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): SOLUTION

Signature: ____

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

		1	2	3	4	5	6	7	8	9	10
200	0.25	1.33	1.40	1.38	1.36	1.34	1.32	1.30	1.29	1.28	1.27
200	0.1	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63
200	0.05	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88
200	0.025	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18	2.11
200	0.01	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41
200	0.005	8.06	5.44	4.41	3.84	3.47	3.21	3.01	2.86	2.73	2.63
200	0.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26	3.12

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?

With more than 2 groups, this should be an *F*-distribution.
$I = 5$, $N = 5 + 4 + 5 + 6 + 4 = 24$; thus the *F*-distribution is $F_{4,\,19}$.

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

The given *p*-value provides weak evidence against H_0. Thus, it is best to not reject the null hypothesis, so there is insufficient evidence of a difference between the 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."

The statement is "false" because random assignment needs to be present to make causal inferences possible.

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

Most conservative.

Question 5 (2 marks) Evan Rachel Wood 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, she computes a 95% 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.

Since the value of 1 is not within the 95% confidence interval for the ratio of medians, the corresponding null hypothesis will be rejected. Thus, the medians are different at $\alpha = 0.05$.

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?

Since $1 - \alpha = 0.95$, then $\alpha = 0.05$ and for a one-tailed test, the appropriate α would be 0.025.

Since the interval above does not contain 1, then $p\text{-value} < \alpha$.

Question 7 (21 marks total) At the Ministry of Official Languages, a prominent Danish linguist (L. Egolandsen) investigates the number of languages known by citizens of the world. He decides to randomly sample 40 citizens from seven random international cities. The tables below summarize: summary statistics of number of languages by city, also listing the respective nation and continent; the ANOVA output. Assume all assumptions hold.

Group	City	Nation	Continent	Sample Mean	Sample S.D.
1	Istanbul	Turkey	Asia	2.596	0.802
2	Lagos	Nigeria	Africa	4.259	1.005
3	Osaka	Japan	Asia	2.116	0.952
4	Vancouver	Canada	N. America	1.945	0.946
5	Edinburgh	U.K.	Europe	1.529	0.486
6	Karachi	Pakistan	Asia	4.266	1.666
7	London	U.K.	Europe	1.549	0.481

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	*p*-value
Between (Extra)	337.683	6	56.280	58.826	0.000
Within (Full)	261.185	273	0.957		
Total (Reduced)	598.868	279			

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 95.8%, what are the corresponding individual confidence levels?

iii) (4 marks) Using the Bonferroni method and the table below that summarizes the *mean differences* between each pair (for example, the mean difference between group 1 and group 2 is -1.663), construct a *means comparison diagram* below the table that joins groups together that are **not** different. Note that the margin of error for all pairs is 0.682.

$$\boxed{m = \frac{I(I-1)}{2} = \frac{7(6)}{2} = 21}$$

$$\boxed{1 - \frac{\alpha_E}{m} = 1 - \frac{0.042}{21} = 0.998 \rightarrow 99.8\%}$$

Row – Column	1	2	3	4	5	6
2	**-1.663**					
3	0.479	**2.142**				
4	0.651	**2.314**	0.172			
5	**1.066**	**2.729**	0.587	0.415		
6	**-1.671**	-0.008	**-2.150**	**-2.322**	**-2.737**	
7	**1.047**	**2.710**	0.568	0.396	-0.020	**2.717**

(Note: The table above identifies differences larger than the margin of error **in this form**.)

1.529	1.549	1.945	2.116	2.596	4.259	4.266
5	7	4	3	1	2	6

The ANOVA output below compares average number of languages by grouping the above cities by continent.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	235.766	3	78.589	59.736	.000
Within Groups	363.102	276	1.316		
Total	598.868	279			

The ANOVA output below compares average number of languages by grouping the above cities by nation.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	337.675	5	67.535	70.846	.000
Within Groups	261.193	274	.953		
Total	598.868	279			

b) (8 marks) Test if the model where all cities have potentially different mean numbers of languages is better than the model that groups cities by continent. Carry out a full analysis in detail, identifying the SSR and df for the respective models. (Hint: Use the table on the first page of the exam to help find the p-value.)

H_0: (4-mean model) $\mu_1 = \mu_3 = \mu_6$, μ_2, μ_4, $\mu_5 = \mu_7$
H_A: (7-mean model) μ_1, μ_2, μ_3, μ_4, μ_5, μ_6, μ_7

SSR(r) = 363.102
df(r) = 276

SSR(f) = 261.185
df(f) = 273

$$F_0 = \frac{(SSR(r) - SSR(f))/(df(r) - df(f))}{SSR(f)/df(f)} = \frac{(363.102 - 261.185)/(276 - 273)}{261.185/273}$$

$$F_0 = \frac{(101.917)/(3)}{261.185/273} = 35.509 \sim F_{df(f)}^{df(r)-df(f)} = F_{273}^{3} \approx F_{200}^{3}$$

$$p\text{-value} = P\left(F_{df(f)}^{df(r)-df(f)} > F_0\right) = P\left(F_{273}^{3} > 35.509\right) \approx P\left(F_{200}^{3} > 35.509\right) \in (0, 0.001)$$

JA: Strong to convincing evidence against H_0. Reject H_0. Thus, the seven-mean model is better than the four-mean model.

Contrast Coefficients

Contrast	Type 1	2	3	4	5	6	7
1	1	-3	1	0	0	1	0
2	1	0	1	-3	0	1	0
3	2	0	2	0	-3	2	-3
4	0	0	0	2	-1	0	-1
5	1	1	1	-6	1	1	1
6	1	-6	1	1	1	1	1

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Gross	Assume equal variances	1	-3.798	.536	-7.089	273	.000
		2	3.144	.536	5.869	273	.000
		3	13.369	.847	15.782	273	.000
		4	2.360	.379	6.230	273	.000
		5	3.099	1.002	3.092	273	.002
		6	-13.100	1.002	-13.070	273	.000

c) (7 marks) Do Asian nations have a different average number of languages than North American nations? Construct a confidence interval to answer this question.

i) (2 marks) First, define a linear combination of means contrasting the average number of languages for the Asian nations vs. the North American nations. Fill in the blanks below for the coefficients in your contrast.

$$\gamma = \boxed{\left(\frac{1}{3}\right)}\mu_1 + \boxed{(0)}\mu_2 + \boxed{\left(\frac{1}{3}\right)}\mu_3 + \boxed{(-1)}\mu_4 + \boxed{(0)}\mu_5 + \boxed{\left(\frac{1}{3}\right)}\mu_6 + \boxed{(0)}\mu_7$$

ii) (2 marks) Give the estimate and standard error for this contrast.
See contrast 2 for 3*g*.

g = 3.144/3 = 1.048 $S.E.(g)$ = 0.536/3 = 0.179

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

$g \pm t_{N-I,\,\alpha/2} \times S.E.(g)$ → 1.048 ± (1.660)(0.179) → 1.048 ± 0.296 → (0.752, 1.345)

With 90% confidence, the contrast is between 0.752 and 1.345.

Question 9 (10 marks) While *The Simpsons* is still a somewhat popular TV show, a television analyst (C.B. Guy) noticed that the quality of the Halloween episodes tended to be greater, so they wondered if internet users agreed. Using the film's rating at IMDB.com (measured on a scale of 0 to 10, with 10 being the highest), the analyst compared the sample of Halloween episodes from each of 26 seasons to the average of the remaining episodes within each season, finding the following summary statistics. Assume rating follows a normal distribution.

Summary statistic	Halloween	Rest of season	Difference
Average	7.646	7.335	0.311
Standard Deviation	0.631	0.612	0.225

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, do Halloween episodes have greater quality?

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

Paired samples

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

Let $d = y_{\text{H}} - y_{\text{Rest}}$

H_0: $\mu_d \leq 0$ H_A: $\mu_d > 0$

c) (3 marks) Suppose the standard error of the appropriate estimate is 0.0440. Calculate the test statistic, state the distribution of the test statistic, and determine the range of the *p*-value.

$$t_0 = \frac{\bar{d} - D_0}{s_d / \sqrt{n}} = \frac{0.311 - 0}{0.0440} = 7.053$$

$$t_0 = t_{n-1} = t_{26-1} = t_{25}$$

$$3.725 < t_0 = 7.053 < \infty$$
$$0.0005 > p\text{-value} > 0$$

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

Strong to convincing evidence against H_0. → Reject H_0.
→ Halloween episodes have greater quality, obviously because of James Earl Jones.
(Quoth the raven: "Nevermore.")