Student Performance Q&A:

2002 AP® Statistics Free-Response Questions

The following comments are provided by the Chief Reader about the 2002 free-response questions for AP Statistics. *They are intended to assist AP readers as they develop training sessions to help teachers better prepare their students for the AP Exams*. They give an overview of each question and its performance, 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 included. Readers are encouraged to use their expertise to create strategies for teachers to improve student performance in specific areas.

Question 1

What was intended by the question?

The purpose of Question 1 was to assess the student's ability to interpret a graphical display and to reason using interval estimates. To receive full credit for this question, the student needed to comment on the increasing precision of the estimates over time and to use the intervals in the graphical display to discuss the degree of support provided for each of the two competing theories. Good communication was particularly important on this question.

How well did students perform?

There were many good, concise responses to this question, but many students had difficulty using statistical terminology correctly in the discussion.

What were common errors or omissions?

Common errors in answering Question 1 are listed below.

- Many students incorrectly used statistical terminology in critical places. This often resulted in the response being either ambiguous or incorrect as written.
- Many students used the term "margin of error" incorrectly. They did not seem to understand that the margin of error is a number and used the term as if it was an interval.
- In part (a), many students used the terms "experimental values", "observations", "data", and "estimates" interchangeably.
- In parts (b) and (c), many students focused only on the point estimates, ignoring the intervals altogether.
- In parts (b) and (c), many students used the terms "impossible", "certain", and "proved" inappropriately.
- Some students mistakenly thought the graphical display was a residual plot from a regression analysis, and tried to interpret the plot in this context.

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

Students need to be more careful with the use of statistical terminology. Providing more opportunities for students to engage in written communication would be helpful.

Question 2

What was intended by the question?

The purpose of Question 2 was to assess understanding of some basic principles of experimental design, including pairing/blocking, randomization, and blinding. A response describing a matched pairs design with appropriate randomization and a correct discussion of double-blinding received full credit for this question.

How well did students perform?

Most students produced a completely randomized design. While this design would allow comparison of the two treatments, it is not as good as a matched pairs design.

What were common errors or omissions?

Common errors in answering Question 2 are listed below.

- Many students provided a diagram of their design with no explanation, even though the question specifically stated, "Include a few sentences on how (the design) would be implemented."
- The vast majority of students gave a completely randomized design with two treatment groups. Although this design allows the two treatments to be compared, it is not as good as a matched pairs design.
- Some students identified potential blocking variables, such as gender, occupation, climate, etc. Students then formed blocks based on one or more of these variables. While not incorrect, this still produces a design that is inferior to the matched pairs design.
- Many students failed to describe a randomization scheme appropriate to their design.
- Some students did not understand that the design was to use the 100 volunteers given and tried to devise a way to randomly select volunteers from some population.
- In part (b), many students realized that double-blinding involved two parties being unaware of the treatment assignments, but most were confused about who besides the volunteers was to be blinded. Most students did not adequately convey that the second party that must be blinded was the person who would be evaluating the boots for wear.
- Some students neglected to mention that the volunteers themselves must also be blinded, and only addressed the blinding of some other person.
- Some students incorrectly indicated that randomization eliminates the need for blinding.

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

Many students relied solely on diagrams to show their design, even though the question specifically asked them to write a few sentences. Students should be encouraged to read the

questions carefully. In teaching experimental design, do not rely too much on diagrams for describing a design. Some designs, like the matched pairs design of this question, are very difficult to diagram in a way that conveys all of the necessary information. Encourage students to explain experimental designs in words as well as in pictures.

Question 3

What was intended by the question?

The purpose of Question 3 was to evaluate the student's ability to compute probabilities based on the normal distribution and to evaluate the student's knowledge of properties of the distribution of a sum of independent random variables. To receive full credit for this question, the student was required to compute and interpret a probability, correctly determine the mean and standard deviation of a sum of independent random variables, and then use the computed mean and standard deviation to correctly compute a probablility.

How well did students perform?

Students generally did well on part (a) of this problem, but often were unable to determine the mean and standard deviation in part (b).

What were common errors or omissions?

Common errors in answering Question 3 are listed below.

- In part (a) some students claimed that the event was unlikely because it was more than two standard deviations from the mean, without appealing to the normality of the distribution of times.
- Some students claimed that it is impossible for a variable to take on a value more than one standard deviation from its mean.
- Many students did not know how to compute the standard deviation of a sum of independent random variables.
- Some students tried to turn this problem into an inference problem and attempted to carry out some sort of hypothesis test.

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

Again, students should be encouraged to read questions carefully. Many students were unable to compute the standard deviation in part (b), and may not have recognized the question as one that involved a sum of independent random variables in context. Providing examples of this type in a variety of different contexts may help students to recognize this type of problem. Also, students who drew pictures of the probability distributions were much less likely to compute the wrong probability. Encourage students to draw pictures where appropriate.

Question 4

What was intended by the question?

The purpose of Question 4 was to determine if the student could read standard statistics computer output and to assess understanding of correlation and influential points in a regression analysis. To receive full credit for this question, the response had to include the correct equation for the least-squares regression line, a correct interpretation of the correlation coefficient, and a reasonable discussion of whether it was reasonable to use the given line over a restricted range of plane sizes.

How well did students perform?

Students who could read computer output and who knew the difference between the correlation coefficient (r) and the coefficient of determination (r^2) generally did well on this question.

What were common errors or omissions?

Common errors in answering Question 4 are listed below.

- Many students could not read the given computer output. Often students misinterpreted the value of s (the standard deviation about the regression line) as the slope of the line.
- Many students did not define the variables, even though the question requested this be done.
- Many students confused the correlation coefficient and the coefficient of determination.
- Some students incorrectly used the adjusted r^2 value to compute the value for the correlation coefficient.
- Some students described a correlation coefficient of 0.755 as weak, fairly weak, or extremely weak, and did not recognize that a value this large is indicating a moderately strong relationship.
- Some students reported values for the correlation coefficient that were clearly unreasonable, such as 7.55 or 4.02. These students did not recognize that the correlation coefficent must be between –1 and +1.
- In part (c), many students decided the question was asking them about extrapolation, rather than prediction over a restricted range.
- Some students indicated that influential points would be removed from the analysis, but then did not describe the effect of removing them.

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

Many students did not appear to be comfortable reading computer output. Students should have the opportunity to work with computer output throughout the course. Students should also be encouraged not to be careless in the use of statistical terminology.

Question 5

What was intended by the question?

The purpose of Question 5 was to evaluate whether the student could carry out a test of hypotheses and state conclusions in context. To receive full credit for this question, the student needed to identify two distinct pairs of hypotheses in part (a) and then in part (b) to identify an appropriate test procedure, check (not just state) any necessary conditions for the test, and then, based on the result of the test, give an appropriate conclusion in context.

How well did students perform?

Student performance on hypothesis testing questions has been improving in recent years. More students (but still not the majority) are addressing necessary conditions and students are doing a better job of stating a correct conclusion in context.

What were common errors or omissions?

Common errors in answering Question 5 are listed below.

- Many students failed to address all of the required conditions (assumptions) for the test.
- Some students stated the required conditions for the test, but did not check to determine if these conditions were met.
- Many students did not provide any link between the computations performed as part of the test procedure and the eventual conclusion.
- Many students did not provide a conclusion in context.
- In part (a), many students did not give two new pairs of hypotheses. Some gave only one, and others merely rewrote the hypotheses given in the problem in a different form.
- Too many students stated the hypotheses for the test in terms of sample rather than in terms of population characterisitics.

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

Emphasize the importance of checking required conditions for statistical procedures. Encourage students to be explicit about what led them to reject or fail to reject the null hypothesis, and to provide a conclusion in the context of the question. Emphasize the difference between sample values and population characteristics, and especially emphasize that hypotheses must be in terms of population characteristics.

Question 6

What was intended by the question?

Question 6 was the exam's investigative task. As such, its purpose is to evaluate the student's understanding in several course topic areas and to assess ability to integrate statistical ideas and apply them in a new context. This year's investigative task involved confidence intervals and hypothesis testing, as well as some concepts from sampling.

How well did students perform?

Student performance on this question was disappointing, as parts (a), (b), and (c) addressed material central to the AP Statistics course. Correct computions, by themselves, were not sufficient for earning credit in any of the parts of this question. This was the lowest scoring question on this year's exam. The surprise was that students tended to do well on the investigative part (d), but not on parts (a), (b), and (c), which should have been routine for students.

What were common errors or omissions?

Common errors in answering Question 6 are listed below.

- In part (a), most students failed to recognize that there are required conditions (assumptions) that need to be checked prior to computing a confidence interval.
- In part (a), many students failed to provide an interpreation of the computed interval, even though one was specifically requested.
- In part (a) students often incorrectly interpreted the confidence interval as meaning that 95 percent of the population is between 0.517 and 0.625, or by saying that 95 percent of the time the sample proportion (or population proportion) was between 0.517 and 0.625.

- In part (b), many students could not give a correct interpretation of the confidence level.
- In part (c), many students had difficulty with notation, often stating hypotheses in terms of sample values rather than population characteristics.
- In part (c) many students failed to state and check required conditions (assumptions) before performing the test.
- In part (d), some students failed to recognize that the difference in sample sizes created an imbalance in the pooled estimate. Students who recognized the need to balance the sample sizes were generally successful in part (d).

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

Emphasize that there are required conditions that need to be checked prior to using a confidence interval or hypothesis testing procedure. Provide students with the opportunity to practice interpreting both the confidence interval and the associated confidence level.

General Comments on Exam Performance

Exam performance this year (and in past years) was strongest in the area of describing data and weakest in the area of statistical inference. This was apparent in both the free-response inference questions as well as in the multiple-choice questions dealing with inference. In general, students were much stronger on the mechanical and computational aspects of problems than on parts that required interpretation or conceptual understanding. Communication of results continues to be a weakness.

Areas identified in previous exam reports that continue to be problematic are listed below.

- Many students failed to read questions carefully and, as a result, answered a question different from the one that was asked.
- Many students did not answer questions **in context**. Explanations and conclusions in context are always necessary for a complete answer.
- More students stated assumptions when carrying out a hypothesis test, but few understood that assumptions must also be checked.
- A disappointingly large number of students still seemed to believe that it is OK to draw conclusions by "just looking at the data", and did not seem to understand the need to employ inferential procedures, even when asked to provide statistical evidence to support conclusions.

What Can Teachers Do To Improve Performance In Specific Problem Areas?

- Emphasize conceptual understanding and communication over mechanics.
- Have students practice communicating conclusions and interpreting results throughout the course, and not just at the end when inferential techniques are covered.
- Allow sufficient time to cover the entire course outline.
- Integrate computer use if possible and, at a minimum, be sure that students are comfortable reading computer output.