

AP[®] STATISTICS
2011 SCORING GUIDELINES (Form B)

Question 6

Intent of Question

The primary goals of this question were to assess students' ability to (1) interpret the slope of a regression line in context; (2) decide whether or not a model should be used for prediction; (3) describe the sampling distribution of a sample mean; (4) use the sampling distribution of a sample mean to obtain an interval of plausible values; (5) compare two different study plans and decide which one would provide a better estimator of the slope; (6) propose a different study plan to check an assumption.

Solution

Part (a):

For each additional foot that is added to the width of the grass buffer strip, an additional 3.6 parts per hundred of nitrogen is removed on average from the runoff water.

Part (b):

No. This is extrapolation beyond the range of data from the experiment. Buffer strips narrower than 5 feet or wider than 15 feet were not investigated.

Part (c):

Because the distribution of nitrogen removed for any particular buffer strip width is normally distributed with a standard deviation of 5 parts per hundred, the sampling distribution of the mean of four observations when the buffer strips are 6 feet wide will be normal with mean $33.8 + 3.6 \square 6 = 55.4$ parts per hundred and a standard deviation of $\frac{s}{\sqrt{n}} = \frac{5}{\sqrt{4}} = 2.5$ parts per hundred.

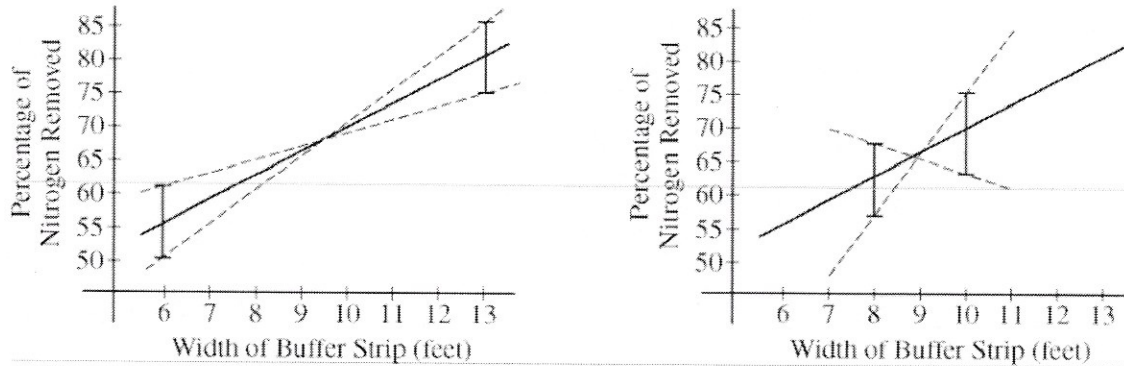
Part (d):

The distribution of the sample mean is normal, so the interval that has probability 0.95 of containing the mean nitrogen content removed from four buffer strips of width 6 feet extends from $55.4 - 1.96 \square 2.5 = 50.5$ parts per hundred to $55.4 + 1.96 \square 2.5 = 60.3$ parts per hundred.

AP[®] STATISTICS
2011 SCORING GUIDELINES (Form B)

Question 6 (continued)

Part (e):



If we think that the sample mean nitrogen removed at a particular buffer width might reasonably be any value in the intervals shown, a sample regression line will result from connecting any point in the interval above 6 to any point in the interval above 13. With this in mind, the dashed lines in the plots above represent extreme cases for possible sample regression lines. From these plots, we can see that there is a wider range of possible slopes in the second plot (on the right) than in the first plot (on the left). Because of this, the variability in the sampling distribution of b , the estimator for the slope of the regression line, will be smaller for the first study plan (with four observations at 6 feet and four observations at 13 feet) than it would be for the second study plan (with four observations at 8 feet and four observations at 10 feet). Therefore, the first study plan (on the left) would provide a better estimator of the slope of the regression line than the second study plan (on the right).

Part (f):

To assess the linear relationship between width of the buffer strip and the amount of nitrogen removed from runoff water, more widths should be used. To detect a nonlinear relationship it would be best to use buffer widths that were spaced out over the entire range of interest. For example, if the range of interest is 6 to 13 feet, eight buffers with widths 6, 7, 8, 9, 10, 11, 12 and 13 feet could be used.

Scoring

This question is scored in four sections. Section 1 consists of parts (a) and (b); section 2 consists of parts (c) and (d); section 3 consists of part (e); section 4 consists of part (f). Each of the four sections is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

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

1. The response in part (a) is correct, as evidenced by the correct interpretation of the slope, in context.
2. The response in part (b) is correct, as evidenced by the identification of extrapolation as the reason that the model should not be used AND the response is in context.

Partially correct (P) if only one of the two components listed above is correct.

AP[®] STATISTICS
2011 SCORING GUIDELINES (Form B)

Question 6 (continued)

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes

- Part (a) is incorrect if the interpretation is not in context or if the interpretation does not acknowledge uncertainty (for example, does not include “on average” or “about” or “approximately” or “predicted” when referring to the increase in nitrogen removed).
- Ideally a correct solution would also include units and make it clear that it is the approximate increase for each additional foot added to the buffer, but in the context of this larger investigative task, failure to do so is not sufficient to make part (a) incorrect.
- Part (b) is incorrect if extrapolation is not identified as the reason or if the response is not in context.

Section 2 is scored as follows:

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

1. The response in part (c) states that the sampling distribution is normal AND provides a correct mean and standard deviation.
2. The response in part (d) uses the correct mean and standard deviation of the sampling distribution or incorrect values carried over from part (c) AND a correct critical value (1.96 or 2) to compute a correct interval.

Partially correct (P) if only one of two components listed above is correct.

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes

- Stating that the sampling distribution is approximately normal is acceptable for part (c).
- Part (c) is incorrect if the response does not state that the distribution is normal or if an incorrect mean or standard deviation is given.
- Part (d) is incorrect if an incorrect critical value (for example a t critical value) is used or if an incorrect mean or standard deviation other than incorrect values from part (c) is used in the computation of the interval.

Section 3 is scored as follows:

Essentially correct (E) if study plan 1 is chosen in part (e), and the response demonstrates awareness of sampling variation in the estimates of the slopes of the regression lines, AND this is clearly communicated in the context of the two study plans.

Partially correct (P) if study plan 1 is chosen in part (e), and the response demonstrates awareness of sampling variation in the estimates of the slopes of the regression lines, BUT the justification of the choice of study plan 1 is not clearly communicated.

Incorrect (I) if the response fails to meet the criteria for E or P.

AP[®] STATISTICS
2011 SCORING GUIDELINES (Form B)

Question 6 (continued)

Section 4 is scored as follows:

Essentially correct (E) if the response specifies another study plan that uses eight buffer strips of at least three different specified widths, and the response indicates in how many locations each width will be used, OR the response makes it clear that at least three different buffer widths will be used and indicates that the buffer widths to be used will be spread out over the range of interest.

Note: Specifying eight different widths is sufficient for an E in section 4.

Partially correct (P) if the response does not meet criteria for E, BUT the stated plan uses at least three different widths. Widths need not be specified for a P.

Incorrect (I) if the plan does not use at least three different widths.

Each essentially correct (E) section counts as 1 point. Each partially correct (P) section counts as $\frac{1}{2}$ point.

- | | |
|---|----------------------|
| 4 | Complete Response |
| 3 | Substantial Response |
| 2 | Developing Response |
| 1 | Minimal Response |

If a response is between two scores (for example, 2 $\frac{1}{2}$ points), use a holistic approach to decide whether to score up or down, depending on the overall strength of the response and communication.