Stat 3011 First Midterm Exam **October 18, 1999**

Name _____ Student ID _____

The exam is closed book. You may use a calculator, and one $8\frac{1}{2}$ by 11 sheet of paper with formulas (or anything else) on it, but no other notes. Put all of your work on this test form (use the back if necessary). Show your work or give an explanation of your answer. No credit for numbers with no indication of where they came from.

The points for the questions total to 100. There are 3 pages and 6 problems.

1. [20 pts.] The following data are household incomes for 10 households in thousands of dollars

   15 8 30 41 33 150 61 46 38 44

   (a) Calculate the sample mean for these data.

   (b) Calculate the sample median for these data.

   (c) Calculate the 10% trimmed mean for these data.

   (d) Give one good reason why the mean is not as good an estimate of the "center" of the data as the median or the trimmed mean.

2. [10 pts.] Explain what the phrase "correlation is not causation" means, and explain what this implies about linear regression.

3. [15 pts.] The following is a frequency distribution table for weights of widgets in ounces for a sample of size 100. Only the first two columns filled in.

Class	Frequency	Relative Frequency	Cumulative Relative Frequency
$30 -< 35$	20		
$35 -< 40$	30		
$40 -< 50$	40		
$50 -< 60$	10		

(a) Fill in the last two columns of the table: relative frequency and cumulative relative frequency.

(b) Draw a histogram using these class intervals. Include correct axes and axis labels.

4. [20 pts.] Suppose the scores on the first and second midterms in a Stat 3011 class had the following summary statistics

	mean	standard deviation
first midterm	85	6
second midterm	83	8

and the correlation between first and second midterm scores was 0.6.

(a) Find the equation of the least-squares regression line for predicting the score on the second midterm from the score on the first midterm.

(b) Use the regression equation to predict the second midterm score for a student who got 80 on the first midterm.

(c) What is the fraction of variance explained by this regression?

5. [15 pts.] Suppose the probability a Cheapo brand 100 watt light bulb lasts more than 1000 hours before burning out is 0.25, and suppose light bulb failures are statistically independent. Suppose I install four such light bulbs in my house.

(a) What is the probability all four last more than 1000 hours?

(b) What is the probability at least one lasts more than 1000 hours?

6. [20 pts.] Let z be a random variable with a standard normal distribution.

(a) What is $P(z < -2.345)$?

(b) What is $P(.67 < z < .89)$?

(c) What is the 20th percentile of the standard normal distribution?