2017

AP CollegeBoard

AP Statistics

Sample Student Responses and Scoring Commentary

Inside:

- ☑ Free Response Question 1
- ☑ Scoring Guideline
- ☑ Student Samples
- ☑ Scoring Commentary

© 2017 The College Board. College Board, Advanced Placement Program, AP, AP Central, and the acorn logo are registered trademarks of the College Board. Visit the College Board on the Web: www.collegeboard.org.

AP Central is the official online home for the AP Program: apcentral.collegeboard.org

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) explain statistical terms used when describing the relationship between two variables; (2) interpret the slope of a linear regression equation; and (3) calculate a value of y when given a regression equation, a value of x, and a residual.

Solution

Part (a):

In the context of a scatterplot in which y represents weight and x represents length, the following are defined.

A positive relationship means that wolves with higher values of length also tend to have higher weights.

A linear relationship means that as length increases by one meter, weight tends to change by a constant amount, on average.

A strong relationship means that the data points fall close to a line (or curve).

Part (b):

The slope of 35.02 indicates that two wolves that differ by one meter in length are predicted to differ by 35.02 kilograms in weight, with the longer wolf having the greater weight.

Part (c):

In general, a residual is equal to actual weight minus predicted weight, or equivalently,

$$\text{actual weight} = \text{predicted weight} + \text{residual}.$$

For the wolf with length 1.4 meters and residual of -9.67, the predicted weight is

$$-16.46 + 35.02(1.4) = 32.568 \text{ kilograms}.$$

Therefore, the actual weight of the wolf is $32.568 + (-9.67) = 22.898$ kilograms.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. A reasonable definition of positive
2. A reasonable definition of linear
3. A reasonable definition of strong
4. At least one definition in context

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

Question 1 (continued)

Partially correct (P) if the response includes only three of the four components.

Incorrect (I) if the response does not meet the conditions for E or P.

Notes:

- The description of a positive relationship should clearly indicate that relatively low values of one variable tend to appear with relatively low values of the other variable, and relatively high values of the first variable tend to appear with relatively high values of the other variable.
  - Examples of acceptable responses:
    - As length increases, so does weight.
    - Longer wolves weigh more.
    - The points on the graph go up as you move from left to right.
  - Examples of unacceptable responses:
    - As length goes up, weight changes.
    - Both length and weight get bigger.
    - The correlation is greater than 0.
- The description of a linear relationship can take one of two approaches: the data pattern (data points exhibit the pattern of a line in the graph) or the constant rate of change (as the explanatory variable changes, the response variable exhibits a constant rate of change).
  - Examples of acceptable responses:
    - The points generally follow a straight line.
    - The relationship between x and y is straight.
    - Length and weight have a constant slope.
  - Examples of unacceptable responses:
    - The points all line up.
    - You can draw a straight line through the points.
    - There is a positive correlation.
    - Every increase in x yields a 35.02 increase in y.
- The description of strong should indicate how close points are to a line.
  - Examples of acceptable responses:
    - Observed values are close to predicted values.
    - Deviations from the least-squares regression line are small.
    - The correlation coefficient is close to 1.
  - Examples of unacceptable responses:
    - All the points are close together.
    - The scatterplots are clustered together.
    - There is a high positive correlation.
- Context can be shown by referring to length and weight or by using meters and kilograms.
- Sketches and graphs can be used to help clarify definitions, but a sketch alone cannot satisfy a definition component.

Part (b) is scored as follows:

Essentially correct (E) if the response includes the following three components:

1. The correct value of 35.02 for the slope.
2. An interpretation that includes an increase of a specified amount of weight for each unit increase in length.
3. An indication that the relationship is not exact by using words such as “on average” or “predicted weight.”

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

Question 1 (continued)

Partially correct (P) if the response includes only two of the three components.

Note: If the response identifies the slope as -16.46 (the intercept value), the second component is satisfied only if the response states that for each one-meter increase in length there is a *decrease* in predicted or average weight of 16.46 kilograms.

Incorrect (I) if the response does not meet the criteria for E or P.

Part (c) is scored as follows:

Essentially correct (E) if the response includes the following two components:

1. A correct computation for the predicted value 32.568 kilograms.
2. A correct computation for the actual weight 22.9 kilograms using the given residual and the predicted value.

Partially correct (P) if the response provides a correct computation for the predicted value but is not able to complete the correct calculation of the actual weight, including if the residual is defined in the wrong direction as (predicted weight) – (actual weight) to give an answer of 42.24 kilograms;

OR

if the response provides an incorrect value for the predicted weight, but then uses that value correctly to determine the actual weight as (predicted weight) + residual = (predicted weight) $+(-9.67)$;

OR

if the response provides a correct answer for the actual weight but does not give sufficient information to determine how it was calculated.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- The expression $-16.46 \quad 35.02(1.4)$ is enough to satisfy the first component.
- The equation $-16.46 + 35.02(1.4) - 9.67 = 22.9$ satisfies both components.
- Arithmetic mistakes are overlooked if they do not lead to an unreasonable answer (such as a negative value). For example, $32.568 + (-9.67) = 21.9$ satisfies the second component.

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

Question 1 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

STATISTICS
SECTION II
Part A
Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II score—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. Researchers studying a pack of gray wolves in North America collected data on the length x, in meters, from nose to tip of tail, and the weight y, in kilograms, of the wolves. A scatterplot of weight versus length revealed a relationship between the two variables described as positive, linear, and strong.

   (a) For the situation described above, explain what is meant by each of the following words.

   (i) Positive:

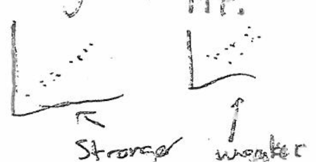
   In general, as the length of the gray wolves increases the general trend of the data suggests that weight will increase as well. A positive association means that the data is in general going upwords from left to right when graphing length, x, vs weight, y.

   (ii) Linear:

   Linear means that the data appears to be clustered around a line, and the data is not better followed by another curve. the graph between length and weight is best described by a line.

   (iii) Strong:

   This means that the data is well modelled by a line The residuals are generally fairly small, and the line is a good fit.



Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.

1A 2

The data collected from the wolves were used to create the least-squares equation $\hat{y} = -16.46 + 35.02x$.

(b) Interpret the meaning of the slope of the least-squares regression line in context.

As the length of a gray wolf increases by one meter, the least squares regression line predicts that the weight in kilograms of a wolf would increase by approximately 35.02.

(c) One wolf in the pack with a length of 1.4 meters had a residual of −9.67 kilograms. What was the weight of the wolf?

$y - \hat{y} = -9.67$ from the residual

$x = 1.4$m, the actual length of the wolf.

$\hat{y} = -16.46 + 35.02 \cdot 1.4 = 32.57$ kg

$y = \hat{y} - 9.67 = 32.57 - 9.67 = \boxed{22.90 \text{ kg}}$

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

© 2017 The College Board.
Visit the College Board on the Web: www.collegeboard.org.