Stats Review
Chapter 13

Revised 8/16

Note:

This review is meant to highlight basic concepts from the course.   It does not cover all concepts presented by your instructor. Refer back to your notes, unit objectives, handouts, etc. to further prepare for your exam.

The questions are displayed on one slide followed by the answers are displayed in red on the next.

This review is available in alternate formats upon request.

ANOVA Table

Fill in the ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-test Statistic
Treatment	837	4		
Error	9150	15		
Total				

ANOVA Table

Fill in the ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-test Statistic
Treatment	837	4		
Error	9150	15		
Total				

1. Total Sum of Squares: Add the sum of squares 837+9150=9987
2. Total Degrees of Freedom: Add the degrees of freedom 4+15=19
3. Find the Mean Squares of the treatment (MST)= $\frac{Sum\ of\ Squares\ Treatment}{Degrees\ of\ Freedom\ Treatment} = \frac{837}{4} = 209.25$
3. Find the Mean Squares of the Error (MSE) = $\frac{Sum\ of\ Squares\ Error}{Degrees\ of\ Freedom\ Error} = \frac{9150}{15} = 610$
4. F-Test Statistic = $\frac{MST}{MSE} = \frac{209.25}{610} = .343$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-Test Statistic
Treatment	837	4	**209.5**	**.343**
Error	9150	15	**610**	
Total	**9987**	**19**		

Computing F Statistic By Hand

Determine the F-Test Statistic given the summary statistics.

Population	Sample Size	Sample Mean	Sample Variance
1	10	38	27
2	10	41	30
3	10	56	23

Computing F Statistic By Hand

Determine the F-Test Statistic given the summary statistics.

Population	Sample Size	Sample Mean	Sample Variance
1	10	38	27
2	10	41	30
3	10	56	23

1. Find $\bar{x}$ using $\bar{x} = \frac{\sum n_i \bar{x}_i}{\sum n_i}$

$$\bar{x} = \frac{10 \cdot 38 + 10 \cdot 41 + 10 \cdot 56}{10 + 10 + 10} = 45$$

2. Find SST: $SST = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \cdots + n_k(\bar{x}_k - \bar{x})^2$

$$10(38-45)^2 + 10(41-45)^2 + 10(56-45)^2 = 1860$$

3. Find SSE: $SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2$

$$(10-1)27 + (10-1)30 + (10-1)23 = 720$$

4. Find MST: $\text{MST} = \frac{SST}{k-1}$

$$\frac{1860}{3-1} = 930$$

5. Find MSE: $\text{MSE} = \frac{SSE}{n-k}$

$$\text{MSE} = \frac{720}{30-3} = 26.66667$$

6. Calculate F: $F = \frac{MST}{MSE}$

$$F = \frac{930}{26.66667} = \mathbf{34.87}$$

Comparing Three or More Means

A company wants to make healthier donuts. The company tested 4 different fats to see which one was least absorbed by he donut during frying. Each fat was used for 5 batches of 2 dozen donuts each. The data is shown below. Is there a difference in means of fat absorbed at the $\alpha=.01$ significance level?

	Grams of fat absorbed				
	Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
Fat 1	64	72	68	77	56
Fat 2	78	91	97	82	85
Fat 3	75	93	78	71	63
Fat 4	55	66	49	64	70

Comparing Three or More Means

A company wants to make healthier donuts. The company tested 4 different fats to see which one was least absorbed by he donut during frying. Each fat was used for 5 batches of 2 dozen donuts each. The data is shown below. Is there a difference in means of fat absorbed at the α=.01 significance level?

	Grams of fat absorbed				
	Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
Fat 1	64	72	68	77	56
Fat 2	78	91	97	82	85
Fat 3	75	93	78	71	63
Fat 4	55	66	49	64	70

Conditions are met (see page 622)

1) State the null and alternative hypothesis

$H_o: \mu_1 = \mu_2 = \mu_3 = \mu_4$

H_1: At least 1 population mean is different.

2) Select the significance level. α=.01 was given

3) Compute the F-static/Make an ANOVA Table using technology

The following was given from an Anova Table using StatCrunch: F-Test Statistic; p-value =.0019

4) Reject H_o if p-value< α

.0019<.01 so reject H_o

5) State the Conclusion.

Reject the null hypothesis. There is sufficient evidence to conclude that at least one population mean of fat absorbed is different.

Tukey's Test

For the donut problem on the previous slides, we rejected the null hypothesis. Use Tukey's test to determine which pairwise mean differ using a familywise error rate of α=.01. The following table gives you the means and the mean square estimate (MSE).

$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	MSE (s^2)
67.4	86.6	76	60.8	78.95

This is from the ANOVA table

Tukey's Test

For the donut problem on the previous slides, we rejected the null hypothesis. Use Tukey's test to determine which pairwise mean differs using a familywise error rate of α=.01. The following table gives you the means and the mean square estimate (MSE).

$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	MSE (s^2)
67.4	86.6	76	60.8	78.95

This is from the ANOVA table

1) Compute the pairwise differences. Subtract the smaller mean from the larger mean.
2) Compute the test statistic

$q_0 = \frac{Difference\ in\ means}{\sqrt{\frac{s^2}{2}(\frac{1}{n_1}+\frac{1}{n_2})}}$ where s^2 =78.95, and the n's are the sample size of each fat

Comparison	Differences (step 1)	Test Statistic (step 2)
Fat 1 vs. Fat 2	86.6-67.4=19.2	$\frac{19.2}{\sqrt{\frac{78.95}{2}(\frac{1}{4}+\frac{1}{4})}} = 4.32$
Fat 1 vs. Fat 3	76-67.4=8.6	1.94
Fat 1 vs. Fat 4	67.4-60.8=6.6	1.49
Fat 2 vs. Fat 3	86.6-76=10.6	2.39
Fat 2 vs. Fat 4	86.6-60.8=25.8	5.81
Fat 3 vs. Fat 4	76-60.8=15.2	3.42

Continues

Tukey's Test

3) Use table IX to find the critical value $q_{\alpha,v,k}$.

α=.01 , v=n-k=20-4=16 and k=4. From the table $q_{.01,16,4}$ =5.192

4) Reject H_o if $q_0 > q_{\alpha,v,k}$

Comparison	Differences (step 1)	Test Statistic (step 32)	Critical Value (step 3)	Decision (step 4)
Fat 1 vs. Fat 2	19.2	4.32	5.192	4.32<5.192 Do not reject
Fat 1 vs. Fat 3	8.6	1.94	5.192	1.94<5.192 Do not reject
Fat 1 vs. Fat 4	6.6	1.49	5.192	1.49<5.192 Do not reject
Fat 2 vs. Fat 3	10.6	2.39	5.192	2.39<5.192 Do not reject
Fat 2 vs. Fat 4	25.8	5.81	5.192	5.81>5.192 Reject
Fat 3 vs. Fat 4	15.2	3.42	5.192	3.42<5.192 Do not reject

5) State the Conclusion:

There is sufficient evidence to conclude that fats 2 and 4 do not have the same mean amount of fat absorbed. This is denoted $\underline{\mu_2 \mu_4}$.

Completely Randomized Block Design

A psychologist wanted to test the effect of different types of music has on secretaries' typing efficiency. Each subject was given 3 types in random order: no music, hard rock, and classical. The subjects were given a typing test. Assume the requirement for normality is satisfied.

a) Verify that the requirement of equal population variances for each treatment is met.

b) Is there sufficient evidence that the mean typing scores are different among the 3 music types at the $\alpha=.05$ level of significance.

c) If the null hypothesis was rejected, use Tukey's test to determine which pairwise means differ using a familywise error rate of $\alpha=.05$.

		Subjects					
		1	2	3	4	5	6
Music Type	No Music	28.3	27.1	26.4	26.1	28.4	25.3
	Hard Rock	28.4	26.9	26.1	26.4	28.9	25.1
	Classical	28.7	27.2	26.8	27.3	29.1	25.8

Completely Randomized Block Design

a) Verify that the requirement of equal population variances for each treatment is met.
b) Is there sufficient evidence that the mean typing score are different among the 3 music types at the α=.05 level of significance.
c) If the null hypothesis was rejected, use Tukey's test to determine which pairwise means differ using a familywise error rate of α=.05 .

a) Calculate all the standard deviations

$$s_{no} = 1.24, s_{rock} = 1.44, s_{classic} = 1.23$$

Since none of the standard deviations are more than twice the size of any other standard deviation, the equal population standard deviation is met.

b) Using technology (see directions on page 655) we get the following ANOVA table.

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Subjects	5	25.016111	5.0032222	92.843299	<0.0001
Music Type	2	1.1411111	0.57055556	10.587629	0.0034
Error	10	0.53888889	0.053888889		
Total	17	26.696111			

The music type (the treatment) was p-value is .003 which is less than α. **We reject the null hypothesis. There is sufficient evidence to conclude that the mean test score is different among the three music types at the α=.05 level of significance.**

Continues

Completely Randomized Block Design

c) With the null hypothesis being rejected in part b, we can use Tukey's test to determine which pairwise means differ using a familywise error rate of α=.05.

Using technology (see directions on page 655) we get the following

Tukey HSD results (95% level) for Music Type:

Classical subtracted from

	Difference	Lower	Upper	P-value
Hard Rock	-0.51666667	-0.88407143	-0.1492619	0.0081
No Music	-0.55	-0.91740476	-0.18259524	0.0055

Hard Rock subtracted from

	Difference	Lower	Upper	P-value
No Music	-0.033333333	-0.4007381	0.33407143	0.9666

The confidence intervals for the difference between none and classical and the difference

Between rock and classical do not contain 0. **This means there is sufficient evidence to conclude that the means of the test scores with between rock music and classical are different. There is sufficient evidence to conclude that mean test scores between no music and classical is different.**

Two-Way ANOVA

A researcher wanted to determine if various types of music on the agitation of people with different stages of Alzheimer's disease. Three forms of music were chosen: soft rock, classical, and jazz. While listening to the music, a test was conducted to 3 patients within each stage of Alzheimer's. The results of the test are in the below table.

Stage	Music		
	Soft rock	Classical	Jazz
Early Stage Alzheimer's	104	111	95
	82	104	84
	81	74	104
Middle Stage Alzheimer's	111	109	106
	112	115	110
	82	97	99
Late Stage Alzheimer's	112	108	92
	108	99	129
	117	104	120

a) Determine if there is significant interaction between music and Alzheimer's stage.

b) If there is no significant interaction, determine whether there is significant difference in the means for the levels of music and Alzheimer's stage (separately)

c) If there was significant differences in the means for each level of each factor, use Tukey's test to determine which pairwise means differ using a family wise error rate of α=.05.

Two-Way ANOVA

a)Determine if there is significant interaction between music and Alzheimer's stage.

b)If there is no significant interaction, determine whether there is significant difference in the means for the levels of music and Alzheimer's stage (separately)

a) To determine if there is significant interaction between music and Alzheimer's stage, we use technology to determine the ANOVA table. See page 669 for instructions.

From the below ANOVA table, we can see by the interaction p-value of .812, **there is not a significant interaction.**

ANOVA table

Source	DF	SS	MS	F-Stat	P-value
Music	2	50.666667	25.333333	0.15468114	0.8578
Stage	2	1304	652	3.9810041	0.037
Interaction	4	256	64	0.39077341	0.8125
Error	18	2948	163.77778		
Total	26	4558.6667			

b) Since the p-value for the music factor is large (.8578), there is **no sufficient evidence of a difference in means from the music factor**. Since the p-value for the stage factor is small (.0237), there **is sufficient evidence of a difference in means from the stage factor**.

Continues

Two-Way ANOVA

c) If there was significant differences in the means for each level of each factor, use Tukey's test to determine which pairwise means differ using a family wise error rate of $\alpha=.05$.

The Tukey's test for the music factor is unnecessary because there was no evidence of a difference in means (see part b)).

Using technology, we get the following confidence intervals. (e=early, m=middle, l=late)

Tukey HSD results (95% level) for Stage:

e subtracted from

	Difference	Lower	Upper	P-value
l	16.666667	1.2698952	32.063438	0.0327
m	11.333333	-4.0634381	26.730105	0.1736

l subtracted from

	Difference	Lower	Upper	P-value
m	-5.3333333	-20.730105	10.063438	0.6569

We can see that the only confidence interval that does not contain zero is the difference between the late and early stages. **There is sufficient evidence to conclude that there is a difference in means of the heart rates between the late and early stages of Alzheimer's.**