_1, \cdots, \alpha_k)$ for $\underset{\sim}{p}$ whose density function is given by

(2. 3) $$g(\underset{\sim}{p})=\frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)}\,p_1^{\alpha_1-1}\cdots p_k^{\alpha_k-1},$$
$$\alpha=\alpha_1+\cdots+\alpha_k,\quad \alpha_i>0,\quad i=1,\cdots,k.$$

3. Bayes sequential selection rule.

In this section, the Bayes sequential selection rule with respect to the Dirichlet prior $D(\alpha_1, \cdots, \alpha_k)$ and the loss function given by (2. 2) is derived, and the simplification of it is given.

Since the Bayes sequential selection rule consists of the terminal selection rule and the stopping rule, we first consider the terminal selection rule. It follows from the loss function given by (2. 2) that the Bayes rule for fixed sample size n is determined by minimizing the posterior expected value of p_i, i.e.

(3. 1) $$E(p_i\,|\,\underset{\sim}{X}_1=\underset{\sim}{x}_1,\ \cdots,\ \underset{\sim}{X}_{1n}=\underset{\sim}{x}_n)=(\alpha_i+t_{in})/(\alpha+n),\quad i=1,\cdots,\ k.$$

Therefore, the next result follows from a theorem in Ferguson (1967), p. 314.

Lemma 1. After observing $\underset{\sim}{x}_1=\underset{\sim}{X}_1,\cdots,\ \underset{\sim}{X}_n=\underset{\sim}{x}_n$, the terminal selection rule selects the cell π_i if

(3. 2) $$\alpha_i+t_{in}=\min(\alpha_1+t_{1n},\ \cdots,\ \alpha_k+t_{kn})$$

where $$(t_{1n},\ \cdots,\ t_{kn})=\sum_{j=1}^{n}\underset{\sim}{x}_j.$$

Now we consider a Bayes sequential selection problem truncated at J as usually done in a sequential decision problem. First, we introduce some necessary notations. Define

(3. 3) $$U_n(\underset{\sim}{x}_1,\ \cdots,\ \underset{\sim}{x}_n)=\inf_{1\le i\le k}E\{L((n,\ d_i),\ \underset{\sim}{p})\,|\,\underset{\sim}{X}_1=\underset{\sim}{x}_1,$$
$$\cdots,\ \underset{\sim}{X}_n=\underset{\sim}{x}_n\},\quad n=0,1,\ \cdots,\ J.$$

It follows from Lemma 1 that $U_n(\underset{\sim}{x}_1,\ \cdots,\ \underset{\sim}{x}_n)$ can be written as

(3. 4) $$U_n(\underset{\sim}{x}_1,\ \cdots,\ \underset{\sim}{x}_n)=nc+\min(\alpha_i+t_{in})/(\alpha+n)$$

$$-E(p_{[1]} \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{X}_n = \underset{\sim}{x}_n).$$

Also, to use the backward induction, we define $V_n^{(J)}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)$ for $n=J, \ J-1, \ \cdots, \ 0$ inductively by

$$\text{(3.5)} \qquad \begin{cases} V_J^{(J)}(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_J) = U_J(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_J) \\ V_n^{(J)}(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n) = \min\{U_n(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n), \\ E(V_{n+1}^{(J)}(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n, \underset{\sim}{X}_{n+1}) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \ \cdots, \underset{\sim}{X}_n = \underset{\sim}{x}_n)\}, \end{cases}$$
$$n = J-1, \ J-2, \ \cdots, \ 0.$$

It is clear that $V_n^{(J)}(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n)$ denotes the minimum conditional Bayes risk in the truncated problem after having observed $\underset{\sim}{X}_1 = \underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{X}_n = \underset{\sim}{x}_n$. Obviously, $V_0^{(J)}$ denotes the minimum Bayes risk for the truncated problem. The next result gives a bound on J to obtain the Bayes sequential selection rule for the general problem.

Lemma 2. Let $V_0^{(\infty)}$ denote the minimum Bayes risk for the general problem specified by (2.2) and (2.3). Then,

$$V_0^{(J_0)} = V_0^{(\infty)}$$

where J_0 is the greatest integer such that $J_0 < \max\{c^{-1} - \alpha, 1\}$.

Proof. Since $p_i - p_{[1]}$ is bounded, it follows from a theorem in Ferguson (1967), p.318 that $V_0^{(J)} \to V_0^{(\infty)}$ as $J \to \infty$. Furthermore, it follows from(3.4) that

$$\begin{aligned} & U_{n-1}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}) - E(U_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}, \underset{\sim}{X}_n) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \cdots, \ \underset{\sim}{X}_{n-1} = \underset{\sim}{x}_{n-1}) \\ & = \min(\alpha_i + t_{i,n-1})/(\alpha + n - 1) - E(\min(\alpha_i + t_{i,n-1} + X_{in}) \mid \\ & \quad \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_{n-1} = \underset{\sim}{x}_{n-1})/(\alpha + n) - c \\ & \leqslant \min(\alpha_i + t_{i,n-1})/(\alpha + n - 1) - \min(\alpha_i + t_{i,n-1})/(\alpha + n) - c \\ & \leq (\alpha + n)^{-1} - c. \end{aligned}$$

Therefore, for all $n > J_0$,

$$\begin{aligned} & U_{n-1}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}) - E(U_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}, \underset{\sim}{X}_n) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \cdots, \\ & \underset{\sim}{X}_{n-1} = \underset{\sim}{x}_{n-1}) \leq 0. \end{aligned}$$

Thus the result follows from a theorem in Ferguson (1967) p. 322.

It follows from Lemma 1 and Lemma 2 that the Bayes sequential selection rule is completely determined by the backward induction, which can be summarized as follows:
(a) Stopping rule: Stop sampling after taking n $(0 \leq n \leq J_0 - 1)$ observations $\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n$ if and only if

$$U_n(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n) < E(V_{n+1}^{(J_0)}(\underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{x}_n, \underset{\sim}{X}_{n+1}) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \ \cdots, \ \underset{\sim}{X}_n = \underset{\sim}{x}_n)$$

and stop sampling for $n=J_0$.

(b) Terminal selection rule: After stopping with observations $\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n$, select the cell π_i if

$$\alpha_i+t_{in}=\min(\alpha_1+t_{1n}, \cdots, \alpha_k+t_{kn})$$

where $t_{1n}, \cdots, t_{kn}$ are the cell frequencies.

However, it is difficult to implement the above stopping rule due to the difficulty involved in the computation of $E(p_{[1]}|\underset{\sim}{X}_1=\underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_n=\underset{\sim}{x}_n)$.

Therefore we need to simplify the stopping rule. In order to do this, we introduce some notations as follows: Define

$$W_n^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)=nc+\min(\alpha_i+t_{in})/(\alpha+n), \quad n=0, 1, \cdots, J_0. \tag{3.6}$$

$$W_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)=E(p_{[1]}|\underset{\sim}{X}_1=\underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_n=\underset{\sim}{x}_n), \quad n=0, 1, \cdots, J_0. \tag{3.7}$$

$$\begin{cases} V_{J_0}^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{J_0})=W_{J_0}^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{J_0}). \\ V_n^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)=\min\{W_n^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n), \\ E(V_{n+1}^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n, \underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=\underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_n=\underset{\sim}{x}_n)\} \end{cases} \tag{3.8}$$

$$n=J_0-1, \; J_0-2, \; \cdots, \; 0.$$

It follows from (3.4) that

$$U_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)=W_n^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)-W_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n), \tag{3.9}$$

$$n=0, 1, \cdots, J_0.$$

Lemma 3. Let $V_n^{(J_0)}(\underset{\sim}{x}_1, \cdots \underset{\sim}{x}_n)$, $V_n^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)$ and $W_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)$ be defined by (3.5), (3.8) and (3.7), respectively. Then we have

$$V_n^{(J_0)}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)=V_n^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)-W_n(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n),$$
$$n=0, 1, \cdots, J_0.$$

Proof. We shall prove this lemma by induction. By (3.5), (3.8) and (3.9), the result holds for $n=J_0$. Suppose that it holds for a given $n\geq 1$. Then,

$$\begin{aligned} & V_{n-1}^{(J_0)}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}) \\ & =\min\{U_{n-1}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}), E(V_n^{(J_0)}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}, \underset{\sim}{X}_n)| \\ & \qquad \underset{\sim}{X}_1=\underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_{n-1}=\underset{\sim}{x}_{n-1})\} \\ & =\min\{W_{n-1}^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1})-W_{n-1}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}), \\ & \quad E(V_n^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_{n-1}, \underset{\sim}{X}_n)|\underset{\sim}{X}_1=\underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_{n-1}=\underset{\sim}{x}_{n-1}) \end{aligned}$$

$$-E(E(p_{[1]}|\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_{n-1}=x_{n-1},\underset{\sim}{X}_n))\}$$
$$=\min\{W^*_{n-1}(x_1,\cdots,x_{n-1}),E(V^{(J_0)^*}_n(x_1,\cdots,x_{n-1},\underset{\sim}{X}_n)|$$
$$\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_{n-1}=x_{n-1})\}-W_{n-1}(x_1,\cdots,x_{n-1})$$
$$=V^{(J_0)*}_{n-1}(x_1,,\cdots x_{n-1})-W_{n-1}(x_1,\cdots,x_{n-1}).$$

Hence, it holds for $n-1$, which completes the proof.

Lemma 4. For $U_n(x_1,\cdots,x_n)$ and $V^{(J_0)}_n(x_1,\ \cdots,\ x_n)$ given by (3.4) and (3.5), the following relation holds:

$$U_n(x_1,\cdots,x_n)<E(V^{(J_0)}_{n+1}(x_1,\cdots,x_n,X_{n+1})\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)\}$$

if and only if

$$W^*_n(x_1,\cdots,x_n)<E(V^{(J_0)*}_{n+1}(x_1,\cdots,x_n,\underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=x_1,,\cdots,\underset{\sim}{X}_n=x_n)$$

Proof. From Lemma 3, we have

$$E(V^{(J_0)}_{n+1}(x_1,\cdots,x_n,\underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)$$
$$=E(V^{(J_0)*}_{n+1}(x_1,\cdots,x_n,\underset{\sim}{X}_{n+1})-W_{n+1}(x_1,\cdots,x_n,\underset{\sim}{X}_{n+1})|$$
$$\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)$$
$$=E(V^{(J_0)*}_{n+1}(x_1,\cdots,x_n,\underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)-W_n(x_1,\cdots,x_n),$$

which completes the proof by (3.9).

Note that $E(V^{(J_0)*}_{n+1}(x_1,\cdots,\ x_n,\underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=x_1,\cdots,\ \underset{\sim}{X}_n=x_n)$ can be computed as follows: By (2.1), (2.3), the marginal distribution of $\underset{\sim}{X}_1,\cdots,\ \underset{\sim}{X}_{n+1}$ is given by

(3.10) $$f(x_1,\cdots,x_{n+1})=\frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)}\cdot\frac{\Gamma(\alpha_1+t_{1,n+1})\cdots\Gamma(\alpha_k+t_{k,n+1})}{\Gamma(\alpha+n+1)},\quad n=0,1,2,\cdots$$

so that the conditional distribution of $\underset{\sim}{X}_{n+1}$, given $\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n$, is represented as

(3.11) $$f(x_{n+1}|\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)=(\alpha_i+t_{in})/(\alpha+n)\ \text{ if } x_{n+1}=\underset{\sim}{e}_i$$

where $\underset{\sim}{e}_i$ is the k-dimensional unit vector of which the i^{th} element is 1.

Thus, by (3.11),

(3.12) $$E(V^{(J_0)*}_{n+1}(x_1,\cdots,\ x_n,\underset{\sim}{X}_{n+1})|\underset{\sim}{X}_1=x_1,\cdots,\underset{\sim}{X}_n=x_n)=\sum_{i=1}^{k}\frac{\alpha_i+t_{in}}{\alpha+n}V^{(J_0)^*}_{n+1}(x_1,\cdots,x_n,\underset{\sim}{e}_i).$$

Now, we can summarize all the results so far in the following theorem.

Theorem 1. The Bayes sequential selection rule for the general problem specified by (2.2) and (2.3) is given as follows:

(a) Stopping rule: Stop sampling after taking n $(0\leq n\leq J_0-1)$observations $x_1,\cdots,x_n$ if and only if

$$W_n^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n) < E(V_{n+1}^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n, \underset{\sim}{X}_{n+1}) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_n = \underset{\sim}{x}_n)$$

where $W_n^*(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n)$ and $E(V_{n+1}^{(J_0)*}(\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n, \underset{\sim}{X}_{n+1}) \mid \underset{\sim}{X}_1 = \underset{\sim}{x}_1, \cdots, \underset{\sim}{X}_n = \underset{\sim}{x}_n)$ are given by (3.6) and (3.12), and stop sampling for $n = J_0$.

(b) Terminal selection rule: After stopping with observations $\underset{\sim}{x}_1, \cdots, \underset{\sim}{x}_n$, select the cell π_i if

$$\alpha_i + t_{in} = \min(\alpha_1 + t_{1n}, \cdots, \alpha_k + t_{kn})$$

where $t_{1n}, \cdots, t_{kn}$ are the cell frequencies.

Note that, for a symmetric prior $D(\alpha_0, \cdots, \alpha_0)$, the implementation of the Bayes sequential selection rule can be simplified considerably. For a symmetric prior, the Bayes sequential selection rule is invariant under the permutation of $(t_{1n}, \cdots, t_{kn})$. Moreover, the truncation number J_0 turns out to be much smaller. In fact, the inequality

$$\min(\alpha_0 + t_{i,n-1}) \leq \alpha_0 + (n-1)/k$$

can be used to prove the next result just as Lemma 2 was proved.

Lemma 5. For the symmetric Dirichlet prior $D(\alpha_0, \cdots, \alpha_0)$, the truncation number J_0 in Theorem 1 is the largest integer such that

(3.13) $$J_0 < \max\{(kc)^{-1} - k\alpha_0, 1\}.$$

Lemma 5 is used to find the continuation sets, that is, the sets of the cell counts where the sampling should be continued, for $k = 2(1)5$, $c = 0.01, 0.008$ and $\alpha_0 = 1$ in Table I.

4. Comparative study for the vague prior.

As pointed out in Section 1, many selection rules have been proposed for the problem of selecting the most probable cell. Mainly, they are different in sampling rules-fixed sample size, inverse and Bayes sequential sampling rules. The results of the comparative study have been reported in Ramey and Alam (1980). However, for selecting the least probable cell, it seems clear that the inverse sampling scheme cannot work. In fact, no inverse sampling rule has been devised yet for the problem of selecting the least probable cell. Now, for such a problem, we have two selection rules at hand, namely, the Bayes sequential selection rule (BSSR) proposed in this paper and the fixed sample size rule of Alam and Thompson (1972). of course, the fixed sample size rule has been proposed in a different framework, i.e., the indifference-zone approach, which is different from that in this paper, i.e., the Bayesian approach. However, it would be interesting to compare them and see how much saving can be obtained by BSSR.

In this section, we make comparisons of BSSR and the best fixed sample size Bayes selection rule (BFBSR) with respect to vague prior $D(1,\cdots,1)$. Let V_F^n denote the Bayes risk of the fixed sample size Bayes selection rule where the sample size is n, and let $V_F=\min_n V_F^n$ denote the Bayes risk of BFBSR. Then, as was done in Ramey and Alam (1980), we consider ϕ defined by

$$(4.1)\qquad \phi=\frac{V_F-V^{(J_0)}}{V_F}\times 100$$

where $V_0^{(J_0)}$ denotes the risk of BSSR.

The value of ϕ represents the savings by BSSR in terms of percentage.

To obtain the value of ϕ, We consider the computation of $V_0^{(J_0)}$ in the first place. It follows from Lemma 3 that $V_0^{(J_0)}$ is given by

$$(4.2)\qquad V_0^{(J_0)}=V_0^{(J_0)*}-W_0$$

where W_0 and $V_0^{(J_0)*}$ are defined by (3.7) and (3.8), respectively. In order to compute W_0, note that if $\underset{\sim}{p}\sim D(1,\cdots,1)$, then we can represent $p_1,\cdots,p_k$ as

$$p_i=Z_i/(Z_1+\cdots+Z_k),\quad i=1,\cdots,k$$

where $Z_1,\cdots,Z_k$ are IID exponential random variables with mean 1. It follows from the independence between p_i and $Z_1+\cdots+Z_k$ that

$$(4.3)\qquad \begin{aligned} W_0&=E(\min p_i)\\ &=E(\min Z_i)/E(Z_1+\cdots+Z_k)\\ &=k^{-2}. \end{aligned}$$

Thus, from (4.2), we have

$$(4.4)\qquad V_0^{(J_0)}=V_0^{(J_0)*}-k^{-2}$$

where $V_0^{(J_0)*}$ can be computed easily. Next, we turn to the computation of $V_F=\min V_F^n$. It follows from (2.2) and (3.1) that V_F^n is given by

$$(4.5)\qquad V_F^n=nc-\ W_0+\{1+E(\min_i\ T_{in})\}/(n+k)$$

where the expectation is taken with respect to the marginal distribution of $T_{1n},\ \cdots,\ T_{kn}$. It is obvious by (3.10) that the marginal distribution of $T_{1n},\ \cdots,\ T_{kn}$ is given by

$$p(T_{1n},=t_{1n},\ \cdots,\ T_{kn}=t_{kn})=({}_kH_n)^{-1}$$

where $\quad {}_kH_n=(n+k-1)!/\{(k-1)!n!\}.$

Thus, we have

$$(4.6)\qquad V_F^n = nc - \frac{1}{k^2} + \frac{1}{k+n}\left\{1 + \frac{1}{{}_kH_n}\left\{\sum_{m=0}^{[\frac{n}{k}]-1} m\cdot({}_kH_{n-km} - {}_kH_{n-k(m+1)}) + \left[\frac{n}{k}\right]{}_kH_{n-k[\frac{n}{k}]}\right\}\right\}$$

where $[x]$ denotes the greatest integer function. Hence, we can compute $V_0^{(J_0)}$ by (4.4) and $V_F = \min_n V_F^n$ by (4.6).

Table Ⅱ below gives the values of $V_0^{(J_0)}$, V_F and ϕ for $c=0.002(0.001)0.01$, $k=2(1)5$ with respect to the vague prior $D(1, \cdots, 1)$.

One final remark should be in order. We assumed in this comparative study the vague prior for computational simplicity. However, the same method can be applied to the symmetric prior $D(\alpha_0, \cdots, \alpha_0)$, if desired. In such a case, we need the table by Gupta(1960) for the values of W_0.

Table Ⅰ. **Continuation sets for the uniform prior** ($\alpha_0=1$).

(* means "No continuation sample point for n".) (A) $c=0.01$

n \ k	2	3	4	5
0	(0, 0)	(0, 0, 0)	(0, 0, 0, 0)	(0, 0, 0, 0, 0)
1	(0, 1)	(0, 0, 1)	(0, 0, 0, 1)	(0, 0, 0, 0, 1)
2	(1, 1)	(0, 1, 1) (0, 0, 2)	(0, 0, 1, 1), (0, 0, 0, 2)	(0, 0, 0, 1, 1), (0, 0, 0, 0, 2)
3	(1, 2)	(1, 1, 1) (0, 0, 3)	(0, 0, 1, 2), (0, 0, 0, 3)	(0, 0, 1, 1, 1), (0, 0, 0, 1, 2) (0, 0, 0, 0, 3)
4	(2, 2)	(1, 1, 2) (0, 0, 4)	(0, 0, 2, 2), (0, 0, 1, 3) (0, 0, 0, 4)	(0, 0, 1, 1, 2), (0, 0, 0, 2, 2) (0, 0, 0, 1, 3), (0, 0, 0, 0, 4)
5	(2, 3)	(1, 2, 2)(1, 1, 3)(0, 0, 5)	(0, 0, 2, 3), (0, 0, 1, 4) (0, 0, 0, 5)	*
6	(3, 3)	(2, 2, 2) (1, 1, 4) (0, 0, 6)	*	
7	(3, 4)	(2, 2, 3) (1, 1, 5)		
8	(4, 4)	(2, 2, 4) (1, 1, 6)		
9	(4, 5)	(2, 2, 5) (1, 1, 7)		
10	(5, 5)	(2, 2, 6) (1, 1, 8)		
11	(5, 6)	(2, 2, 7)		
12	(6, 6)	(2, 2, 8)		
13	*	(2, 2, 9)		
14		*		

Table Ⅰ (Continued). (B) $C=0.008$

n \ k	2	3	4	5
0	(0, 0)	(0, 0, 0)	(0, 0, 0, 0)	(0, 0, 0, 0, 0)
1	(0, 1)	(0, 0, 1)	(0, 0, 0, 1)	(0, 0, 0, 0, 1)

2	(1, 1)	(0, 1, 1), (0, 0, 2)	(0, 0, 1, 1), (0, 0, 0, 2)	(0, 0, 0, 1, 1), (0, 0, 0, 0, 2)
3	(1, 2)	(1, 1, 1), (0, 0, 3)	(0, 0, 1, 2), (0, 0, 0, 3)	(0, 0, 1, 1, 1), (0, 0, 0, 1, 2) (0, 0, 0, 0, 3)
4	(2, 2)	(1, 1, 2), (0, 0, 4)	(0, 0, 2, 2), (0, 0, 1, 3) (0, 0, 0, 4)	(0, 0, 1, 1, 2), (0, 0, 0, 2, 2) (0, 0, 0, 1, 3), (0, 0, 0, 0, 4)
5	(2, 3)	(1, 2, 2), (1, 1, 3) (0, 0, 5)	(0, 0, 2, 3), (0, 0, 1, 4) (0, 0, 0, 5)	(0, 0, 1, 2, 2), (0, 0, 1, 1, 3) (0, 0, 0, 2, 3), (0, 0, 0, 1, 4) (0, 0, 0, 0, 5)
6	(3, 3)	(2, 2, 2), (1, 1, 4), (0, 0, 6)	(0, 0, 3, 3), (0, 0, 2, 4) (0, 0, 1, 5), (0, 0, 0, 6)	*
7	(3, 4)	(2, 2, 3), (1, 1, 5), (0, 0, 7)	*	
8	(4, 4)	(2, 3, 3), (2, 2, 4), (1, 1, 6)		
9	(4, 5)	(3, 3, 3), (2, 2, 5), (1, 1, 7)		
10	(5, 5)	(3, 3, 4), (2, 2, 6), (1, 1, 8)		
11	(5, 6)	(3, 3, 5), (2, 2, 7), (1, 1, 9)		
12	(6, 6)	(3, 3, 6), (2, 2, 8), (1, 1, 10)		
13	(6, 7)	(3, 3, 7), (2, 2, 9)		
14	(7, 7)	(3, 3, 8), (2, 2, 10)		
15	(7, 8)	(3, 3, 9), (2, 2, 11)		
16	(8, 8)	(3, 3, 10)		
17	*	(3, 3, 11)		
18		(3, 3, 12)		
19		*		

Table Ⅱ. Values of $V_0^{(J_0)}$, V_F and ϕ for the vague prior ($\alpha_0=1$).

c	$k=2$			$k=3$			$k=4$			$k=5$		
	$V_0^{(J_0)}$	V_F	ϕ	$V_0^{(J_0)}$	V_F	ϕ	$V_0^{(J_0)}$	V_F	ϕ	$V_0^{(J_0)}$	V_F	ϕ
.002	.0290	.0407	28.8	.0396	.0483	17.9	.0455	.0516	11.8	.0496	.0539	8.0
.003	.0362	.0488	25.8	.0489	.0583	16.2	.0560	.0624	10.3	.0610	.0651	6.3
.004	.0420	.0557	24.7	.0567	.0663	14.5	.0645	.0710	9.1	.0706	.0744	5.1
.005	.0475	.0607	21.7	.0634	.0737	13.9	.0721	.0786	8.3	.0789	.0821	3.9
.006	.0528	.0657	19.6	.0694	.0796	2.81	.0788	.0854	7.7	.0864	.0891	3.0
.007	.0574	.0707	18.8	.0749	.0846	11.4	.0851	.0914	7.0	.0931	.0951	2.0
.008	.0605	.0740	18.2	.0768	.0896	10.9	.0908	.0965	6.0	.0993	.1008	1.5
.009	.0637	.0770	17.3	.0841	.0946	11.1	.0961	.1016	5.4	.1051	.1058	0.7
.010	.0668	.0800	16.5	.0882	.0996	11.4	.1011	.1061	4.7	.1101	.1108	0.6

(Received May 1982; Revised October 1982)

REFERENCES

(1) Alam, K. and Thompson, J.R. (1972). "On selecting the least probable multinomial event."

Annals Mathematical Statistcs, 43, 1981—1990.

(2) Bechhofer, R.E., Elmaghraby, S.A. and Morse, N. (1959). "A single sample multiple-decision procedure for selecting the multinomial event which has the largest probability." Annals Mathematical Statistics, 30, 102—119.

(3) Cacoulos, T. and Sobel,M.(1966). "An inverse sampling procedure for selecting the probable event in a multinomial distribution." Multivariate analysis : Proceedings of an International Symposium(P.R.Krishnaiah), Academic Press, New York, 423—455.

(4) Ferguson, T.S.(1967). "Mathematical Statistics - A decision theoretic approach". Academic Press, New York and London.

(5) Gupta, S.S.(1960). "Order Statistics from the Gamma Distribution." Technometrics, 2, 243—262.

(6) Kesten,H. and Morse,N. (1959). "A property of multinomial distribution". annals Mathematical Statistics, 30, 120—127.

(7) Ramey, Jr., J.T. and Alam, K. (1979). "A sequential procedure for selecting the most probable multinomial event." Biometrika, 66, 171—173.

(8) Ramey, Jr., J.T. and Alam,K. (1980), "A Bayes sequetial proceure for selecting the most probable multinomial event." Communications in statistics, A9(3), 265—276.