Statistics 1: Sections 4 and 6
CSUS, Fall 2008

Practice Problems for Exam 1

Note: These practice problems are NOT a comprehensive list of the material covered for Exam 1. You are responsible for all the sections we covered in Chapter 1.1 – 4.6 of the text. These problems are meant to be a sample of the type of problem you <u>might</u> see on the exam. Also, the length of this set of problems is longer than the exam will be. As a rough guide, you should allow 70-75 minutes for this set of problems.

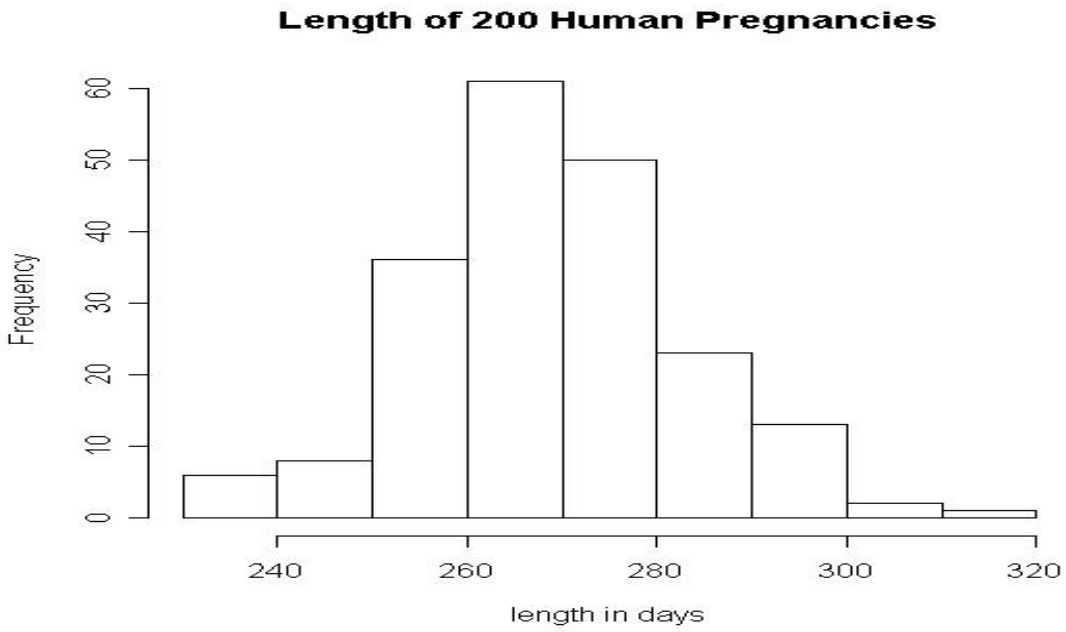
One 8.5"x11" sheet of handwritten notes allowed. Calculators allowed. The actual exam will have space for your work on the exam paper itself.

1) For the following dataset: 10,6,2,7,100

a) Calculate the mean and median.
b) Which is a better measure of central tendency, the mean or median? Why?
c) Calculate the standard deviation using the formula $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

d) If you added a fixed number, k, to each value in the data set, would the standard deviation change? If so, how?
e) A portion of the ordered exam scores for a class of 100 students is shown below. Will removing the lowest score affect the standard deviation? If so, how?
6, 59, 62, 63, 67, 70,…, 96, 99

2) Use the histogram to answer the questions below.



a) The value 240 represents the upper boundary of the left-most rectangle in the histogram. What is the upper boundary of the second rectangle from the left?

b) Use the histogram to estimate the *relative frequency* of the category having upper bound 260.

c) Use the histogram to estimate the total number of pregnancies lasting 260 days or less.

d) The mean and standard deviation for this data are 269 days and 14.2 days, respectively. Calculate mean +/- two standard deviations for this data.

e) Use the histogram to approximate the percent of data falling in the interval calculated in part (d).

f) Which would give you the best estimate of the percent of data falling in the interval in part (d) -- Chebychev's Rule or the Empirical Rule? Give reasons for your choice then use the chosen rule to estimate the percent of data in the interval from part (e).

3) (20 points) A random sample of voters in Clark County were classified by political party and gender. The results are shown below. What is the probability a randomly selected voter is

	Democrat	Republican	Other
Male	27	36	7
Female	18	10	2

a) A democrat?

b) A female given the person is a democrat?

c) A democrat or a female?

d) Are the events voter is a democrat and voter is a republican mutually exclusive? Why or why not?

e) If two voters are randomly selected without replacement, what is the probability

i) They are both democrats?
ii) At least one is a democrat?

4) The following data give exam scores for 28 students in a CSUS mathematics course, arranged from lowest to highest.

47	49	49	53
57	61	63	66
66	72	73	75
75	76	79	81
82	85	85	85
85	88	88	89
90	95	95	100

a) Find the lower quartile, median and upper quartile of these scores.

b) Construct a stem-and-leaf diagram for this data.

5) There are three people on a committee: Al, Bob and Carl. One person is selected to bring donuts for the next committee meeting and a person is selected to take minutes at the meeting. The same person may be assigned both tasks. Assume these selections are made at random.

a) List the outcomes in the sample space, S.
b) What is the probability Bob brings donuts and Carl takes minutes?
c) What is the probability the same person is assigned both tasks?
d) What is the probability Bob is assigned at least one task?

e) What is the probability Bob is assigned no task?

6) Randomly selected girls are given the Wide Range Achievement Test. Their ages are listed along with their test scores as well as some calculations useful for this problem.

	Age, x	Score, y	xy	x^2	y^2
	6.1	17.8	108.58	37.21	316.84
	7.2	47.4	341.28	51.84	2,246.76
	5.9	25.8	152.22	34.81	665.64
	6.3	24.3	153.09	39.69	590.49
	10.5	66.6	699.30	110.25	4,435.56
	11.0	91.4	1,005.40	121.00	8,353.96
Total	47.0	273.3	2,459.87	394.80	16,609.25

a) Calculate the linear correlation coefficient, r, and interpret it.
b) Calculate the best fit regression line.
c) What does the slope of the regression line tell us about the relationship between age and score?
d) Predict the exam score for a girl who is 9 years old.

7) Identify the following variables as qualitative or quantitative. Classify the quantitative variables as discrete or continuous.
a) Marital status
b) Time until an alkaline battery wears out
c) Brand of an alkaline battery
d) Whether or not a subject has disease X
e) Number of votes a political candidate receives
f) The pounds of sugar consumed by a person in a week