STOR 155, Section 1, Final Examination
Thursday, April 30, 2009

Name: _____

Pledge: I have neither given nor received aid on this examination.

Signature: _____

Instructions: Do not do any actual numerical calculations. Answers in a form that you would type into an Excel field, such as "=28*SQRT(82)^2", with a *working* answer, are expected.

1. A company makes 20% of its cars at factory A, and the rest at factory B. Factory A produces 1% lemons, and Factory B produces 2% lemons. A car is chosen at random. What is the probability that:

   a. It came from Factory B?

   b. It is a lemon, if it came from Factory B?

   c. It is a lemon, from Factory B?

   d. It is a lemon?

   e. It came from Factory B, if it is a lemon?

2. A survey of 2000 student loan borrowers found that 200 had loans totaling more than $40,000.

   a. Give a 99% best guess Confidence Interval for the proportion of all loans totaling more than $40,000.

   b. Give an Excel expression for the exact p-value for concluding that the proportion of all loans is more than $40,000.

   c. Use the Normal approximation to give an alternate answer to (b).

   d. Why is the approximation used in (c) appropriate?

   e. What is the 98% conservative margin of error in estimating the proportion of all loans?

3. Scores on tests for a class were:

	A	B	C	D	E
1	1st Exam	153	144	162	127
2	2nd Exam	145	140	143	130
3	Difference	8	4	19	-3

a. Assuming each column represents one student, give a formula for the p-value to show that scores on the 1st exam are significantly higher than those on the 2nd exam.

b. Again assuming each column represents one student, give an 80% Confidence Interval for the difference between the mean scores.

c. Assuming the exam scores come from two different classes, give a formula for the p-value to assess whether exam scores are significantly different between the two exams.

d. Write the equation of the least squares regression line, of the 2nd score as a function of the 1st score, in terms of Excel commands

e. Write an Excel command to calculate the correlation between exam scores. Will the answer be positive, 0 or negative?

4. For a random variable with distribution:

y	-1	0	1	3
f(y)	0.5	0.2	0.2	0.1

Find:

a. $P\{-1 < Y \leq 2\}$

b. $P\{Y = 1 \mid Y > 0\}$

c. $P\{Y = 1 \mid Y < 0\}$

d. The expected value of X

e. The standard deviation of X

5. A TV ad claims that at most 30% of people prefer Brand X. Suppose that 6 out of 10 randomly selected people prefer Brand X.

   a. Give an exact p-value to decide whether or not we should dispute the claim.

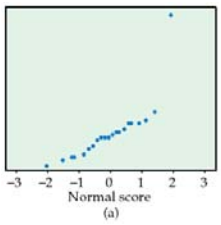
   b. If the p-value in part a turns out to be 0.03, give a "yes-no" conclusion.

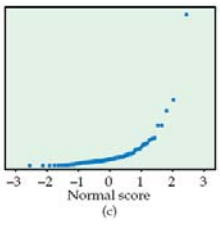
   c. If the p-value in part a turns out to be 0.03, give a gray level conclusion.

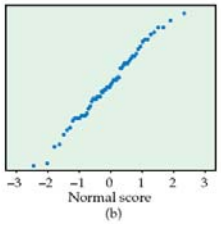
   d. Give a 98% conservative confidence interval for the proportion of people that prefer Brand X.

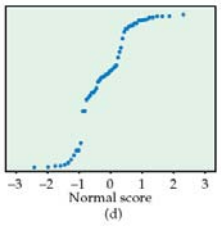
   e. How large a sample (in the best guess sense) is needed so that with probability 90%, the estimated proportion of people that prefer Brand X is within 0.01 of the actual number?

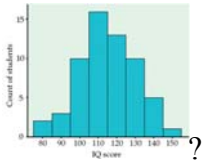
6. A set of 4 Normal Quantile plots (in scrambled order, watch the labels) are:



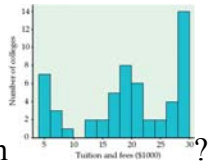






i. Which most likely is from a data set of IQ scores with histogram  ?

ii. Which most likely is from a data set of 21 car mileages of 2-seater cars?

iii. Which most likely is from a data set of tuition charges with histogram  ?

iv. Which most likely is from a data set of 80 phone call lengths with histogram  ?

v. Which most likely is from a Normally distributed data set?

vi. Which most likely is from a Normal distribution, with a single outlier?

vii. Which most likely is from a distribution with multiple clusters?

7. Length of horse pregnancies vary according to a roughly Normal distribution, with mean 340, and standard deviation 5.

   a. Use the 68-95-99.7 rule to indicate the range which contains 95 % of the data.

   b. Use the 68-95-99.7 rule to indicate which % of pregnancies last at least 335 days.

   c. Give an Excel command to answer part (a).

   d. Give an Excel command to answer part (b).

   e. How large a sample should be used to be 98% sure of estimating the true mean within 0.1?

8. Gas mileages for a vehicle, after a random sample of fill-ups are:

	A	B	C	D	E	F
1	41.5	50.7	36.8	44.2	45	37.4

a. Find the sample mean and standard deviation.

b. Find the 60% margin of error in estimation of the population mean.

c. Give a 60% Confidence Interval for the population mean.

d. Find the p-value to test whether the population mean is less than 40.

e. Briefly state (5 words or less) the needed assumptions in parts (c) and (d).