STAT 651 Midterm Test (50 minutes) - 1:50pm - 2:40pm November, 2010

NAME: **Total number of Marks: /25**

There are 6 questions in this paper, do not be deterred, they are all straightforward. Read each question carefully. There are questions on both side of the page. The number of marks for each question are given in brackets. Be smart about how you answer. If you can't answer one question move on the to next and return to the questions you could not do after answering all the other questions!

Rubric: This exam is an open book exam you can use all written materials that you want, normal tables and a calculator.

Write your solutions in the question paper.

GOOD LUCK!

(1a) Suppose you want to test the equality of two population means against the alternative that the two population means are different (we will assume that both populations have the same variance). You do not observe the whole population, but you do observe samples of size m from population 1 and n from population 2. Given the information below, which sample sizes should I choose.

(A) The population variances is 1, $m = 50$ and $n = 50$.

(B) The population variances is 0.5 $m = 40$ and $n = 40$

(C) The population variance is 0.5 $m = 40$ and $n = 60$.

(D) The population variance is 1 $m = 45$ and $n = 55$.

[2]

(b) Xuan draws 300 samples, each sample is of size 30. For each sample, Xuan constructs a 95% CI, on average how many of these confidence intervals will contain the true mean?

[1]

(c) Jake reads about an opinion poll, that says the proportion of Americans who were optimistic about their future is $[0.58 \pm 0.1]$ with 95% confidence.

Jake asks you want they mean by 'the proportion of Americans who were optimistic about their future is $[0.58 \pm 0.1]$ with 95% confidence', and what would happen to this interval if the number of people sampled in the opinion poll increased. Answer his questions in four lines or less.

[2]

(2) Two brothers, Daniel and Joseph are having a debate about Daniel's frog. Daniel claims that his frog has psychic powers, because his frog has correcly predicted the outcome of the past four baseball matches (assume here that each match has only two teams, team A and team B, taking part, and the outcome can only be team A or team B winning - a draw is not allowed). Joseph disputes his brother's claim.

Help Joseph win the argument. By stating the null and alternative hypothesis for this scenario, use the ideas behind statistical testing to show that based on the frog's predictions, there is no evidence that Daniel's frog is psychic.

[3]

(3) Suppose I do a hypothesis test and I am unable to reject the null at the 5% level. Which statement is correct? [2]

(A) A type I error could have been committed, the p-value is less than 5% and I know the type I error.

(B) A type II error could have been committed, and the p-value is greater than 5%.

(C) A type II was commited, and the p-value is greater than 5%.

(D) A type II error could have been committed, the p-value is greater than 5% and I know the type II error.

(4) Let μ denote the population mean and $\bar{X}$ denote the sample mean. Suppose we observe a sample of 25 individuals, and the sample mean is $\bar{X} = 3$. Bill wants to test the research hypothesis that the mean is greater than two. He states the hypothesis as $H_0 : \bar{X} \leq 2$ against the alternative $H_A : \bar{X} > 2$. Did he write his hypothesis correctly? Give a reason for your answer.

[1]

(5) Ibrahim wants to investigate the time students spend on homework everyday. He draws a random sample of size n different students and asks how much time they spend on homework. He collects the data and puts it into JMP and runs a one sample test. The JMP output is given on the adjacent page. Based on the output, answer the following questions. Do the tests at the 5% level.

(i) Test the hypothesis $H_0 : \mu \leq 0.9$ hours against the alternative $H_A : \mu > 0.9$ hours. Do the test by calculating the p-value. State any assumptions you have made in order to do the test.

[3]

(ii) How large a sample size does Ibrahim need to use to ensure that the probability he will reject the null of the hypothesis in (ii) will be 0.85 when the true average is $\mu = 1.3$.

[3]

(6) Roopa is doing research on the amount of iron gained or lost after going on two different types of diets. She took of a sample of 20 volunteers and randomly allocated them into two groups (ten people in each group). Group 1 she put on a high vitamin C diet, the amount of iron gained or lost is given in the first row below (if the number is positive this is a gain, if the number is negative it is a loss). Group 0 she put on a high calcium diet, the iron gained or lost is given in the second row of the table below.

group										
1	0.51	2.75	0.79	4.41	-1.23	1.06	1.98	2.32	1.59	-18.41
0	-0.18	0.92	-0.25	1.56	-0.38	-0.21	-0.62	-1.68	-3.15	-0.33

(i) Roopa's research hypothesis is that there is a <u>difference</u> between the average amount of iron gained for the two diets. Let μ_1 denote the mean amount of iron gained on the vitamin C diet and μ_0 denote the mean amount of iron gained on the high calcium diet.

State Roopa's null and alternative hypothesis?

[1]

(ii) State precisely the statistical test(s) that you would you use to test this hypothesis. Give a reason for your answer, and the assumptions required.

[2]

(iii) Roopa put the data in JMP, and runs some tests. Her JMP output is given on the adjacent page. The tests are done at the 5% level.

What are the results of her tests, when she uses the t-test and the Wilcoxon test?

[2]

(iv) Do the two tests give different results? If this is the case, by looking at the data and the output, explain why this may be the case. By using the data and the output, which test gives the most reliable answer (give a reason for your answer).

[3]