www.mymathscloud.com

CollegeBoard
Advanced Placement Program

AP® Statistics
2006 Scoring Guidelines

The College Board: Connecting Students to College Success

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 5,000 schools, colleges, universities, and other educational organizations. Each year, the College Board serves seven 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.

© 2006 The College Board. All rights reserved. College Board, AP Central, APCD, Advanced Placement Program, AP, AP Vertical Teams, Pre-AP, SAT, and the acorn logo are registered trademarks of the College Board. Admitted Class Evaluation Service, CollegeEd, connect to college success, MyRoad, SAT Professional Development, SAT Readiness Program, and Setting the Cornerstones are trademarks owned by the College Board. PSAT/NMSQT is a registered trademark of the College Board and National Merit Scholarship Corporation. All other products and services may be trademarks of their respective owners. Permission to use copyrighted College Board materials may be requested online at: www.collegeboard.com/inquiry/cbpermit.html.

Visit the College Board on the Web: www.collegeboard.com.
AP Central is the official online home for the AP Program: apcentral.collegeboard.com.

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 1

Intent of Question

The primary goals of this question are: (1) to assess a student's ability to use simple graphical displays (dotplots in this case) to compare and contrast two distributions; and (2) to evaluate a student's ability to recognize what statistical information is most useful in making different practical decisions.

Solution

Part (a):

Both distributions of distances are roughly symmetric and somewhat mound-shaped. The center of the distances for catapult A (median A = 136 cm) is slightly lower than the center of the distances for catapult B (median B = 138 cm). There is more variability in the distances traveled by the Ping-Pong balls launched with catapult A. There are distances that are extreme enough to be called (potential) outliers in the catapult A distribution, but there are no outliers among the catapult B distances.

Part (b):

Catapult B would be best because the distances vary less about the center of the distribution for catapult B. If catapult B is properly placed, the balls launched will have a higher probability of landing in the narrow (only 5 cm wide) target band.

Part (c):

The catapult should be placed 138 cm from the target line. Since the distribution of distances for catapult B seems to be fairly symmetric and somewhat mound-shaped, the median (138 cm) is a good representation of the center of the distribution. Placing catapult B at this location would have resulted in a high proportion (30/40 = 0.75) of Ping-Pong balls from this sample of launches landing in the target band.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct (E) if the student correctly identifies similarities and differences in center, spread, and shape for the two distributions.

Part (a) is partially correct (P) if the student correctly identifies similarities and differences in two of the three characteristics (center, shape, and spread) for the two distributions.

Part (a) is incorrect (I) if the student correctly identifies no more than one similarity or difference of the three characteristics (center, shape, and spread) for the two distributions.

Notes:

- Correct comments regarding outliers should be viewed as a positive. However, comments about outliers do not count as one of the three required characteristics.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 1 (continued)

- Describing catapult A's distribution as "normal" or "skewed left" or "uniform" is not acceptable for the shape characteristic. Describing either distribution as "approximately normal" is acceptable.

- Giving separate lists of measures of center and/or spread for the two distributions with no linkage between them is not an acceptable discussion of similarities and differences for these characteristics.

Part (b) is essentially correct (E) if catapult B is chosen using a rationale based on the variability in the distances.

Part (b) is partially correct (P) if catapult B is chosen, but the explanation does not refer to the variability in the distances.

Part (b) is incorrect (I) if catapult B is chosen and no explanation is provided OR catapult A is chosen.

Part (c) is essentially correct (E) if:
the catapult is placed at the median (or mean) of the distances traveled by the Ping-Pong balls, and the explanation addresses why the median (or mean) was selected based on a property of the chosen statistic that relates to the context of the problem;
OR
the catapult is placed at a distance of 137.5-139.5 cm from the target line, and the explanation indicates that the chosen distance resulted in a high proportion of the balls in the sample landing in the target band.

Part (c) is partially correct (P) if the catapult is placed at an acceptable distance from the target line, but the explanation is incomplete or incorrect.

Part (c) is incorrect (I) if the catapult is placed less than 137.5 centimeters or more than 139.5 centimeters from the target line.

Notes:
- Simply saying "because it's the median (or mean)" is an incomplete explanation.
- Some students may confuse the 5 cm band as meaning 5 cm on either side of the target line. If the student chooses the median (or mean) and satisfactorily addresses why the median (or mean) was selected OR chooses a value of 137-140 cm and the explanation indicates that the chosen distance resulted in a high proportion of the balls in the sample landing in the target band, score the response as partially correct.
- If a student gives the distance from the catapult to the front or back of the shaded band rather than the distance to the target line, but gives an otherwise correct response, score part (c) as partially correct.
- If a student picks catapult A in part (b) and follows through correctly in part (c), then part (c) should be scored as essentially correct.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 1 (continued)

4 Complete Response

All three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct
OR
One part essentially correct and two parts partially correct
OR
Three parts partially correct

1 Minimal Response

One part essentially correct and either zero or one part partially correct
OR
No parts essentially correct and two parts partially correct

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 2

Intent of Question

The primary goal of this question is to assess a student's ability to identify the estimated regression line and to identify and interpret important statistics from regression output provided by statistical software in the context of a practical problem.

Solution

Part (a):

The regression line is $\hat{y} = -2.679 + 9.5x$, where $\hat{y}$ represents the estimated (or predicted) mean height of the soapsuds and x represents the amount of detergent added to the pan.

Part (b):

The value $s = 1.99821$ mm is the standard deviation of the residuals. This statistic measures a typical amount of variability in the vertical distances from the observed height of the soapsuds to the regression line.

OR

The value $s = 1.99821$ mm is a measure of variation in the height of soapsuds for a given amount of detergent.

Part (c):

The standard error of the estimated slope parameter is 0.7553 mm per gram. Thus, the standard deviation of the estimated slope for predicting the height of soapsuds by using an amount of detergent is estimated to be 0.7553 mm per gram. This value estimates the variability in the sampling distribution of the estimated slope (i.e., how much we would expect sample slopes to vary from experiment to experiment).

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct (E) if the least squares regression line is correctly identified and the variables are correctly defined.

Part (a) is partially correct (P) if:

the least squares regression line is correctly identified and either of the two variables are not correctly defined;

OR

the least squares regression line is not presented using estimated or predicted notation, or $\hat{y}$, AND both variables are correctly defined;

OR

only one of the two values is correctly identified from the table and both variables are correctly defined.

Part (a) is incorrect (I) if the least squares regression line is incorrectly identified or not identified, and the variables are not correctly defined.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 2 (continued)

Notes:

- If y is identified as the height of the soapsuds and x is identified as the amount of detergent, then the student should get credit for defining the variables. However, y must be identified as an estimated height somewhere in the student response in order to get this part essentially correct.
- If $\hat{y}$ (or *estimated height*) is used to specify the regression line and y is identified as the height, the response should be scored as essentially correct. For example, a response of $\text{estimated height} = -2.679 + 9.5\text{amount}$, where y represents the height of suds an x represents the amount of detergent should be scored essentially correct.
- If the regression line is specified using y (or height) instead of $\hat{y}$ (or estimated height), the response is scored as partially correct. For example, a response of $y = -2.679 + 9.5x$ where y = height of suds and x = amount of detergent should be scored as partially correct.
- If the estimates of the intercept and slope are reversed ($\hat{y} = 9.5 - 2.679x$), then the response should be scored as incorrect.

Part (b) is essentially correct (E) if the standard deviation is correctly interpreted in the context of this study.

Part (b) is partially correct (P) if the standard deviation is correctly interpreted in general terms without context.

Part (b) is incorrect (I) if the response indicates that s is any other standard deviation (e.g., univariate distribution of y), or slope.

Notes:

- If s is interpreted as the estimated standard deviation of the differences between the observed values for the height of soapsuds and the values predicted from the regression line, the response should be scored essentially correct.
- If s is interpreted as a "typical" prediction error for estimating height from the amount of detergent, then the response should be scored essentially correct.

Part (c) is essentially correct (E) if the standard error is identified and interpreted correctly.

Part (c) is partially correct (P) if standard error is identified but interpretation is weak (e.g., the standard error is a standard deviation of the slope). The major idea of sampling variability is not included.

Part (c) is incorrect (I) if the standard error is not correctly identified, identified with no interpretation, or an incorrect interpretation is provided.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 2 (continued)

4 Complete Response

All three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct and either zero or one part partially correct

OR

No parts essentially correct and two parts partially correct

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question are to assess a student's ability to: (1) recognize the random variable of interest, identify its probability distribution, and calculate a probability for a linear combination of a normal random variable and a constant; (2) use basic probability rules to find a different probability; and (3) use the sampling distribution of the sample mean to find a probability about the mean of three observations.

Solution

Part (a):

Since $M = D + E$ (a normal random variable plus a constant is a normal random variable), we know that M is normally distributed with a mean of 2 feet and a standard deviation of 1.5 feet. Thus,

$P(M < 0) = P\left(Z < \frac{0-2}{1.5}\right) < P(Z < -1.33) = 0.0918$, where $Z = \frac{M-\mu}{\sigma}$.

Part (b):

$$\begin{aligned} P(\text{at least one measurement} < 0) &= 1 - P(\text{all three measurements} \geq 0) \\ &= 1 - (1 - 0.0918)^3 \\ &= 1 - (0.9082)^3 \\ &= 1 - 0.7491 \\ &= 0.2509 \end{aligned}$$

Part (c):

Let $\bar{X}$ denote the mean of three independent depth measurements taken at a point where the true depth is 2 feet. Since each measurement comes from a normal distribution, the distribution of $\bar{X}$ is normal with a mean of 2 feet and a standard deviation of $\frac{1.5}{\sqrt{3}} = 0.8660$ feet. Thus,

$P(\bar{X} < 0) = P\left(Z < \frac{0-2}{\frac{1.5}{\sqrt{3}}}\right) < P(Z < -2.31) = 0.0104$, where $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 3 (continued)

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct (E) if the student clearly does ALL three of the following:
- identifies the distribution as normal;
- specifies BOTH μ and σ; AND
- calculates the correct probability.

Part (a) is partially correct (P) if the student:
- calculates the correct probability but fails to identify the distribution as normal with BOTH μ and σ specified;
  OR
- correctly identifies the distribution as normal with BOTH μ and σ specified but fails to calculate the correct probability.

Part (a) is incorrect (I) if any of the following occur:
- the student indicates the probability is 0.5 because the random error is symmetric about zero;
  OR
- the student uses a mean of zero and a standard deviation of 1;
  OR
- the student conducts a hypothesis test.

Notes:
- The student may use the distribution of the error, E, to solve the problem. That is, finding the area below -2 for a normal distribution with mean 0 and standard deviation 1.5 should be scored essentially correct (E).
  Thus $P(E < -2) = P\left(Z < \frac{-2-0}{1.5}\right) < P(Z < -1.33) = 0.0918$, where $Z = \frac{E-\mu}{\sigma}$.
- If only the calculator command normalcdf $(-\infty, 0, 2, 1.5)$ is provided along with the probability 0.0912, then the response should be scored as partially correct (P).

Part (b) is essentially correct (E) if the student calculates the correct probability AND:
- correctly applies complement and probability rules using the value obtained in part (a);
  OR
- clearly identifies the distribution as binomial AND specifies BOTH n and p using the value obtained in part (a).

Part (b) is partially correct (P) if the student:
- clearly identifies the distribution as binomial AND specifies BOTH n and p, using the value obtained in part (a), but does not calculate the correct probability;
  OR

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 3 (continued)

- clearly identifies the distribution as binomial AND specifies BOTH n and p using a value of p that is unrelated to the value obtained in part (a) and calculates the correct probability based on their value of p;
  OR
- calculates the correct probability using the value obtained in part (a) but fails to correctly identify the distribution as binomial with BOTH n and p specified;
  OR
- recognizes the solution as the sum of the product of the probabilities of successes and failures, using the answer from part (a), but omits only the binomial coefficients.

Part (b) is incorrect (I) if the student calculates $P(\text{at least one measurement} < 0) = 1 - p^3$, where p is the solution to part (a).

Notes:

- The solution using the binomial distribution with $p = 0.0918$ is:

$$\begin{aligned} P(\text{at least one measurement} < 0) &= P(B = 1) + P(B = 2) + P(B = 3) \\ &= \binom{3}{1} 0.0918^1 (1 - 0.0918)^2 + \\ &\quad \binom{3}{2} 0.0918^2 (1 - 0.0918)^1 + \binom{3}{3} 0.0918^3 \\ &= 0.2272 + 0.0230 + 0.0008 \\ &= 0.2510 \end{aligned}$$

- If only the calculator command $1 - \text{binomcdf}(3, 0.0918, 0)$ is provided along with the probability 0.2509, then the response should be scored as partially correct (P).

Part (c) is essentially correct (E) if the student clearly does ALL three of the following:

- identifies the distribution of the sample mean as normal;
- specifies BOTH $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$; AND
- calculates the correct probability.

Part (c) is partially correct (P) if the student:

- calculates the correct probability, but fails to identify the distribution of the sample mean as normal with BOTH $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ specified;
  OR
- correctly identifies the distribution of the sample mean as normal with BOTH $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ specified, but fails to calculate the correct probability.

Part (c) is incorrect (I) if any of the following occur:

- the student uses the same calculation as in part (a);
  OR

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 3 (continued)

- the student uses an incorrect standard deviation (e.g., $\frac{1.5}{\sqrt{2}}$ or $\sqrt{3(1.5)}$;

  OR
- the student conducts a hypothesis test.

Notes:

- An alternate solution using the sum instead of the mean is:
  Let $\bar{X}$ denote the mean of three independent depth measurements taken at a point where the true depth is 2 feet. Since each measurement comes from a normal distribution, the distribution of the sum of the three measurements, $S = (X_1 + X_2 + X_3)$, is normal with a mean $\mu_S = 6$ feet and a standard deviation $\sigma_S = 2.598$ feet $\left(\sigma_S = 3\left(\frac{1.5}{\sqrt{3}}\right)\right.$, often calculated as $\left.\sqrt{(1.5)^2 + (1.5)^2 + (1.5)^2}\right)$.

  Thus $P(S < 0) = P\left(Z < \frac{0 - 6}{2.598}\right) = P(Z < -2.31) = 0.0104$, where $Z = \frac{S - \mu_S}{\sigma_S}$.

- If only the calculator command normalcdf $(-\infty, 0, 2, 0.866)$ is provided along with the probability 0.01046, then the response should be scored as partially correct (P).
- If the student does not consistently specify a correct μ and σ from the same distribution, i.e., for the mean or the sum, the response should be scored at most partially correct (P).

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 3 (continued)

4 Complete Response

All three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct
OR
One part essentially correct and two parts partially correct
OR
Three parts partially correct

1 Minimal Response

One part essentially correct and one part partially correct
OR
One part essentially correct and no parts partially correct
OR
No parts essentially correct and two parts partially correct

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 4

Intent of Question

The primary goals of this question are to evaluate a student's ability to: (1) identify and compute an appropriate confidence interval, after checking the necessary conditions; (2) interpret the interval in the context of the question; and (3) use the confidence interval to conduct an appropriate test of significance.

Solution

Part (a):

Step 1: Identifies the appropriate confidence interval by name or formula and checks appropriate conditions.

Two sample *t* interval for $\mu_A - \mu_S$, the difference in mean waiting times, or

$$(\bar{x}_A - \bar{x}_S) \pm t^*_{df}\sqrt{\frac{s_A^2}{n_A} + \frac{s_S^2}{n_S}}$$ [See the next page for possible values of *df*.]

Conditions: 1. Independent random samples
2. Large samples or normal population distributions

One sample of 150 patients divided into two groups after sampling does not meet the condition of two independent random samples with fixed sample sizes. Nevertheless, it is reasonable to assume that mode of transportation splits the patients into two independent groups. Secondly, use of the two sample t interval is reasonable because each sample size is large (e.g., $n_A = 77 > 30$ and $n_S = 73 > 30$).

Alternatively, we could assume that the waiting times are (at least approximately) normally distributed, but we have no way to check this assumption with the information provided.

Step 2: Correct Mechanics

Degrees of freedom = $\min\{(77-1),(73-1)\} = 72$.

$$(6.04 - 8.30) \pm 2.6459\sqrt{\frac{4.30^2}{77} + \frac{5.16^2}{73}}$$

$$-2.26 \pm 2.6459 \cdot (0.7777)$$

$$-2.26 \pm 2.0577$$

$$(-4.3177, -0.2023)$$

Step 3: Interpretation

Based on this sample, we are 99 percent confident that the true difference in the populations' mean waiting times (ambulance – self) is between –4.3177 minutes and –0.2023 minutes.

Equivalently,

With 99 percent confidence, the true mean wait time for those who arrive by ambulance is shorter than those who are self transported by somewhere between 0.2 and 4.3 minutes.

Part (b):

Since zero is not in the 99 percent confidence interval of plausible values for the true difference in means, we can reject $H_0 : \mu_A - \mu_S = 0$ in favor of the alternative $H_a : \mu_A - \mu_S \neq 0$ at the $\alpha = .01$ significance level.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 4 (continued)

Thus, we have statistically significant evidence that there is a difference in the mean wait times for the two groups.

Scoring

Part (a) is scored according to the number of correct steps. Each step is scored as essentially correct (E) or incorrect (I). Part (b) is scored as essentially correct (E) or incorrect (I).

Step 1: Identification of method and check of conditions.
A score of essentially correct (E) requires each of the following; otherwise, the score is I.

- A correct confidence interval procedure should be named or a correct formula given.
- Normality must be assessed by either checking that *EACH* sample size is sufficiently large (e.g., >30) OR by stating that each population needs to be normally distributed but that the data are not given to check that.
- It must be stated that the two groups are assumed to be independent random samples.

Step 2: Mechanics
An identifiable minor error in Step 2 will not necessarily change a score from essentially correct to incorrect. The following confidence intervals are all scored E.

Solutions to Step 2			
Procedure	**d.f.**	**t***	**Confidence Interval**
Unequal Variances	140.37	2.61140	(–4.2910, –0.2291)
Large samples	∞	2.576	(–4.26, –0.26)
Pooled variance	148	2.6095	(–4.2797, –0.2404)
Conservative Approach	72	2.6459	(–4.3177, –0.2023)
Unequal Variances	100 (table)	2.626	(–4.302, –0.218)
Unequal Variances	1000 (table)	2.581	(–4.267, –0.253)
Conservative Approach	60 (table)	2.660	(–4.329, –0.191)
Conservative Approach	80 (table)	2.639	(–4.312, –0.208)

Step 3: Interpretation of Confidence Interval
For a score of essentially correct (E), the interpretation must be about a difference of population means AND be in context AND mention the 99 percent level of confidence.

Note: If the interpretation is correct and also explains the meaning of "confidence level," then that extra explanation must be correct for a score of E. If it is not, the score is I.
Note: The correct interpretation of the confidence interval for Step 3 of part (a) may be found in part (b).

Part (b) is essentially correct (E) if the student concludes that the mean times differ because zero is not contained in the 99 percent confidence interval. The confidence level or the significance level must be stated and the conclusion must be stated in context.

Part (b) is incorrect (I) if the student only says the mean wait time for patients transported via ambulance is significantly lower without indicating the significance level adjustment needed for a one-sided test.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 4 (continued)

4 Complete Response

All three steps of the confidence interval in part (a) are essentially correct and part (b) is essentially correct.

3 Substantial Response

All three steps of the confidence interval in part (a) are essentially correct and part (b) is incorrect.
OR
Two steps of the confidence interval in part (a) are essentially correct and part (b) is essentially correct.

2 Developing Response

Two steps of the confidence interval in part (a) are essentially correct and part (b) is incorrect.
OR
One step of the confidence interval in part (a) is essentially correct and part (b) is essentially correct.

1 Minimal Response

One step of the confidence interval in part (a) is essentially correct and part (b) is incorrect.
OR
Part (b) is essentially correct.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this question are to evaluate a student's ability to: (1) identify the treatments in a biological experiment; (2) present a completely randomized *design* to address the research question of interest; (3) describe the benefit of limiting sources of variability; and (4) describe the limitations to the scope of inference for the biologist.

Solution

Part (a):

The three different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high) yield a total of $3 \times 2 = 6$ different treatment combinations for this experiment.

Treatment Combination	Nutrient	Salinity Level
1	A	Low
2	A	High
3	B	Low
4	B	High
5	C	Low
6	C	High

Part (b):

Since 10 tiger shrimps have already been randomly placed into each of 12 similar tanks in a controlled environment, we must randomly assign the treatment combinations to the tanks. Each treatment combination will be randomly assigned to 2 of the 12 tanks. One way to do this is to generate a random number for each tank. The treatment combinations are then assigned by sorting the random numbers from smallest to largest.

Treatment Combination	Nutrient	Salinity Level	Tanks with
1	A	Low	Smallest and second smallest random numbers
2	A	High	Third and fourth smallest random numbers
3	B	Low	Fifth and sixth smallest random numbers
4	B	High	Seventh and eighth smallest random numbers
5	C	Low	Ninth and tenth smallest random numbers
6	C	High	Next to largest and largest random numbers

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 5 (continued)

After three weeks the weight gain (after – before) is computed for each tank, and the treatments are compared using appropriate averages.

Part (c):

Using only tiger shrimp will reduce a source of variation in the experimental units, the tanks of shrimp in this experiment. By eliminating this possible source of variation, type of shrimp, we are better able to isolate the variability due to the factors of interest to us (nutrient and salinity level). This will make it easier to identify any treatment effects that may be present.

Part (d):

Using only tiger shrimp will limit the scope of inference for the biologist. Ideally, the biologist would like to identify the treatment combination that leads to the most growth for all shrimp. However, the biologist will only be able to identify the best treatment combination for tiger shrimp because other types of shrimp may respond differently to the treatments.

Scoring

Part (a) is scored as essentially correct (E) or incorrect (I). Parts (b), (c), and (d) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is essentially correct (E) if all six treatments are correctly listed. This may be done in a 2 x 3 table or tree diagram but must clearly indicate the six treatments. A correct but incomplete listing of treatments in part (a) can be recovered in part (b) if the six treatments are clearly stated.

Listing the factors (nutrients A, B, C and salinity high, low) is incorrect and cannot be recovered in part (b).

Part (b) is essentially correct (E) if:

- each treatment combination is randomly assigned to 2 of the 12 tanks
  AND
- a correct procedure for randomization is described (so that two knowledgeable statistics users would use the same method to assign treatments to the tanks).

Part (b) is partially correct (P) if only one of these components is present. For example,

- Each treatment is randomly assigned to 2 of the 12 tanks, but the method of randomization is not fully described (i.e., just say randomly assign each treatment to 2 of the 12 tanks).
  OR
- A correct procedure for randomization of the treatments to the tanks is described, but each treatment does not necessarily appear twice.

Part (b) is incorrect (I) if there is no randomization or randomization of treatments is applied to the shrimps only (not the tanks).

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 5 (continued)

Notes:

- If the randomization has been correctly applied to the tanks, additionally randomizing the shrimps or treatments will be regarded as extraneous.
- Because the stem indicates shrimp growth is to be compared, students are not required to identify a response variable in part (b) as was done in the model solution.

Part (c) is essentially correct (E) if

- the statistical advantage of reduced variability is identified
  AND
- an appropriate explanation that relates reduced variability to increasing the likelihood of determining differences among treatments is clearly provided.

Part (c) is partially correct (P) if only one of the two components is correct.

Part (c) is incorrect (I) if neither of the two components is present.

Notes:

- In this completely randomized design, confounding is not possible. Therefore a reference to confounding or lurking variables always incurs a penalty.

Part (d) is essentially correct (E) if

- the statistical disadvantage of limited scope of inference is identified
  AND
- an explanation that different species of shrimp may respond differently to treatments is provided.

(If the different responses to the treatments by other species of shrimp have been established in part (c), then it need not be repeated in part (d).)

Part (d) is partially correct (P) if only one of the two parts of the essentially correct response is provided.

Part (d) is incorrect (I) if neither of the two parts of the essentially correct response is provided,

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points) use a holistic approach to determine whether to score up or down depending on the strength of the response and communication. The strength of the responses in parts (b) and (c) may be most important in making this choice.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

AP® STATISTICS
2006 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question are to evaluate a student's ability to apply the concepts of significance testing to a new setting, in particular to: (1) state hypotheses for a parameter of interest, given a research question; (2) evaluate a new test statistic and use the probability distribution associated with that statistic to test the hypotheses of interest; (3) identify the values of the test statistic that would lead to rejection of the null hypothesis on a graph; and (4) interpret simulated sampling distributions for different populations.

Solution

Part (a):

Let σ^2 denote the variance in the temperatures measured by the thermostats recently produced by this manufacturer.

$H_0 : \sigma^2 = 1.52(^\circ F)^2$ OR Recently produced thermostats are not more variable than thermostats produced in the past.

$H_a : \sigma^2 > 1.52(^\circ F)^2$ OR Recently produced thermostats are more variable than thermostats produced in the past.

Part (b):

$$\frac{(n-1)s^2}{1.52} = \frac{9 \times (1.4277)^2}{1.52} = \frac{9 \times (2.0383)}{1.52} = 12.069$$

Part (c):

The test statistic has a χ^2 distribution with 9 degrees of freedom under H_0. The chance of exceeding the observed value of 12.069, under H_0, is

$$p\text{–value} = P(\chi_9^2 \geq 12.069) = 0.2094 .$$

(or, from the table, .20 < *p*-value < .25). Since the *p*-value is greater than 0.05, we cannot reject the null hypothesis. That is, we do not have statistically significant evidence that recent thermostats are less reliable (more variable) than in the past.

Part (d):

The smallest value that would have led to the rejection of the null hypothesis is the 95th percentile of the χ^2 distribution with 9 degrees of freedom, which is 16.92. The rejection region contains all values greater than or equal to 16.92. This region should be identified on the graph by indicating the approximate location of 16.92 on the axis and shading the region that is bounded by the vertical line through 16.92, the horizontal axis, and the χ^2 curve.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 6 (continued)

Part (e)

Indicate the region to the right of 16.92 on all three histograms.

Part (f)

The population with the largest variance will tend to produce the largest values of s^2 in the simulation and hence the largest test statistics. Histogram III has the largest probability of producing a sample that would lead to the rejection of H_0 so Histogram III corresponds to the population with the largest variance.

Similarly, the test statistics will tend to be smallest for the population with variance closest to 1.52. Histogram II has the smallest probability of producing a sample that would lead to the rejection of H_0 so Histogram II corresponds to the population with the smallest variance.

Scoring

Each of four components are scored as essentially correct (E), partially correct (P), or incorrect (I).

I. **Parts (a) and (b)** are combined into one component and scored as essentially correct (E) if both part (a) and part (b) are correct.

Parts (a) and (b) are partially correct (P) if one of the two parts is correct.

Notes*:*

1. If a two-sided alternative is used or the hypotheses involve a mean, then part (a) is not correct.
2. Nonstandard notation for the population variance must be defined.
3. If the value of s (or of s^2) is not shown in part (b), then part (b) is incorrect.

II. **Part (c)** is scored as essentially correct (E) if both:

- The p-value is given (or the test statistic compared to the critical value).
- The conclusion is written in context and linked to the p-value.

Part (c) is partially correct (P) if one of the two bulleted items is correct.

Notes*:*

1. Conditions (SRS, normal population) are given in the problem so it is not necessary to restate them. However, if incorrect conditions are given, the first bullet is incorrect.
2. If the null hypothesis is "accepted" or equivalent, the second bullet is incorrect.
3. If both an α and a p-value are given, the linkage is implied. If no α is given, the solution must be explicit about the linkage by giving a correct interpretation of the p-value or explaining how the conclusion follows from the p-value.

III. **Parts (d) and (e)** are combined into one component and scored as essentially correct (E) if both:

- The critical value is identified as 16.92.
- The region to the right of a cut-off point of between 15 and 20 is identified in part (d). AND the same region is identified in each of the three histograms in part (e).

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).

www.mymathscloud.com

Question 6 (continued)

Parts (d) and (e) are partially correct (P) if one of the two bulleted items above is correct.

IV. **Part (f)** is essentially correct (E) if both:

- Histograms III and II are identified as the simulated sampling distributions for the populations with the largest and smallest variances, respectively.
- The justification refers to fact that Histogram III came from the population with the largest variance because the identified region is largest, and so it will be more likely to reject the null hypothesis. (Similarly for the smallest variance.)

Part (f) is partially correct (P) if both:

- Histograms III and II are identified as the simulated sampling distributions for the populations with the largest and smallest variances, respectively.
- The justification says only that Histogram III represents the population with the largest variance because the identified region is largest. OR The justification refers to the fact that the simulated sampling distribution for the population with the largest variance should result in sample variances—and hence test statistics—that are centered at the largest values. (Similarly for the smallest variance.) OR The justification refers to the fact that the simulated sampling distribution for the population with the largest variance should result in sample variances—and hence test statistics—that are more variable and Histogram III has the more variable values of the test statistic. (Similarly for the smallest variance.)

Part (f) is incorrect (I) if

- Histograms III and II are identified as the simulated sampling distributions for the populations with the largest and smallest variances, respectively, but the justification refers only to the fact that these histograms themselves have the largest and smallest variability.

Note*:*

1. If only one of Histogram III or Histogram II is identified and correctly justified, the response is scored partially correct.

For each of the four components,

Essentially Correct (E) = 1
Partially Correct (P) = 1/2
Incorrect (I) = 0

4 Complete Response

3 Substantial Response

2 Developing Response

1 Minimal Response

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.

© 2006 The College Board. All rights reserved.
Visit apcentral.collegeboard.com (for AP professionals) and www.collegeboard.com/apstudents (for students and parents).