

2004 AP[®] STATISTICS FREE-RESPONSE QUESTIONS

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 of your results and explanation.

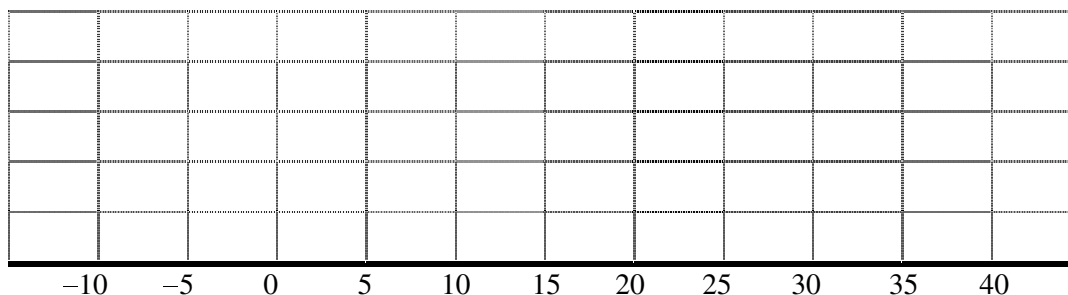
1. A consumer advocate conducted a test of two popular gasoline additives, A and B. There are claims that the use of either of these additives will increase gasoline mileage in cars. A random sample of 30 cars was selected. Each car was filled with gasoline and the cars were run under the same driving conditions until the gas tanks were empty. The distance traveled was recorded for each car.

Additive A was randomly assigned to 15 of the cars and additive B was randomly assigned to the other 15 cars. The gas tank of each car was filled with gasoline and the assigned additive. The cars were again run under the same driving conditions until the tanks were empty. The distance traveled was recorded and the difference in the distance with the additive minus the distance without the additive for each car was calculated.

The following table summarizes the calculated differences. Note that negative values indicate less distance was traveled with the additive than without the additive.

Additive	Values Below Q_1	Q_1	Median	Q_3	Values Above Q_3
A	-10, -8, -2	1	3	4	5, 7, 9
B	-5, -3, -3	-2	1	25	35, 37, 40

- (a) On the grid below, display parallel boxplots (showing outliers, if any) of the differences of the two additives.



- (b) Two ways that the effectiveness of a gasoline additive can be evaluated are by looking at either
 - the proportion of cars that have increased gas mileage when the additive is used in those cars
 - or
 - the mean increase in gas mileage when the additive is used in those cars.
 - i. Which additive, A or B, would you recommend if the goal is to increase gas mileage in the highest proportion of cars? Explain your choice.
 - ii. Which additive, A or B, would you recommend if the goal is to have the highest mean increase in gas mileage? Explain your choice.

2004 AP[®] STATISTICS FREE-RESPONSE QUESTIONS

2. Researchers who are studying a new shampoo formula plan to compare the condition of hair for people who use the new formula with the condition of hair for people who use the current formula. Twelve volunteers are available to participate in this study. Information on these volunteers (numbered 1 through 12) is shown in the table below.

Volunteer	Gender	Age
1	Male	21
2	Female	20
3	Male	47
4	Female	60
5	Female	62
6	Male	61
7	Male	58
8	Female	44
9	Male	44
10	Female	24
11	Male	23
12	Female	46

- (a) These researchers want to conduct an experiment involving the two formulas (new and current) of shampoo. They believe that the condition of hair changes with age but not gender. Because researchers want the size of the blocks in an experiment to be equal to the number of treatments, they will use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.
- (b) Other researchers believe that hair condition differs with both age and gender. These researchers will also use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.
- (c) The researchers in part (b) decide to select three of the six blocks to receive the new formula and to give the other three blocks the current formula. Is this an appropriate way to assign treatments? If so, describe a method for selecting the three blocks to receive the new formula. If not, describe an appropriate method for assigning treatments.

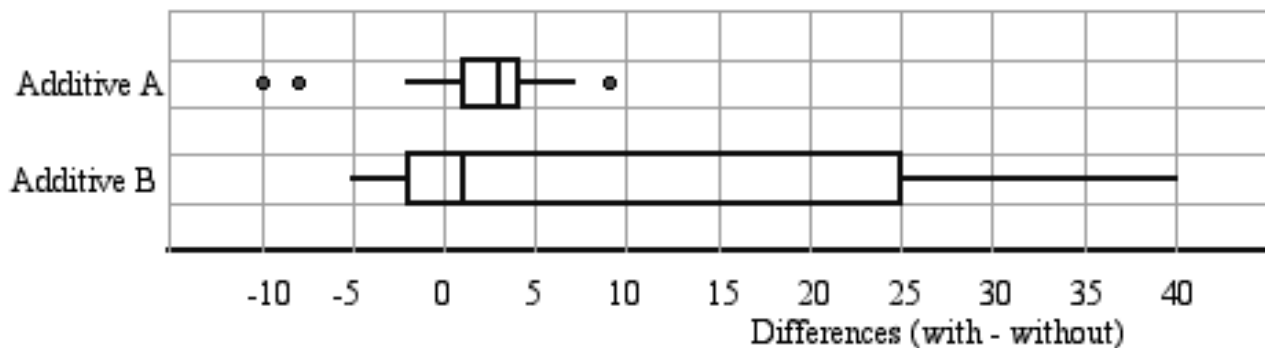
**AP[®] STATISTICS
2004 SCORING GUIDELINES**

Question 1

Solution

Part (a):

	Additive A	Additive B
IQR	$4 - 1 = 3$	$25 - (-2) = 27$
$1.5 \times \text{IQR}$	4.5	40.5
$Q_1 - 1.5 \times \text{IQR}$	$1 - 4.5 = -3.5$	$-2 - 40.5 = -42.5$
$Q_3 + 1.5 \times \text{IQR}$	$4 + 4.5 = 8.5$	$25 + 40.5 = 65.5$
	3 outliers	no outliers



Part (b i):

Additive A is better at increasing the mileage in the greatest number of cars. The mileage increased for at least seventy-five percent of the cars with additive A, whereas the mileage decreased for more than twenty-five percent of the cars with additive B.

Part (b ii):

Additive B appears to produce a higher mean mileage gain than additive A. The boxplot for additive B clearly shows that the distribution of differences is skewed to the right, which will pull the average towards the larger values. The mean difference for additive B will be substantially greater than the median of 1. On the other hand, the distribution of differences for additive A has much less variability, as seen by comparing the lengths of the two boxes, and appears to be skewed to the left. The mean difference for additive A will be less than the median of 3.

**AP[®] STATISTICS
2004 SCORING GUIDELINES**

Question 1 (cont'd.)

Scoring

Each part is scored as essentially correct, partially correct, or incorrect.

Part (a) is essentially correct if parallel boxplots are correctly drawn and the display includes labels. The labels for additives A and B are required, but if the label for either axis (or both axes) is the only item missing then part (a) is still scored as essentially correct.

Part (a) is partially correct if only one of the following occurs:

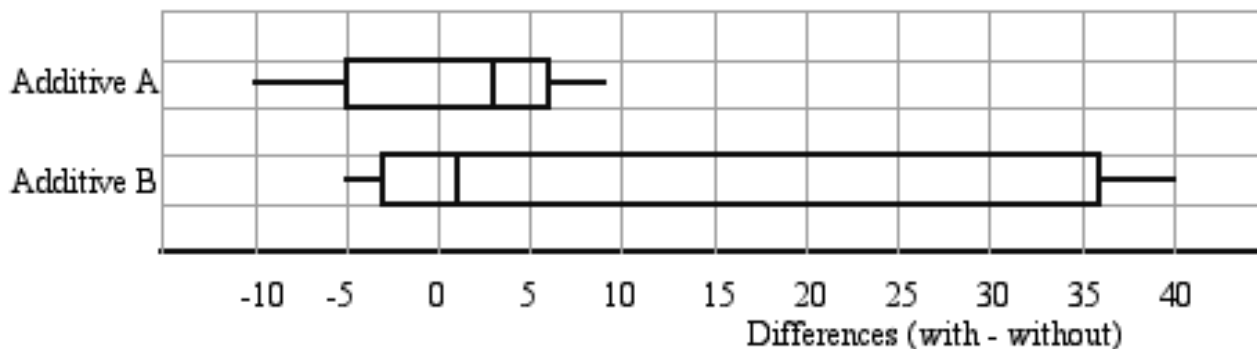
- if neither additive A nor B is labeled, or the labels are reversed, OR
- there is an error in the construction of the boxplots, OR
- outliers are not correct, but otherwise the boxplots are correct.

Note: If the same error is present in the boxplot for Additive A as in the boxplot for Additive B, count this as one error, not two.

Part (a) is incorrect if either of the following occur:

- 2 or 3 of the errors above are committed OR
- the numbers provided in the stem of the problem are used as a data set to form the boxplots below.

INCORRECT BOXPLOTS USING SUMMARY STATISTICS AS DATA



Note: If the labeling is reversed for boxplots A and B, score parts (b i) and (b ii) consistently with the reversed labels.

AP[®] STATISTICS
2004 SCORING GUIDELINES

Question 1 (cont'd.)

Part (b i) is essentially correct if additive A is selected as the additive that will increase the mileage in the greatest number of cars and the explanation involves a comparison implicitly or explicitly using the first quartiles from **both** distributions.

Part (b i) is partially correct if any of the following occur:

- additive A is selected, but the only comparison is between the lower quartile of additive A and the median of additive B, OR
- additive A is selected as the additive that will increase the mileage in the greatest number of cars but the comparison is weak or incomplete, such as relying only on variability or only on the median, OR
- additive A is selected and justified by addressing either A or B, but there is no explicit comparison to the other additive.

Part (b i) is scored as incorrect if any of the following occur:

- additive A is selected, but no explanation is provided, OR
- additive A is selected, but the explanation is incorrect, OR
- additive B is selected, OR
- summary statistics are used correctly in part (a), but used as data in part (b).

Part (b ii) is essentially correct if additive B is selected as the additive that will produce a higher mean mileage gain and the explanation involves a comparison using the shapes of **both** distributions.

For additive A it is sufficient to refer to the effect of the outliers on the mean.

Part (b ii) is partially correct if any of the following occur:

- additive B is selected as the additive that will produce a higher mean mileage gain but the comparison is weak or incomplete, OR
- additive B is selected and justified by addressing either A or B, but there is no explicit comparison to the other additive, OR
- additive A is selected, because the student argues that medians should be used to measure center instead of means because of the outliers or because the distributions are skewed, OR
- a comparison is made between A and B, but only one side, either left or right, is addressed for both plots.

Part (b ii) is scored as incorrect if any of the following occur:

- additive B is selected, but no explanation is provided,
- additive B is selected, but the explanation is incorrect,
- additive A is selected for any reason other than that stated above for partially correct,
- summary statistics are used correctly in part (a), but used as data in part (b), OR
- the student claims that the means cannot be determined from the information provided.