Journal of the Korean
Statistical Society
Vol. 23, No. 1, 1994

Multiple Comparisons with the Best in the Analysis of Covariance†

Young-Hoon Lee [1]

ABSTRACT

When a comparison is made with respect to the unknown best treatment, Hsu (1984, 1985) proposed the so called multiple comparisons procedures with the best in the analysis of variance model. Applying Hsu's result to the analysis of covariance model, simultaneous confidence intervals for multiple comparisons with the best in a balanced one-way layout with a random covariate are developed and are applied to a real data example.

KEYWORDS : Multiple comparison with the best, Analysis of covariance, Simultaneous confidence interval, Balanced one-way layout, Random covariate.

1. INTRODUCTION

In the analysis of covariance (ANCOVA), a most commonly used method to obtain simultaneous confidence intervals for all pairwise differences of treatment effects is the one proposed by Tukey (1953) and Kramer (1957). Tukey-Kramer method is applicable when covariates are fixed constants. When there

[1] Department of Statistics, Kangnŭng National University, Kangnŭng, 210-702, Korea.
† Research supported by Korea Research Foundation 1987-1989 at Seoul National University under the direction of Professor Woo-Chul Kim.

is a random covariate, Thigpen and Paulson (1974) gave exact simultaneous confidence intervals for all pairwise differences of treatment effects. Bryant and Paulson (1976) and Bryant and Bruvold (1980) extended the result by Thigpen and Paulson to more general settings.

When the comparison is made with respect to the unknown best treatment, Hsu (1984, 1985) proposed the so called multiple comparisons procedures with the best (MCB) in the analysis of variance model. There has not been any result regarding MCB type inference in the analysis of covariance. This paper develops MCB type confidence intervals in a balanced one-way layout with a random covariate.

Section 2 introduces a model and states some preliminary results. Section 3 derives simultaneous MCB type confidence intervals, and a real data example is considered in Section 4. All the proofs are given in the appendix.

2. NOTATIONS AND SOME PRELIMINARY FACTS

Consider the following ANCOVA model

$$y_{ij} = \mu_i + \beta(x_{ij} - \xi) + \varepsilon_{ij}, \quad i = 1, \cdots, k \ ; \ j = 1, \cdots, n \tag{2.1}$$

where (x_{ij}, y_{ij}) are assumed to be independently distributed with bivariate normal distributions having the mean vectors $(\xi, \mu_i)'$, the variances σ_x^2, σ_y^2 and the covariance σ_{xy}. Here, $\mu_1, \cdots, \mu_k$ denote the unknown treatment effects, β is the unknown regression coefficient, and ε_{ij}'s are unobservable errors with mean 0 and variance σ^2. It should be noted that under the model (2.1) the following relations hold:

$$\beta = \sigma_{xy}/\sigma_x^2 \ \text{ and } \ \sigma^2 = \sigma_y^2 - \sigma_{xy}^2/\sigma_x^2. \tag{2.2}$$

The unbiased estimators of the parameters involved are

$$\hat{\xi} = \bar{x}_{\cdot\cdot}, \quad \hat{\beta} = S_{xy}/S_{xx}, \quad \hat{\mu}_i = \bar{y}_{i\cdot} - \hat{\beta}(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot}), \quad \hat{\sigma}^2 = (S_{yy} - S_{xy}^2/S_{xx})/\nu$$

where $\bar{x}_{i\cdot}$, $\bar{x}_{\cdot\cdot}$, and $\bar{y}_{i\cdot}$ denote the sample means, S_{xx}, S_{yy}, and S_{xy} are the usual sums of squares and the sum of products, respectively, and $\nu = k(n-1) - 1$.

It is well-known (for example, Rao (1973) pp. 201-209) that

$$T_1 = (S_{xx})^{1/2}(\hat{\beta} - \beta)/\sigma, \quad T_2 = \nu\hat{\sigma}^2/\sigma^2, \quad \text{and} \quad T_3 = S_{xx}/\sigma_x^2$$

are independently distributed with the standard normal distribution, chi-square distributions with degrees of freedom ν and $(\nu + 1)$, respectively. Furthermore $\hat{\beta}$, $\hat{\sigma}^2$ and $(\bar{x}_{i\cdot}, \bar{y}_{i\cdot})$ $(i = 1, \cdots, k)$ are independent. It follows that

$$T_4 = \sigma_x(\nu + 1)^{1/2}(\hat{\beta} - \beta)/\sigma = T_1/\left\{T_3/(\nu + 1)\right\}^{1/2}$$

has a t-distribution with $(\nu + 1)$ degrees of freedom, and is independent of $T_2 = \nu\hat{\sigma}^2/\sigma^2$. Consequently,

$$T_5 = \left(1 + T_4^2/(\nu + 1)\right)^{-1}$$

has a beta distribution with parameters $(\nu + 1)/2$ and $1/2$.

In deriving an MCB procedure under the model (2.1), we shall need the pdf of

$$W = (T_2 T_5/2)^{1/2} = \left\{\frac{1}{2}\frac{\nu\hat{\sigma}^2/\sigma^2}{1 + \sigma_x^2(\hat{\beta} - \beta)^2/\sigma^2}\right\}^{1/2} \tag{2.3}$$

in Section 3. The following lemma gives the pdf of W.

Lemma 2.1. The pdf of W in (2.3) is given by

$$g_\nu(w) = \frac{2\nu\sqrt{\pi}}{\Gamma\left((\nu + 1)/2\right)} w^{\nu-1}\left\{1 - \Phi(\sqrt{2}\,w)\right\}, \quad w > 0 \tag{2.4}$$

where $\Gamma(\cdot)$ is the gamma function and $\Phi(\cdot)$ is the cdf of the standardized normal distribution.

3. MULTIPLE COMPARISONS WITH THE BEST

Hsu (1984, 1985) constructed simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j (i = 1, \cdots, k)$ in the case of one-way balanced ANOVA model. Following his idea, we develop simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j$ under the ANCOVA model (2.1). Suppose that a positive $d = d(k, \nu, \alpha)$ is chosen so that

$$\mathrm{P}_{\underline{0}}\left(\hat{\mu}_k > \max_{j\neq k} \hat{\mu}_j - d\hat{\sigma}/\sqrt{n}\right) = 1-\alpha \tag{3.1}$$

for a given $0 < \alpha < 1$. Here, $\mathrm{P}_{\underline{0}}(\cdot)$ means that the probability is computed when $\mu_1 = \cdots = \mu_k = 0$. The explicit determination of d is given in the next lemma.

Lemma 3.1. The positive constant d in (3.1) is determined by solving

$$1-\alpha = \int_0^\infty \left[\int_{-\infty}^\infty \Phi^{k-1}(z + d\sqrt{2/\nu}\, w)\, d\Phi(z)\right] g_\nu(w)\, dw \tag{3.2}$$

where $g_\nu(w)$ is given by (2.4).

Once a positive constant d is determined by (3.1), simultaneous confidence intervals for $\mu_i - \max_{j\neq i}\mu_j$ $(i = 1, \cdots, k)$ can be obtained in a manner exactly similar to Hsu (1984, 1985)'s. For the sake of completeness, the proof of the next result is included in the appendix.

Theorem 3.1. The $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_i - \max_{j\neq i}\mu_j$ $(i = 1, \cdots, k)$ are given as follows:

$$\mu_i - \max_{j\neq i}\mu_j \in \pm(\hat{\mu}_i - \max_{j\neq i}\hat{\mu}_j \pm d\hat{\sigma}/\sqrt{n}\,)^{\pm}, \quad i = 1, \cdots, k \tag{3.3}$$

where d is determined by (3.2) and $x^+(x^-)$ denotes the positive (negative, respectively) part of x.

To implement the simultaneous confidence bounds in (3.3), we need to find the values $d = d(k, \nu, \alpha)$ by (3.2). The values of d are tabulated in Table 2 at the end of the paper for k = 2, 3, 4, 5, ν = 5(1)20, 26, 30, 60, 90, 120, ∞, and $\alpha = 0.01$.

The defect of the confidence intervals in (3.3) is that they always contain 0. This means that even if a treatment is inferred to be the best by those intervals, no positive lower bound on how much it is better than the rest can be given by the intervals in (3.3). Such a defect was noted by Hsu (1985), and he called the intervals of the type (3.3) the constrained confidence intervals, while he devised the so-called unconstrained confidence intervals by employing slightly larger cutoff points in the case of ANOVA model. Such a modification to the confidence intervals in (3.3) can be carried out in exactly the same manner as that in Hsu (1985). Thus the results will be stated without proofs.

Let $c = c(k, \nu, \alpha)$ be a positive constant satisfying

$$P_{\underline{0}}\left(\max_{j\neq k} \hat{\mu}_j - c\hat{\sigma}/\sqrt{n} < \hat{\mu}_k < \hat{\mu}_{k-1} + c\hat{\sigma}/\sqrt{n}\right) = 1 - \alpha \tag{3.4}$$

for a given α. Methods similar to those in Lemma 3.1 can be used to show that (3.4) is equivalent to the following.

$$\int_0^\infty \left[\int_{-\infty}^\infty \Phi^{k-2}(z + c\sqrt{2/\nu}\, w)\left\{\Phi(z + c\sqrt{2/\nu}\, w) - \Phi(z - c\sqrt{2/\nu}\, w)\right\} d\Phi(z)\right] g_\nu(w)\, dw = 1 - \alpha \tag{3.5}$$

where $g_\nu(w)$ is given by (2.4). The cutoff points c satisfying (3.5) are also tabulated in Table 2 at the end of paper for selected values of k, ν and $\alpha = 0.01$. Note that c-values are larger than the d-values, reflecting more restriction in the probability statement in (3.4) than (3.1). The table of the values of c and d for $\alpha = 0.05$ has been omitted now to save space, although it is available from the author.

Let (k) denote the random index corresponding to the largest $\hat{\mu}_i$ ($i = 1, \cdots, k$), i.e. $\hat{\mu}_{(k)} = \max_{1\leq i\leq k} \hat{\mu}_i$. Define the confidence bounds as follows:

$$L_i = \hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j - c\hat{\sigma}/\sqrt{n} \quad (i = 1, \cdots, k), \tag{3.6}$$

$$U_i = \begin{cases} \min\left\{(\hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j + c\hat{\sigma}/\sqrt{n})^+, -L_{(k)}\right\}, & i \neq (k) \\ \hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j + c\hat{\sigma}/\sqrt{n}, & i = (k). \end{cases} \tag{3.7}$$

The next result gives the so-called unconstrained confidence intervals.

Theorem 3.2. The $100(1 - \alpha)\%$ simultaneous confidence intervals for $\mu_i - \max_{j\neq i} \mu_j$ ($i = 1, \cdots, k$) are given by

$$L_i \leq \mu_i - \max_{j\neq i} \mu_j \leq U_i \quad i = 1, \cdots, k \tag{3.8}$$

where L_i and U_i ($i = 1, \cdots, k$) are defined by (3.6) and (3.7).

As noted previously, the lower bounds given by (3.3) are always *non-positive* while the lower bounds provided by (3.8) can be *positive*. By employing the confidence bounds in (3.8), we can provide a positive lower bound for $\mu_{(k)} - \max_{j \neq (k)} \mu_j$ when $\mu_{(k)} > \max_{j \neq (k)} \mu_j + c\hat{\sigma}/\sqrt{n}$.

4. A REAL DATA EXAMPLE

The data in Table 1 below are cited in Wildt and Ahtola (1978). They are the records of traffic accidents in 12 rural municipalities in Finland for two years.

Table 1. Traffic Accidents During the Test Period (y) and Traffic Accidents During Preceding Year (x)

Type of Liquor Licensing (Treatments or Groups)					
No Store or Restaurant (Group 1)		Package Store Only (Group 2)		Restaurant & Package Store (Group 3)	
x	y	x	y	x	y
190	177	252	226	206	226
261	225	228	196	239	229
194	167	240	198	217	215
217	176	246	206	177	188

We assume the ANCOVA model in (2.1) for this data set. The ordinary F-test for the hypothesis $\mathrm{H}_0 : \mu_1 = \mu_2 = \mu_3$ rejects H_0 at $\alpha = 0.01$, where μ_i is the average impact of liquor licensing in the i-th group on traffic accidents. It does not, however, give any idea about which group has higher impacts.

The cutoff values of the confidence intervals in (3.3) and (3.8) can be found from Table 2 for $\alpha = 0.01$, which are given by

$$d(3, 8, 0.01) = 4.92 \quad \text{and} \quad c(3, 8, 0.01) = 5.41.$$

Using these values, we obtain the following 99% simultaneous confidence intervals from (3.3), (3.6), and (3.7):

99%	Constrained	Unconstrained
$\mu_1 - \max(\mu_2, \mu_3)$	[-56.24, 0]	[-58.61, -4.97]
$\mu_2 - \max(\mu_1, \mu_3)$	[-54.94, 0]	[-57.31, -4.97]
$\mu_3 - \max(\mu_1, \mu_2)$	[0, 54.94]	[4.97, 57.31]

From the constrained confidence intervals, we can conclude that group 3 has the largest impact on traffic accidents, while the unconstrained confidence intervals give additional information about how much higher impacts group 3 has than the other groups. Finally, we remark that Thigpen and Paulson (1974)'s method does not separate group 3 from the other groups at 99%.

Table 2. The cutoff points $d(k,\nu,\alpha)$ and $c(k,\nu,\alpha)$ for $\alpha = 0.01$

	$d(k,\nu,\alpha)$				$c(k,\nu,\alpha)$			
k / ν	2	3	4	5	2	3	4	5
5	5.25	6.12	6.62	6.97	6.35	6.89	7.26	7.52
6	4.80	5.55	5.98	6.28	5.72	6.17	6.48	6.71
7	4.52	5.18	5.56	5.82	5.32	5.71	5.98	6.18
8	4.33	4.92	5.27	5.51	5.04	5.41	5.64	5.82
9	4.18	4.74	5.06	5.29	4.84	5.18	5.40	5.56
10	4.07	4.60	4.90	5.11	4.69	5.00	5.21	5.37
11	3.98	4.49	4.78	4.98	4.57	4.87	5.07	5.21
12	3.91	4.41	4.68	4.87	4.48	4.76	4.95	5.09
13	3.84	4.33	4.60	4.78	4.40	4.67	4.85	4.99
14	3.80	4.27	4.53	4.71	4.33	4.60	4.77	4.91
15	3.76	4.22	4.47	4.65	4.27	4.53	4.71	4.84
16	3.73	4.17	4.42	4.59	4.23	4.48	4.66	4.78
17	3.70	4.13	4.38	4.55	4.19	4.43	4.61	4.73
18	3.67	4.10	4.34	4.50	4.15	4.39	4.56	4.69
19	3.65	4.07	4.30	4.47	4.13	4.36	4.52	4.65
20	3.63	4.04	4.27	4.43	4.09	4.33	4.49	4.61
26	3.53	3.93	4.15	4.30	3.97	4.20	4.35	4.46
30	3.49	3.88	4.10	4.24	3.92	4.14	4.28	4.39
60	3.37	3.73	3.92	4.06	3.76	3.96	4.09	4.19
90	3.33	3.68	3.87	4.00	3.71	3.91	4.04	4.13
120	3.31	3.65	3.84	3.97	3.68	3.88	4.01	4.10
∞	3.25	3.58	3.77	3.89	3.62	3.80	3.92	4.01

REFERENCES

(1) Bryant, J. L. and Bruvold, N. T. (1980). Multiple Comparison Procedures in the ANCOVA, *Journal of the American Statistical Association*, 75, 874-880.

(2) Bryant, J. L. and Paulson, A. S. (1976). An Extension of Tukey's Method of Multiple Comparisons to Experimental Designs with Random Concomitant Variables, *Biometrika*, 63, 631-638.

(3) Hsu, J. C. (1984). Constrained Simultaneous Confidence Intervals for Multiple Comparisons with the Best, *The Annals of Statistics*, 12, 1136-1144.

(4) Hsu, J. C. (1985). A Method of Unconstrained Multiple Comparisons with the Best, *Communications in Statistics*, A 14, 2009-2028.

(5) Kramer, C. Y. (1957). Extension of Multiple Range Tests to Group Correlated Adjusted Means, *Biometrics*, 13, 13-18.

(6) Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, New York: Wiley.

(7) Thigpen, C. C. and Paulson, A. S. (1974). A Multiple Range Test for Analysis of Covariance, *Biometrika*, 61, 479-484.

(8) Tukey, J. W. (1953). The Problem of Multiple Comparisons, *Unpublished Manuscript*, Department of Statistics, Princeton University, Princeton, New Jersey 08544.

(9) Wildt, A. R. and Ahtola, O. T. (1978). *Analysis of Covariance*, Beverly Hills: Sage Publications, Inc..

APPENDIX

Proof of Lemma 2.1 It follows from the joint distribution of T_2 and T_5 that the joint pdf of $W = (T_2T_5/2)^{1/2}$ and $Z = T_2/2$ is given by

$$f(w,z) = \frac{\nu}{\sqrt{\pi}\,\Gamma\left((\nu+1)/2\right)} \frac{w^\nu e^{-z}}{z\sqrt{z-w^2}}, \quad z > w^2, \quad w > 0.$$

Thus the pdf of W is given by

$$g_\nu(w) = \frac{\nu}{\sqrt{\pi}\,\Gamma\left((\nu+1)/2\right)} w^\nu \int_{w^2}^{\infty} \frac{e^{-z}}{z\sqrt{z-w^2}}\,dz. \qquad \text{(A.1)}$$

The integral in (A.1) can be evaluated by considering the Laplace transform $\mathcal{L}\{h\}$ of

$$h(t) = \int_{w^2}^{\infty} \frac{e^{-tz}}{z\sqrt{z-w^2}}\,dz, \quad t > 0.$$

Simple integration shows that

$$\mathcal{L}\{h\}(s) = \frac{\pi}{s}\left(\frac{1}{w} - \frac{1}{\sqrt{w^2+s}}\right), \quad s > 0.$$

Thus, inverting this Laplace transform, we obtain

$$h(t) = \frac{2\pi}{w}\left\{1 - \Phi(\sqrt{2t}\,w)\right\}, \quad t > 0.$$

Taking $t = 1$, we obtain the integral in (A.1) so that the pdf of W can be written in the form of (2.4) ■.

Proof of Lemma 3.1 First, note that

$$\hat{\mu}_k - \hat{\mu}_j = \left\{\bar{y}_{k\cdot} - \hat{\beta}(\bar{x}_{k\cdot} - \xi)\right\} - \left\{\bar{y}_{j\cdot} - \hat{\beta}(\bar{x}_{j\cdot} - \xi)\right\}, \; j = 1, \cdots, k-1$$

and that $(\bar{x}_{i\cdot}, \bar{y}_{i\cdot})$ $(i = 1, \cdots, k)$ are independently distributed with $\hat{\beta}$ and $\hat{\sigma}^2$. Furthermore, the conditional distribution of $\bar{y}_{i\cdot} - \hat{\beta}(\bar{x}_{i\cdot} - \xi)$, given $\hat{\beta}$, is the normal distribution with mean μ_i and variance

$$\left(\sigma_y^2 - 2\hat{\beta}\sigma_{xy} + \hat{\beta}^2\sigma_x^2\right)/n = \left\{\sigma^2 + \sigma_x^2(\hat{\beta} - \beta)^2\right\}/n. \tag{A.2}$$

The identity in (A.2) follows from (2.2). Therefore, when $\mu_1 = \cdots = \mu_k = 0$, the conditional distribution of $\hat{\mu}_k - \hat{\mu}_j$ $(j = 1, \cdots, k-1)$ given $\hat{\beta}$ and $\hat{\sigma}^2$ can be represented as follows: for $j = 1, \cdots, k-1$,

$$\sqrt{n}\,(\hat{\mu}_k - \hat{\mu}_j)/\hat{\sigma} \overset{\mathrm{d}}{\equiv} \left\{\sigma^2 + \sigma_x^2(\hat{\beta} - \beta)^2\right\}^{1/2}(z_k - z_j)/\hat{\sigma} \tag{A.3}$$

where $z_1, \cdots, z_k$ are independent standard normal random variables independent of $\hat{\beta}$ and $\hat{\sigma}$, and $\overset{\mathrm{d}}{\equiv}$ means that the distributions of both sides are the same. Note that the right hand side of (A.3) can be written as $(z_k - z_j)/(\sqrt{2/\nu}\,W)$ where W is given in (2.3). Thus, the probability in (3.1) can be written as

$$\mathrm{P}_{\underline{0}}\left(z_k > \max_{j\neq k} z_j - d\sqrt{2/\nu}\, W\right)$$

where $z_1, \cdots, z_k$ are independent standard normal random variables, and are independent of W. Hence the equation (3.2) follows ■.

Proof of Theorem 3.1 Consider a pivotal event

$$E = (\hat{\mu}_{[k]} - \mu_{[k]} > \max_{i\neq[k]}(\hat{\mu}_i - \mu_i) - d\hat{\sigma}/\sqrt{n}\,)$$

where $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$ denote the ordered $\mu_1, \mu_2, \cdots, \mu_k$. Note that by (3.1)

$$\mathrm{P}_{\underline{\theta}}(E) = 1 - \alpha \quad \text{for any} \quad \underline{\theta} = (\mu_1, \cdots, \mu_k, \xi, \beta, \sigma_x^2, \sigma_y^2, \sigma_{xy}).$$

Thus, it suffices to show that the event E implies the lower and upper confidence bounds in (3.3). The upper confidence bounds in (3.3) are obtained from the following:

$$\begin{aligned}
E &= (\mu_{[k]} - \mu_i < \hat{\mu}_{[k]} - \hat{\mu}_i + d\hat{\sigma}/\sqrt{n} \;\text{ for all }\; i \neq [k]) \\
&\subset (\mu_{[k]} - \max_{j\neq[k]} \mu_j < \hat{\mu}_{[k]} - \hat{\mu}_j + d\hat{\sigma}/\sqrt{n} \;\text{ for all }\; j \neq [k]) \\
&= (\mu_{[k]} - \max_{j\neq[k]} \mu_j < \hat{\mu}_{[k]} - \max_{j\neq[k]} \hat{\mu}_j + d\hat{\sigma}/\sqrt{n} \;\text{ and} \\
&\qquad \mu_i - \max_{j\neq i} \mu_j \leq (\hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j + d\hat{\sigma}/\sqrt{n}\,)^+ \;\text{ for all }\; i \neq [k]) \\
&\subset (\mu_i - \max_{j\neq i} \mu_j \leq (\hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j + d\hat{\sigma}/\sqrt{n}\,)^+ \;\text{ for all }\; i).
\end{aligned}$$

The lower confidence bounds in (3.3) can be obtained similarly as follows:

$$\begin{aligned}
E &= (\mu_i - \mu_{[k]} < \hat{\mu}_i - \hat{\mu}_{[k]} - d\hat{\sigma}/\sqrt{n} \;\text{ for all }\; i \neq [k]) \\
&\subset (\mu_i - \mu_{[k]} > \hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j - d\hat{\sigma}/\sqrt{n} \;\text{ for all }\; i \neq [k]) \\
&= (\mu_{[k]} - \max_{j\neq[k]} \mu_j \geq -(\hat{\mu}_{[k]} - \max_{j\neq[k]} \hat{\mu}_j - d\hat{\sigma}/\sqrt{n}\,)^- \;\text{ and} \\
&\qquad \mu_i - \max_{j\neq i} \mu_j > \hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j - d\hat{\sigma}/\sqrt{n} \;\text{ for all }\; i \neq [k]) \\
&\subset (\mu_i - \max_{j\neq i} \mu_j \geq -(\hat{\mu}_i - \max_{j\neq i} \hat{\mu}_j - d\hat{\sigma}/\sqrt{n}\,)^- \;\text{ for all }\; i),
\end{aligned}$$

which completes the proof ■.