



**AP[®] Statistics
2004 Scoring Commentary
Form B**

The materials included in these files are intended for noncommercial use by AP teachers for course and exam preparation; permission for any other use must be sought from the Advanced Placement Program[®]. Teachers may reproduce them, in whole or in part, in limited quantities, for face-to-face teaching purposes but may not mass distribute the materials, electronically or otherwise. This permission does not apply to any third-party copyrights contained herein. These materials and any copies made of them may not be resold, and the copyright notices must be retained as they appear here.

The College Board is a not-for-profit membership association whose mission is to connect students to college success and opportunity. Founded in 1900, the association is composed of more than 4,500 schools, colleges, universities, and other educational organizations. Each year, the College Board serves over three million students and their parents, 23,000 high schools, and 3,500 colleges through major programs and services in college admissions, guidance, assessment, financial aid, enrollment, and teaching and learning. Among its best-known programs are the SAT[®], the PSAT/NMSQT[®], and the Advanced Placement Program[®] (AP[®]). The College Board is committed to the principles of excellence and equity, and that commitment is embodied in all of its programs, services, activities, and concerns.

For further information, visit www.collegeboard.com

Copyright © 2004 College Entrance Examination Board. All rights reserved. College Board, Advanced Placement Program, AP, AP Central, AP Vertical Teams, APCD, Pacesetter, Pre-AP, SAT, Student Search Service, and the acorn logo are registered trademarks of the College Entrance Examination Board. PSAT/NMSQT is a registered trademark of the College Entrance Examination Board and National Merit Scholarship Corporation. Educational Testing Service and ETS are registered trademarks of Educational Testing Service. Other products and services may be trademarks of their respective owners.

For the College Board's online home for AP professionals, visit AP Central at apcentral.collegeboard.com.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 1

Sample: 1 of 3

Score: 4

The response to part (a) clearly identifies the decrease in impact rate as the age of the craters increases. It also describes the curvature by referring to a negative exponential relationship with asymptotes. The response to part (b) correctly interprets the value of r^2 as the percentage of the variability of the observed values of $\ln(\text{impact rate})$ that is accounted for by changes in $\ln(\text{age})$, although reference to the log-transformed variables could have been more clearly stated. The response to part (c) describes the quadratic pattern in the residual plot and concludes that the proposed regression model is not appropriate and some other transformation of the variables would be needed to obtain a straight-line relationship. This response earns a score of 4.

Sample: 2 of 3

Score: 3

This response clearly describes the negative association between impact and age by noting that impact rate decreases as age increases. It refers to the relationship as moderately strong and also refers to a curved pattern. The response to part (b) incorrectly ignores the transformed variables and interprets the value of r^2 as the percentage of the variability in impact rate values that can be attributed to changes in age values. The response to part (c) uses the curved pattern in the residual plot to argue that the linear regression model is inappropriate. This response earns a score of 3.

Sample: 3 of 3

Score: 3

The response to part (a) correctly refers to a negative association between impact rate and age, but it fails to point out the curvature in the relationship. The response to part (b) clearly interprets the value of r^2 as the percentage of the variability of the observed values of $\ln(\text{impact rate})$ that is accounted for by observed changes in $\ln(\text{age})$. The response to part (c) points out a curved pattern in the residual plot and concludes that the proposed regression model is not appropriate for modeling the relationship between the transformed variables. This response earns a score of 3.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 2

Sample: 1 of 3

Score: 4

The response to part (a) indicates that bias may have been introduced by selecting the first 100 students entering the dining hall, because students who arrive early may enjoy the food more than other students. It also indicates that a simple random sample of 100 students would have been better, because that sampling procedure gives every possible sample of 100 students an equal chance of being selected. The response to part (b) correctly identifies both the potential sources of bias response and presents a new way to ask the question that avoids both problems. This response earns a score of 4.

Sample: 2 of 3

Score: 4

The response to part (a) indicates that students entering the dining hall early may differ from other students with respect to tastes or other characteristics. This shows that the opinions of food quality held by the first 100 students entering the dining hall might not be representative of the larger population of students living in the dormitories. It goes on to describe how a simple random sample of students could be obtained. The response to part (b) identifies only one of the two potential sources of bias, the leading statement that “Many students believe that the food served in the dining hall needs improvement,” and correctly declares that the problem could be fixed by deleting the first sentence of the survey item. This response earns a score of 4.

Sample: 3 of 3

Score: 3

The response to part (a) suggests that some of the first 100 students entering the dining hall may not live in the dormitories and some students who live in the dormitories may not eat in the dining hall. Mentioning the possibility that the manager’s sampling procedure could include students who do not live in the dormitories or exclude students who do live in the dormitories is an incomplete response, however, because it does not indicate how this might induce bias in the estimation of the proportion of students living in the dormitories who think food service should be improved. It does not indicate that the opinions about food quality held by early arrivals at the dining hall might differ from the opinions of the other students who live in the dormitories. The response to part (a) correctly describes how one could obtain a simple random sample of students who live in the dormitories. The response to part (b) is essentially correct even though it identifies only one of the two potential sources of bias and provides an alternative question that fixes both sources of bias. This response earns a score of 3.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 3

Sample: 1 of 3

Score: 4

The response to part (a) clearly shows how the probability is obtained using both a probability formula and a diagram. The calculation is done correctly, but the answer is reported as a percentage instead of a probability. Using both the computed percentage from part (a) and the standard deviation, the student reaches an appropriate conclusion in part (b). The response to part (c) clearly shows how the standard deviation of the average weight of ten hopper cars is computed, and it also shows how the probability of observing an average weight larger than 70.7 tons is computed. The computations are correct and used to reach an appropriate conclusion in part (d). The conclusion is stated in the context of the ore-loading problem. This response earns a score of 4.

Sample: 2 of 3

Score: 4

The response to part (a) clearly shows how the z -value is obtained and describes the probability in words. The calculation is done correctly. The response to part (b) is presented as a test of the null hypothesis that the filling equipment produces a mean filling weight of 70 tons per car, and an appropriate conclusion is reached and stated in the context of the ore loading problem. The response to part (c) clearly shows how the standard deviation of the average weight of ten hopper cars and the appropriate z -value are obtained. The probability is accurately evaluated. The response to part (d) is also presented as a test of the null hypothesis that the filling equipment produces a mean filling weight of 70 tons per car, using the average weight of a sample of 10 cars and a known population standard deviation. An appropriate conclusion is reached and stated in the context of the ore-loading problem. This response earns a score of 4.

Sample: 3 of 3

Score: 3

The response to part (a) shows how the z -value is obtained and describes the probability. The calculation is done correctly, but the probability is reported as a percentage. The response to part (b) is incorrect. It suggests that observing a car weighing at least 70.7 tons would be more likely to occur when the filling mechanism is working properly than when it is overfilling. If the mechanism was overfilling and produced a mean weight of 70.7 tons, for example, the probability of observing a car weighing more than 70.7 tons would be 0.5, which is greater than 0.218. The response to part (c) clearly shows how the standard deviation of the average weight of ten hopper cars and the appropriate z -value are obtained. The probability is accurately evaluated. The small probability computed in part (c) is used in part (d) to reach an appropriate conclusion. This response earns a score of 3.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 4

Sample: 1 of 3

Score: 4

In part (a), the student addresses the assumption of independent random samples from two normal distributions that provides the foundation for the construction of a confidence interval for the difference in two population means. The two-sample t confidence interval is identified by both words and formula. The confidence interval is correctly evaluated and it is correctly interpreted with respect to the difference in the mean times spent studying for the populations of sixth and seventh graders. The response in part (b) attempts to explain that the procedure proposed by the assistant principal is incorrect because it artificially matches pairs of students from independent samples, when no relationship exists between students from different independent samples. This response earns a score of 4.

Sample: 2 of 3

Score: 3

In the middle of the response to part (a), this student assesses the assumptions of independent random samples from two normal distributions using the information provided. The formula for a two-sample confidence interval is displayed on the second page. The confidence interval is correctly evaluated, but it is not interpreted with respect to the difference in the mean times spent studying for the populations of sixth and seventh graders. The interpretation only refers to the “true difference in time spent.” The response to part (b) explains that the procedure proposed by the assistant principal is not better than the procedure used in part (a) because the samples are independent. It goes on to show how a before-after type of matched pairs study could be performed. This response earns a score of 3.

Sample: 3 of 3

Score: 3

At the beginning of the response to part (a), this student indicates that the t -distribution will be used to construct a confidence interval and addresses the assumption of independent random samples. The student notes that the sample standard deviations are about the same size and decides to use a pooled estimate of variance in estimating the standard deviation for the difference between the two sample means. The formula for a two-sample confidence interval is displayed and the confidence interval is correctly evaluated. The confidence interval is correctly interpreted with respect to the difference in the mean times spent studying for the populations of sixth and seventh graders. The response to part (b) suggests that the student knows that pairing can be used to reduce the standard deviation of the difference in two sample means, but it does not address the question posed in part (b). This response provides nothing to indicate that the student knows that matching on the response variable is inappropriate. This response earns a score of 3.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 5

Sample: 1 of 3

Score: 4

The boxplots are correctly drawn in part (a). The outliers in the sample of golden jackals are displayed, and the labeling enables us to determine which plot displays the data for modern Thai dogs and which plot displays the golden jackal data. The student compares the medians, ranges, and the interquartile ranges of the data from the two samples. The right skewness in the golden jackal data is noted. In part (b), the student infers from the boxplots that the distribution of modern Thai dog mandible lengths is roughly normal and indicates that it would be appropriate to construct a confidence interval. The response to part (c) concludes that it would be inappropriate to perform a two-sample t -test because the data indicate that the distribution of golden jackal mandible lengths is not symmetric. This response earns a score of 4.

Sample: 2 of 3

Score: 4

The boxplots are correctly drawn and labeled in part (a). The outliers in the sample of golden jackal mandible lengths are displayed. The student compares the medians and the interquartile ranges of the data from the two samples. The indication of right skewness in the distribution of golden jackal mandible lengths was not discussed in this part. In addition to indicating that there are no outliers in part (b), the student indicates that approximate symmetry of the distribution is implied because the sample median is approximately the same distance from each quartile. This leads to the conclusion that it would be appropriate to construct a confidence interval. The response to part (c) concludes that it would be inappropriate to perform a two-sample t -test because the golden jackal mandible data contain outliers and the data indicate that the distribution of golden jackal mandible lengths is heavily skewed to the right. This response earns a score of 4.

Sample: 3 of 3

Score: 3

The boxplots are correctly drawn and labeled in part (a). The response only compares the shapes of the distributions of mandible lengths for modern Thai dogs and golden jackals. It fails to compare either estimates of location or spread of the distributions. The student compares the medians and the interquartile ranges of the data from the two samples. In part (b), the student addresses the need to use the t -distribution to construct the confidence interval, because the sample size is moderate and indicates that the boxplot for the modern Thai dog data suggests that the normality assumption is plausible. The response to part (c) concludes that it would be inappropriate to perform a two-sample t -test because the boxplot indicates that the distribution of golden jackal mandible lengths is highly skewed to the right. This response earns a score of 3.

AP[®] STATISTICS
2004 SCORING COMMENTARY (Form B)

Question 6

Sample: 1 of 3

Score: 4

In part (a), the sample proportions are clearly labeled, which provides the necessary labeling of the population proportions used in the statement of the null and alternative hypotheses (step 1). The two-sample z -test is indicated and sample size conditions are checked (step 2). This student used the unpooled estimate of the variance of a difference in two proportions in the z -test. The test statistic and p -value are correctly evaluated (step 3). The conclusion is linked to the computed p -value and correctly stated in the context of the difference in the proportions of banded birds on the two islands (step 4). The total number of birds on Island A is correctly estimated in part (b). In part (c), the student lists concerns about the possible effects of banding on survival rates and whether weaker or sick birds might be more easily captured. This response earns a score of 4.

Sample: 2 of 3

Score: 4

In part (a), the null and alternative hypotheses are stated using unlabeled parameters p_1 and p_2 (step 1). The two-sample z -test is indicated and sample size conditions are checked (step 2). This student used the pooled estimate of the variance of a difference in two proportions in the z -test, as provided by most calculators. The test statistic and p -value are correctly evaluated (step 3). The conclusion is linked to the computed p -value and a specific significance level, and it is correctly stated in the context of the difference in the proportions of banded birds on the two islands (step 4). The total number of birds on Island A is correctly estimated in part (b). In part (c), the student lists concerns about the possible effects of banding on the probability of recapture and whether weaker or sick birds might be more easily captured. Although the student neglects to label the parameters used to state the hypotheses in part (a), there is still enough correct work present for this response to earn a score of 4.

Sample: 3 of 3

Score: 3

The boxplots are correctly drawn and labeled in part (a). The response only compares the shapes of the distributions of mandible lengths for modern Thai dogs and golden jackals. It fails to compare either estimates of location or spread of the distributions. The student compares the medians and the interquartile ranges of the data from the two samples. In part (b), the student addresses the need to use the t -distribution to construct the confidence interval because the sample size is moderate. The student indicates that the boxplot suggests that the normality assumption is plausible. The response to part (c) concludes that it would be inappropriate to perform a two-sample t -test because the boxplot indicates that the distribution of golden jackal mandible lengths is highly skewed to the right. This response earns a score of 3.