Name: _____
Date:

Business Data Analysis
201-316-VA

In Class Exercise #7: Regression & Basic Probability

1. **Crime Rate**

Let x be a random variable representing percentage change in neighborhood population in the past few years, and let y be a random variable representing crime rate (crimes per 1000 population). A random sample of 6 Denver neighborhoods gave the following information

x	29	2	11	17	7	6
y	173	35	132	127	69	53

(a) Construct a table to compute $\Sigma x, \Sigma y, \Sigma x^2, \Sigma y^2, \Sigma xy$

x	y	x^2	y^2	xy
29	173	841	29,929	5017
⋮				
Σ 72	589	1340	72,277	9499

$n = 6$

(b) Find the equation of the least-squares line

slope: $b = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{6(9499) - 72 \cdot 589}{6 \cdot 1340 - (72)^2} = \frac{14586}{2856}$

≈ 5.1071

intercept: $a = \bar{y} - b\bar{x} = \frac{589}{6} - 5.1071 \cdot \frac{72}{6} \approx 36.8810$

$\hat{y} = a + bx \approx 36.8810 + 5.1071x$

(c) Find the sample correlation coefficient r. What does it tell us in this case?

$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}} = \frac{6(9499) - 72 \cdot 589}{\sqrt{6 \cdot 1340 - 72^2}\sqrt{6 \cdot 72,277 - 589^2}} \approx 0.927$

Strong positive correlation

(d) What percentage of variation in y is explained by the least-squares model?

$r^2 \approx 85.9\%$ of the variation is explained by the model

(e) For a neighborhood with a 12% increase in population in the past few years, predict the change in the crime rate.

if $x = 12$ $\quad \hat{y} \approx 36.8810 + 5.1071(12) \approx 98.17$

The crime rate will increase by 98.17 per 1000 pop

2. **Customer Purchases**

John runs a computer software store. Yesterday, 58 people entered the store, and 25 of them bought at least one item.

(a) Estimate the probability that a person who walks into the store today buys something.

$\frac{25}{58}$

(b) Estimate the probability that a person who walks into the store today does not buy anything.

$1 - \frac{25}{58} = \frac{33}{58}$

3. **Marbles**

A bag contains 12 red marbles, 5 blue marbles and 3 green marbles. Without looking into the bag, you draw one at random.

20 marbles total

(a) What is the probability that the marble you picked is red? Blue? Green?

$P(R) = \frac{12}{20}$

$P(B) = \frac{5}{20}$

$P(G) = \frac{3}{20}$

(b) You draw a marble from the bag, then place it back into the bag before drawing a new marble. Are the probabilities of drawing marbles of the different colors the same as in part a)? Explain.

Yes, since the marbles in the bag are the same.

(c) You draw a marble from the bag, then draw a new marble without replacing the first one into the bag. Are the probabilities of drawing marbles of the different colors the same as in part a)? Explain.

No, since the marbles in the bag are different.