Name: _____
Student No.: _____

University of Alberta

Department of Mathematical and Statistical Sciences

Statistics 252 - Midterm Examination Version A (Solution)

Date: March 5, 2009

Instructor: Alireza Simchi

Time: 11:00-12:20

Instructions: (READ ALL INSTRUCTIONS CAREFULLY.)

1. This is a closed book exam. You are permitted to use a non-programmable calculator. Please turn off your cellular phones or pagers.

2. The exam consists of **three** parts. In the first two parts there are 15 multiple-choice questions. For each multiple-choice question choose the answer that is closest to being correct. Circle one of the letters (a)-(e) corresponding to your chosen answer for each question. All answers will be graded right or wrong (no partial credit) in this part. Each single question is worth 1 point. All numerical answers are rounded. In the third part there is one long-answer problems. **Show all your work to get full credit**. In fact, answers must be accomplished by adequate justification. If you run out of space, use the back of any page for answers as needed. Clearly direct the marker to answers that you provide on the back of a page.

3. This exam has **8** pages including this cover. Please ensure that you have all pages and write your name and your student ID at the top of each page.

4. The statistical tables and formula sheet are provided in a separate booklet.

5. The exam is graded out of a total of **25** points.

Name: _____
Student No.:_____

Circle one answer for each question on the following table. Each question is worth 1 mark.

Question	Answer
1	a b **c** d e
2	**a b c d e**
3	**a** b c d e
4	**a b c d e**
5	a b c **d** e
6	**a b c d e**
7	a b c **d** e
8	**a b c d e**
9	**a** b c d e
10	**a b c d e**
11	a b **c** d e
12	**a b c d e**
13	**a** b c d e
14	**a b c d e**
15	a b c d **e**

Name: _____
Student No.:_____

PART 1

Weight, in grams, of the 15 male house sparrows that survived and the 14 that perished from the sever winter storm were recorded. Some summary statistics were given in the following. Assume all the required assumptions are satisfied. Is there enough evidence to conclude that the average weight of sparrows who survived in the sever winter storm is less than the average weight of sparrows who perished? Questions **1** to **3** are related to this question.

Group Statistics

	Status	N	Mean	Std. Deviation	Std. Error Mean
Weight	Survived	15	25.1400	1.12745	.29111
	Perished	14	26.1571	1.03828	.27749

1. What is the distribution of the test statistic under the null hypothesis?

a) t(13) **b)** t(14) **c)** **t(27)** **d)** t(28) **e)** t(29)

2. What is the absolute value of the test statistic?

a) 0.52 **b)** 1.52 **c)** **2.52** **d)** 3.52 **e)** 4.52

3. The range of p-value can be describe as:

a) **less than 0.01**.
b) between 0.01 and 0.025.
c) between 0.025 and 0.05.
d) between 0.05 and 0.10.
e) greater than 0.10

4. In determining a 95% confidence interval for the mean difference for two groups, what is the upper limit of the confidence interval $\mu_2 - \mu_1$, where μ_1 is the mean of weight for survived sparrows and μ_2 is the mean of weight for perished sparrows?

a) 0.189 **b)** 0.845 **c)** 1.189 **d)** **1.845** **e)** 2.189

PART 2

A researcher conducted an experiment to compare the effectiveness of four new weight-reducing agents to that of an existing agent. The standard agent is labeled agent S, and the four new agents were labeled A1, A2, A3, and A4. The researcher randomly divided a random sample of 50 males into five equal groups, with preparation A1 assigned to the first group, A2 to the second group, and so on. Then, they gave a pre-study physical to each person in the experiment and told him how many pounds overweight he was. A comparison of the mean number of pounds overweight for the groups showed no significant differences. The researchers then began he study program, and each group took the prescribed preparation for a fixed period of time. Then the weight losses (WL) recorded at the end of the study period. Assume all required assumptions are satisfied. The summary statistics and ANOVA table for these data are given in the following:

Name: _____
Student No.:_____

WL	N	Mean	Std. Deviation
S or 0	10	9.27	1.15859
A1 or 1	10	10.27	1.02637
A2 or 2	10	11.02	1.12131
A3 or 3	10	12.05	.82899
A4 or 4	10	12.24	.75601
Total	50	10.97	1.46959

ANOVA

wl

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	61.618	?	?	?	?
Within Groups	?	?	?		
Total	105.825				

Suppose we would like to carry out a test to determine if there are any significant differences among the five weight-reducing agents. Questions **5** to **8** are related to this test.

5. In the test for any mean differences, what is the distribution of the test statistic under the null hypothesis?

a) F(3,46) **b)** F(4, 46) **c)** F(3, 45) **d)** **F(4,45)** **e)** F(5, 45)

6. In the test for any mean differences, what is the best estimate for the common standard deviation of the five weight-reducing agents?

a) 0.98 **b)** **0.99** **c)** 1.00 **d)** 1.1 **e)** 1.2

7. In the test for any mean differences, what is the value of the test statistic (approximately)?

a) 12.5 **b)** 13.1 **c)** 14.3 **d)** **15.7** **e)** 16.9

8. In the test of any mean differences, the range of p-value can be describe as:

(a) **less than 0.001**.
(b) between 0.001 and 0.01.
(c) between 0.01 and 0.025.
(d) between 0.025 and 0.05.
(e) greater than 0.05

Consider a contrast to compare the mean of standard agent to the average of the four agent means. Questions **9** and **10** are related to this contrast.

9. What is the estimate of the contrast?

a) **2.125** **b)** 3.125 **c)** 4.52 **d)** 8.50 **e)** 12.5

10. What is the standard error for the contrast?

a) **0.35** **b)** 1.40 **c)** 2.35 **d)** 3.40 **e)** 5.6

Name: _____
Student No.:_____

Suppose we would like to determine if there are any differences among the four new weight-reducing agents A1, A2, A3, and A4. Use the following ANOVA tables to answer questions **11** to **13**.

ANOVA Table (Comparing all Five Groups):

ANOVA

wl

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	61.618	?	?	?	?
Within Groups	?	?	?		
Total	105.825				

ANOVA Table (Ignoring New Agents or New Agents Versus Control):

ANOVA

wl

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	36.125	1	?	?	?
Within Groups			?		
Total	105.825				

11. What is the distribution of the test statistic under the null hypothesis?

a) F(1,48) **b)** F(2, 46) **c)** **F(3, 45)** **d)** F(4,45) **e)** F(5, 45)

12. What is the value of the test statistic (approximately)?

a) 2.81 **b)** **8.65** **c)** 12.53 **d)** 15.68 **e)** 24.85

13. In the test of any mean differences, the range of p-value can be describe as:

a) **less than 0.001**.
b) between 0.001 and 0.01.
c) between 0.01 and 0.025.
d) between 0.025 and 0.05.
e) greater than 0.05

Suppose we would like to use the Bonferroni method for all simultaneous inference with family-wise confidence of 95%. Questions **14** and **15** are related to this method.

14. What is (approximately) the critical value for 95% family-wise confidence intervals?

a) 1.3 **b)** 1.7 **c)** 2.1 **d)** 2.4 **e)** **2.9**

15. What is (approximately) the margin of error for all 95% family-wise confidence intervals?

a) 0.57 **b)** 0.75 **c)** 0.93 **d)** 1.06 **e)** **1.29**

PART 3

16. (10 Marks in Total) Previous studies suggest that vegetarians may not receive enough zinc in their diets. As the zinc requirement is particularly important during pregnancy, researchers conducted a study to determine whether vegetarian pregnant women are at greater risk from low zinc levels than are non-vegetarian pregnant women. Twenty-nine women were monitored: twelve vegetarians who were pregnant, six non-vegetarians who were pregnant, five vegetarians who were not pregnant, and six non-vegetarians who were not pregnant. None of these women were smokers, and none of the non-pregnant women were taking oral contraceptives. The zinc content in hair was measured for each woman.

Define:

μ_1: Average zinc content for non-vegetarians pregnant women (NV-P)

μ_2: Average zinc content for vegetarians and pregnant women (VP)

μ_3: Average zinc content for vegetarians and non-pregnant women (V-NP)

μ_4: Average zinc content for non-vegetarians and non-pregnant women (NV-NP)

When needed use Tables 1 to 4 below to answer questions **(a)** and **(b)** (you won't necessarily need all of the tables).

Table 1: The summary statistics of zinc content for the 4 groups are:

Descriptive

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
1) NV-P	6	178.0000	14.47757	5.91044	150.00	189.00
2) VP	12	177.0833	20.87390	6.02578	125.00	207.00
3) V-NP	5	192.8000	13.29286	5.94475	177.00	212.00
4) NV-NP	6	201.1667	16.61826	6.78438	180.00	220.00
Total	29	184.9655	19.68227	3.65491	125.00	220.00

Table 2: The ANOVA table for the comparison of average zinc content for the 4 groups:

ANOVA

Zinc

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2918.416	3	972.805	3.067	.046
Within Groups	7928.550	25	317.142		
Total	10846.966	28			

Name: _____
Student No.:_____

Table 3: The ANOVA table for the comparison of average zinc content between pregnant and non-pregnant women (ignoring vegetarian status):

ANOVA

Zinc

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2724.142	1	2724.142	9.055	.006
Within Groups	8122.823	27	300.845		
Total	10846.966	28			

Table 4: The ANOVA table for the comparison of average zinc content between vegetarians and non-vegetarians women (ignoring pregnancy status):

ANOVA

Zinc

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	436.519	1	436.519	1.132	.297
Within Groups	10410.446	27	385.572		
Total	10846.966	28			

a) **(5 marks)** Consider the vegetarian effect after accounting for pregnancy status. Does there appear to be a significant vegetarian affect within either of the pregnancy status? In other words, is there a significant difference in zinc content between vegetarians and non-vegetarians women with the same type of pregnancy status? Carry out **a single overall test** to determine if there is a difference in average zinc content between vegetarians and non-vegetarians women who have the same pregnancy status (either pregnant or non-pregnant). In terms of the parameters defined earlier, state clearly the models in the null and alternative hypothesis. Also, identify the residuals sum of squares and degrees of freedom for the models in the null and alternative hypothesis. Calculate the test-statistic, the p-value and identify the distribution of the test-statistic under the null hypothesis. What do you conclude?

If there is no vegetarian effect for pregnant women, then we have $\mu_1 = \mu_2$.
If there is no vegetarian effect for non-pregnant women, then we have $\mu_3 = \mu_4$.

$$\begin{cases} H_0 : \mu_1 = \mu_2 \quad and \quad \mu_3 = \mu_4 \quad (\text{Reduced model :} \\ \quad \text{Two means model only for pregnant and/or nonpregnant women)} \ \textbf{(0.5 marks)} \\ H_1 : \mu_1, \mu_2, \mu_3, \mu_4 (\text{Full model : Four means model}) \end{cases}$$

From table 2:
$SS_{\text{Res}}(\text{Four means model}) = 7928.550$ and $df_{\text{Res}}(\text{Four means model}) = 25$ **(1 mark)**

From table 3:
$SS_{\text{Res}}(\text{Two means model}) = 8122.823$ and $df_{\text{Res}}(\text{Two means model}) = 27$ **(1 mark)**

Hence, Extra SS = 8122.823 – 7928.550 = 194.273 and Extra df = 27 -25 = 2.

Hence, the value of the test statistic is:

$$TS = \frac{Extra \quad SS / Extra \quad df}{MS_{\text{Res}}(\text{Full} \quad \text{model})} = \frac{194.273/2}{7928.550/25} = 0.306 \text{ . } \textbf{(1 mark)}$$

If the null hypothesis is true, then the test statistic has an F-distribution with degrees of freedom $df_1 = 2, df_2 = 25$ **. (0.5 marks)**

So p-value is:

$$p - value = P(TS > 0.306) \Rightarrow p - value > 0.10 \quad \textbf{(0.5 marks)}$$

Conclusion: The p-value greater than 0.10 indicates weak evidence against null hypothesis. Therefore, there is not enough evidence to conclude that there is a difference in average zinc content between vegetarians and non-vegetarians women who have the same pregnancy status (either pregnant or non-pregnant). **(0.5 marks)**

b) (5 marks) Use the Bonferroni method to calculate two simultaneous 96% confidence intervals for the difference in average zinc content for vegetarians and non-vegetarians at each of the pregnancy status (one for vegetarian pregnant women vs. non-vegetarian pregnant women, and one for vegetarian non-pregnant women vs. non-vegetarians women who are not pregnant). What is your final conclusion, based on your results from parts **(a)** and **(b)**?

We want 95% family-wise confidence interval for:

a. vegetarian effect for pregnant women: $\gamma_P = \mu_2 - \mu_1$
b. vegetarian effect for non-pregnant women: $\gamma_{NP} = \mu_3 - \mu_4$

$\alpha_E = 0.04, m = 2 \Rightarrow \alpha_I = \frac{\alpha_E}{m} = \frac{0.04}{2} = 0.02$. We should find two individual 98% confidence intervals to obtain a 96% family-wise confidence interval. Hence:

$\frac{\alpha_I}{2} = \frac{\alpha_E}{2m} = \frac{0.04}{2(2)} = 0.01, df = N - I = 29 - 4 = 25 \Rightarrow t^*_{25,0.01} = 2.485$ **(1 mark)**

$$\hat{\gamma}_P = \bar{y}_2 - \bar{y}_1 = 177.0833 - 178.000 = -0.9167,$$
$$S.E.(\hat{\gamma}_P) = \sqrt{MS(within)}\sqrt{\frac{(1)^2}{n_2} + \frac{(-1)^2}{n_1}} = \sqrt{317.142}\sqrt{\frac{1}{12} + \frac{1}{6}} = 8.9042$$

(1 mark)

$$\hat{\gamma}_{NP} = \bar{y}_3 - \bar{y}_4 = 192.8000 - 201.1667 = -8.3667,$$
$$S.E.(\hat{\gamma}_{NP}) = \sqrt{MS(within)}\sqrt{\frac{(1)^2}{n_2} + \frac{(-1)^2}{n_1}} = \sqrt{317.142}\sqrt{\frac{1}{5} + \frac{1}{6}} = 10.7836$$

(1 mark)

An individual 98% confidence interval for $\gamma_i, i = P, NP$, is $\hat{\gamma}_i \pm 2.485 S.E.(\hat{\gamma}_i)$. The results are summarized in the following table. If the confidence interval does not include zero, then we reject the null hypothesis $H_0 : \gamma_i = 0$ versus $H_1 : \gamma_i \neq 0$. Therefore, γ_i is significant.

Name: _____
Student No.:_____

Parameter	Estimate	Standard Error of Estimate	**Lower Bound**	**Upper Bound**	**Significant**
$\gamma_P = \mu_2 - \mu_1$	-0.9167	8.9042	**-23.0436**	**21.2102**	**No**
$\gamma_{NP} = \mu_3 - \mu_4$	-8.3667	10.7836	**-35.1639**	**18.4305**	**No**

Conclusion: (0.5 marks for each confidence interval and its conclusion. Total of 1 mark)

- **It is estimated with 98% confidence that the average of zinc content is between 23.0436 units smaller to 21.2102 units larger for vegetarian pregnant women than the average of the zinc content for non-vegetarian women who are pregnant. Since confidence interval include zero, there is no difference between vegetarian and non-vegetarian women who are pregnant.**

- **It is estimated with 98% confidence that the average of zinc content is between 35.1639 units smaller to 18.4305 units larger for vegetarian non-pregnant women than the average of the zinc content for non-vegetarian women who are not pregnant. Since confidence interval include zero, there is no difference between vegetarian and non-vegetarian women who are not pregnant.**

.

Final Conclusion: There is not difference in average of zinc content between vegetarian and non-vegetarian women. (1 mark)