Statistics 8
Sample Midterm 1

NAME:_____

Student ID#:_____

One page of notes allowed. *You should have 8 problems on 3 pages.* Each part of each problem is 5 points.

1. Eleven students were asked to measure their pulses for 30 seconds and multiply by two to get their one-minute pulse rates. The results were: 32, 60, 62, 66, 70, 72, 74, 74, 78, 80, 84.
   a. Create a five-number summary for these pulse rates.

   b. The pulse rate of 32 is an outlier. The student forgot to multiply by two. Which one of the three reasons for outliers discussed in class and in the book applies here?

   c. Correct the outlier and redo the five-number summary.

   d. For each of the following statistics, specify whether it would increase, decrease or remain the same after the outlier is corrected.

      Mean: _____ Median: _____

      Standard deviation: _____ Interquartile range: _____

      Standardized score for a pulse rate equal to the mean of the current data: _____

2. A study done by the Center for Academic Integrity at Rutgers University asked 2116 students at 21 colleges and universities a series of questions. Some of the schools had an "honor code" and others did not. Of the students at schools with an honor code, 7% reported having plagiarized a paper via the Internet, while at schools with no honor code, 13% did so. (*Sacramento Bee*, Feb 29, 2000, D1.)
   a. Was this a randomized experiment or an observational study?

   b. What are the explanatory and response variables for this study?

   c. Give an example of a possible confounding variable and explain why it fits the criteria for a confounding variable.

3. Over many years, the average rainfall during the month of November in San Francisco, California, is 2.62 inches. The standard deviation is 2.79 inches. Based on this information, explain how you can tell that the distribution of rainfall values cannot be bell-shaped.

4. Math SAT scores for students admitted to a university are bell-shaped with a mean of 520 and a standard deviation of 60.
   a. Draw a picture of these SAT scores, indicating the cutoff points for the middle 68%, 95% and 99.7% of the scores.

   b. A student had a math SAT score of 490. Find the standardized score for this student *and* draw where her score would fall on your picture in part (a).

5. The table below shows the opinions of 908 respondents in the General Social Survey to the question "Do you believe there is life after death?" The purpose of examining the data is to see if there is a gender difference in how people would respond to this question.

	Yes	No	Total
Male	282	109	391
Female	408	109	517
Total	690	218	908

   a. Write the null and alternative hypotheses for this study.

   b. Find the expected count for the number of males who believe there is life after death.

   c. The chi-square statistic for this situation is 5.63 and the *p*-value is 0.018. State the appropriate conclusion in the context of this situation.

6. In the previous question, there were 517 females in the sample and 408 of them said they believe there is life after death.
   a. What is the margin of error for this survey (for females only)?

   b. Compute a 95% confidence interval for the <u>percent</u> of females in the population who believe there is life after death.

7. In a study of acupuncture for treating pain, 100 volunteers were recruited. Half were randomly assigned to receive acupuncture and the other half to receive a sham acupuncture treatment. The patients were followed for 6 months and the treating physician measured their degree of pain relief. The patients did not know which treatment they actually received, but the treating physicians were aware of who was getting acupuncture and who wasn't. For the following list of terms, *circle* the ones that apply to this study and *cross out* the ones that do not apply to this study:

   Randomized experiment, Observational study, Blocking, Single blind, Double blind

8. The percent of high school graduates who took the SAT exam in 1998 varied widely from state-to-state (including Washington DC), with a high of 83% in Washington DC and a low of 4% in Mississippi and Utah. The regression equation relating the average 1998 verbal SAT score in the states to the percent of graduates who took the SAT in the state is:

$$\text{Verbal Average} = 573 - 1.08\ (\text{PercentTook})$$

   a. Is the association between the percent who took the exam and the average verbal SAT score positive or negative? Explain why, logically, that would be expected.

   b. In California 47% of high school graduates took the exam. What is the *predicted* average verbal SAT score for a state in which 47% of high school graduates took the exam?

   c. The average verbal SAT score for California was 497. What is the *residual* for California?

   d. Does the *intercept* of 573 have a logical interpretation in this situation? Explain.