AMS 315
Data Analysis
Chapter Nine Lecture Notes
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. We discussed the protected-t confidence intervals in the lecture notes for Chapter 8. This procedure is often called Fisher's Least Significant Difference (LSD) and was discussed in earlier editions of your text. Carmer and Swanson [Carmer, SG; Swanson, MR (1973). An evaluation of ten pairwise multiple comparison procedures for Monte Carlo Methods. Journal of the American Statistical Association, 68, 66-74. Available on JSTOR] 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.

Bonferroni's (Boole's) Inequality

Let R_1 be the event that null hypothesis one (H_{01}) is rejected, and let R_2 be the event that null hypothesis one (H_{02}) is rejected. Then from Chapter 4, $\Pr\{R_1 \cup R_2\} = \Pr\{R_1\} + \Pr\{R_2\} - \Pr\{R_1 \cap R_2\}$. When R_1 and R_2 are independent, $\Pr\{R_1 \cap R_2\} = \Pr\{R_1\} \times \Pr\{R_2\}$. Otherwise, it may be difficult to calculate $\Pr\{R_1 \cap R_2\}$. Boole's inequality compares $\Pr\{R_1 \cup R_2\} = \Pr\{R_1\} + \Pr\{R_2\} - \Pr\{R_1 \cap R_2\}$ to $\Pr\{R_1\} + \Pr\{R_2\}$. The result is $\Pr\{R_1 \cup R_2\} \le \Pr\{R_1\} + \Pr\{R_2\}$. Although this result was found by Boole, virtually all researchers refer to is as Bonferroni's inequality.

In statistics, $\Pr_0\{R_1 \cup R_2\} = \Pr_0\{\text{At least one Type I error}\} = \alpha_{\text{Overall}}$. It is natural to let $\alpha_1 = \Pr_0\{R_1\}$ and let $\alpha_2 = \Pr_0\{R_2\}$. Then Boole's (Bonferroni's) inequality is that $\alpha_{\text{Overall}} \le \alpha_1 + \alpha_2$. The bound on the overall level of significance is the sum of the levels of significance of each individual test. This property is described as inflation of the significance level.

Let R_3 be the event that null hypothesis three (H_{03}) is rejected with level of significance α_3, …, and let R_N be the event that null hypothesis N (H_{0N}) is rejected with level of significance α_N. Then the inequality generalizes, and

$\Pr_0\{R_1 \cup \cdots \cup R_N\} = \Pr_0\{\text{At least one Type I error}\} = \alpha_{\text{Overall}} \le \alpha_1 + \cdots + \alpha_N$.
Conventionally, researchers set the individual levels of significance to be less than some relatively small value $\varepsilon > 0$ such as 0.01 or 0.05. Then, the inequality is
$\Pr_0\{R_1 \cup \cdots \cup R_N\} = \Pr_0\{\text{At least one Type I error}\} = \alpha_{\text{Overall}} \le \alpha_1 + \cdots + \alpha_N \le \varepsilon.$
Researchers may set $\alpha_1 = \cdots \alpha_N = \alpha$ so that
$\Pr_0\{R_1 \cup \cdots \cup R_N\} = \Pr_0\{\text{At least one Type I error}\} = \alpha_{\text{Overall}} \le N\alpha \le \varepsilon.$ That is, researchers set each individual significance level to $\frac{\varepsilon}{N}$. Equivalently, in a multiple regression analysis, one can multiply the p-value of a regression coefficient by *N*, the number of independent variables considered.

9.2 Linear Contrasts

Hypotheses are typically stated using contrasts linear in the treatment means. For example, in a balanced one-way layout with $I = 4$, let
$\lambda = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4$, where $\mu_i = E(Y_{ij}) = E(Y_{i\bullet})$. The $a_i,\ i = 1,\ldots,I$ are fixed constants with $\sum_{i=1}^{I} a_i = 0$. Here $\lambda = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4$ is a linear combination of parameters and hence is itself a parameter. We need a statistic $\hat{\lambda}$ such that $E(\hat{\lambda}) = \lambda.$
That statistic is $\hat{\lambda} = a_1Y_{1\bullet} + a_2Y_{2\bullet} + a_3Y_{3\bullet} + a_4Y_{4\bullet}$ The expected value is
$E(\hat{\lambda}) = a_1E(Y_{1\bullet}) + a_2E(Y_{2\bullet}) + a_3E(Y_{3\bullet}) + a_4E(Y_{4\bullet}) = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4 = \lambda$. The variance is $\text{var}(\hat{\lambda}) = a_1^2 \text{ var}(Y_{1\bullet}) + a_2^2 \text{ var}(Y_{2\bullet}) + a_3^2 \text{ var}(Y_{3\bullet}) + a_4^2 \text{ var}(Y_{4\bullet}) = \sum_{i=1}^{I} a_i^2 \frac{\sigma_{1W}^2}{J}$, where *J* is the number of observations in each treatment. For example,
$\lambda_L = (-3)\mu_1 + (-1)\mu_2 + (1)\mu_3 + (3)\mu_4$ is a contrast that is called the “linear contrast.” A vector space description of the linear contrast would be as the inner product of two vectors (-3, -1, 1, 3) and $(\mu_1, \mu_2, \mu_3, \mu_4)$

The statistic estimating the linear contrast is $\hat{\lambda}_L = (-3)Y_{1\bullet} + (-1)Y_{2\bullet} + (1)Y_{3\bullet} + (3)Y_{4\bullet}$. Then
$E(\hat{\lambda}_L) = \lambda_L$, and $\text{var}(\hat{\lambda}_L) = \sum_{i=1}^{I} a_i^2 \frac{\sigma_{1W}^2}{J} = [(-3)^2 + (-1)^2 + (1)^2 + (3)^2]\frac{\sigma_{1W}^2}{J} = \frac{20\sigma_{1W}^2}{J}.$

The sum of squares due to a contrast is $SS_\lambda = \frac{(\hat{\lambda})^2}{\sum_{i=1}^{I} \frac{a_i^2}{J}}$. For example, the sum of squares due to the linear contrast for $I = 4$ is $SS_L = \frac{(\hat{\lambda}_L)^2}{20/J}$, with one degree of freedom.

The set of linear contrasts in I mean parameters is a vector space with dimension $I-1$. Another contrast for a balanced one way layout with $I=4$ is $\lambda_Q=(1)\mu_1+(-1)\mu_2+(-1)\mu_3+(1)\mu_4$, which is called the "quadratic contrast." Of course, the quadratic contrast is the inner product of the vector (1,-1,-1,1) and $(\mu_1,\mu_2,\mu_3,\mu_4)$. The inner product of the vector of coefficients of the linear contrast (-3, -1, 1, 3) and the vector of coefficients of the quadratic contrast (1,-1,-1,1) is 0. Considered as vectors, these two contrasts are orthogonal. The statistic estimating the quadratic contrast is $\hat{\lambda}_Q=(1)Y_{1\bullet}+(-1)Y_{2\bullet}+(-1)Y_{3\bullet}+(1)Y_{4\bullet}$, and

$$SS_Q=\frac{(\hat{\lambda}_Q)^2}{[1^2+(-1)^2+(-1)^2+(1)^2]/J}=\frac{(\hat{\lambda}_Q)^2}{4/J},$$ with one degree of freedom.

The dimension of the set of linear contrasts in 4 mean parameters is still a vector space with dimension $4-1=3$, so that there is a third orthogonal contrast $\lambda_C=(-1)\mu_1+(3)\mu_2+(-3)\mu_3+(1)\mu_4$, called the cubic contrast. The cubic contrast is the inner product of the vector (-1,3,-3,1) and $(\mu_1,\mu_2,\mu_3,\mu_4)$. The statistic estimating the cubic contrast is $\hat{\lambda}_Q=(-1)Y_{1\bullet}+(3)Y_{2\bullet}+(-3)Y_{3\bullet}+(1)Y_{4\bullet}$, and

$$SS_C=\frac{(\hat{\lambda}_C)^2}{[(-1)^2+(3)^2+(-3)^2+(1)^2]/J}=\frac{(\hat{\lambda}_C)^2}{20/J},$$ with one degree of freedom.

Each pair of these three contrasts is orthogonal. Each has an associated sum of squares based on 1 degree of freedom, and $SS_L+SS_Q+SS_C=SS_{Treatments}$.

Lack of fit test

There was a problem in the discussion of Chapter 8 that I wish to complete. That problem was:

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 6 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 6 observations for each treatment. They observe that $y_{1.}=264.5$, $y_{2.}=255.9$, $y_{3.}=216.2$, and $y_{4.}=263.8$, where $y_{i.}$ is the average of the observations taken on the *i*th level. They also observe that $s_1^2=411.9$, $s_2^2=522.2$, $s_3^2=631.8$, and $s_4^2=521.9$, 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. What is your conclusion? Use significance levels set to 0.10, 0.05, and 0.01. Make sure that you discuss the optimal setting of the concentration level and how you could document it.

Answer:

Analysis of Variance Table

Source	Sum of Squares	Degrees of Freedom	Mean Square	
Treatments	9467.4	3	3155.8	F=6.046
Error	10439.0	20	521.95	
Total	19906.4	23		

The critical values for the F-test are 2.38 for the 0.10 level, 3.10 for the 0.05 level, and 4.94 for the 0.01 level. Reject the null hypothesis that the mean responses of the four treatments are equal. The optimum setting is to use 20% as the concentration setting. This can be confirmed with Fisher's protected confidence intervals.

New questions:

c. Compute the sums of squares for the linear, quadratic, and cubic contrasts.
d. Complete the analysis of variance table for the linear regression of the dependent variable (area affected by surface flaws) on the chemical concentration. That is, use the sum of squares for the linear contrast as the sum of squares due to the linear regression. Test the null hypothesis that the dependent variable has no linear association with the independent variable. Use the 0.10, 0.05, and 0.01 levels of significance.
e. Compute the sum of squares for lack of fit for the linear regression model. Complete the analysis of variance table that includes the sum of squares due to the linear regression, the sum of squares due to lack of fit for the linear regression, and the sum of squares for pure error. What is your conclusion? Use the 0.10, 0.05, and 0.01 levels of significance.

Answers:

c. For the contrasts:

$$\hat{\lambda}_L = (-3)Y_{1\bullet} + (-1)Y_{2\bullet} + (1)Y_{3\bullet} + (3)Y_{4\bullet} = (-3)\times 264.5 + (-1)\times 255.9 + (1)\times 216.2 + (3)\times 263.8 = -41.8;$$

$$\hat{\lambda}_Q = (1)Y_{1\bullet} + (-1)Y_{2\bullet} + (-1)Y_{3\bullet} + (1)Y_{4\bullet} = (1)\times 264.5 + (-1)\times 255.9 + (-1)\times 216.2 + (1)\times 263.8 = 56.2;$$

$$\hat{\lambda}_C = (-1)Y_{1\bullet} + (3)Y_{2\bullet} + (-3)Y_{3\bullet} + (1)Y_{4\bullet} = (-1)\times 264.5 + (3)\times 255.9 + (-3)\times 216.2 + (1)\times 263.8 = 118.4.$$

Then the sums of squares of the contrasts are:

$SS_L = \frac{(\hat{\lambda}_L)^2}{20/J} = \frac{(-41.8)^2}{20/6} = 524.172$; $SS_Q = \frac{(\hat{\lambda}_Q)^2}{4/J} = \frac{(56.2)^2}{4/6} = 4737.66$; and

$SS_C = \frac{(\hat{\lambda}_C)^2}{20/J} = \frac{(118.4)^2}{20/6} = 4205.568$. An important check is to calculate

$SS_L + SS_Q + SS_C = 524.172 + 4737.66 + 4205.568 = 9467.4 = SS_{Treatments}$.

d. The analysis of variance table based on the calculations of Chapter 11 is:

Analysis of Variance Table
Linear Regression

Source	Sum of Squares	Degrees of Freedom	Mean Square	
Regression on Chemical Concentration	524.172	1	524.172	F=0.595
Error	19382.228	22	881.010	
Total	19906.4	23		

Since the Chapter 11 F-test has the value 0.595, which is less than 1, we will accept the null hypothesis that there is no linear relation between dependent and independent variables. More formally, the critical value for a central F distribution with 1 numerator and 22 denominator degrees of freedom and level of significance 0.10 is 2.949. Since the F statistic (which has value 0.595) is less than 2.949, the null hypothesis is accepted at the 0.10 level (and of course at the 0.05 and 0.01 levels).

For part e, the lack of fit test analysis of variance table is below:

Analysis of Variance Table
Linear Regression and Lack of Fit Sums of Squares

Source	Sum of Squares	Degrees of Freedom	Mean Square	F test
Regression on Chemical Concentration	524.172	1	524.172	
Lack of (linear) fit	8943.228	2	4471.614	8.57
Pure Error	10439.0	20	521.95	
Total	19906.4	23		

There are two ways of calculating the sum of squares for lack of fit. The first is to add the sum of squares for the quadratic contrast and the sum of squares for the cubic contrast (that is, $4737.66 + 4205.568 = 8943.228$ with 2 degrees of freedom). The second is to subtract the sum of squares for the linear contrast from the sum of squares for treatments (that is, $9467.4 - 524.172 = 8943.228$ with 3-1 degrees of freedom). The lack of fit null hypothesis is that the linear model is adequate. The test statistic is the ratio of the mean square for lack of fit divided by the mean square for pure error. When the linear model is adequate, this test statistic should be equal to 1, modulo statistical variability. When the test statistic is larger than one, the linear model is not adequate. That is a more complex model is required. Here that more complex model may have both a linear and quadratic term.

Here, $F_{LOF} = \frac{MS_{LOF}}{MS_{PE}} = \frac{4471.614}{521.95} = 8.57$, with 2 numerator and 20 denominator degrees of freedom. The critical value for a central F distribution with 2 numerator and 20 denominator degrees of freedom and level of significance 0.01 is 5.850. Since $F_{LOF} = \frac{MS_{LOF}}{MS_{PE}} = 8.57 > 5.850$, we reject the null hypothesis that the linear model is adequate at the 0.01 level.

Comments

One should always take advantage of observations from repeated independent variable values. The lack of fit test is an objective data-based indicator of the adequacy of a model. A research team should design its study so that there are repeated independent variable values. One can also group observations into sets that have approximately the same independent variable values so that one can calculate an approximate lack of fit test.

We can continue our example with the techniques from Chapter 12. The linear model is that $Y_i = \beta_0 + \beta_1 x_i + \sigma_{L|x} Z_i$. We accepted $H_{0L} : \beta_1 = 0$ rather than $H_{1L} : \beta_1 \neq 0$ because the relevant F statistic was 0.595, and we rejected the null hypothesis that the linear model was adequate (because the lack of fit F statistic was 8.57, significant at the 0.01 level). It is natural to consider a quadratic model $Y_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \sigma_{Q|x} Z_i$. The analysis of variance table is shown below. The F-test for the quadratic coefficient given that the linear coefficient is in the model is $F_{Q|L} = \frac{SS_Q/1}{MSE} = \frac{4737.66}{697.36} = 6.79$ with 1 numerator and 21 denominator degrees of freedom. The critical values are 2.961 for the 0.10 level, 4.325 for the 0.05 level and 8.017 for the 0.01 level. We would reject $H_{0Q} : \gamma_2 = 0$ rather than $H_{1Q} : \gamma_2 \neq 0$ at the 0.10 and 0.05 levels. We would accept $H_{0Q} : \gamma_2 = 0$ at the 0.01 level.

Analysis of Variance Table
Linear and Quadratic Regression

Source	Sum of Squares	Degrees of Freedom	Mean Square	F test
Linear Regression on Chemical Concentration	524.172	1	524.172	
Quadratic regression given linear regression	4737.66	1	4737.66	6.79
Error	14644.568	21	697.36	
Total	19906.4	23		

In the event that we use the linear and quadratic model, we can test for lack of fit. The analysis of variance table is shown below.

Analysis of Variance Table
Linear and Quadratic Regression and Lack of Fit Sums of Squares

Source	Sum of Squares	Degrees of Freedom	Mean Square	F test
Linear Regression on Chemical Concentration	524.172	1	524.172	
Quadratic regression given linear regression	4737.66	1	4737.66	
Lack of (quadratic) fit	4205.568	1	4205.568	8.057
Pure Error	10439.0	20	521.95	
Total	19906.4	23		

The lack of fit F-test for the quadratic model is
$F_{Q\,LOF} = \frac{SS_{LOF\,Q}/1}{MSE} = \frac{4205.568}{521.95} = 8.057$ with 1 numerator and 20 denominator degrees of freedom. The critical values are 2.975 for the 0.10 level, 4.351 for the 0.05 level and 8.096 for the 0.01 level. We would reject the adequacy of the quadratic model at the 0.10 and 0.05 levels. We would accept the adequacy of the quadratic model at the 0.01 level (just barely).

Scheffe Confidence Intervals for All Contrasts

In a balanced one-way layout with $I = 4$, let $\lambda = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4$, where $\mu_i = E(Y_{ij}) = E(Y_{i\bullet})$. The $a_i,\ i = 1,\ldots,I$ are fixed constants with $\sum_{i=1}^{I} a_i = 0$. Here $\lambda = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4$ is a linear combination of parameters and hence is itself a parameter. We need a statistic $\hat{\lambda}$ such that $E(\hat{\lambda}) = \lambda$.
That statistic is $\hat{\lambda} = a_1 Y_{1\bullet} + a_2 Y_{2\bullet} + a_3 Y_{3\bullet} + a_4 Y_{4\bullet}$ The expected value is $E(\hat{\lambda}) = a_1 E(Y_{1\bullet}) + a_2 E(Y_{2\bullet}) + a_3 E(Y_{3\bullet}) + a_4 E(Y_{4\bullet}) = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4 = \lambda$. The variance is $\operatorname{var}(\hat{\lambda}) = a_1^2 \operatorname{var}(Y_{1\bullet}) + a_2^2 \operatorname{var}(Y_{2\bullet}) + a_3^2 \operatorname{var}(Y_{3\bullet}) + a_4^2 \operatorname{var}(Y_{4\bullet}) = \sum_{i=1}^{4} a_i^2 \frac{\sigma_{1W}^2}{J}$, where J is the number of observations in each treatment. A 99% confidence interval for λ, with σ_{1W} known, is then $\hat{\lambda} \pm 2.576\sqrt{\sum_{i=1}^{4} a_i^2 \frac{\sigma_{1W}^2}{J}}$. As usual, σ_{1W}^2 is not known, and we use the *MSE* to estimate it. The 99% confidence interval for λ, with σ_{1W} unknown, is then $\hat{\lambda} \pm t_{2.576,n-I}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}}$, where *MSE* has $n - I = n - 4$ degrees of freedom.

This calculation assumes that there is only one contrast that we want a confidence interval for. Scheffe's method gives confidence intervals for all contrasts. Mathematically, the set of contrasts using $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ is a vector space with dimension $4 - 1 = 3$. The 99% Scheffe confidence interval for λ, with σ_{1W} unknown, is $\hat{\lambda} \pm \sqrt{3F_{0.99,3,n-4}}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}}$, where *MSE* has $n - 4$ degrees of freedom. We can calculate a 99% Scheffe confidence interval for any linear contrast. The confidence is then 99% that all of the confidence intervals are simultaneously correct.

For the example that we have been studying, $\sqrt{3F_{0.99,3,24-4}} = \sqrt{(3 \times 4.938)} = 3.849$. The 99% Scheffe confidence interval for
$\lambda_L = (-3)\mu_1 + (-1)\mu_2 + (1)\mu_3 + (3)\mu_4$ is
$\hat{\lambda}_L \pm \sqrt{3F_{0.99,3,20}}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}} = -41.8 \pm \sqrt{14.814}\sqrt{\frac{20 \times 521.95}{6}} = -41.8 \pm 160.55.$ Similarly, the 99% Scheffe confidence interval for
$\lambda_Q = (1)\mu_1 + (-1)\mu_2 + (-1)\mu_3 + (1)\mu_4$ is
$\hat{\lambda}_Q \pm \sqrt{3F_{0.99,3,20}}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}} = 56.2 \pm \sqrt{14.814}\sqrt{\frac{4 \times 521.95}{6}} = 56.2 \pm 71.80.$ The 99% Scheffe confidence interval for
$\lambda_C = (-1)\mu_1 + (3)\mu_2 + (-3)\mu_3 + (1)\mu_4$ is

$\hat{\lambda}_C \pm \sqrt{3F_{0.99,3,20}}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}} = 118.4 \pm \sqrt{14.814}\sqrt{\frac{20\times 521.95}{6}} = 118.4 \pm 160.55.$ The confidence that these three intervals are simultaneously correct is 99%. There is no need to adjust the significance level for the multiple comparisons made, as in the Bonferroni inequality.

One can form Scheffe confidence intervals for any number of contrasts. For example, the 99% Scheffe confidence interval for $\mu_3 - \mu_2$ is

$$y_{3\bullet} - y_{2\bullet} \pm \sqrt{3F_{0.99,3,20}}\sqrt{\sum_{i=1}^{4} a_i^2 \frac{MSE}{J}} = 216.2 - 255.9 \pm \sqrt{14.814}\sqrt{\frac{2MSE}{6}} = -39.7 \pm 50.7.$$

This Scheffe interval includes zero, while the protected t-confidence interval excluded zero. This is an example of the greater strictness of the Scheffe confidence interval.

Tukey's W procedure

To use this procedure, one first calculates the Tukey sampling margin of error $W = q_\alpha(I,\nu)\sqrt{\frac{MSE}{J}}$, where I is the number of treatments in the one-way layout, $\nu = n - I$ is the degrees of freedom of the MSE, and $q_\alpha(I,\nu)$ is a percentile extracted from Table 10 (percentage points of the Studentized range). For the current example with $\alpha = 0.01,\ W = q_\alpha(I,\nu)\sqrt{\frac{MSE}{J}} = 5.02\sqrt{\frac{521.95}{6}} = 46.82.$ The 99% Tukey confidence interval for $\mu_3 - \mu_2$ is $y_{3\bullet} - y_{2\bullet} \pm W = 216.2 - 255.9 \pm 46.82 = -39.7 \pm 46.82.$ Since this includes zero, we would conclude that the expected value of the second treatment is equal to the expected value of the third treatment at the 0.01 level of significance. Similarly, the 99% Tukey confidence interval for $\mu_3 - \mu_1$ is $y_{3\bullet} - y_{1\bullet} \pm W = 216.2 - 264.5 \pm 46.82 = -48.3 \pm 46.82.$ The 99% Tukey confidence interval for $\mu_3 - \mu_1$ excludes zero, suggesting that these two treatments have different expected values. In general, Tukey confidence intervals for $\mu_i - \mu_{i'}$ are narrower than the Scheffe confidence intervals.

Working-Hotelling Confidence Band for $E[\hat{Y}(x)]$

An important Chapter 11 result was that $\hat{Y}(x)$ has the normal distribution $N(\beta_0 + \beta_1 x, \sigma_{Y|x}^2(\frac{1}{n} + \frac{(x-\bar{x}_n)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}))$.

The 95% confidence interval for $E[\hat{Y}(x)] = \beta_0 + \beta_1 x$, where x is a single value is then $\hat{Y}(x) \pm t_{1.960,n-2}\sqrt{MSE(\frac{1}{n} + \frac{(x-\bar{x}_n)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2})}$. The same multiple comparison issues hold when one seeks confidence intervals for $\hat{Y}(x)$ for multiple values of *x*. The application of the theory of the Scheffe confidence intervals to the problem of finding confidence intervals for multiple independent variable values generates the "Working-Hotelling Confidence Region for $E[\hat{Y}(x)]$.

This region is $\hat{Y}(x) \pm \sqrt{2F_{0.05,2,n-2}} \sqrt{MSE(\frac{1}{n} + \frac{(x-\bar{x}_n)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2})}, -\infty < x < \infty.$ There is then 95% confidence that the regression line is completely covered by this region.