Set B #8

(2007 #4)

AP® STATISTICS
2007 SCORING GUIDELINES (Form B)

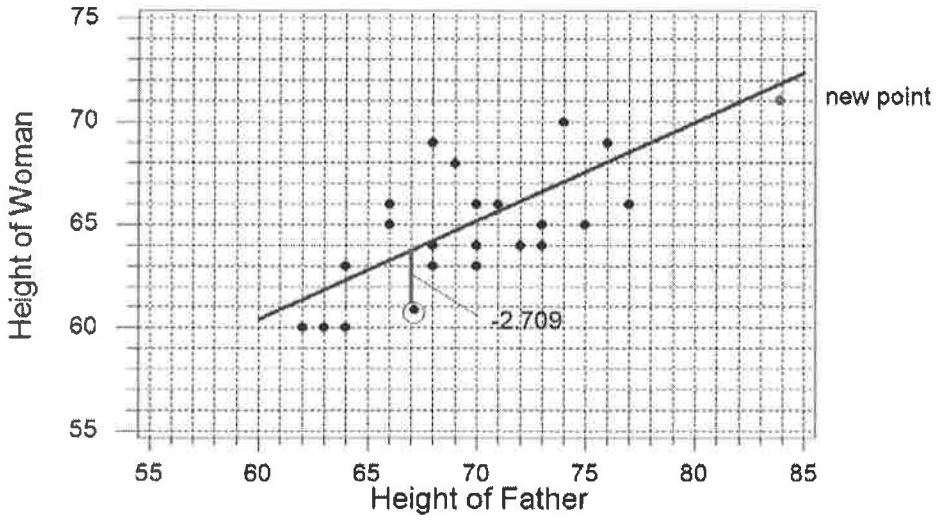
Question 4

Intent of Question

The goals of this question are to assess a student's ability to: (1) plot a least squares regression line; (2) examine a residual; and (3) discuss the effect of an additional observation on an estimated correlation coefficient and on the least squares estimate of the slope of a line.

Solution

Parts (a) and (b):



When $x = 67$, $\hat{y} = 35.1 + 0.427(67) = 63.709$

and the residual $= y - \hat{y} = 61 - 63.709 = -2.709$.

Part (c):

See the new point indicated in the plot above. The slope would remain about the same since the new point is consistent with the linear pattern in the original plot (i.e., close to the line).

The correlation coefficient would increase. We know that $b = r\frac{s_y}{s_x}$. The added point will increase s_x more than it will increase s_y so $\frac{s_y}{s_x}$ will be less than 1. If the slope is to stay the same, r must increase.

OR

This point fits the pattern well and has an x value that is far from $\bar{x}$.

AP® STATISTICS
2007 SCORING GUIDELINES (Form B)

Question 4 (continued)

Scoring

This problem is scored in 4 sections. Section 1 consists of the graphical parts of (a) and (b) together. Section 2 consists of the numerical parts of (b). Section 3 consists of the first part of (c). Section 4 consists of the second part of (c).

Each section is scored as either essentially correct (E), partially correct (P), or incorrect (I).

Section 1 (graphical parts of a and b) is essentially correct (E) if:

1. the regression line is drawn correctly on the scatterplot;
2. the point $(67, 61)$ is circled and the vertical segment corresponding to the residual is drawn on the scatterplot.

Section 1 is partially correct (P) if the response includes one of the above two elements.

Section 2 (numerical part of b) is essentially correct (E) if the residual is correctly computed as –2.709;
OR
the response states that the residual was approximated using the graph, a reasonable value for the residual is given, and the sign of the residual is correct.

Section 2 is partially correct (P) if the magnitude of the residual is correct but the sign is wrong.

Section 3 (first part of (c)) is essentially correct (E) if it:

1. states that the slope will remain about the same (or change slightly);
2. provides an explanation based on the new point fitting the pattern in the original plot.

Section 3 is partially correct (P) if it states that the slope will be about the same, but the explanation is missing or incorrect.

NOTE: If the line is drawn incorrectly in part (a), and the answer to this part is consistent with the line drawn, section 3 is essentially correct (E).

Section 4 (second part of (c)) is essentially correct (E) if it:

1. states that the value of the correlation coefficient will increase;
2. provides an explanation based on the relative changes in s_x and s_y
   OR
   based on the fact that the new point fits the pattern AND is far out in the x direction,
   OR
   because the linear pattern is stronger.

Section 4 is partially correct (P) if it states that the value of the correlation coefficient will increase, but the explanation is missing or incorrect.

NOTE: If the response just says that the correlation coefficient will increase because the point is close to the line, section 4 is partially correct.

AP® STATISTICS
2007 SCORING GUIDELINES (Form B)

Question 4 (continued)

4 **Complete Response**

All four sections essentially correct

3 **Substantial Response**

Three sections essentially correct and no sections partially correct

OR

Two sections essentially correct and two sections partially correct

2 **Developing Response**

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and two sections partially correct

OR

Four parts partially correct

1 **Minimal Response**

One section essentially correct and no sections partially correct

OR

No sections essentially correct and two sections partially correct

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.