

Statistics 252 – Midterm Exam A

Instructor: Paul Cartledge

Instructions:

1. Read all the instructions carefully.
2. This is a closed book exam.
3. You may use the formula sheets and the tables provided and a calculator only.
4. You have 50 minutes to complete the exam.
5. The exam is out of a total of 42 marks.
6. Show your work in all sections to receive full credit. Final numerical answers should have THREE significant decimal places.
7. Use the backs of the pages for scrap work.
8. Make sure your name and signature are on the front and that your ID number is on the top of page two.
9. When referring to “log”, I am always referring to the natural log.
10. If no significance level is given, use the “judgment approach”.
11. When asked for a “confidence interval”, state the estimate, the standard error, and the critical value. Then, calculate and interpret the interval.
12. When asked to “carry out a full analysis in detail”, set up the hypotheses, calculate the test statistic, state the distribution of the test statistic (such as t_9 or $F_{3,10}$), approximate the p-value, and state your conclusion in plain English.

Name (last name, first name): _____

Signature: _____

Component	Notes	Worth	Mark
Short Answer	6 questions	11	
Long Answer			
Question 7	3 parts	21	
Question 8	4 parts	10	
Total		42	

ID: _____

Question 1 (2 marks) Starting their midterm, a student sees a test statistic of 2.21 in a test for any differences between several means. From a data structure where the groups have respective sample sizes of 5, 4, 5, 6, and 4, what is the distribution of the test statistic?

Question 2 (2 marks) Using information from Question 1 and a p -value of 0.107, apply the “judgment approach” and make a decision *and* conclusion about the rejection (OR non-rejection) of the null hypothesis that claims equality among the several means.

Question 3 (2 marks) Is the following statement true or false? Defend your answer either way in one or two sentences. Simply an answer of “true” or “false” will not receive any credit. “If random sampling is present, it is possible to make causal inferences.”

Question 4 (1 mark) What is the advantage of Scheffe multiple comparisons?

Question 5 (2 marks) Bono attempts to analyse two independent random samples. Putting the first sample first in the order, some skewness in the sample distributions requires a log transformation of the data. Thus, he reports a 99% confidence interval of the ratio of the medians on the original scale to be (0.522, 0.864). What can you say about the possible rejection of the null hypothesis testing to see if the medians on the original scale are the same? Explain your answer in one or two sentences. No calculation required.

Question 6 (2 marks) Using the confidence interval from Question 5, find the corresponding significance level to test if the second median is larger than the first median. Also, how would the p -value compare to this significance level?

Question 7 (21 marks total) At Greendale Community College, an instructor is getting concerned about student performance in the night classes. Professor P. Professorson decided to accumulate information on the following classes. The tables below summarize: summary statistics of final exam percentage (measured in decimals); the ANOVA output; selected linear combinations. Assume all assumptions hold.

Group	Class	n_i	Sample Mean	Sample S.D.
1	History of Something	10	0.847	0.0224
2	Introduction to Basics	20	0.851	0.0265
3	Studyology	10	0.905	0.0310
4	Theoretical Physical Education	10	0.792	0.0309
5	Math 1-2-3	20	0.742	0.0542
6	Class 101	20	0.881	0.0515

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p -value
Between (Extra)	0.289			34.372	0.000
Within (Full)					
Total (Reduced)	0.430	89			

a) (6 marks)

i) (1 mark) If you had to fully analyse “unplanned comparisons” upon the groups, how many unique pairings of the groups will you need?

ii) (1 mark) Using the Bonferroni method, if the experiment-wise confidence level is 94%, what are the corresponding individual confidence levels?

iii) (4 marks) Using the Bonferroni method and the table below that summarizes the *margin of error* between each pair (for example, the margin of error for group 1 vs. group 2 is 0.0505), construct a visual diagram below the table that joins groups together that are **not** different.

Row vs. Column	1	2	3	4	5
2	0.0505				
3	0.0583	0.0505			
4	0.0583	0.0505	0.0583		
5	0.0505	0.0413	0.0505	0.0505	
6	0.0505	0.0413	0.0505	0.0505	0.0413

The ANOVA output below compares average final exam percentage by grouping the above groups into two categories: classes that average over 80% and those that average less.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups			.245	116.628	.000
Within Groups					
Total	.430	89			

The ANOVA output below compares average final exam percentage by grouping the above groups into three categories: classes that average < 80%, between 80 – 90%, and > 90%.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups			.130	66.837	.000
Within Groups					
Total	.430	89			

b) (8 marks) Test if the model where all classes have potentially different mean percentages is better than the model that groups classes by comparing to an 80% boundary. Carry out a full analysis in detail, identifying the SSR and df for the respective models. (Hint: You can assume the p -value will be less than 0.01.)

Contrast Coefficients

Contrast	Type					
	1	2	3	4	5	6
1	1	1	1	-1	-1	-1
2	1	-1	1	1	-1	-1
3	1	0	-.5	-.5	0	0
4	1	-.5	1	1	-.5	-.5
5	0	1	0	0	-.5	-.5

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Gross	Assume equal variances	1	.188	.028	6.837	84	.000
		2	.069	.028	2.509	84	.014
		3	-.002	.016	-.126	84	.900
		4	1.306	.024	54.842	84	.000
		5	.040	.011	3.536	84	.001

c) (7 marks) Do the smaller classes have different final exam percentages than the larger classes? Construct a confidence interval to answer this question.

- i) (2 marks) First, define a linear combination of means contrasting the average percentage for the smaller classes vs. the larger classes. Fill in the blanks below for the coefficients in your contrast.

$$\gamma = \text{---} \mu_1 + \text{---} \mu_2 + \text{---} \mu_3 + \text{---} \mu_4 + \text{---} \mu_5 + \text{---} \mu_6$$

- ii) (2 marks) Give the estimate and standard error for this contrast.

$$g = \text{---} \quad S.E.(g) = \text{---}$$

- iii) (3 marks) Calculate a 90% confidence interval for the contrast.

Question 9 (10 marks) Recent events in the past year have got people wondering about the marketability of double-decker couches in certain locations. Investigating the business, a marketing expert takes a random sample of 45 department stores in Bricksburgh and a sample of the same size in Clown Town. He measures the monthly income of each store's furniture department to obtain the following summary statistics (units are in thousands of \$US).

Summary statistic	Bricksburgh	Clown Town	Difference
Average	193.40	172.51	20.89
Standard Deviation	39.80	20.14	44.60

NOTE: Please note that the third column summarize the differences from the original observations. By choosing a test, you will be using certain columns of the above table, but not all of them.

Based on statistical evidence, is Bricksburgh selling more double-decker couches?

a) (3 marks) Is the above situation two independent samples or a paired sample?

b) (2 marks) Write the appropriate null and alternative hypotheses.

c) (3 marks) Suppose the test statistic is $t_0 = 3.142$. State the distribution of the test statistic and determine the range of the p -value.

d) (2 marks) Make a decision and state a conclusion in plain English.