AMS 315
Data Analysis
Chapter Nine Study Guide
Multiple Comparisons
Spring 2023

Context

The procedures in this chapter provide tools to deal with the inflation of the level of significance due to multiple testing.

Chapter Nine

9.1. Introduction and Abstract of Research Study

The authors give a good introduction to the issues of multiple testing.

9.2 Linear Contrasts

Hypotheses are typically stated using linear contrasts. Definition 9.1 specifies what a contrast is. Definition 9.2 deals with orthogonal contrasts. Since estimators corresponding to orthogonal contrasts are independent when dealing with normally distributed data, orthogonal contrasts are useful in statistical analyses. In a one-way analysis of variance with t treatments, there are $t-1$ orthogonal contrasts, one for each degree of freedom in the between group sum of squares. Note the formula for the sum of squares for a contrast. Each of the $t-1$ orthogonal contrasts has a sum of squares based on 1 degree of freedom. The sum of these $t-1$ sums of squares is equal to the between group sum of squares. There is an F-test for the null hypothesis that each linear contrast is equal to zero; this F-test has 1 numerator and $n-t$ denominator degrees of freedom.

9.3 Which Error Rate Is Controlled

This section discusses the Bonferroni (Boole) inequality, which is the simplest tool for dealing with multiple comparisons. To maintain an experiment-wide level of significance of α_E when there are m tests, test each individual hypothesis at the α_E / m level of significance.

Fisher's Least Significant Difference (not in text)

Carmer and Swanson (1973) reported that this was the most powerful approach to dealing with the multiple comparisons issue in analyses of variance. The approach is first to test the null hypothesis that all treatment means are equal in an analysis of variance. If this hypothesis is accepted, there is no further comparison of individual means. That is, the overall test of equality of treatment means is serving as protection for individual comparisons. If the null hypothesis is rejected, then compare two treatment means with a t-test using the mean square for error as the estimate of the common treatment variance.

9.4. Scheffe's S Method

Although this method is slightly less powerful than the Tukey W procedure, it generalized to regression problems very naturally. You should master it.

9.5 Tukey's W Procedure

This is the most powerful procedure for comparing two treatment means in a one-way analysis of variance situation. It was covered in class. Examinations may include questions on this procedure. It is based on the Studentized range of a sample from a normal distribution.

9.6. Dunnett's Procedure: Comparison of Treatments to a Control

This was not covered in lectures. It is an effective procedure for dealing with an important problem in the analysis of clinical trials.

9.7 A Nonparametric Multiple Comparison Procedure

This was not covered in lecture, and I will not ask about it in examinations.

9.8 Research Study: Are Interviewers' Decisions Affected by Different Handicap Types

The case study develops issues that occur frequently in scientific studies of the evaluation of participants.

9.9 Summary and Key Formulas

This is fundamental review material.

Example Past Examination Questions

1. A research team studied *Y*, the blemished area on the skin of a laboratory animal, and how *Y* was affected by the dose of medicine. They used three doses of medicine: 0, 1, and 2 units respectively. They randomly assigned 20 animals to each dosage and observed that the average values of *Y* at each dosage were $y_{0\bullet} = 26.3, y_{1\bullet} = 23.7$, and $y_{2\bullet} = 15.8$ where $y_{i.}$ was the average of the observations taken with dosages $i = 0,1,2$, respectively. They also observed that

$s_0^2 = 0.72, s_1^2 = 0.96$, and $s_2^2 = 0.82$ where s_i^2 was the unbiased estimate of the variance for the observations taken with dosages $i = 0,1,2$, respectively.

a. Complete the analysis of variance table for these results; that is, be sure to specify the degrees of freedom, sum of squares, mean square, and F-test. Test the null hypothesis that the expected blemish areas are the same for the three dosages at the 0.10, 0.05, and 0.01 levels of significance.
b. Find the sum of squares due to the linear and quadratic contrasts. Test whether they are significant at the 0.10, 0.05, and 0.01 levels of significance. The coefficients of the linear contrast are $-1,0,1$; and the coefficients of the quadratic contrast are $1,-2,1$.
c. Find the 99% Scheffe confidence intervals for the linear and quadratic contrasts.
d. What is your recommendation of dosage setting to minimize the expected blemish area?

Solution:

a. Analysis of Variance Table

Source	DF	SS	MS	F
Treatment	2	1196.134	598.1	720.6
Error	57	47.5	0.83	
Total	59	1243.634		

Since $F_{\alpha=0.1,2,57} = 2.398$, $F_{\alpha=0.05,2,57} = 3.159$, and $F_{\alpha=0.01,2,57} = 4.998$, reject H0 at each level of significance.

b. The estimated linear contrasts are $\hat{\lambda}_{Linear} = -10.5$ and $\hat{\lambda}_{Quadratic} = -5.3$ with $SS_{Linear} = 1102.5$ and $SS_{Quadratic} = 93.63$. Then, $F_{Linear} = 1328.3$, and $F_{Quadratic} = 112.8$. Since $F_{\alpha=0.1,1,57} = 2.796$, $F_{\alpha=0.05,1,57} = 4.010$, and $F_{\alpha=0.01,1,57} = 7.102$, reject the null hypothesis that the linear contrast is 0 against the two-sided alternative that it is not equal to 0 at each level of significance. Similarly, reject the null hypothesis that the quadratic contrast is 0 against the two-sided alternative that it is not equal to 0 at each level of significance. Note that the critical F value for $\alpha = 10^{-14}$ is 107.40. Even after allowing for the multiple comparisons, both tests are highly significant. Also, note the confidence intervals calculated in part c.

c. The 99% Scheffe confidence interval for λ_{Linear} is -10.5 ± 0.91, which excludes 0. The 99% Scheffe confidence interval for $\lambda_{Quadratic}$ is -5.3 ± 1.58, which also excludes 0.

d. The best observed result was for the dosage of 2 units. The average for 2 units was significantly lower than the average for 1 unit as documented by the 99% Scheffe confidence interval for $E(Y_{2j}) - E(Y_{1j'})$, which was -7.9 ± 0.91. Since the observed averages decreased with increasing dosage, researchers should consider studying the response to dosages higher than 2 units.

2. A research team wishes to specify a manufacturing process so that *Y*, the area in a product affected by surface flaws is as small as possible. They have four levels of concentration of a chemical used to wash the product before the final manufacturing step and want to determine whether the concentration level causes a change in $E(Y)$. They run a balanced one-way layout with 10 observations for

each concentration with level 1 set at 10%, level 2 set at 15%, level 3 set at 20%, and level 4 set at 25%. They run a balanced one-way layout with 10 observations for each treatment. They observe that

$y_{1.} = 14.5,\ y_{2.} = 15.9,\ y_{3.} = 14.2,$ and $y_{4.} = 13.8,$ where $y_{i.}$ is the average of the observations taken on the *i*th level. They also observe that

$s_1^2 = 0.98,\ s_2^2 = 1.23,\ s_3^2 = 0.87,$ and $s_4^2 = 0.85,$ where s_i^2 is the unbiased estimate of the variance for the observations taken on the *i*th level.

a. Complete the analysis of variance table for these results; that is, be sure to specify the degrees of freedom, sum of squares, mean square, and F-test .
b. State the 99% Scheffe confidence interval for $E(Y_{4j} - Y_{3j'})$.
c. Partition the sum of squares due to the concentrations into sum of squares due to the linear contrast, the quadratic contrast, and the cubic contrast. The vector of coefficients for the linear contrast is $(-3,-1,1,3)$, for the quadratic ($(1,-1,-1,1)$, and for the cubic $(-1,3,-3,1)$. Evaluate the importance of each of these sums of squares. That is, test the significance of each using the 0.10, 0.05, and 0.01 levels of significance.
d. What is your conclusion? Make sure that you discuss the optimal setting of the concentration level.

Solution:

a. Analysis of variance table

Source	DF	SS	MS	F
Treatment	3	25	8.33	8.48
Error	36	35.37	0.9825	
Total	39	60.37		

Since $F_{.01,3,36} = 4.377$, reject the null hypothesis that the four treatment means are equal at the 0.01 level of significance.

b. The 99% Scheffe confidence interval for $E(Y_{4j} - Y_{3j'})$ is (-2.01, 1.21)

c. The estimated linear contrasts are $\hat{\lambda}_{Linear} = -3.8$, $\hat{\lambda}_{Quadratic} = -1.8$, and $\hat{\lambda}_{Cubic} = 4.4$. The decomposition of the sum of squares is $SS_{Linear} = 7.22$, $SS_{Quadratic} = 8.1$, and $SS_{Cubic} = 9.68$ Then, $F_{Linear} = 7.3486$, $F_{Quadratic} = 8.244$. and $F_{Cubic} = 9.85$ Since $F_{\alpha=0.1,1,36} = 2.850$, $F_{\alpha=0.05,1,36} = 4.113$, $F_{\alpha=0.01,1,36} = 7.396$, $F_{\alpha=0.005,1,36} = 8.943$, and $F_{\alpha=0.001,1,36} = 12.832$, reject the null hypothesis that the linear contrast is 0 against the two-sided alternative that it is not equal to 0 at the .05 but not .01 level of significance. Reject the null hypothesis that the quadratic contrast is 0 against the two-sided alternative that it is not equal to 0 at the 0.01 level of significance. Similarly, reject the null hypothesis that the cubic contrast is 0 at the 0.01 level of significance. The p-value for the linear test is 0.0102; after Bonferroni adjustment (i.e., multiplying the p-value by 3), the Bonferroni corrected p-value is 0.0306. The conclusion should be tempered to rejection of the linear contrast hypothesis at the 0.05 level adjusting for the three tests made. Similarly, the Bonferroni adjustment of the quadratic p-value (0.0068) is .0204, and the Bonferroni adjustment of the cubic p-value ((0.00338) is 0.0101.

d. The concentration of the washing chemical is associated with the extent of flawed area as shown by the result of section a. The lowest average occurred with concentration 25%, but the average with 25% concentration was not significantly different from the average using 20% concentration (see result b). The strength of the decreasing pattern suggests that researchers examine the effect of using concentrations even larger than 25%.