

Ag 400 - Quiz 5
 Makeup - Fall 2008

- A. Suppose that a study was conducted to test the relative effectiveness of three different weight loss programs. Ten overweight subjects were randomly assigned to each program, and their weight loss measured at the end of the 12 week program. The following data were compiled.

Weight Loss Program	Number of Participants	Mean Number of Pounds Lost Per Participant
1 (Diet only)	10	15.0
2 (Exercise only)	10	10.0
3 (Diet and exercise combined)	10	20.0

- (24) 1. Test the statistical significance of the differences among the three weight loss programs. Use the .05 level to determine statistical significance. Specify the following items:

- a. State the Null Hypothesis and the Alternative Hypothesis.

H_0 : there are no differences among the 3 weight loss programs in regard to number of pounds lost.

H_a : there are differences . . .

- b. Complete the following table. Assume that the information given in the table is correct.

AOV Table

Sources of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Total	1850			
Among	500	2	250	5.00
Within	1350	27	50	

- c. Specify degrees of freedom for testing the significance of the calculated "F."

d. What is the critical value of the test statistic at the .05 level?

$$3.35$$

e. Reject or do not reject the Null Hypothesis.

Reject H_0

f. What is the probability that you have made a Type I error in (e) above?

$$.01 < \alpha < .05$$

g. What is the probability that you have made a Type II error in (e) above?

$$\beta = 0$$

h. Interpret the meaning of your conclusion in (e) above precisely in terms of the problem.

There are differences among the three weight loss programs in # pounds lost. Diet and exercise combined was associated with the greatest weight loss, followed by Diet alone. Exercise only was assoc. with the least weight loss

(8) 2. Use a Scheffe's Test to test the statistical significance of the difference between the mean weight loss of persons in the "Diet Only" weight loss program, and those in the "Exercise Only" weight loss program. Use the .05 level to determine significance.

a. Show your work.

$$\text{Sig diff} = \sqrt{50 \left(\frac{1}{10} + \frac{1}{10} \right)} \sqrt{3.35(2)} = 8.18$$

$$\text{Actual diff} = 5 \quad \therefore \text{n.s. } .05$$

b. Make your conclusions.

Cannot conclude that the amount of weight lost is different for Diet only or Exercise Only programs

(20) 3. Indicate whether each of the following statement is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F).

F a. The Null Hypothesis for the above analysis could have been stated as follows:

H_0 : In the population none of the three weight loss programs are effective in helping subjects to lose weight.

T b. The test in (A1b) above is a two-tailed test.

F c. In the above analysis, weight loss program is measured by three nominally scaled variables.

F d. Based on the analysis in (A1b) presented above, you should conclude that weight loss program 3 is significantly more effective than are programs 1 and 2 in helping people to lose weight.

F e. If you had access to the original data on which the above compilation was based, including the exact number of pounds lost by each subject in the study and information on the weight loss program in which each subject was enrolled, you could have tested the statistical significance of the relationship between weight loss program and pounds lost using a bivariate correlation-regression analysis.

F f. The analysis of variance carried out in A1b above tests the statistical significance of the differences among the variances in number of pounds lost for the three weight loss programs.

F g. The statistical unit in the above analysis is a pound of weight loss because that is what is being analyzed.

F h. Eta could be calculated here to measure the strength of the relationship between weight loss program and number of pounds lost. Here Eta would be positive. Therefore you can conclude that as weight loss program increases, number of pounds lost increases.

F i. The statistical interaction of weight loss program and number of pounds lost is statistically significant (.05 level) in these data. Therefore we can conclude that the effect of weight loss program on number of pounds lost differs for those in the "Diet only" category, those in the "Exercise only" group, and those who had both "Diet and Exercise."

F j. On the basis of the F-test in A1 above, you should conclude that the assumption of homogeneous variances is violated.

- B. Suppose that you are interested in the changes that occur in students' self esteem during their first semester in college. To this end, you assess the Self Esteem of six freshmen students using a standardized test designed to measure this concept, during the first week of classes, again at mid-semester, and again the end of their first semester. You compile the following data:

Student	Time Period			Totals
	1 st Week	Mid-Semester	End of Semester	
	-----Self Esteem Score-----			
1	2	3	4	9
2	5	4	3	12
3	3	5	7	15
4	5	6	7	18
5	3	5	7	15
6	<u>6</u>	<u>7</u>	<u>8</u>	<u>21</u>
Totals	24	30	36	90

- (10) 1. Test the statistical significance of the changes in Self Esteem scores that occurred across the semester.
- a. Show your work in the table below. Use the .05 level to determine significance. You may assume that the figures given in the table are correct.

AOV Table

Sources of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Total	54	17		
Among Time Periods	12	2	6	5.00 sig .05
Among students	30	5	6	
Error	12	10	1.2	

- b. Make your conclusions. Is there evidence that Self Esteem scores change during the students' freshmen year? If yes, how do they change? If no, what can you conclude?

Self esteem score increase during the year

- (8) 2. Indicate whether each of the following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F).

F a. The above analysis deals with matched cases and therefore could have been carried out using a paired t-test.

F b. The interactive effect of time period and self esteem is statistically significant at the .05 level in these data.

T c. The Null Hypothesis for the test carried out in B1 above can be stated as follows:

H_0 : there are no differences among the three time periods in regard to mean Self Esteem Scores.

T d. If you reject the Null Hypothesis in B1 above, you cannot make a Type II error (i.e. $\beta = 0.00$).

- (30) C. A random sample of Mudville residents was surveyed to obtain information on the relationship between AGE and subjects' feelings of Happiness or Satisfaction with his/her current life situation. "Happiness" was measured by an attitude scale such that the higher the score the "Happier" was the individual.

The attached output uses the following variables.

HAPPY = Happiness Score
(an interval scale)

AGE

- 1 < 30
- 2 30 - 59
- 3 60 and over

Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F).

- F 1. The relationship of AGE to HAPPY in this sample can be described as follows.

As age increases, Happiness Scores decrease, then increase, then decrease again.

- T 2. The sum of the squares of the deviations of the 436 happiness scores about 7.6498 equals 496.510.

- T 3. The variance in happiness scores for the entire sample is about 1.14.

- F 4. The F-value on Page 1 tests the following Null Hypothesis:

H_0 : there is no relationship between the people less than 30 years of age, those between 30 and 59 years, and those 60 years and over in regard to Happiness Score.

- F 5. Based on Levene's Statistic you should conclude that the variances in Happiness Scores for each of the age categories are the same.

- T 6. The figures reported on the output are statistics; "F" is an inferential statistic, while the Means and Standard Deviations are descriptive statistics.

- T 7. Both the Levene's Test and the "F" statistic use two-tailed tests on this output.

- T 8. Eta calculated for this analysis equals about (.29). This is statistically significant at the .001 level.

- F 9. In the sample, persons over 60 years of age have the highest Happiness Scores; those between 30 and 59 years have the lowest Happiness Scores. Those who are less than 30 years of age have an intermediate mean score. However, none of these differences are statistically significant.
- F 10. This output leads to the conclusion that the Homogeneity of Variance assumption for the F-test reported on Page 1 is violated. Therefore, we should be cautious in interpreting the results of the F statistic.
- T 11. Approximately 92% of the variation in Happiness Scores in the sample is not associated with differences among the three age categories used as the independent variable in this analysis.
- F 12. Persons who are less than 30 years of age are significantly (.05 level) happier than 30-59 year olds and significantly less happy than are those who are 60 years old or older.
- T 13. Based on the analysis on Page 2, you can conclude that persons <30 years of age do not differ significantly from those who are 60 years of age or older in regard to Happiness Scores.
- F 14. The statistical unit on Page 1 of the output is a person, but on Page 2, the statistical unit is an age category.
- I 15. Based on this analysis, you can conclude that age is related to happiness scores with people in the 30-59 years of age category expressing the least happiness and both younger and older persons reporting significantly greater happiness than those in the "middle years."

Oneway

Descriptives

HAPPY

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
< 30	119	7.6731	1.08000	.09900	7.4771	7.8692	4.00	9.60
30-59	142	7.2366	1.06893	.08970	7.0593	7.4140	3.60	9.30
60 & oner	175	7.9691	.94569	.07149	7.8280	8.1102	5.60	9.40
Total	436	7.6498	1.06836	.05117	7.5492	7.7503	3.60	9.60

Test of Homogeneity of Variances

HAPPY

Levene Statistic	df1	df2	Sig.
.754	2	433	.471

ANOVA

HAPPY

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	42.153	2	21.077	20.086	.000
Within Groups	454.357	433	1.049		
Total	496.510	435			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: HAPPY
Scheffe

(I) Age	(J) Age	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
< 30	30-59	.4365*	.12731	.003	.1238	.7492
	60 & oner	-.2960	.12171	.053	-.5950	.0029
30-59	< 30	-.4365*	.12731	.003	-.7492	-.1238
	60 & oner	-.7325*	.11570	.000	-1.0167	-.4483
60 & oner	< 30	.2960	.12171	.053	-.0029	.5950
	30-59	.7325*	.11570	.000	.4483	1.0167

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

HAPPY

Scheffe^{a,b}

Age	N	Subset for alpha = .05	
		1	2
30-59	142	7.2366	
< 30	119		7.6731
60 & oner	175		7.9691
Sig.		1.000	.053

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 141.778.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.