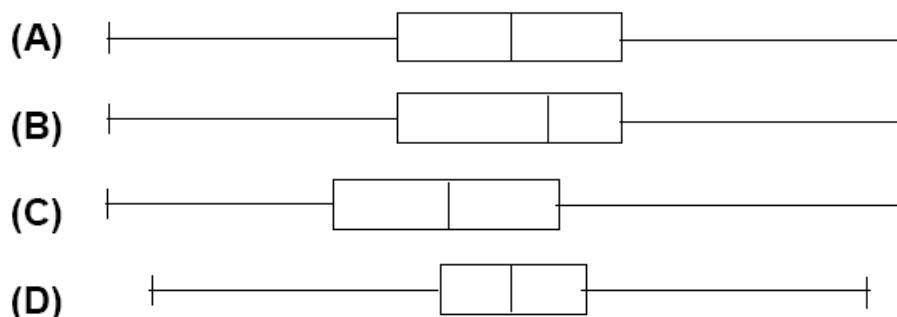


# Classroom Voting Questions: Elementary Statistics

## Density Curves and Normal Distributions

1. If a large sample were drawn from a normal distribution and accurately represented the population, which of the following is most likely to be a box plot of that sample?



**(E) Two from (A)-(D) are correct.**

**(F) Three from (A)-(D) are correct.**

**(G) All from (A)-(D) are correct.**

*Answer: (e)* (A) This response is one of two appropriate box plots; the other is (D). This box plot might represent a normal distribution because it is symmetrical both within the box and with the length of the whiskers. The box in (A) is wider than that of (D) indicating a larger standard deviation.

(B) A box plot for a normal distribution would be completely symmetrical. In this box plot, the median is not symmetrical within the box although the whiskers are of the same length.

(C) The whiskers are of different lengths so this represents a skewed rather than a symmetrical distribution.

(D) This response is one of two appropriate box plots; the other is (A). This box plot might represent a normal distribution because it is symmetrical both within the box and with the length of the whiskers. The box in (D) is narrower than that of (A) indicating a smaller standard deviation

(E)\* correct (A) and (D) are both box plots for normal distributions.

(F), (G) Only (A) and (D) are box plots for normal distributions.

by Murphy, McKnight, Richman, and Terry

STT.01.03.010

CC KC MA207 F09: 32/3/3/10/**42**/10/0 time 2:00

AS DH MA3321 Su12: 15/0/0/0/**77**/8/0 time 2:00

AS DH MA1333 010 F12: 0/0/0/11/**22**/17/50 time 2:00

AS DH MA1333 020 F12: 5/0/0/0/**14**/19/62 time 2:30

AS DH 1333 010 S13: 0/0/0/9/**57**/4/30 time 2:30

AS DH 1333 020 S14: 0/0/0/0/**13**/16/71 time 2:40 ,

AS DH 1333 010 F14: 12/0/0/0/**85**/3/0 time 2:30 ,

AS DH 1333 020 S15: 18/9/0/9/**27**/14/23 time 2:40 ,

2. Consider the continuous random variable  $X$  = the weight in pounds of a randomly selected newborn baby born in the United States last year. Suppose that  $X$  can be modeled with a normal distribution with mean  $\mu = 7.57$  and standard deviation  $\sigma = 1.06$ . If the standard deviation were  $\sigma = 1.26$  instead, how would that change the graph of the pdf of  $X$ ?
- (a) The graph would be narrower and have a greater maximum value.
  - (b) The graph would be narrower and have a lesser maximum value.
  - (c) The graph would be narrower and have the same maximum value.
  - (d) The graph would be wider and have a greater maximum value.
  - (e) The graph would be wider and have a lesser maximum value.
  - (f) The graph would be wider and have the same maximum value.

*Answer: (e).* Increasing the standard deviation increases the spread of the random variable, which means the graph will be wider. Since the area under the curve must be 1, making the graph wider means making it shorter as well so that the area remains constant.

by Derek Bruff

STT.01.03.020

AS DH MA3321 Su12: 15/0/0/54/**31** time 2:20

AS DH 1333 010 S13: 0/0/0/67/**7**/27 time 3:00

AS DH 1333 020 S14: 4/8/0/17/**13**/58 time 3:30 ,

AS DH 1333 010 F14: 0/0/0/47/**23**/30 time 3:00 ,

AS DH 1333 020 S15: 0/0/0/39/**33**/28 time 3:10 ,

3. Consider the continuous random variable  $X$  = the weight in pounds of a randomly newborn baby born in the United States during 2006. Suppose that  $X$  can be modeled with a normal distribution with mean  $\mu = 7.57$  and standard deviation  $\sigma = 1.06$ . If the mean were  $\mu = 7.27$  instead, how would that change the graph of the pdf of  $X$ ?

- (a) The graph would be shifted to the left.
- (b) The graph would be shifted to the right.
- (c) The graph would become more negatively skewed.
- (d) The graph would become more positively skewed.
- (e) The graph would have a greater maximum value.
- (f) The graph would have a lesser maximum value.

*Answer: (a).* Decreasing the mean without changing the standard deviation moves the entire graph to the left, 0.3 units in this case.

by Derek Bruff

STT.01.03.030

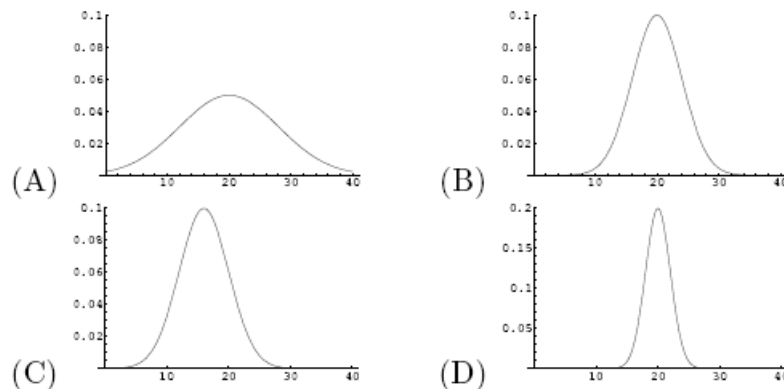
AS DH MA3321 Su12: **92**/0/0/0/0/8 time 1:30

AS DH 1333 020 S14: **83**/4/0/0/4/8 time 2:00 ,

AS DH 1333 010 F14: **79**/0/0/0/6/15 time 2:20 ,

AS DH 1333 020 S15: **100**/0/0/0/0/0 time 2:30 ,

4. If  $X$  is a normal random variable with mean  $\mu = 20$  and standard deviation  $\sigma = 4$ , which of the following could be the graph of the pdf of  $X$ ?



*Answer: (b).* Note that the vertical axis for the fourth graph is different. The third choice has the wrong mean, the first choice has too great a standard deviation, and the fourth choice has too small a standard deviation. Note that 68% of the area under the curve should fall between  $\mu - \sigma$  and  $\mu + \sigma$ . Also, one can show that the graph changes concavity at  $\mu \pm \sigma$ .

by Derek Bruff

STT.01.03.040

CC HZ MA207 F09: 52/**43**/0/5 time 2:00

CC KC MA207 F09: 39/**26**/6/29 time 2:00

AS DH MA3321 Su12: 0/**100**/0/0 time 2:00

5. Find  $z_{0.15}$ .

- (a) 1.04
- (b)  $-1.04$

*Answer: (a).*

by David A. Huckaby

STT.01.03.045 DH 90

AS DH MA1333 010 F12: **11**/89 time 2:30

AS DH MA1333 020 F12: **35**/65 time 2:00

AS DH 1333 010 S13: **76**/24 time 2:30

AS DH 1333 020 S14: **96**/4 time 2:20 ,

AS DH 3321 010 S14: **74**/26 time 2:20 ,

AS DH 1333 010 F14: **89**/11 time 2:00 ,

AS DH 3321 010 F14: **74**/26 time 2:10 ,

AS DH 1333 020 S15: **47**/53 time 3:00 ,

AS DH 3321 010 S15: **87**/13 time 2:00 ,

6. Yogurt is sold in cartons labeled as containing 6 oz, but the actual contents vary slightly from container to container. Suppose that the content distribution is approximately normal in shape with a mean of 6 oz and a standard deviation of 0.05 oz. What can be said about the percentage of cartons that have actual contents less than 5.95 oz?

- (a) The percentage is approximately 68%
- (b) The percentage is approximately 34%
- (c) The percentage is approximately 32%
- (d) The percentage is approximately 16%

*Answer: (d).*

by Roxy Peck for the textbooks: Roxy Peck and Jay Devore, Statistics: The Exploration and Analysis of Data, 6th Edition, Brooks/Cole Cengage Learning 2008 and Roxy Peck, Chris Olsen and Jay Devore, Introduction to Statistics and Data Analysis, 3rd Edition, Brooks/Cole Cengage Learning 2008.

STT.01.03.050

CC KC MA207 F09: 23/15/20/**42** time 2:00

AS DH MA3321 Su12: 8/0/8/**85** time 2:20

AS DH MA1333 010 F12: 0/8/46/**46** time 3:30

AS DH MA1333 020 F12: 0/8/46/**46** time 3:40  
 AS DH 1333 010 S13: 0/25/6/**69** time 4:30  
 AS DH 1333 020 S14: 21/21/25/**33** time 3:40 ,  
 AS DH 3321 010 S14: 0/9/4/**87** time 3:00 ,  
 AS DH 1333 010 F14: 3/33/10/**53** time 3:30 ,  
 AS DH 3321 010 F14: 37/16/5/**42** time 4:00 ,  
 AS DH 1333 020 S15: 5/15/10/**70** time 3:10 ,

7. The University of Oklahoma has changed its admission standards to require an ACT-score of 26. We know the ACT is normally distributed with a mean of 21 and an SD of 5. If we sample 100 students who took the ACT at random, how many would be expected to qualify for admission to OU?
- (a) 5
  - (b) 16
  - (c) 34
  - (d) 84
  - (e) none of the above

*Answer: (b).* (A) Students giving this answer are probably guessing. This number is the standard deviation of the ACT distribution. Coincidentally, it is also the distance between the OU admission score of 26 and the ACT mean of 21.

(B)\* correct A systematic solution to this problem requires the student to recognize the need to use the 68-95-99.7 rule in order to determine the proportion of kids who would be eligible under the new admission standards. Since the rule applies to standard ( $z$ ) scores, the first step is to make a raw-score to  $z$ -score conversion.

Step 1: ACT to  $z$ -score conversion  $z = \frac{26-21}{5} = +1$ .

Step 2: What percentage of a random sample would be expected to score above  $z = +1$  if the data were normally distributed? Using the Empirical Rule, we note that approximately 68% percent of the data are expected to fall within one standard deviation of the mean, with 34

Step 3. 16% of 100 randomly sampled students is 16 students.

(C) Students neglected to subtract 34% from 50% to get the proportion to the right of  $z = +1$ .

(D) Students added the 50% below the mean to the 34% above the mean but below  $z = +1$ .

(E) Ask students what they think the correct answer should be. They often have interesting ideas about the context/question.

by Murphy, McKnight, Richman, and Terry

STT.01.03.060

CC HZ MA207 F09: 0/**63**/26/0/11 time 1:30  
AS DH MA3321 Su12: 0/**100**/0/0/0 time 2:10  
AS DH MA1333 010 F12: 0/**75**/25/0/0 time 4:00  
AS DH MA1333 020 F12: 0/**92**/0/8/0 time 4:00  
AS DH 1333 020 S14: 4/**63**/4/11/19 time 2:20 ,  
AS DH 1333 010 F14: 0/**97**/0/3/0 time 3:40 ,  
AS DH 3321 010 F14: 0/**29**/71/0/0 time 2:30 ,  
AS DH 1333 020 S15: 0/**95**/0/5/0 time 3:00 ,

8. A colleague has collected 1000 old VW vans for resale. The colleague, an old stats professor, will only sell a van to those who can answer the following question: The  $-2$  SD sales price for one of these vans is set at \$550; and  $+2$  SD sales price is set at \$1100. He tells you the distribution of sales prices is approximately normal. What is the expected number of vans for sale between \$550 and \$1100?
- (a) 500
  - (b) 680
  - (c) 750
  - (d) 888
  - (e) 950

*Answer: (e).* (A), (C), (D) These answers all come from using Chebyshev's rule.

(B) Students correctly used the Empirical Rule, but incorrectly used 1 SD instead of 2.

(E)\* correct Since the distribution is approximately normal, students should use the Empirical Rule, which states that approximately 95

by Murphy, McKnight, Richman, and Terry

STT.01.03.070

CC HZ MA207 F09: 0/14/10/0/**76** time 1:50  
AS DH MA3321 Su12: 0/8/8/0/**85** time 1:40  
AS DH MA1333 010 F12: 0/7/19/13/**63** time 2:10  
AS DH MA1333 020 F12: 8/0/8/8/**77** time 3:30  
AS DH 1333 010 S13: 0/17/25/8/**50** time 4:00  
AS DH 1333 020 S14: 0/7/7/7/**78** time 3:00 ,  
AS DH 1333 010 F14: 0/0/21/12/**67** time 3:10 ,

9. The heights of women are normally distributed with a mean of 65 inches and an SD of 2.5 inches. The heights of men are also normal with a mean of 70 inches. What percent of women are taller than a man of average height?

- (a) 0.15%
- (b) 2.5%
- (c) 5%
- (d) 16%
- (e) insufficient information

*Answer: (b).* (A) Students incorrectly used 3 SD instead of 2, but did look only at one tail.

(B)\* correct Students must recognize the need to use the 68-95-99.7 rule (normally distributed data) and the need to convert to a z-score.

Step 1: How many SD units is a male height of 70 above the female average height of 65?  $\frac{70-65}{2.5} = +2$  SD.

Step 2. What is the percentage of women taller than 70 ( $z = +2$ )? Using the 68-95-99.7 rule, we recognize that 95% fall between  $\pm 2$  SD, with 2.5% in each tail. So 2.5% of all women are taller than 70 inches (the average male height).

(C) Students correctly used 2 SD, but did not look only at the upper tail.

(D) Students incorrectly used 1 SD instead of 2, but did look only at one tail.

(E) Students may believe that they need the SD for the distribution of mens heights to answer the question.

by Murphy, McKnight, Richman, and Terry

STT.01.03.080

CC HZ MA207 F09: 14/**62**/19/0/0 time 2:00

CC KC MA207 F09: 4/**71**/14/7/0 time 3:00

AS DH MA3321 Su12: 0/**92**/8/0/0 time 2:40

AS DH MA1333 010 F12: 0/**81**/13/0/6 time 2:00

AS DH MA1333 020 F12: 0/**77**/23/0/0 time 2:20

AS DH 1333 010 S13: 6/**88**/6/0/0 time 3:10

AS DH 1333 020 S14: 12/**53**/35/0/0 time 2:50 ,

AS DH 3321 010 S14: 0/**77**/15/8/0 time 3:20 ,

AS DH 1333 010 F14: 0/**44**/50/6/0 time 4:30 ,

AS DH 1333 020 S15: 0/**50**/25/0/25 time 2:50 ,

AS DH 3321 010 S15: 0/**100**/0/0/0 time 2:30 ,

10. Many psychological disorders (e.g. Depression, ADHD) are based on the application of the 2 SD rule assuming a normal distribution of reported symptoms. This means that anyone who reports a symptom count that is greater than the 2 SD point in a normal population can be considered to be “abnormal” or “disordered”.

Given this definition of “disorder”, what is expected prevalence rate of these disorders based on the 2 SD rule?

- (a) 0.15%
- (b) 2.5%
- (c) 5%
- (d) 16%
- (e) 95%

*Answer: (b).* (A) Students incorrectly used 3 SD instead of 2 (99.7% of the data), but did look only at one tail.

(B)\* correct This is a straightforward application of the Empirical Rule. If a person is given a positive diagnosis ONLY if their reported symptoms exceed 2 SD, then we can expect 1.2 of 5% to be positively diagnosed. Hence, the correct answer is 2.5%.

(C) Students correctly used 2 SD, but incorrectly did not look only at the upper tail.

(D) Students incorrectly used 1 SD instead of 2 (68% of the data), but did look only at one tail.

(E) Students recognize that 95% of the data are expected to fall within 2 SD of the mean, but do not understand that the question is asking for the proportion in the upper tail.

by Murphy, McKnight, Richman, and Terry

STT.01.03.090

CC HZ MA207 F09: 15/**15**/55/0/10 time 1:30

AS DH MA3321 Su12: 0/**75**/25/0/0 time 1:30

AS DH 1333 010 S13: 0/**46**/15/23/15 time 3:00

AS DH 1333 020 S14: 0/**72**/21/0/7 time 3:00 ,

AS DH 1333 010 F14: 0/**43**/26/13/17 time 3:20 ,

11. The ACT has a mean of 21 and an SD of 5. The SAT has a mean of 1000 and a SD of 200. Joe Bob Keith took the ACT and he needs a score of 1300 on the SAT to get into UNC-Chapel Hill and a score of 1400 on the SAT to get into Duke. UNC and Duke both told Joe Bob Keith that they will convert the ACT to the SAT using a z-score (or standard-score) transformation. Joe Bob Keith has decided to go to the school with the highest standards that will accept him. If he doesn't qualify for either Duke or UNC, then it's Faber College for Joe Bob Keith. As it turns out, Joe Bob Keith got a 30 on the ACT, but he cannot figure out what that means for his choice of college. Help Joe Bob Keith out. Where is he going to school?
- (a) UNC
  - (b) Duke
  - (c) Faber



*Answer: (a).* Note: Students need to recognize the need to convert to  $z$ -scores so that a comparable scale can be used.

Joe Bob Keiths ACT  $z$ -score is  $\frac{30-21}{5} = \frac{9}{5} = +1.8$ .

(A)\* correct The  $z$ -score associated with UNC's minimum SAT score is  $\frac{1300-1000}{200} = +1.5$ . Since Joe Bob Keiths  $z$ -score is higher than this minimum, he qualifies for admission to UNC. As calculated in (B), he does not qualify for Duke, so UNC is the best school that will accept him, so that's where he will go.

(B) The  $z$ -score associated with Duke's minimum SAT score is  $\frac{1400-1000}{200} = +2$ . Since Joe Bob Keiths  $z$ -score is lower than this minimum, he does not qualify for admission to Duke.

(C) Duke will accept Joe Bob Keith, but so will UNC so the answer is (A).

by Murphy, McKnight, Richman, and Terry

STT.01.03.100

CC HZ MA207 F09: **71**/24/5 time 4:30

CC KC MA207 F09: **67**/33/0 time 5:00

AS DH MA3321 Su12: **75**/17/8 time 4:20

AS DH MA1333 020 F12: **78**/17/6 time 4:50

12. Let  $Z$  be a standard normal random variable. Which of the following probabilities is the smallest?

- (a)  $P(-2 < Z < -1)$
- (b)  $P(0 < Z < 2)$
- (c)  $P(Z < 1)$
- (d)  $P(Z > 2)$

*Answer: (d).* Looking at the graph of the pdf for  $Z$  quickly narrows the choices down. Adding in the rules of thumb about standard deviations (e.g. 68% of the population lies within one standard deviation of the mean) helps us finalize our answer.

by Derek Bruff

STT.01.03.110

CC KC MA207 F09: 43/11/18/**29** time 3:20

AS DH MA3321 Su12: 25/0/17/**58** time 2:20

AS DH MA1333 010 F12: 42/8/33/**17** time 3:30

AS DH 1333 010 S13: 9/64/23/**5** time 4:30

AS DH 1333 020 S14: 30/4/15/**52** time 4:00 ,

AS DH 1333 010 F14: 6/0/0/**94** time 4:40 ,

AS DH 1333 020 S15: 40/5/0/**55** time 4:00 ,

13. Let  $Z$  be a standard normal random variable. Which of the following probabilities is the smallest?

- (a)  $P(0 \leq Z \leq 2.07)$
- (b)  $P(-0.64 \leq Z \leq -0.11)$
- (c)  $P(Z > -1.06)$
- (d)  $P(Z < -0.88)$

*Answer: (d).* Looking at the graph of the pdf for  $Z$  narrows the choices down to b or d. Using a table of the normal distribution, we can compute these and determine which is smaller.

by Derek Bruff

STT.01.03.120

CC KC MA207 F09: 11/25/18/**46** time 3:30

AS DH MA3321 Su12: 0/79/0/**21** time 4:00

AS DH MA1333 010 F12: 0/45/0/**55** time 3:30

AS DH 1333 010 S13: 0/81/0/**19** time 4:20

AS DH 1333 010 F14: 6/68/6/**19** time 4:00 ,

14. 77% of the area under a normal curve lies to the left of what  $z$ -score?

- (a) 0.74
- (b) 0.77
- (c)  $z_{0.77}$
- (d) 0.78

*Answer: (a).* This question addresses common points of confusion concerning  $z$ -scores and areas.

- (a) Correct.
- (b) 0.77 is simply the area given in the problem.
- (c)  $z_\alpha$  is the  $z$ -score having an area of  $\alpha$  to the *right*. So  $z_{1-0.77} = z_{0.23}$  would be correct.
- (d) This confuses  $z$ -scores and areas. (0.78 is the area to the left of  $z = 0.77$ .)

by David A. Huckaby

STT.01.03.130

15. Intelligence quotients (IQs) are normally distributed with a mean of 100 and a standard deviation of 16. What percentage of the population has an IQ between 112 and 116?

- (a) 4%
- (b) 7%
- (c) 9%
- (d) 25%

*Answer: (b).*

- (a) 4 is simply the difference in the IQs.
- (b) Correct. The  $z$ -scores are 1.00 and 0.75. Subtracting the areas to the left of these  $z$ -scores gives  $0.8413 - 0.7734 = 0.07$  (to two decimal places).
- (c) While “formula-maximalist” students might use  $z = \frac{IQ-100}{16}$  even for  $IQ = 116$ , “formula-minimalist” students might try to reduce every problem like this to an invocation of the empirical rule. The difference in the  $z$ -scores,  $1.00 - 0.75 = 0.25$  is one-eighth of the difference between  $z = -1$  and  $z = 1$ , the  $z$ -scores that are known from the empirical rule to bound an area of about 0.68.  $\frac{0.68}{8} = 0.09$  (to two decimal places). (A quick sketch can quickly dispel the misconception.)
- (d) The difference in the  $z$ -scores is  $1.00 - 0.75 = 0.25$ . (Problems like this are understandably difficult for students who are still unclear on the distinction between  $z$ -scores and areas.)

by David A. Huckaby

STT.01.03.140

16. Intelligence quotients (IQs) are normally distributed with a mean of 100 and a standard deviation of 16. Find  $Q_3$  for IQ.
- (a) 111
  - (b) 112
  - (c) A continuous distribution does not have quartiles.

*Answer: (a).*

- (a) Correct. The  $z$ -score that has an area of 0.75 to its left is  $z = 0.67$ . So  $Q_3 = 100 + 0.67(16) = 111$ .
- (b) This answer is obtained by using the area instead of the  $z$ -score:  $100 + 0.75(16) = 112$ . (Problems like this are understandably difficult for students who are still unclear on the distinction between  $z$ -scores and areas.)
- (c) Some students’ conception of  $Q_3$  is that it is the result of a procedure of finding two medians of a discrete data set. This question can reinforce the idea that  $Q_3$  is ideally the 75th percentile.

by David A. Huckaby

STT.01.03.150

17. Jeannie works at the drive-through window at a local fast-food restaurant. In the middle of the afternoon, the mean time between customers arriving at the window is 5 minutes with a standard deviation of 5 minutes. As a customer drives up to the window, Jeannie is wondering what the probability is that the next customer will arrive more than 10 minutes from now. TRUE or FALSE: Jeannie should convert 10 into a  $z$ -score and then find the area to the right of that  $z$ -score under the standard normal curve.
- (a) True, and I am very confident.
  - (b) True, and I am not very confident.
  - (c) False, and I am not very confident.
  - (d) False, and I am very confident.

*Answer: (d).* The variable “wait time” takes only non-negative values. With a mean and standard deviation of 5, the distribution must therefore be right-skewed. Treating the distribution as if it were normal will produce inaccurate results. [Follow-up question: If Jeannie were to treat the distribution as if it were normal in determining the probability that the next customer will arrive more than 10 minutes from now, will her answer be too high or too low? Why? Answer: Too low. The actual right-skewed distribution has a fatter right tail than a normal distribution has. Another follow-up question: Does anyone—perhaps from another course you took—know a famous distribution that might fit the given parameters? Answer: The exponential distribution.]

by David A. Huckaby

STT.01.03.160

18. Which of the following is not a reason for constructing a normal probability plot or a normal quantile plot?
- (a) You are about to perform an inferential statistics procedure.
  - (b) You have computed a five-number summary and would like to display the results.
  - (c) You are concerned about outliers.
  - (d) You are concerned about skewness.

*Answer: (b).* Some inferential statistics procedures—including many of the procedures learned in a first course—require that the population be normally distributed and that the data have no outliers. A normal probability plot or a normal quantile plot is often constructed prior to performing such an inferential statistics procedure in order to check for normality and for outliers. A *boxplot* displays the results of a five-number summary.

by David A. Huckaby  
STT.01.03.170