Journal of the KSQC
Vol. 16, No.2 1988

Confidence Bounds for Superiority [+]

Jong-Woo Jeon*
Woo-Chul Kim*
Gyu-Jin Jeong**

ABSTRACT

The problem of making confidence statements is considered about the means of treatments with t largest sample values among k available treatments. These confidence bounds are used in selecting a fixed number of superior treatments. An illustrative example is also provided.

1. Introduction

Suppose experimenters observe k independent statistics Y_i with pdf $f(y-\theta_i)$ $(i=1, \cdots, k)$, where θ_i is the average treatment effect of the treatment i. They often question whether they can rank the treatments according to the ordered values $Y_{(1)} < \cdots < Y_{(k)}$ of $Y_1, \cdots, Y_k$.

One way of answering to this question is to construct lower confidence bounds for $\theta_{(k-t+1)} - \max_{1 \leq i \leq k-t} \theta_{(i)} (t=1, \cdots, k-1)$. Bofinger(1983) considered a lower confidence bound for $\theta_{(k)} - \max_{1 \leq i \leq k-1} \theta_{(i)}$, and Gutmann and Maymin(1985) provided another lower confidence bound for $\theta_{(k)} - \max_{1 \leq i \leq k-1} \theta_{(i)}$ which was shown to be better than Bofinger's.

Hsu's(1984) method can provide *simultaneous* lower confidence bounds for $\theta_{(k-t+1)} - \max_{1 \leq i \leq k-t} \theta_{(i)}$ $(t=1, \cdots, k-1)$. Using this idea, Mengerson and Bofinger(1986) obtained a lower confidence bound for $\min_{k-t+1 \leq j \leq k} \theta_{(j)} - \max_{1 \leq i \leq k-t} \theta_{(i)}$.

+ 본 연구는 1987 대학발전기금 대우학술연구비의 지원으로 수행되었음.
* 서울대학교 계산통계학과
** 한남대학교 응용통계학과

The purpose of this article is to extend Gutmann and Maymin's result so that *simultaneous* lower confidence bounds for $\theta_{(k-t+1)}-\max_{1\le i\le k-t}\theta_{(i)}\,(t=1,\cdots,k-1)$ can be constructed. In Section 2, we extend the result of Gutmann and Maymin(1985) to obtain a lower confidence bound for $\theta_{(k-t+1)}-\max_{1\le i\le k-t}\theta_{(i)}$ for fixed t. Section 3 treats the case of normal means problem. Based on the result in Section 2, a method of constructing simultaneous lower confidence bounds is explained in Section 3 with an illustrative example. Comparisons with previous results are also provided.

Finally, it should be remarked that there are several other approaches to ranking the treatments, and a good reference can be made to Gupta and Panchapakesan(1979).

2. Lower Confidence Bound For Superiority

We assume that $Y_1,\cdots,Y_k$ are independent with pdf's $f(y-\theta_i)$ $(i=1,\cdots,k)$, and that $f(y-\theta)$ has the monotone likelihood ratio (MLR) property in y and θ.

In this section, we will construct a lower confidence bound of the following form ;

$$\theta_{(k-t+1)}-\max_{1\le i\le k-t}\theta_{(i)}\geqq L(Y_{(k-t+1)}-Y_{(k-t)}) \qquad (2.1)$$

where L is a suitably chosen function according to a given level $1-\alpha(0<\alpha<1)$, and $Y_{(1)}<\cdots<Y_{(k)}$ are the ordered $Y_1,\cdots,Y_k$.

In the construction of a lower confidence bound, the following lemma by Wijsman(1985) is useful.

Lemma 2.1. Let g_i and $h_i\,(i=1,2)$ be non-negative functions defined on some real intervals and the four integrals $\int g_i(x)h_j(x)dx$ be positive for $i=1,2$ and $j=1,2$. Then inequality

$$\int g_1(x)h_1(x)dx/\int g_1(x)h_2(x)dx\geqq\int g_2(x)h_1(x)dx/\int g_2(x)h_2(x)dx \qquad (2.2)$$

holds if g_1/g_2 and h_1/h_2 are monotonic in the same direction.

Now, to define the function L in (2.1), let H denote the cdf of $(Y_1-\theta_1)-(Y_2-\theta_2)$. Then for a given $1-\alpha(0<\alpha<1)$, the function L is defined by

$$H(L(w)-w)/H(L(w))=\alpha \qquad (2.3)$$

for each $w>0$. The existence of such a function L satisfying (2.3) is proved in the next lemma.

Lemma 2.2. Under the assumption of the MLR property of $f(y-\theta)$, there exists a non-decreasing function L satisfying(2.3) for a given $1-\alpha(0<\alpha<1)$.

Proof. For each $w>0$, let us consider a function G_w given by

$$G_w(a)=H(a-w)/H(a)$$

where H is the cdf of $(Y_1-\theta_1)-(Y_2-\theta_2)$. Since the MLR property of $f(y-\theta)$ implies the same property for the cdf $F(y-\theta)$, it follows from Lemma 2.1 that for $a_1 \leqq a_2$,

$$G_w(a_1)=\int_{-\infty}^{\infty} F(y-w) f(y-a_1)\,dy \Big/ \int_{-\infty}^{\infty} F(y) f(y-a_1)\,dy$$

$$\leqq \int_{-\infty}^{\infty} F(y-w) f(y-a_2)\,dy \Big/ \int_{-\infty}^{\infty} F(y) f(y-a_2)\,dy$$

$$=G_w(a_2).$$

Thus, for a given α, the existence of the function $L(w)$ follows from the fact that

$$\lim_{a\to-\infty} G_w(a)=0 \text{ and } \lim_{a\to+\infty} G_w(a)=1.$$

The monotonicity of L is rather clear from the observation that $G_w(a)$ is non-increasing in w for each fixed a.

The next result is an extension of Gutmann and Maymin's (1985) result.

Theorem 2.1 Assume that the pdf $f(y-\theta)$ has the MLR property in y and θ. Then we have

$$\inf_{\underline{\theta}} P_{\underline{\theta}}[\theta_{(k-t+1)}-\max_{1\leqq i\leqq k-t} \theta_{(i)} > L(Y_{(k-t+1)}-Y_{(k-t)})] \geqq 1-\alpha.$$

where L is defined by (2.3) for a given level $(1-\alpha)$.

Proof. Let g denote the inverse function of L so that $g(\Delta)>0$ for any Δ. To maximize the uncoverage probability, we consider the conditional uncoverage probability

$$P_{\underline{\theta}}[g(\theta_{(i-t+1)}-\max_{1\leqq i\leqq k-t} \theta_{(i)}) \leqq Y_{(k-t+1)}-Y_{(k-t)}|A]$$

where $A=\{Y_i \geqq Y_{k-t+1} \geqq Y_j,\ \ i=k-t+2,\cdots,k,\ \ j=1,\cdots,\ k-t\}$. Note that such a conditional uncoverage probability is symmetric in $\theta_1,\cdots,\theta_{k-t}$, and in $\theta_{k-t+2},\cdots,\theta_k$, respectively. Thus, we may assume that $\theta_{k-t} \geqq \cdots \geqq \theta_1$.

Repeated applications of Lemma 2.1 yield that, with $\bar{F}$ denoting $1-F$,

$$P_{\underline{\theta}}[g(\theta_{(k-t+1)}-\max_{1\leqq i\leqq k-t} \theta_{(i)}) \leqq Y_{(k-t+1)}-Y_{(k-t)}|A]$$

$$=\frac{\int_{-\infty}^{\infty} \prod_{j=1}^{k-t} F(y-\theta_j-g(\theta_{k-t+1}-\theta_{k-t})) \prod_{i=k-t+2}^{k} \bar{F}(y-\theta_i) f(y-\theta_{k-t+1})\,dy}{\int_{-\infty}^{\infty} \prod_{j=1}^{k-t} F(y-\theta_j) \prod_{i=k-t+2}^{k} \bar{F}(y-\theta_i) f(y-\theta_{k-t+1})\,dy}$$

$$\leq \frac{\int_{-\infty}^{\infty} \prod_{j=1}^{k-t} F(y-\theta_j - g(\theta_{k-t+1}-\theta_{k-t}))f(y-\theta_{k-t+1})\,dy}{\int_{-\infty}^{\infty} \prod_{j=1}^{k-t} F(y-\theta_j)f(y-\theta_{k-t+1})\,dy}$$

$$\leq \frac{\int_{-\infty}^{\infty} F(y+\theta_{k-t+1}-\theta_{k-t}-g(\theta_{k-t+1}-\theta_{k-t}))f(y)\,dy}{\int_{-\infty}^{\infty} F(y+\theta_{k-t+1}-\theta_{k-t})f(y)\,dy}$$

It follows that the conditional uncoverage probability is bounded above by

$$H(L(w)-w)/H(L(w))=\alpha$$

where $w=g(\theta_{k-t+1}-\theta_{k-t})$. Therefore the result follows by observing that the conditioning event can be replaced by any other permutation of $Y_1, \cdots, Y_k$.

It should be remarked that

$$L(w)=0 \text{ for } w=x_{\alpha/2},$$

where $x_{\alpha/2}$ is the upper $\alpha/2$ quantile of the cdf H. Hence we can conclude, with confidence $(1-\alpha)$, that

$$\theta_{(k-t+1)} > \max_{1\leq i\leq k-t} \theta_{(i)} \text{ whenever } Y_{(k-t+1)} - Y_{(k-t)} \geqq x_{\alpha/2}.$$

3. Application to Normal Means Problem

In this section, we assume the usual one-way balanced model as follows ;

$$X_{ij}=\mu_i+\varepsilon_{ij},\quad i=1,\cdots,k\ ;\ j=1,\cdots,n$$

where μ_i's are the treatment effects and ε_{ij}'s are independent and normally distributed with mean 0 and unknown variance σ^2.

It can be easily seen that the resuflt of Section 2 can be appleid by taking $Y_i=\bar{X}_i$, $\theta_i=\mu_i/\sigma$ and by considering the conditional coverage probability given the pooled estimator $\hat{\sigma}^2$ of σ^2.

Thus the following $100(1-\alpha)\%$ confidence statement can be made ;

$$(\mu_{(k-t+1)} - \max_{1\leq i\leq k-t} \mu_{(i)})/\sigma > \sqrt{2/n}\, h_\nu(\sqrt{n}(\bar{X}_{(k-t+1)} - \bar{X}_{(k-t)})/\sqrt{2}\hat{\sigma}) \qquad (3.1)$$

where $h_\nu(w)$ for $w>0$ is given by

$$\int_0^{\infty} \Phi(h_\nu(w)-wu)/\Phi(h_\nu(w))\,dQ_\nu\ (u)=\alpha. \qquad (3.2)$$

Here Q_ν is the cdf of $\hat{\sigma}/\sigma$ and $\nu = k(n-1)$. Note that $h_\nu(w) = 0$ for $w = t_{\alpha/2}(\nu)$, and that the case of known common variance can be obtained from (3.1) and (3.2) by taking $\nu = +\infty$.

To implement the confidence statement in (3.1), the values of the function $h_\nu(w)$ have been computed for selected values of ν and w and for $\alpha = 0.05$ and 0.01, which are available upon request. For selected cases, the shapes of $h_\nu(w)$ are given in Figure I.

The confidence lower bound in (3.1) reduces to the one by Gutmann and Maymin (1985) for $t = 1$. In fact the function $h_\nu(w)$ defined by (3.2) is the same as that in Gutmann and Maymin, and it was not tabulated by them.

The values of $h_\nu(w)$ in (3.2) vs w for $\alpha = 0.05$

A : $\nu = 20$ B : $\nu = 60$

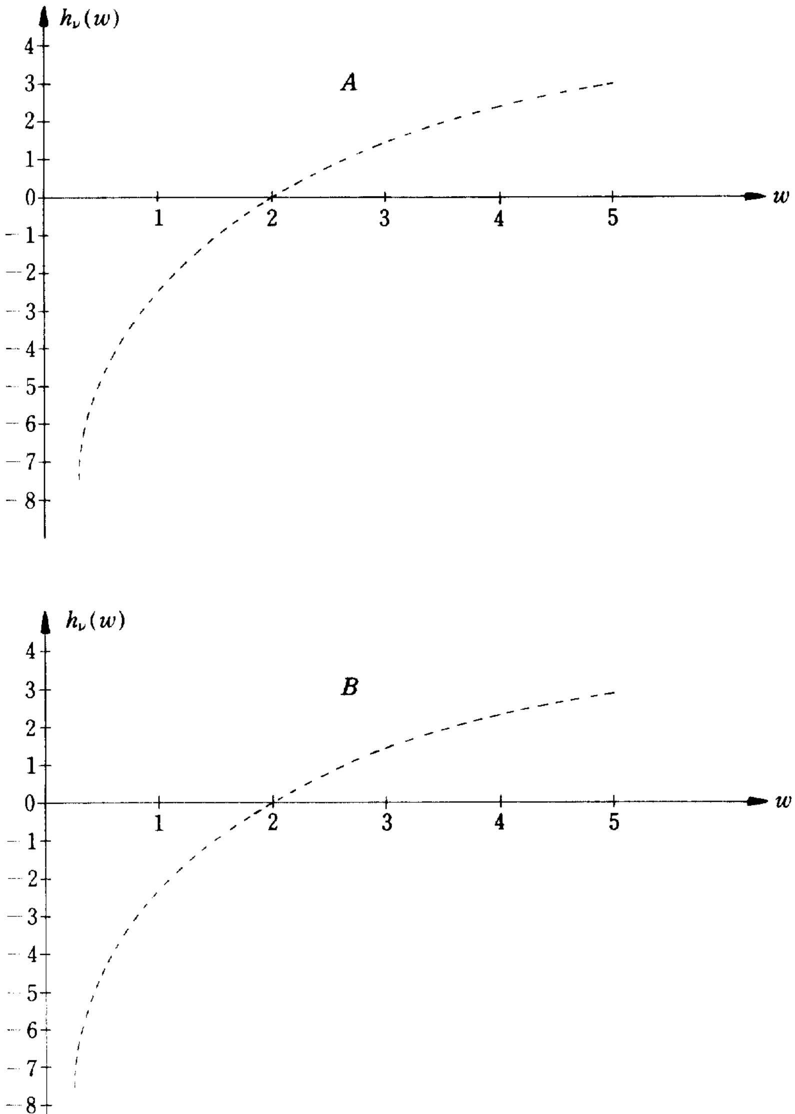

Figure I

As an illustrative example, we consider the simulated data by Kleijnen, Naylor and Seaks (1972). They considered a selection problem, in which a firm is interested in selecting production plans with more profits among $k=5$ possible plans. They ran simulation experiments with a sample of size $n=50$ for each plan and assumed that the profit using each plan has a normal distribution with a common unknown variance. The summary of data is given in Table 3.1.

Table 3.1 : Summary of Profit Data

plan	Mean Profit	Standard Deviation
1=(2)	2976.44	175.83
2=(3)	2992.30	202.20
3=(1)	2675.20	250.51
4=(5)	3265.30	221.81
5=(4)	3131.90	277.04

From Table 3.1, we observe that the pooled sample standard deviation is $\hat{\sigma}=228.26$ with $\nu=5\times 49=245$ degrees of freedom and that plan 4 and 5 correspond to the largest and the second largest sample mean profits, respectively.

First, we consider a situation where only one plan with the largest profit is desired. We observe from Table 3.1 that

$$\sqrt{n}\,(\bar{X}_{(5)}-\bar{X}_{(4)})/\sqrt{2}\hat{\sigma}=2.92.$$

For $\alpha=0.05$, it was found to be $\sqrt{2/n}\ h_{\nu}(2.92)=0.24$. Thus the 95% lower confidence bound in (3.1) is given as follows :

$$(\mu_{(5)}-\max_{1\leq i\leq 4}\mu_{(i)})/\sigma>0.24. \qquad (3.3)$$

Mengerson and Bofinger's (1986) lower confidence bound is given as follows :

$$\min_{k-t+1\leq j\leq k}\mu_{(j)}-\max_{1\leq i\leq k-t}\mu_{(i)}\geqq(\bar{X}_{(k-t+1)}-\bar{X}_{(k-t)}-d^{MB}\sqrt{2/n}\,\hat{\sigma})^{-}$$

where $a^{-}=\min(a,0)$, and a positive constant d^{MB} is tabulated by them. When $k=5$, $t=1$ and $\alpha=0.05$, the constant is found to be $d^{MB}=2.16$. Thus in this case the 95% lower confidence bound by Bofinger and Mengerson is given as follows :

$$\mu_{(5)}-\max_{1\leq i\leq 4}\mu_{(j)}\geqq 0,$$

which is not quite useful as a confidence bound since it does not tell how much $\mu_{(5)}$ is better than others. Note that the confidence bound in (3.3) is strictly positive.

Next, consider the case where we are interested in top two plans with larger profits. Mengerson and Bofinger's result can be applied to make the following statement with 95% confidence.

$$\min(\mu_{(5)},\ \mu_{(4)}) - \max_{1\leq i\leq 3} \mu_{(i)} \geqq 0 \qquad (3.4)$$

from which we can make an inference that plans 4 and 5 are *at least as good as* plans 1, 2 *and* 3.

Now, consider a method of making *joint* inference on $\mu_{(5)} - \max_{1\leq i\leq 4} \mu_{(i)}$ and $\mu_{(4)} - \max_{1\leq i\leq 3} \mu_{(i)}$. To make a 95% *simultaneous* confidence statements, we start with 97.5% confdience lower bounds individually. In this example, the values of

$$h_\nu(\sqrt{n}(\bar{X}_{(5)} - \bar{X}_{(4)})/\sqrt{2}\hat{\sigma}) = h_\nu(2.92) \text{ and}$$

$$h_\nu(\sqrt{n}(\bar{X}_{(4)} - \bar{X}_{(3)})/\sqrt{2}\hat{\sigma}) = h_\nu(3.06)$$

corresponding to $\alpha^* = \alpha/2 = 0.025$ need to be computed. These values have been found to be, for $\alpha^* = 0.025$,

$$h_\nu(2.92) = 0.87 \text{ and } h_\nu(3.06) = 1.03.$$

Therefore, we can make a joint inference with 95% confidence as follows :

$$(\mu_{(5)} - \max_{1\leq i\leq 4} \mu_{(i)})/\sigma > 0.17 \text{ and } (\mu_{(4)} - \max_{1\leq i\leq 3} \mu_{(i)})/\sigma > 0.21.$$

Thus we can claim the strict superiorities of plan $(5) = 4$ over plan $(4) = 5$, and of plan $(4) = 5$ over plans 1, 2, 3 jointly with 95% confidence. Note that this is a much stronger conclusion than the inference form(3.4).

The above method of making a joint inference from marginal inferences is the so-called Bonferroni method. It can be generalized to make a joint inference at level $(1-\alpha)$ as follows : for all $t = 1, 2, \cdots, k-1$

$$(\mu_{(k-t+1)} - \max_{1\leq i\leq k-t} \mu_{(i)})/\sigma > \sqrt{2/n}\ h_\nu(\sqrt{n}(\bar{X}_{(k-t+1)} - \bar{X}_{(k-t)})/\sqrt{2}\hat{\sigma})$$

where h_ν is computed with $\alpha^* = \alpha/(k-1)$ replacing α in(3.2).

Finally, it should be remarked that the values of $h_\nu(w)$ were computed numerically by finding a root of (3.2) via the bisection method with accuracy up to 10^{-5}. In these computations, evaluation of the intergral was carried out by using IMSL's subroutine MDNOR and 32 points Gauss-Laguerre quadrature formula.

REFERENCES

1. Bofinger, E. (1983), *Multiple comparisons and selection*. Aust. J. Statist, 25(2), 198–207.
2. Gupta, S. S. and Panchapakesan, S. (1979), *Multiple Decision Procedures : Theory and Methodology of Selecting and Ranking Populations*. John Wiley & Sons, Inc., New York.
3. Gutmann, S. and Maymin, Z. (1985), *Testing retrospective hypotheses : is the selected population the best?* Unpublished manuscript.
4. Hsu, J. C. (1984), *Constrained simultaneous confidence intervals for multiple comparisons with the best*. Ann. Statist., 12. 1136–1144.
5. Kleijnen, J. P. C. Naylor, T. H. and Seaks T. G. (1972), *The use of Multiple ranking procedures to analyse simulations of management systems : a tutorial*. Management Science Application Series, 18 B, 245–257.
6. Mengerson, K. and bofinger, E. (1986) *Confidence bounds and selection of the best populations*. Unpublished manuscript.
7. Wijsman, R. A. (1985), *A useful inequality on ratios of integrals, with application to maximum likelihood estimation*. J. Am Statist. Assoc., 80, 472–475.