SUBSET SELECTION PROCEDURES FOR THE BINOMIAL MODELS BASED ON A CLASS OF PRIORS AND SOME APPLICATIONS *

by

Shanti S. Gupta and Yuning Liao
Purdue University

Technical Report #90-56C

Department of Statistics
Purdue University

October 1990
Revised May 1991

* Research supported in part by the NSF Grants DMS–8702620 and DMS–8717799 at Purdue University.

⋆

Subset Selection Procedures for the Binomial Models based on A Class of Priors and Some Applications *

Shanti S. Gupta and Yuning Liao
Department of Statistics, Purdue University,
West Lafayette, Indiana 47907, U.S.A.

Abstract

The problem of selecting a subset containing the best of $k(\geq 2)$ binomial populations is studied. The approach is more general than the classical subset selection procedures studied by Gupta and Sobel (1960) and Gupta and McDonald (1986). In these preceding papers, the infimum of the probability of a correct selection occurs when all the parameters are equal (in the limit) to the largest unknown parameter. Thus it is natural to formulate the problem on the assumption that the largest unknown parameter follows a prior. Several priors have been considered and the associated procedures have been evaluated. Performance comparisons are made between the classical and the new procedures. Applications of this approach to the control problem and Poisson models are provided.

1 Introduction

In general, in order to solve a ranking and selection problem, we usually need to find the least favorable configuration (LFC) (more details can be found in Bechhofer (1954), Gupta (1956)

*Research supported in part by the NSF Grants DMS-8702620 and DMS-8717799 at Purdue University.

and Gupta and Panchapakesan (1979), etc). Fortunately, the LFC is easy to find for many distributions under some assumptions. Then a lower bound for the probability of a correct selection can be calculated under the LFC and the preassigned probability $P^*(1/k < P^* < 1)$ of a correct selection(CS) is guaranteed by choosing the lower bound under the LFC to be at least equal the required P^* value, where the computation of the lower bound on the probability of a correct selection under the LFC does not depend on the parameter space. The location and scale-type parameter models are two well formulated examples.

However, if the LFC does not provide a usable lower bound for the probability of a correct selection (like the decision problem where we want to pick the population associated with the largest parameter from several Poisson populations), or if it is very difficult to determine (like the decision problem of the Binomial populations), then the computation of the lower bound on the probability of a correct selection turns out to be very difficult, and then the classical selection approach even can not be applied. The main reason for this difficulty in the computation of any usable lower bound for the probability of a correct selection is its dependence on the parameter $\underline{\theta}$ itself.

In the following, we study the subset selection approach for the problem of selecting the best population from among k binomial populations, where we introduce the prior information into the problem. We formulate the problem based on the prior distribution of the largest unknown parameter $\theta_{[k]}$.

2 Binomial Populations

The classical subset selection procedure for selecting the population associated with the largest success probability from several binomial populations has been studied by Gupta and Sobel (1960).

Let $\pi_1, \pi_2, \cdots, \pi_k$ denote k independent binomial populations. For each $i = 1, 2, \cdots, k$, let X_i denote the observed number of successes based on n independent observations from population π_i. Then X_i follows a binomial distribution with probability function $f(x|p_i)$, where

$$f(x|p_i) = \binom{n}{x} p_i^x (1-p_i)^{n-x},\ x = 0, 1, \cdots, n;\ 0 < p_i < 1. \tag{1}$$

Let $p_{[1]} \leq p_{[2]} \leq \cdots \leq p_{[k]}$ be the ranked values of the unknown p_i values, $i = 1, 2, \cdots, k$. The best population is the one associated with the largest parameter $p_{[k]}$. Again, the goal is to select a nonempty subset of $\pi_1, \pi_2, \cdots, \pi_k$, which contains the best population with a minimum guaranteed probability $P^*(1/k < P^* < 1)$.

In Gupta and Sobel(1960)'s paper, the following selection rule was used

$$R^{max}: \quad \text{Select } \pi_i, \text{ if and only if}: X_i \geq \max_{1 \leq j \leq k} X_j - d,$$
$$\text{for some integer } d(0 \leq d \leq n).$$

Let $X_{(i)}$ be the unknown observation associated with $p_{[i]}$, $i = 1, 2, \cdots, k$. Then

$$\begin{aligned}
& P(CS|R^{max}) = P\left(X_{(k)} \geq X_{(j)} - d,\ j = 1, 2, \cdots, k-1\right) \\
= & \sum_{x=0}^{n} \binom{n}{x} p_{[k]}^x \left(1 - p_{[k]}\right)^{n-x} \prod_{j=1}^{k-1} \left\{ \sum_{t=0}^{x+d} \binom{n}{t} p_{[j]}^t \left(1 - p_{[j]}\right)^{n-t} \right\} \\
\geq & \sum_{x=0}^{n} \binom{n}{x} p_{[k]}^x \left(1 - p_{[k]}\right)^{n-x} \left\{ \sum_{t=0}^{x+d} \binom{n}{t} p_{[k]}^t \left(1 - p_{[k]}\right)^{n-t} \right\}^{k-1}.
\end{aligned} \tag{2}$$

In order to meet the requirement that $P(CS|R^{\max}) \geq P^*$, it suffices to find an integer $d(0 \leq d \leq n)$ such that

$$\inf_{0 \leq p \leq 1} P(CS|p, d) \geq P^*,$$

where

$$P(CS|p, d) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \left\{ \sum_{j=0}^{x+d} \binom{n}{j} p^j (1-p)^{n-j} \right\}^{k-1}. \tag{3}$$

Recently, the so-called Gupta-Sobel rule R^{max} have been extensively studied by Gupta and Mcdonald (1986). Many useful values have been tabulated by the authors. Here, we will show that the classical subset selection procedure is a special procedure resulting from a new approach described in the paper.

Carefully examining the inequality (2), we can see that the infimum of the probability of a correct selection occurs when all the parameters are equal (in the limit) to the largest unknown parameter. Therefore, it is natural to formulate the selection problem on the assumption that the largest unknown parameter follows a prior.

Let $G(\cdot)$ be the prior distribution on the space of the largest parameter $p_{[k]}$. The $P(CS)$, using the natural subset selection rule R^{max}, has the following lower bound

$$\begin{aligned}
& P(CS|d, G, R^{max}) \\
= & \int_0^1 P\left(X_{(k)} \geq X_{(j)} - d,\ j = 1, 2, \cdots, k-1 | p_{[k]}\right) dG(p_{[k]}) \\
= & \int_0^1 \left(\sum_{x=0}^{n} \binom{n}{x} p_{[k]}^x \left(1 - p_{[k]}\right)^{n-x} \prod_{j=1}^{k-1} \left\{ \sum_{t=0}^{x+d} \binom{n}{t} p_{[j]}^t \left(1 - p_{[j]}\right)^{n-t} \right\} | p_{[k]} \right) dG(p_{[k]}) \\
\geq & \int_0^1 \sum_{x=0}^{n} \binom{n}{x} p_{[k]}^x \left(1 - p_{[k]}\right)^{n-x} \left\{ \sum_{t=0}^{x+d} \binom{n}{t} p_{[k]}^t \left(1 - p_{[k]}\right)^{n-t} \right\}^{k-1} dG(p_{[k]}). \qquad (4)
\end{aligned}$$

Note that this inequality is similar to that of (2).

In the following, we classify the prior information about the largest (unknown) parameter $p_{[k]}$ into three categories

(a) the prior distribution of $p_{[k]}$ is completely known,

(b) the prior distribution of $p_{[k]}$ is totally unknown,

(c) the prior distribution of $p_{[k]}$ is partly known.

Where, the assumption (a) is though unrealistic, can still be used to obtain results for the case (b) and (c).

3 Results for the Prior Distribution $G(\cdot)$

From (4), we see that the prior distribution of $p_{[k]}$ determines a lower bound on the $P(CS)$. We will consider the problem, when $G(\cdot)$ falls into three categories.

(a) The distribution function $G(\cdot)$ is completely known.

For any given $P^*(1/k < P^* < 1)$, if $G(\cdot)$ is a proper distribution, then we can find the smallest integer $d (0 \leq d \leq n)$ satisfying

$$P(CS|d, G, R^{max}) \geq P^*.$$

Actually, we need only to find the smallest integer $d (0 \le d \le n)$ satisfying

$$P(n,k,d,G) \ge P^*,$$

where

$$P(n,k,d,G) = \int_0^1 \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \left\{ \sum_{t=0}^{x+d} \binom{n}{t} p^t (1-p)^{n-t} \right\}^{k-1} dG(p). \tag{5}$$

For a specific choice of $G(\cdot)$, we give the following proposition.

Proposition 3.1 *Let*

$$\mathcal{B}_e = \left\{ b_{\alpha,\beta}(p) = \frac{p^{(\alpha-1)}(1-p)^{(\beta-1)}}{B(\alpha,\beta)} \ : \ \alpha, \ \beta > 0 \right\},$$

where $B(\alpha,\beta) = [\Gamma(\alpha)\Gamma(\beta)]/\Gamma(\alpha+\beta)]$ is the beta function. Then, for any beta distribution $b_{\alpha,\beta} \in \mathcal{B}_e$, we have

$$P(n,k,d,b_{\alpha,\beta}) = \sum_{x=0}^{n} \binom{n}{x} \sum_{h=0}^{(k-1)(x+d)} \mathcal{H}(h) \mathcal{R}_{\alpha,\beta}(h,x),$$

where

$$\mathcal{H}(h) = \prod_{j_1,j_2,\cdots,j_{k-1}<h}^{j_1+j_2+\cdots+j_{k-1}=h} \binom{n}{j_l},$$

and

$$\mathcal{R}_{\alpha,\beta}(h,x) = \frac{B(h+x+\alpha, nk-h-x+\beta)}{B(\alpha,\beta)}.$$

(b) The distribution function $G(\cdot)$ is totally unknown.

When G is unknown, then

$$\min_{G \in \mathcal{G}} P(n,k,d,G) = \min_{0<p<1} P(CS|p,d),$$

where $\mathcal{G}$ is the class of all distributions on the interval $(0,1)$. Then it is exactly the lower bound used in Gupta and Sobel (1960). Or one may use the "*non-informative*" priors, one of them is

$$G(p) = 1_{\{0<p<1\}},$$

which is a member of $\mathcal{B}_e$, if $\alpha = \beta = 1$. Then, we have

$$P(n,k,d,\mathcal{U}) = \frac{\sum_{x=0}^{n}\sum_{j_1=0}^{x+d}\cdots\sum_{j_{k-1}}^{x+d}\binom{n}{x}\binom{n}{j_1}\cdots\binom{n}{j_{k-1}}}{(nk+1)\binom{nk}{x+j_1+\cdots+j_{k-1}}}.$$

Another "noninformative" prior is the so-called Jeffreys prior with the following density function

$$\begin{aligned} G_J(p) &= \frac{[p(1-p)]^{-\frac{1}{2}}}{B(\frac{1}{2},\frac{1}{2})} \\ &= \frac{1}{\pi}[p(1-p)]^{-\frac{1}{2}}. \end{aligned}$$

Using the uniform prior and Jeffreys prior, we have computed the values of $P(n,k,d,G)$, for $n = 2(1)10$, $k = 2(1)5$, $d = 0,1,\cdots,n-1$ in Table 1. From the tables, we notice that the bounds are much larger than that of the classical procedure provided, for instance, when $n = 5$, $k = 3$, $d = 1$, the lower bound of the guarantee P(CS) will increase from 0.7265 for the classical procedure to 0.8188 for $G = U$(uniform), and to 0.8675 for $G = G_J$(Jeffreys)(see the column of $d = 1$ in Table 2). For other choices of n, k, d, the values of $P(n,k,d,G)$, when $G = U(G_J)$, can also be easily calculated from formula (5).

(c) The distribution function $G(\cdot)$ is partly known.

(i) The support (say $A \subset [0,1]$) of the distribution $G(\cdot)$ is known, but the exact form of the distribution function $G(\cdot)$ is unknown.

Then, either the infimum of $P(CS|d,p)$ on A, or a uniform prior on A can be used. Actually, it is reasonable for the experimenter to know that the support of G is not the whole parameter space. For instance, we may know that all the populations have their success rate very high, say at least larger than $\frac{3}{4}$, so the support of the largest (unknown) $p_{[k]}$ is not the whole interval $[0,1]$, but a subset of it (i.e. $[\frac{3}{4},1]$).

(ii) $G(\cdot)$ is in a ε-contamination class of a known distribution G_0, say, G belongs to the class Γ, where

$$\Gamma = \{G: \; G = (1-\varepsilon)G_0 + \varepsilon Q, \; Q \in \mathcal{Q}\}. \tag{6}$$

The distribution G_0 can be any known proper function. For the choosing of $\mathcal{Q}$, one may consider the case (b) and case (c)(i) of above.

(iii) Another interesting choice of the $G(\cdot)$ is: $G \in \mathcal{Q}_{a_0,b_0}$, where

$$\mathcal{Q}_{a_0,b_0} =$$
$$\{Q: \ Q \text{ is uniform on } (max\{0, a_0 - \varepsilon\}, min\{1, b_0 + \varepsilon\}), \ 0 \le a_0 \le b_0 \le 1, \varepsilon > 0\}.$$

It can be used in situations when the practitioner can specify an interval (or a point), say $[a_0, b_0]$, within which one is confident that the parameter $p_{[k]}$ most likely falls in.

Figure 1 and Figure 2 show some behaviors of $P(n, k, d, Q^{\varepsilon}_{a_0})$ against ε, when $a_0 = 0(1)$ and $a_0 = \frac{1}{4}(\frac{3}{4})$, where $Q^{\varepsilon}_{a_0}$ is the uniform distribution function on the interval $(max\{0, a_0 - \varepsilon\}, min\{1, a_0 + \varepsilon\})$.

Table 2 shows how different procedures (for different priors) perform. Where we can see that if the largest (unknown) parameter $p_{[k]}$ is believed to be larger than .75, the lower bounds of P(CS) provided by the classical procedure will definitely need to be replaced by $\inf_{.75<p<1} P(CS|p, d)$, which is much better than the one the classical procedure provided.

The above approach can be easily applied to the following problems.

4 Selecting Populations Better than A Control

Let $p_0 \in (0, 1)$ denote the unknown probability of a unit being defective in the control population π_0. Population π_i is said to be *good* if $p_i \le p_0$, and *bad* if $p_i > p_0$, $i = 1, 2, \cdots, k$ (see Gupta and Sobel (1958)). Here, we assume that $\pi_1, \pi_2, \cdots, \pi_k$ are k independent binomial populations, and for each i, $i = 1, 2, \cdots, k$, we have a random observation X_i, which is the number of defectives observed in the sample of n independent observations arising from the population π_i.

For the parameter p_0 associated with the control population π_0, we assume that there is a proper prior distribution $G(\cdot)$ for p_0, In general, if $G(\cdot)$ is not degenerated, then we have to assume that there is a sample X_0 from population π_0 which follows a distribution $F(\cdot)$, where $F(\cdot)$ may not be a binomial distribution, but must be characterized only by the control parameter p_0 (i.e. p_0 is the only parameter of F). And p_0 can be estimated by X_0, for example, X_0 is the unbiased estimate of p_0. Here, we only concern the case where F, again, is binomial with parameter n and p_0.

Then, using the natural subset selection rule R^{max} defined as follows:

$$R^{max}: \quad \text{Retain } \pi_i, \quad \text{if and only if}: \quad X_i < X_0 + d, \quad \text{for some } d.$$

Like in Gupta and Sobel(1958)'s paper, we define a correct selection(CS) to be that all $k_1(\leq k)$ good populations are selected in the selected subset. So, the event CS will be that all k_1 good populations have their X_i values less than $X_0 + d$, therefore

$$\begin{aligned}
&P(CS|R^{max}, G, F, d) \\
&= \int_{\mathcal{P}_0}\int_{\mathcal{X}_0} P\{X_i < X_0 + d,\ i = 1, 2, \cdots, k_1 |\ p_0, x_0\}\, dF(x_0)dG(p_0) \\
&= \int_{\mathcal{P}_0}\int_{\mathcal{X}_0} \prod_{i=1}^{k_1}\left[\sum_{t=0}^{x_0+d}\binom{n}{t}p_i^t(1-p_i)^{n-t} \mid p_0, x_0\right] dF(x_0)dG(p_0) \\
&\geq \int_{\mathcal{P}_0}\int_{\mathcal{X}_0}\left[\sum_{t=0}^{x_0+d}\binom{n}{t}p_0^t(1-p_0)^{n-t}\right]^{k_1} dF(x_0)dG(p_0) \\
&\geq \int_{\mathcal{P}_0}\int_{\mathcal{X}_0}\left[\sum_{t=0}^{x_0+d}\binom{n}{t}p_0^t(1-p_0)^{n-t}\right]^{k} dF(x_0)dG(p_0), \qquad (7)
\end{aligned}$$

where $k = k_1 + k_2$, with k_1 being the number of the populations that is better than the standard. $\mathcal{P}_0$ and $\mathcal{X}_0$ are supports of p_0 and x_0, respectively.

To meet the so-called P^*-requirement, one need only to find the smallest integer $d(0 \leq d \leq n)$ satisfying

$$\int_{\mathcal{P}_0}\int_{\mathcal{X}_0}\left[\sum_{t=0}^{x_0+d}\binom{n}{t}p_0^t(1-p_0)^{n-t}\right]^{k} dF(x_0)dG(p_0) \geq P^*.$$

Clearly, d exists, when both G and F are proper distributions.

In our case, since F is also a binomial distribution $(B(n, p_0))$, so (7) becomes

$$\begin{aligned}
&P(CS|R^{max}, G, F, d) \\
&\geq \int_{\mathcal{P}_0}\sum_{x=0}^{n}\binom{n}{x}p_0^x(1-p_0)^{n-x}\left[\sum_{t=0}^{x+d}\binom{n}{t}p_0^t(1-p_0)^{n-t}\right]^{k} dG(p_0). \qquad (8)
\end{aligned}$$

Furthermore, if the prior G is chosen to minimize the probability function on the right hand side of (8), then the whole selection process coincides with that of the classical subset selection approach (Gupta and Sobel (1958)). That is: the desired value of d is the smallest integer for which

$$\min_{0\le p_0\le 1}\left\{\sum_{x=0}^{n}\binom{n}{x}p_0^x(1-p_0)^{n-x}\left[\sum_{t=0}^{x+d}\binom{n}{t}p_0^t(1-p_0)^{n-t}\right]^k\right\}\ge P^*. \tag{9}$$

For other situations, the following remark applies.

Remark 4.1 *For choosing the prior distribution G, one can proceed as in Section 3.*

5 Poisson Populations

It has been shown that the usual type of the selection procedures for selecting the largest parameter from among k Poisson populations do not exist for some values of the probability $P^*(1/k < P^* < 1)$ of a correct selection (Goel (1972)). Moreover, Leong and Wong (1977) showed that the infimum of the probability of a correct selection, using the classical type of selection procedure, is $1/k$ (see also Gupta, Leong and Wong (1979)).

Since then, some different approaches and techniques have been proposed by Gupta and Huang (1975), Goel (1975) and Gupta and Wong (1977) etc. In particular, in the paper of Gupta and Huang (1975), they successfully avoided the difficulty of finding

$$P(CS|R^{max})\ge \inf_{0<\lambda<\infty}P(CS|\lambda,d),$$

where

$$P(CS|\lambda,d)=\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}\left\{\sum_{t=0}^{x+d}\frac{e^{-\lambda}\lambda^t}{t!}\right\}^{k-1},$$

by using the scale-type subset selection rule and the following inequality

$$\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}\left\{\sum_{t=0}^{[\frac{x+1}{c}]}\frac{e^{-\lambda}\lambda^t}{t!}\right\}^{k-1}\ge\left\{\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}\sum_{t=0}^{[\frac{x+1}{c}]}\frac{e^{-\lambda}\lambda^t}{t!}\right\}^{k-1}.$$

In the following, we use the same technique which had been used in the previous section to handle the Poisson model by considering the prior information about the largest (unknown) parameter.

5.1 Formulation of the Problem

Let $\pi_1, \pi_2, \cdots, \pi_k$ be k Poisson populations with parameter $\lambda_1, \lambda_2, \cdots, \lambda_k$ respectively. For each π_i, let X_i denote a random observation arising from the ith population. It is assumed that X_i follows a Poisson distribution with probability mass function $f(x|\lambda_i)$, where

$$f(x|\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^x}{x!}, \qquad x = 0, 1, 2, \cdots; \quad \lambda_i > 0.$$

Let $\lambda_{[1]} \leq \lambda_{[2]} \leq \cdots \leq \lambda_{[k]}$ denote the ordered values of the unknown parameter λ_i's, $i = 1, 2, \cdots, k$. We assume that the largest (unknown) parameter $\lambda_{[k]}$ has a proper prior distribution $G(\cdot)$. Our goal is to select a nonempty subset which contains the population associated with the largest parameter $\lambda_{[k]}$ with at least a guaranteed probability $P^*(1/k < P^* < 1)$.

For the natural subset selection rule $R^{\max}$:

$$R^{\max}: \quad \text{Select } \pi_i, \quad \text{if and only if}: \quad X_i \geq \max_{1 \leq j \leq k} X_j - d, \quad \text{for some } d\ (> 0).$$

Let $X_{(i)}$ be the random variable associated with the unknown parameter $\lambda_{[i]}$, where $\lambda_{[1]} \leq \lambda_{[2]} \leq \cdots \leq \lambda_{[k]}$ is the ordered values of λ_i's, $i = 1, 2, \cdots, k$, then

$$\begin{aligned}
& P(CS|d, G, R^{\max}) \\
= & \int_0^\infty P\left(X_{(k)} \geq X_{(j)} - d, \ \forall\, j \neq k | \lambda_{[k]}\right) dG(\lambda_{[k]}) \\
= & \int_0^\infty \left[\sum_{x=0}^\infty \frac{e^{-\lambda_{[k]}}\lambda_{[k]}^x}{x!} \prod_{i=1}^{k-1} \left\{\sum_{t=0}^{x+d} \frac{e^{-\lambda_{[i]}}\lambda_{[i]}^t}{t!}\right\} |\lambda_{[k]}\right] dG(\lambda_{[k]}) \\
\geq & \int_0^\infty \sum_{x=0}^\infty \frac{e^{-\lambda}\lambda^x}{x!} \left\{\sum_{t=0}^{x+d} \frac{e^{-\lambda}\lambda^t}{t!}\right\}^{k-1} dG(\lambda),
\end{aligned}$$

where the last inequality follows from property (i) of the following lemma:

Lemma 5.1 *Let*

$$P(CS|\underline{\lambda}, d) = \sum_{x=0}^\infty \frac{e^{-\lambda_{[k]}}\lambda_{[k]}^x}{x!} \prod_{i=1}^{k-1} \left\{\sum_{t=0}^{x+d} \frac{e^{-\lambda_{[i]}}\lambda_{[i]}^t}{t!}\right\},$$

then the function $P(CS|\underline{\lambda}, d)$ *is* *(i) non-increasing in $\lambda_{[i]}$, keeping other components of $\underline{\lambda}$ fixed, for $i = 1, 2, \cdots, k-1$; and*

(ii) non-decreasing in $\lambda_{[k]}$, keeping other components of $\underline{\lambda}$ fixed.

Proof: *We have*

$$\frac{\partial}{\partial\lambda}\left(\sum_{i=0}^{x+d}\frac{e^{-\lambda}\lambda^i}{i!}\right) = -\frac{e^{-\lambda}\lambda^{x+d}}{(x+d)!} < 0,$$

therefore

$$P(CS|\underline{\lambda}, d) \geq P(CS|\lambda_{[k]}\underline{1}, d), \quad \textit{for all} \ \ \underline{\lambda} \in (0, \infty)^k.$$

Hence property (i) follows.

For property (ii), consider

$$a(x+d) = \prod_{i=1}^{k-1}\left\{\sum_{t=0}^{x+d}\frac{e^{-\lambda_{[i]}}\lambda_{[i]}^t}{t!}\right\},$$

which is increasing in $x+d$. Now

$$P(CS|\underline{\lambda}, d) = \sum_{x=0}^{\infty}\frac{e^{-\lambda_{[k]}}\lambda_{[k]}^x}{x!}a(x+d), \quad \textit{for fixed} \ \ \lambda_{[i]}, \ i = 1, 2, \cdots, k-1. \ \textit{So}$$

$$\frac{\partial P(CS|\underline{\lambda}, d)}{\partial\lambda_{[k]}} = \sum_{x=0}^{\infty}\frac{e^{-\lambda_{[k]}}\lambda_{[k]}^x}{x!}[a(x+1+d) - a(x+d)] \geq 0,$$

hence the result. □

Proposition 5.1 *: The conjugate priors for the problem is the gamma family, and for any gamma distribution $g_{\alpha,\beta} \in \Gamma_0, \ \alpha, \beta > 0$, i.e.*

$$g_{\alpha,\beta}(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda} \ \ (\lambda > 0),$$

we have

$$P(CS|d, g_{\alpha,\beta}, R^{\max}) \geq \sum_{x=0}^{\infty}\sum_{j_1=0}^{x+d}\cdots\sum_{j_{k-1}=0}^{x+d}\frac{\mathcal{H}_{\alpha,\beta}(x, k, h)}{x!j_1!\cdots j_{k-1}!},$$

where $h = j_1 + j_2 + \cdots + j_{k-1}$, and

$$\mathcal{H}_{\alpha,\beta}(x, h, k) = \frac{\Gamma(x+h+\alpha)\beta^\alpha}{\Gamma(\alpha)(\beta+k)^{x+h+\alpha}}.$$

Remark 5.1 *One disadvantage in this approach is that there does not exist any proper noninformative prior for Poisson distribution. The only possible choice we can think of is:*

$$g_{1/2,\beta}(\lambda) = \frac{\beta^{1/2}}{\Gamma(1/2)}\lambda^{-1/2}e^{-\beta\lambda} \quad (\lambda > 0), \tag{10}$$

where $\Pi(\lambda) \propto \lambda^{-1/2}$ *is Jeffreys noninformative prior. The value* β *has to be determined in advance. Note that*

$$g_{1/2,\beta}(\lambda) \propto \lambda^{-1/2} \quad \textit{as} \quad \beta \to 0.$$

Table 1: $P(n, k, d, G)$ Values for $G = U(G_J)$ and $n = 2(1)10$, $k = 2(1)5$, $0 \le d < n$.

n	d	k = 2		k = 3		k = 4		k = 5	
2	0	0.7667	(0.8203)	0.6571	(0.7355)	0.5929	(0.6839)	0.5504	(0.6486)
	1	0.9667	(0.9766)	0.9429	(0.9600)	0.9246	(0.9471)	0.9100	(0.9368)
3	0	0.7286	(0.7871)	0.6060	(0.6898)	0.5350	(0.6320)	0.4884	(0.5929)
	1	0.9357	(0.9541)	0.8929	(0.9236)	0.8613	(0.9010)	0.8367	(0.8832)
	2	0.9929	(0.9951)	0.9869	(0.9911)	0.9818	(0.9876)	0.9773	(0.9846)
4	0	0.7032	(0.7640)	0.5726	(0.6586)	0.4979	(0.5968)	0.4491	(0.5554)
	1	0.9095	(0.9346)	0.8522	(0.8931)	0.8111	(0.8632)	0.7798	(0.8401)
	2	0.9825	(0.9879)	0.9687	(0.9785)	0.9573	(0.9706)	0.9476	(0.9639)
	3	0.9984	(0.9989)	0.9970	(0.9980)	0.9957	(0.9971)	0.9945	(0.9963)
5	0	0.6847	(0.7465)	0.5488	(0.6357)	0.4716	(0.5708)	0.4214	(0.5278)
	1	0.8875	(0.9178)	0.8188	(0.8675)	0.7708	(0.8320)	0.7346	(0.8049)
	2	0.9711	(0.9799)	0.9493	(0.9647)	0.9317	(0.9525)	0.9170	(0.9422)
	3	0.9953	(0.9968)	0.9912	(0.9941)	0.9876	(0.9916)	0.9844	(0.9894)
	4	0.9996	(0.9998)	0.9993	(0.9995)	0.9990	(0.9993)	0.9987	(0.9991)
6	0	0.6705	(0.7327)	0.5307	(0.6172)	0.4518	(0.5505)	0.4007	(0.5064)
	1	0.8686	(0.9032)	0.7910	(0.8456)	0.7376	(0.8057)	0.6979	(0.7757)
	2	0.9596	(0.9717)	0.9301	(0.9510)	0.9069	(0.9346)	0.8878	(0.9211)
	3	0.9911	(0.9939)	0.9836	(0.9888)	0.9771	(0.9844)	0.9714	(0.9805)
	4	0.9988	(0.9991)	0.9976	(0.9984)	0.9965	(0.9977)	0.9956	(0.9970)
	5	0.9999	(0.9999)	0.9998	(0.9999)	0.9998	(0.9998)	0.9997	(0.9998)
7	0	0.6591	(0.7213)	0.5163	(0.6023)	0.4362	(0.5341)	0.3845	(0.4892)
	1	0.8524	(0.8904)	0.7674	(0.8268)	0.7098	(0.7832)	0.6674	(0.7508)
	2	0.9484	(0.9636)	0.9119	(0.9377)	0.8836	(0.9176)	0.8607	(0.9013)
	3	0.9861	(0.9905)	0.9748	(0.9827)	0.9652	(0.9761)	0.9568	(0.9704)
	4	0.9973	(0.9982)	0.9949	(0.9966)	0.9927	(0.9951)	0.9907	(0.9937)
	5	0.9997	(0.9998)	0.9994	(0.9996)	0.9991	(0.9994)	0.9988	(0.9992)
	6	1.0000	(1.0000)	1.0000	(1.0000)	0.9999	(1.0000)	0.9999	(1.0000)

Table 1 continued.

n	d	2		3		4		5	
8	0	0.6498	(0.7118)	0.5046	(0.5899)	0.4236	(0.5204)	0.3715	(0.4748)
	1	0.8382	(0.8791)	0.7470	(0.8102)	0.6861	(0.7637)	0.6416	(0.7293)
	2	0.9377	(0.9558)	0.8947	(0.9251)	0.8620	(0.9017)	0.8358	(0.8828)
	3	0.9807	(0.9867)	0.9654	(0.9761)	0.9525	(0.9673)	0.9415	(0.9596)
	4	0.9954	(0.9969)	0.9919	(0.9942)	0.9878	(0.9917)	0.9845	(0.9895)
	5	0.9992	(0.9995)	0.9985	(0.9990)	0.9978	(0.9985)	0.9971	(0.9981)
	6	0.9999	(0.9999)	0.9998	(0.9999)	0.9998	(0.9998)	0.9997	(0.9998)
	7	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)
9	0	0.6419	(0.7035)	0.4948	(0.5793)	0.4131	(0.5087)	0.3606	(0.4626)
	1	0.8257	(0.8689)	0.7293	(0.7956)	0.6656	(0.7465)	0.6194	(0.7105)
	2	0.9276	(0.9483)	0.8787	(0.9132)	0.8421	(0.8868)	0.8131	(0.8657)
	3	0.9750	(0.9827)	0.9556	(0.9692)	0.9396	(0.9582)	0.9260	(0.9487)
	4	0.9931	(0.9953)	0.9871	(0.9913)	0.9819	(0.9877)	0.9771	(0.9845)
	5	0.9985	(0.9990)	0.9972	(0.9981)	0.9959	(0.9973)	0.9947	(0.9965)
	6	0.9998	(0.9999)	0.9996	(0.9997)	0.9994	(0.9996)	0.9991	(0.9994)
	7	1.0000	(1.0000)	1.0000	(1.0000)	0.9999	(0.9999)	0.9999	(0.9999)
	8	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)
10	0	0.6351	(0.6964)	0.4865	(0.5701)	0.4042	(0.4989)	0.3515	(0.4522)
	1	0.8145	(0.8597)	0.7137	(0.7825)	0.6476	(0.7313)	0.6000	(0.6939)
	2	0.9181	(0.9412)	0.8639	(0.9020)	0.8238	(0.8729)	0.7923	(0.8498)
	3	0.9692	(0.9786)	0.9458	(0.9623)	0.9268	(0.9490)	0.9108	(0.9378)
	4	0.9904	(0.9935)	0.9823	(0.9880)	0.9753	(0.9832)	0.9690	(0.9789)
	5	0.9976	(0.9984)	0.9954	(0.9969)	0.9934	(0.9956)	0.9916	(0.9943)
	6	0.9995	(0.9997)	0.9991	(0.9994)	0.9987	(0.9991)	0.9983	(0.9988)
	7	0.9999	(0.9999)	0.9999	(0.9999)	0.9998	(0.9999)	0.9998	(0.9998)
	8	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)	1.0000	(1.0000)

Table 2: Values of $P(\cdot)$, for various procedures, for $n = 5,\ k = 3$.

	d					
Procedures	0	1	2	3	4	5
$\inf_{0<p<1} P(CS\|p, d)$	0.4673	0.7265	0.9040	0.9797	0.9981	1.0000
$\inf_{.75<p<1} P(CS\|p, d)$	0.5039	0.8045	0.9528	0.9937	0.9997	1.0000
$\inf_{\varepsilon>0} P(n, k, Q^{\varepsilon}_{\frac{1}{2}})$	0.4676	0.7269	0.9041	0.9798	0.9981	1.0000
$\inf_{\varepsilon>0} P(n, k, Q^{\varepsilon}_{\frac{1}{4}})$	0.4844	0.7718	0.9378	0.9888	0.9991	1.0000
$\inf_{\varepsilon>0} P(n, k, Q^{\varepsilon}_{\frac{3}{4}})$	0.5039	0.7975	0.9396	0.9890	0.9991	1.0000
$\inf_{\varepsilon>0} P(n, k, Q^{\varepsilon}_{0.0})$	0.5143	0.7901	0.9378	0.9888	0.9991	1.0000
$\inf_{\varepsilon>0} P(n, k, Q^{\varepsilon}_{1.0})$	0.5300	0.7975	0.9396	0.9890	0.9991	1.0000
$P(n, k, d, U)$	0.5488	0.8188	0.9493	0.9912	0.9993	1.0000
$P(n, k, d, G_J)$	0.6357	0.8675	0.9647	0.9941	0.9995	1.0000

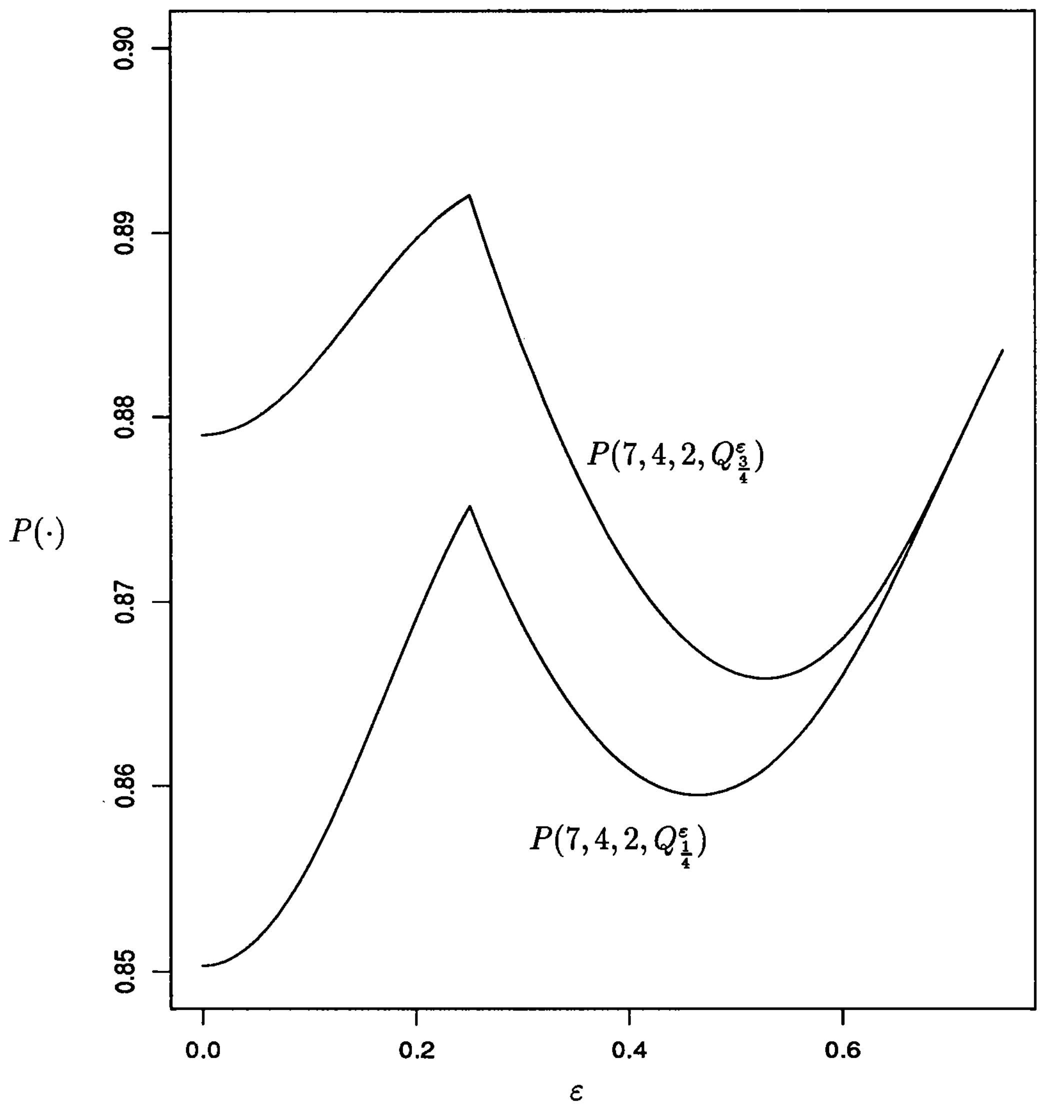

Figure 1: $P(7,4,2,Q^{\varepsilon}_{a_0})$, for $a_0 = \frac{1}{4}(\frac{3}{4})$.

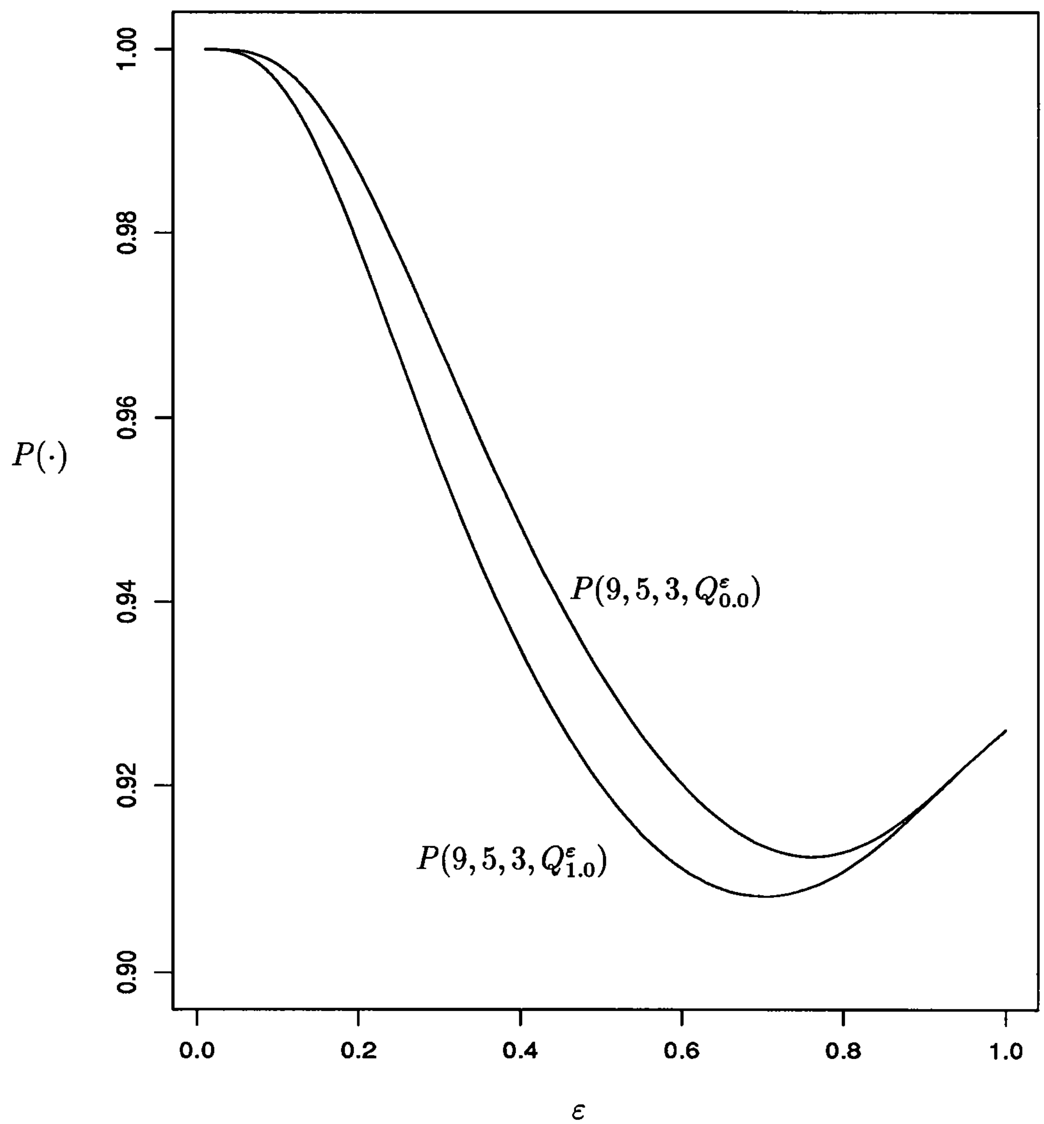

Figure 2: $P(9,5,3,Q^{\varepsilon}_{a_0})$, for $a_0 = 0.0(1.0)$.

References

[1] Bechhofer, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, **25**, 16-39.

[2] Goel, P. K. (1972). A Note on the Nonexistence of Subset Selection Procedures for Poisson Populations. Mimeograph series No. 303, Department of Statistics, Purdue University, West Lafayette, Indiana.

[3] Goel, P. K. (1975). A note on Subset Selection Procedure with Poisson Processes. Unpublished Report.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeograph Series No. 150), Institute of Statistics, University of North Carolina, Chapl Hill, North Carolina.

[5] Gupta, S. S. and Huang, D. -Y. (1975). On Subset Selection Procedures for Poisson Populations and Some Applications to the Multinomial Selection Problems. in: *Applied Probability*, (R. P. Gupta ed.), North-Holland, Amsterdam, 97-109.

[6] Gupta, S. S., Leong, Y. K. and Wong, W. -Y. (1979). On Subset Selection Procedures for Poisson Populations. *Bulletin of the Malaysian Mathematical Society*, (2) **2**, No. 2, 89-110.

[7] Gupta, S. S. and Mcdonald, G. C. (1986). A Statistical Selection Approach to Binomial Models. *Journal of Quality Technology*, **18**(No. 2), 103-115.

[8] Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, John Wiley & Sons, inc., New York.

[9] Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. *Annals of Mathematical Statistics*, **29**, 235-244.

[10] Gupta, S. S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. in: *Contributions to Probability and Statistics*, (I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow and H. B. Mann ed.), Stanford University Press, Stanford, California, **Chapter 20**, 224-248.

[11] Gupta, S. S. and Wong, W. -Y. (1977). On subset selection procedures for Poisson Processes and some applications to the binomial and multinomial problems. in: *Operations Research Verfahren*, (R. Henn, P. Kall, B. Korte, O. Krafft, W. Oettli, K. Ritter, J. Rosenmüller, N. Schmitz, H. Schubert and W. Vogel eds.), Verlag Anton Hain, Meisenheim am Glan, West Germany, 49-70.

[12] Leong, Y. K. and Wong, W. -Y. (1977). On Poisson selection problems. Research Report 23/77, Department of Mathematics, University of Malaya, K. Lumpur, Malaysia.