AP® STATISTICS
2005 SCORING GUIDELINES

Question 3

Solution

Part (a):

Yes, the linear model is appropriate for these data. The scatterplot shows a strong, positive, linear association between the number of railcars and fuel consumption, and the residual plot shows a reasonably random scatter of points above and below zero.

Part (b):

According to the regression output, fuel consumption will increase by 2.15 units for each additional railcar. Since the fuel consumption cost is \$25 per unit, the average cost of fuel per mile will increase by approximately $\$25 \times 2.15 = \$53.75$ for each railcar that is added to the train.

Part (c):

The regression output indicates that $r^2 = 96.7\%$ or 0.967. Thus, 96.7% of the variation in the fuel consumption values is explained by using the linear regression model with number of railcars as the explanatory variable.

Part (d):

No, the data set does not contain any information about fuel consumption for any trains with more than 50 cars. Using the regression model to predict the fuel consumption for a train with 65 railcars, known as extrapolation, is not reasonable.

Scoring

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

Part (a) is essentially correct (E) if the model is deemed appropriate AND the explanation clearly indicates:
- There is a linear pattern in the scatterplot; OR
- There is no pattern in the residual plot.

Part (a) is partially correct (P) if the:
- Model is deemed appropriate AND the student refers to the scatterplot or residual plot but fails to state the relevant characteristic of the plot; OR
- Student refers to the relevant characteristic of the scatterplot or residual plot without deeming model appropriate.

Part (a) is incorrect (I) if the student:
- States that the model is appropriate without an explanation; OR
- States that the model is inappropriate; OR
- Makes a decision based only on numeric values from the computer output.

Question 3 (continued)

Part (b) is essentially correct (E) if the point estimate for the slope (2.15 or 2.1495) and the fuel consumption cost per unit ($25) are used to calculate the correct point estimate ($53.75 or $53.7375 ≈ $53.74).

Part (b) is partially correct (P) if only the point estimate for the slope (2.15 or 2.1495) is stated with a supporting calculation or interpretation.

Part (c) is essentially correct (E) if the student states:
- 96.7% of the variation in fuel consumption is explained by the linear regression model; OR
- 96.7% of the variation in fuel consumption is explained by the number of railcars.

Part (c) is partially correct (P) if the student makes one of the above statements using R-Sq(adj) = 96.3%.

Part (d) is essentially correct (E) if the student states that this is unreasonable due to extrapolation.

Part (d) is partially correct (P) if the student states this is:
- Unreasonable but provides a weak explanation; OR
- Reasonable even though it is considered a slight extrapolation.

Note: Any answer appearing without supporting work is scored as incorrect (I).

Each essentially correct (E) response counts as 1 point, each partially correct (P) response counts as ½ point.

4 Complete Response

3 Substantial Response

2 Developing Response

1 Minimal Response

Note: If a response is in between two scores (for example, 2 ½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.