

2003 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. Since Hill Valley High School eliminated the use of bells between classes, teachers have noticed that more students seem to be arriving to class a few minutes late. One teacher decided to collect data to determine whether the students' and teachers' watches are displaying the correct time. At exactly 12:00 noon, the teacher asked 9 randomly selected students and 9 randomly selected teachers to record the times on their watches to the nearest half minute. The ordered data showing minutes after 12:00 as positive values and minutes before 12:00 as negative values are shown in the table below.

Students	-4.5	-3.0	-0.5	0	0	0.5	0.5	1.5	5.0
Teachers	-2.0	-1.5	-1.5	-1.0	-1.0	-0.5	0	0	0.5

- (a) Construct parallel boxplots using these data.
- (b) Based on the boxplots in part (a), which of the two groups, students or teachers, tends to have watch times that are closer to the true time? Explain your choice.
- (c) The teacher wants to know whether individual student's watches tend to be set correctly. She proposes to test $H_0: \mu = 0$ versus $H_a: \mu \neq 0$, where μ represents the mean amount by which all student watches differ from the correct time. Is this an appropriate pair of hypotheses to test to answer the teacher's question? Explain why or why not. Do not carry out the test.

2003 AP® STATISTICS FREE-RESPONSE QUESTIONS

2. When a law firm represents a group of people in a class action lawsuit and wins that lawsuit, the firm receives a percentage of the group's monetary settlement. That settlement amount is based on the total number of people in the group—the larger the group and the larger the settlement, the more money the firm will receive.

A law firm is trying to decide whether to represent car owners in a class action lawsuit against the manufacturer of a certain make and model for a particular defect. If 5 percent or less of the cars of this make and model have the defect, the firm will not recover its expenses. Therefore, the firm will handle the lawsuit only if it is convinced that more than 5 percent of cars of this make and model have the defect. The firm plans to take a random sample of 1,000 people who bought this car and ask them if they experienced this defect in their cars.

- Define the parameter of interest and state the null and alternative hypotheses that the law firm should test.
- In the context of this situation, describe Type I and Type II errors and describe the consequences of each of these for the law firm.

3. Men's shirt sizes are determined by their neck sizes. Suppose that men's neck sizes are approximately normally distributed with mean 15.7 inches and standard deviation 0.7 inch. A retailer sells men's shirts in sizes S, M, L, XL, where the shirt sizes are defined in the table below.

Shirt size	Neck size
S	$14 \leq \text{neck size} < 15$
M	$15 \leq \text{neck size} < 16$
L	$16 \leq \text{neck size} < 17$
XL	$17 \leq \text{neck size} < 18$

- Because the retailer only stocks the sizes listed above, what proportion of customers will find that the retailer does not carry any shirts in their sizes? Show your work.
- Using a sketch of a normal curve, illustrate the proportion of men whose shirt size is M. Calculate this proportion.
- Of 12 randomly selected customers, what is the probability that exactly 4 will request size M ? Show your work.

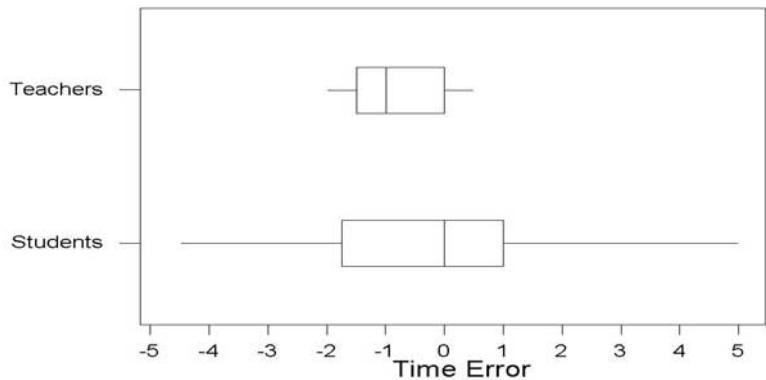
**AP[®] STATISTICS
2003 SCORING GUIDELINES**

Question 1

Solution

Part (a):

	Students	Teachers
Min	-4.5	-2.0
Q_1	-1.75 (or -0.5)	-1.5
Median	0	-1.0
Q_3	1.0 (or 0.5)	0
Max	5.0	0.5
	no outliers	no outliers



Part (b):

The teachers' watches tend to have times that are closer to the true time. Although the teachers' watches tend to be slow, the times are less variable than the student times and so more teachers' watches had times that were close to the true time.

Part (c):

Testing this pair of hypotheses will not answer the teacher's question. The mean amount of time by which student watches differ from the true time could be zero even when student watch times differ greatly from the true time if large positive differences are offset by large negative differences.

**AP[®] STATISTICS
2003 SCORING GUIDELINES**

Question 1 (cont'd)

Scoring:

Each part is essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct (E) if parallel boxplots are correctly drawn (using either version of students' quartiles), includes correct labels (students vs. teachers) and shows a common scale. If only one group is labeled, we'll infer the correct label for the other group. A label on the scale is not needed, but should be correct if included. The five-number summary is not required, but may be used to compensate for an incorrectly placed point on a boxplot.

Part (a) is partially correct (P) if there is only one error (e.g., omitting labels and/or scale, showing outliers, highly inconsistent scale, one misplaced point, an incorrect method of calculating quartiles, or reversing labels). Minor arithmetic errors will be overlooked.

NOTE: Two separate boxplots are acceptable as parallel plots providing the scales for the two plots are the same so that comparisons can be easily made.

Part (b) is essentially correct (E) if teachers are picked based on some form of variability (e.g., range, compactness, IQR, spread, min-to-max, etc.).

Part (b) is partially correct (P) if either "Teachers" are selected with an incorrect explanation or no explanation, OR "Students" are selected based on center.

Note: If $Q1 = -0.5$ and $Q3 = 0.5$ are being used for the students, then a choice of students based on IQR OR a choice of teachers based on range is scored essentially correct (E). An argument for students based on center stills scores as partially correct (P). A choice of one group over the other without indicating which measure of variability has been used is incorrect (I), but considering both issues without making a choice is an example of parallel solutions – each of which is essentially correct (E).

Part (c) is essentially correct (E) if the response says that testing these hypotheses will not answer the question posed and a plausible explanation is given. Plausible explanations should focus on things such as: positives and negatives canceling out, individual times versus group average, or variability. Reference to the boxplots is not expected.

Part (c) is partially correct (P) if the answer given is "No" and either there is no explanation and part (b) was scored as essentially correct (E),

OR

There is an incomplete or incorrect explanation that addresses the connection (or lack of connection) between the given hypotheses and the teacher's question. [Note: Answers that discuss conditions needed to *perform* the test do not address the issue of connection. Such responses are scored as incorrect (I), as are responses that attempt to change the teacher's question.]

**AP[®] STATISTICS
2003 SCORING GUIDELINES**

Question 1 (cont'd)

4 Complete Response (3E)

All three parts essentially correct

3 Substantial Response (2E 1P)

Two parts essentially correct and 1 part partially correct

2 Developing Response (2E 0P or 1E 2P or 3P)

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and 2 parts partially correct

OR

two parts partially correct

1 Minimal Response (1E 1P or 1E 0P or 0E 2P)

One part essentially correct and either 0 or 1 parts partially correct

OR

No parts essentially correct and 2 parts partially correct