Stat 311 Midterm   Name: **KEY**

13 February 2008   Section:_________________

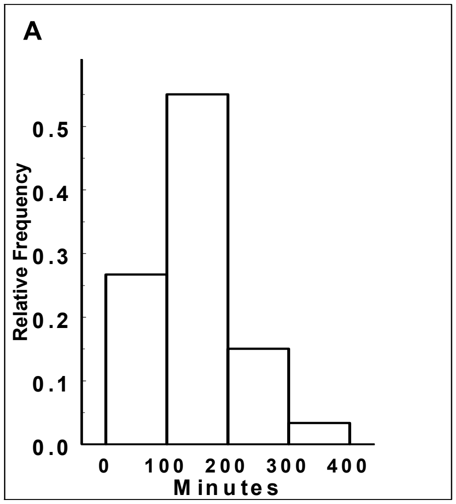
This is a closed book exam. You are allowed, however, to use the formula card that came with the textbook and to have one sheet (double-sided) of 8.5 x 11 paper with notes, either handwritten or typed. You may also use a calculator, although be sure to show your work. The exam consists of five problems worth a total of 100 points. Point values for each part of a question are designated in parentheses at the beginning of the problem.

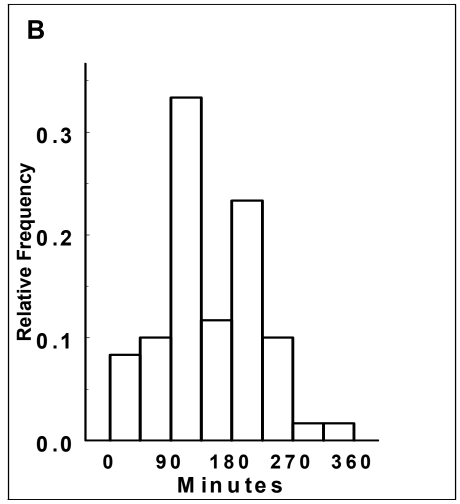
Be sure to show your work as indicated in order to receive credit.

Problem 1 (5+5+5+5 points): Thirty women and thirty men were randomly selected from a large first-year college class. Each student in the sample was asked how many minutes they studied on a typical weeknight. Summary statistics from the sample responses, by gender, are shown below.

Gender	Min	Q1	Median	Mean	Q3	Max
Female	60.0	120.0	175.0	165.2	180.0	360.0
Male	0.0	60.0	120.0	117.2	150.0	300.0

A couple of histograms generated from the sample data are shown below. Histogram A is on the left; histogram B is on the right. Both histograms include all 60 data values. Histogram A was the default plot generated by a statistical software package. Histogram B was customized by requesting that the plot have eight bars.





a) Describe the shape of the distribution shown in histogram A.

Unimodal with a slight right-skew.

b) Describe the shape of the distribution shown in histogram B.

Bimodal with the same slight right-skew or approximately symmetric.

c) How might you explain the differences between the distributions shown in histogram A and B?

The bimodality in histogram B may be due to a difference in minutes studied between males and females. The mean and median minutes studied for females are considerably higher than minutes studied for males. By using more bins, this difference comes through in the plot.

d) Fill in the blank: using the summary statistics for females, a female student exceeding 270 minutes of study time per typical weeknight would be considered a potential outlier. **Show your work.**

IQR(females) = 180 -120 = 60
1.5(IQR) = 1.5(60) = 90
Upper limit = 180 + 90 = 270

Problem 2 (10+10 points): Based on 2007 statistics collected by a small liberal arts college, we know that the probability that a student is female is 60%. The probability that a student is studying statistics is 7%. The conditional probability that a student is female, given that the student is studying statistics is 71%.

a) Are the events "being a woman" and "studying statistics" independent? **Explain using an appropriate equation.**

Let F be the event a student is female and S be the event a student studies statistics

Given: $P(F)=0.6$; $P(S)=0.07$; $P(F|S)=0.71$

Since $P(F|S)=0.71 \neq P(F)=0.6$, F and S are not independent

Or show $P(F \cap S) \neq P(F)P(S)$

b) For a student chosen at random, what is the conditional probability that the student is studying statistics, given the student is female? **Interpret your probability in one sentence.**

Want $P(S|F)=\dfrac{P(F|S)P(S)}{P(F)}=\dfrac{0.71(0.07)}{0.60}=0.083$

Approximately 8% of females study statistics.

Problem 3 (8+7 points): A group of high school students are conducting a group project for their psychology class to determine teen attitudes toward the 18-year-old voting age. The school has 840 students under the age of 18 and 120 students over the age of 18. The group wants to randomly select a sample of students to interview and decides to randomly select 84 students under the age of 18 and separately, randomly select 12 students over the age of 18.

a) Why does the sample not qualify as a simple random sample?

This is not a SRS because not all sample of size 96 are equally likely. For example, you cannot have a sample of 96 students under the age of 18.

b) If the group had decided on selecting a simple random sample, how many different simple random samples are possible? **Just set up an expression for the number without solving for it.**

$\binom{960}{96}=\dfrac{960!}{96!(864!)}=$ a big number!

Problem 4 (6+8+6 points): The TV remote is dead and needs new batteries. In a group of 20 batteries lying around in your closet, you know that 6 are dead. You choose 4 batteries at random for your TV remote. Let *X* denote the number of good batteries selected. Assume *X* can be modeled as a binomial random variable.

a) What are the parameter values for *X* (i.e. what are *n* and *p*)?

$n = 4$; $p = 1 - (6/20) = 14/20 = 0.7$

b) Calculate the exact probability that the TV remote will NOT work once the batteries are replaced (assume that all four batteries must be good for the remote to work).

$P(\text{remote does not work}) = P(X \le 3) = 1 - P(X = 4).$

$$P(X=4) = \binom{4}{4} 0.7^4 (0.3)^0 = 0.2401$$

$$P(X \le 3) = 1 - .02401 = 0.7599$$

Also accept exact probability = 1 – (14/20 x 13/19 x 12/18 x 11/17) = 1 - 0.2066 = 0.7934

c) Calculate the mean and standard deviation of *X*.

$$\mu_X = np = 4(.7) = 2.8$$

$$\sigma_X = \sqrt{npq} = \sqrt{4(.7)(.3)} = 0.92$$

Problem 5 (5+5+5+10 points): Carbon monoxide (CO) emissions for a certain kind of car vary with a mean 2.9 g/mi and standard deviation 0.4 g/mi. Assume that CO emissions follow a normal distribution.

a) For a randomly selected car of the type described above, what is the probability that CO emissions exceed the EPA limit of 3.4 g/mi?

$$P(X > 3.4) = P\left(Z > \frac{3.4 - 2.9}{0.4}\right) = P(Z > 1.25) = 1 - P(Z < 1.25) = 1 - 0.8944 = 0.1056$$

b) What is the probability that the CO emissions for a randomly selected car fall between 2.9 and 3.5 g/mi?

$$P(2.9 < X < 3.5) = P\left(0 < Z < \frac{3.5 - 2.9}{0.4}\right) = P(0 < Z < 1.5) = P(Z < 1.5) - P(Z < 0) = 0.9332 - 0.5000 = 0.4332$$

c) What is the probability that the CO emission reading for a randomly selected car is 2.9 g/mi?

$$P(X = 2.9) = 0$$

d) Fill in the blank: only 3.6% of cars will be expected to have CO emissions above <u>3.62</u> g/mi. **Show your work.**

$$P(X > x_0) = 0.036 \Rightarrow z_0 = 1.80$$

$$P(Z > z_0) = P\left(Z > \frac{x_0 - 2.9}{0.4}\right) = 0.036$$

Thus, $1.80 = \dfrac{x_0 - 2.9}{0.4} \Rightarrow x_0 = 1.80(0.4) + 2.9 \Rightarrow x_0 = 3.62$