*

SELECTING THE MOST RELIABLE POISSON
POPULATION PROVIDED IT IS BETTER THAN A CONTROL:
A NONPARAMETRIC EMPIRICAL BAYES APPROACH

by

Shanti S. Gupta   and   TaChen Liang
Purdue University   Wayne State University

Technical Report #97-9C

*

SELECTING THE MOST RELIABLE POISSON POPULATION PROVIDED IT IS BETTER THAN A CONTROL: A NONPARAMETRIC EMPIRICAL BAYES APPROACH

by

Shanti S. Gupta   and   TaChen Liang

Purdue University   Wayne State University

Abstract

We study the problem of selecting the most reliable Poisson population from among k competitors provided it is better than a control using the nonparametric empirical Bayes approach. An empirical Bayes selection procedure is constructed based on the isotonic regression estimators of the posterior means of failure rates associated with the k Poisson populations. The asymptotic optimality of the empirical Bayes selection procedure is investigated. Under certain regularity conditions, we have shown that the proposed empirical Bayes selection procedure is asymptotically optimal and the associated Bayes risk converges to the minimum Bayes risk at a rate of order $O(\exp(-cn))$ for some $c > 0$, where n denotes the number of historical data at hand when the present selection problem is considered.

1

1. Introduction:

In the research and development stage, an experimenter often confronts the problem of selecting the most reliable design from among several competing designs. Usually, the populations that are compared are the life length distributions of the designs. The most reliable design is defined as the one with the longest mean life. The problem of selecting the most reliable design has been studied in the literature. The readers are referred to Gupta and Panchapakesan (1988) for a comprehensive survey of selection procedures in reliability models.

Consider k types of competing designs $\pi_1, \ldots, \pi_k$, which are put on life tests. Suppose in case of a failure, the failed design is immediately replaced by the same type of design. It is assumed that the failure times of type i design are exponentially distributed with an unknown failure rate θ_i, and these failure times are mutually independently distributed. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered values of the failure rate $\theta_1, \ldots, \theta_k$. The design associated with the smallest failure rate $\theta_{[1]}$ is called the most reliable design. Over a fixed common time interval $[0, T]$, we let X_i denote the number of failures of type i design. Then, X_i follows a Poisson distribution with occurrence rate θ_i. With this sampling scheme, Dixon and Bland (1971) derived a Bayes solution to the problem of ranking the failure rates $\theta_1, \ldots, \theta_k$. Gupta, Leong and Wong (1979) have developed a subset selection procedure for selecting a subset containing the most reliable design. Alam (1971) and Alam and Thompson (1973) have also studied selection procedures for selecting the most reliable design based on inverse sampling observations.

Let θ_0 be a specified control level. Design π_i is said to be better than the control θ_0 and acceptable if $\theta_i < \theta_0$; otherwise, π_i is said to be bad and should be excluded. In many practical situations, an experimenter makes a selection only when the most reliable design is better that the control θ_0. For example, let θ_0 be the failure rate of the currently used design. An experimenter may make a selection from among the k competitors and replace the currently used design by the newly selected one provided the newly selected design is better than the level θ_0. Otherwise, the experimenter may select none and continue using the current design.

Consider a situation in which one will be dealing with repeated independent selection

problems. In such instances, it is reasonable to formulate the component problem in the sequence as a Bayes decision problem with respect to an unknown prior distribution on the parameter space. One then uses the accumulated historical data to improve the decision procedures at each stage. This is the so-called empirical Bayes approach (see Robbins (1956, 1964)). Empirical Bayes procedures have been derived for subset selection goals by Deely (1965). For recent developments of empirical Bayes selection procedures on the research area of ranking and selection, the readers are referred to Gupta, Liang and Rau (1994, 1995), and the references quoted there.

In this paper, we study the problem of selecting the most reliable Poisson population provided it is better than a control using the nonparametric empirical Bayes approach. The paper is organized as follows. The selection problem is formulated in Section 2. For the loss function (2.1), we derive a Bayes selection procedure, which is based on the posterior means of the failure rates $\theta_i, i = 1, \ldots, k$. The empirical Bayes framework of the selection problem is described in Section 3. Based on the isotonic regression estimators of the posterior means of the failure rates, an empirical Bayes selection procedure is constructed. We study the asymptotic optimality of this procedure in Section 4. It is shown that under certain regularity conditions, the proposed empirical Bayes selection procedure is asymptotically optimal and its Bayes risk converges to the minimum Bayes risk at a rate of order $O(\exp(-cn))$ for some positive number c, where n denotes the number of the accumulated historical data at hand when the present selection problem is considered.

2. Formulation of the Selection Problem and a Bayes Selection Procedure

Consider $k(\geq 2)$ independent Poisson populations $\pi_1, \ldots, \pi_k$, with unknown occurrence rates $\theta_1, \ldots, \theta_k$, respectively. Let $\theta_{[1]} \leq \ldots \ldots \leq \theta_{[k]}$ denote the ordered values of the parameters $\theta_1, \ldots, \theta_k$. It is assumed that the exact pairing between the ordered and the unordered parameters is unknown. The occurrence rate θ_i may be viewed as the failure rate of type i design, as described previously. Therefore, in the following, a population π_i with $\theta_i = \theta_{[1]}$ is considered as the most reliable design. Let θ_0 be a known pre-specified control level. A population π_i with $\theta_i < \theta_0$ is said to be better than the control θ_0 and is acceptable. Otherwise, population π_i is said to be bad and not acceptable. Our goal is to derive empirical Bayes procedures to select the most reliable Poisson population which

should also be better than the control θ_0. If there is no such population, we select none and exclude all k competitors as bad.

Let $\Omega = \{\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k) |\ \theta_i > 0,\ i = 1, \ldots, k\}$ be the parameter space and let $\mathcal{A} = \{\underset{\sim}{a} = (a_0, a_1, \ldots, a_k) |\ a_i = 0, 1; i = 0, 1, \ldots, k,$ and $\sum_{i=0}^{k} a_i = 1\}$ be the action space. For an action $\underset{\sim}{a}$, when $a_i = 1$ for some $i = 1, \ldots, k$, it means that population π_i is selected as the most reliable design and considered to be better than the control θ_0; when $a_0 = 1$, it means that all the k populations are excluded as bad compared with the control θ_0 and no selection is made. We consider the following loss function: For $\underset{\sim}{\theta} \in \Omega$ and $\underset{\sim}{a} \in \mathcal{A}$,

$$L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=0}^{k} a_i \theta_i - min(\theta_{[1]}, \theta_0). \tag{2.1}$$

For each $i = 1, \ldots, k$, let X_i denote a random observation arising from a Poisson population π_i with occurrence rate θ_i. Without loss of generality, we assume that the fixed common time interval is $[0, 1]$. That is, $T = 1$. Thus, conditioning on θ_i, X_i has a probability function $f_i(x|\theta_i)$, where

$$f_i(x|\theta_i) = \exp(-\theta_i)\theta_i^x / x!,\ x = 0, 1, 2, \ldots \tag{2.2}$$

It is assumed that θ_i is a realization of a random variable Θ_i which has an unknown prior distribution G_i over $(0, \infty)$. The random vectors $(X_i, \Theta_i), i = 1, \ldots, k$, are assumed to be mutually independent.

Let $\mathcal{X}$ be the sample space generated by $\underset{\sim}{X} = (X_1, \ldots, X_k)$. A selection procedure $\underset{\sim}{\delta} = (\delta_0, \delta_1, \ldots, \delta_k)$ is defined to be a mapping from the sample space $\mathcal{X}$ into the product space $[0, 1]^{k+1}$, such that for each $\underset{\sim}{x} \epsilon \mathcal{X}$, the function $\underset{\sim}{\delta}(\underset{\sim}{x}) = (\delta_0(\underset{\sim}{x}), \delta_1(\underset{\sim}{x}), \ldots, \delta_k(\underset{\sim}{x}))$ satisfies that $0 \leq \delta_i(\underset{\sim}{x}) \leq 1, i = 0, 1, \ldots, k$ and $\sum_{i=0}^{k} \delta_i(\underset{\sim}{x}) = 1$. That is, for each $i = 1, \ldots, k, \delta_i(\underset{\sim}{x})$ is the probability of both selecting π_i as the most reliable population and also as being better than the control θ_0; and $\delta_0(\underset{\sim}{x})$ is the probability of excluding all the k populations as bad and hence selecting none.

Let $\mathcal{C}$ be the class of all selection procedures. For each $\underset{\sim}{\delta} \in \mathcal{C}$, let $R(G, \underset{\sim}{\delta})$ denote its associated Bayes risk, where $G(\underset{\sim}{\theta}) = \prod_{i=1}^{k} G_i(\theta_i)$. Then $R(G) = \inf_{\underset{\sim}{\delta} \epsilon C} R(G, \underset{\sim}{\delta})$ is the minimum Bayes Risk among the class $\mathcal{C}$. A selection procedure, say $\underset{\sim}{\delta}_G$, such that $R(G, \underset{\sim}{\delta}_G) = R(G)$ is

4

called a Bayes selection procedure. We consider only these priors for which $\int_0^\infty \theta dG_i(\theta) < \infty$ for each $i = 1, \ldots, k$ so that for each selection procedure $\underset{\sim}{\delta}, R(G, \underset{\sim}{\delta})$ is always finite, which insures the selection problem to be meaningful.

Based on the preceding statistical model, the Bayes risk associated with the selection procedure $\underset{\sim}{\delta}$ is:

$$
\begin{aligned}
R(G, \underset{\sim}{\delta}) &= \sum_{\underset{\sim}{x} \epsilon \mathcal{X}} \sum_{i=0}^{k} \delta_i(\underset{\sim}{x}) \int_\Omega \theta_i f(\underset{\sim}{x}|\underset{\sim}{\theta}) dG(\underset{\sim}{\theta}) - C \\
&= \sum_{\underset{\sim}{x} \epsilon \mathcal{X}} \left[\sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \psi_i(x_i) + \delta_0(\underset{\sim}{x}) \theta_0 \right] f(\underset{\sim}{x}) - C
\end{aligned}
\tag{2.3}
$$

where $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = \prod_{i=1}^{k} f_i(x_i|\theta_i), f(\underset{\sim}{x}) = \prod_{i=1}^{k} f_i(x_i), C = \int_\Omega min(\theta_{[1]}, \theta_0) dG(\underset{\sim}{\theta})$ and $f_i(x_i) = \int f_i(x_i|\theta) dG_i(\theta) = h_i(x_i)(x_i!)^{-1}$ is the marginal probability function of $X_i, h_i(x_i) = \int \exp(-\theta)\theta^{x_i} dG_i(\theta)$,

$$
\psi_i(x_i) = E\left[\Theta_i | X_i = x_i\right] = h_i(x_i + 1)/h_i(x_i).
\tag{2.4}
$$

Note that $\psi_i(x_i)$ is the posterior mean of Θ_i given $X_i = x_i$.

From (2.3), a Bayes selection procedure $\underset{\sim}{\delta}_G = (\delta_{G0}, \ldots, \delta_{Gk})$ can be obtained as follows: For each $\underset{\sim}{x} \ \epsilon \ \mathcal{X}$, let

$$
A(\underset{\sim}{x}) = \{i = 0, 1, \ldots, k | \psi_i(x_i) = \min_{0 \leq j \leq k} \psi_j(x_j)\},
\tag{2.5}
$$

where $\psi_0(x_0) \equiv \theta_0$. Define

$$
i_G \equiv i_G(\underset{\sim}{x}) = \min\{i | i \epsilon A(\underset{\sim}{x})\},
\tag{2.6}
$$

and for each $i = 0, \ldots, k$, define

$$
\delta_{Gi}(\underset{\sim}{x}) = \begin{cases} 1, & \text{if } i = i_G, \\ 0 & \text{otherwise.} \end{cases}
\tag{2.7}
$$

From (2.3), one can see that $\underset{\sim}{\delta}_G$ is a Bayes selection procedure. Also, it should be noted that any selection procedure $\underset{\sim}{\delta}$ satisfying $\sum_{i \epsilon A(\underset{\sim}{x})} \delta_i(\underset{\sim}{x}) = 1$ for each $\underset{\sim}{x} \ \epsilon \ \mathcal{X}$ is always

5

a Bayes selection procedure. The (minimum) Bayes risk is $R(G, \underset{\sim}{\delta}_G)$ since this risk is the minimum, by construction, where

$$
\begin{aligned}
R(G, \underset{\sim}{\delta}_G) &= \sum_{\underset{\sim}{x} \in \mathcal{X}} \left[\sum_{i=0}^{k} \delta_{Gi}(\underset{\sim}{x}) \psi_i(x_i) \right] f(\underset{\sim}{x}) - C \\
&= \sum_{\underset{\sim}{x} \in \mathcal{X}} \psi_{i_G}(x_{i_G}) f(\underset{\sim}{x}) - C.
\end{aligned}
\tag{2.8}
$$

3. Empirical Bayes Selection Procedures

It should be noted that the Bayes selection procedure $\underset{\sim}{\delta}_G$ depends on the prior distribution G. Since G is unknown, it is not possible to implement the Bayes procedure $\underset{\sim}{\delta}_G$ for the selection problem at hand. In the following, it is assumed that certain historical data from each of the k populations are available. In such a situation, the empirical Bayes approach is adopted.

3.1. Empirical Bayes Framework

For each $i = 1, \ldots, k$, let $(X_{ij}, \Theta_{ij}), j = 1, 2, \ldots$ be random vectors associated with population π_i, where X_{ij} is observable while Θ_{ij} is unobservable. Note that Θ_{ij} stands for the failure rate of the design belonging to π_i at stage j. It is assumed that Θ_{ij} has a prior distribution G_i, for all $j = 1, 2, \ldots$, and conditioning on $\Theta_{ij} = \theta_{ij}, X_{ij}$ follows a Poisson distribution with occurrence rate θ_{ij}; and $(X_{ij}, \Theta_{ij}), i = 1, \ldots, k, j = 1, 2, \ldots$ are mutually independent. At the present stage, say stage $n + 1$, let $\underset{\sim}{X}_i(n) = (X_{i1}, \ldots, X_{in})$ denote the accumulated historical data associated with π_i, and let $X_i = X_{i,n+1}$ be the present random observation arising from π_i, and $\theta_i = \theta_{i,n+1}$ be a realization of $\Theta_{i,n+1}, i = 1, \ldots, k$. Let $\underset{\sim}{X}(n) = (\underset{\sim}{X}_1(n), \ldots, \underset{\sim}{X}_k(n))$ and $\underset{\sim}{X} = (X_1, \ldots, X_k)$. At stage $n + 1$, we want to select the population associated with $\theta_{[1]}$ provided that $\theta_{[1]} < \theta_0$ using the loss function (2.1).

By (2.5)-(2.7), a natural empirical Bayes selection procedure can be derived as follows. For each $i = 1, \ldots, k$, based on the accumulated past data $\underset{\sim}{X}_i(n)$ and the present observation $X_i = x_i$, let $\psi_{in}(x_i) \equiv \psi_{in}(x_i, \underset{\sim}{X}_i(n))$ be an empirical Bayes estimator of $\psi_i(x_i)$. Then, let

$$
A_n(\underset{\sim}{x}) = \{i = 0, 1, \ldots, k | \psi_{in}(x_i) = \min_{0 \le j \le k} \psi_{jn}(x_j)\}
\tag{3.1}
$$

where $\psi_{0n}(x_0) \equiv \theta_0$. Define

$$i_n \equiv i_n(\underset{\sim}{x}) = \min\{i | i \ \epsilon \ A_n(\underset{\sim}{x})\}. \tag{3.2}$$

Analogous to (2.7), an empirical Bayes selection procedure $\underset{\sim}{\delta}_n = (\delta_{n0}, \dots, \delta_{nk})$ can be obtained as follows: For each $\underset{\sim}{x} \ \epsilon \ \mathcal{X}$,

$$\delta_{ni}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } i = i_n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Let $R(G, \underset{\sim}{\delta}_n | \underset{\sim}{X}(n))$ denote the conditional Bayes risk of the selection procedure $\underset{\sim}{\delta}_n$ given $\underset{\sim}{X}(n)$ and let $R(G, \underset{\sim}{\delta}_n)$ be the overall Bayes risk of $\underset{\sim}{\delta}_n$. That is,

$$R(G, \underset{\sim}{\delta}_n | X(n)) = \sum_{\underset{\sim}{x} \epsilon \mathcal{X}} \psi_{i_n}(x_{i_n}) f(\underset{\sim}{x}) - C, \tag{3.4}$$

and

$$R(G, \underset{\sim}{\delta}_n) = E_{\underset{\sim}{X}(n)} R(G, \underset{\sim}{\delta}_n | \underset{\sim}{X}(n)), \tag{3.5}$$

where the expectation $E_{\underset{\sim}{X}(n)}$ is taken with respect to the probability measure generated by $\underset{\sim}{X}(n)$.

Note that $R(G, \underset{\sim}{\delta}_n | X(n)) - R(G, \underset{\sim}{\delta}_G) \geq 0$ for all $\underset{\sim}{X}(n)$ and for all n, since $\underset{\sim}{\delta}_G$ is a Bayes selection procedure. Hence $R(G, \underset{\sim}{\delta}_n) - R(G, \underset{\sim}{\delta}_G) \geq 0$ for all n. The nonnegative regret Bayes risk $R(G, \underset{\sim}{\delta}_n) - R(G, \underset{\sim}{\delta}_G)$ can be used as a measure of performance of the empirical Bayes selection procedure $\underset{\sim}{\delta}_n$.

A sequence of empirical Bayes selection procedures $\{\underset{\sim}{\delta}_n\}_{n=1}^{\infty}$ is said to be asymptotically optimal relative to the prior distribution G if $R(G, \underset{\sim}{\delta}_n) - R(G, \underset{\sim}{\delta}_G) \to 0$ as $n \to \infty$. Further, $\{\underset{\sim}{\delta}_n\}_{n=1}^{\infty}$ is said to be asymptotically optimal of order $\{\alpha_n\}_{n=1}^{\infty}$ relative to the prior distribution G if $R(G, \underset{\sim}{\delta}_n) - R(G, \underset{\sim}{\delta}_G) = O\left(\alpha_n\right)$ where $\{\alpha_n\}_{n=1}^{\infty}$ is a sequence of positive numbers such that $\lim_{n \to \infty} \alpha_n = 0$.

In the following, we seek a sequence of empirical Bayes selection procedures possessing the asymptotic optimality.

3.2 Construction of Empirical Bayes Selection Procedure

To construct an empirical Bayes selection procedure as described in (3.1)-(3.3), we first need to construct an empirical Bayes estimator for $\psi_i(x_i)$. Since $\psi_i(x_i)$ is increasing in x_i we desire a monotone estimator for $\psi_i(x_i)$.

Based on $X_i(n)$, for each $x = 0, 1, \ldots$, define

$$
\begin{cases}
f_{in}(x) = \frac{1}{n} \sum\limits_{j=1}^{n} I_{\{x\}}(X_{ij}), \\
h_{in}(x) = f_{in}(x)/a(x).
\end{cases}
\tag{3.6}
$$

where $a(x) = 1/(x!)$. Note that $f_{in}(x)$ and $h_{in}(x)$ are unbiased and consistent estimators of $f_i(x)$ and $h_i(x)$, respectively. From (2.4), $h_{in}(x_i + 1)/h_{in}(x_i)$ is a naive estimator of the posterior mean $\psi_i(x_i)$. However, this estimator may not possess the monotonicity property. Thus, we consider an isotonic regression version of $h_{in}(x_i + 1)/h_{in}(x_i)$.

Let $N_{i1} = \min(X_{i1}, \ldots, X_{in})$ and $N_{in} = \max(X_{i1}, \ldots, X_{in})$. For each $y = 0, 1, \ldots$, define

$$
\begin{cases}
\Psi_{in}(y) = \sum\limits_{x=0}^{y} h_{in}(x+1)a(x+1), \ \Psi_i(y) = \sum\limits_{x=0}^{y} h_i(x+1)a(x+1), \\
H_{in}(y) = \sum\limits_{x=0}^{y} h_{in}(x)a(x+1), \ H_i(y) = \sum\limits_{x=0}^{y} h_i(x)a(x+1),
\end{cases}
\tag{3.7}
$$

and $\Psi_{in}(-1) = \Psi_i(-1) = H_{in}(-1) = H_i(-1) = 0$.

Next, define

$$
\psi_{in}^*(N_{i1}) = \min_{N_{i1} \le y \le N_{in}} \left\{ \frac{\Psi_{in}(y) - \Psi_{in}(N_{i1} - 1)}{H_{in}(y) - H_{in}(N_{i1} - 1)} \right\},
\tag{3.8}
$$

where $b/a = \infty$ when $a = 0$; and for each x between (including) $N_{i1}+1$ and N_{in}, recursively, define

$$
\psi_{in}^*(x) = \min_{x \le y \le N_{in}} \left\{ \frac{\Psi_{in}(y) - \sum\limits_{z=N_{i1}}^{x-1} \psi_{in}^*(z)h_{in}(z)a(z+1)}{H_{in}(y) - H_{in}(x-1),} \right\}.
\tag{3.9}
$$

Note that since $H_{in}(N_{in}) - H_{in}(x - 1) > 0$ for $N_{i1} \le x \le N_{in}, \psi_{in}^*(x) < \infty$. $\{\psi_{in}^*(x)\}_{x=N_{i1}}^{N_{in}}$ is the isotonic regression of $\{h_{in}(x+1)/h_{in}(x)\}_{x=N_{i1}}^{N_{in}}$ with random weights $\{h_{in}(x)a(x+1)\}_{x=N_{i1}}^{N_{in}}$, see Puri and Singh (1990). Hence, $\psi_{in}^*(x)$ is nondecreasing in x for $x = N_{i1}, \ldots, N_{in}$. By Barlow, et. al. (1972), for $N_{i1} < x \le N_{in}$,

$$
\begin{aligned}
\sum_{z=N_{i1}}^{x-1} \psi_{in}^*(z)h_{in}(z)a(z+1) &\le \sum_{z=N_{i1}}^{x-1} \frac{h_{in}(z+1)}{h_{in}(z)} h_{in}(z)a(z+1) \\
&= \sum_{z=N_{i1}}^{x-1} h_{in}(z+1)a(z+1) \\
&\le \Psi_{in}(x-1).
\end{aligned}
\tag{3.10}
$$

8

Therefore, from (3.8) - (3.10), we have, for each $x = N_{i1}, \ldots, N_{in}$,

$$\psi_{in}^*(x) \geq \min_{x \leq y \leq N_{in}} \left\{ \frac{\Psi_{in}(y) - \Psi_{in}(x-1)}{H_{in}(y) - H_{in}(x-1)} \right\}. \tag{3.11}$$

For $x = 0, 1, \ldots, N_{i1} - 1$, define $\psi_{in}^*(x) = \psi_{in}^*(N_{i1})$; for $x > N_{in}$, define $\psi_{in}^*(x) = \psi_{in}^*(N_{in})$. Therefore, we see that $\psi_{in}^*(x)$ is a monotone function of the nonnegative integers x.

Now, we let $\underset{\sim}{\delta}_n^* = (\delta_{n0}^*, \ldots, \delta_{nk}^*)$ be the empirical Bayes selection procedure defined through (3.1)-(3.3) by replacing $\psi_{in}(x_i)$ by $\psi_{in}^*(x_i)$. We also denote its associated $A_n(\underset{\sim}{x})$ and i_n by $A_n^*(\underset{\sim}{x})$ and i_n^*, respectively. Then, conditioning on $\underset{\sim}{X}(n)$, the conditional Bayes risk of $\underset{\sim}{\delta}_n^*$ is:

$$R(G, \underset{\sim}{\delta}_n^* | \underset{\sim}{X}(n)) = \sum_{\underset{\sim}{x} \epsilon \mathcal{X}} \psi_{i_n^*}(x_{i_n^*}) f(\underset{\sim}{x}) - C, \tag{3.12}$$

and its overall Bayes risk is:

$$R(G, \underset{\sim}{\delta}_n^*) = E_{\underset{\sim}{X}(n)} R(G, \underset{\sim}{\delta}_n^* | \underset{\sim}{X}(n)). \tag{3.13}$$

4. Asymptotic Optimality of $\underset{\sim}{\delta}_n^*$

In this section, we evaluate the asymptotic optimality of the empirical Bayes selection procedure $\underset{\sim}{\delta}_n^*$. We assume the prior distribution G_i being nondegenerate so that $\psi_i(x_i)$ is strictly increasing in x_i. Let $B_i(1) = \{x_i | \psi_i(x_i) < \theta_0\}, B_i(2) = \{x_i | \psi_i(x_i) = \theta_0\}$ and $B_i(3) = \{x_i | \psi_i(x_i) > \theta_0\}$. Define

$$
\begin{aligned}
m_i &= \begin{cases} \sup B_i(1) & \text{if } B_i(1) \neq \phi, \\ -1 & \text{otherwise;} \end{cases} \\
M_i &= \begin{cases} \inf B_i(3) & \text{if } B_i(3) \neq \phi, \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}
\tag{4.1}
$$

By the increasing property of $\psi_i(x), m_i \leq M_i$. When $B_i(3) \neq \phi, m_i < M_i < \infty$. We may have either $M_i = m_i + 1$ for which $B_i(2) = \phi$, or $M_i = m_i + 2$ for which $\psi_i(m_i + 1) = \theta_0$ and $B_i(2) = \{m_i + 1\}$.

Let E_i denote the event that $N_{i1} = 0$ and $N_{in} \geq M_i + 2, i = 1, \ldots, k$, and $E = \bigcap_{i=1}^{k} E_i$; let E_i^c and E^c denote the complements of the events E_i and E, respectively. Also, let $A_i = \{\underset{\sim}{x} \epsilon \mathcal{X} | i_G(\underset{\sim}{x}) = i\}, i = 0, 1, \ldots, k$.

9

4.1 Analysis of Regret Bayes Risk

From (2.8) and (3.12), given $\underset{\sim}{X}(n)$, the conditional regret Bayes risk of $\underset{\sim}{\delta}_n^*$ is:

$$R(G, \underset{\sim}{\delta}_n^* | \underset{\sim}{X}(n)) - R(G, \underset{\sim}{\delta}_G) = \sum_{\underset{\sim}{x}\epsilon\mathcal{X}} [\psi_{i_n^*}(x_{i_n^*}) - \psi_{i_G}(x_{i_G})]I(E)f(\underset{\sim}{x})$$
$$+ \sum_{\underset{\sim}{x}\epsilon\mathcal{X}} [\psi_{i_n^*}(x_{i_n^*}) - \psi_{i_G}(x_{i_G})]I(E^c)f(\underset{\sim}{x})$$

(4.2)

where $I(S)$ denotes the indicator function of the event S. Now,

$$\sum_{\underset{\sim}{x}\epsilon\mathcal{X}} [\psi_{i_n^*}(x_{i_n^*}) - \psi_{i_G}(x_{i_G})]I(E)f(\underset{\sim}{x})$$

$$= \sum_{\underset{\sim}{x}\epsilon\mathcal{X}} \sum_{i=0}^{k} \sum_{j=0}^{k} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j \text{ and } E\}[\psi_i(x_i) - \psi_j(x_j)]f(\underset{\sim}{x})$$

$$= \sum_{\underset{\sim}{x}\epsilon A_0} \sum_{i=1}^{k} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = 0, E\}[\psi_i(x_i) - \theta_0]f(\underset{\sim}{x})$$

(4.3)

$$+ \sum_{j=1}^{k} \sum_{\underset{\sim}{x}\epsilon A_j} I\{i_n^*(\underset{\sim}{x}) = 0, i_G(\underset{\sim}{x}) = j, E\}[\theta_0 - \psi_j(x_j)]f(\underset{\sim}{x})$$

$$+ \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underset{\sim}{x}\epsilon A_j} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, E\}[\psi_i(x_i) - \psi_j(x_j)]f(\underset{\sim}{x}).$$

On A_0, for $i \ \epsilon \ A(\underset{\sim}{x}), \psi_i(x_i) = \theta_0$. So,

$$I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = 0, E\}[\psi_i(x_i) - \theta_0]$$
$$= I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = 0, i\notin A(\underset{\sim}{x}), E\}[\psi_i(x_i) - \theta_0].$$

(4.4)

For each $j = 1, \ldots, k$, on A_j, for $i \ \epsilon \ A(\underset{\sim}{x}), \psi_i(x_i) = \psi_j(x_j)$. So,

$$I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, E\}[\psi_i(x_i) - \psi_j(x_j)]$$
$$= I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, i\notin A(\underset{\sim}{x}), E\}[\psi_i(x_i) - \psi_j(x_j)].$$

(4.5)

Combining (4.2)–(4.6), the regret Bayes risk of δ_n^* can be written as:

$$R(G, \delta_n^*) - R(G, \delta_G)$$

$$= \sum_{\underset{\sim}{x} \in A_0} \sum_{i=1}^{k} E_{\underset{\sim}{X}(n)} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = 0, i \notin A(\underset{\sim}{x}), E\}[\psi_i(x_i) - \theta_0] f(\underset{\sim}{x})$$

$$+ \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in A_j} E_{\underset{\sim}{X}(n)} I\{i_n^*(\underset{\sim}{x}) = 0, i_G(\underset{\sim}{x}) = j, 0 \notin A(\underset{\sim}{x}), E\}[\theta_0 - \psi_j(x_j)] f(\underset{\sim}{x})$$

$$\hspace{8cm} (4.6)$$

$$+ \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in A_j} E_{\underset{\sim}{X}(n)} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, i \notin A(\underset{\sim}{x}), E\}[\psi_i(x_i) - \psi_j(x_j)] f(\underset{\sim}{x})$$

$$+ E_{\underset{\sim}{X}(n)}[\sum_{\underset{\sim}{x} \in \mathcal{X}} [\psi_{i_n^*}(x_{i_n^*}) - \psi_{i_G}(x_{i_G})] I(E^c) f(\underset{\sim}{x})]$$

$$\equiv I_n + II_n + III_n + IV_n.$$

4.2 Rate of Convergence

The main result of this paper is regarding the rate of convergence of the regret Bayes risk of the empirical Bayes selection procedure δ_n^*. This result is stated as a theorem as follows.

Theorem 4.1 Let δ_n^* be the empirical Bayes selection procedure constructed in Section 3. Suppose that

(a) $\int_0^\infty \theta dG_i(\theta) < \infty$ for each $i = 1, \ldots, k$, and

(b) $M_i < \infty$ for each $i = 1, \ldots, k$.

Then, $R(G, \delta_n^*) - R(G, \delta_G) = O(\exp\{-\gamma n\})$ for some constant $\gamma > 0$.

Proof: To prove this theorem, it suffices to investigate the asymptotic behaviors of the four terms I_n, II_n, III_n and IV_n given in (4.6). First, we introduce the following notations:

$$b_i(M_i, \theta_0) = [h_i(M_i + 1) - \theta_0 h_i(M_i)] a(M_i + 1),$$

$$c_j(z, \theta_0) = [h_j(z + 1) - \theta_0 h(z)] / (\frac{1}{a(z + 1)} + \frac{\theta_0}{a(z)}),$$

$$s(x_i, x_j) = [\psi_i(x_i) + \psi_j(x_j)]/2,$$

$$C_1 = \theta_0 + \sum_{i=1}^{k} \int \theta dG_i(\theta)$$

11

(I) Let $\gamma_{1i} = 2[b_i(M_i, \theta_0) \min(1, \theta_0^{-1})/4]^2$ and $\gamma_1 = \min_{1 \leq i \leq k} \gamma_{1i}$. Note that $\gamma_{1i} > 0$ for each $i = 1, \ldots, k$ and therefore $\gamma_1 > 0$. By Proposition A.1,

$$I_n \leq \sum_{\underset{\sim}{x} \epsilon A_0} \sum_{i=1}^{k} d \exp\{-n\gamma_{1i}\} \psi_i(x_i) f(\underset{\sim}{x})$$

$$\leq d \exp\{-n\gamma_1\} \sum_{\underset{\sim}{x} \epsilon \mathcal{X}} \sum_{i=1}^{k} \psi_i(x_i) f(\underset{\sim}{x})$$

$$= d \exp\{-n\gamma_1\} [\sum_{i=1}^{k} \int_0^\infty \theta dG_i(\theta)].$$

(II) Let $\gamma_{2j} = 2 \min_{0 \leq z \leq m_j} c_j^2(z, \theta_0)$ and $\gamma_2 = \min_{1 \leq j \leq k} \gamma_{2j}$. Then $\gamma_{2j} > 0$ for each $j = 1, \ldots, k$, and therefore, $\gamma_2 > 0$. By Proposition A.2,

$$\mathrm{II}_n \leq \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \epsilon A_j} \sum_{z=0}^{m_j} \exp\{-n\gamma_{2j}\} \theta_0 f(\underset{\sim}{x})$$

$$\leq \theta_0 \exp\{-n\gamma_2\} [\max(m_1, \ldots, m_k) + 1] \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \epsilon A_j} f(\underset{\sim}{x})$$

$$\leq \theta_0 \exp\{-n\gamma_2\} [\max(m_1, \ldots, m_k) + 1]$$

since $\sum_{j=1}^{k} \sum_{\underset{\sim}{x} \epsilon A_j} f(\underset{\sim}{x}) \leq 1$.

(III) For each $i \neq j$, let $A_j(i) = \{\underset{\sim}{x} \epsilon A_j | x_i \leq M_i - 1\}$, $A_j^c(i) = \{\underset{\sim}{x} \epsilon A_j | x_i \geq M_i\}$. Thus,

III_n

$$= \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \epsilon A_j(i)} E_{\underset{\sim}{X}(n)} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, i \notin A(\underset{\sim}{x}), x_i \leq M_i - 1, E\} [\psi_i(x_i) - \psi_j(x_j)] f(\underset{\sim}{x})$$

$$+ \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \epsilon A_j^c(i)} E_{\underset{\sim}{X}(n)} I\{i_n^*(\underset{\sim}{x}) = i, i_G(\underset{\sim}{x}) = j, i \notin A(\underset{\sim}{x}), x_i \geq M_i, E\} [\psi_i(x_i) - \psi_j(x_j)] f(\underset{\sim}{x})$$

$$\equiv III_{n1} + III_{n2}.$$

For $\underset{\sim}{x} \epsilon A_j(i), x_j \leq m_j, x_i \leq M_i - 1$. Let

$$\gamma_{3ij1}(\underset{\sim}{x}) = 2[b_i(x_i, s(x_i, x_j)) \min(1, 1/s(x_i, x_j))/4]^2.$$

12

Note that $\gamma_{3ij1}(\underline{x}) > 0$. Also, $\gamma_{3ij1}(\underline{x})$ depends on $\underline{x}$ only through x_i and x_j, for which $0 \leq x_i \leq M_i - 1$, and $0 \leq x_j \leq m_j$. Therefore, $\gamma_{3ij1} = \min\limits_{\underline{x} \epsilon A_j(i)} \gamma_{3ij1}(\underline{x}) > 0$.

Let $\gamma_{3ij2}(\underline{x}, z) = 2c_j^2(z, s(x_i, x_j)), 0 \leq z \leq x_j$. Note that $\gamma_{3ij2}(\underline{x}, z) > 0$ for each $0 \leq z \leq x_j$. Therefore, $\gamma_{3ij2}(\underline{x}) \equiv \min\limits_{0 \leq z \leq x_j} \gamma_{3ij2}(\underline{x}, z) > 0, \gamma_{3ij2} \equiv \min\limits_{\underline{x} \epsilon A_j(i)} \gamma_{3ij2}(\underline{x}) > 0, \gamma_{3ij} \equiv$ $\min(\gamma_{3ij1}, \gamma_{3ij2}) > 0, \gamma_{3i} = \min\limits_{\substack{1 \leq j \leq k \\ j \neq i}} \gamma_{3ij} > 0$ and $\gamma_3 = \min\limits_{1 \leq i \leq k} \gamma_{3i} > 0$.

By Proposition A.3(b),

$$III_{n1} \leq \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underline{x} \epsilon A_j(i)} [d \exp\{-n\gamma_3\} + (x_j + 1) \exp\{-n\gamma_3\}] \psi_i(x_i) f(\underline{x})$$

$$\leq [d + \max(m_1, \ldots, m_k) + 1] \exp\{-n\gamma_3\} \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underline{x} \epsilon A_j(i)} \psi_i(x_i) f(\underline{x})$$

$$\leq [d + \max(m_1, \ldots, m_k) + 1] \exp\{-n\gamma_3\} k [\sum_{i=1}^{k} \int_0^\infty \theta dG_i(\theta)].$$

By Proposition A.3 (a) and the definition of γ_{1i},

$$III_{n2} \leq \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underline{x} \epsilon A_j^c(i)} d \exp\{-n\gamma_{1i}\} \psi_i(x_i) f(\underline{x})$$

$$\leq d \exp\{-n\gamma_1\} \sum_{j=1}^{k} \sum_{i=1}^{k} \sum_{\underline{x} \epsilon A_j^c(i)} \psi_i(x_i) f(\underline{x})$$

$$\leq d \exp\{-n\gamma_1\} k [\sum_{i=1}^{k} \int_0^\infty \theta dG_i(\theta)].$$

(IV) Let $\gamma_{4i} = \min(\ell n[1 - f_i(0)]^{-1}, \ell n[F_i(M_i + 1)]^{-1})$ and $\gamma_4 = \min\limits_{1 \leq i \leq k} \gamma_{4i}$.

Then $\gamma_{4i} > 0$ for each $i = 1, \ldots, k$ and therefore $\gamma_4 > 0$. By Proposition A.4,

$$IV_n \leq 2kC_1 \exp\{-n\gamma_4\},$$

where $C_1 = \theta_0 + \sum\limits_{i=1}^{k} \int_0^\infty \theta dG_i(\theta)$.

Now, $\gamma \equiv \min(\gamma_1, \gamma_2, \gamma_3, \gamma_4) > 0$. From the preceding analysis, we have: $I_n = O(\exp(-\gamma n)), II_n = O(\exp(-\gamma n)), III_n = O(\exp(-\gamma n))$ and $IV_n = O(\exp(-\gamma n))$. Therefore, from (4.6), we conclude that $R(G, \underline{\delta}_n^*) - R(G, \underline{\delta}_G) = O(\exp\{-\gamma n\})$.

13

Appendix

In the following, we introduce certain preliminary results without providing the proofs for presenting a concise proof for Theorem 4.1. The interested readers are referred to Gupta and Liang (1997) for the details of the proofs.

Proposition A.1 For $x \in A_0$ and $i \notin A(x)$,

$$E_{X(n)} I\{i_n^*(x) = i, i_G(x) = 0, i \notin A(x) \text{ and } E\}$$
$$\leq d \exp\{-2n[b_i(M_i, \theta_o) \min(1, \theta_0^{-1})/4]^2\},$$

where $b_i(M_i, \theta_0) = [h_i(M_i + 1) - \theta_0 h_i(M_i)] a(M_i + 1)$ and d is a constant independent of the distribution of $X(n)$.

Proposition A.2. For each $j = 1, \ldots, k$ and for $x \in A_j$,

$$E_{X(n)} I\{i_n^*(x) = 0, \ i_G(x) = j \text{ and } E\} \leq \sum_{z=0}^{m_j} \exp\{-2n c_j^2(z, \theta_0)\},$$

where $c_j(z, \theta_0) = [h_j(z + 1) - \theta_0 h_j(z)] / (\frac{1}{a(z+1)} + \frac{\theta_0}{a(z)})$.

For each $j = 1, \ldots, k$, each $x \in A_j$ and $i \notin A(x), \psi_j(x_j) = \min_{0 \leq l \leq k} \psi_\ell(x_\ell) < \min(\theta_0, \psi_i(x_i))$. There are two cases regarding the value of x_i : either $x_i \geq M_i$ for which $\psi_i(x_i) > \theta_0$ or $x_i \leq M_i - 1$. Thus,

$$E_{X(n)} I\{i_n^*(x) = i, i_G(x) = j, i \notin A(x) \text{ and } E\}$$
$$= E_{X(n)} I\{i_n^*(x) = i, i_G(x) = j, i \notin A(x), x_i \geq M_i \text{ and } E\}$$
$$+ E_{X(n)} I\{i_n^*(x) = i, i_G(x) = j, i \notin A(x), x_i \leq M_i - 1 \text{ and } E\}.$$

Proposition A.3

(a) $E_{X(n)} I\{i_n^*(x) = i, i_G(x) = j, i \notin A(x), x_i \geq M_i \text{ and } E\}$

$\leq d \exp\{-2n[b_i(M_i, \theta_0) \min(1, \theta_0^{-1})/4]^2\}.$

(b) $E_{X(n)} I\{i_n^*(x) = i, i_G(x) = j, i \notin A(x), x_i \leq M_i - 1 \text{ and } E\}$

$\leq d \exp\{-2n[b_i(x_i, s(x_i, x_j)) \min(1, 1/s(x_i, x_j))/4]^2\}$

14

$$+ \sum_{z=0}^{x_j} \exp\{-2nc_j^2(z, s(x_i, x_j))\},$$

where $s(x_i, x_j) = [\psi_i(x_i) + \psi_j(x_j)]/2$.

Proposition A.4

$$E_{\underset{\sim}{X}(n)}\{[\sum_{\underset{\sim}{x} \in \mathcal{X}}[\psi_{i_n^*}(x_{i_n^*}) - \psi_{i_G}(x_{i_G})]f(x)]I(E^c)\}$$

$$\leq C_1 \sum_{i=1}^{k} [\exp\{-n \ \ell n[1 - f_i(0)]^{-1}\} + \exp\{-n \ \ell n[F_i(M_i + 1)]^{-1}\}],$$

where $C_1 = \theta_0 + \sum_{i=1}^{k} \int \theta dG_i(\theta)$.

References

Alam, K. (1971). Selection from Poisson processes. *Ann. Inst. Statist. Math.*, **23**, 411–418.

Alam, K. and Thompson, J. R. (1973). A problem of ranking and estimation with Poisson processes. *Technometrics*, **15**, 801–808.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley, New York.

Deely, J. J. (1965). Multiple decision procedures from an empirical Bayes approach. Ph.D. Thesis (Mimeo. Ser. No. 45), Department of Statistics, Purdue University, West Lafayette, Indiana.

Dixon, D. O. and Bland, R. P. (1971). A Bayes solution to the problem of ranking Poisson parameters. *Ann. Inst. Statist. Math.*, **23**, 119–124.

Gupta, S. S., Leong, Y. K. and Wong, W. Y. (1979). On subset selection procedures for Poisson populations. *Bull. Malaysian Math. Soc.*, **2**, 89–110.

Gupta, S. S., and Liang, T. (1997). Selecting the most reliable Poisson population provided it is better than a control: A nonparametric empirical Bayes approach. Technical Report # 97-9C, Department of Statistics, Purdue University, West Lafayette, Indiana.

Gupta, S. S., Liang, T. and Rau, R. B. (1994). Empirical Bayes two-stage procedures for selecting the best Bernoulli population compared with a control. In : *Statistical Decision Theory and Related Topics–V*, (Eds. S. S. Gupta and J. O. Berger), Springer-Verlag, Berlin, New York, 277–292.

Gupta, S. S., Liang, T. and Rau, R.-B. (1995). Empirical Bayes two-stage procedures for selecting the best normal population compared with a control. In : *Symposia Gaussiana Conf. B*, (Eds. Mammitzsch and Schneeweib), Walter de Gruyter & Co., Berlin, New York, 171–196.

Gupta, S. S. and Panchapakesan, S. (1988). Selection and Ranking procedures in reliability models. In: *Handbook of Statistics*, **Vol. 7**, (Eds. P. R. Krishnaiah and C. R. Rao), Elsevier Science Publishers B. V., 131–156.

Puri, P. S. and Singh, H. (1990). On recursive formulas for isotonic regression useful for statistical inference under order restrictions. *J. Statist. Plann. Inference*, **24**, 1–11.

Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkeley SympMath. Statist. Probab.*, **1**, 157–163, University of California Press.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist..* **35**, 1–20.