

**Student Performance Q&A:
2008 AP[®] Statistics Free-Response Questions**

The following comments on the 2008 free-response questions for AP[®] Statistics were written by the Chief Reader, Christine Franklin of the University of Georgia, Athens. They give an overview of each free-response question and of how students performed on the question, including typical student errors. General comments regarding the skills and content that students frequently have the most problems with are included. Some suggestions for improving student performance in these areas are also provided. Teachers are encouraged to attend a College Board workshop to learn strategies for improving student performance in specific areas.

Question 1

What was the intent of this question?

The primary goals of this question were to assess a student's ability to (1) compare two distributions; (2) reevaluate shape, center, and spread for comparing the two distributions after one of the distributions is transformed by multiplying each of the data points by a constant; and (3) make a prediction about the means of the two distributions based on information derived about the behavior of the distributions from the boxplots.

How well did students perform on this question?

The mean score was 1.01 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

- Not understanding what is required to compare distributions
- Failing to *compare* measures of center and spread (instead, just listing statistics)
- Mentioning the shape of only one distribution
- Using mean for median in a boxplot
- Making a statement about standard deviations based on boxplots
- Incorrectly using "IQR" (interquartile range) to mean the quartiles themselves or the box portion of a boxplot
- Trying to infer normality or unimodality from a symmetric boxplot

- Confusing skewed to the left and skewed to the right
- Incorrectly describing range as “4–14” or “5–20”

Part (b)

- Not understanding the effect of multiplying observations by $\frac{4}{3}$ (linear transformation)
- Not understanding that this linear transformation will affect both center and spread
- Failing to *compare* the center and spread of the three-quarter-cup boxplot in part (a) to the center and spread of the three-quarter-cup boxplot in part (b)
- Just listing new statistics without comparisons

Part (c)

- Answering the question about means correctly but failing to provide an adequate justification based on the left skewness of the distribution for the one-cup serving size
- Using “low values” to explain that the mean is below the median for one cup instead of left skewness
- Making qualitative judgments about cereals themselves or some motive of the cereal companies to lure children into eating sugary cereals

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

This question demonstrated the importance of having students *practice communicating information in writing*. In many cases it appeared that students knew the content necessary to answer the question but were unable to synthesize and communicate the required information in their answers. Students often simply listed summary statistics and did not put into words the necessary comparisons and justifications for why a particular answer was given.

Question 2

What was the intent of this question?

The primary goals of this question were to assess a student’s ability to (1) identify a potential source of nonresponse bias and recognize a possible consequence for interpreting the results of a survey; (2) recognize that increasing sample size does not remove bias; and (3) recommend an appropriate course of action to solve a practical problem with the use of a survey.

How well did students perform on this question?

The mean score was 1.49 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

- Failing to provide statistical reasoning in explanations (instead, just listing a consequence)
- Failing to link a potential source of bias to a consequence with a specific direction in overestimating or underestimating the proportion

- Confusing nonresponse bias with undercoverage, validity, and question wording
- Talking about “eliminating” bias instead of a consequence
- Confusing “consequence” with the source or definition of response bias

Part (b)

- Arguing that nonresponse would occur in the second sample instead of realizing that nonresponse in the first sample would contribute to bias in the combined sample regardless of the outcome for the second sample
- Citing that some families could be selected by both samples
- Arguing that increasing the sample size would make the results more accurate

Part (c)

- Failing to give an explicit strategy for increasing response rate (instead, just stating that the survey should be “mandatory”)
- Using a new sampling method (e.g., SRS, cluster, stratified) without indicating how nonresponse would be reduced
- Confusing nonresponse bias with undercoverage, validity, and question wording
- Suggesting a strategy that creates another source of bias
- Changing the population (districtwide vote)

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

Poor communication (lack of clarity in students’ writing) was a major issue with responses to this question. Also, many students failed to provide statistical *justification* and *reasoning* in their answers.

Question 3

What was the intent of this question?

The primary goals for this question were to assess a student’s ability to (1) recognize and calculate the mean as the expected value of a probability distribution; (2) demonstrate how to use two distributions to form all possible ways a specific difference may occur; (3) calculate a probability for this specific difference occurring; and (4) calculate a probability from the probability distribution of all possible differences.

How well did students perform on this question?

The mean score was 2.2 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

- Not showing work or a formula to support written calculations
- Rounding expected values to integers

- Using nonuniversal calculator notation
- Showing only partial work (not showing key multiplications or additions)
- Showing work, but with minor arithmetic errors

Part (b)

- Not showing work or a formula to support written calculations
- Listing only two of the three score combinations
- Thinking that all combinations were equally likely
- Adding probabilities instead of multiplying them, or using other inappropriate formulas
- Showing work, but with minor arithmetic errors

Part (c)

- Not showing work or a formula to support written calculations
- Not adding the probabilities (but multiplying or averaging them instead, for example)
- Showing work, but with minor arithmetic errors

Part (d)

- Not showing work or a formula to support written calculations
- Finding the probability that Josephine will have a higher score (the positive differences) instead of Crystal having a higher score (the negative differences)
- Not using the answer obtained in part (c) to complete the table
- Failing to realize that the probabilities in the table have a sum of 1
- Showing work, but with minor arithmetic errors

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

It is important to help students understand the necessity of *always showing complete work* when arriving at a numerical answer. As evidenced by the common errors in all parts, (a) through (d), students so often only gave what we call a bald answer: one with no work or justification. Some students earned a score of 0 on this question; although they computed all of the final answers, they showed no supporting work. Also stress with students the importance of providing written statements instead of just calculator syntax as part of their explanations. Calculator syntax without written supporting statements earned students little or no credit.

Question 4

What was the intent of this question?

The primary goals for this question were to assess a student's ability to (1) create and interpret a scatterplot and (2) estimate a proportion and the associated standard error.

How well did students perform on this question?

The mean score was 1.42 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

- Constructing a scatterplot of counts but not recognizing that the differing number of devices tested at each temperature makes this type of plot not useful (however, the labeling and scaling of the plots were generally well done)
- Using the term “probability” synonymously with “proportion”

Part (b)

- Failing to address the strength of the relationship and the linearity (however, the direction of the relationship was often given in context)

Part (c)

- Calculating a regression equation (using all four points) and using the equation to construct an estimate at 40°C
- Not knowing how to calculate the standard error (very often students would use the standard error of the difference of the two proportions or of a single proportion)

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

A careful reading of the question might have helped students understand the inherent problems with graphically representing the differing numbers of devices at various temperatures. In addition, students should have a basic understanding of where and when to use the formulas. With respect to the formulas for standard errors, some understanding of the components that contribute to the variation would have aided students in identifying the appropriate formulas or making minor modifications to a standard formula.

Question 5

What was the intent of this question?

The primary goals of this question were to assess a student’s ability to (1) state the appropriate hypotheses; (2) identify and compute the appropriate test statistic; (3) make a conclusion in the context of the problem; and (4) compare two sets of proportions to identify the preferred habitat.

How well did students perform on this question?

The mean score was 0.76 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

Step 1: Hypotheses

- Stating hypotheses in terms of sample data rather than population
- Stating an alternative hypothesis that indicates *all* proportions must differ
- Using unclear language (e.g., “a moose in an area”)

Steps 2 and 3: Test and Mechanics

- Omitting degrees of freedom
- Naming the wrong test or not naming a test at all

Step 4: Conclusion

- Stating a weak linkage or no linkage to the p -value/rejection region
- Confusing the sample and population

Part (b)

- Using chi-square contribution as the only justification for choice
- Choosing habitat type 3 due to “more moose than expected” without indication of largest difference

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

Part of the poor performance on this inference question may have been due to the statistical test needed for answering the question (chi-square goodness of fit), which is one that teachers often omit if instructional time is limited. Pacing one’s instruction in the course is vital to ensure that time is available to cover all of the statistical tests listed in the Topic Outline in the *AP Statistics Course Description*. Also, the evaluation of student responses for test-of-hypotheses questions has used the same standard scoring scheme since 1999. More information about scoring can be found in the scoring guidelines for past exams, which are posted on AP Central®. There, teachers will also find numerous test-of-hypotheses questions from past exams that can be used to help prepare students for this type of question.

Question 6

What was the intent of this question?

The primary goals of this investigative task were to assess a student’s ability to (1) identify and conduct an appropriate inference based on the differences in the posttest and pretest scores; (2) identify and interpret appropriate information from statistical software; (3) make an inference based on separate regression analyses; and (4) recognize and explain the additional information provided from the different analyses.

How well did students perform on this question?

The mean score was 1.13 out of a possible 4 points.

What were common student errors or omissions?

Part (a)

Component 1: Hypotheses

- Omitting the μ notation and/or not mentioning mean
- Using \bar{x} instead of μ
- Using a two-sided alternative hypothesis

Component 2: Test and Conditions

- Not checking normality (at least not well enough)
- Mentioning only sample size
- Omitting to check conditions entirely
- Checking only for one of the two groups (this was a less common error)
- Neglecting to cite independence

Component 3: Mechanics

- Reporting only p -value, not the test statistic or degrees of freedom

Component 4: Conclusion

- Not citing means or “on average” in conclusion
- Not providing linkage (this was a less common error)

Part (b)

Equations, Slope Interpretations

- Not including randomness/variability aspect (this was a very common error)
- Not defining variables or writing equations in context
- Not saying “increase” or “additional” for both variables
- Not including “hat” on y in the equation
- Misreading output

Part (c)

P -values, Conclusions for Slope Test

- Reporting p -values for intercepts instead of intercept values
- Presenting an insufficient interpretation in (i) that amounts to “accept the null hypothesis”
- Making a statement like “Conclude that there is no correlation ...”
- Exhibiting confusion about when to reject or not reject the null hypothesis

Part (d)

Additional Information Provided by Regression Analyses

- Not using information from the scatterplot
- Not being able to see in context connections across various parts of the question
- Providing a laundry list of regression items (e.g., slope, R-squared value) without referring to the context of the question
- Confusing degree of improvement (post - pre) with steepness of slope/correlation
- Contradicting the conclusion from part (a) (e.g., “They show us that students at the original school are showing greater improvement because the slope is steeper.”)

Based on your experience of student responses at the AP Reading, what message would you like to send to teachers that might help them to improve the performance of their students on the exam?

The good news with respect to this year’s investigative task is that for the hypothesis-testing parts, many hypotheses were nearly acceptably stated (though sometimes ill defined); many students chose the correct test to use; many students drew correct conclusions expressed in the context of the question; and many students identified p -values correctly and made the correct decision based on the p -value. Many students also provided the necessary linkage between the conclusion and the hypotheses. In other words, students

were able to perform reasonably well on the "textbook" parts of the investigative task. However, in part (d) very, very few students were able to think outside the box and go beyond the listing of the standard answers one would give to a question of this type. *More experience with investigative task types of questions* will help students integrate concepts and statistical tools, allowing them to apply their knowledge in a new setting. As noted in the preceding list of common mistakes, few students made an effort to see the connection of the different parts of the investigative task. Also, students should be reminded to allow plenty of time for the investigative task, since it counts for more than any of the other five free-response questions.