MATH 160A   Introduction to Applied Statistics   Fall 2008   Exam #4

Instructions: This exam is a tool to help me (and you) assess how well you are learning the course material. As such, you should report enough written detail for me to understand how you are thinking about each problem.   (100 points total)

---

1. You and a friend are debating whether women favor universal health insurance more than men (in the U.S.). Your friend finds the results of a sample survey study in which the proportion of women who favor universal health insurance is 56% and the proportion of men who favor universal health insurance is 52%. (The survey was done with appropriately selected random samples.) Your friend claims that these values from the sample survey give conclusive evidence that the proportion who favor universal health insurance is greater for women than for men. Explain to your friend what essential element is missing here and what additional information from the study is needed to give a meaningful interpretation of the survey results. Use languague your friend can understand assuming that your friend has never had a statistics course.   (9 points)

2. You work for a polling company that has been hired to get information on the current proportion of adults older than 25 in the United States who have earned a bachelor's degree. You know that this proportion has never been more than 30% in the past. The goal for this study is to produce a 99% confidence interval having a margin of error less than ±1.5%. What sample size would you recommend?   (9 points)

3. An institute that runs a four-week language immersion program for high school French teachers in summers wants to assess the impact of the program. In a particular summer, 20 teachers are participating. A standardized test of understanding for spoken French was given to each teacher at the beginning of the program and again at the end of the program. The test has 36 items. The "post-test minus pre-test" scores for the 20 teachers are

   2   0   6   6   3   3   2   3   1   6   6   6   3   0   1   1   0   2   3   3

   For this data, the mean is 2.850 and the standard deviation is 2.134.

   (a) What broad question is being asked here?   (3 points)

   (b) Identify the specific population and parameter of interest in this situation.   (4 points)

   (c) Give your judgment on whether or not the methods we have available are appropriate for computing a confidence interval for the mean score in this situation. Briefly explain how you reach your judgment.   (6 points)

   (d) Compute a 95% confidence interval for the mean score change.   (12 points)

   (e) Give a brief (one or two sentence) conclusion using non-techincal langauge.   (5 points)

4. Observational studies have indicated a connection between calcuium intake and blood pressure so researchers design an experiment to test whether or not increasing calcium intake reduces blood pressure in men. In the double-blind experiment, each subject will be randomly assigned to take either a daily calcium supplement or a placebo for 12 weeks. For each subject, systolic blood pressure will be measured at the beginning of the study and at the end. The change in blood pressure (final minus initial) for each subject will then be computed.

   (a) Identify the populations and parameters of interest in this situation.   (4 points)

   (b) Set up null and alternative hypotheses for this situation.   (4 points)

   (c) The researchers carried out the experiment with a total of 21 subjects; 10 were assigned to take calcium and 11 were assigned to take the placebo. The change in blood pressure (final minus initial) is given in the table for subjects in each group along with the mean and the standard deviation.

   Calcium group

7	−4	18	17	−3	−5	1	10	11	−2

   mean = 5.00   standard deviation = 8.74

   Placebo group

−1	12	−1	−3	3	−5	5	2	−11	−1	−3

   mean = −0.27   standard deviation = 5.90

   Use this data to carry out a significance test on the hypotheses you set up in (b).
   (12 points)

   (d) Give a brief (one or two sentence) conclusion using non-technical language.   (5 points)

5. Every two years, the Josephson Institute does a study of ethics among high school students based on a survey given to a sample of high school students. One of the questions on the survey asks "How many times have you cheated on a test at school in the past year?" Response options are "Never", "Only once", and "Two or more times". Here are the sample sizes and counts of "Never" for the two most recent surveys:

Year	Sample size	Count of "Never"
2006	34,105	13,586
2008	28,718	10,413

   (a) Carry out some relevant statistical inference (confidence interval or significance test, your choice) on this data to address the question of whether high school student cheating has changed in the last two years. Include a check of conditions that the method you choose is appropriate. Give a non-technical conclusion written as a complete sentence (or two).   (22 points)

   (b) The survey was conducted by randomly selecting 100 high schools in the United States and then having all students in those schools complete the survey during class. The students were assured that their responses would be anonymous. Given the methodology of this study, describe one source of potential bias that might affect the question above. Also describe the possible direction of bias for this source.   (5 points)