www.mymathscloud.com

CollegeBoard
Advanced Placement Program

AP® Statistics

Practice Exam

The questions contained in this AP® Statistics Practice Exam are written to the content specifications of AP Exams for this subject. Taking this practice exam should provide students with an idea of their general areas of strengths and weaknesses in preparing for the actual AP Exam. Because this AP Statistics Practice Exam has never been administered as an operational AP Exam, statistical data are not available for calculating potential raw scores or conversions into AP grades.

This AP Statistics Practice Exam is provided by the College Board for AP Exam preparation. Teachers are permitted to download the materials and make copies to use with their students in a classroom setting only. To maintain the security of this exam, teachers should collect all materials after their administration and keep them in a secure location. Teachers may not redistribute the files electronically for any reason.

© 2008 The College Board. All rights reserved. College Board, Advanced Placement Program, AP, AP Central, SAT, and the acorn logo are registered trademarks of 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. Visit the College Board on the Web: www.collegeboard.com.

www.mymathscloud.com

Contents

Directions for Administration ii
Section I: Multiple-Choice Questions 1
Section II: Free-Response Questions 46
Student Answer Sheet for Multiple-Choice Section 64
Multiple-Choice Answer Key 65
Free-Response Scoring Guidelines 66

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.

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
Directions for Administration

The AP Statistics Exam is three hours in length and consists of a multiple-choice section and a free-response section.

- The 90-minute multiple-choice section contains 40 questions and accounts for 50 percent of the final grade.
- The 90-minute free-response section contains 6 questions and accounts for 50 percent of the final grade.

A 10-minute break should be provided after Section I is completed.

The actual AP Exam is administered in one session. Students will have the most realistic experience if a complete morning or afternoon is available to administer this practice exam. If a schedule does not permit one time period for the entire practice exam administration, it would be acceptable to administer Section I one day and Section II on a subsequent day.

Total scores on the multiple-choice section are based only on the number of questions answered correctly. No points are deducted for incorrect answers and no points are awarded for unanswered questions.

- The use of approved calculators is permitted during the exam. More information on calculators that are approved for use on this exam can be found on AP Central.
- It is suggested that the practice exam be completed using a pencil to simulate an actual administration.
- Teachers will need to provide paper for the students to write their free-response answers. Teachers should provide directions to the students indicating how they wish the responses to be labeled so the teacher will be able to associate the student's response with the question the student intended to answer.
- Remember that students are not allowed to remove any materials, including scratch work, from the testing site.

www.mymathscloud.com

Section I

Multiple-Choice Questions

www.mymathscloud.com

Formulas begin on page 3.
Questions begin on page 6.
Tables begin on page 42.

www.mymathscloud.com

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1}\sum \left(x_i - \bar{x}\right)^2}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum \left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$r = \frac{1}{n-1}\sum \left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\sum \left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\sum \left(x_i - \bar{x}\right)^2}}$$

www.mymathscloud.com

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$E(X) = \mu_x = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_x = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If $\bar{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

www.mymathscloud.com

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\bullet$ (standard deviation of statistic)

Single-Sample

Statistic	**Standard Deviation of Statistic**
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

Statistic	**Standard Deviation of Statistic**
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\ \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

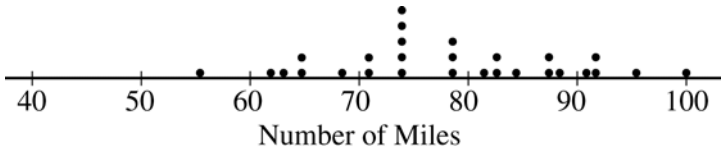
Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

www.mymathscloud.com

STATISTICS
SECTION I
Time—1 hour and 30 minutes
Number of questions—40
Percent of total grade—50

Directions: Solve each of the following problems, using the available space for scratch work. Select the answer choice that is best in each case and place the letter of your choice in the corresponding box on the student answer sheet. Do not spend too much time on any one problem.

1. The dotplot below displays the total number of miles that the 28 residents of one street in a certain community traveled to work in one five-day workweek.



Which of the following is closest to the percentile rank of a resident from this street who traveled 85 miles to work that week?

(A) 60

(B) 70

(C) 75

(D) 80

(E) 85

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

2. A large simple random sample of people aged nineteen to thirty living in the state of Colorado was surveyed to determine which of two MP3 players just developed by a new company was preferred. To which of the following populations can the results of this survey be safely generalized?

   (A) Only people aged nineteen to thirty living in the state of Colorado who were in this survey

   (B) Only people aged nineteen to thirty living in the state of Colorado

   (C) All people living in the state of Colorado

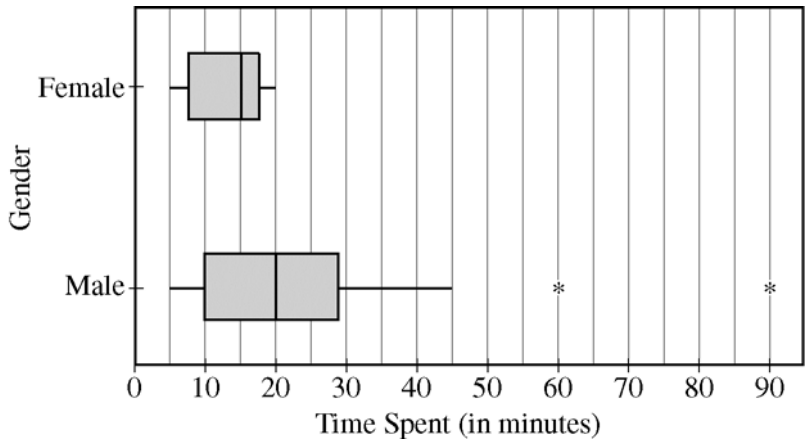
   (D) Only people aged nineteen to thirty living in the United States

   (E) All people living in the United States

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

3. Janelle collected data on the amount of time in minutes each person in a large sample of customers spent in a local store. The data also included recording the gender of each customer. These data were used to generate the boxplots shown below.



Which of the following statements is true?

(A) The range in the amount of time in minutes males in the sample of customers spent in the store is approximately 40 minutes.

(B) The mean amount of time in minutes males in the sample of customers spent in the store is approximately 20 minutes.

(C) The third quartile of the amount of time in minutes males in the sample of customers spent in the store is approximately 45 minutes.

(D) The interquartile range of the amount of time in minutes females in the sample of customers spent in the store is 15 minutes.

(E) Approximately half of the males in the sample of customers spent at least as much time in the store as any female in the sample of customers.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

4. The probability of obtaining a head when a certain coin is flipped is about 0.65. Which of the following is clo to the probability that heads would be obtained 15 or fewer times when this coin is flipped 25 times?

   (A) 0.14

   (B) 0.37

   (C) 0.39

   (D) 0.60

   (E) 0.65

---

5. A study was conducted using data collected on the birth weights of a random sample of 10 pairs of identical twins to determine whether the twin born first tends to weigh more than the twin born second. Let μ_F represent the average birth weight of all twins born first, μ_S represent the average birth weight of all twins born second, and μ_D represent the average difference in birth weight (weight of first minus weight of second) for all pairs of twins. Which of the following would be the null and alternative hypotheses for this study?

   (A) $H_0 : \mu_F = \mu_S$ and $H_a : \mu_F < \mu_S$

   (B) $H_0 : \mu_F = \mu_S$ and $H_a : \mu_F \neq \mu_S$

   (C) $H_0 : \mu_D = 0$ and $H_a : \mu_D > 0$

   (D) $H_0 : \mu_F - \mu_S = \mu_D$ and $H_a : \mu_F - \mu_S > \mu_D$

   (E) $H_0 : \mu_F - \mu_S = \mu_D$ and $H_a : \mu_F - \mu_S \neq \mu_D$

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

6. The buyer for an electronics store wants to estimate the proportion of defective wireless game controllers in a shipment of 5,000 controllers from the store's primary supplier. The shipment consists of 200 boxes each containing 25 controllers. The buyer numbers the boxes from 1 to 200 and randomly selects six numbers in that range. She then opens the six boxes with the corresponding numbers, examines all 25 controllers in each of these boxes, and determines the proportion of the 150 controllers that are defective. What type of sample is this?

   (A) Biased random sample

   (B) Nonrandom sample

   (C) Simple random sample

   (D) Stratified random sample

   (E) Cluster random sample

---

7. A distribution of test scores is not symmetric. Which of the following is the best estimate of the z-score of the third quartile?

   (A) 0.67

   (B) 0.75

   (C) 1.00

   (D) 1.41

   (E) This z-score cannot be estimated from the information given.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

8. Sophomore, junior, and senior students at a high school will be surveyed regarding a potential increase in the extracurricular student activities fee. There are three possible responses to the survey question—agree with the increase, do not agree with the increase, or no opinion. A chi-square test will be conducted to determine whether the response to this question is independent of the class in which the student is a member. How many degrees of freedom should the chi-square test have?

   (A) 9

   (B) 6

   (C) 4

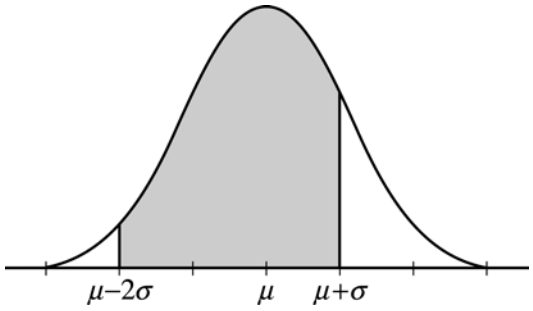
   (D) 2

   (E) 1

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

9. A certain type of remote-control car has a fully charged battery at the time of purchase. The distribution of running times of cars of this type, before they require recharging of the battery for the first time after its period of initial use, is approximately normal with a mean of 80 minutes and a standard deviation of 2.5 minutes. The shaded area in the figure below represents which of the following probabilities?



(A) The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 75 minutes and 82.5 minutes.

(B) The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 75 minutes and 85 minutes.

(C) The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 77.5 minutes and 82.5 minutes.

(D) The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 77.5 minutes and 85 minutes.

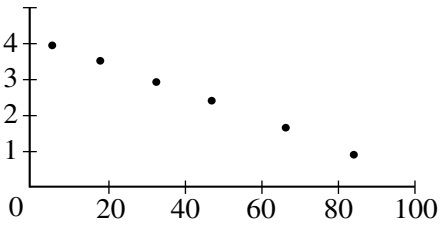
(E) The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 77.5 minutes and 87.5 minutes.

GO ON TO THE NEXT PAGE.

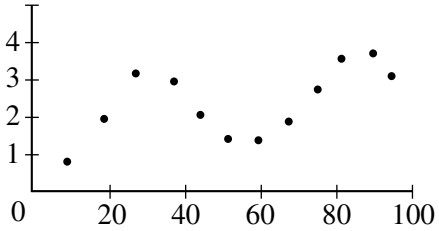
www.mymathscloud.com

10. Which of the following scatterplots could represent a data set with a correlation coefficient of $r = -1$?

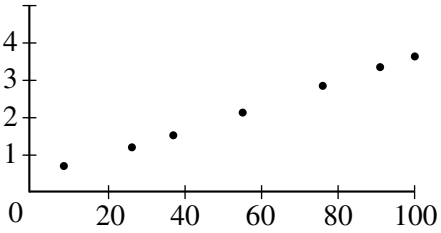
(A)



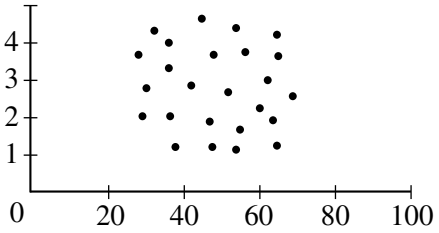
(B)



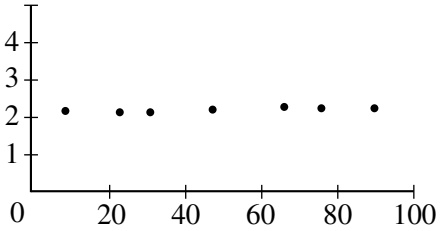
(C)



(D)



(E)



GO ON TO THE NEXT PAGE.

www.mymathscloud.com

11. Publishers of a magazine wish to determine what proportion of the magazine's 50,000 subscribers are pleased with their subscription. The publishers intend to mail a survey to 1,000 subscribers randomly selected from those who have received the magazine for 5 years or more. This introduces selection bias, since long-subscribing customers are more likely to be pleased with their subscription. Which of the following would best eliminate selection bias?

(A) Mail surveys to 2,000 subscribers randomly selected from those who have received the magazine for 5 years or more.

(B) Mail surveys to 1,000 subscribers randomly selected from those who have received the magazine for 1 year or less.

(C) Mail surveys to 1,000 subscribers randomly selected from all subscribers.

(D) Mail surveys to 1,000 subscribers randomly selected from those who have received the magazine for 5 years or more and 1,000 subscribers randomly selected from those who have received the magazine for 1 year or less.

(E) Mail surveys to 500 subscribers randomly selected from a group who have received a free six-month subscription within the past year.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

12. An environmental scientist wants to test the null hypothesis that an antipollution device for cars is not effectiv
Under which of the following conditions would a Type I error be committed?

(A) The scientist concludes that the antipollution device is effective when it actually is not.

(B) The scientist concludes that the antipollution device is not effective when it actually is.

(C) The scientist concludes that the antipollution device is effective when it actually is.

(D) The scientist concludes that the antipollution device is not effective when it actually is not.

(E) A Type I error cannot be committed in this situation.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

13. A data set of test scores is being transformed by applying the following rule to each of the raw scores.

Transformed score = 3.5(raw score) + 6.2

Which of the following is NOT true?

(A) The mean transformed score equals 3.5(the mean raw score) + 6.2.

(B) The median transformed score equals 3.5(the median raw score) + 6.2.

(C) The range of the transformed scores equals 3.5(the range of the raw scores) + 6.2.

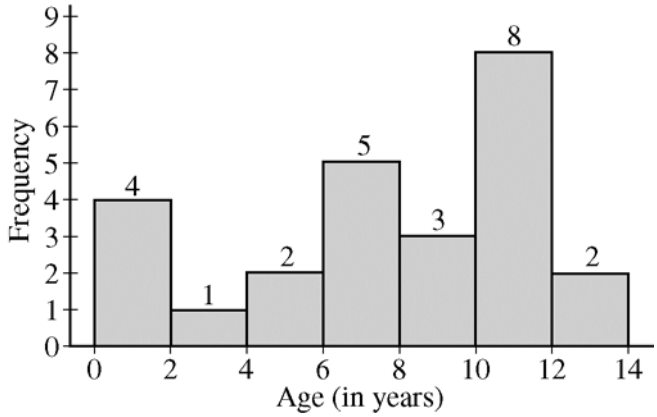
(D) The standard deviation of the transformed scores equals 3.5(the standard deviation of the raw scores).

(E) The IQR of the transformed scores equals 3.5(the IQR of the raw scores).

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

14. A random sample of 25 households from the Mountainview School District was surveyed. In this survey, data were collected on the age of the youngest child living in each household. The histogram below displays the data collected in the survey.



In which of the following intervals is the median of these data located?

(A) 0 years old to less than 2 years old

(B) 4 years old to less than 6 years old

(C) 6 years old to less than 8 years old

(D) 8 years old to less than 10 years old

(E) 10 years old to less than 12 years old

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

15. Based on a random sample of 50 students, the 90 percent confidence interval for the mean amount of money students spend on lunch at a certain high school is found to be ($3.45, $4.15). Which of the following statements is true?

(A) 90% of the time, the mean amount of money that all students spend on lunch at this high school will be between $3.45 and $4.15.

(B) 90% of all students spend between $3.45 and $4.15 on lunch at this high school.

(C) 90% of all random samples of 50 students obtained at this high school would result in a sample mean amount of money students spend on lunch between $3.45 and $4.15.

(D) 90% of all random samples of 50 students obtained at this high school would result in a 90% confidence interval that contains the true mean amount of money students spend on lunch.

(E) Approximately 45 of the 50 students in the random sample will spend between $3.45 and $4.15 on lunch at this high school.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

16. Ten percent of all Dynamite Mints candies are orange and 45 percent of all Holiday Mints candies are orange. Two independent random samples, each of size 25, are selected—one from Dynamite Mints candies and the other from Holiday Mints candies. The total number of orange candies in the two samples is observed. What are the expected total number of orange candies and the standard deviation for the total number of orange candies, respectively, in the two samples?

(A) 7 and 2.905

(B) 7 and 3.987

(C) 13.75 and 2.233

(D) 13.75 and 2.905

(E) 13.75 and 3.987

---

17. In a random sample of 60 shoppers chosen from the shoppers at a large suburban mall, 36 indicated that they had been to a movie in the past month. In an independent random sample of 50 shoppers chosen from the shoppers in a large downtown shopping area, 31 indicated that they had been to a movie in the past month. What significance test should be used to determine whether these data provide sufficient evidence to reject the hypothesis that the proportion of shoppers at the suburban mall who had been to a movie in the past month is the same as the proportion of shoppers in the large downtown shopping area who had been to a movie in the past month?

(A) One proportion z-test

(B) One sample z-test for a mean

(C) One sample t-test for a mean

(D) Two proportion z-test

(E) Two sample t-test for means

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

Questions 18-19 refer to the following scenario and set of data.

In the 1830s, land surveyors began to survey the land acquired in the Louisiana Purchase. Part of their task was to note the sizes of trees they encountered in their surveying. The table of data below is for bur oak trees measured during the survey.

Tree Diameter (in inches)	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
4	2	0.005	2	0.005
5	5	0.012	7	0.016
6	19	0.044	26	0.060
7	5	0.012	31	0.072
8	41	0.095	72	0.167
9	12	0.028	84	0.195
10	53	0.123	137	0.318
11	3	0.007	140	0.325
12	64	0.148	204	0.473
13	4	0.009	208	0.483
14	46	0.107	254	0.589
15	11	0.026	265	0.615
16	48	0.111	313	0.726
17	2	0.005	315	0.731
18	47	0.109	362	0.840
19				
20	30	0.070	392	0.910
21				
22	4	0.009	396	0.919
23				
24	18	0.042	414	0.961
25				
26	1	0.002	415	0.963
27				
28	2	0.005	417	0.968
29				
30	8	0.019	425	0.986
31				
32				
33				
34	1	0.002	426	0.988
35	1	0.002	427	0.991
36	3	0.007	430	0.998
37				
38	1	0.002	431	1.000

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

18. Which of the following differences in cumulative relative frequencies gives the proportion of trees that are 12 inches to 16 inches, inclusive, in diameter?

(A) 0.615 – 0.325

(B) 0.615 – 0.473

(C) 0.726 – 0.325

(D) 0.726 – 0.473

(E) 0.731 – 0.325

---

19. An outlier may be defined as a data point that is more than 1.5 times the interquartile range below the lower quartile or is more than 1.5 times the interquartile range above the upper quartile. According to this definition, what is the diameter, in inches, of the smallest tree that is an outlier?

(A) 4
(B) 28
(C) 30
(D) 34
(E) 36

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

20. The National Honor Society at Central High School plans to sample a random group of 100 seniors from all h[illegible] schools in the state in which Central High School is located to determine the average number of hours per week spent on homework. A 95 percent confidence interval for the mean number of hours spent on homework will then be constructed using the sample data. Before selecting the sample, the National Honor Society decides that it wants to decrease the margin of error. Which of the following is the best way to decrease the margin of error?

(A) Increase the confidence level to 99%

(B) Use the population standard deviation

(C) Use the sample standard deviation

(D) Increase the sample size

(E) Decrease the sample size

---

21. Ninety percent of the people who have a particular disease will have a positive result on a given diagnostic test. Ninety percent of the people who do not have the disease will have a negative result on this test. If 5 percent of a certain population has the disease, what percent of that population would test positive for the disease?

(A) 4.5%
(B) 5%
(C) 10%
(D) 14%
(E) 90%

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

22. A recent survey concluded that the proportion of American teenagers who have a cell phone is 0.27. The true population proportion of American teenagers who have a cell phone is 0.29. For samples of size 1,000 that are selected at random from this population, what are the mean and standard deviation, respectively, for the sampling distribution of the sample proportion of American teenagers who have a cell phone?

(A) $0.27, \sqrt{1000(0.27)(0.73)}$

(B) $0.27, \sqrt{\dfrac{(0.29)(0.71)}{1000}}$

(C) $0.27, \sqrt{\dfrac{(0.27)(0.73)}{1000}}$

(D) $0.29, \sqrt{\dfrac{(0.29)(0.71)}{1000}}$

(E) $0.29, \sqrt{1000(0.29)(0.71)}$

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

23. A marketing company wants to estimate the proportion of consumers in a certain region of the country who would react favorably to a new marketing campaign. Further, the company wants the estimate to have a margin of error of no more than 5 percent with 90 percent confidence. Of the following, which is closest to the minimum number of consumers needed to obtain the estimate with the desired precision?

(A) 136

(B) 271

(C) 385

(D) 542

(E) 769

---

24. Traffic data revealed that 35 percent of automobiles traveling along a portion of an interstate highway were exceeding the legal speed limit. Using highway cameras and license plate registrations, it was also determined that 52 percent of sports cars were also speeding along the same portion of the highway. What is the probability that a randomly selected car along the same portion of the highway was a speeding sports car?

(A) 0.870

(B) 0.673

(C) 0.182

(D) 0.170

(E) It cannot be determined from the information given.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

25. In a large school district, 16 of 85 randomly selected high school seniors play a varsity sport. In the same district, 19 of 67 randomly selected high school juniors play a varsity sport. A 95 percent confidence interval for the difference between the proportion of high school seniors who play a varsity sport in the school district and high school juniors who play a varsity sport in the school district is to be calculated. What is the standard error of the difference?

(A) 0.0347

(B) 0.0695

(C) 0.1362

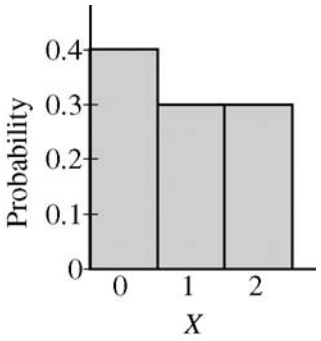
(D) 0.9800

(E) 1.6900

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

26. The number of points, X, scored in a game has the probability distribution below.



The number of points obtained in one game is independent of the number of points obtained in a second game. When the game is played twice, the sum of the number of points for both times could be 0, 1, 2, 3, or 4. If Y represents the sampling distribution of the sum of the scores when the game is played twice, for which value of Y will the probability be greatest?

(A) 0

(B) 1

(C) 2

(D) 3

(E) 4

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

27. Each person in a random sample of adults indicated his or her favorite color. The results, shown in the table below, are reported by age group of the respondents.

	Red	Green	Blue	Other	Total
Under 30	20	42	16	36	114
30-50	24	35	24	25	108
Over 50	25	22	35	10	92
Total	69	99	75	71	314

If choice of color is independent of age group, which of the following expressions is equal to the expected number of respondents who are aged 30 to 50, inclusive, and prefer green?

(A) $\frac{(99)(108)}{314}$

(B) $\frac{(69)(108)}{314}$

(C) $\frac{(35)(99)}{108}$

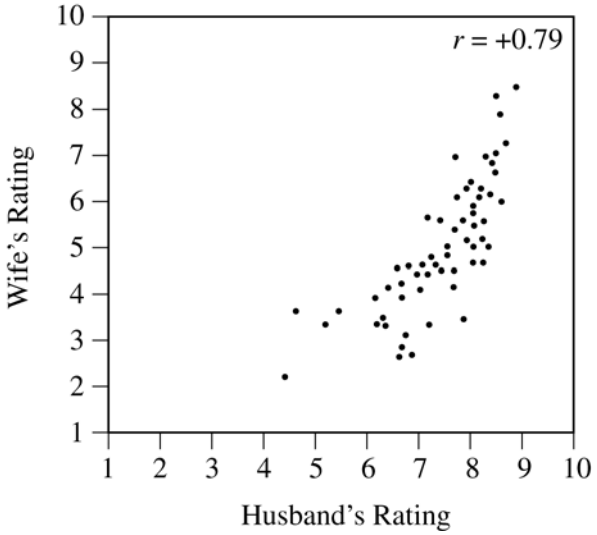
(D) $\frac{(35)(108)}{314}$

(E) $\frac{(35)(99)}{314}$

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

28. In a recent survey, 60 randomly selected married couples from the same town were asked to rate the overall quality of living in their town on a scale from 1 (very poor) to 10 (excellent) on twenty different attributes such as accessibility to major highways, availability of entertainment, services provided by tax dollars, etc. For each couple, the husband's individual ratings on the twenty attributes were averaged to produce an overall quality rating, and that process was repeated for the wife. Each point on the scatterplot below displays the overall rating of one of the 60 couples with the husband's rating represented by the horizontal axis and the wife's rating represented by the vertical axis.



Based on the scatterplot, which of the following statements is true?

(A) Husbands tended to rate the quality of living higher than their wives did.

(B) More overall ratings of 7 or less were assigned by husbands than by wives.

(C) The range in the husbands' overall ratings is greater than the range in the wives' overall ratings.

(D) The difference in overall ratings between a husband and wife was not more than 3 for any couple.

(E) For each couple, the overall rating assigned by the husband was the same as the overall rating assigned by the wife.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

29. A 90 percent confidence interval is to be created to estimate the proportion of television viewers in a certain area who favor moving the broadcast of the late weeknight news to an hour earlier than it is currently. Initially, the confidence interval will be created using a simple random sample of 9,000 viewers in the area. Assuming that the sample proportion does not change, what would be the relationship between the width of the original confidence interval and the width of a second 90 percent confidence interval that is created based on a sample of only 1,000 viewers in the area?

(A) The second confidence interval would be 9 times as wide as the original confidence interval.

(B) The second confidence interval would be 3 times as wide as the original confidence interval.

(C) The width of the second confidence interval would be equal to the width of the original confidence interval.

(D) The second confidence interval would be $\frac{1}{3}$ as wide as the original confidence interval.

(E) The second confidence interval would be $\frac{1}{9}$ as wide as the original confidence interval.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

30. A new restaurant is interested in determining the best time-temperature combination for roasting a five-pound cut of lamb. The times to be tested are 45 minutes, 60 minutes, and 90 minutes at temperatures of 350 degrees Fahrenheit and 425 degrees Fahrenheit for each time, with the exception of the 90 minute–425 degree combination. That combination is being eliminated because it will overcook the lamb, which leaves five combinations remaining. From 10 identical cuts of lamb, 2 are randomly selected to roast using each of the time-temperature combinations in the same oven. The quality of the finished product is evaluated for each roast. Which of the following is true?

(A) The explanatory variable is the quality of the finished lamb.

(B) The response variable is the roasting temperature for the lamb.

(C) If the experiment is repeated, identical results will be expected.

(D) There should be a control group (i.e., a group in which no treatment is given).

(E) The two cuts that are being roasted for each time-temperature combination are an example of replication.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

31. A local real estate magazine used the median instead of the mean when it reported the SAT score of the average student who attends Groveland High School. A graphical display of SAT scores of students who attend Groveland High School indicated that the data were strongly skewed to the right. Which of the following explains why, in this situation, the median is a more accurate indicator of the SAT score of the average student than the mean is?

(A) The mean is affected by the skewness, whereas the median is not.

(B) The median is always the preferred statistic.

(C) The mean will be less than the median when the data are strongly skewed to the right.

(D) The mean should be used only when data are strongly skewed to the left.

(E) The median is equal to one-half the sum of the maximum and minimum SAT scores at Groveland High.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

32. The director of a fitness center wants to examine the effects of two exercise classes (spinning and aerobics) on body fat percentage. A six-week spinning class and a six-week aerobics class are offered at the same time and on the same days, so that a person can enroll in only one of them. A new class of each is about to begin, and each class has 25 people in it. Ten people are randomly selected from each class. Each person's body fat percentage is measured at the beginning and again at the end of the six-week class. Using the change in body fat percentage as the response variable and conducting a test at the $\alpha = 0.01$ level, the director determines that there is a significant difference between the treatment means. Which of the following is a confounding variable in the study?

(A) The director's choice of spinning and aerobics classes as the types for use in the study

(B) The random sample of 10 people from each class

(C) The participants' choice of which class to take

(D) The use of body fat percentage as the measure of effectiveness of the treatment

(E) The fact that both classes were conducted three times each week at the same time of day and for the same amount of time each day

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

33. Raoul performed an experiment using 16 windup rubber band single-propeller airplanes. He wound up the propeller a different number of times and recorded the amount of time (in seconds) that the airplane flew for each number of rotations that the propeller was wound. A regression analysis was performed and the partial computer output is given below.

```
The regression equation is

Time = 0.924 + 0.0462 Rotations

Predictor         Coef     SE Coef           T         P
Constant        0.9241      0.6413        1.44     0.172
Rotation       0.04625     0.01565        2.96     0.010

S = 0.5426      R-Sq = 38.4%     R-Sq(adj) = 34.0%
```

Which of the following is a 95 percent confidence interval for the slope of the regression line that relates the number of rotations the rubber band is wound and the plane's flight time?

(A) $0.0462 \pm (2.145)(0.01565)$

(B) $0.924 \pm (2.145)(0.6413)$

(C) $0.0462 \pm (2.96)(0.01565)$

(D) $0.924 \pm (2.96)(0.6413)$

(E) $0.0462 \pm (2.131)(0.01565)$

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

34. A recent study was conducted to investigate the duration of time required to complete a certain manual dexter task. The reported mean was 10.2 seconds with a standard deviation of 16.0 seconds. Suppose the reported values are the true mean and standard deviation for the population of subjects in the study. If a random sample of 144 subjects is selected from the population, what is the approximate probability that the mean of the sample will be more than 11.0 seconds?

(A) 0.1151

(B) 0.2743

(C) 0.7257

(D) 0.8849

(E) Based on the values of the true mean and true standard deviation, it can be concluded that the population distribution is not normal and therefore the probability cannot be calculated.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

35. To determine whether employees at Site X have higher salaries, on average, than employees at Site Y of the same company do, independent random samples of salaries were obtained for the two groups. The data are summarized below.

	Site X	Site Y
Mean	\$61,234	\$60,529
Standard Deviation	\$4,352	\$3,456
n	235	183

Based on the data, which of the following statements is true?

(A) At the 5% significance level, employees at Site Y have a significantly higher mean salary than employees at Site X do.

(B) At the 1% significance level, employees at Site Y have a significantly higher mean salary than employees at Site X do.

(C) At the 5% significance level, employees at Site X have a significantly higher mean salary than employees at Site Y do.

(D) At the 1% significance level, employees at Site X have a significantly higher mean salary than employees at Site Y do.

(E) At the 10% significance level, there is no significant difference in salaries between the employees at the two sites.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

36. The director of a technical school was curious about whether there is a relationship between students who complete one of the school's most popular health sciences certificate programs and whether those students go on to complete more advanced studies in the health sciences within two years of completing the certificate program. She randomly selected 100 students who completed the program. Data collected on these students are shown in the table below.

		Completed More Advanced Studies		
		Yes	No	Total
Completed Most Popular Health Sciences Certificate Program	Yes	35	25	60
	No	5	35	40
	Total	40	60	100

Which of the following statements is true for these 100 students?

(A) Being a person who completed more advanced studies is more likely than being a person who did not complete more advanced studies.

(B) Being a person who completed the program is less likely than being a person who did not complete the program.

(C) Being a person who completed the program and completed more advanced studies is less likely than being a person who did not complete the program and did not complete more advanced studies.

(D) Being a person who did not complete the program but completed more advanced studies is less likely than being a person who completed the program and completed more advanced studies.

(E) Being a person who completed the program but did not complete more advanced studies is more likely than being a person who did not complete the program and did not complete more advanced studies.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

37. On their birthdays, employees at a large company are permitted to take a 60-minute lunch break instead of the usual 30 minutes. Data were obtained from 10 randomly selected company employees on the amount of time that each actually took for lunch on his or her birthday. The company wishes to investigate whether these data provide convincing evidence that the mean time is greater than 60 minutes. Of the following, which information would NOT be expected to be a part of the process of correctly conducting a hypothesis test to investigate the question, at the 0.05 level of significance?

(A) Being willing to assume that the distribution of actual birthday lunch times for all employees at the company is approximately normal

(B) Knowing that there are no outliers in the data as indicated by the normal probability plot and boxplot

(C) Using a t-statistic to carry out the test

(D) Using 9 for the number of degrees of freedom

(E) Given that the p-value is greater than 0.05, rejecting the null hypothesis and concluding that the mean time was not greater than 60 minutes

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

38. A dog food company wishes to test a new high-protein formula for puppy food to determine whether it promo faster weight gain than the existing formula for that puppy food. Puppies participating in an experiment will be weighed at weaning (when they begin to eat puppy food) and will be weighed at one-month intervals for one year. In designing this experiment, the investigators wish to reduce the variability due to natural differences in puppy growth rates. Which of the following strategies is most appropriate for accomplishing this?

(A) Block on dog breed and randomly assign puppies to existing and new formula groups within each breed.

(B) Block on geographic location and randomly assign puppies to existing and new formula groups within each geographic area.

(C) Stratify on dog breed and randomly sample puppies within each breed. Then assign puppies by breed to either the existing or the new formula.

(D) Stratify on geographic location of the puppies and randomly sample puppies within each geographic area. Then assign puppies by geographic area to either the existing or the new formula.

(E) Stratify on gender and randomly sample puppies within gender groups. Then assign puppies by gender to either the existing or the new formula.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

39. A statistics student wants to compare the mean times needed to access flight information for two major airlines. Twenty randomly selected students accessed one airline's Web site, and the time required to locate the flight information using the Web site had a mean of 2.5 minutes and a standard deviation of 0.8 minute. Twenty different randomly selected students accessed the other airline's Web site, and the time required to locate the flight information using the Web site had a mean of 2.1 minutes and a standard deviation of 1.1 minutes. Assuming that the conditions for inference are met, which of the following statements about the *p*-value obtained from the data and the conclusion of the significance test is true?

(A) The *p*-value is less than 0.01; therefore, there is a significant difference in mean search times on the two Web sites.

(B) The *p*-value is greater than 0.01 but less than 0.05; therefore, there is a significant difference in mean search times on the two Web sites.

(C) The *p*-value is greater than 0.05 but less than 0.10; therefore, there is a significant difference in mean search times on the two Web sites.

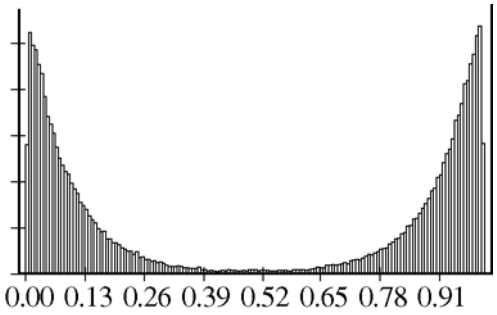
(D) The *p*-value is greater than 0.10; therefore, there is no significant difference in mean search times on the two Web sites.

(E) Since this is a matched-pairs situation, additional information is needed to perform a test of significance.

GO ON TO THE NEXT PAGE.

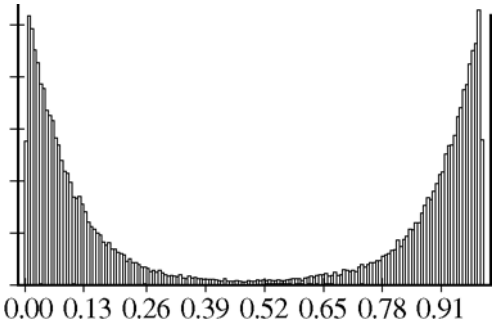
www.mymathscloud.com

40. The histogram below represents data obtained after the census of an entire population was conducted.

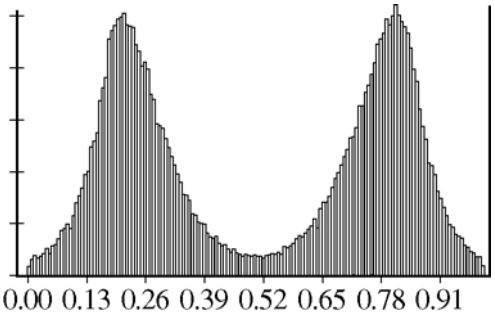


The sampling distribution of the sample mean based on samples of size 2 for the population was simulated, and a histogram of the results was produced. Which of the following histograms is most likely the histogram of that sampling distribution?

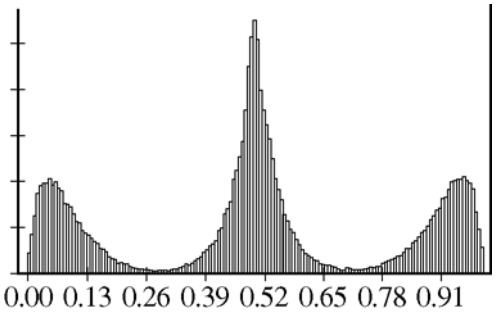
(A)



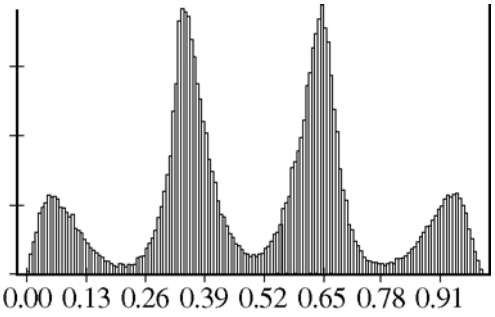
(B)



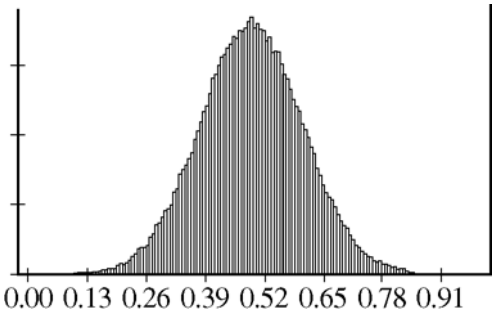
(C)



(D)



(E)



GO ON TO THE NEXT PAGE.

www.mymathscloud.com

END OF SECTION I

IF YOU FINISH BEFORE TIME IS CALLED, YOU MAY CHECK YOUR WORK ON THIS SECTION.

DO NOT GO ON TO SECTION II UNTIL YOU ARE TOLD TO DO SO.

www.mymathscloud.com

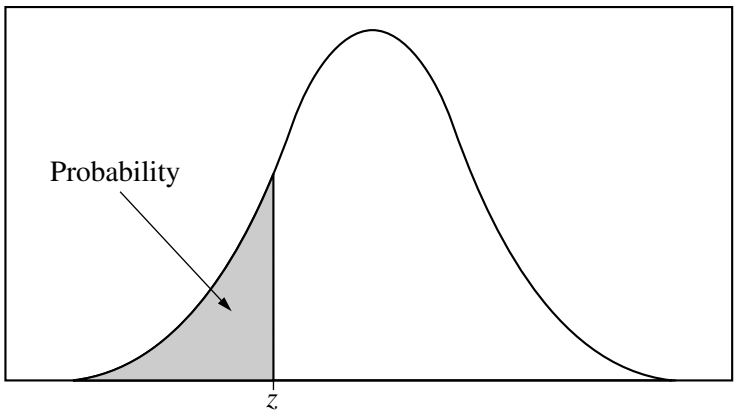


Table entry for z is the probability lying below z.

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

www.mymathscloud.com

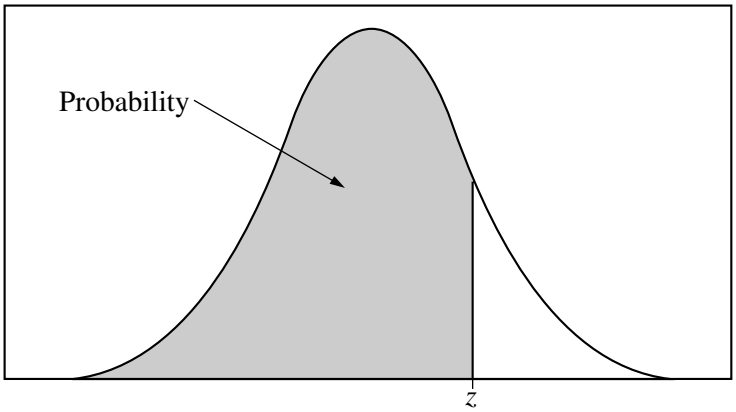


Table entry for z is the probability lying below z.

Table A (***Continued***)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

www.mymathscloud.com

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

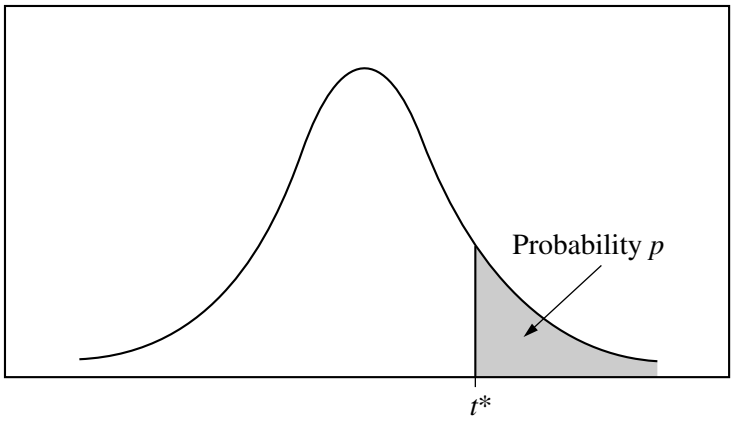


Table B t **distribution critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

www.mymathscloud.com

Table entry for p is the point (χ^2) with probability p lying above it.

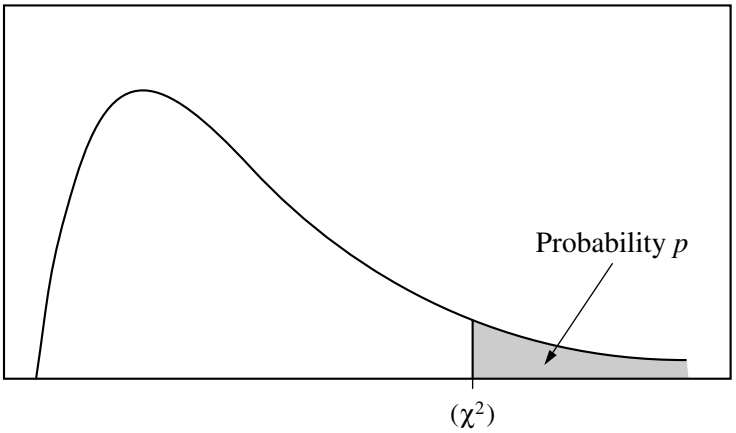


Table C χ^2 critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

www.mymathscloud.com

Section II

Free-Response Questions

www.mymathscloud.com

Formulas begin on page 48.
Questions begin on page 51.
Tables begin on page 60.

www.mymathscloud.com

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\Sigma x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1}\Sigma\left(x_i - \bar{x}\right)^2}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\Sigma\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\Sigma\left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$r = \frac{1}{n-1}\Sigma\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\Sigma\left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\Sigma\left(x_i - \bar{x}\right)^2}}$$

www.mymathscloud.com

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$E(X) = \mu_x = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_x = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If $\bar{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

www.mymathscloud.com

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\bullet$ (standard deviation of statistic)

Single-Sample

Statistic	**Standard Deviation of Statistic**
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

Statistic	**Standard Deviation of Statistic**
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

www.mymathscloud.com

STATISTICS
SECTION II
Part A
Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II score—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. Caffeine, a chemical found in many popular beverages, is known for reducing fatigue. A student wanted to investigate the caffeine content in popular beverages, such as soft drinks, energy drinks, tea, and coffee. The following data collected by the student show the amounts of caffeine (in milligrams per 12-ounce serving) for twelve popular beverages.

72	55	34	45	38	70	7.5	165	80	105	40	35

   (a) Construct an appropriate graphical display of the amounts of caffeine found in the twelve beverages.

   (b) Use the graph in part (a) to write a few sentences describing the distribution of caffeine content for the twelve beverages.

   (c) A 12-ounce cup of one popular gourmet coffee contains over 300 milligrams of caffeine. If this value was added to the data set of twelve numbers above, how would the mean and median of the data set above compare with the mean and median of the new data set with the thirteen numbers? Explain how this comparison could be made without performing any computations.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

2. Members of the research and development division of a bicycle tire manufacturer are investigating tread life of rubber bicycle tires. They have suggested that a study be conducted to determine whether bicycle tires produced using a new synthetic rubber compound have a longer tread life than the tread life of bicycle tires produced using the standard rubber compound.

   A researcher in the division suggested the study be designed in the following way. Select 60 identical bicycles and randomly assign 30 of those bicycles to one group, A, and the rest to a second group, B. All 60 bicycles will be equipped with front tires produced using the standard rubber compound. However, the bicycles in group A will be equipped with rear tires produced using the new synthetic rubber compound, while the bicycles in group B will be equipped with rear tires produced using the standard rubber compound.

   A total of 60 bicyclists will be randomly selected from the population of students at a local university who regularly ride a bicycle. The 60 bicycles will be randomly assigned to the 60 students (with a different bicycle assigned to each student), and the students will be asked to ride the bicycles for a six-month period. At the end of the six-month period, the researcher will compare the mean amounts of rear tire tread wear for the bicycles in the two groups.

   (a) What type of design has been proposed for the study?

   What is the response variable in the design?

   (b) Other than using a larger sample size, describe a better design for the study than the one proposed by the researcher. Explain why your design is better.

   (c) For your design in part (b), identify a statistical test that could be conducted to determine whether tires produced using the new compound have longer tread life than tires produced using the standard compound. (You do not have to carry out the test.)

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

3. An important method for controlling the spread of the H6N2 influenza (bird flu) virus in chickens is having a procedure to determine whether chickens are infected with the virus. It is common to apply a procedure, called an ELISA test, to measure the concentration of anti–bird flu antibodies in a blood sample taken from a chicken. If the ELISA test reveals a high-enough concentration of antibodies, the chicken is said to test positive, and it is classified as infected with the virus. Otherwise, the chicken is said to test negative, and it is classified as not infected. However, the ELISA test is a complex procedure that is not always accurate. One type of mistake, a false positive result, occurs when the ELISA test gives a positive result for a chicken that is not infected with the virus. A second type of mistake, a false negative result, occurs when the ELISA test gives a negative result for an infected chicken.

   Considering the possibility of false positives and false negatives for tests on individual chickens, veterinarians have developed the following procedure for determining if the H6N2 virus is present in a large flock of chickens.

   - Randomly select 10 chickens from the flock.
   - Perform the ELISA test on a blood sample from each of the 10 chickens.
   - Conclude that the H6N2 virus is present in the flock if at least 3 out of the 10 chickens have positive ELISA test results.

   Suppose a veterinarian applies the procedure to a flock of 100,000 chickens at a commercial egg production farm. The ELISA test is known to have probability 0.05 of producing a false positive result and probability 0.10 of producing a false negative result for a single chicken.

   (a) If no chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is not present in the flock? Show how you found your answer.

   (b) If no chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer.

   (c) If every chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer.

   (d) If 20 percent of the chickens in the flock are infected with the H6N2 virus and the other 80 percent are <u>not</u> infected, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

4. The department of parks and recreation of a certain city conducts summer programs for residents of its six districts. The summer programs include operating and maintaining community swimming pools in each of the districts as well as offering sports and recreational programs for school-age children, young adults, and older adults.

   The table below shows the proportion of households by district out of all households that participated in the summer programs, based on annual data that were collected from simple random samples each summer over a 10-year period, ending in the year 2000. The proportions are being used by the city for planning purposes and for more efficiently targeting the introduction of future programs.

District	A	B	C	D	E	F
Proportion of Households	0.32	0.12	0.10	0.27	0.05	0.14

   City leaders want to test if the proportions that are being used by the city are still valid. Data collected by a statistician from a simple random sample this past summer indicated that the following number of households participated in each district.

District	A	B	C	D	E	F
Number of Households	100	35	40	22	12	31

   (a) The statistician claims that the data for this past summer provide evidence that the proportions that are being used by the city are no longer valid. Give statistical evidence to justify the claim.

   (b) Which one of the six districts had the greatest change in participation since the year 2000 ? Use the information from part (a) to explain your choice.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

5. Bone mineral density (BMD) is a measure of bone strength. It is defined as the ratio of bone mass to the cross sectional area of the bone that is scanned, and it is expressed in units of grams per square centimeter (g/cm^2). Recent studies suggest that peak BMD in women is achieved between ages 15 and 40, and BMD declines after age 45. Decreased BMD is associated with increased risk of bone fracture.

   In a recent study, the impact of regular physical exercise on women in differing stages of BMD development was examined. A simple random sample of 59 women between the ages of 41 and 45 and with no major health problems were enrolled in the study. The women were classified into one of the two following groups, based on their level of exercise activity.

   - Sedentary: minimal participation in physical exercise in the past three years (This group contained 31 women.)
   - Walkers: walk at an aerobic pace at least 135 minutes per week during the past three years (This group contained 28 women.)

   (a) The table below shows the mean BMD and corresponding standard deviation for each of the two groups of women.

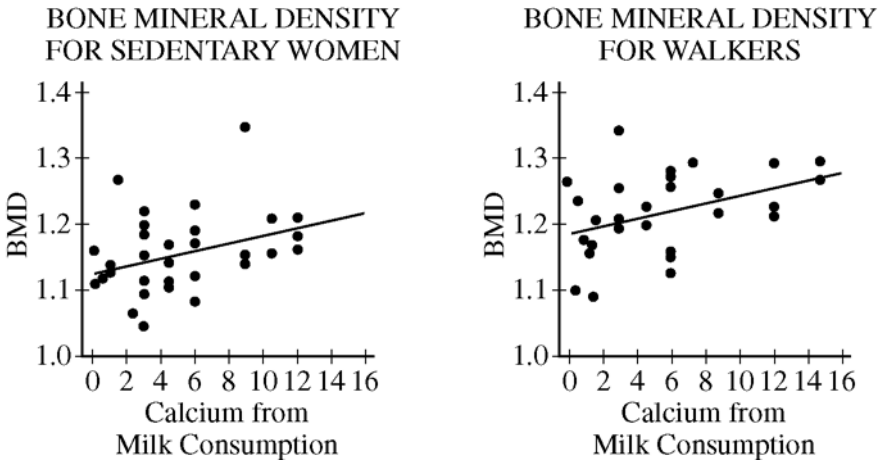
Exercise Group	Number of Women	Mean BMD	Standard Deviation
Sedentary	31	1.151	0.068
Walkers	28	1.209	0.067

   A *t*-test was conducted to compare the mean BMD levels for sedentary women and walkers. The results of the test showed a significant difference at the 0.01 level of significance. Based on the results, can it be concluded that converting sedentary women into walkers would necessarily increase their BMD levels? Explain.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

(b) There was some concern that women in the two groups may have different dietary habits that could affect BMD. For example, higher intake of milk or other foods and supplements that provide additional calcium to the body could increase BMD. To examine this possibility, the researchers also asked each woman in the study to report on weekly milk consumption when she was age 20 through age 29. The data were then used to compute a value of calcium obtained from milk consumption for each woman in the study. BMD levels were plotted against the calcium intake from consumption of milk for women in each of the groups. The plots are shown below, with the least squares estimate of a regression line on each plot.



What do the plots indicate about the relationship between BMD and calcium intake from milk consumption from age 20 through age 29 ?

(c) The line graphs in part (b) suggest that sedentary women tend to have lower calcium intake from milk consumption than walkers do. Assuming that this is true, describe the impact, if any, that it would have on conclusions that can be reached from *t*-tests for comparing mean BMD levels for the two groups, such as those conducted in part (a).

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

STATISTICS
SECTION II
Part B
Question 6

Spend about 25 minutes on this part of the exam.

Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. An engineer is developing a polymer material and is concerned that the mean density, d, of the material is not sufficiently close to the desired target value of 1.37 kilograms per milliliter (kg/mL). Nine different samples of the material were prepared. The volumes, in milliliters, for the samples were 10, 20, 30, 40, 50, 60, 70, 80 and 90. The engineer carefully measured the mass (in kilograms) of each sample. For the range of volumes, the true density, d, (in kg/mL) of the material can be estimated by the slope of the least-squares regression line fit to the resulting data. (Recall that density is defined as mass divided by volume.) Computer output for the regression analysis is shown below.

Mass = (Intercept) + d*Volume + Error

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	0.9759	0.49276	1.98
Volume	1.3812	0.00876	157.73

s = 0.6783

(a) Since the process the engineer used to measure the mass of samples does not always provide the true density value, the regression model shown above contains a random error term. To use the t-distribution to perform a test of hypotheses or construct a confidence interval for the slope of the least-squares regression line, the random errors must conform to some model assumptions. These include the assumption that the random errors are independent of each other. In the context of this experiment, this means that the error that the engineer makes in measuring the mass of one sample has no influence on the error made in determining the density of any other sample.

Describe two other assumptions about the distribution of the random errors that are needed to use the t-distribution to perform a test of hypotheses or construct a confidence interval for the slope of the least squares regression line.

(b) Assuming that all of the assumptions that you considered in part (a) are satisfied, construct a 95 percent confidence interval for d, the true density of the material. With respect to the target density of 1.37 kg/mL, what conclusion can be reached?

(c) In context, what does the intercept in the least squares regression line that is associated with the computer output above represent?

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

(d) Use your answer from part (c) to explain why it might be reasonable to set the intercept equal to 0 and consider the resulting alternative model

Mass = d*Volume + Error

as a model for the true density of the material.

The engineer fit a regression model with an intercept of 0 to the data from the nine samples and obtained the following results for the least squares estimate of the slope.

Mass = d*Volume + Error

Coefficients:

	Estimate	Std. Error	t value
Volume	1.3966	0.00469	297.49

s = 0.7925

The engineer then wanted to use the results to construct a 95 percent confidence interval for the slope, but could not decide if it should be constructed as

$1.3966 \pm (2.306)(0.00469)$

or as

$1.3966 \pm (2.262)(0.00469)$.

The first expression uses the critical value 2.306, the 97.5 percentile of a *t*-distribution with 8 degrees of freedom, and the second expression uses the critical value 2.262, the 97.5 percentile of a *t*-distribution with 9 degrees of freedom. The engineer decided to perform a simulation study to determine the appropriate formula to use to construct a 95 percent confidence interval for the slope.

(e) To perform the simulation study, the engineer will simulate samples of mass observations using the model

Mass = 1.3966*Volume + Error

for the nine volumes of material (10, 20, 30, 40, 50, 60, 70, 80, and 90) used in the study. The standard deviation of the random errors is assumed to be 0.7925. Explain how the engineer can simulate a sample of observations for the nine amounts using a computer program that can generate random samples from a standard normal distribution.

(f) Given the method of simulating samples described in part (e), explain how the engineer can determine which of the two expressions provides an appropriate method of constructing a 95 percent confidence interval for the slope of a least squares regression line, using a model with an intercept of 0.

GO ON TO THE NEXT PAGE.

www.mymathscloud.com

STOP

END OF EXAM

---

www.mymathscloud.com

Table entry for z is the probability lying below z.

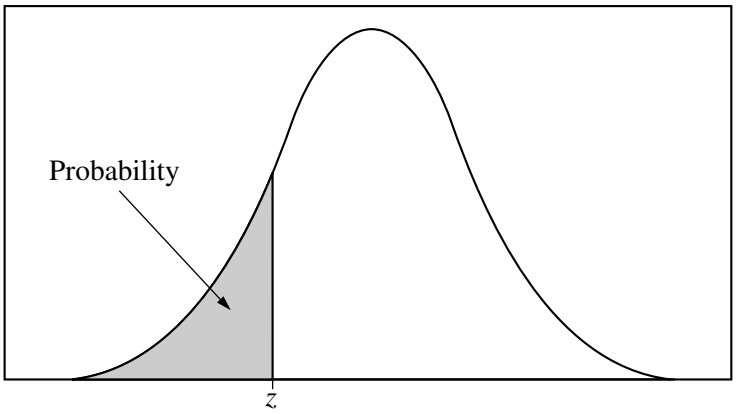


Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

www.mymathscloud.com

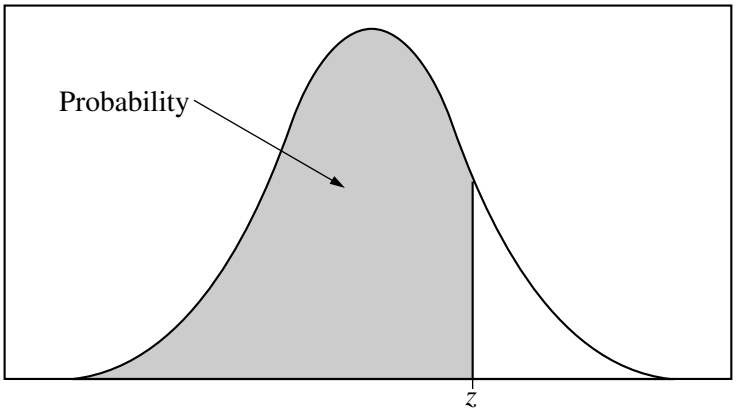


Table entry for z is the probability lying below z.

Table A (***Continued***)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

www.mymathscloud.com

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

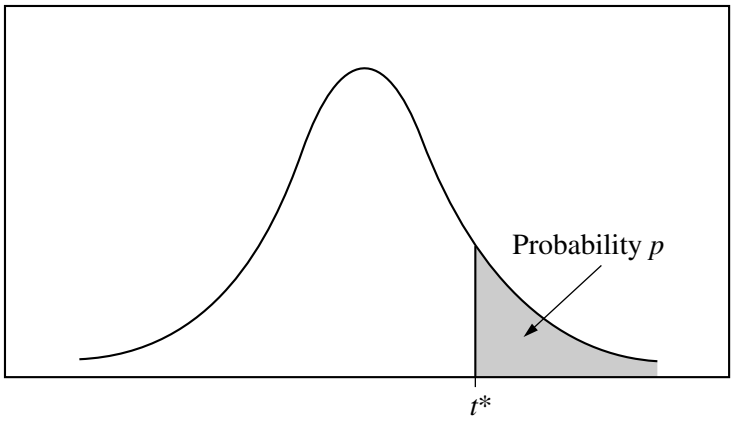


Table B t **distribution critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

www.mymathscloud.com

Table entry for p is the point (χ^2) with probability p lying above it.

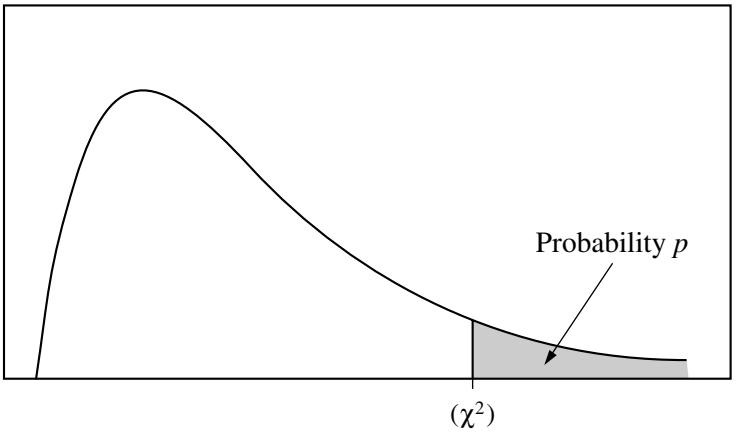


Table C χ^2 critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

www.mymathscloud.com

Name: _____

AP® Statistics
Student Answer Sheet for Multiple-Choice Section

No.	Answer
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	

No.	Answer
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	

www.mymathscloud.com

AP® Statistics
Multiple-Choice Answer Key

No.	Correct Answer
1	B
2	B
3	E
4	B
5	C
6	E
7	E
8	C
9	A
10	A
11	C
12	A
13	C
14	D
15	D
16	D
17	D
18	C
19	D
20	D
21	D
22	D
23	B
24	E
25	B
26	C
27	A
28	A
29	B
30	E

No.	Correct Answer
31	A
32	C
33	A
34	B
35	C
36	D
37	E
38	A
39	D
40	C

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

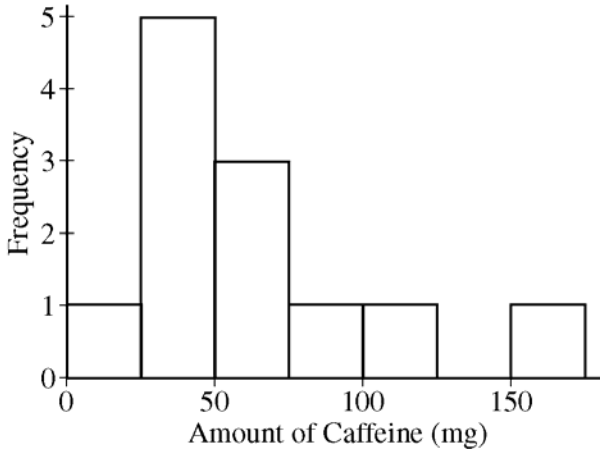
Question 1

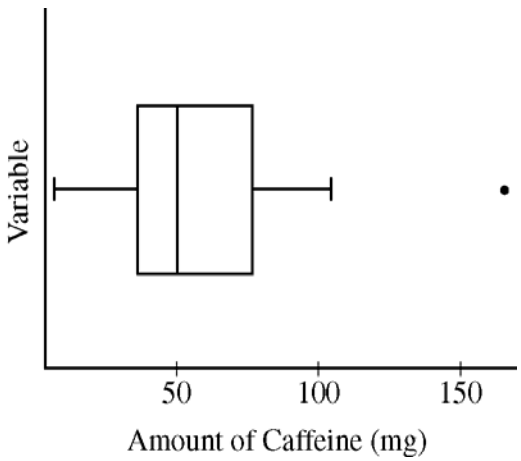
Intent of Question:

The primary goals of this question are to assess the student's ability to: (a) graphically display a small data set in a meaningful way; (b) describe the distribution of the data; and (c) evaluate how the addition of an outlier affects the measures of central tendency of the data.

Sample Correct Response:

a. Two different graphical displays that would be possible correct responses are shown below.





b. The median amount of caffeine in popular beverages is 50 mg. The IQR of caffeine is 39.5 mg. (Note: The standard deviation of caffeine is 41.37 mg, but it is not expected that the student will obtain this statistic directly from the graph.) The distribution for the amount of caffeine in beverages is skewed right with an outlier at 165 mg.

c. The mean would increase for the larger data set, since it is affected by outliers. However, the median would not substantially change since it is resistant to outliers. (Note: The median may shift ever so slightly, since the number of values in the data set changes.)

Scoring:

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

Part (a):

- is essentially correct if the student presents a correct graphical display such as a boxplot or histogram, with complete labeling of that display.
- is partially correct if the student displays the data correctly but does not label the graph completely *OR* labels the graph completely but makes one or two minor errors on a graphical display that is otherwise correct.
- is incorrect if the student uses an incorrect graphical display (or no graphical display), with or without correct labels.

www.mymathscloud.com

Question 1 (continued)

Part (b):

- is essentially correct if the student correctly addresses the center, shape, and spread of the data in context.
- is partially correct if the student correctly addresses two of the three attributes (center, shape, and spread) of the data.
- is incorrect if the student addresses fewer than two of the three attributes.

Part (c):

- is essentially correct if the student presents correct comparisons between the original (i.e., without the 300 data point included) and new (i.e., with the 300 data point included) sets for both the mean and median.
- is partially correct if the student presents a correct comparison between the original (i.e., without the 300 data point included) and new (i.e., with the 300 data point included) sets for the mean *OR* the median.
- is incorrect if the student does not present a correct comparison or presents no comparison.

4 Complete Response

All three parts are essentially correct.

3 Substantial Response

Two parts are essentially correct, and one part is partially correct.

2 Developing Response

Two parts are essentially correct, and no part is partially correct.

OR

One part is essentially correct, and two parts are partially correct.

OR

Three parts are partially correct.

1 Minimal Response

One part is essentially correct, and either one part or no part is partially correct.

OR

No parts are essentially correct, and two parts are partially correct.

0 No Credit

Response does not meet criteria for at least the minimal level.

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

Question 2

Intent of Question:

This question assesses the student's ability to: (a) identify the type of design and response variable for a statistical study; (b) propose a better design for the study than the one presented; and (c) identify an appropriate statistical test to use in the analysis of the data.

Sample Correct Response:

a. This is a completely randomized design. The response variable is the amount of tread wear, which can be found by taking the difference in the tread on the rear tire at the outset and the tread on the rear tire after six months of use.

b. Place one tire with the new compound and one tire with the standard compound on each bicycle. Randomly choose 30 of the 60 bicycles, and place the tire with the new compound on the front for those bicycles and the tire with the new compound on the rear for the remaining 30 bicycles. Randomly distribute the bicycles to the students. Since students use bicycles differently, this design will ensure that each tire is exposed to similar conditions. (Note: This design would be stronger if the front and rear positions of the new and standard tires on each bicycle are interchanged after the first three months of the study.)

c. A matched-pairs *t*-test would determine whether tires produced using the new compound have longer tread life than tires produced using the standard compound.

Scoring:

Parts (a) and (b) are each scored as essentially correct, partially correct, or incorrect. Part (c) is scored as essentially correct or incorrect.

Part (a):

- is essentially correct if the student identifies both the type of design and the response variable correctly.
- is partially correct if the student identifies either the type of design or the response variable correctly.
- is incorrect if the student identifies neither the type of design nor the response variable correctly.

Part (b):

- is essentially correct if the student suggests a better design than the one proposed by the researcher in part (a) and describes it sufficiently.
- is partially correct if the student suggests a better design with a less-than-sufficient explanation.
- is incorrect if the student does not suggest a better design.

Part (c):

- is essentially correct if the student identifies a correct test.
- is incorrect if the student does not identify a correct test.

www.mymathscloud.com

Question 2 (continued)

4 Complete Response

All three parts are essentially correct.

3 Substantial Response

Two parts are essentially correct, and one part is partially correct.

2 Developing Response

Two parts are essentially correct, and no part is partially correct.

OR

One part is essentially correct, and two parts are partially correct.

1 Minimal Response

One part is essentially correct, and either one part or no part is partially correct.

OR

No parts are essentially correct, and two parts are partially correct.

0 No Credit

Response does not meet criteria for at least the minimal level.

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

Question 3

Intent of Question:

This question requires the student to apply knowledge of a binomial distribution in the computation of probabilities.

Sample Correct Response:

a. The veterinarian will correctly conclude that the H6N2 virus is not present in the flock if no more than two chickens have false positive results. The probability that this occurs is

$$\binom{10}{0}(0.05)^0(0.95)^{10} + \binom{10}{1}(0.05)^1(0.95)^9 + \binom{10}{2}(0.05)^2(0.95)^8 = 0.9884964$$

b. Using the answer from part (a), the probability that the veterinarian will incorrectly conclude that the H6N2 virus is present in the flock is

$$1 - 0.9884964 = 0.0115036$$

c. The veterinarian will correctly conclude that the H6N2 virus is present in the flock if at least three chickens show positive results. The probability that this occurs is

$$\sum_{k=3}^{10}\binom{10}{k}(0.90)^k(0.10)^{10-k} = 1 - \left[\binom{10}{0}(0.90)^0(0.10)^{10} + \binom{10}{1}(0.90)^1(0.10)^9 + \binom{10}{2}(0.90)^2(0.10)^8\right]$$

$$= 0.9999996$$

d. The veterinarian will correctly conclude that the H6N2 virus is present in the flock if at least three chickens show positive results. The probability that a randomly selected chicken shows a positive result is

$$\text{P(true positive)} \times \text{P(infected chicken)} + \text{P(false positive)} \times \text{P(non-infected chicken)}$$
$$= (0.90)(0.20) + (0.05)(0.80) = 0.22$$

Then the probability that at least three chickens show positive results is

$$\sum_{k=3}^{10}\binom{10}{k}(0.22)^k(0.78)^{10-k} = 1 - \left[\binom{10}{0}(0.22)^0(0.78)^{10} + \binom{10}{1}(0.22)^1(0.78)^9 + \binom{10}{2}(0.22)^2(0.78)^8\right]$$

$$= 0.3831197$$

Scoring:

Parts (a), (b), (c), and (d) are each scored as essentially correct, partially correct, or incorrect.

www.mymathscloud.com

Question 3 (continued)

Parts (a) through (d):

- are each essentially correct if the student computes a correct probability and fully supports the probability with correct work. Although the numerical results of intermediate computations are not required in the supporting work, students will need to show enough evidence of the correct process and numbers utilized in that process to arrive at the final numerical probability value in each part.
- are each partially correct if the student computes a correct probability but the supporting work is incomplete (there must be some correct supporting work to earn partial credit), *OR* the computed probability is incorrect but the supporting work is complete and a correct process is presented. In the latter instance, the error in computed probability should be the result of a computation mistake or minor transcription error only.
- are each incorrect if the student does not compute a correct probability and does not present a sufficient amount of supporting work.

Each essentially correct response in each part is worth 1 point, each partially correct response is worth ½ point, and each incorrect response is worth 0 points.

4 Complete Response

A total of 4 points are earned.

3 Substantial Response

A total of 3 points are earned.

2 Developing Response

A total of 2 points are earned.

1 Minimal Response

A total of 1 point is earned.

0 No Credit

A total of 0 points are earned.

Note: If the total score earned for the four parts is not a whole number of points (i.e., 3½ points, 2½ points, 1½ points, or ½ point), then the quality of the overall work in all four parts collectively needs to be evaluated. If the overall quality warrants, the total score should be moved up ½ point to the next higher whole number score; if that is unwarranted, then the total score should be moved down ½ point to the next lower whole number score. The general student performance expectations found in the *AP Statistics Course Description* are helpful in making this judgment. In the May 2007, May 2008 version of that publication, the performance expectations can be found on pages 27–28.

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

Question 4

Intent of Question:

This question assesses the student's ability to identify and conduct the appropriate statistical test to investigate a claim and to use the results to answer a follow-up question.

Sample Correct Response:

a. To test the statisticians' claim, a chi-square goodness of fit test should be performed.

Null hypothesis: Population proportions for the past summer are consistent with the proportions that are being used by the city for planning purposes and for more efficiently targeting the introduction of future programs.

Alternative hypothesis: Population proportions for the past summer are not consistent with the proportions that are being used by the city for planning purposes and for more efficiently targeting the introduction of future programs.

Check of conditions:

- Data were collected from a simple random sample.
- Expected cell counts are each greater than 5. (See column three of the table below.)

District	Observed, O	Expected, E	$\frac{(O-E)^2}{E}$
A	100	76.8	7.01
B	35	28.8	1.33
C	40	24	10.67
D	22	64.8	28.27
E	12	12	0.00
F	31	33.6	0.20
	Sum		$\chi^2 = 47.48$

p-value < 0.001, with 5 degrees of freedom

Since the p-value is so small, we have convincing evidence that the population proportions for the past summer are not consistent with the proportions that are being used by the city.

b. Based on the chi-square table, it appears that District D showed the greatest change in participation since 2000, because it contributed the greatest amount toward the calculation of the chi-square value.

www.mymathscloud.com

Question 4 (continued)

Scoring:

The part (a) score is comprised of three separate components—stating the null and alternative hypotheses and verifying the conditions for inference; carrying out the mechanics of the test; and stating a correct conclusion in the context of the question. Each component is scored as essentially correct or incorrect. Part (b) of the question is comprised of a single component that is scored as essentially correct or incorrect.

Part (a)—Hypotheses/Conditions:

- is essentially correct if the student states the correct null and alternative hypotheses and verifies the conditions for inference for this test.
- is incorrect if the student does not present correct information for both the hypotheses to be tested and a verification of the conditions for inference.

Part (a)—Mechanics:

- is essentially correct if the student carries out the mechanics of the test. There must be a sufficient level of detail in the supporting computations in order to earn credit for this component.
- is incorrect if the student does not display correct mechanics for the test.

Part (a)—Conclusion:

- is essentially correct if the student presents a correct conclusion in the context of the question linked to the *p*-value.
- is incorrect if the student is unable to present a correct conclusion, or no conclusion is presented.

Part (b):

- is essentially correct if the student identifies the correct district and supports the choice of that district.
- is incorrect if the student identifies an incorrect district or does not adequately support the choice of the correct district.

4 Complete Response

Four components are essentially correct.

3 Substantial Response

Three components are essentially correct.

2 Developing Response

Two components are essentially correct.

1 Minimal Response

One component is essentially correct

0 No Credit

No components are essentially correct.

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

Question 5

Intent of Question:

The purpose of this question is to assess the student's ability to determine whether a cause-and-effect conclusion is justified in a particular situation and to make appropriate conclusions using available data.

Sample Correct Response:

a. Since this was an observational study, a cause-and-effect conclusion is not justified. Sedentary women could differ from walkers with respect to other factors that affect BMD. For example, walkers may be more likely to avoid smoking and may be more careful about including sources of calcium in their diets. Simply encouraging sedentary women to walk more may not achieve the improvement in average BMD level shown in this study, if other differences between the two groups are ignored.

b. The plots of the data and the estimated regression lines suggest that increased calcium intake from milk consumption by women in their 20s is associated with higher average BMD levels. (Note: Higher levels of milk consumption by women in their 20s could be associated with a history of higher levels of milk consumption as children and teenagers and later as adults.)

c. This limits the conclusions that can be made about possible effects of exercise on BMD. If higher levels of milk consumption provide calcium needed to promote higher BMD levels, and sedentary women tend to have lower levels of milk consumption than walkers do, then all or part of the difference in BMD means for the two groups could be due to higher levels of calcium intake among walkers and not the physical activity.

Scoring:

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

Part (a):

- is essentially correct if the student clearly identifies another factor (besides walking) that could be contributing to the difference in BMD levels between the two groups *OR* explains the observational nature of the study.
- is partially correct if the student correctly realizes that there are other contributing factors but does not address at least one of them clearly in the response.
- is incorrect if the student's response does not meet at least the criteria described in the partially correct category.

Part (b):

- is essentially correct if the student presents a correct statement of the relationship between calcium intake from milk consumption and BMD levels.
- is partially correct if the student's response suggests the correct relationship between calcium intake and BMD levels, but the description of the relationship is incomplete or is not clearly stated.
- is incorrect if the student's response does not meet at least the criteria described in the partially correct category.

www.mymathscloud.com

Question 5 (continued)

Part (c):

- is essentially correct if the student uses the information in part (b) in an explanation that identifies the limitations associated with the conclusion in part (a).
- is partially correct if the student makes correct observations regarding the limitations of the conclusion in part (a) but does not link those observations strongly enough to the information presented in part (b).
- is incorrect if the student's response does not meet at least the criteria described in the partially correct category.

4 Complete Response

All three parts are essentially correct.

3 Substantial Response

Two parts are essentially correct, and one part is partially correct.

2 Developing Response

Two parts are essentially correct, and no part is partially correct.

OR

One part is essentially correct, and two parts are partially correct.

OR

Three parts are partially correct.

1 Minimal Response

One part is essentially correct, and either one part or no part is partially correct.

OR

No parts are essentially correct, and two parts are partially correct.

0 No Credit

Response does not meet criteria for at least the minimal level.

www.mymathscloud.com

AP® Statistics
Free-Response Scoring Guidelines

Question 6

Intent of Question:

The purpose of this question is to assess the student's ability to work with a linear regression model; describe how a simulation could be performed in a study; and use that simulation method to determine which of two possible expressions provides an appropriate method for constructing a confidence interval for the slope of a least-squares regression line with an intercept of 0.

Sample Correct Response:

a. One assumption is that the random errors at each volume have the same variance. The level of variation in the mass measurements about the true mass does not vary with the volume. A second assumption is that the random errors at each volume are normally distributed.

b. A 95 percent confidence interval for the slope is $1.3812 \pm (2.365)(0.00876)$ or (1.3605, 1.4019) kg/mL. Since a 95 percent confidence interval has a probability of 0.95 of containing the true value, and the interval provided by the data includes 1.37, this result is consistent with a material with a true density of 1.37 kg/mL.

c. The intercept represents the mean mass in kilograms for a sample with a volume of 0.

d. Since a sample with no volume should have no mass, it would be very reasonable to consider a regression model with an intercept of 0.

e. The basic steps for simulating a data set with nine observations are:

- Use a calculator or computer to generate a random observation z from a standard normal distribution.
- For a specific volume, compute a simulated mass measurement as $y = 1.3966 \times \text{volume} + 0.7925 \times z$.
- Repeat each step in the previous two bulleted statements for each of the nine amounts—10, 20, 30, 40, 50, 60, 70, 80, and 90—to obtain a sample of nine mass measurements.

f. The engineer would need to simulate a large number of data sets, say 10,000, using the procedure described in part (e). For each simulated data set, the engineer would need to fit the least-squares regression line with an intercept of 0 to obtain the estimated slope and its standard error and evaluate both of the proposed formulas for constructing a 95 percent confidence interval for the slope. For the appropriate method, about 95 percent of the simulated confidence intervals should contain 1.3966, the slope of the "true" line used to simulate the data sets.

Scoring:

There are six parts for this question. Parts (a) and (b) are scored together as a single score—essentially correct, partially correct, or incorrect. Parts (c) and (d) are scored together as a single score—also as essentially correct, partially correct, or incorrect. Part (e) is scored as essentially correct, partially correct, or incorrect. Part (f) is scored as essentially correct, partially correct, or incorrect.

Each essentially correct response in each part is worth 1 point, each partially correct response is worth ½ point, and each incorrect response is worth 0 points.

www.mymathscloud.com

Question 6 (continued)

Parts (a) and (b):

- are essentially correct if the student presents two correct assumptions about the distribution of the random errors in part (a) and constructs a correct confidence interval in part (b) with a corresponding correct conclusion in context.
- are partially correct if the student correctly responds to part (a) in its entirety but only partially correctly to part (b), *OR* correctly responds to part (b) in its entirety but only partially correctly to part (a), *OR* responds partially correctly to each of the two parts. In order to earn credit for a partially correct response in either part, the part (a) response must have one correct assumption stated, and the part (b) response must have a correct confidence interval with an incorrect or incomplete conclusion *OR* a confidence interval that contains only a computational error with a corresponding conclusion that is correct based on that interval.
- are incorrect if the student's response does not meet at least the criteria described in the partially correct category.

Parts (c) and (d):

- are essentially correct if the student presents a correct interpretation for the intercept in part (c) and a correct explanation in part (d) that supports the use of an intercept of 0 for the alternative model.
- are partially correct if the student answers either part (c) or part (d) correctly.
- are incorrect if the student's response does not meet at least the criteria described in the partially correct category.

Part (e)

- is essentially correct if the student presents a correct method for simulating a data set with nine observations.
- is partially correct if the student presents an incomplete method for simulating the data set. The method must convey some information beyond the general level (i.e., use a random number table or use the random number generator on the calculator).
- is incorrect if the student's response does not meet at least the criteria described in the partially correct category.

Part (f)

- is essentially correct if the student fully explains the process for determining which of the two provided expressions will result in a confidence interval for the slope of a least-squares regression line for the model with an intercept of 0.
- is partially correct if the student utilizes the information in part (e) to fit the least-squares regression line with an intercept of 0 for the purpose of obtaining an estimated slope and standard error but does not complete the process *OR* makes a mistake in carrying out this process.
- is incorrect if the student's response does not meet at least the criteria described in the partially correct category.

www.mymathscloud.com

Question 6 (continued)

4 Complete Response

A total of 4 points are earned.

3 Substantial Response

A total of 3 points are earned.

2 Developing Response

A total of 2 points are earned.

1 Minimal Response

A total of 1 point is earned.

0 No Credit

A total of 0 points are earned.

Note: If the total score earned for the four parts is not a whole number of points (i.e., 3½ points, 2½ points, 1½ points, or ½ point), then the quality of the overall work in all four parts collectively needs to be evaluated. If the overall quality warrants, the total score should be moved up ½ point to the next higher whole number score; if that is unwarranted, then the total score should be moved down ½ point to the next lower whole number score. The general student performance expectations found in the *AP Statistics Course Description* are helpful in making this judgment. In the May 2007, May 2008 version of that publication, the performance expectations can be found on pages 27–28.