2.10 Multiple Comparison Procedures

This section ia compilation of material from your text, the *SAS/STAT User's Guide* (1992), Hochberg and Tamhane (1987), Toothaker (1991), Miller (1981), and Milliken and Johnson (1992).

- When performing hypothesis tests on a single mean, on differences between means, or on contrasts, we must realize the following limitations:
  - Individual tests using a specified α level are only appropriate if the tests were not suggested by the data.
  - If we run a series of hypothesis tests each at a specified α level, then the α level applies to the individual tests and not for a set of tests.
- A procedure that enables us to make more than one comparison among two or more means is called a **multiple comparison procedure (MCP)**.
- The most common situation occurs when the researcher suspects that there may be differences among the a means and it is important to determine which means can be considered significantly different from each other. The typical approach is to look at all $a(a-1)/2$ pairwise comparisons of the form $\mu_i - \mu_j$ and test for significant differences.
- Failure to reject $H_0 : \mu_i = \mu_j$ in a MCP should not lead you to conclude that H_0 is true. Failure to reject H_0 implies only that the difference between population means, if any, is not large enough to be detected for the given n_i sample sizes.
- Warnings: A MCP can lead to counter-intuitive results when the sample sizes are unequal and there are large disparities among the n_i sample sizes. For example, suppose there are 4 treatments (A, B, C, D) such that $\overline{y}_{A\cdot} > \overline{y}_{B\cdot} > \overline{y}_{C\cdot} > \overline{y}_{D\cdot}$. Suppose $n_A = n_D = 3$ while $n_B = n_C = 100$. Then the difference between μ_B and μ_C may be significant while the difference between μ_A and μ_D is not.
- Generating a set of confidence intervals may be more informative than the results of the set of significance tests in a MCP. Confidence intervals (i) show the degree of uncertainty in each comparison, (ii) indicate the statistical significance and (iii) may prevent nonstatisticians from falsely concluding that means that are not significantly different implies equality.
- Details will be presented for the following multiple MCPs: Fisher's Least Significant Difference (LSD), Bonferroni, Sidak, Tukey, and Dunnett.

Two Types of Hypotheses

The following two types of hypotheses describe two possible relationships that may exist among the population means. These will be related to different types of MCP error rates.

- A **complete null hypothesis** exists if all of the a means are equal. This corresponds to H_0 for the F-test in the ANOVA. This is also referred to as the 'full' or 'overall' null hypothesis.
- A **partial null hypothesis** exists when the complete null hypothesis is false, but at least one pair of means are equal.
  - For example, suppose there are 4 treatments (A, B, C, D) and that $\mu_A = \mu_B = \mu_C$ but that $\mu_i \neq \mu_D$ for $i = A, B, C$. Then, there is a partial null hypothesis in the three equal means.
- For Fisher's Least Significant Difference (LSD), Bonferroni, Sidak, and Tukey MCPs, we only need to consider the complete null hypothesis.

Types of Errors

- A **Type I error** is a rejection of the null hypothesis H_0 when it is true. It is typical to let α be the probability of a Type I error.
- A **Type II error** is a failure to reject the null hypothesis H_0 when it is false. For a pairwise comparison $\mu_i - \mu_j$, the probability of a Type II error, denoted β, depends on the value of $\mu_i - \mu_j$.
- The **power of a test** $= 1 - \beta =$ the probability of rejecting H_0 when it is false.

Error Rates

- Situation: Suppose that a MCP requiring a total of C pairwise comparisons or contrasts is to be performed.
- When choosing a MCP, the researcher (whether he or she realizes it) is selecting a method of controlling the error rate. That is, for each MCP there is a value of α^* assigned to each of the C comparisons. This choice also affects the power of the chosen testing procedure.
- The **comparisonwise error rate (CER)** is the Type I error rate for each comparison. We will use α^* to denote the CER.
- As an alternative to CER, we can control the **experimentwise error rate for the complete null hypothesis (EERC)** . That is, we may want to control the probability of making at least one Type I error among the C comparisons (and not per comparison) to be at most some prespecified value α when the complete null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ is true. In this case the CER $\alpha^* < \alpha$.
- In addition to the CER and the EERC, there is one more error rate that will be useful in evaluating the performance of the various MCPs:
  - The **maximum experimentwise error rate (MEER)** under any complete or partial null hypothesis is the probability of making at least Type I error if any complete or partial null hypothesis is true (at least one pair and possibly all pairs of means are equal).
- Suppose$C$ pairwise comparisons or contrasts are being performed with CER $= \alpha^*$. Then

$$\begin{aligned} \Pr(\text{of at least one Type I error}) &= 1 - \Pr(\text{of no Type I errors}) \\ &\leq \qquad (1) \\ &< \qquad (2) \end{aligned}$$

- Equality in (1) will occur if the C comparisons consist of orthogonal contrasts. Empirical studies have shown that the probability of at least one Type I error is fairly close to the bound $1 - (1 - \alpha^*)^C$ even when the contrasts are not orthogonal.
- Fixing the CER α^* can increase the power but at the expense of a larger EERC. Specifically, assuming α^* is fixed, the EERC among the C comparisons increases as C increases.
- By fixing the EERC at some value α (e.g., $\alpha = .05$), the experimenter wants to prevent declaring too many comparisons significant by chance alone. Whenever the experimenter is trying to answer many questions using MCPs in a single experiment, it is often recommended that the experimenter control the EERC.
- It is possible to control the MEER (and hence the EERC) by setting the CER α^* at a value smaller than the desired overall α level. Two ways this can be accomplished are by reconsidering the inequalities in (1) and (2).

Method 1: Set the CER $\alpha^* = \alpha$. Then from (1)

$$\Pr(\text{of at least one Type I error}) \leq 1 - (1-\alpha)^C .$$

Method 2: Set the CER $\alpha^* = \alpha/C$. Then from (2)

$$\Pr(\text{of at least one Type I error}) < C\alpha^* = \alpha.$$

This method assures that the MEER $< \alpha$. This corresponds to the **Bonferroni** MCP.

Method 3: Using (1), find α^* such that $\alpha = 1-(1-\alpha^*)^C$. This yields CER $\alpha^* = 1-(1-\alpha)^{1/C}$. Thus,

$$\Pr(\text{of at least one Type I error}) \leq 1-(1-\alpha^*)^C = \alpha.$$

This method assures that the MEER $\leq \alpha$. These corresponds to the **Sidak** MCP.

- **Method 4**: If all of the sample sizes are equal, the EERC can also be controlled by using the distribution of a **studentized range** (which is the range of a independent standard normal random variables (e.g., $(\overline{y}_i - \mu_i)/(\sigma/\sqrt{n})$) divided by the square root of an independent χ^2/ν variable with ν d.f. (e.g., $\sqrt{MS_E/\sigma^2}$ with $\nu = N - a$ d.f.). This approach corresponds to Tukey's test.

2.10.1 MCP I: Fisher's Least Significant Difference (LSD)

- Situation: Given that F_0 from the ANOVA indicates rejection of $H_0 : \mu_1 = \cdots = \mu_a$, we want to consider ALL possible pairwise comparisons among the a treatment means:

$$H_0 : \mu_i = \mu_j \qquad \text{vs.} \qquad H_1 : \mu_i \neq \mu_j \qquad \text{while fixing the CER } = \alpha.$$

Simultaneous Testing Using Test Statistics

To perform Fisher's LSD MCP:

- If $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ is not rejected, then no means are considered significantly different.
- If $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ is rejected, then perform individual t-tests for each of the $a(a-1)/2$ pairwise comparisons for a specified α.
  - Reject H_0 if $|t| \geq t^*$ where the test statistic $t = \dfrac{\widehat{D}_{ij}}{se(\widehat{D}_{ij})}$ and $t^* =$ is the critical t-value from the t-distribution with $N-a$ d.f..
- The confidence interval for $\mu_i - \mu_j$ is $\widehat{D}_{ij} \pm t^* se(\widehat{D}_{ij})$.

Comments

- It is up to the user to first check the ANOVA F-test result to decide whether or not to proceed to look at the pairwise comparisons.
- Using the results of the preliminary ANOVA F-test will control the EERC. That is, the probability of rejecting $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ when it is true, is α.
- Once the researcher proceeds to the set of pairwise comparisons given rejection from the ANOVA F-test, the MEER is no longer controlled. That is, as the number of treatments a gets larger, the MEER may be considerably larger than α.

- This test is also called the *protected t-test* because the individual t-tests are computed only if the ANOVA F-test indicates significance (i.e., the ANOVA F is seen as 'protecting' the usual t- tests). 'LSD' corresponds to the fact that α-level critical t-value is the smallest critical value that the t-statistic must exceed to be significant when considering only a single comparison.
- Because of the poor control of the MEER, the LSD procedure is generally not recommended. However, because the CER is controlled at α for each comparison. Thus, this procedure has the highest power among the four MCPs.
- If, however, the experiment was exploratory in nature, then the results for the LSD MCP could suggest which levels to consider for future experimentation.
- Fisher's LSD procedure does not require equal sample sizes.

2.10.2 MCP II: Bonferroni

- Situation: We are considering a particular set of contrasts (e.g., pairwise comparisons) for the a treatment means that are specified by the researcher prior to data analysis while bounding MEER at α.

Testing Using Simultaneous Bonferroni Confidence Intervals

- Suppose we are interested in a family of C contrasts (e.g., C pairwise comparisons).
- The Bonferroni multiple comparison confidence limits for each contrast Γ are:

$$\widehat{\Gamma} \quad \pm \quad t^* \, se(\widehat{\Gamma})$$

where

$$\widehat{\Gamma} = \sum_{i=1}^{a} c_i \overline{y}_{i\cdot} \qquad se(\widehat{\Gamma}) = \sqrt{MS_E \sum (c_i^2/n_i)} \qquad t^* =$$

and $t(\alpha/2C, N-a)$ is a critical value associated with the t-distribution with $N-a$ degrees of freedom.

- If a confidence interval for Γ does not contain 0 then we reject $H_0 : \Gamma = 0$ in favor of $H_1 : \Gamma \neq 0$.

Simultaneous Testing Using a Test Statistic

- To perform Bonferroni's MCP for Pairwise Comparisons:
  1. For each comparison of means $(\mu_i - \mu_j)$, calculate $\widehat{D}_{ij} = \overline{y}_{i\cdot} - \overline{y}_{j\cdot}$ and $se(\widehat{D}_{ij})$.
  2. Calculate $b_d = t(\alpha/2C, N-a) se(\widehat{D}_{ij})$.
  3. Decision rule: Reject $H_0 : \mu_i = \mu_j$ if $|\widehat{D}_{ij}| \geq b_d$.

Comments

- The MEER $< \alpha$ for the Bonferroni MCP.
- The Bonferroni MCP uses the actual number of comparisons C in the selection of critical t-values. Because of the reliance on C, the Bonferroni procedure will have relatively good power for small sets of planned comparisons and relatively low power for large sets of planned comparisons.
- The Bonferroni procedure is applicable whether or not the sample sizes are equal.

2.10.3 MCP III: Sidak

- Situation: We want to consider a particular set of pairwise comparisons or contrasts among the a treatment means that are specified by the researcher prior to data analysis while bounding MEER at α.
- The procedure is the same as the Bonferroni procedure except that $t^* =$ where $\alpha^* = 1 - (1-\alpha)^{1/C}$.

Comments

- The Bonferroni MCP comments also apply to the Sidak MCP except that the MEER $\leq \alpha$.
- Although the critical t-values for the Bonferroni and Sidak tests are close, the Sidak values will always be smaller than the Bonferroni values. Thus, the Sidak procedure will have both a slightly higher MEER and power than the Bonferroni procedure.

2.10.4 MCP IV: Tukey's Honest Significant Difference (HSD)

- Situation: Given that F_0 from the ANOVA indicates rejection of $\text{H}_0 : \mu_1 = \cdots = \mu_a$, we now want to find which means are different. That is for all $i \neq j$, test

$$H_0 : \mu_i = \mu_j \qquad \text{vs.} \qquad H_1 : \mu_i \neq \mu_j$$

while keeping the MEER $\leq \alpha$.

Testing Using Simultaneous Tukey Confidence Intervals

- The Tukey multiple comparison confidence limits for all pairwise comparisons $D = \mu_i - \mu_j$ are:

$$\widehat{D}_{ij} \;\; \pm \;\; T\, se(\widehat{D}_{ij})$$

where

$$\widehat{D}_{ij} = \overline{y}_{i\cdot} - \overline{y}_{j\cdot} \qquad se(\widehat{D}_{ij}) = \sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \qquad T =$$

and $q_\alpha(a, N-a)$ is the critical value from the corresponding studentized range distribution.

- If a confidence interval for $\mu_i - \mu_j$ does not contain 0 then we reject $H_0 : \mu_i = \mu_j$ in favor of $H_1 : \mu_i \neq \mu_j$.

Simultaneous Testing Using Critical Studentized Range Values

- To perform Tukey's MCP:

  1. For each comparison of means $(\mu_i - \mu_j)$, calculate

$$\widehat{D}_{ij} = \overline{y}_{i\cdot} - \overline{y}_{j\cdot} \qquad \text{and} \qquad se(\widehat{D}_{ij}) \;=\; \sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

  2. Calculate $q_d = \dfrac{q_\alpha(a, N-a)}{\sqrt{2}} se(\widehat{D}_{ij})$.
  3. Decision rule: Reject $H_0 : \mu_i = \mu_j$ if $|\widehat{D}_{ij}| \geq q_d$.

Comments

- When the treatment sample sizes are equal, then the EERC $= \alpha$ and the MEER $< \alpha$.
- When the treatment sample sizes are not equal, then the EERC is less than α. Thus, the Tukey method is Type I error conservative when the sample sizes are unequal.
- The Tukey procedure is more powerful than the Bonferroni or Sidak procedures when considering <u>all</u> pairwise comparisons.
- However, unlike the Bonferroni and Sidak methods, the Tukey procedure depends on the number of treatments a and not on the number of comparisons C. Thus, when only C pairwise comparisons are of interest ($C < a(a-1)/2$), the Tukey procedure can have lower power for smaller C.

2.10.5 MCP V: Dunnett's MCP for Comparisons to a Control

- <u>Situation:</u> The experimenter is comparing $a-1$ experimental treatments to a control.
- In this situation, the desired inference may be directional or one-sided. The question is often 'Is one of the treatments better than the control?'. When this is the case, one-tailed critical values are used.
- In other cases, the researcher just wants to detect if the treatment means are different than the control mean. For such a nondirectional (or two-sided) hypothesis, two-tailed critical values are used.
- Then for each experimental treatment i, we
  1. Calculate the $a-1$ differences $D_i = \overline{y}_{i\cdot} - \overline{y}_{control}$.
  2. For a one-sided alternative, calculate

$$D_\alpha = d_\alpha(a-1, N-a)\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_{control}}\right)}.$$

<u>Decision Rule</u> for $H_1 : \mu_i > \mu_{control}$, reject $H_0 : \mu_i = \mu_{control}$ if $D_i > D_\alpha$.
<u>Decision Rule</u> for $H_1 : \mu_i < \mu_{control}$, reject $H_0 : \mu_i = \mu_{control}$ if $D_i < -D_\alpha$.

  3. For a two-sided alternative, calculate

$$D_\alpha = d_\alpha(a-1, N-a)\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_{control}}\right)}.$$

<u>Decision Rule</u> for $H_1 : \mu_i \neq \mu_{control}$, reject $H_0 : \mu_i = \mu_{control}$ if $|D_i| > D_\alpha$.

- The one and two-sided $d_\alpha(a-1, N-a)$ values are contained in tables in the Appendix of the textbook.
- For Dunnett's test, the MEER $< \alpha$.

References

- Hochberg, Yosef and Tamhane, Ajit C. (1987), *Multiple Comparison Procedures*, Wiley, New York.
- Miller, Rupert G. (1981), *Simultaneous Statistical Inference*, Springer-Verlag, New York.
- Milliken, George A. and Johnson, Dallas E. (1992), *The Analysis of Messy Data, Volume I: Designed Experiments*, Van Nostrand Reinhold, New York.
- Toothaker, Larry E. (1991), *Multiple Comparisons for Researchers*, SAGE Publications, Newbury Park, CA.
- *SAS/STAT User's Guide Volume 2* Fourth Edition (1992), SAS Institute Inc., Cary NC.

2.10.6 Multiple Comparison Procedure Examples

Example: A single-factor CRBD was run with $a = 5$ treatments and $n = 5$ replications per treatment. The data are given below:

Treatment				
A	B	C	D	E
19.08	22.04	18.68	16.99	15.34
17.07	21.44	19.86	13.18	13.52
18.91	18.82	19.68	16.97	15.23
15.09	20.49	17.78	12.90	15.63
17.00	19.34	17.86	15.00	13.21

Perform Fisher's LSD, Bonferroni, Sidak, Tukey, and Dunnett MCPs.

SAS Code for MCP Examples

```
DM 'LOG; CLEAR; OUT; CLEAR;';
ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\mct1.PDF';
OPTIONS NODATE NONUMBER LS=76;

*****************************************;
*** Multiple Comparison Test Example  ***;
*****************************************;

DATA in;
     DO trt = 'A', 'B', 'C', 'D', 'E';
     DO rep = 1 to 5;
        INPUT y @@; OUTPUT;
     END; END;
LINES;
19.08 17.07 18.91 15.09 17.00
22.04 21.44 18.82 20.49 19.34
18.68 19.86 19.68 17.78 17.86
16.99 13.18 16.97 12.90 15.00
15.34 13.52 15.23 15.63 13.21
;
* PROC PRINT DATA=in;

PROC GLM DATA=in ; * PLOTS = (ALL);
     CLASS trt;
     MODEL y = trt / SS3 SOLUTION;
     MEANS trt / LSD TUKEY BON SIDAK ALPHA=.05 CLDIFF LINES;
     MEANS trt / DUNNETT('C') DUNNETTL('C') DUNNETTU('C');
TITLE 'Multiple Comparison Test Example';
RUN;
```

Multiple Comparison Test Example

The GLM Procedure

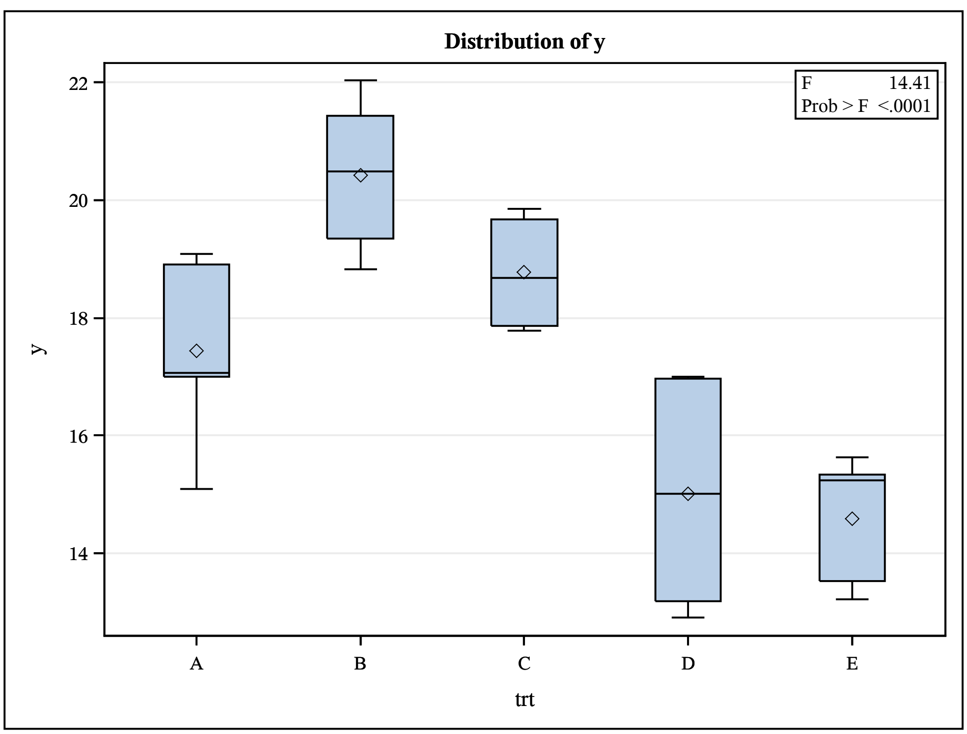
Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	122.7958160	30.6989540	14.41	<.0001
Error	20	42.5974000	2.1298700		
Corrected Total	24	165.3932160			

R-Square	Coeff Var	Root MSE	y Mean
0.742448	8.463080	1.459407	17.24440

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	4	122.7958160	30.6989540	14.41	<.0001

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		14.58600000	B	0.65266684	22.35	<.0001
trt	**A**	2.84400000	B	0.92301029	3.08	0.0059
trt	**B**	5.84000000	B	0.92301029	6.33	<.0001
trt	**C**	4.18600000	B	0.92301029	4.54	0.0002
trt	**D**	0.42200000	B	0.92301029	0.46	0.6525
trt	**E**	0.00000000	B	.	.	.



Fisher's LSD MCP

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of t	2.08596
Least Significant Difference	1.9254

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	**Difference Between Means**	**95% Confidence Limits**		
B - C	1.6540	-0.2714	3.5794	
B - A	2.9960	1.0706	4.9214	***
B - D	5.4180	3.4926	7.3434	***
B - E	5.8400	3.9146	7.7654	***
C - B	-1.6540	-3.5794	0.2714	
C - A	1.3420	-0.5834	3.2674	
C - D	3.7640	1.8386	5.6894	***
C - E	4.1860	2.2606	6.1114	***
A - B	-2.9960	-4.9214	-1.0706	***
A - C	-1.3420	-3.2674	0.5834	
A - D	2.4220	0.4966	4.3474	***
A - E	2.8440	0.9186	4.7694	***
D - B	-5.4180	-7.3434	-3.4926	***
D - C	-3.7640	-5.6894	-1.8386	***
D - A	-2.4220	-4.3474	-0.4966	***
D - E	0.4220	-1.5034	2.3474	
E - B	-5.8400	-7.7654	-3.9146	***
E - C	-4.1860	-6.1114	-2.2606	***
E - A	-2.8440	-4.7694	-0.9186	***
E - D	-0.4220	-2.3474	1.5034	

Tukey's MCP

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of Studentized Range	4.23186
Minimum Significant Difference	2.762

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
B - C	1.6540	-1.1080	4.4160	
B - A	2.9960	0.2340	5.7580	***
B - D	5.4180	2.6560	8.1800	***
B - E	5.8400	3.0780	8.6020	***
C - B	-1.6540	-4.4160	1.1080	
C - A	1.3420	-1.4200	4.1040	
C - D	3.7640	1.0020	6.5260	***
C - E	4.1860	1.4240	6.9480	***
A - B	-2.9960	-5.7580	-0.2340	***
A - C	-1.3420	-4.1040	1.4200	
A - D	2.4220	-0.3400	5.1840	
A - E	2.8440	0.0820	5.6060	***
D - B	-5.4180	-8.1800	-2.6560	***
D - C	-3.7640	-6.5260	-1.0020	***
D - A	-2.4220	-5.1840	0.3400	
D - E	0.4220	-2.3400	3.1840	
E - B	-5.8400	-8.6020	-3.0780	***
E - C	-4.1860	-6.9480	-1.4240	***
E - A	-2.8440	-5.6060	-0.0820	***
E - D	-0.4220	-3.1840	2.3400	

Sidak's MCP

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of t	3.14330
Minimum Significant Difference	2.9013

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
B - C	1.6540	-1.2473	4.5553	
B - A	2.9960	0.0947	5.8973	***
B - D	5.4180	2.5167	8.3193	***
B - E	5.8400	2.9387	8.7413	***
C - B	-1.6540	-4.5553	1.2473	
C - A	1.3420	-1.5593	4.2433	
C - D	3.7640	0.8627	6.6653	***
C - E	4.1860	1.2847	7.0873	***
A - B	-2.9960	-5.8973	-0.0947	***
A - C	-1.3420	-4.2433	1.5593	
A - D	2.4220	-0.4793	5.3233	
A - E	2.8440	-0.0573	5.7453	
D - B	-5.4180	-8.3193	-2.5167	***
D - C	-3.7640	-6.6653	-0.8627	***
D - A	-2.4220	-5.3233	0.4793	
D - E	0.4220	-2.4793	3.3233	
E - B	-5.8400	-8.7413	-2.9387	***
E - C	-4.1860	-7.0873	-1.2847	***
E - A	-2.8440	-5.7453	0.0573	
E - D	-0.4220	-3.3233	2.4793	

Bonferroni's MPC

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of t	3.15340
Minimum Significant Difference	2.9106

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
B - C	1.6540	-1.2566	4.5646	
B - A	2.9960	0.0854	5.9066	***
B - D	5.4180	2.5074	8.3286	***
B - E	5.8400	2.9294	8.7506	***
C - B	-1.6540	-4.5646	1.2566	
C - A	1.3420	-1.5686	4.2526	
C - D	3.7640	0.8534	6.6746	***
C - E	4.1860	1.2754	7.0966	***
A - B	-2.9960	-5.9066	-0.0854	***
A - C	-1.3420	-4.2526	1.5686	
A - D	2.4220	-0.4886	5.3326	
A - E	2.8440	-0.0666	5.7546	
D - B	-5.4180	-8.3286	-2.5074	***
D - C	-3.7640	-6.6746	-0.8534	***
D - A	-2.4220	-5.3326	0.4886	
D - E	0.4220	-2.4886	3.3326	
E - B	-5.8400	-8.7506	-2.9294	***
E - C	-4.1860	-7.0966	-1.2754	***
E - A	-2.8440	-5.7546	0.0666	
E - D	-0.4220	-3.3326	2.4886	

Fisher's LSD MCP

Means with the same letter are not significantly different.				
t Grouping		Mean	N	trt
	A	20.4260	5	B
	A			
B	A	18.7720	5	C
B				
B		17.4300	5	A
	C	15.0080	5	D
	C			
	C	14.5860	5	E

Tukey's MCP

Means with the same letter are not significantly different.				
Tukey Grouping		Mean	N	trt
	A	20.4260	5	B
	A			
B	A	18.7720	5	C
B				
B	C	17.4300	5	A
	C			
D	C	15.0080	5	D
D				
D		14.5860	5	E

Sidak's MCP

Means with the same letter are not significantly different.				
Sidak Grouping		Mean	N	trt
	A	20.4260	5	B
	A			
B	A	18.7720	5	C
B				
B	C	17.4300	5	A
	C			
	C	15.0080	5	D
	C			
	C	14.5860	5	E

Bonferroni's MPC

Means with the same letter are not significantly different.				
Bon Grouping		Mean	N	trt
	A	20.4260	5	B
	A			
B	A	18.7720	5	C
B				
B	C	17.4300	5	A
	C			
	C	15.0080	5	D
	C			
	C	14.5860	5	E

Dunnett's MCP (Two-Sided)

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of Dunnett's t	2.65103
Minimum Significant Difference	2.4469

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
B - C	1.6540	-0.7929	4.1009	
A - C	-1.3420	-3.7889	1.1049	
D - C	-3.7640	-6.2109	-1.3171	***
E - C	-4.1860	-6.6329	-1.7391	***

Dunnett's MCP (Lower One-Sided)

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of Dunnett's t	2.30443
Minimum Significant Difference	2.127

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
B - C	1.6540	-0.4730	Infinity	
A - C	-1.3420	-3.4690	Infinity	
D - C	-3.7640	-5.8910	Infinity	
E - C	-4.1860	-6.3130	Infinity	

Dunnett's MCP (Upper One-Sided)

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	2.12987
Critical Value of Dunnett's t	2.30443
Minimum Significant Difference	2.127

Comparisons significant at the 0.05 level are indicated by ***.				
trt Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
B - C	1.6540	-Infinity	3.7810	
A - C	-1.3420	-Infinity	0.7850	
D - C	-3.7640	-Infinity	-1.6370	***
E - C	-4.1860	-Infinity	-2.0590	***

R code for various MCPs

```
library(multcomp)
treatment <-
as.factor(c(rep("A",5),rep("B",5),rep("C",5),rep("D",5),rep("E",5)))
treatment
y <- c(19.08,17.07,18.91,15.09,17.00,22.04,21.44,18.82,20.49,19.34,
18.68,19.86,19.68,17.78,17.86,16.99,13.18,16.97,12.90,15.00,
15.34,13.52,15.23,15.63,13.21)

aovmodel <- aov(y~treatment)
summary(aovmodel)

# Run this line for any MCP
mcp <- glht(aovmodel,linfct=mcp(treatment="Tukey"))

# To output Fisher's LSD Test results
summary(mcp,test=univariate())

# To output Tukey's Test results
summary(mcp)

# To output Bonferroni's Test results
summary(mcp,test=adjusted(type="bonferroni"))

# By default the first level is the control. To make a different level
# the control level, redefine the treatment labels by putting
# an underscore at the beginning of the control level (e.g., "_C").

trt <- as.factor(c(rep("A",5),rep("B",5),rep("_C",5),rep("D",5),rep("E",5)))
trt
aovdunnett <- aov(y~trt)

# Dunnett's Test
dunnett2 <- glht(aovdunnett,linfct=mcp(trt="Dunnett"))
summary(dunnett2)
dunnettL <- glht(aovdunnett,linfct=mcp(trt="Dunnett"),alternative="less")
summary(dunnettL)
dunnettU <- glht(aovdunnett,linfct=mcp(trt="Dunnett"),alternative="greater")
summary(dunnettU)
```

R output for various MCPs

```
> # To output Fisher's LSD Test results
> summary(mcp,test=univariate())

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
B - A == 0    2.996      0.923   3.246 0.004048 **
C - A == 0    1.342      0.923   1.454 0.161477
D - A == 0   -2.422      0.923  -2.624 0.016259 *
E - A == 0   -2.844      0.923  -3.081 0.005890 **
C - B == 0   -1.654      0.923  -1.792 0.088282 .
D - B == 0   -5.418      0.923  -5.870 9.65e-06 ***
E - B == 0   -5.840      0.923  -6.327 3.56e-06 ***
D - C == 0   -3.764      0.923  -4.078 0.000586 ***
E - C == 0   -4.186      0.923  -4.535 0.000202 ***
E - D == 0   -0.422      0.923  -0.457 0.652455
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Univariate p values reported)
```

```
> # To output Tukey's Test results
> summary(mcp)

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
B - A == 0    2.996      0.923   3.246  0.02952 *
C - A == 0    1.342      0.923   1.454  0.60205
D - A == 0   -2.422      0.923  -2.624  0.10333
E - A == 0   -2.844      0.923  -3.081  0.04167 *
C - B == 0   -1.654      0.923  -1.792  0.40508
D - B == 0   -5.418      0.923  -5.870  < 0.001 ***
E - B == 0   -5.840      0.923  -6.327  < 0.001 ***
D - C == 0   -3.764      0.923  -4.078  0.00477 **
E - C == 0   -4.186      0.923  -4.535  0.00171 **
E - D == 0   -0.422      0.923  -0.457  0.99030
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


> # To output Bonferroni's Test results
> summary(mcp,test=adjusted(type="bonferroni"))

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
B - A == 0    2.996      0.923   3.246  0.04048 *
C - A == 0    1.342      0.923   1.454  1.00000
D - A == 0   -2.422      0.923  -2.624  0.16259
E - A == 0   -2.844      0.923  -3.081  0.05890 .
C - B == 0   -1.654      0.923  -1.792  0.88282
D - B == 0   -5.418      0.923  -5.870 9.65e-05 ***
E - B == 0   -5.840      0.923  -6.327 3.56e-05 ***
D - C == 0   -3.764      0.923  -4.078  0.00586 **
E - C == 0   -4.186      0.923  -4.535  0.00202 **
E - D == 0   -0.422      0.923  -0.457  1.00000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- bonferroni method)
```

```
> trt
 [1] A A A A A B B B B B _C _C _C _C _C D D D D D E E E E E
Levels: _C A B D E

> # Dunnett's Test

> summary(dunnett2)

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Dunnett Contrasts

Linear Hypotheses:                               "Two-sided" Dunnett
            Estimate Std. Error t value Pr(>|t|)
A - _C == 0   -1.342      0.923  -1.454  0.41925
B - _C == 0    1.654      0.923   1.792  0.25024
D - _C == 0   -3.764      0.923  -4.078  0.00223 **
E - _C == 0   -4.186      0.923  -4.535  < 0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


> summary(dunnettL)

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Dunnett Contrasts

Linear Hypotheses:                                "Less than" Dunnett
            Estimate Std. Error t value  Pr(<t)
A - _C >= 0   -1.342      0.923  -1.454 0.21286
B - _C >= 0    1.654      0.923   1.792 0.99743
D - _C >= 0   -3.764      0.923  -4.078 0.00116 **
E - _C >= 0   -4.186      0.923  -4.535 < 0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


> summary(dunnettU)

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Dunnett Contrasts

Linear Hypotheses:                                 "Greater than" Dunnett
            Estimate Std. Error t value Pr(>t)
A - _C <= 0   -1.342      0.923  -1.454  0.993
B - _C <= 0    1.654      0.923   1.792  0.126
D - _C <= 0   -3.764      0.923  -4.078  1.000
E - _C <= 0   -4.186      0.923  -4.535  1.000
(Adjusted p values reported -- single-step method)
```