

Correlation and Regression

Assignment: Heiman, Chapters 7 and 8

Terms you should know.

Correlation
.....
.....

Positive Correlation
.....
.....

Negative Correlation
.....
.....

Zero Correlation
.....
.....

Scatterplot
.....
.....

Linear Relationship
.....
.....

Curvilinear Relationship
.....
.....

Restriction of Range
.....
.....

Homoscedasticity
.....
.....

Heteroscedasticity
.....
.....

Linear Regression
.....
.....

Predictor Variable
.....
.....

Criterion variable
.....
.....

Coefficient of Determination
.....
.....

Coefficient of Alienation
.....
.....

Standard Error of Estimate
.....
.....

Least Squares Criterion
.....
.....

Formulas and Symbols You Should Know

r
.....
.....

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

$Z'_y = r_{xy} \times Z_x$
.....
.....

r_{xy}^2
.....
.....

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

S_Y
.....
.....

$S_{y'} = \sqrt{\sum \frac{(Y - Y')^2}{N}}$
.....
.....

$S_{y'} = S_y * \sqrt{1 - r_{xy}^2}$
.....
.....

Concepts You Should Master

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. How do you interpret this correlation?
 - b. 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?
 - c. 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 television 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?

Homework #14: Calculations You Should Master

Name: _____ (This is my work, and my work alone.)

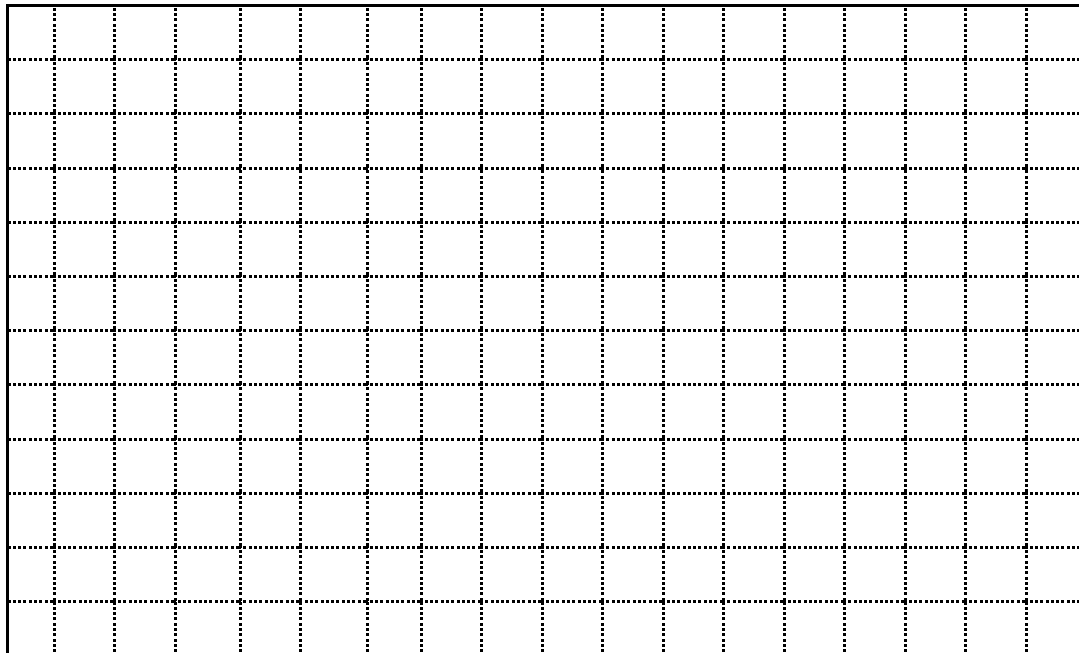
1. The following data represent pairs of test scores for ten students.

Test 1	Test 2	Z_{T1}	Z_{T2}	$Z_{T1} \cdot Z_{T2}$
30	72	-1.89	-1.18	
34	70	-0.96	-1.60	
35	76	-0.72	-0.32	
36	80	-0.49	0.53	
39	73	0.21	-0.96	
39	79	0.21	0.32	
40	76	0.44	-0.32	
40	83	0.44	1.18	
42	85	0.91	1.60	
46	81	1.85	0.75	

Test 1 Mean = 38.1 Standard deviation = 4.28

Test 2 Mean = 77.5 Standard deviation = 4.67

Draw a scatter diagram of these test scores using the chart below.



2. Compute the correlation between the two test scores in item 1 using the z-score method.
3. Drawn the regression line in the figure in item 1.
4. Determine the critical value of r .
 - a. What is df ?
 - b. At $p < .05$
 - c. At $p < .01$
5. Compute a predicted value of one score given another score using z-scores. What are the predicted Z values for Test 1 given the following z score values for Test 2?
 - a. 1.00
 - b. -2.00
 - c. -.98
6. What are the predicted values for Test 2 given the following z-score values for Test 1?
 - a. 1.00
 - b. -2.00
 - c. -3.24
7. What are the predicted raw score values in test 2 given the following raw scores for test 1?
 - a. 30
 - b. 46

8. What are the predicated raw score values for test 1 given the following values for test 2?
- a. 70
 - b. 87
9. Compute the coefficient of determination, coefficient of alienation, and standard error of estimate.
- a. What is the coefficient of determination for the above problem?
 - b. What is the coefficient of alienation for the above problem?
 - c. What is the standard error of estimate in predicting Test 2 for the above problem?
 - d. What is the standard error of estimate for predicting Test 1 for the above problem?
10. Compute the 90% confidence intervals for the following predictions.
- a. The predicted value of a score on Test 2 is 80.
 - b. The predicted value of a score on Test 1 is 34.

11. Using SPSS calculate the correlation between the following variables in your data set. Use the C-scores: Assertiveness, Sociability, Calmness, Conformity, Restraint.
- a. Which variables had the strongest correlation?

 - b. What is the correlation between
 - i. Assertiveness and Sociability?

 - ii. Is this correlation statistically significant?

 - iii. Is this correlation practically significant?