Math 160 K

FINAL EXAM

NAME_____

General Notes:

1. Show work. A correct answer without supporting work may not be given credit.
2. Look over the test first, and then begin.
3. Calculators are permitted on this exam, but only for basic arithmetic (i.e., no statistical calculations)

Monday, December. 14, 2009
200 pts.

I.   The first thing we do with a new data set is to examine it closely and see what sort of stories it tells us. This first batch of questions considers this task.

1.   Consider the following series of numbers:
     67 / 75 / 79 / 85 / 85 / 95 / 97 / 98 / 98 / 99

     a.   (10 pts.)   Give the five-number summary for these data

     b.   (10 pts.)   Sketch a box plot for the preceding five-number summary.

(problem 1 continued)

c.　　(5 pts.) What is an outlier?

d.　　(10 pts.)　　What condition in this set of data would we use to determine if some score is an outlier? (i.e., above what value or below what value)

e.　　(5 pts.) Calculate the average of this set of numbers

(problem 1 continued)

f.   (15 pts.)   Give a formula for computing the standard deviation of a set of numbers, and use the formula to calculate the standard deviation of this set of numbers.

2.   One very important way to look at data is to construct a picture of it (i.e., to graph it). What sorts of graphs do we use (5 pts. each)

    a. For categorical variables?

    b. For quantative variables?

II.   We frequently look for a relationship (or lack thereof) between two variables.

1.   One way to compare two quantative variables to see if there is a possible relationship
         between them is to test for correlation and compute a least-squares regression line.

         a.   (5 pts.) What is **r** and how is it calculated? (i.e., what is the formula for
                  calculating **r**?)

         b.   (5 pts.) What information does r^2 give us?

         c.   (5 pts.) What is a least-squares regression line? What does it minimize?

(problem 1 continued)

d. (15 pts.) Suppose that $r = 0.7$, $s_y = 3$, that $s_x = 1.5$, and that $\bar{x} = 1, \bar{y} = 3$. Give formulas for the y-intercept and slope of the least-squares regression line and calculate the equation for the least-squares regression line in this case.

e. (5 pts.) What is the predicted y-value for an x value of 2 in the above?

III.   Distributions

1.   (5 pts.) What is a distribution / density function?

2.   (10 pts. each). Over the course of the semester we have discussed several different
         distributions. Briefly describe each of the following. As a *part* of your answers, say what
         the parameters mean and say when each distribution is appropriate to use.

         a.   Binomial distribution B(n, p)

(problem 2 continued)

b.   Normal distribution N(μ, σ)

c.   Student's t distribution T(n)

IV.	Sampling

1.	(10 ps.) What is a simple random sample (SRS) and how can we get one?

2.	(5 pts.) A simple random sample of 5 students from a class of 30 students numbered 01, 02, ..., 30 is to be selected. Use Table B starting at line 151 to select these five students.

V.   Confidence

1.   (10 pts.)   Suppose that we are interested in the average weight of a honey bee, and collect a SRS of 25 bees from a hive and measure the weight of the bees in the sample. Suppose further that we find the average weight of the bees in this sample to be 0.22 grams is taken from a population with unknown mean but with a standard deviation of 0.05. Suppose that the standard deviation of the weights of the population of honey bees is 0.05 (I'm guessing on the standard deviation: the weight is from Wikipedia). Give a 95% confidence interval, expressed either as an interval or a margin of error (your choice) for the population average. The average (from Wikipedia) is 0.218. Is that within your confidence interval?

2.   (10 pts.)   Suppose now that I do not know the actual standard deviation, but I calculate it from my sample to be (surprisingly enough) to be 0.05. I need to use Student's t-distribution. How many degrees of freedom do I have? What is my 95% confidence interval for the average in this case?

VI.   Hypothesis testing

1.   (10) Explain what Type I and Type II errors are (give definitions)

2.   (5 pts). If the power of a particular test at the $\alpha = 0.05$ level is 80%, what is the likelihood of a Type I error? Of a Type II error?

VII   Essay questions   (short answers and incomplete sentences not given very much credit).
       Pick one of the following two questions and respond to it in a paragraph or so. (15 pts.)

1..   What is the study of statistics all about?

2.   What is an IRB and what does it do? Do we have any evidence that ethical practices need
       to be specified?