

**AP[®] STATISTICS
2006 SCORING GUIDELINES**

Question 2

Intent of Question

The primary goal of this question is to assess a student's ability to identify the estimated regression line and to identify and interpret important statistics from regression output provided by statistical software in the context of a practical problem.

Solution

Part (a):

The regression line is $\hat{y} = -2.679 + 9.5x$, where \hat{y} represents the estimated (or predicted) mean height of the soapsuds and x represents the amount of detergent added to the pan.

Part (b):

The value $s = 1.99821$ mm is the standard deviation of the residuals. This statistic measures a typical amount of variability in the vertical distances from the observed height of the soapsuds to the regression line.

OR

The value $s = 1.99821$ mm is a measure of variation in the height of soapsuds for a given amount of detergent.

Part (c):

The standard error of the estimated slope parameter is 0.7553 mm per gram. Thus, the standard deviation of the estimated slope for predicting the height of soapsuds by using an amount of detergent is estimated to be 0.7553 mm per gram. This value estimates the variability in the sampling distribution of the estimated slope (i.e., how much we would expect sample slopes to vary from experiment to experiment).

Scoring

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

Part (a) is essentially correct (E) if the least squares regression line is correctly identified and the variables are correctly defined.

Part (a) is partially correct (P) if:

the least squares regression line is correctly identified and either of the two variables are not correctly defined;

OR

the least squares regression line is not presented using estimated or predicted notation, or \hat{y} , AND both variables are correctly defined;

OR

only one of the two values is correctly identified from the table and both variables are correctly defined.

Part (a) is incorrect (I) if the least squares regression line is incorrectly identified or not identified, and the variables are not correctly defined.

AP[®] STATISTICS
2006 SCORING GUIDELINES

Question 2 (continued)

Notes:

- If y is identified as the height of the soapsuds and x is identified as the amount of detergent, then the student should get credit for defining the variables. However, y must be identified as an estimated height somewhere in the student response in order to get this part essentially correct.
- If \hat{y} (or *estimated height*) is used to specify the regression line and y is identified as the height, the response should be scored as essentially correct. For example, a response of estimated height = $-2.679 + 9.5\text{amount}$, where y represents the height of suds and x represents the amount of detergent should be scored essentially correct.
- If the regression line is specified using y (or height) instead of \hat{y} (or estimated height), the response is scored as partially correct. For example, a response of $y = -2.679 + 9.5x$ where $y = \text{height of suds}$ and $x = \text{amount of detergent}$ should be scored as partially correct.
- If the estimates of the intercept and slope are reversed ($\hat{y} = 9.5 - 2.679x$), then the response should be scored as incorrect.

Part (b) is essentially correct (E) if the standard deviation is correctly interpreted in the context of this study.

Part (b) is partially correct (P) if the standard deviation is correctly interpreted in general terms without context.

Part (b) is incorrect (I) if the response indicates that s is any other standard deviation (e.g., univariate distribution of y), or slope.

Notes:

- If s is interpreted as the estimated standard deviation of the differences between the observed values for the height of soapsuds and the values predicted from the regression line, the response should be scored essentially correct.
- If s is interpreted as a “typical” prediction error for estimating height from the amount of detergent, then the response should be scored essentially correct.

Part (c) is essentially correct (E) if the standard error is identified and interpreted correctly.

Part (c) is partially correct (P) if standard error is identified but interpretation is weak (e.g., the standard error is a standard deviation of the slope). The major idea of sampling variability is not included.

Part (c) is incorrect (I) if the standard error is not correctly identified, identified with no interpretation, or an incorrect interpretation is provided.

**AP[®] STATISTICS
2006 SCORING GUIDELINES**

Question 2 (continued)

4 Complete Response

All 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 two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct and either zero or one part partially correct

OR

No parts essentially correct and two parts partially correct