Journal of the Korean
Statistical Society
Vol. 20, No. 2, 1991

Empirical Bayesian Multiple Comparisons with the Best†

Woo-Chul Kim* and Hyung Tae Hwang**

ABSTRACT

A parametric empirical Bayes procedure is proposed and studied to compare treatments simultaneously with the best. Minimum Bayes risk lower bounds are derived for an additive loss function, and their relationship with Bayesian simultaneous confidence lower bounds is given. For the proposed empirical Bayes procrdure, the nominal confidence level both in Bayesian sense and in frequentist's sense is shown to be controlled asymptotically. For practical implementation, a measure of significance similar to p-value is suggested with an illustrative example.

1. Introduction

Let $\bar{y}_1, \cdots, \bar{y}_k$ denote treatement means based on n replications each, such that $\bar{y}_i \sim N(\theta_i, \sigma^2/n)$, $i=1, \cdots, k$, independently, and let s^2 be an estimate of the error variance σ^2 such that $rs^2/\sigma^2 \sim \chi_r^2$, independently of $\bar{y}_1, \cdots, \bar{y}_k$. This setting often represents a reduction of many balanced designs for the comparison of treatments.

Following Gupta's(1956) pioneering work, many different formulations have been proposed to detect the *best* treatment associated with the largest θ_i, *i.e.*, $\max_{1 \le i \le k} \theta_i$. Hsu(1981, 1984 a) was the first to consider such a problem in the framework of simultaneous confidence intervals. Hsu(1984 a) obtained $100(1-\alpha)\%$ simultaneous confidence intervals for $\theta_i - \max_{j \ne i} \theta_j$, $i=1, \cdots, k$ as follows:

$$\theta_i - \max_{j \ne i} \theta_j \in [(\bar{y}_i - \max_{j \ne i} \bar{y}_j - d\sqrt{2/n}\, s)^-, \ (\bar{y}_i - \max_{j \ne i} \bar{y}_j + d\sqrt{2/n}\, s)^+], \qquad (1.1)$$

where $a^- = \min(a, 0)$, $a^+ = \max(a, 0)$ and d is a suitably chosen constant. The confidence intervals in (1.1) are in the spirit of frequentist's.

† Research supported by the Korean Science and Engineering Foundation Grant 88-07-23-02
* Department of Computer Science and Statistics, Seoul National University, Seoul, Korea
** Department of Computer Science and Statistics, Dankook University, Seoul, Korea

Some attemps have been made to provide Bayesian solutions to the problem of pairwise multiple comparisons, for example, by Dixon and Duncan (1975) among others. No attempt, however, has been made to obtain Bayesian analogue to (1.1) for the multiple comparisons with the *best*, up to the authors' knowledge. See Hochberg and Tamhane (1987) for detailed references.

The purpose of this article is to study the problem of multiple comparison with the best (MCB) from a Bayesian view point. First we derive minimum Bayes risk lower confidence bounds with regard to the best for an additive loss funcition, and then propose Baysian simultaneous lower confidence bounds analogously to (1.1). Next, we study the asymptotic confidence level of the empirical Bayes procedure obtained by substituting estimates for prior parameters. Finally, the extension to the unbalanced case is given along with an illustrative example.

2. Bayesian Procedure for Multiple Comparisons with the Best

The prior distribution of $\theta_1, \cdots, \theta_k$ for given σ^2 is taken to be

$$\theta_i \mid \sigma^2 \sim N(\psi, \sigma^2/c), i=1, \cdots, k, \text{ independently},$$

and the prior of σ^2 is assumed to be

$$p(\sigma^2)=1/\sigma^2.$$

Then the joint posterior distribution of $\theta_1, \cdots, \theta_k$, given $\bar{y}_1, \cdots, \bar{y}_k$ and s^2 is as follows:

$$(\theta_1, \cdots, \theta_k) \mid \bar{y}_1, \cdots, \bar{y}_k, s^2 \sim T_k(f, \mathbf{x}, s_0^2 I_k), \tag{2.1}$$

where $T_k(f, \mathbf{x}, s_0^2 I_k)$ denotes the k-variate t distribution with pdf proportional to

$$\left[1+\frac{1}{f}\frac{\Sigma_{i=1}^k (\theta_i-x_i)^2}{s_0^2}\right]^{-\frac{f+k}{2}},$$

$\mathbf{x}$ denotes $(x_1, \cdots, x_k)$ and

$$\begin{cases} f=r+k, \quad x_i=\dfrac{n\bar{y}_i+c\psi}{n+c}, i=1, \cdots, k \\ s_0^2=\dfrac{1}{f}\left[\dfrac{rs^2}{n+c}+\dfrac{nc\Sigma_{i=1}^k (\bar{y}_i-\psi)^2}{(n+c)^2}\right] \end{cases} \tag{2.2}$$

Under this setting, consider a decision problem in which one of θ_i is selected as a candidate for the best and the lower bounds $l_j (j=1, \cdots, k, j\neq i)$ are provided for $\theta_i-\theta_j (j\neq i)$. Then the decision set D can be described as follows:

$$D=\{(i ; l_j, j\neq i) : 1\leq i\leq k, \ -\infty<l_j<\infty, \ j\neq i\}.$$

For this decision problem assume the following loss structure for fixed $\beta \in (0, \frac{1}{2})$:

$$L[(i\,;\,l_j, j\neq i) : \theta]=\sum_{\substack{j=1\\ j\neq i}}^{k}\{-l_j+\beta^{-1}(l_j-(\theta_i-\theta_j))^{+}\}, \tag{2.3}$$

where θ denotes $(\theta_1, \cdots, \theta_k)$. By this loss function, we mean that the loss gets larger as the lower bound l_j gets smaller as long as the lower bound l_j is truely less than $\theta_i-\theta_j$. This loss functaion also reflects the penalty $l_j-(\theta_i-\theta_j)$ when l_j is not truely a lower bound on $\theta_i-\theta_j$.

Theorem 2.1 *The minimum Bayes risk procedure with respect to the loss (2.3) is given by*

$$((k)\,;\,(x_{(k)}-x_j)-\sqrt{2}\ t_f^{(\beta)}s_0,\ j\neq(k)) \tag{2.4}$$

where $x_{(k)}$ denotes the largest posterior mean in (2.2), and $t_f^{(\beta)}$ denotes the upper β quantile of the t distribution with f degrees of freedom.

proof Fix i. Then from the additivity of the loss, it follows that the Bayes risk for the decision problem is minimized by minimizing the subcomponent risks. It follows from (2.1) that the subcomponent posterior loss with respect to the prior of $\Delta=\theta_i-\theta_j$ is given by

$$-l_j+\frac{1}{\beta}\int_{-\infty}^{l_j}(l_j-\Delta)\frac{1}{\sqrt{2}\ s_0}\ t_f\Big(\frac{\Delta-(x_i-x_j)}{\sqrt{2}\ s_0}\Big)d\Delta, \tag{2.5}$$

where t_f denotes the pdf of the t distribution with f degrees of freedom. By differentiating (2.5) with respect to l_j, it can be shown that (2.5) is minimized when

$$l_j^o=x_i-x_j-\sqrt{2}t_f^{(\beta)}s_0. \tag{2.6}$$

Substituting l_j^o for l_j in (2.5) and summing over $j(\neq i)$, we can see that

$$E\{L[(i\,;\,l_j^o, j\neq i) : \theta]\mid \mathbf{y},\ s^2\}$$

$$=-\sum_{j\neq i}^{k}(x_i-x_j)-\frac{\sqrt{2}(k-1)}{\beta}s_0\int_{-\infty}^{-t_f^{(\beta)}} u t_f(u)du,$$

which is obviously minimized when $i=(k)$. Hence the proof is completed■.

The minimum Bayes risk procedure in (2.4) is not quite satisfactory in the sense that it controls *the comparisonwise error* only, not *the experimentwise error*. Namely, due to the assumed additive loss structure of (2.3), it does not possess the property of truely simultaneous comparisons with the best.

Therefore we seek for Bayesian simultaneous confidence lower bounds for $\theta_{(k)}-\theta_{(1)}, \cdots, \theta_{(k)}-\theta_{(k-1)}$ where $x_{(1)}<x_{(2)}< \cdots < x_{(k)}$ denote the ordered posterior means $x_1, \cdots, x_k$ in (2.2). Note that the selection of $\theta_{(k)}$ as the candidate for the best can be justified by Theorem 2.1. It follows from (2.1) that

$$\Big(\frac{(\theta_{(k)}-\theta_{(1)})-(x_{(k)}-x_{(1)})}{\sqrt{2}\ s_0}, \cdots, \frac{(\theta_{(k)}-\theta_{(k-1)})-(x_{(k)}-x_{(k-1)})}{\sqrt{2}\ s_0}\Big)\mid \mathbf{y},\ s^2 \sim T_{k-1}(f,\ \mathbf{o},\ \mathrm{R}), \tag{2.7}$$

where the associated correlation matrix $R=\{\rho_{ij}\}$ is given by $\rho_{ij}=0.5$ for $i\neq j$ and $\rho_{ii}=1$. Letting $T^{(\alpha)}_{k-1,f,0.5}$ be the upper α equicoordinate quantile of the multivariate t distribution $T_{k-1}(f, \mathbf{o}, R)$, we have

$$P_{\psi,c}[\theta_{(k)}-\theta_{(i)}\geq x_{(k)}-x_{(i)}-\sqrt{2}T^{(\alpha)}_{k-1,f,0.5}\ s_0\ i=1, \cdots, k-1 \mid \mathbf{y}, s^2]=1-\alpha, \tag{2.8}$$

which provides $100(1-\alpha)\%$ simultaneous lower confidence bounds for $\theta_{(k)}-\theta_{(i)}, i=1, \cdots, k-1$, to be called the *Bayesian MCB procedure*.

It should be remarked that (2.8) guarantees posterior coverage probability while Hsu's procedure in (1.1) is designed to guarantee the frequentist's coverage probability. Also note that the Bayesian *MCB* procedure in (2.8) becomes minimum Bayes risk procedure in (2.4) if α and β are related by $t_f^{(\alpha)}=T^{(\alpha)}_{k-1,f,0.5}$.

As a final remark, the values of $T^{(\alpha)}_{p,v,\rho}$ (in our case, $p=k-1$, $v=f$, $\rho=0.5$) can be found in the tables of Gupta, Panchapakesan and Sohn (1985) for selected values of α, p, ρ and $v\geq 15$.

3. Empirical Bayes Procedure

When it is difficult to specify the prior parameters ψ and c, empirical estimation of them can be done using the marginals. Note that for any fixed σ^2,

$$\bar{y}_i \mid \sigma^2 \sim N(\psi, \sigma^2(\frac{1}{n}+\frac{1}{c})), i=1, \cdots, k, \textit{ independently}.$$

Hence, letting

$$\bar{y}=\frac{1}{k}\sum_{i=1}^{k}\bar{y}_i \text{ and } MS_T=\frac{n}{k-1}\sum_{i=1}^{k}(\bar{y}_i-\bar{y})^2,$$

$\bar{y}$ and $\frac{1}{n}MS_T$ may be considered as the ordinary unbiased estimates of ψ and $\sigma^2(\frac{1}{n}+\frac{1}{c})$ respectively for fixed σ^2. Since s^2 is unbiased for σ^2, c can be estimated by $n/(F-1)$, where $F=MS_T/s^2$, the ordinary F-ratio in analysis of variance.

After making some corrections for negative values of $n/(F-1)$, we propose the estimates of ψ and c as follows:

$$\begin{cases} \hat{\psi}=\bar{y} \\ \hat{c}=\begin{cases} n/(F-1), & \text{if } F>1 \\ \infty, & \text{if } F\leq 1. \end{cases} \end{cases}$$

By substituting these estimates in (2.2), we have

$$\begin{cases} \hat{x}_i=(1-\frac{1}{F})^+(\bar{y}_i-\bar{y})+\bar{y} \\ \hat{s}_0^2=(1-\frac{1}{F})^+(1-\frac{1}{r+k})s^2/n \end{cases} \tag{3.1}$$

Replacing x_i and s_0^2 by $\hat{x}_i$ and $\hat{s}_0^2$ in (2.8), we have the following procedure which will be called the *parametric empirical Bayes (PEB) MCB procedure*: For $i=1, \cdots, k-1$,

$$\theta_{(k)}-\theta_{(i)}\geq(1-\frac{1}{F})^+(\bar{y}_{(k)}-\bar{y}_{(i)})-\sqrt{2}\ T^{(\alpha)}_{k-1,r+k,0.5}\left((1-\frac{1}{F})^+(1-\frac{1}{r+k})\right)^{1/2}\frac{s}{\sqrt{n}} \quad (3.2)$$

To study the coverage probability of the *PEB MCB* procedure in (3.2), let the event *coverage* denote the event (3.2) and $\theta_{[1]}\leq\theta_{[2]}\leq\cdots\leq\theta_{[k]}$ denote the ordered $\theta_1, \theta_2, \cdots, \theta_k$. The next result shows that the nominal coverage probability, in frequentist's sense, is guaranteed as $n\to\infty$ be the *PEB MCB* procedure.

Theorem 3.1 *Suppose that* $\theta_{[k]}>\theta_{[k-1]}$, *and that* $r\to\infty$ *as* $n\to\infty$. *Then*

$$\lim_{n\to\infty} P[\textit{coverage} \mid \theta_1, \cdots, \theta_k, \sigma^2]=1-\alpha.$$

proof First, note that as $n\to\infty$,

$$P((k)=[k] \mid \theta_1, \cdots, \theta_k, \sigma^2)\to 1$$

and

$$\frac{1}{n}F\xrightarrow{p}\frac{\sum_{i=1}^k(\theta_i-\theta)^2}{(k-1)\sigma^2}$$

Therefore the coverage probability can be computed on the event $((k)=[k]),\ F>1)$ for sufficiently large n.

Futhermore, on the event $((k)=[k]),\ F>1)$, the event *coverage* can be expressed as

$$B_n\max_{i\neq k}\left\{\frac{(\bar{y}_{[k]}-\theta_{[k]})-(\bar{y}_i-\theta_{[i]})}{s/\sqrt{n}}-\frac{\sqrt{n}}{(F-1)}\frac{(\theta_{[k]}-\theta_{[i]})}{s}\right\}\leq\sqrt{2}\ T^{(\alpha)}_{k-1,\infty,0.5}$$

where $B_n=(T^{(\alpha)}_{k-1,\infty,0.5}\ /\ T^{(\alpha)}_{k-1,r+k,0.5})\left((1-\frac{1}{F})/(1-\frac{1}{r+k})\right)^{1/2}$.

Thus the result follows from the fact that, as $n\to\infty$,

$$B_n\xrightarrow{p}1,\quad \frac{\sqrt{n}}{(F-1)}\frac{\theta_{[k]}-\theta_{[i]}}{s}\xrightarrow{p}0$$

and

$$\frac{(\bar{y}_{[k]}-\theta_{[k]})-(\bar{y}_{[i]}-\theta_{[i]})}{s/\sqrt{n}}\xrightarrow{d}Z_k-Z_i \qquad (i=1, \cdots, k-1)$$

where $Z_1, \cdots, Z_k$ are independently distributed as $N(0, 1)$■.

Now we consider the asymptotic posterior coverage probability as the number of treatments k gets larger. For this we need the following lemma.

Lemma 3.1 *Let* $(t_1, \cdots, t_{k-1})' \sim T_{k-1}(r+k, o, R)$ with R *in* (2.7), *and* $T=\min_{1\le i\le k-1} t_i$. *Then, as* $k\to\infty$,

$$\sqrt{2}(T+(\log(k-1))^{1/2}) \xrightarrow{d} N(0, 1),$$

and

$$T^{(\alpha)}_{k-1,r+k,0.5} - (\log(k-1))^{1/2} = \frac{1}{\sqrt{2}} z_\alpha + o(1),$$

where z_α *denotes the upper* α quantile of N(0, 1).

proof The distribution of T is the same as that of $(Z_1 - \max_{2\le i\le k} Z_i) / (\sqrt{2}S)$ where Z_i's are *iid* N(0, 1) and S^2 is the average of $r+k$ *iid* $\chi^2(1)$ random variables independent of Z_i's. Then the result follows from the well known fact that $\max_{2\le i\le k} Z_i = (2\log(k-1))^{1/2} + o(1)$ *a.s. as* $k\to\infty$ ■.

Theorem 3.2 *For any finite* ψ *and* c *and under the assumption of* $\frac{1}{r} = O(\frac{1}{k})$, *we have*

$$\lim_{k\to\infty} P_{\psi,c}[coverage \mid \mathbf{y}, s^2] = 1-\alpha \qquad a.e..$$

proof First, note that

$$P_{\psi,c}[coverage \mid \mathbf{y}, s^2]$$
$$= P_{\psi,c}\left[\frac{\min_{1\le i\le k-1}\{(\theta_{(k)} - \hat{x}_{(k)}) - (\theta_{(i)} - \hat{x}_{(i)})\}}{\sqrt{2}\ \hat{s}_0} \ge -T^{(\alpha)}_{k-1,r+k,0.5} \mid \mathbf{y}, s^2\right]. \qquad (3.3)$$

It is easy to see that the probability in (3.3) is bounded by the following two probabilities:

$$P_{\psi,c}\left[\frac{\min_{1\le i\le k-1}\{(\theta_{(k)} - x_{(k)}) - (\theta_{(i)} - x_{(i)})\}}{\sqrt{2}\ s_0} + (\log(k-1))^{1/2}\right.$$
$$\left. \ge -\frac{\hat{s}_0}{s_0} T^{(\alpha)}_{k-1,r+k,0.5} \pm \frac{\max_{i\ne j} |(\hat{x}_i - \hat{x}_j) - (x_i - x_j)|}{s_0} + (\log(k-1))^{1/2} \mid \mathbf{y}, s^2\right]. \qquad (3.4)$$

Furthermore, it follows from Lemma 3.1 and a lemma in Kim and Hwang (1991) that the probabilites in (3.4) converges a.e. to $P[Z_1 > -z_\alpha] = 1-\alpha$ ■.

The results in this section indicate that *PEB* procedure can guarantee the coverage probability, in both frequentist's and Bayesian sense, at least asymptotically.

4. Extension to Unbalanced Designs

In the case of unbalanced designs, we have $\bar{y}_i \sim N(\theta_i, \sigma^2/n_i)$, $i=1, \cdots, k$, independently, and $rs^2/\sigma^2 \sim \chi^2_r$ independently of $\bar{y}_1, \cdots, \bar{y}_k$. The prior distributions of $\theta_1, \cdots, \theta_k$ and σ^2 are just the same as in Section 2.

In a similar way to Section 2, *the extended Bayes MCB procedure* is proposed as follows: For $i=1, \cdots, k-1$,

$$\theta_{(k)} - \theta_{(i)} \ge x_{(k)} - x_{(i)} - T^{(\alpha)}_{k-1,r+k,\{\rho_{ij}\}} \sqrt{s^2_{(k)} + s^2_{(i)}}, \qquad (4.1)$$

where $x_{(1)}<x_{(2)}< \cdots <x_{(k)}$ denotes the ordered $x_1, x_2, \cdots, x_k$, and

$$\begin{cases} x_i=\dfrac{n\bar{y}_i+c\psi}{n_i+c}, \\ s_i^2=\dfrac{1}{(r+k)(n_i+c)}\left[\sum_{j=1}^{k}\dfrac{cn_j(\bar{y}_j-\psi)^2}{n_j+c}+rs^2\right] \end{cases} \tag{4.2}$$

In (4.1), $T^{(\alpha)}_{k-1,r+k,\{\rho ij\}}$ denotes the upper α equicooridinate quantile of the multivariate t distribution $T_{k-1}(r+k, \mathbf{o}, \{\rho_{ij}\})$ with

$$\rho_{ij}=\{(1+\frac{s_{(i)}{}^2}{s_{(k)}{}^2})(1+\frac{s_{(j)}{}^2}{s_{(k)}{}^2})\}^{-1/2}, \quad (i\neq j). \tag{4.3}$$

For the empirical Bayesian approach, we estimate ψ and c as follows:

$$\begin{cases} \hat{\psi}=\bar{y} \\ \hat{c}=\begin{cases} k_0/(F-1) & \textit{for } F>1 \\ \infty & \textit{for } F\leq 1, \end{cases} \end{cases}$$

where

$$\bar{y}=(\sum_{i=1}^{k}n\bar{y}_i)/(\sum_{i=1}^{k}n_i), k_0=\{(\sum_{i=1}^{k}n_i)^2-\sum_{i=1}^{k}n_i^2\}/\{(k-1)\sum_{i=1}^{k}n_i\},$$

and F denotes the ordinary F-ratio in analysis of variance. Substituting these into (4.2), we have

$$\begin{cases} \hat{x}_i=\dfrac{n\bar{y}_i+\hat{c}\hat{\psi}}{n_i+\hat{c}}, \quad i=1, \cdots, k \\ \hat{s}_i^2=\dfrac{1}{(n_i+\hat{c})(r+k)}\left[\sum_{j=1}^{k}\dfrac{\hat{c}n_j(\bar{y}_j-\hat{\psi})^2}{n_j+\hat{c}}+rs^2\right], \quad i=1, \cdots, k. \end{cases}$$

Substituting these for x_i and s_i^2 in (4.1), we propose *the extended PEB MCB procedure* as follows: For $i=1, \cdots, k-1$,

$$\theta_{(k)}-\theta_{(i)}\geq\hat{x}_{(k)}-\hat{x}_{(i)}-T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}ij\}}\sqrt{\hat{s}^2_{(k)}+\hat{s}^2_{(i)}}. \tag{4.4}$$

Note that the value of $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}ij\}}$ is the solution of the following equation for t:

$$\int_0^\infty\int_{-\infty}^{\infty}\prod_{i=1}^{k-1}\Phi\left[\frac{\lambda_i z+tu}{\sqrt{1-\lambda_i^2}}\right]d\Phi(z)dF_{r+k}(u)=1-\alpha, \tag{4.5}$$

where Φ denotes the standard normal cdf and F_{r+k} denotes the cdf of a $\sqrt{\chi^2_{r+k}/(r+k)}$ *random variable, and*

$$\lambda_i=(1+\frac{n_{(k)}+\hat{c}}{n_{(i)}+\hat{c}})^{-1/2}, \quad i=1, \cdots, k-1. \tag{4.6}$$

It is not difficult to solve (4.5) numerically. In fact, a computer program for solving (4.5) is available upon request from the authors.

However, it should be noted that the values of $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}$ must be evaluated separately for each combination of the values of $\hat{s}^2_{(1)}, \cdots, \hat{s}^2_{(k)}$. Thus it would be desirable to find suitable approximations to $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}$ that can be readily obtained from the available tables.

Natural approximations to $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}$ are provided by replacing the $\hat{\rho}_{ij}$'s by some common value ρ since the tables are widely available for this equicorrelated case. One of the easiest way is to use $\rho=\min\{\hat{\rho}_{ij}\}>-(k-2)^{-1}$, whose conservatism can be achieved by the well known inequality of Slepian (1962). On the other hand Dunnett (1985) has suggested the use of the arithmetic mean $\bar{\rho}$ of the $\hat{\rho}_{ij}$'s for less conservative approximation. His extensive numerical study showed that $\bar{\rho}$ always provides upper bounds on the values of $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}$. Thus the above two methods may be used to approximate the values of $T^{(\alpha)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}$ conservatively.

Note that if we set $n_i=n$ for all i, then all of the above three methods reduce to the PEB MCB procedure in the balanced case.

Following Hsu's(1984 b) idea, we define S_B−*value* as the smallest α for which the treatment (k) can be declared, *a posteriori*, as the best treatment by the extended PEB MCB procedure: In other words

$$\min_{1\le i\le k-1}\{\hat{x}_{(k)}-\hat{x}_{(i)}-T^{(S_B)}_{k-1,r+k,\{\hat{\rho}_{ij}\}}\sqrt{\hat{s}_{(k)}{}^2+\hat{s}_{(i)}{}^2}\}=0.$$

Then S_B−value can be computed as follows:

$$S_B=1-\int_0^\infty\int_{-\infty}^\infty\prod_{i=1}^{k-1}\Phi[\frac{\lambda_i z+tu}{\sqrt{1-\lambda_i^2}}]d\Phi(z)dF_{r+k}(u),$$

where $t=\min_{i\neq k}\{(\hat{x}_{(k)}-\hat{x}_{(i)})/(\hat{s}^2_{(k)}+\hat{s}^2_{(i)})^{1/2}\}$. Note that we need the same program for computing S_B as that to solve (4.5). The next example illustrates how to apply the extended PEB MCB procedure, and to interpret the S_B−value.

Example 4.1 Hsu(1984 b) cited the data in Table 4.1, which are the results of an experiment to compare seven brands of filters. These brands are compared in their ability to reveal the microorganism fecal coliform in river water.

Table 1. Counts of colonies on different brands of filters

Brand	counts	$\bar{y}_i$
1	69 122 95	95.3
2	118 154 102	124.7
3	171 132 182	161.7
4	122 119 −	120.5
5	204 225 190	206.3
6	140 130 127	132.3
7	170 165 −	167.5

For this data set the pooled sample variance is $s^2=411.194$ with $r=12$ degrees of freedom, and the F−ratio is F=9.499 which is significantly larger than the corresponding critical value even at 1% level.

For $\alpha=0.1$, the solution of (4.5) was numerically found to be

$$T^{(0.1)}_{6.19,\{\hat{\rho}ij\}}=2.1025,$$

which yields the confidence lower bounds in Table 4.2.

It should be noted that the two conservative approximations yields $T^{(0.1)}_{6.19,min\{\hat{\rho}ij\}}=2.1307$ and $T^{(0.1)}_{6.19,\hat{\rho}}=2.1029$, which are close to the numerical solution 2.1025.

Table 2. 90% confidence lower bounds by PEB MCB procedure

differences	lower bounds
$(\theta_5-\theta_1)$	68.432
$(\theta_5-\theta_2)$	41.524
$(\theta_5-\theta_3)$	8.454
$(\theta_5-\theta_4)$	41.427
$(\theta_5-\theta_6)$	34.977
$(\theta_5-\theta_7)$	0.873

We have computed the S_B−value numerically ;

$$S_B-value=0.0913$$

which confirms that brand 5 can be declared as the best with posterior error probability less than 10%.

Note that, even though the approach is different, Hsu's(1984 b) MCB procedure can not declare the brand 5 as the best at frequentist's error probability $\alpha=0.1$.

Acknowledgements

Thanks are due to referees for making valuable suggestions.

References

1. Dixon, D.O. and Duncan, D.B. (1975). Minimum Bayes risk t-intervals for multiple comparisons. *Journal of the American Statistical association*, 70, 822-831.
2. Dunnett, C.W. (1985). Multiple comparisons between several treatments and a specified treatment. Invited talk at the Spring ENAR Meeting, Raleigh, North Carolina.
3. Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Mimeo Series No. 150, Institute of Statistics, University of North Carolina, Chapel Hill.
4. Gupta, S.S., Panchapakesan, S. and Sohn, J.K. (1985). On the distribution of the Studentized maximum of equally correlated normal random variables. *Communications in Statistics*, Series B, 14, 103-135.
5. Hochberg, Y. and Tamhane, A.C. (1987). *Multiple comparison procedures*. John Wiley & Sons, New York.

6. Hsu, J.C. (1981). Simultaneous confidence intervals for all distances from the 'best'. *Annals of Statistics*, 9, 1026-1034.
7. Hsu, J.C. (1984 a). Constrained simultaneous confidence intervals for multiple comparisons with the best. *Annals of Statistics*, 12, 1136-1144.
8. Hsu, J.C. (1984 b). Ranking and selection and multiple comparisons with the best. *Design of Experiments : Ranking and Selection*, New York : Marcel Dekker, 23-33.
9. Kim, W.C. and Hwang, H.T. (1991). A parametric empirical Bayesian method for multiple comparisons. *Journal of Korean Statistical Society*, 20, 44-56.
10. Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Technical Journal*, 41, 463-501.