Correlation and Regression

Assignment.

G&W Chapter 16, Chapter 17 Sections 17.1 & 17.2
Kachigan, Chapter 4, pps 160 - 180

Terms you should know.

Correlation

. . . .
. . . .
. . . .

Positive Correlation

. . . .
. . . .
. . . .

Negative Correlation

. . . .
. . . .
. . . .

Zero Correlation

. . . .
. . . .
. . . .

Scattergram

. . . .
. . . .
. . . .

Linear Relationship

. . . .
. . . .
. . . .

Curvilinear Relationship

. . . .
. . . .
. . . .

Homoscedasticity

. . . .
. . . .
. . . .

Heteroscedasticity

....
....
....

Regression

....
....
....

Coefficient of Determination

....
....
....

Coefficient of Alienation

....
....
....

Standard Error of Estimate

....
....
....

Least Squares Criterion

....
....
....

Formulas and Symbols You Should Know

r

....
....
....

$$r_{xy} = \frac{\Sigma(Z_x * Z_y)}{N}$$

....
....
....

Z'_y

....
....
....

$Z_y = r_{xy} * Z_x$

....
....
....

r_{xy}^{2}

. . . .
. . . .
. . . .

$1 - r_{xy}^{2}$

. . . .
. . . .
. . . .

$S_{Y'}$

. . . .
. . . .
. . . .

$$S_y * \sqrt{1 - r_{xy}^2}$$

. . . .
. . . .
. . . .

$$S_X * \sqrt{1 - r_{xy}^2}$$

. . . .
. . . .
. . . .

Computations You Should be Able to Perform.

1. The following data represent scores for the first annual wind-spitting contest conducted by Sal Liva, the former Spanish champion. Contestants teach take turn spitting into the wind. Both the wind speed and spitting distance for ten contestants are sown in this table.

Contestant	Wind Speed (mph)	Spit Distance (feet)	Z_{WS}	Z_{SD}	$Z_{T1} \cdot Z_{T2}$
1	5.00	4.00	-1.34	0.95	-1.27
2	10.00	4.50	-0.68	1.28	-0.87
3	13.00	2.50	-0.28	-0.05	0.01
4	5.00	2.75	-1.34	0.12	-0.16
5	18.00	3.50	0.38	0.62	0.24
6	20.00	2.00	0.65	-0.38	-0.25
7	20.00	0.50	0.65	-1.38	-0.90
8	25.00	0.00	1.31	-1.72	-2.25
9	25.00	2.00	1.31	-0.38	-0.50
10	10.00	4.00	-0.68	0.95	-0.64
Mean	15.10	2.58	0.00	0.00	
Std Dev	7.55	1.50	1.00	1.00	

a. Draw a scatter diagram of these scores and draw in the regression line.
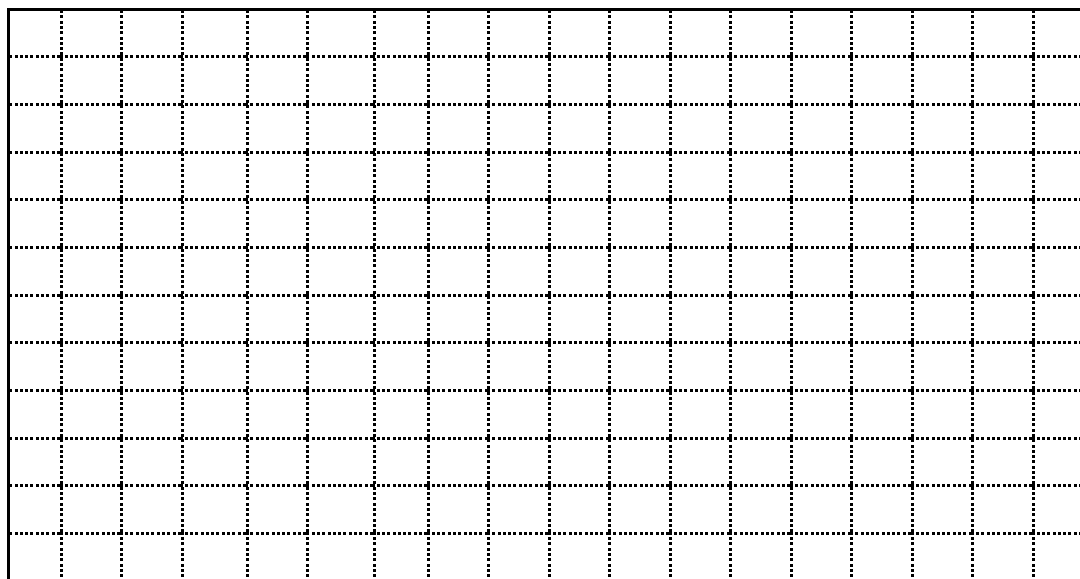
b. Compute the correlation coefficient using the z-score method.
2. Compute a predicted value of one score given another score using z-scores.

    a. What are the predicted Z values for Distance given the following z score values for Wind Speed?

        i.   1.00

        ii. -2.00

        iii. 1.40

        iv. -.98

3. What are the predicted raw score values in Distance given the following raw scores for Wind Speed?

    a. 10

    b. 4

4. What is the coefficient of determination for the above problem?

5. What is the coefficient of alienation for the above problem?

6. What is the standard error of estimate in predicting Distance for the above problem?

7. Compute the 90% confidence intervals for the following predictions.

    a. The predicted spitting distance is 5 feet.

    b. The predicted spitting distance is 1 foot..

Concepts and Interpretation

1. A school psychologist did some research and noted that there was a correlation of -.38 between the number of hours a student spent watching television and high school grade point average.

   a. In this study, what was the null hypothesis?

   b. How do you interpret this correlation?

   c. If you were making predictions of a student's grade point average based on knowledge of his/her TV habits, how confident would you be in the accuracy of these predictions? Why?

   d. Suppose that this researcher concluded that watching TV caused a student's grades to diminish. Would you agree with this? Why?

2. A researcher gathered data from a group of college students. The researcher coded their class as follows; 1 for freshman, 2 for sophomore, 3 for junior, and 4 for senior. In a later analysis the teacher found a correlation of +.33 between class and GPA. How would you interpret this?

3. A school psychologist conducted a study of the relationship between number of hours spent watching televison and grades. The psychologist divided a group of students into the top 10% and the bottom 10% on the basis of their grades. Then this researcher correlated their overall grades with the number of hours they spent watching television. The result was a correlation of -.20. How would you interpret this result?

4. Another researcher decided to redo the study of grades and hours spent watching television. This researcher rank ordered the students on the basis of their grades and then computed a Pearson correlation between grades and hours spent watching tv. What did this person do wrong?

SPSS Assignment #5

NAME: _____

1. Using SPSS, compute the correlations between the C-scores for the five AVA variables (V-1, V-2, V-3, V-4, and V-5).

   a. Which two variables have the highest correlation?

   b. Is this a significant relationship?

   c. What is the Coefficient of Determination for this relationship?

   d. Based on this Coefficient of Determination, would you say that this correlation has practical value?