Extra Chapter 3 Test Review

AP Stat

Paulson

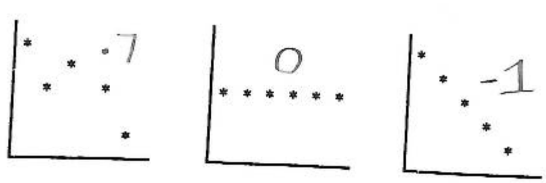
5. Which of the following statements about the correlation coefficient are true?
   I. The correlation coefficient and the slope of the regression line may have opposite signs. X
   II. A correlation of 1 indicates a perfect cause-and-effect relationship between the variables. X
   III. Correlations of +.87 and −.87 indicate the same degree of clustering around the regression line. ✓

   (A) I only
   (B) II only
   (C) III only
   (D) I and II
   (E) I, II, and III

25. Which of the following statements about residuals are true?
   I. The mean of the residuals is always zero. ✓
   II. Influential scores have large residuals. X
   III. A definite pattern in the residual plot is an indication that a nonlinear model should be tried. ✓

   (A) II only
   (B) I and II
   (C) I and III
   (D) II and III
   (E) I, II, and III

27. Consider the following three scatterplots:



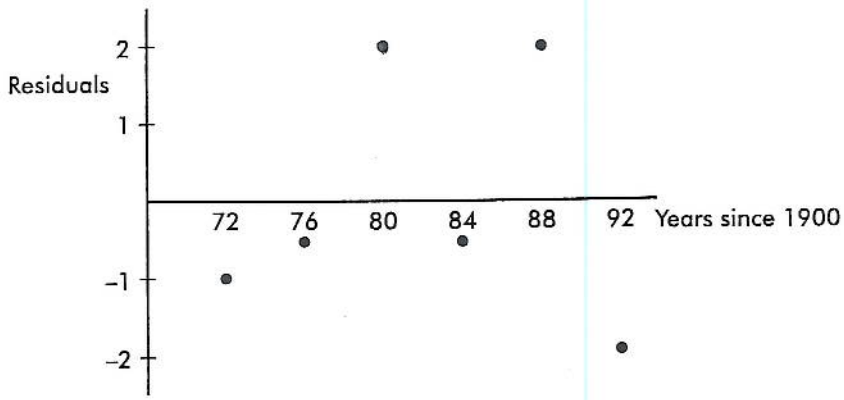
What is the relationship among r_1, r_2, and r_3, the correlations associated with the first, second, and third scatterplots, respectively?

(A) $r_1 < r_2 < r_3$
(B) $r_1 < r_3 < r_2$
(C) $r_2 < r_3 < r_1$
(D) $r_3 < r_1 < r_2$
(E) $r_3 < r_2 < r_1$

1. Data show a trend in winning long jump distances for an international competition over the years 1972–92. With jumps recorded in inches and dates in years since 1900, a least squares regression line is fit to the data. The computer output and a graph of the residuals are as follows:

R squared = 92.1%

Variable	Coefficient	SE of Coeff	t-ratio	Prob
Constant	256.576	11.59	22.1	0.0001
Year	0.95893	0.141	6.81	0.0024



(a) Does a line appear to be an appropriate model? Explain.
(b) What is the slope of the least squares line? Give an interpretation of the slope.
(c) What is the correlation?
(d) What is the predicted winning distance for the 1980 competition?
(e) What was the actual winning distance in 1980?

(a) Yes. The residual graph is not curved and appears to be random or scattered.

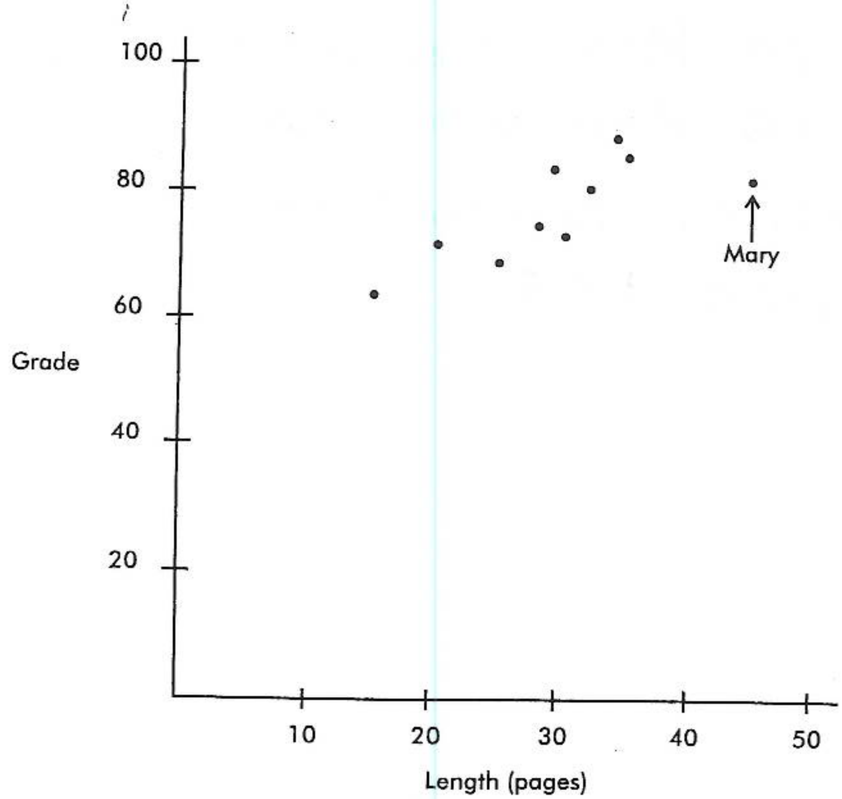
(b) The slope is 0.95893. The winning jump on average improves 0.95893 inches per year.

(c) $r^2 = .921$ $r = .96$

(d) $\hat{y} = 256.576 + .95893(80) = 333.3$ inches

(e) Residual for 1980 was 2 so $333.3 + 2 = 335.3$ inches

2. The following scatterplot shows the grades for research papers for a sociology professor's class plotted against the lengths of the papers (in pages).



Mary turned in her paper late and was told by the professor that her grade would have been higher if she had turned it in on time. A computer printout fitting a straight line to the data (not including Mary's score) by the method of least squares gives

```
Grade = 46.51 + 1.106 Length
R-sq = 74.6%
```

(a) Find the correlation coefficient for the relationship between grade and length of paper based on these data (excluding Mary's paper).

(b) What is the slope of the regression line and what does it signify?

(c) How will the correlation coefficient change if Mary's paper is included? Explain your answer.

(d) How will the slope of the regression line change if Mary's paper is included? Explain your answer.

(e) What grade did Mary receive? Predict what she would have received if her paper had been on time.

(a) $r^2 = .746$ $r = .864$

(b) The slope is 1.106. For each additional page, the average grade raises by 1.106.

(c) Including Mary's paper will lower the correlation coefficient because her result seem far off the regression line through the other points.

→

(d) Including Mary's paper will swing the regression line down and lower the value of the slope.

(e) From the graph, Mary received an 82. From the regression line, Mary would have received $\hat{y} = 46.51 + 1.106(45) = 96.3$ if she had turned in her paper on time.