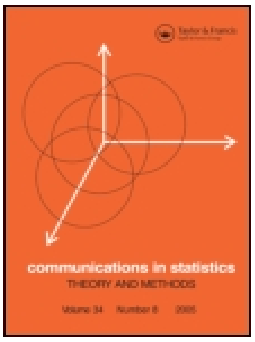

Taylor & Francis
Taylor & Francis Group

Communications in Statistics—Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Selecting the Best Process Based on Capability Index via Empirical Bayes Approach

Wen-Tao Huang & Yao-Tsung Lai

Communications in Statistics—Theory and Methods, 38: 1576–1588, 2009

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920802715024

Ranking and Selection Theory

Selecting the Best Process Based on Capability Index via Empirical Bayes Approach

WEN-TAO HUANG[1] AND YAO-TSUNG LAI[2]

[1]Department of Management Sciences and Decision Making, Tamkang University, Taiwan, ROC
[2]Department of Statistics and Actuarial Sciences, Aletheia University, Taiwan, ROC

Consider k $(k \geq 2)$ manufacturing processes whose mean θ_i, variance σ_i^2 and process capability index $C_{pw}(i)$, $i = 1, \ldots, k$, are all unknown. For two given control values $C_{pw}(0)$ and σ_0^2, we are interested in selecting some process whose capability index is no less than $C_{pw}(0)$ and is the largest in the qualified subset in which each process variance is no larger than σ_0^2. Under a Bayes framework, we consider the normally distributed manufacturing processes taking normal-gamma as its conjugate prior. A Bayes approach is set up and an empirical Bayes procedure is proposed which has been shown to be asymptotically optimal. A simulation study is carried out for the performance of the proposed procedure and it is found practically useful.

Keywords Asymptotic optimality; Empirical Bayes rule; Process capability index; Ranking and selection; C_{pw}.

Mathematics Subject Classification Primary 62C12; Secondary 90B50.

1. Introduction

To understand and evaluate a process, one of effective methods is to consider some quantitative measure to estimate the performance of the process under study. The well-known measure of product quality in industry is the capability index. It is a dimensionless measure based on some parameters and specifications that are involved in the process.

In most literature related to capability index, it is mainly focused on the estimation and testing problems. In many practical applications, instead of these topics, there occurs a problem that relates to select the most desirable manufacturing

Received May 1, 2008; Accepted December 29, 2008
Address correspondence to Wen-Tao Huang, Department of Management Sciences and Decision Making, Tamkang University, Taipei, Taiwan, ROC; E-mail: akenwt@yahoo.com.tw

process among several available processes. Suppose a new product is under study for production, and there are k processes to produce it. We are interested in identifying one of them as the most desirable process to produce the product.

For selecting the best manufacturing process, Tseng and Wu (1991) considered the selection problem in terms of capability index C_p which is introduced by Kane (1986a,b). Since the difference between the upper and lower specification limit is a known quantity, the problem considered in Tseng and Wu (1991) is equivalent to select the process which is corresponding to the smallest variance. There have been several capability indices such as C_{pm} (see Chan et al., 1988), C_{pk} (see Gunter, 1989), C_{pmk} (see Pearn et al., 1992), among others. However, mostly C_{pm} and C_{pk} are widely used. Spiring (1997) modified C_{pm} and proposed C_{pw} which included C_p, C_{pm}, and C_{pk} as special case. So in this article, we consider selecting the best process in terms of C_{pw} which is a modified quantity of C_{pm} taking weight between the variance and the square difference between mean and target. Furthermore, we consider another criterion so that the capability index of the process selected should be larger than a prefixed value which can be considered as a control. Some related selection problem of identifying the best under some other multicriteria has been studied in Huang and Lai (1999), among others.

The main purpose of this article is to propose a new selection rule for practical applications of selecting the best manufacturing process in the area of industrial statistics. We formulate the problem in a Bayes framework and apply the empirical Bayes method which is pioneered by Robbins (1956, 1964). We focus on practical computation and implement an alternative selection rule in the field of manufacturing process. In this article, we do not intend to develop any new theoretical results in the area of empirical Bayes methodology. Instead of evaluating the convergence rate of regret, we carry out some simulation study to see its concrete behavior of correct selection which maybe more helpful to applied statisticians working in this area.

Since we consider our selection based on multicriteria, construction of a perfect loss function confronts a dilemma, because it involves the priorities of criteria. It is reasonable to question why penalty for committing error A is more serious than that of committing error B. However, if one tries to change this situation, another analogous problem would also happen.

In this Bayes setup, we consider normal-gamma for its prior. As is well known, conjugate prior has its own justification and it has been widely applied in many areas. As previously mentioned, we are mainly interested in practical applications in selection problem. We have no intention to use more general nonparametric prior for its mathematical generalization.

In Sec. 2, we formulate the problem and develop the Bayes framework. In Sec. 3, we propose an empirical Bayes procedure and which is shown to be asymptotically optimal. In the last section some Monte Carlo simulation results are given to show the performance of the proposed procedure.

2. Formulation of Problem and a Bayes Decision Rule

In this article, we utilize the process capability index proposed by Spiring (1997) to evaluate the effectiveness of a manufacturing process. This index is defined as follows.

C_{pw} : Let π be a normal manufacturing process with unknown mean θ and unknown variance σ^2. T is the target value, and USL and LSL are, respectively, the upper and lower specification limit, and which are all prefixed. A process capability index C_{pw} of π is defined as the following:

$$C_{pw} = \frac{USL - LSL}{6\sqrt{\sigma^2 + w(\theta - T)^2}},$$

where w $(0 \leq w \leq 1)$ is a weight.

According to process capability index introduced above, we define the best manufacturing process as follows.

Criteria for Selection. Let $\pi_1, \ldots, \pi_k$, be k normal manufacturing processes such that π_i has mean θ_i, variance σ_i^2, and process capability index $C_{pw}(i)$, $i = 1, \ldots, k$. Let $C_{pw}(0)$ and σ_0^2 be two known control values (prefixed). Define $S = \{\pi_i \mid \sigma_i \leq \sigma_0, i = 1, \ldots, k\}$. A manufacturing process π_i is considered as the best, if $C_{pw}(i) \geq C_{pw}(0)$ and it simultaneously satisfies the following conditions:

(i) $\pi_i \in S$, and
(ii) $C_{pw}(i) = \text{Max}_{\pi_j \in S} C_{pw}(j)$.

Let $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)$, $\underset{\sim}{\sigma} = (\sigma_1, \ldots, \sigma_k)$, and $\Omega = \{(\theta_i, \sigma_i) | -\infty < \theta_i < +\infty, \sigma_i > 0, i = 1, \ldots, k\}$ be the parameter space. Let $\underset{\sim}{a} = (a_0, a_1, \ldots, a_k)$ denote an action, where $a_i = 0$, or 1; $i = 0, 1, \ldots, k$, and $\sum_{i=0}^{k} a_i = 1$. If $a_i = 1$, for some $i = 1, \ldots, k$, it means that manufacturing process π_i is selected as the best. When $a_0 = 1$, it means that no manufacturing process is considered as the best, i.e., none in k manufacturing processes satisfies the selection criteria. Let $\mathscr{A} = \{\underset{\sim}{a}\}$ denote the action space.

For the sake of convenience, corresponding to $C_{pw}(i)$, $\forall i = 0, 1, \ldots, k$, we define a new quantity $C'_{pw}(i)$ as follows. For a given positive $C^*_{pw} < C_{pw}(0)$ and for $i = 0, 1, \ldots, k$, define

$$C'_{pw}(i) = C_{pw}(i) I_{\{\sigma_i \leq \sigma_0\}} + C^*_{pw} I_{\{\sigma_i > \sigma_0\}}. \tag{1}$$

It is clear to see that those manufacturing processes which do not meet the requirement (i) will also fail to meet the requirement (ii) in selection criteria in terms of the associated quantity $C'_{pw}(i)$.

In a decision-theoretic approach, we consider the following.

Loss Function. For a given control value $C_{pw}(0)$, and parameter vectors $\underset{\sim}{\theta}, \underset{\sim}{\sigma}$, if action $\underset{\sim}{a}$ is taken, a loss $L(\underset{\sim}{\theta}, \underset{\sim}{\sigma}; \underset{\sim}{a})$ is incurred and which is defined by

$$L(\underset{\sim}{\theta}, \underset{\sim}{\sigma}; \underset{\sim}{a}) = \sum_{i=0}^{k} a_i C'^{-2}_{pw}(i) - C'^{-2}_{pw[k]}, \tag{2}$$

where $C'_{pw[k]} = \text{Max}_{0 \leq i \leq k} C'_{pw}(i)$.

Note that here we are treating selection problem rather than estimation. For instance, if we take an alternative loss such as $L(\underset{\sim}{\theta}, \underset{\sim}{\sigma}; \underset{\sim}{a}) = C'_{pw[k]} - \sum_{i=0}^{k} a_i C'_{pw}(i)$.

Then, this loss essentially is equivalent to that defined by (2) in the sense that they both select the same process under same observed data. Since for a given data set, the ith process which corresponds to the minimum loss in (2) is the same as that corresponding to the minimum in the alternative loss. Yet, their respective penalties corresponding to same incorrect selection may be quite different. This means that the quantity of a penalty due to an incorrect selection (thus incurring a loss) is rather in the sense of "relativeness" than that of "absoluteness". Furthermore, we note that two different incorrect selections may cause two quite different penalties and it is difficult to judge which is more serious than the other in many situations, because it concerns the priority order of multicriteria.

The so defined loss $L(\underset{\sim}{\theta}, \underset{\sim}{\sigma}; \underset{\sim}{a})$ has reflected penalty for a wrong action. However, it may not reflect reasonable penalty in all cases. For its simplicity, we propose the loss (2) for our selection problem. In this article, we consider an empirical Bayes approach for the problem of selecting the best manufacturing process which is normally distributed.

We would like to point out that in many occasions the random observation X of some characteristic under study is not exactly normally distributed. So the usual normality assumption may not be satisfied under certain circumstances. To modify this difficulty, we consider a Bayes framework and this expands in some extent the family of distributions of X under consideration.

For each $i = 1, \ldots, k$, let $X_{i1}, \ldots, X_{iM}$ be an independently random sample of size M from π_i with mean θ_i and variance σ_i^2 and its observed value is denoted by $x_{i1}, \ldots, x_{iM}$, respectively. For its convenience, let $\tau_i = 1/\sigma_i^2$, $i = 1, \ldots, k$. It is assumed that (θ_i, τ_i) is a realization of a random vector (Θ_i, Γ_i) with a normal-gamma prior distribution which is the product of conditional normal distribution $N(\mu_i, [(2\alpha_i - 1)\tau_i]^{-1})$ of Θ_i given τ_i, and a marginal gamma prior distribution $Gamma(\alpha_i, \beta_i)$ of Γ_i.

As we have previously mentioned, we consider this selection problem for practical application as our main purpose. So it is important to consider the problem of applicability for practical application when sample size is small. If a nonparametric prior is considered, though it covers a larger family of priors, it still needs to impose some conditions on the prior density and one may still ask what happens if the conditions are not satisfied. However, the key point to consider a parametric prior is that it concerns the applicability for the situation that the sample is small. As can be seen in Table 1 in Sec. 4, the rate of correct selection is more than 85% even when the sample size is as small as 10. It may have difficulty to achieve it when a nonparametric prior is considered. So we restrict ourselves to a conjugate priors in our set up.

For convenience of notation, we denote $\underset{\sim}{x}_i = (x_{i1}, \ldots, x_{iM})$, $x_i = \frac{1}{M}\sum_{j=1}^{M} x_{ij}$, and $s_i^2 = \frac{1}{M-1}\sum_{j=1}^{M}(x_{ij} - x_i)^2$, for $i = 1, \ldots, k$. It has been shown that (see Raiffa and Schlaifer, 1961) the conditional posterior distribution of Θ_i given $\underset{\sim}{x}_i$ and τ_i is a normal distribution $N(\varphi_i(x_i), [(2\alpha_i + M - 1)\tau_i]^{-1})$, where

$$\varphi_i(x_i) = \mathrm{E}[\Theta_i \mid \underset{\sim}{x}_i, \tau_i] = \frac{(2\alpha_i - 1)\mu_i + Mx_i}{2\alpha_i + M - 1}, \tag{3}$$

Table 1
Behavior of empirical Bayes rules with respect to various sample sizes ($w = 1$)

n	f_n	$\overline{D}_n$	$n\overline{D}_n$	$SE(\overline{D}_n)$
10	0.8556	1.7125E-02	1.7125E-01	3.3902E-03
20	0.8917	9.4061E-03	1.8812E-01	1.3359E-03
30	0.9110	5.9942E-03	1.7983E-01	6.4524E-04
40	0.9231	4.8998E-03	1.9599E-01	5.0000E-04
50	0.9339	3.6897E-03	1.8449E-01	3.4546E-04
60	0.9431	2.8443E-03	1.7066E-01	2.3656E-04
70	0.9435	2.5919E-03	1.8144E-01	2.0210E-04
80	0.9461	2.3654E-03	1.8923E-01	1.6538E-04
90	0.9522	1.9827E-03	1.7845E-01	1.4753E-04
100	0.9491	1.9496E-03	1.9496E-01	1.2354E-04
200	0.9674	9.4790E-04	1.8958E-01	4.5927E-05
300	0.9753	5.6293E-04	1.6888E-01	2.1498E-05
400	0.9775	4.4768E-04	1.7907E-01	1.4806E-05
500	0.9808	3.0333E-04	1.5166E-01	8.2669E-06
600	0.9788	3.4662E-04	2.0797E-01	9.3995E-06
700	0.9802	2.7777E-04	1.9444E-01	6.6305E-06
800	0.9830	2.0734E-04	1.6587E-01	4.0981E-06
900	0.9818	2.2765E-04	2.0488E-01	4.8548E-06
1000	0.9849	1.6283E-04	1.6283E-01	3.3842E-06

and the marginal posterior distribution of Γ_i given $\underset{\sim}{x}_i$ is a gamma distribution $Gamma(\alpha'_i, \eta_i)$, where

$$\begin{aligned} \alpha'_i &= 2\alpha_i + \frac{M}{2} - 1, \quad \text{and} \\ \eta_i &= \beta_i + \frac{(M-1)s_i^2}{2} + \frac{(2\alpha_i - 1)M(x_i - \mu_i)^2}{2(2\alpha_i + M - 1)}. \end{aligned} \tag{4}$$

The random vectors $(\Theta_1, \Gamma_1), \ldots, (\Theta_k, \Gamma_k)$ are assumed to be mutually independent.

Let $\underset{\sim}{x} = (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_k)$ and χ be the sample space generated by $\underset{\sim}{x}$. A selection rule $\underset{\sim}{d} = (d_0, d_1, \ldots, d_k)$ is a mapping defined on the sample space χ into the $k+1$ product space $[0, 1] \times [0, 1] \times \cdots \times [0, 1]$ such that $\sum_{i=0}^{k} d_i(\underset{\sim}{x}) = 1$, for all $\underset{\sim}{x} \in \chi$. For every $\underset{\sim}{x} \in \chi$, $d_i(\underset{\sim}{x})$ denotes the probability of selecting manufacturing process π_i as the best, $i = 1, \ldots, k$; and $d_0(\underset{\sim}{x})$ denotes the probability that none is selected as the best.

For ease of notation, let $\underset{\sim}{\tau} = (\tau_1, \ldots, \tau_k)$, $\underset{\sim}{\mu} = (\mu_1, \ldots, \mu_k)$, $\underset{\sim}{\alpha} = (\alpha_1, \ldots, \alpha_k)$, $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_k)$, $\underset{\sim}{\Theta} = (\Theta_1, \ldots, \Theta_k)$, and $\underset{\sim}{\Gamma} = (\Gamma_1, \ldots, \Gamma_k)$. Let $h(\underset{\sim}{\theta}, \mid \underset{\sim}{x}, \underset{\sim}{\tau}; \underset{\sim}{\mu}, \underset{\sim}{\alpha})$ be the joint conditional posterior probability density function of $\underset{\sim}{\Theta}$ given $\underset{\sim}{x}$ and $\underset{\sim}{\tau}$, and $g(\underset{\sim}{\tau} \mid \underset{\sim}{x}; \underset{\sim}{\alpha}, \underset{\sim}{\beta})$ be the joint conditional posterior probability density function of $\underset{\sim}{\Gamma}$ given $\underset{\sim}{x}$. Let $h_i(\theta_i \mid \underset{\sim}{x}_i, \tau_i; \mu_i, \alpha_i)$ and $g_i(\tau_i \mid \underset{\sim}{x}_i; \alpha_i, \beta_i)$ be the conditional posterior probability density function of Θ_i and Γ_i, respectively. Under the preceding formulation, the Bayes risk of a selection rule $\underset{\sim}{d}$, denoted by $r(\underset{\sim}{d})$, can be then computed.

According to notations of $h(\cdot)$ and $g(\cdot)$ which have been defined previously, we compute the Bayes risk as follows.

Define

$$\phi_i(\underset{\sim}{x_i}) = \frac{36}{(USL-LSL)^2}\left\{\frac{(w+(2\alpha_i+M-1))(1-G(\tau_0\,|\,\alpha_i'-1,\eta_i))}{(2\alpha_i+M-1)(\alpha_i'-1)\eta_i} + w(\varphi_i(x_i)-T)^2(1-G(\tau_0\,|\,\alpha_i',\eta_i))\right\} + C_{pw}^{*-2}G(\tau_0\,|\,\alpha_i',\eta_i), \quad (5)$$

where $G(x\,|\,b,c)$ is a Gamma cumulative distribution function (cdf) with parameters b and c, respectively. For convenience of notation, we define $\phi_0(\underset{\sim}{x_0}) = C_{pw}^{-2}(0)$.

For some constant C and denoting the joint marginal density of $\underset{\sim}{X}$ by $f(\underset{\sim}{x})$, the Bayes risk is given by

$$r(\underset{\sim}{d}) = \int_\chi \sum_{i=0}^{k} d_i(\underset{\sim}{x})\phi_i(\underset{\sim}{x_i})f(\underset{\sim}{x})d\underset{\sim}{x} - C. \quad (6)$$

Bayes Rule. For each $\underset{\sim}{x} \in \chi$, let

$$Q(\underset{\sim}{x}) = \left\{i\,|\,\phi_i(\underset{\sim}{x_i}) = \operatorname*{Min}_{0\le j\le k}\phi_j(\underset{\sim}{x_j}),\, i=0,1,\ldots,k\right\}. \quad (7)$$

Define

$$i^* = i^*(\underset{\sim}{x}) = \begin{cases} 0 & \text{if } Q(\underset{\sim}{x}) = \{0\}, \\ \mathrm{Min}\{i\,|\,i\in Q(\underset{\sim}{x}),\, i\neq 0\} & \text{otherwise.}\end{cases} \quad (8)$$

Then, according to (5), (7), and (8), it can be derived that a Bayes selection rule $\underset{\sim}{d}^B = (d_0^B, d_1^B, \ldots, d_k^B)$ is given as follows:

$$\begin{cases} d_{i^*}^B(\underset{\sim}{x}) = 1, \\ d_j^B(\underset{\sim}{x}) = 0, \quad \text{for } j\neq i^*. \end{cases} \quad (9)$$

3. The Empirical Bayes Selection Rule

In the problem formulated in Sec. 2, we consider that $\alpha_1, \ldots, \alpha_k$ are all known with $\alpha_i > 1$ for all i. Since $\phi_i(\underset{\sim}{x_i})$ still involves the unknown parameters $\mu_i, \beta_i, i = 1, \ldots, k$, hence, the proposed Bayes selection rule $\underset{\sim}{d}^B$ is not applicable. However, based on the past data, these unknown parameters can be estimated and a decision can be made if one more observation is taken. For $i = 1, \ldots, k$, let X_{ijt} denote a sample of size M from π_i with a normal distribution $N(\theta_{it}, \tau_{it}^{-1})$ at time t $(t = 1, \ldots, n)$, $j = 1, \ldots, M$, and (θ_{it}, τ_{it}) is a realization of a random vector $(\Theta_{it}, \Gamma_{it})$ which is an independent copy of (Θ_i, Γ_i) with a normal-gamma distribution described in preceding section. It is assumed that $(\Theta_{it}, \Gamma_{it})$, $i = 1, \ldots, k$, $t = 1, \ldots, n$, are mutually independent.

For our convenience, we denote the current random sample X_{ijn+1} by X_{ij}, for $j = 1, \ldots, M$, $i = 1, \ldots, k$.

For each π_i, $i = 1, \ldots, k$, we estimate the unknown parameters μ_i and β_i based on the past data X_{ijt}, $j = 1, \ldots, M$, $t = 1, \ldots, n$. We denote

$$\begin{cases} X_{i.t} = \dfrac{1}{M}\displaystyle\sum_{j=1}^{M} X_{ijt}, \quad X_i(n) = \dfrac{1}{n}\displaystyle\sum_{t=1}^{n} X_{i.t}, \\ W_{i.t}^2 = \dfrac{1}{M-1}\displaystyle\sum_{j=1}^{M}(X_{ijt} - X_{i.t})^2, \quad W_i^2(n) = \dfrac{1}{n}\displaystyle\sum_{t=1}^{n} W_{i.t}^2. \end{cases} \tag{10}$$

For ease of notation, we define μ_{in} and β_{in} as estimators of μ_i and β_i, respectively, by the following:

$$\begin{cases} \mu_{in} = X_i(n), \\ \beta_{in} = (\alpha_i - 1)W_i^2(n). \end{cases} \tag{11}$$

Also, for $i = 1, \ldots, k$, we define

$$\phi_{in}(\underset{\sim}{x_i}) = \frac{36}{(USL - LSL)^2}\left\{\frac{(w + (2\alpha_i + M - 1))(1 - G(\tau_0 \mid \alpha_i' - 1, \eta_{in}))}{(2\alpha_i + M - 1)(\alpha_i' - 1)\eta_{in}} + w(\varphi_{in}(x_i) - T)^2(1 - G(\tau_0 \mid \alpha_i', \eta_{in}))\right\} + C_{pw}^{*-2}G(\tau_0 \mid \alpha_i', \eta_{in}). \tag{12}$$

where

$$\eta_{in} = \beta_{in} + \frac{(M-1)s_i^2}{2} + \frac{(2\alpha_i - 1)M(x_i - \mu_{in})^2}{2(2\alpha_i + M - 1)}, \tag{13}$$

and

$$\varphi_{in}(x_i) = \frac{(2\alpha_i - 1)\mu_{in} + Mx_i}{2\alpha_i + M - 1}. \tag{14}$$

For convenience, we may define $\phi_{0n}(\underset{\sim}{x_0}) = C_{pw}^{-2}(0)$. We consider $\phi_{in}(\underset{\sim}{x_i})$ to be an estimator of $\phi_i(\underset{\sim}{x_i})$.

For each $\underset{\sim}{x} \in \chi$, let

$$Q_n(\underset{\sim}{x}) = \left\{i \mid \phi_{in}(\underset{\sim}{x_i}) = \underset{0 \le j \le k}{\text{Min}}\, \phi_{jn}(\underset{\sim}{x_j}), i = 0, 1, \ldots, k\right\}. \tag{15}$$

Again, define

$$i_n^* = i_n^*(\underset{\sim}{x}) = \begin{cases} 0 & \text{if } Q_n(\underset{\sim}{x}) = \{0\}, \\ \text{Min}\{i \mid i \in Q_n(\underset{\sim}{x}), i \neq 0\} & \text{otherwise.} \end{cases} \tag{16}$$

Then, according to (12), (15), and (16), we conclude that an empirical Bayes selection rule $\underset{\sim}{d}^{*n} = (d_0^{*n}, d_1^{*n}, \ldots, d_k^{*n})$ is given as follows:

$$\begin{cases} d_{i_n^*}^{*n}(\underset{\sim}{x}) = 1, \\ d_j^{*n}(\underset{\sim}{x}) = 0, \quad \text{for } j \neq i_n^*. \end{cases} \tag{17}$$

According to Robbins (1964), a sequence of empirical Bayes selection rule $\{\underset{\sim}{d}^n\}_{n=1}^{\infty}$ is said to be asymptotically optimal, if $\lim_{n\to\infty}[\mathrm{E}_n[r(\underset{\sim}{d}^n)] - r(\underset{\sim}{d}^B)] = 0$.

Note that the estimators $W_i^2(n)$ (defined in (10)), μ_{in}, β_{in} (both defined in (11)), η_{in} (defined in (13)), $\varphi_{in}(x_i)$ (defined in (14)), and $\phi_{in}(\underset{\sim}{x}_i)$ (defined in (12)) are, respectively, consistent estimators of $\beta_i/(\alpha_i - 1)$, μ_i, β_i, η_i, $\varphi_i(x_i)$, and $\phi_i(\underset{\sim}{x}_i)$, $i = 1, \ldots, k$.

Checking conditions (C) and (D) of Robbins (1964) and applying the Dominated Convergence Theorem, we can conclude the following.

Proposition 3.1. *The empirical Bayes selection rule* $\underset{\sim}{d}^{*n}(\underset{\sim}{x})$, *defined in* (15)–(17), *is asymptotically optimal.*

4. Simulation Study

Since we emphasize applicability of our selection rule for small sample size, and also we are interested in asking if the result of simulation study is acceptable for small size of sample though it is asymptotic optimal. So a simulation study is important and worthwhile.

To investigate the performance of proposed empirical Bayes selection rule $\underset{\sim}{d}^{*n}(\underset{\sim}{x})$ defined in Sec. 3, we carry out some simulation studies which are summarized in Tables 1 and 2 and Figs. 1–5. The quantity $E_n[r(\underset{\sim}{d}^{*n})] - r(\underset{\sim}{d}^B)$ is used as a measure of performance of the empirical Bayes selection rule $\underset{\sim}{d}^{*n}(\underset{\sim}{x})$.

For a given current observation $\underset{\sim}{x}$ and given past observation x_{ijt}, let

$$D_n(\underset{\sim}{x}) = \sum_{i=0}^{k} \left[d_i^{*n}(\underset{\sim}{x}) - d_i^B(\underset{\sim}{x})\right]\phi_i(\underset{\sim}{x}_i) = \phi_{i_n^*}(\underset{\sim}{x}_{i_n^*}) - \phi_{i^*}(\underset{\sim}{x}_{i^*}).$$

Then

$$E_n\left[r(\underset{\sim}{d}^{*n})\right] - r(\underset{\sim}{d}^B) = E[E_n[D_n(\underset{\sim}{X})]].$$

Therefore, the sample mean of $D_n(\underset{\sim}{x})$ based on the observations of $\underset{\sim}{x}$ and x_{ijt}, $i = 1, \ldots, k$, $j = i, \ldots, M$, $t = 1, \ldots, n$, can be used as an estimator of $E_n[r(\underset{\sim}{d}^{*n})] - r(\underset{\sim}{d}^B)$.

We briefly explain the simulation scheme as follows.

(1) For each time t, $t = 1, \ldots, n$, and each manufacturing process π_i, $i = 1, \ldots, k$, generate observations $x_{i1t}, \ldots, x_{iMt}$, by the following way.

(1.a) Take a value τ_{it} according to distribution $Gamma(\alpha_i, \beta_i)$.

Table 2
The frequency of the process selected as the best under various weights for Group 2 ($n = 100$)

Weight	CD	Process 0	1	2	3	4	5
0.0	9914	0	0	0	1609	38	8353
		(0)	(0)	(0)	(1695)	(37)	(8268)
		[0]	[0]	[0]	[1609]	[37]	[8268]
0.1	9454	0	8	2714	7268	10	0
		(0)	(8)	(2772)	(7208)	(12)	(0)
		[0]	[7]	[2471]	[6966]	[10]	[0]
0.2	9437	0	11	2641	7334	14	0
		(0)	(13)	(2695)	(7275)	(17)	(0)
		[0]	[8]	[2390]	[7027]	[12]	[0]
0.3	9414	0	8	2687	7294	11	0
		(0)	(9)	(2746)	(7232)	(13)	(0)
		[0]	[8]	[2424]	[6972]	[10]	[0]
0.4	9442	23	9	2708	7255	5	0
		(28)	(8)	(2789)	(7170)	(5)	(0)
		[19]	[6]	[2475]	[6939]	[3]	[0]
0.5	9250	1204	0	1662	7133	1	0
		(1233)	(0)	(1699)	(7067)	(1)	(0)
		[1014]	[0]	[1425]	[6810]	[1]	[0]
0.6	9364	3324	0	219	6457	0	0
		(3362)	(0)	(235)	(6403)	(0)	(0)
		[3049]	[0]	[182]	[6133]	[0]	[0]
0.7	9286	4394	0	32	5574	0	0
		(4362)	(0)	(34)	(5603)	(1)	(0)
		[4023]	[0]	[26]	[5237]	[0]	[0]
0.8	9283	5056	0	12	4932	0	0
		(5095)	(0)	(12)	(4893)	(0)	(0)
		[4718]	[0]	[11]	[4554]	[0]	[0]
0.9	9323	5615	0	2	4383	0	0
		(5606)	(0)	(2)	(4392)	(0)	(0)
		[5272]	[0]	[2]	[4049]	[0]	[0]
1.0	9326	6073	0	2	3925	0	0
		(6053)	(0)	(2)	(3945)	(0)	(0)
		[5726]	[0]	[2]	[3598]	[0]	[0]

(1.b) Take a value θ_{it} according to distribution $N(\mu_i, [(2\alpha_i - 1)\tau_{it}]^{-1})$.

(1.c) Let $\sigma_{it}^2 = \tau_{it}^{-1}$. For given θ_{it} and σ_{it}^2, generate random samples $x_{i1t}, \ldots, x_{iMt}$, according to distribution $N(\theta_{it}, \sigma_{it}^2)$.

(2) Based on the observations x_{ijt}, $i = 1, \ldots, k$, $j = i, \ldots, M$, $t = 1, \ldots, n$, estimate the unknown parameters μ_i and β_i according to (11) and they are denoted by μ_{in} and β_{in}, respectively.

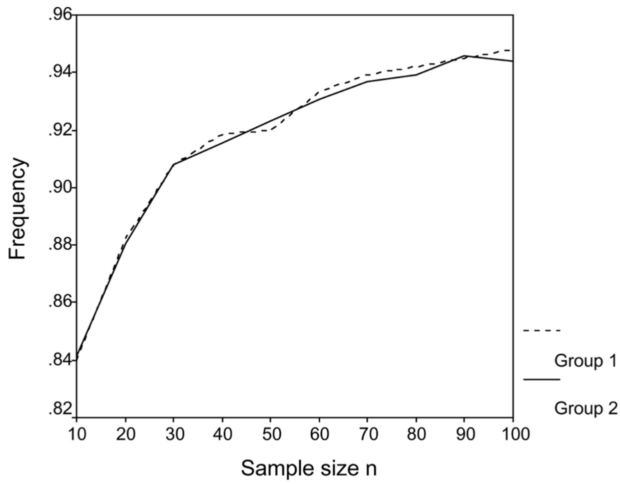

Figure 1. Relative frequency of correct decision with respect to sample sizes for both groups of manufacturing processes ($w = 0.5$).

(3) For each manufacturing process π_i, $i = 1, \ldots, k$, repeat Step (1) with $t = n + 1$ and $j = 1, \ldots, M$, and take its sample mean as our current sample x_i. Thus, the current sample vector is given by $\underset{\sim}{x} = (x_1, \ldots, x_k)$.
(4) For given value of weight w and control value $C_{pw}(0)$, based on the current sample vector, determine the Bayes selection rule $\underset{\sim}{d}^B$ and the empirical Bayes selection rule $\underset{\sim}{d}^{*_n}$ according to (9) and (17), respectively. Then compute $D_n(\underset{\sim}{x})$.
(5) Repeat Steps (1)–(4) 10,000 times, and then take its average denoted by $\overline{D}_n$ which is used as an estimate of $E[r(\underset{\sim}{d}^{*_n})] - r(\underset{\sim}{d}^B)$. Let $SE(\overline{D}_n)$ denote the estimated standard deviation.

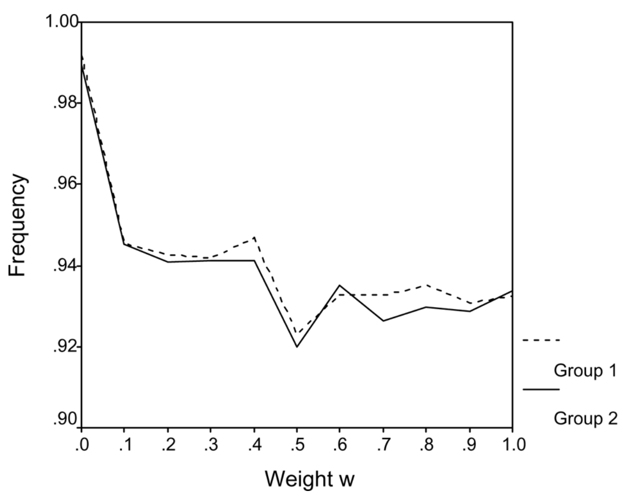

Figure 2. Relative frequency of correct decision with respect to different weight for both groups of manufacturing processes ($n = 50$).

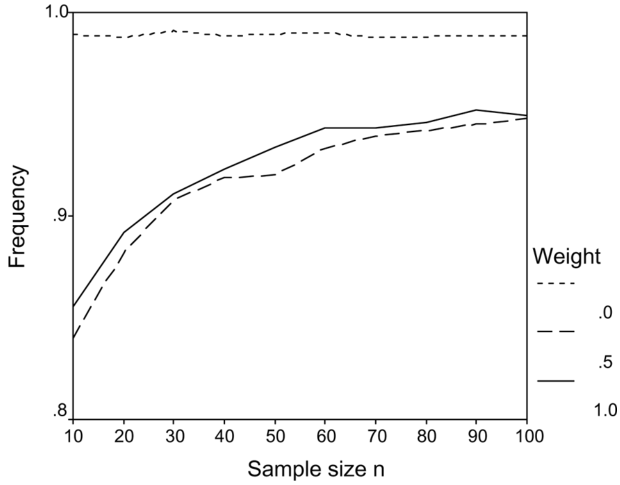

Figure 3. Relative frequency of correct decision with respect to sample size for three different weight values for Group 1.

In this simulation study, we consider two groups of manufacturing processes, called Group 1 and Group 2. For both groups, we take $M = 5$, $T = 100$, $USL - LSL = 8$, $C_{pw}(0) = 1.5$, and $C^*_{pw} = 1.499$. In Group 1, we take $k = 3$, $(\mu_1 = 98.75, \alpha_1 = 4, \beta_1 = 4)$, $(\mu_2 = 101, \alpha_2 = 4, \beta_2 = 0.1)$, and $(\mu_3 = 101.5, \alpha_3 = 4, \beta_3 = 6)$. For Group 2, $k = 5$, $(\mu_1 = 98.5, \alpha_1 = 4, \beta_1 = 6)$, $(\mu_2 = 98.75, \alpha_2 = 4, \beta_2 = 4)$, $(\mu_3 = 101, \alpha_3 = 4, \beta_3 = 0.1)$, $(\mu_4 = 101.5, \alpha_4 = 4, \beta_4 = 6)$, and $(\mu_5 = 102, \alpha_5 = 4, \beta_5 = 8)$.

The relative frequency that an action taken according to the proposed empirical decision rule coincides with that of the Bayes decision rule is denoted by f_n.

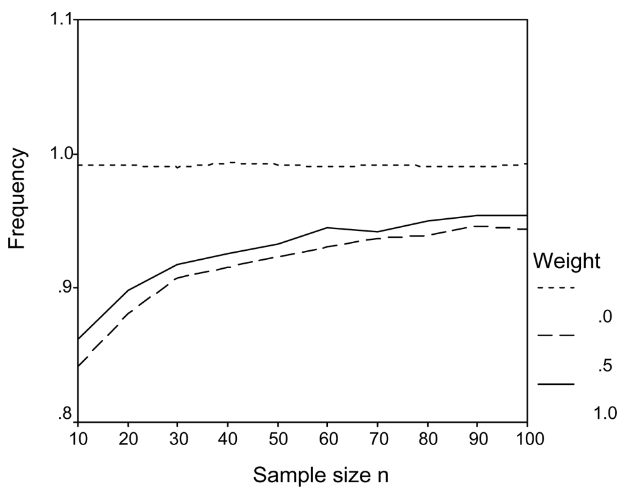

Figure 4. Relative frequency of correct decision with respect to sample size for three different weight values for Group 2.

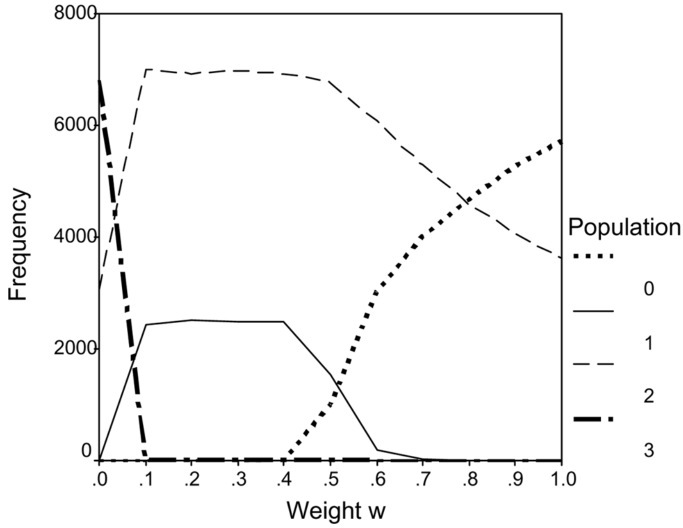

Figure 5. Total times that each process is selected as the best with respect to different weight for Group 1. *Process 0 means none is selected as the best.

In Table 1, we consider Group 1 with $w = 1$. It can be seen that the value of D_n decreases quite rapidly as n increases and on the other hand, that relative frequency f_n increases rapidly. In Table 2, we consider the case of Group 2. For given weight, the first entry in the column of the ith process denotes the number of times that process i has been selected according to the Bayes rule, and the second entry (in parenthesis) denotes the number of times that is selected according to the empirical Byes rule. The third entry (in bracket) shows total times that the ith process has been selected simultaneously by the Bayes rule and the empirical Bayes rules. CD denotes correct decision which means the decision made by the empirical Bayes rule coincides with that of the Bayes rule. For given weight, the entry in column of CD denotes the total times of correct decision in 10,000 simulations for Group 2. This table clearly shows that the factor of weight value has significant influence on the behaviors of correct decision for each process in Group 2, especially, when weight is zero, or near both extreme values of 0 and 1.

Acknowledgment

We are grateful to two referees for their patience of careful reading and many helpful comments which greatly lead to improvement of this presentation.

References

Chan, L. K., Cheng, S. W., Spiring, F. A. (1988). A new measure of process capability: C_{pm}. *J. Qual. Technol.* 20:160–175.

Gunter, B. H. (1989). The use and abuse of C_{pk}. *Qual. Progr.* 22(1):72–73.

Huang, W. T., Lai, Y. T. (1999). Empirical Bayes procedure for selecting the best population with multiple criteria. *Ann. Inst. Statist. Math.* 51(2):281–299.

Kane, V. E. (1986a). Process capability indices. *J. Qual. Technol.* 18:41–52.

Kane, V. E. (1986b). Corrigenda. *J. Qual. Technol.* 18:255.

Pearn, W. L., Kotz, S., Johnson, N. L. (1992). Distributional and inferential properties of process capability indices. *J. Qual. Technol.* 24:216–231.

Raiffa, H., Schlaifer, R. (1961). *Applied Statistical Decision Theory*. Boston: Harvard University.

Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkley Symp. Mathemat. Statist. Probab. I*. Berkley: University of California Press, pp. 157–163.

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

Spiring, F. A. (1997). A unifying approach to process capability indices. *J. Qual. Technol.* 29:49–58.

Tseng, S. T., Wu, T. Y. (1991). Selecting the best manufacturing process. *J. Qual. Technol.* 23:53–62.