Math 207
Introduction to Statistics

Exam II

Please show your work and any R commands that you use.

1. **Standard Normal Curve**: Sketch and compute the specified area under the normal curve. Write down the R command that you use. (15 points)

a) $z < -1$

b) $z > 1.75$

c) $-1 < z < 1.75$

d) What z is greater than 51% of all z scores?

2. **Normal Curve**: Suppose that for a particular pet population subjects' weights average out to 52 pounds with an sd of 5 pounds and have a normal distribution. (15 points)

a) What percent of the weights were below 47 pounds?

b) What percent were above 62 pounds?

c) What percent were between 47 and 62 pounds?

d) What weight was greater than 51% of all the weights?

3. **Correlation**: (10 points)

a) For a representative sample of cars, would the correlation between the age of the car and its gasoline economy (miles per gallon) be positive or negative? Explain

b) The correlation between gasoline economy and income of owner turns out to be positive. How do you account for this association?

4. **RMS Error for a Line**: (15 points) Consider the following data.

x	y
0	2
1	4
2	3

a) Make a scatter plot of the data.

b) An investigator uses the line $y = 2\,x + 1$ to predict y from x. Add the line to your sketch.

c) Calculate the RMS error for the line: $\sqrt{\frac{1}{n}\sum (\text{predicted } y - \text{actual } y)^2}$.

5. **Regression**: (25 points) For people age 25 and over in the U.S. in 2005, the relationship between age and educational level (years of schooling completed) can be summarized as follows.

average age $\approx$ 50 years,	sd $\approx$ 16 years	
average educational level $\approx$ 13.2 years,	sd $\approx$ 3.0 years,	$r \approx -0.2$

a) Write down the equation for the regression line to predict educational level from age.

b) Predict the educational level of a 75 year old subject.

c) About 68% of the 75 year old subjects had educational levels in what range?

d) True or false: For people age 25 and older in the U.S. in 2005, older people tended to be less educated than younger ones. Explain.

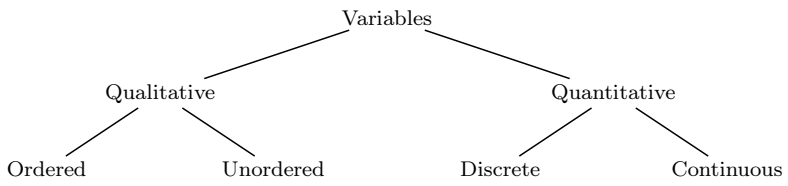
e) True or false, and explain: as you get older, you become less educated. If this statement is false, what could account for the negative correlation?

6. z-**Scores**: (10 points) Test scores on a certain exam average out to 75 with a standard deviation of 3. One student scored 84. Was that a very high score? Moderately high? Or about average? Justify your answer.

7. **NHANESIII Data Set**: (10 points)

a) In the exam data set, what information does the variable `PEP13E2A` give? What kind of variable is it?

b) In the lab data set, what information does the variable `HXPAXTMR` give? What kind of variable is it?



Math 207
Introduction to Statistics

Exam II

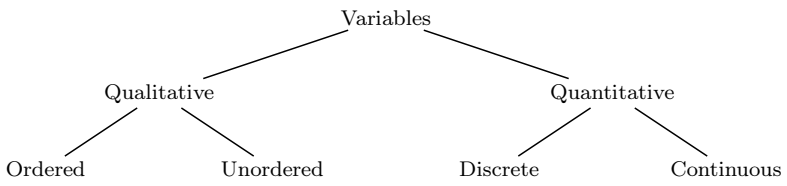
Please show your work and any R commands that you use.

1. z**-Scores**: (10 points) Test scores on a certain exam average out to 78 with a standard deviation of 2. One student scored 88. Was that a very high score? Moderately high? Or about average? Justify your answer.

2. **NHANESIII Data Set**: (10 points)

   a) In the exam data set, what information does the variable `PEP8A` give? What kind of variable is it?

   b) In the lab data set, what information does the variable `DMPCREGN` give? What kind of variable is it?



3. **Standard Normal Curve**: Sketch and compute the specified area under the normal curve. Write down the R command that you use. (15 points)

a) $z > -1$

a) $z < -1.2$

c) $-1.2 < z < -1$

d) What z is <u>less than</u> 30% of all z scores?

4. **Normal Curve**: Suppose that for a particular pet population subjects' weights average out to 8 pounds with an sd of 2 pounds and have a normal distribution. (15 points)

a) What percent of the weights were below 12 pounds?

b) What percent were above 10 pounds?

c) What percent were between 10 and 12 pounds?

d) What weight was <u>less than</u> 30% of all the weights?

5. **RMS Error for a Line**: (15 points) Consider the following data.

x	y
0	3
1	4
2	2

a) Make a scatter plot of the data.

b) An investigator uses the line $y = 5 - x$ to predict y from x. Add the line to your sketch.

c) Calculate the RMS error for the line: $\sqrt{\frac{1}{n}\sum (\text{predicted } y - \text{actual } y)^2}$.

6. **Correlation**: (10 points)

a) For a representative sample of cars, would the correlation between the age of the car and its gasoline economy (miles per gallon) be positive or negative? Explain

b) The correlation between gasoline economy and income of owner turns out to be positive? How do you account for this association?

7. **Regression**: (25 points) For people age 25 and over in the U.S. in 2005, the relationship between age and educational level (years of schooling completed) can be summarized as follows.

$$\text{average age} \approx 52 \text{ years}, \qquad \text{sd} \approx 10 \text{ years}$$

$$\text{average educational level} \approx 12.5 \text{ years}, \qquad \text{sd} \approx 2.5 \text{ years}, \qquad r \approx -0.3$$

a) Write down the equation for the regression line to predict educational level from age.

b) Predict the educational level of a 60 year old subject.

c) About 95% of the 60 year old subjects had educational levels in what range?

d) True or false: For people age 25 and older in the U.S. in 2005, older people tended to be less educated than younger ones. Explain.

e) True or false, and explain: as you get older, you become less educated. If this statement is false, what could account for the negative correlation?

Math 207
Introduction to Statistics

Exam II

Please show your work and any R commands that you use.

1. **RMS Error for a Line**: (15 points) Consider the following data.

x	y
0	1
1	3
2	6

a) Make a scatter plot of the data.

b) An investigator uses the line $y = 1 + x$ to predict y from x. Add the line to your sketch.

c) Calculate the RMS error for the line: $\sqrt{\frac{1}{n}\sum (\text{predicted } y - \text{actual } y)^2}$.

2. **Correlation**: (10 points)

a) For a representative sample of computers, would the correlation between the amount of memory and the age (in years) of the computer be positive or negative? Explain.

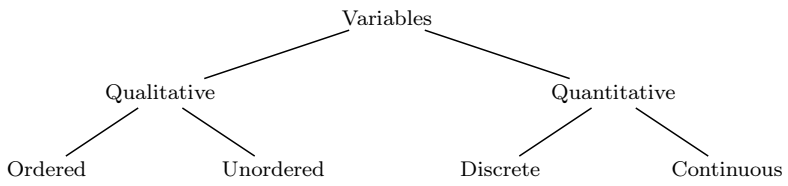
b) The correlation between amount of computer memory and the income of the owner turns out to be positive. How do you account for this association?

3. z**-Scores**: (10 points) Test scores on a certain exam average out to 70 with a standard deviation of 10. One student scored 85. Was that a very high score? Moderately high? Or about average? Justify your answer.

4. **NHANESIII Data Set**: (10 points)

a) In the exam data set, what information does the variable `PEP5R` give? What kind of variable is it?

b) In the lab data set, what information does the variable `PHPFAST` give? What kind of variable is it?



5. **Standard Normal Curve**: Sketch and compute the specified area under the normal curve. Write down the R command that you use. (15 points)

a) $z > 0.8$

a) $z < 2$

c) $0.8 < z < 2$

d) What z is greater than 40% of all z scores?

6. **Normal Curve**: Suppose that for a particular pet population subjects' ages average out to 8 years with an sd of 2 years and have a normal distribution. (15 points)

a) What percent of the pets were older than 3 years?

b) What percent were younger that 12 years?

c) What percent were between 3 and 12 years?

d) What age was greater than 40% of all the ages?

7. **Regression**: (25 points) For people age 25 and over in the U.S. in 2005, the relationship between age and educational level (years of schooling completed) can be summarized as follows.

$$\text{average age} \approx 60 \text{ years}, \qquad \text{sd} \approx 12 \text{ years}$$

$$\text{average educational level} \approx 11.5 \text{ years}, \qquad \text{sd} \approx 3.1 \text{ years}, \qquad r \approx -0.3$$

a) Write down the equation for the regression line to predict educational level from age.

b) Predict the educational level of a 55 year old subject.

c) About 95% of the 55 year old subjects had educational levels in what range?

d) True or false: For people age 25 and older in the U.S. in 2005, older people tended to be less educated than younger ones. Explain.

e) True or false, and explain: as you get older, you become less educated. If this statement is false, what could account for the negative correlation?