Statistics 160 Midterm #2 – November 20, 2008

This exam has 14 problems on 8 numbered pages and is worth a total of 60 points.

You have 75 minutes to complete this exam. Please read all instructions carefully, and check your answers. Show all work neatly and in order, and clearly indicate your final answers. Answers must be justified whenever possible in order to earn full credit. **Unless otherwise specified, no credit will be given for unsupported answers, even if your final answer is correct.** *Points will be deducted for incoherent, incorrect, and/or irrelevant statements.*

This exam is closed-book, except that one $8\frac{1}{2} \times 11$ double-sided page of handwritten notes is permitted as well as a calculator. No other aids are allowed. Copies of Table A and Table C will be provided.

You must answer all of the questions in the space provided. Note that blank space is NOT an indication of a question's difficulty.

Name: _____

Instructor: Michael Kozdron

Problem	Score
1	
2	
3	
4 and 5 and 6	
7	
8 and 9 and 10	
11 and 12 and 13	
14	

TOTAL: _____

1. (*8 points*) A simple random sample of Internet users found that 736 out of 981 teenagers (ages 12 to 17) use instant messaging online, whereas 511 out of 1217 adults (ages 18 and older) use instant messaging online.

(a) Construct a 95% confidence interval using the plus four method for the true proportion of teenagers who use instant messaging online.

(b) Construct a 95% confidence interval using the plus four method for the true proportion of adults who use instant messaging online.

(c) By comparing your confidence intervals in (a) and (b), is there sufficient evidence to conclude that teenagers are much more likely to use instant messaging online than adults. Why?

2. (*8 points*) A fictitious poll of how many hours of housework people do per week got responses from 81 men and 92 women. The two sample histograms were approximately symmetric and the data contained no outliers. Furthermore, the responses of the men gave a sample mean of 3.2 hours with a sample standard deviation of 0.5 hours. For women, the numbers were 3.6 and 0.52, respectively.

(a) Estimate the difference in hours of housework that men and women do, with 98% confidence.

(b) In what sense do you put 98% confidence in your intervals?

(c) Based on your confidence interval from **(a)**, is there significant evidence to conclude that women do more housework than men? If so, at what significance level?

3. (*8 points*) A popular magazine has published an article where it states that, in Canada, a married male is usually older than his wife. Statistics Canada collects data on the ages of married people. Suppose that 40 couples are randomly selected and their ages are recorded.

How would you analyse the collected data to see whether or not the claim of the magazine holds? State clearly

- the null hypothesis and the alternative hypothesis,

- the statistical procedure you would use,

- the experimental design upon which this analysis would be based, and

- the assumptions necessary for the analysis you propose to hold.

Comment on the adequacy (or inadequacy) of this design.

For each of the following problems, circle the BEST answer.

4. (*2 points*) Bob and Carol would both like to estimate the mean weight gain of beef cattle fed a new *Lean Cuisine* diet. They both believe that the weight gain standard deviation is about 30 kg, and they will both perform their statistical inferences at the same confidence level. However, Carol desires a margin of error for her estimate which is only half as big as the margin of error that Bob wants. So, compared with Bob's sample size, Carol's sample size should be

   A half as big;

   B twice as big;

   C four times as big;

   D impossible to tell.

5. (*2 points*) A medical researcher wants to estimate the mean serum cholesterol level of middle-aged males in each of several provinces (e.g., Ontario, Nova Scotia, Saskatchewan) based on random samples. The researcher believes that the standard deviations are about the same in each of the provinces, and wants the standard errors to be about the same. If the sample size chosen for Nova Scotia is 0.1% of its middle-aged male population, what should the corresponding sample size for Ontario be?

   A Exactly 0.1% of the middle-aged male population of Ontario.

   B A lot more than 0.1% of the middle-aged male population of Ontario since Ontario's population is much bigger than Nova Scotia's.

   C A lot less than 0.1% of the middle-aged male population of Ontario since Ontario's population is much bigger than Nova Scotia's.

6. (*2 points*) Based on a simple random sample from a population with unknown mean μ, one obtains $[33, 45]$ as a 95% confidence interval for μ. Which of the following statements is true?

   A The probability that the sample mean falls between 33 and 45 is 0.95.

   B The probability that μ falls between 33 and 45 is 0.95.

   C The probability that the sample mean for a future sample falls between 33 and 45 is 0.95.

   D If a million independent random samples were taken from the population and a 95% confidence interval for μ calculated using each sample, then the proportion of intervals covering μ would be close to 0.95.

7. (*10 points*)

(a) A monkey is sitting at a specially designed keyboard which has 26 keys, one for each letter in the alphabet. Assuming that the monkey types at random, and ignoring punctuation and capitalization, what is the probability that he produces a copy of *Hamlet*? You may assume that *Hamlet* is 70 pages long at 30 lines per page and that each line has 80 letters. *You do not have to compute a value; just set up as much of the computation as you can.*

(b) Prof. Mobed of the Physics Department here at the U of R noted that the mass of a nucleon (a proton or neutron) is about 1.67×10^{-27} kg and that the mass of the sun is about 2×10^{30} kg, suggesting that there are about 1.2×10^{57} nucleons in the sun. Which event is more likely: randomly choosing a particular nucleon in the sun or a monkey producing *Hamlet*?

(c) After the monkey has typed as many characters as there are in *Hamlet*, it is noted that the proportion of pure vowels (a,e,i,o,u) that he typed is 3.2%. What parameter does 0.032 estimate? What does this tell you about the randomness of the monkey's typing?

For each of the following problems, circle ALL answers that are correct.

8. (*2 points*) The sampling distribution of the sample mean is

    A approximately normally distributed for large sample sizes;

    B normally distributed;

    C t-distributed when σ is unknown;

    D derived from the population distribution;

    E its distribution.

9. (*2 points*) The sample size required to estimate a population mean to a specified degree of accuracy

    A increases with the population size;

    B depends on the standard deviation of the population;

    C depends on the confidence level you want;

    D is obtained from the P-value;

    E is usually an approximate procedure.

10. (*2 points*) The hypothesis test $H_0 : \mu = 0$ against $H_a : \mu > 0$

    A can be based on a t-distribution if σ is unknown and the population is nearly normal;

    B can be based on a normal distribution if σ is known and the population is nearly normal;

    C always gives a P-value between zero and one;

    D always gives a P-value less than 0.5;

    E has a probability of a Type I error which we generally have ignored.

For each of the following problems, circle ALL answers that are correct.

11. (*2 points*) In a hypothesis test,

   A Type I errors are the same as Type II errors;

   B Type I errors are usually more serious than Type II errors;

   C Type II errors occur with probability equal to the P-value;

   D Type I errors occur with probability equal to the P-value;

   E Type I errors are controlled by using a level of significance.

12. (*2 points*) A t-statistic for a hypothesis test

   A always increases with the sample size;

   B decreases as the sample standard deviation decreases;

   C has degrees of freedom determine by the sample size(s);

   D is always positive;

   E is always positive when you can reject the null hypothesis.

13. (*2 points*) Every valid hypothesis test

   A has a test statistic;

   B has a P-value;

   C has hypotheses which are disjoint;

   D will reject the null hypothesis;

   E controls for the probabilities of Type I errors.

14. (*8 points*) Michael is a pizza inspector for the *Saskatchewan Health Authority*. He has received numerous complaints about a certain pizzeria for failing to comply with its advertisements. The pizzeria claims, on the average, that each of its pizza is topped with 200 grams of pepperoni. The dissatisfied customers feel that the actual amount of pepperoni used is considerably less. To settle the matter, Michael decides to do a hypothesis test. He assumes that the distribution of pepperoni weights on a pizza is normal with standard deviation $\sigma = 25$ grams and wants to draw conclusions about the mean weight μ. His objective is to test $H_0 : \mu = 200$ against $H_a : \mu < 200$. For data he will collect a sample of 4 large pizzas and find the mean weight of pepperoni. If the mean weight of these 4 pizzas is less than 180 grams, he will reject H_0. At what significance level α is Michael is conducting his test?