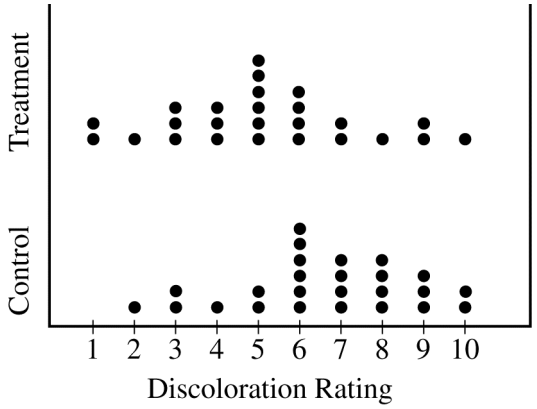
STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II grade—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. The department of agriculture at a university was interested in determining whether a preservative was effective in reducing discoloration in frozen strawberries. A sample of 50 ripe strawberries was prepared for freezing. Then the sample was randomly divided into two groups of 25 strawberries each. Each strawberry was placed into a small plastic bag.

   The 25 bags in the control group were sealed. The preservative was added to the 25 bags containing strawberries in the treatment group, and then those bags were sealed. All bags were stored at 0°C for a period of 6 months. At the end of this time, after the strawberries were thawed, a technician rated each strawberry's discoloration from 1 to 10, with a low score indicating little discoloration.

   The dotplots below show the distributions of discoloration rating for the control and treatment groups.



   (a) The standard deviation of ratings for the control group is 2.141. Explain how this value summarizes variability in the control group.

   (b) Based on the dotplots, comment on the effectiveness of the preservative in lowering the amount of discoloration in strawberries. (No calculations are necessary.)

   (c) Researchers at the university decided to calculate a 95 percent confidence interval for the difference in mean discoloration rating between strawberries that were not treated with preservative and those that were treated with preservative. The confidence interval they obtained was (0.16, 2.72). Assume that the conditions necessary for the *t*-confidence interval are met.

   Based on the confidence interval, comment on whether there would be a difference in the population mean discoloration ratings for the treated and untreated strawberries.

GO ON TO THE NEXT PAGE.

2. As dogs age, diminished joint and hip health may lead to joint pain and thus reduce a dog's activity level. Such a reduction in activity can lead to other health concerns such as weight gain and lethargy due to lack of exercise. A study is to be conducted to see which of two dietary supplements, glucosamine or chondroitin, is more effective in promoting joint and hip health and reducing the onset of canine osteoarthritis. Researchers will randomly select a total of 300 dogs from ten different large veterinary practices around the country. All of the dogs are more than 6 years old, and their owners have given consent to participate in the study. Changes in joint and hip health will be evaluated after 6 months of treatment.

   (a) What would be an advantage to adding a control group in the design of this study?

   (b) Assuming a control group is added to the other two groups in the study, explain how you would assign the 300 dogs to these three groups for a completely randomized design.

   (c) Rather than using a completely randomized design, one group of researchers proposes blocking on clinics, and another group of researchers proposes blocking on breed of dog. How would you decide which <u>one</u> of these two variables to use as a blocking variable?

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.

   (a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?

   - A random sample of 15 fish having a mean length that is greater than 10 inches

   or

   - A random sample of 50 fish having a mean length that is greater than 10 inches

   Justify your answer.

   (b) Suppose the standard deviation of the sampling distribution of the sample mean for random samples of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.

   (c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b) ? Justify your answer.

GO ON TO THE NEXT PAGE.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 1

Intent of Question

The goals of this question are to assess a student's ability to: (1) explain how a commonly used statistic measures variability; (2) use a graphical display to address the research question of interest in a simple comparative experiment; and (3) use a confidence interval to make an appropriate inference.

Solution

Part (a):

Roughly speaking, the standard deviation $(s = 2.141)$ measures a "typical" or "average" distance between the individual discoloration ratings and the mean discoloration rating for the strawberries in the control group.

Part (b):

The preservative does appear to have been effective in lowering the amount of discoloration in strawberries. The discoloration ratings for the strawberries that received the preservative, shown in the top dotplot, are clearly centered at a value that is lower than the center of the discoloration rating distribution for the control group, shown in the bottom dotplot. In addition, the dotplots can be used to find all five statistics in the five-number summary (min, Q1, median, Q3, and max) for both groups. In fact, four of the five statistics (the maximum is the only exception) are lower for the strawberries that received the preservative.

Part (c):

Since zero is not contained in the 95 percent confidence interval for the difference $\mu_c - \mu_t$, we can conclude that there is a significant difference between the mean ratings for the two groups at the $\alpha = 0.05$ level. The population mean discoloration rating for untreated strawberries is estimated to be between 0.16 and 2.72 units higher than the population mean discoloration rating for treated strawberries. Thus, we think there would be a difference in the population mean discoloration ratings for treated and untreated strawberries.

Scoring

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

Part (a) is scored as essentially correct (E) if the standard deviation is interpreted correctly in the context of this experiment.

Part (a) is scored as partially correct (P) if:
a standard textbook description of the standard deviation is provided without any reference to context, e.g., the standard deviation is described as the square root of an average squared deviation, or a "typical" or "average" deviation from the mean;
OR
the student provides evidence that the distribution of discoloration ratings in the control group is approximately normal, then correctly applies the 68-95-99.7 rule.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 1 (continued)

Part (a) is scored as incorrect (I) if:
the formula for the standard deviation is copied from the formula sheet and no further explanation is provided;
OR
the student uses the 68-95-99.7 rule without justifying that the distribution of discoloration ratings in the control group is approximately normal.

Part (b) is scored as essentially correct (E) if the student indicates that the preservative appears to be effective and explicitly links this decision to comparison of a characteristic of relative standing from the dotplots for the two groups.

Part (b) is scored as partially correct (P) if:
the student says that the preservative appears to be effective because the discoloration ratings appear to be lower for the treatment group, but the student does not explicitly link this decision to comparison of a characteristic of relative standing for the two groups;
OR
the student correctly compares one or more characteristics of relative standing for the two groups but never states that the preservative was effective in lowering discoloration.

Part (b) is scored as incorrect (I) if:
the student says that the preservative is not effective because the centers of the two distributions are roughly the same;
OR
the student says that the preservative is effective, with incorrect or no justification.

Part (c) is scored as essentially correct (E) if the student indicates that zero is not included in the confidence interval, so there is a difference (in population means), AND states the conclusion in the context of this experiment.

Part (c) is scored as partially correct (P) if:
the student indicates that zero is not included in the confidence interval, so there is a difference (in population means), but does not state the conclusion in the context of this experiment;
OR
the student correctly interprets the 95 percent confidence interval in context and indicates that there is a difference (in population means), without indicating that zero is not included in the confidence interval.

Part (c) is scored as incorrect (I) if the student concludes that the preservative is not effective *OR* says that no conclusion can be made based on the confidence interval, *OR* the student states a conclusion that refers to sample means instead of population means.

Notes:
- The student is not required to specify the significance level in part (c), but if it is specified, it must be correct.

Question 1 (continued)

- An adjustment could be made to formally conduct a one-sided test, but in general, confidence intervals are used to conduct two-sided tests. The fact that the lower endpoint of the confidence interval is positive does provide evidence that the preservative is effective in lowering the amount of discoloration in strawberries. The correct formal statement is: The 97.5 percent lower confidence bound for the difference in the means is above zero (0.16), so at the 0.025 level we would conclude that the mean rating for the treated berries is significantly lower than the mean for the untreated strawberries.

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