Shepherd UNIVERSITY

Math 314
Statistics

Exam II

No books or notes may be used on this exam.

1. (30 points) In a certain class, midterm scores average out to 70 with an SD of 15, as do scores on the final. The correlation between midterm scores and final scores is about 0.50.

a) Write down the equation for the regression line.

b) Sketch the regression line.

c) Estimate the average final exam scores for students whose midterm scores were 85, 40 and 70.

d) Estimate the midterm score for a student who scored 90 on the final exam.

2. (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?

3. (10 points) A box contains 5 red marbles and 4 green ones.

a) If three marbles are drawn *with replacement*, what is the chance that two of the three are red?

b) If three marbles are drawn *without replacement*, what is the chance that two of the three are red?

4. (10 points) Cards are drawn at random without replacement from a well-shuffled deck. Calculate the following probabilities.

a) First card is a ♣ or a ♠

b) Tenth card is a ♡ or a face card

c) Second card is an ace given that the first one drawn was a face card

5. (10 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.

$$\text{average age} \approx 50 \text{ years}, \qquad \text{SD} \approx 16 \text{ years}$$

$$\text{average education level} \approx 13.2 \text{ years}, \qquad \text{SD} \approx 3.0 \text{ years}, \qquad r \approx -0.2$$

a) 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.

b) 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. (15 points) Suppose that the regression line of y on x is

$$y = \left(\frac{\sqrt{3}}{2}\right)\left(\frac{8}{\sqrt{3}}\right)(x-1)$$

where $\text{mean}_x = 1$, $\text{mean}_y = 0$, $\text{SD}_x = \sqrt{3}$, $\text{SD}_y = 8$ and $r = \sqrt{3}/2$.

a) Calculate the RMS error for regression $\text{RMS} = \text{SD}_y\sqrt{1-r^2}$. What can you say about the RMS for the SD line?

b) If $x = 4$, then we expect 68% of the y-values to be between _____ and _____ .

(15 points)

7. 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: $\sqrt{\frac{1}{n}\sum (\text{predicted } y - \text{actual } y)^2}$.