

HOMEWORK: REGRESSION

SOLVE THE FOLLOWING PROBLEMS IN TWO WAYS:

- (a) USING CALCULATOR
- (b) USING MS EXCEL

Write out the regression equation. What is the correlation coefficient? What is r-square (coefficient of determination).

Problem 1: A researcher is interested in determining whether there is a relationship between number of packs of cigarettes smoked per day and longevity (in years). $n=10$.

| # packs of cigarettes smoked (X) | Longevity (Y) |
|-------------------------------------|------------------|
| 0 | 80 |
| 0 | 70 |
| 1 | 72 |
| 1 | 70 |
| 2 | 68 |
| 2 | 65 |
| 3 | 69 |
| 3 | 60 |
| 4 | 58 |
| 4 | 55 |

$$\sum X = 20; \sum Y = 667; \sum XY = 1247; \sum X^2 = 60; \sum Y^2 = 44,983$$

Problem 2: A researcher is interested in determining whether there is a relationship between advertising and sales for her firm. $n = 11$ sales areas.

| advertising in \$thousands(X) | Sales in millions(Y) |
|----------------------------------|-------------------------|
| 1 | 0 |
| 1 | 1 |
| 2 | 4 |
| 4 | 3 |
| 5 | 5 |
| 6 | 4 |
| 6 | 7 |
| 6 | 8 |
| 7 | 9 |
| 10 | 9 |
| 10 | 7 |

$$\sum X = 58; \sum Y = 57; \sum XY = 383; \sum X^2 = 404; \sum Y^2 = 391$$

Problem 3: A researcher is interested in determining whether there is a relationship between price and quantity demanded for her firm. $n = 11$.

| <u>Price(X)</u> | <u>Q-demanded(Y)</u> |
|-----------------|----------------------|
| 2 | 95 |
| 3 | 90 |
| 4 | 84 |
| 5 | 80 |
| 6 | 74 |
| 7 | 69 |
| 8 | 62 |
| 9 | 60 |
| 10 | 63 |
| 11 | 50 |
| 12 | 44 |

$$\sum X = 77; \quad \sum Y = 771; \quad \sum XY = 4,864; \quad \sum X^2 = 649; \quad \sum Y^2 = 56,667$$

Problem 4: A researcher is interested in determining whether there is a relationship between shelf space and number of books sold for her bookstore. $n = 11$.

| <u>Shelf Space in feet(X)</u> | <u>Books Sold(Y)</u> |
|-------------------------------|----------------------|
| 7.0 | 280 |
| 3.5 | 140 |
| 4.0 | 170 |
| 4.2 | 200 |
| 4.8 | 215 |
| 3.9 | 190 |
| 4.9 | 240 |
| 7.5 | 295 |
| 3.0 | 125 |
| 5.9 | 265 |
| 5.0 | 200 |

$$\sum X = 53.7; \quad \sum Y = 2,320; \quad \sum XY = 12,070; \quad \sum X^2 = 282.21; \quad \sum Y^2 = 519,700$$

Problem 5: A researcher is interested in determining whether there is a relationship between grades and hours studied for statistics.

| <u>Hours studied(X)</u> | <u>Grade on final(Y)</u> |
|-------------------------|--------------------------|
| 1 | 20 |
| 2 | 30 |
| 4 | 40 |
| 7 | 60 |
| 6 | 65 |
| 7 | 70 |
| 8 | 80 |
| 9 | 90 |
| 8 | 95 |
| 10 | 100 |

$$\Sigma X = 62; \Sigma Y = 650; \Sigma XY = 4,750; \Sigma X^2 = 464; \Sigma Y^2 = 49,150$$

Problem 6: A researcher is interested in determining whether there is a relationship between number of police officer in a district and number of crimes

| Number of Police Officers (X) | Number of Crimes (Y) |
|-------------------------------|----------------------|
| 4 | 49 |
| 6 | 42 |
| 8 | 38 |
| 9 | 31 |
| 10 | 24 |
| 12 | 24 |
| 12 | 28 |
| 13 | 23 |
| 15 | 21 |
| 20 | 19 |
| 26 | 12 |
| 28 | 14 |

$$\Sigma X = 163; \Sigma Y = 325; \Sigma XY = 3,593; \Sigma X^2 = 2,839; \Sigma Y^2 = 10,177$$

Problem 7: A researcher is interested in determining whether there is a relationship between education (in years) and net income (in thousands of dollars) studied for statistics.

| Education in Years | Income (in thousands) |
|--------------------|-----------------------|
| 9 | 20 |
| 10 | 22 |
| 11 | 24 |
| 11 | 23 |
| 12 | 30 |
| 14 | 35 |
| 14 | 30 |
| 16 | 29 |
| 17 | 50 |
| 19 | 45 |
| 20 | 43 |
| 20 | 70 |

$$\sum X = 173; \quad \sum Y = 421; \quad \sum XY = 6,616; \quad \sum X^2 = 2,665; \quad \sum Y^2 = 17,129$$

Problem 8: A researcher is interested in determining whether there is a relationship between high school average and job performance ((the higher the number, the better the performance) at a certain company.

| High school average (X) | Job performance (Y) |
|-------------------------|---------------------|
| 60 | 2 |
| 78 | 5 |
| 98 | 10 |
| 66 | 3 |
| 87 | 8 |
| 77 | 5 |
| 61 | 4 |
| 90 | 6 |
| 91 | 7 |
| 79 | 6 |
| 88 | 7 |
| 99 | 9 |

| | |
|----|---|
| 88 | 4 |
| 85 | 8 |
| 81 | 9 |

$$\sum X = 1,228; \sum Y = 93; \sum XY = 7,932; \sum X^2 = 102,560; \sum Y^2 = 655$$

(Problem 9)

A researcher wants to see whether there is a relationship between the number of colds people get in one year and the average amount of vitamin C they consume.

| Milligrams of Vitamin C (X) | Number of Colds (Y) |
|-----------------------------|---------------------|
| 830 | 0 |
| 900 | 0 |
| 900 | 1 |
| 170 | 1 |
| 230 | 1 |
| 50 | 2 |
| 420 | 2 |
| 280 | 2 |
| 200 | 3 |
| 200 | 4 |
| 80 | 5 |
| 50 | 9 |

$$\sum X = 4,310; \sum Y = 30; \sum XY = 5,050; \sum X^2 = 2,736,900; \sum Y^2 = 146$$

(Problem 10)

A researcher wants to see whether there is a relationship between the number of hours people exercise weekly and how long they live.

| Hours Exercised (X) | Longevity (Y) |
|---------------------|---------------|
| 0 | 70 |
| 0 | 68 |
| 0 | 75 |
| 2 | 66 |
| 2 | 76 |
| 3 | 72 |

| | |
|----|----|
| 4 | 69 |
| 4 | 73 |
| 6 | 72 |
| 6 | 74 |
| 8 | 72 |
| 8 | 77 |
| 10 | 73 |
| 10 | 77 |
| 12 | 76 |
| 17 | 78 |
| 20 | 81 |
| 24 | 82 |
| 30 | 86 |
| 32 | 89 |

$$\sum X = 198; \quad \sum Y = 1506; \quad \sum XY = 15,890; \quad \sum X^2 = 3,782; \quad \sum Y^2 = 114,048$$

ADDITIONAL TOPIC: TIME SERIES

This topic may not be covered. It is useful for those students interested in learning about trend lines.

Problem 10: A researcher is interested in drawing a trend line for the cost of a vital part using regression. Note: 1994 = 0; 1995 = 1; 1996 = 2; etc.

| Time (X) | Cost of a part (Y) |
|-----------|--------------------|
| 0 (1994) | 10 |
| 1 | 12 |
| 2 | 15 |
| 3 | 18 |
| 4 | 18 |
| 5 | 16 |
| 6 (2000) | 19 |
| 7 | 22 |
| 8 | 25 |
| 9 | 30 |
| 10 | 35 |
| 11 | 32 |
| 12 | 31 |
| 13 | 35 |
| 14 (2008) | 40 |

$$\sum X = 105; \quad \sum Y = 358; \quad \sum XY = 3,075; \quad \sum X^2 = 1,015; \quad \sum Y^2 = 9,778$$