Quick Questions 24 Simple Linear Regression Analysis

I. Place the number of the appropriate formula, symbol, or expression next to the concept it describes.

A. The standard error of the estimate __7__
B. The y-intercept __3__
C. The regression equation __1__
D. The estimated value of y given x __4__
E. The slope __5__
F. An interval estimate for the conditional mean of Y __2__
G. An interval estimate for an individual value of Y __6__

1.	$\hat{y}._x = a + bx$
2.	$\hat{y}._x \pm ts_{y.x}\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$
3.	$\bar{Y} - b\bar{x}$
4.	$\hat{y}._x$
5.	$\frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$
6.	$\hat{y}._x \pm ts_{y.x}\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$
7.	$\sqrt{\frac{\sum Y^2 - a(\sum Y) - b(\sum XY)}{n-2}}$

II. The following data was first presented in chapter 23. Estimate the regression line for this scatter using the eyeball method.

See page QQ 150

III. Calculate the regression equation to 3 significant digits.

Data from page QQ 150

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} = \frac{69.6}{247} = .2817813$$

$$a = \bar{Y} - b\bar{x} = \frac{\sum Y}{n} - b\frac{\sum X}{n} = \frac{23.2}{8} - (.2817813)(\tfrac{31}{8}) = 1.8080975$$

$$\hat{y}._x = a + bx$$
$$\hat{y}._x = 1.81 + .282x$$

IV. Estimate the grade point average for people who studied 5 hours per weekend.

$$y._5 = 1.81 + .282x = 1.81 + .282(5) = 1.81 + 1.41 = 3.22$$

V. Draw the regression line on the page 156 scatter diagram.

Easy points to determine are the y-intercept (0,1.81) and question IV coordinates (5, 3.22).

VI. Calculate the 98% confidence interval for students who study 5 hours per weekend.

$$S_{y.x} = \sqrt{\frac{\sum Y^2 - a(\sum Y) - b(\sum XY)}{n-2}} = \sqrt{\frac{70.38 - 1.8080975(23.2) - (.2817813)(98.6)}{8-2}} = .329$$

$$df = 8 - 2 = 6$$
$$\alpha/2 = .02/2 = .01 \rightarrow 3.143 \text{ for } t$$

$$\bar{x} = \frac{\sum x}{n} = \frac{31}{8} = 3.875$$

$$\hat{y}._x \pm ts_{y.x}\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$$

$$\hat{y}._5 = 3.22 \pm 3.143(.329)\sqrt{\frac{1}{8} + \frac{(5-3.875)^2}{151 - \frac{(31)^2}{8}}}$$

$$= 3.22 \pm 3.143(.329)(.407421)$$

$$= 3.22 \pm .421$$

$2.80 \leftrightarrow 3.64$

VII. What procedure should be followed if the range by your answer to question E includes negative numbers?

A negative number is not possible. If the range expresses the possibility of a negative number, the confidence interval may be lowered with a larger sample. Even if someone studied only 2 hours, the lower limit of the data, y is positive. Why? The standard error is low and $\hat{y}._2$ could not be zero.