# STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II score—25

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.

6. Grass buffer strips are grassy areas that are planted between bodies of water and agricultural fields. These strips are designed to filter out sediment, organic material, nutrients, and chemicals carried in runoff water. The figure below shows a cross-sectional view of a grass buffer strip that has been planted along the side of a stream.
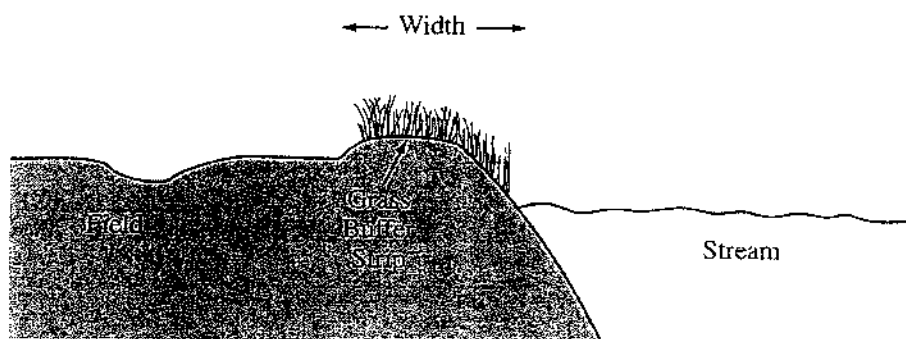
A study in Nebraska investigated the use of buffer strips of several widths between 5 feet and 15 feet. The study results indicated a linear relationship between the width of the grass strip (x), in feet, and the amount of nitrogen removed from the runoff water (y), in parts per hundred. The following model was estimated.

$$\hat{y} = 33.8 + 3.6x$$

(a) Interpret the slope of the regression line in the context of this question.

(b) Would you be willing to use this model to predict the amount of nitrogen removed for grass buffer strips with widths between 0 feet and 30 feet? Explain why or why not.

A scientist in California wants to know if there is a similar relationship in her area. To investigate this, she will place a grass buffer strip between a field and a nearby stream at each of eight different locations and measure the amount of nitrogen that the grass buffer strip removes, in parts per hundred, from runoff water at each location. Each of the eight locations can accommodate a buffer strip between 6 feet and 13 feet in width. The scientist wants to investigate which combination of widths will provide the best estimate of the slope of the regression line.

GO ON TO THE NEXT PAGE.

Suppose the scientist decides to use buffer strips of width 6 feet at each of four locations and buffer strips of width 13 feet at each of the other four locations. Assume the model, $\hat{y} = 33.8 + 3.6x$, estimated from the Nebraska study is the true regression line in California and the observations at the different locations are normally distributed with standard deviation of 5 parts per hundred.

(c) Describe the sampling distribution of the sample mean of the observations on the amount of nitrogen removed by the four buffer strips with widths of 6 feet.

(d) Using your result from part (c), show how to construct an interval that has probability 0.95 of containing the sample mean of the observations from four buffer strips with widths of 6 feet.

For the study plan being implemented by the scientist in California, the graph on the left below displays intervals that each have probability 0.95 of containing the sample mean of the four observations for buffer strips of width 6 feet and for buffer strips of width 13 feet. A second possible study plan would use buffer strips of width 8 feet at four of the eight locations and buffer strips of width 10 feet at the other four locations. Intervals that each have probability 0.95 of containing the mean of the four observations for buffer strips of width 8 feet and for buffer strips of width 10 feet, respectively, are shown in the graph on the right below.
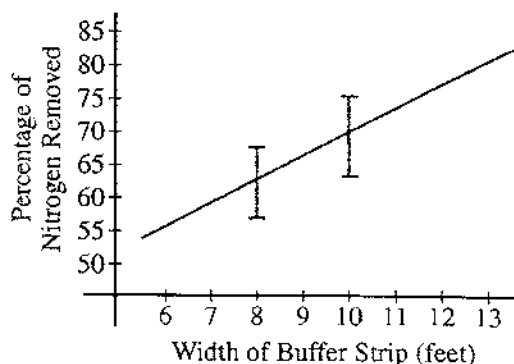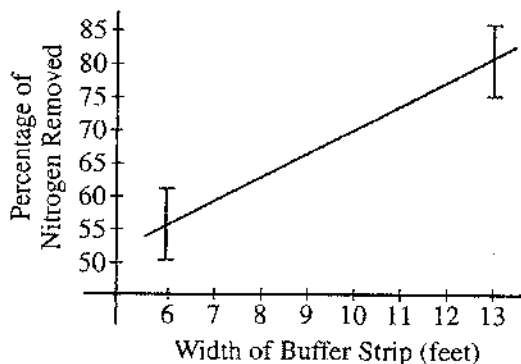
If data are collected for the first study plan, a sample mean will be computed for the four observations from buffer strips of width 6 feet and a second sample mean will be computed for the four observations from buffer strips of width 13 feet. The estimated regression line for those eight observations will pass through the two sample means. If data are collected for the second study plan, a similar method will be used.

(e) Use the plots above to determine which study plan, the first or the second, would provide a better estimator of the slope of the regression line. Explain your reasoning.

(f) The previous parts of this question used the assumption of a straight-line relationship between the width of the buffer strip and the amount of nitrogen that is removed, in parts per hundred. Although this assumption was motivated by prior experience, it may not be correct. Describe another way of choosing the widths of the buffer strips at eight locations that would enable the researchers to check the assumption of a straight-line relationship.

STOP

END OF EXAM

GO ON TO THE NEXT PAGE.

2011 Form B   #6

a) As the width of the buffer strip increases by ONE foot, we would predict the amount of Nitrogen that is removed to increase by 3.6 ppb.

b) Since the model ... was only predicted using strips between 5 and 15 feet, I would NOT be willing to use the model for 0 through 30 feet. That would be an Extrapolation which could cause errors in analysis.

c) Shape, Center, Spread

I would expect the sampling distribution to be approximately normal with a center at $\bar{y} = \mu_{\bar{x}} = E(x) = 55.4$ and a standard deviation of $\sigma_{\bar{x}} = 3.5$

d) $55.4 \pm (3.5)(\text{Invnorm}(.975))$

e) Plan A because ... the sample means would ... the extremes of the possibility, the line would still be close in slope to the Hfd Phi.

f) Instead of 4 measurements at two widths, try ... a greater variety of widths.