Comparisons for all pairs using Tukey-Kramer HSD

Confidence Quantile

q*	Alpha
2.36773	0.05

HSD Threshold Matrix

Abs(Dif)-HSD

	13-18 Years Old	8-12 Years Old	Over 18 Years Old
13-18 Years Old	-10.601	-1.121	26.759
8-12 Years Old	-1.121	-10.601	17.279
Over 18 Years Old	26.759	17.279	-10.601

Positive values show pairs of means that are significantly different.

Connecting Letters Report

Level		Mean
13-18 Years Old	A	137.48000
8-12 Years Old	A	128.00000
Over 18 Years Old	B	100.12000

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	-Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
13-18 Years Old	Over 18 Years Old	37.36000	4.477358	26.7588	47.96118	< .0001*
8-12 Years Old	Over 18 Years Old	27.88000	4.477358	17.2788	38.48118	< .0001*
Over 18 Years Old	8-12 Years Old	9.48000	4.477358	-1.1212	20.08118	0.0898

Figure 15.2.2

It is interesting to note that the two multiple comparison procedures (Fisher's LSD and Tukey's HSD) in the aforementioned example do not produce the same results. This is not uncommon. Specifically, as the number of comparisons increases, the probability of committing a Type I Error increases when using Fisher's LSD. However, regardless of the number of comparisons being made, the probability of committing a Type I Error remains the same (i.e., α) when using Tukey's HSD.

Note

Note that the confidence intervals in the Ordered Differences Report vary slightly from our previous calculations due to the exact q-distribution critical value being used instead of the value from the table with error $df = \infty$.

15.2 Exercises

Basic Concepts

1. What is the purpose of multiple comparison procedures?
2. When should multiple comparison procedures be used?
3. What are the hypotheses tested if there are four population means in the ANOVA?
4. Define the concepts of balanced and unbalanced data when conducting a test to compare the pairwise sample means for a given set of samples.

Exercises

5. How many individual pairwise comparisons would need to be made if there are four population means in the ANOVA? What would be the probability of at least one Type I error if performing individual pairwise comparisons at a 0.01 significance level?
6. A two-sample t-test is conducted to test the pairwise differences in the mean number of candies consumed per family (average size of four family members) per day. The families belong to four different states. The following output is obtained (differences are computed in the given order of the states).

	Null hypothesis	Difference in Means	*t*-Test Statistic	*P*-value
Alabama and Los Angeles	$(H_0: \mu_A - \mu_{LA} = 0)$	5.6378	2.3479	0.046835
New York and Los Angeles	$(H_0: \mu_{NY} - \mu_{LA} = 0)$	−12.5798	7.8741	0.000049
Alabama and Texas	$(H_0: \mu_A - \mu_T = 0)$	41.2156	12.3721	0.000002
New York and Texas	$(H_0: \mu_{NY} - \mu_T = 0)$	0.4132	1.1553	0.281305
Texas and Los Angeles	$(H_0: \mu_T - \mu_{LA} = 0)$	−24.8714	9.2496	0.000015
Alabama and New York	$(H_0: \mu_A - \mu_{NY} = 0)$	32.6741	11.7420	0.000003

Assuming a significance level of $\alpha = 0.05$, answer the following questions.

a. Is there evidence to conclude that, on average, families in Alabama consume more candies per day than families in New York?

b. Which state appears to have the highest candy consumption per family per day according to the output.

7. Fisher's Least Significant Difference method examines the pairwise difference in the mean values of four treatment groups at a 0.05 level of significance. Determine the critical value if the total number of observations in all the samples is 30.

8. The number of paint defects found in a sample of 50 cars produced by three different car manufacturers (labeled A, B and C) are studied. The analysis of variance was significant at the 0.05 level indicating a difference in the average number of paint defects among the car manufacturers. Determine which car manufacturers are different using Fisher's Least Significant Difference method. Assume that the value calculated for Fisher's LSD is 4.4763, which is the same for each pair.

The following table shows the sample mean number of paint defects for each of the manufacturers.

Manufacturer	Mean Number of Paint Defects
A	7
B	12
C	9

9. The mean effect of three treatments on fasting blood sugar levels for three samples of 10 patients are shown below.

Treatments	Mean Fasting Blood Glucose Levels (mg/dL)
A	87.5
B	86.5
C	78.2

The ANOVA output for this experiment using R is as follows.

	df	Sum of Squares	Mean Square	*F*-value	*Pr(>F)*
Treatment	2	521.3	260.6	2.549	0.0968
Residuals	27	2760.6	102.2		

Assuming the level of significance is $\alpha = 0.10$, compare the pairwise differences in the mean blood glucose level for the three treatments using Fisher's Least Significant Difference method.

10. List one advantage of Tukey's HSD method over the two-sample *t*-test when the pairwise differences between the sample means are to be examined.

11. Compute the studentized range value for conducting Tukey's HSD test when the level of significance is equal to 0.05, the number of treatments is equal to 4, and the sample size of each of the four samples is equal to 16.

12. The cholesterol level of a total of 45 subjects is measured. The subjects were randomly divided into three groups and given different doses of medication (0 mg, 5 mg, 10 mg).

The one-way ANOVA table for testing if there is a significant difference in the mean cholesterol level for the different doses of medication is shown below.

	df	Sum of Squares	Mean Square	*F*-value	*Pr*(>*F*)
Dosage	2	53402	26701	3.57566	0.036813
Residuals	42	313632	7467.42857		

Is it wise to conduct a Tukey's HSD test to compare the difference in the mean cholesterol level at the following levels of significance?

a. 1%

b. 5%

13. Consider the test scores of a group of 15 students divided into three samples based on the type of curriculum studied. The following output is obtained after conducting a one-way ANOVA test.

	df	Sum of Squares	Mean Square	*F*-value	*Pr*(>*F*)
Curriculum	2	1301.7	650.9	28.18	2.93 E-05
Residuals	12	277.2	23.1		

The mean scores for the three samples are tabulated below.

Type of Curriculum	Mean Test Score
A	87.2
B	76.6
C	64.4

Determine if the mean tests scores are different for the following curriculum types using the confidence interval approach for Tukey's HSD with a 0.05 level of significance.

a. Sample A and Sample B

b. Sample B and Sample C

15.3 Two-Way ANOVA: The Randomized Block Design

In two-way ANOVA (analysis of variance), there are two independent variables that are being analyzed to determine their effects on a dependent variable. These two independent variables are often referred to as **factors**, which are called **treatments** and **blocks**. Treatments represent the main variable (factor) in the study. Researchers may be interested in the effects of both independent variables but are frequently interested in only one of them (treatments) and use the **randomized block design** to eliminate