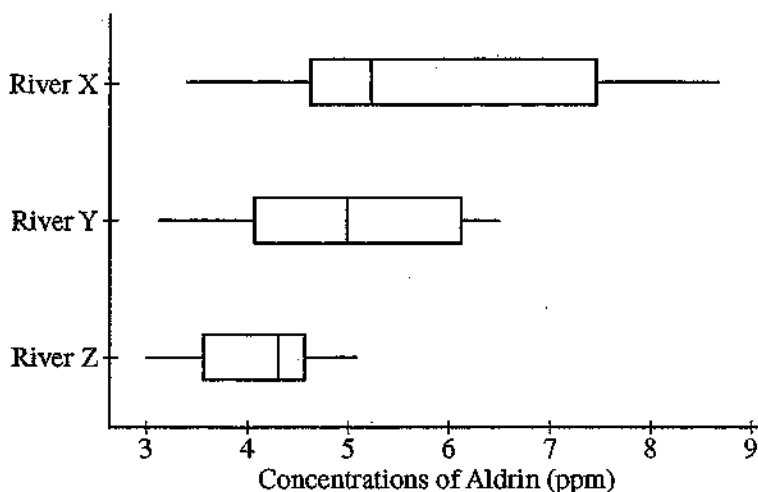


1. As a part of the United States Department of Agriculture's Super Dump cleanup efforts in the early 1990s, various sites in the country were targeted for cleanup. Three of the targeted sites—River X, River Y, and River Z—had become contaminated with pesticides because they were located near abandoned pesticide dump sites. Measurements of the concentration of aldrin (a commonly used pesticide) were taken at twenty randomly selected locations in each river near the dump sites.

The boxplots shown below display the five-number summaries for the concentrations, in parts per million (ppm) of aldrin, for the twenty locations that were sampled in each of the three rivers.



- (a) Compare the distributions of the concentration of aldrin among the three rivers.
- (b) The twenty concentrations of aldrin for River X are given below.

3.4 4.0 5.6 3.7 8.0 5.5 5.3 4.2 4.3 7.3
8.6 5.1 8.7 4.6 7.5 5.3 8.2 4.7 4.8 4.6

Construct a stemplot that displays the concentrations of aldrin for River X.

- (c) Describe a characteristic of the distribution of aldrin concentrations in River X that can be seen in the stemplot but cannot be seen in the boxplot.

5. An advertising agency in a large city is conducting a survey of adults to investigate whether there is an association between highest level of educational achievement and primary source for news. The company takes a random sample of 2,500 adults in the city. The results are shown in the table below.

Primary Source for News	HIGHEST LEVEL OF EDUCATIONAL ACHIEVEMENT			Total
	Not High School Graduate	High School Graduate But Not College Graduate	College Graduate	
Newspapers	49	205	188	442
Local television	90	170	75	335
Cable television	113	496	147	756
Internet	41	401	245	687
None	77	165	38	280
Total	370	1,437	693	2,500

- (a) If an adult is to be selected at random from this sample, what is the probability that the selected adult is a college graduate or obtains news primarily from the internet?
- (b) If an adult who is a college graduate is to be selected at random from this sample, what is the probability that the selected adult obtains news primarily from the internet?
- (c) When selecting an adult at random from the sample of 2,500 adults, are the events "is a college graduate" and "obtains news primarily from the internet" independent? Justify your answer.

2. In response to nutrition concerns raised last year about food served in school cafeterias, the Smallville School District entered into a one-year contract with the Healthy Alternative Meals (HAM) company. Under this contract, the company plans and prepares meals for 2,500 elementary, middle, and high school students, with a focus on good nutrition. The school administration would like to survey the students in the district to estimate the proportion of students who are satisfied with the food under this contract.

Two sampling plans for selecting the students to be surveyed are under consideration by the administration. One plan is to take a simple random sample of students in the district and then survey those students. The other plan is to take a stratified random sample of students in the district and then survey those students.

- (a) Describe a simple random sampling procedure that the administrators could use to select 200 students from the 2,500 students in the district.
 - (b) If a stratified random sampling procedure is used, give one example of an effective variable on which to stratify in this survey. Explain your reasoning.
 - (c) Describe one statistical advantage of using a stratified random sample over a simple random sample in the context of this study.
-

3. A test consisting of 25 multiple-choice questions with 5 answer choices for each question is administered. For each question, there is only 1 correct answer.

- (a) Let X be the number of correct answers if a student guesses randomly from the 5 choices for each of the 25 questions. What is the probability distribution of X ?

This test, like many multiple-choice tests, is scored using a penalty for guessing. The test score is determined by awarding 1 point for each question answered correctly, deducting 0.25 point for each question answered incorrectly, and ignoring any question that is omitted. That is, the test score is calculated using the following formula.

$$\text{Score} = (1 \times \text{number of correct answers}) - (0.25 \times \text{number of incorrect answers}) + (0 \times \text{number of omits})$$

For example, the score for a student who answers 17 questions correctly, answers 3 questions incorrectly, and omits 5 questions is

$$\text{Score} = (1 \times 17) - (0.25 \times 3) + (0 \times 5) = 16.25.$$

- (b) Suppose a student knows the correct answers for 18 questions, answers those 18 questions correctly, and chooses randomly from the 5 choices for each of the other 7 questions. Show that the expected value of the student's score is 18 when using the scoring formula above.
- (c) A score of at least 20 is needed to pass the test. Suppose a student knows the correct answers for 18 questions, answers those 18 questions correctly, and chooses randomly from the 5 choices for each of the other 7 questions. What is the probability that the student will pass the test?