Discrete Random Variables

Chapter **4**

4.2 A discrete variable can assume a countable number of values while a continuous random variable can assume values corresponding to any point in one or more intervals.

4.4 a. The amount of flu vaccine in a syringe is measured on an interval, so this is a continuous random variable.

b. The heart rate (number of beats per minute) of an American male is countable, starting at whatever number of beats per minute is necessary for survival up to the maximum of which the heart is capable. That is, if m is the minimum number of beats necessary for survival, x can take on the values (m, $m + 1$, $m + 2$, ...) and is a discrete random variable.

c. The time necessary to complete an exam is continuous as it can take on any value $0 \le x \le L$, where L = limit imposed by instructor (if any).

d. Barometric (atmospheric) pressure can take on any value within physical constraints, so it is a continuous random variable.

e. The number of registered voters who vote in a national election is countable and is therefore discrete.

f. An SAT score can take on only a countable number of outcomes, so it is discrete.

4.6 The values x can assume are 1, 2, 3, 4, or 5. Thus, x is a discrete random variable.

4.8 Since hertz could be any value in an interval, this variable is continuous.

4.10 The number of prior arrests could take on values 0, 1, 2, .. . Thus, x is a discrete random variable.

4.12 Answers will vary. An example of a discrete random variable of interest to a sociologist might be the number of times a person has been married.

4.14 Answers will vary. An example of a discrete random variable of interest to an art historian might be the number of times a piece of art has been restored.

4.16 The expected value of a random variable represents the mean of the probability distribution. You can think of the expected value as the mean value of x in a *very large* (actually, *infinite*) number of repetitions of the experiment.

4.18 For a mound-shaped, symmetric distribution, the probability that x falls in the interval $\mu \pm 2\sigma$ is approximately .95. This follows the probabilities associated with the Empirical Rule.

4.20 a. This is a valid distribution because $\sum p(x) = .2 + .3 + .3 + .2 = 1$ and $p(x) \ge 0$ for all values of x.

b. This is *not* a valid distribution because $\sum p(x) = .25 + .50 + .20 = .95 \neq 1$.

c. This is *not* a valid distribution because one of the probabilities is negative.

d. This is *not* a valid distribution because $\sum p(x) = .15 + .20 + .40 + .35 = 1.10 \neq 1$.

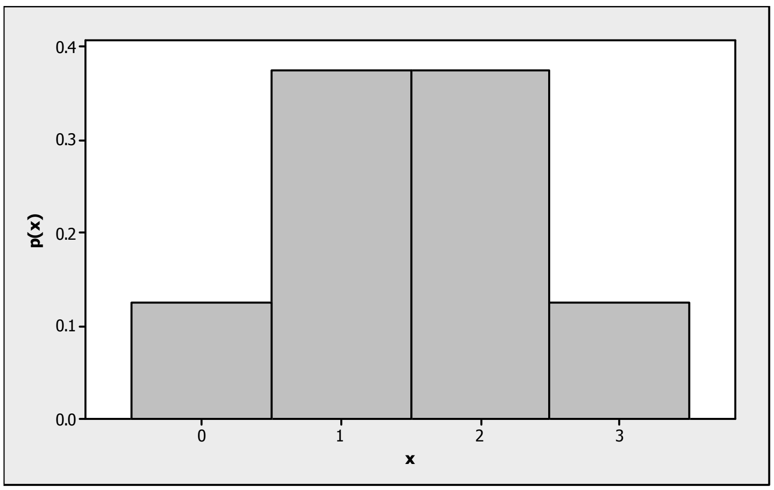
4.22 a. The simple events are (where H = head, T = tail):

	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
x = # heads	3	2	2	2	1	1	1	0

b. If each event is equally likely, then $P(\text{simple event}) = \frac{1}{n} = \frac{1}{8}$

$$p(3) = \frac{1}{8},\ p(2) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8},\ \ p(1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}, \text{ and } p(0) = \frac{1}{8}$$

c.



d. $P(x = 2 \text{ or } x = 3) = p(2) + p(3) = \frac{3}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$

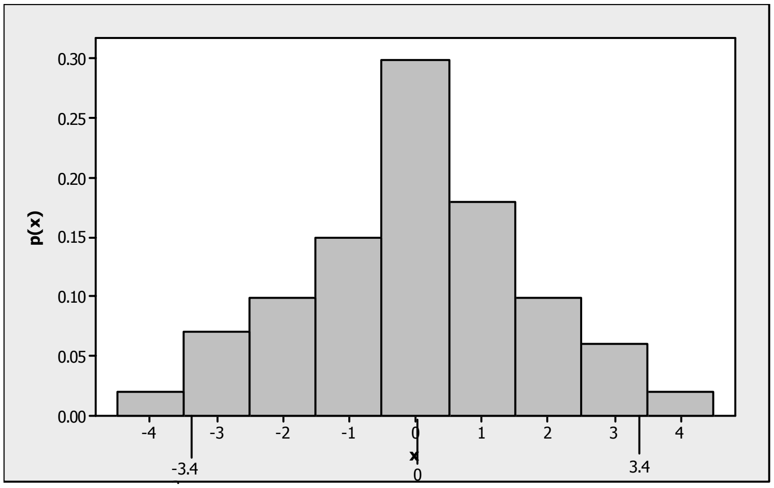
4.24 a.

$$\begin{aligned}\mu = E(x) = \sum xp(x) &= -4(.02) + (-3)(.07) + (-2)(.10) + (-1)(.15) + 0(.3)\\ &\quad +1(.18) + 2(.10) + 3(.06) + 4(.02)\\ &= -.08 - .21 - .2 - .15 + 0 + .18 + .2 + .18 + .08 = 0\end{aligned}$$

$$\begin{aligned}\sigma^2 = E(x-\mu)^2 = \sum (x-\mu)^2 p(x) &= (-4-0)^2(.02) + (-3-0)^2(.07) + (-2-0)^2(.10)\\ &\quad +(-1-0)^2(.15) + (0-0)^2(.30) + (1-0)^2(.18)\\ &\quad +(2-0)^2(.10) + (3-0)^2(.06) + (4-0)^2(.02)\\ &= .32 + .63 + .4 + .15 + 0 + .18 + .4 + .54 + .32 = 2.94\end{aligned}$$

$$\sigma = \sqrt{2.94} = 1.715$$

b.



c. $\mu \pm 2\sigma \Rightarrow 0 \pm 2(1.715) \Rightarrow 0 \pm 3.430 \Rightarrow (-3.430,\ 3.430)$

$$P(-3.430 < x < 3.430) = p(-3) + p(-2) + p(-1) + p(0) + p(1) + p(2) + p(3)$$
$$= .07 + .10 + .15 + .30 + .18 + .10 + .06 = .96$$

4.26 a. To find the probabilities, we divide the percents by 100. In tabular form, the probability distribution for x, the driver-side star rating, is:

x	2	3	4	5
$p(x)$	.0408	.1735	.6020	.1837

b. $P(x = 5) = .1837$

c. $P(x \leq 2) = P(x = 2) = .0408$

d. $E(x) = \sum xp(x) = 2(.0408) + 3(.1735) + 4(.6020) + 5(.1837) = 3.9286$
The average driver-side star rating is 3.9286.

4.28 a. To find probabilities, change the percents given in the table to proportions by dividing by 100. The probability distribution for x is:

x	1	2	3	4
$p(x)$	.40	.54	.02	.04

b. $P(x \geq 3) = P(x = 3) + P(x = 4) = .02 + .04 = .06.$

c. $E(x) = \sum xp(x) = 1(.40) + 2(.54) + 3(.02) + 4(.04) = 1.7$. The average number of delphacid eggs per blade is 1.7.

4.30 a. The number of solar energy cells out of 5 that are manufactured in China, x, can take on values 0, 1, 2, 3, or 5. Thus, x is a discrete random variable.

b. $$p(0)=\frac{(5!)(.35)^0(.65)^{5-0}}{(0!)(5-0)!}=\frac{(5)(4)(3)(2)(1)(.65)^5}{1(5)(4)(3)(2)(1)}=.1160$$

$$p(1)=\frac{(5!)(.35)^1(.65)^{5-1}}{(1!)(5-1)!}=\frac{(5)(4)(3)(2)(1)(.35)(.65)^4}{1(4)(3)(2)(1)}=.3124$$

$$p(2)=\frac{(5!)(.35)^2(.65)^{5-2}}{(2!)(5-2)!}=\frac{(5)(4)(3)(2)(1)(.35)^2(.65)^3}{(2)(1)(3)(2)(1)}=.3364$$

$$p(3)=\frac{(5!)(.35)^3(.65)^{5-3}}{(3!)(5-3)!}=\frac{(5)(4)(3)(2)(1)(.35)^3(.65)^2}{(3)(2)(1)(2)(1)}=.1811$$

$$p(4)=\frac{(5!)(.35)^4(.65)^{5-4}}{(4!)(5-4)!}=\frac{(5)(4)(3)(2)(1)(.35)^4(.65)}{(4)(3)(2)(1)(1)}=.0488$$

$$p(5)=\frac{(5!)(.35)^5(.65)^{5-5}}{(5!)(5-5)!}=\frac{(5)(4)(3)(2)(1)(.35)^5}{(5)(4)(3)(2)(1)(1)}=.0053$$

c. The properties of a discrete probability distribution are: $0 \le p(x) \le 1$ for all values of x and $\sum p(x)=1$. All of the probabilities in part b are greater than 0. The sum of the probabilities is $\sum p(x) = .1160+.3124+.3364+.1811+.0488+.0053=1$.

d. $P(x \ge 4)=p(4)+p(5)=.0488+.0053=.0541$

4.32 The probability distribution of x is:

$p(8.5)=.000123+.008823+.128030+.325213=.462189$

$p(9.0)=.000456+.020086+.153044+.115178=.288764$

$p(9.5)=.001257+.032014+.108400=.141671$

$p(10.0)=.002514+.032901+.034552=.069967$

$p(10.5)=.003561+.021675=.025236$

$p(11.0)=.003401+.006284+.001972=.011657$

$p(12.0)=.000518$

4.34 Let x = winnings in the Florida lottery. The probability distribution for x is:

x	$p(x)$
−\$1	22,999,999/23,000,000
\$6,999,999	1/23,000,000

The expected net winnings would be:

$$\mu = E(x) = (-1)(22{,}999{,}999/23{,}000{,}000) + 6{,}999{,}999(1/23{,}000{,}000) = -\$.70$$

The average winnings of all those who play the lottery is −\$.70.

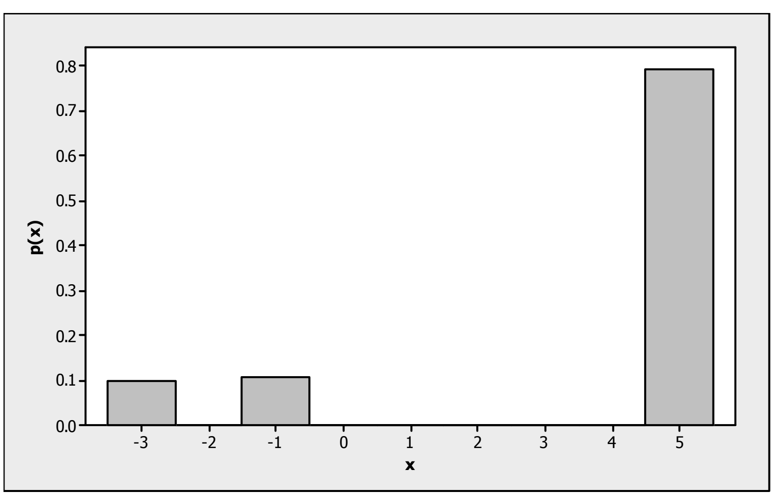
4.36 Assigning points according to the directions is:

OUTCOME OF APPEAL	**Number of cases**	**Points awarded, x**
Plaintiff trial win – reversed	71	−1
Plaintiff trial win – affirmed/dismissed	240	5
Defendant trial win – reversed	68	−3
Defendant trial win – affirmed/dismissed	299	5
TOTAL	678	

To find the probabilities for x, we divide the frequencies by the total sample size. The probability distribution for x is:

x	**p(x)**
−3	68 / 678 = .100
−1	71 / 678 = .105
5	(240+299)/678 = .795
TOTAL	1.000

Using MINITAB, the graph of the probability distribution is:



4.38 a. Since there are 20 possible outcomes that are all equally likely, the probability of any of the 20 numbers is 1/20. The probability distribution of x is:

$P(x = 5) = 1/20 = .05$; $P(x = 10) = 1/20 = .05$; etc.

x	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
$p(x)$	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05

b.
$$
\begin{aligned}
E(x) &= \sum xp(x) = (5-52.5)^2(.05) + (10-52.5)^2(.05) + 15(.05) + 20(.05) \\
&\quad + 25(.05) + 30(.05) + 35(.05) + 40(.05) + 45(.05) + 50(.05) + 55(.05) + 60(.05) \\
&\quad + 65(.05) + 70(.05) + 75(.05) + 80(.05) + 85(.05) + 90(.05) + 95(.05) + 100(.05) \\
&= 52.5
\end{aligned}
$$

c.
$$
\begin{aligned}
\sigma^2 &= E(x-\mu)^2 = \sum (x-\mu)^2 p(x) = (5-52.5)^2(.05) + (10-52.5)^2(.05) \\
&\quad + (15-52.5)^2(.05) + (20-52.5)^2(.05) + (25-52.5)^2(.05) + (30-52.5)^2(.05) \\
&\quad + (35-52.5)^2(.05) + (40-52.5)^2(.05) + (45-52.5)^2(.05) + (50-52.5)^2(.05) \\
&\quad + (55-52.5)^2(.05) + (60-52.5)^2(.05) + (65-52.5)^2(.05) + (70-52.5)^2(.05) \\
&\quad + (75-52.5)^2(.05) + (80-52.5)^2(.05) + (85-52.5)^2(.05) + (90-52.5)^2(.05) \\
&\quad + (95-52.5)^2(.05) + (100-52.5)^2(.05) \\
&= 831.25
\end{aligned}
$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{831.25} = 28.83$$

Since the uniform distribution is not mound-shaped, we will use Chebyshev's theorem to describe the data. We know that at least 8/9 of the observations will fall with 3 standard deviations of the mean and at least 3/4 of the observations will fall within 2 standard deviations of the mean. For this problem,

$\mu \pm 2\sigma \Rightarrow 52.5 \pm 2(28.83) \Rightarrow 52.5 \pm 57.66 \Rightarrow (-5.16,\ 110.16)$. Thus, at least 3/4 of the data will fall between −5.16 and 110.16. For our problem, all of the observations will fall within 2 standard deviations of the mean. Thus, x is just as likely to fall within any interval of equal length.

d. If a player spins the wheel twice, the total number of outcomes will be 20(20) = 400. The sample space is:

5, 5	10, 5	15, 5	20, 5	25, 5...	100, 5
5,10	10,10	15,10	20,10	25,10...	100,10
5,15	10,15	15,15	20,15	25,15...	100,15
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
5,100	10,100	15,100	20,100	25,100...	100,100

Each of these outcomes are equally likely, so each has a probability of 1/400 = .0025.

Now, let x equal the sum of the two numbers in each sample. There is one sample with a sum of 10, two samples with a sum of 15, three samples with a sum of 20, etc. If the sum of the two numbers exceeds 100, then x is zero. The probability distribution of x is:

x	$p(x)$	x	$p(x)$
0	.5250	55	.0250
10	.0025	60	.0275
15	.0050	65	.0300
20	.0075	70	.0325
25	.0100	75	.0350
30	.0125	80	.0375
35	.0150	85	.0400
40	.0175	90	.0425
45	.0200	95	.0450
50	.0225	100	.0475

e. We assumed that the wheel is fair, or that all outcomes are equally likely.

f. $$\mu = E(x) = \sum xp(x) = 0(.5250) + 10(.0025) + 15(.0050) + 20(.0075) + \ldots + 100(.0475)$$
$$= 33.25$$

$$\sigma^2 = E(x-\mu)^2 = \sum (x-\mu)^2 p(x) = (0-33.25)^2(.525) + (10-33.25)^2(.0025)$$
$$+ (15-33.25)^2(.0050) + (20-33.25)^2(.0075) + \ldots + (100-33.25)^2(.0475)$$
$$= 1471.3125$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{1471.3125} = 38.3577$$

g. $P(x = 0) = .525$

h. Given that the player obtains a 20 on the first spin, the possible values for x (sum of the two spins) are 0 (player spins 85, 90, 95, or 100 on the second spin), 25, 30, ..., 100. In order to get an x of 25, the player would spin a 5 on the second spin. Similarly, the player would have to spin a 10 on the second spin order to get an x of 30, etc. Since all of the outcomes are equally likely on the second spin, the distribution of x is:

x	$p(x)$	x	$p(x)$
0	.20	65	.05
25	.05	70	.05
30	.05	75	.05
35	.05	80	.05
40	.05	85	.05
45	.05	90	.05
50	.05	95	.05
55	.05	100	.05
60	.05		

i. The probability that the players total score will exceed one dollar is the probability that x is zero. $P(x = 0) = .20$

j. Given that the player obtains a 65 on the first spin, the possible values for x (sum of the two spins) are 0 (player spins 40, 45, 50, up to 100 on second spin), 70, 75, 80,..., 100. In order to get an x of 70, the player would spin a 5 on the second spin. Similarly, the player would have to spin a 10 on the second spin in order to get an x of 75, etc. Since all of the outcomes are equally likely on the second spin, the distribution of x is:

x	$p(x)$
0	.65
70	.05
75	.05
80	.05
85	.05
90	.05
95	.05
100	.05

The probability that the players total score will exceed one dollar is the probability that x is zero. $P(x = 0) = .65$.

k. There are many possible answers. Notice that $P(x = 0)$ on the second spin is equal to the value on the first spin divided by 100. If a player got a 20 on the first spin, then $P(x = 0)$ on the second spin is $20/100 = .20$. If a player got a 65 on the first spin, then $P(x = 0)$ on the second spin is $65/100 = .65$.

4.40 Suppose we define the following events:

A: {Child has an attached earlobe}
N: {Child does not have an attached earlobe}

From the graph, $P(A) = 1/4 = .25$. Thus, $P(N) = 1 - P(A) = 1 - .25 = .75$

If seven children are selected, there will be $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^7 = 128$ possible outcomes for the 7 children. Of these outcomes, there is one that has all *A*'s, 7 that have 6 *A*'s and 1 *N*, 21 that have 5 *A*'s and 2 *N*'s, 35 ways to get 4 *A*'s and 3 *N*'s, 35 ways to get 3 *A*'s and 4 *N*'s, 21 ways to get 2 *A*'s and 5 *N*'s, 7 ways to get 1 *A* and 6 *N* 's, and 1 way to get 0 *A*'s and 7 *N*'s. The list of the outcomes is:

AAAAAAA	*NAAANAA*	*AAANNNA*	*NNNANAA*	*NAANNAN*	*NNANNAN*
AAAAAAN	*AANNAAA*	*AANANNA*	*NNANNAA*	*ANANNAN*	*NANNNAN*
AAAAANA	*ANANAAA*	*ANAANNA*	*NANNNAA*	*AANNNAN*	*ANNNNAN*
AAAANAA	*NAANAAA*	*NAAANNA*	*ANNNNAA*	*NNAAANN*	*NNNAANN*
AAANAAA	*ANNAAAA*	*AANNANA*	*NNNAANA*	*NANAANN*	*NNANANN*
AANAAAA	*NANAAAA*	*ANANANA*	*NNANANA*	*ANNAANN*	*NANNANN*
ANAAAAA	*NNAAAAA*	*NAANANA*	*NANNANA*	*NAANANN*	*ANNNANN*
NAAAAAA	*AAAANNN*	*ANNAANA*	*ANNNANA*	*ANANANN*	*NNAANNN*
AAAAANN	*AAANANN*	*NANAANA*	*NNAANNA*	*AANNANN*	*NANANNN*
AAAANAN	*AANAANN*	*NNAAANA*	*NANANNA*	*NAAANNN*	*ANNANNN*
AAANAAN	*ANAAANN*	*AANNNAA*	*ANNANNA*	*ANAANNN*	*NAANNNN*
AANAAAN	*NAAAANN*	*ANANNAA*	*NAANNNA*	*AANANNN*	*ANANNNN*
ANAAAAN	*AAANNAN*	*NAANNAA*	*ANANNNA*	*AAANNNN*	*AANNNNN*
NAAAAAN	*AANANAN*	*ANNANAA*	*AANNNNA*	*NNNNNAA*	*ANNNNNN*
AAAANNA	*ANAANAN*	*NANANAA*	*NNNAAAN*	*NNNNANA*	*NANNNNN*
AAANANA	*NAAANAN*	*NNAANAA*	*NNANAAN*	*NNNANNA*	*NNANNNN*
AANAANA	*AANNAAN*	*ANNNAAA*	*NANNAAN*	*NNANNNA*	*NNNANNN*
ANAAANA	*ANANAAN*	*NANNAAA*	*ANNNAAN*	*NANNNNA*	*NNNNANN*
NAAAANA	*NAANAAN*	*NNANAAA*	*NNAANAN*	*ANNNNNA*	*NNNNNAN*
AAAANNA	*ANNAAAN*	*NNNAAAA*	*NANANAN*	*NNNNAAN*	*NNNNNNA*
AANANAA	*NANAAAN*	*NNNNAAA*	*ANNANAN*	*NNNANAN*	*NNNNNNN*
ANAANAA	*NNAAAAN*				

$P(AAAAAAA) = .25^7 = .000061035 = .0001 = P(x = 7)$

$P(AAAAAAN) = (.25)^6(.75) = .000183105$. Since there are 7 ways to select 6 *A*'s and 1 *N*, $P(x = 6) = 7(.000183105) = .0013$

$P(AAAAANN) = (.25)^5(.75)^2 = .000549316$. Since there are 21 ways to select 5 *A*'s and 2 *N*'s, $P(x = 5) = 21(.000549316) = .0115$

$P(AAAANNN) = (.25)^4(.75)^3 = .001647949$. Since there are 35 ways to select 4 *A*'s and 3 *N*'s, $P(x = 4) = 35(.001647979) = .0577$

$P(AAANNNN) = (.25)^3(.75)^4 = .004943847$. Since there are 35 ways to select 3 *A*'s and 4 *N*'s, $P(x = 3) = 35(.004943847) = .1730$

$P(AANNNNN) = (.25)^2(.75)^5 = .014831542$. Since there are 21 ways to select 2 *A*'s and 5 *N*'s, $P(x = 2) = 21(.014831542) = .3115$

$P(ANNNNNN) = (.25)(.75)^6 = .044494628$. Since there are 7 ways to select 1 *A*'s and 6 *N*'s, $P(x = 1) = 7(.044494628) = .3115$

$P(NNNNNNN) = (.75)^7 = .133483886$. Since there is 1 way to select 0 *A*'s and 7 *N*'s, $P(x = 0) = .1335$

4.42 The five characteristics of a binomial random variable are:

a. The experiment consists of n identical trials.

b. There are only two possible outcomes on each trial. We will denote one outcome by S (for Success) and the other by F (for Failure).

c. The probability of S remains the same from trial to trial. This probability is denoted by p, and the probability of F is denoted by q. Note that $q = 1 - p$.

d. The trials are independent.

e. The binomial random variable x is the number of S's in n trials.

4.44 a. There are $n = 5$ trials.

b. The value of p is $p = .7$.

4.46 a. $P(x=1)=\dfrac{5!}{1!4!}(.2)^1(.8)^4=\dfrac{5\cdot4\cdot3\cdot2\cdot1}{(1)(4\cdot3\cdot2\cdot1)}(.2)^1(.8)^4=5(.2)^1(.8)^4=.4096$

b. $P(x=2)=\dfrac{4!}{2!2!}(.6)^2(.4)^2=\dfrac{4\cdot3\cdot2\cdot1}{(2\cdot1)(2\cdot1)}(.6)^2(.4)^2=6(.6)^2(.4)^2=.3456$

c. $P(x=0)=\dfrac{3!}{0!3!}(.7)^0(.3)^3=\dfrac{3\cdot2\cdot1}{(1)(3\cdot2\cdot1)}(.7)^0(.3)^3=1(.7)^0(.3)^3=.027$

d. $P(x=3)=\dfrac{5!}{3!2!}(.1)^3(.9)^2=\dfrac{5\cdot4\cdot3\cdot2\cdot1}{(3\cdot2\cdot1)(2\cdot1)}(.1)^3(.9)^2=10(.1)^3(.9)^2=.0081$

e. $P(x=2)=\dfrac{4!}{2!2!}(.4)^2(.6)^2=\dfrac{4\cdot3\cdot2\cdot1}{(2\cdot1)(2\cdot1)}(.4)^2(.6)^2=6(.4)^2(.6)^2=.3456$

f. $P(x=1)=\dfrac{3!}{1!2!}(.9)^1(.1)^2=\dfrac{3\cdot2\cdot1}{(1)(2\cdot1)}(.9)^1(.1)^2=3(.9)^1(.1)^2=.027$

4.48 a. $P(x = 2) = P(x \le 2) - P(x \le 1) = .167 - .046 = .121$ (from Table II, Appendix A)

b. $P(x \le 5) = .034$

c. $P(x > 1) = 1 - P(x \le 1) = 1 - .919 = .081$

4.50 a. The simple events listed below are all equally likely, implying a probability of 1/32 for each. The list is in a regular pattern such that the first simple event would yield $x = 0$, the next five yield $x = 1$, the next ten yield $x = 2$, the next ten also yield $x = 3$, the next five yield $x = 4$, and the final one yields $x = 5$. The resulting probability distribution is given below the simple events.

$$\begin{bmatrix} FFFFF, FFFFS, FFFSF, FFSFF, FSFFF, SFFFF, FFFSS, FFSFS \\ FSFFS, SFFFS, FFSSF, FSFSF, SFFSF, FSSFF, SFSFF, SSFFF \\ FFSSS, FSFSS, SFFSS, FSSFS, SFSFS, SSFFS, FSSSF, SFSSF \\ SSFSF, SSSFF, FSSSS, SFSSS, SSFSS, SSSFS, SSSSF, SSSSS \end{bmatrix}$$

x	0	1	2	3	4	5
$p(x)$	1/32	5/32	10/32	10/32	5/32	1/32

b. $$P(x=0)=\binom{5}{0}(.5)^0(.5)^{5-0}=\frac{5!}{0!5!}(.5)^0(.5)^5=\frac{5\cdot4\cdot3\cdot2\cdot1}{1\cdot5\cdot4\cdot3\cdot2\cdot1}(1)(.5)^5=.03125$$

$$P(x=1)=\binom{5}{1}(.5)^1(.5)^{5-1}=\frac{5!}{1!4!}(.5)^1(.5)^4=\frac{5\cdot4\cdot3\cdot2\cdot1}{1\cdot4\cdot3\cdot2\cdot1}(.5)^5=.15625$$

$$P(x=2)=\binom{5}{2}(.5)^2(.5)^{5-2}=\frac{5!}{2!3!}(.5)^2(.5)^3=\frac{5\cdot4\cdot3\cdot2\cdot1}{2\cdot1\cdot3\cdot2\cdot1}(.5)^5=.3125$$

$$P(x=3)=\binom{5}{3}(.5)^3(.5)^{5-3}=\frac{5!}{3!2!}(.5)^3(.5)^2=\frac{5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot2\cdot1}(.5)^5=.3125$$

$$P(x=4)=\binom{5}{4}(.5)^4(.5)^{5-4}=\frac{5!}{4!1!}(.5)^4(.5)^1=\frac{5\cdot4\cdot3\cdot2\cdot1}{4\cdot3\cdot2\cdot1\cdot1}(.5)^5=.15625$$

$$P(x=5)=\binom{5}{5}(.5)^5(.5)^{5-5}=\frac{5!}{5!0!}(.5)^5(.5)^0=\frac{5\cdot4\cdot3\cdot2\cdot1}{5\cdot4\cdot3\cdot2\cdot1\cdot1}(.5)^5=.03125$$

4.52 a. For this exercise, a success will be an owner acquiring his/her next dog or cat from a shelter.

b. For this exercise, $n = 10$.

c. For this exercise, $p = .5$.

d. Using Table II, Appendix A, with $n = 10$ and $p = .5$,

$$P(x=7)=P(x\le7)-P(x\le6)=.945-.828=.117$$

e. Using Table II, Appendix A, with n = 10 and p = .5, $P(x \leq 3) = .172$

f. Using Table II, Appendix A, with n = 10 and p = .5,

$P(x > 8) = 1 - P(x \leq 8) = 1 - .989 = .011$

4.54 a. We will check the characteristics of a binomial random variable:

1. This experiment consists of n = 5 identical trials.

2. There are only 2 possible outcomes for each trial. A brand of bottled water can use tap water (S) or not (F).

3. The probability of S remains the same from trial to trial. In this case, $p = P(S) \approx .25$ for each trial.

4. The trials are independent. Since there are a finite number of brands of bottled water, the trials are not exactly independent. However, since the number of brands of bottled water is large compared to the sample size of 5, the trials are close enough to being independent.

5. x = number of brands of bottled water using tap water in 5 trials.

b. The formula for finding the binomial probabilities is:

$p(x) = \binom{5}{x} .25^x (.75)^{5-x}$ for x = 0, 1, 2, 3, 4, 5

c. $P(x = 2) = p(2) = \binom{5}{2} .25^2 (.75)^{5-2} = \frac{5!}{2!3!} .25^2 (.75)^3 = .2637$

d. $P(x \leq 1) = p(0) + p(1) = \binom{5}{0} .25^0 (.75)^{5-0} + \binom{5}{1} .25^1 (.75)^{5-1}$

$= \frac{5!}{0!5!} .25^0 (.75)^5 + \frac{5!}{1!4!} .25^1 (.75)^4 = .2373 + .3955 = .6328$

4.56 a. Let x = number of births in 1,000 that take place by Caesarian section.
$E(x) = np = 1000(.32) = 320.$

b. $\sigma = \sqrt{npq} = \sqrt{1000(.32)(.68)} = 14.7513$

c. Since p is not real small, the distribution of x will be fairly mound-shaped, so the Empirical Rule will apply. We know that approximately 95% of the observations will fall within 2 standard deviations of the mean. Thus,

$\mu \pm 2\sigma \Rightarrow 320 \pm 2(14.7513) \Rightarrow 320 \pm 29.5026 \Rightarrow (290.4974,\ \ 349.5026)$

In a sample of 1000 births, we would expect that somewhere between 291 and 349 will be Caesarian section births.

4.58 a. Let x = number of students initially answering question correctly. Then x is a binomial random variable with n = 20 and p = .5. Using Table II, Appendix A,

$$P(x > 10) = 1 - P(x \le 10) = 1 - .588 = .412$$

b. Let x = number of students answering question correctly after feedback. Then x is a binomial random variable with n = 20 and p = .7. Using Table II, Appendix A,

$$P(x > 10) = 1 - P(x \le 10) = 1 - .048 = .952$$

4.60 $\mu = E(x) = np = 50(.6) = 30.0$

$$\sigma = \sqrt{50(.6)(.4)} = 3.4641$$

Since p is not real small, the distribution of x will be fairly mound-shaped, so the Empirical Rule will apply. We know that approximately 95% of the observations will fall within 2 standard deviations of the mean. Thus,

$$\mu \pm 2\sigma \Rightarrow 30 \pm 2(3.4641) \Rightarrow 30 \pm 6.9282 \Rightarrow (23.0718,\quad 36.9282)$$

4.62 Let x = number of slaughtered chickens in 5 that passes inspection with fecal contamination. Then x is a binomial random variable with n = 5 and p = .01 (from Exercise 3.19.)

$P(x \ge 1) = 1 - P(x = 0) = 1 - .951 = .049$ (From Table II, Appendix A).

4.64 Let n = 20, p = .5, and x = number of correct questions in 20 trials. Then x has a binomial distribution. We want to find k such that:

$$P(x \ge k) < .05 \text{ or } 1 - P(x \le k-1) < .05 \Rightarrow P(x \le k-1) > .95 \Rightarrow k - 1 = 14 \Rightarrow k = 15$$
(from Table II, Appendix A)

Note: $P(x \ge 14) = 1 - P(x \le 13) = 1 - .942 = .058$
$P(x \ge 15) = 1 - P(x \le 14) = 1 - .979 = .021$

Thus, to have the probability less than .05, the lowest passing grade should be 15.

4.66 Let x = Number of boys in 24 children. Then x is a binomial random variable with n = 24 and p = .5.

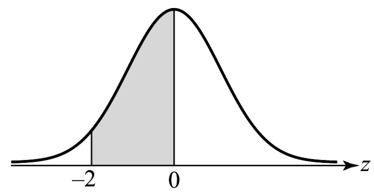
$$\mu = E(x) = np = 24(.5) = 12$$

$$\sigma = \sqrt{npq} = \sqrt{24(.5)(.5)} = \sqrt{6} = 2.4495$$

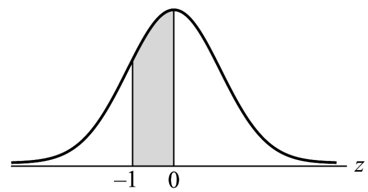
A value of 21 boys out of 24 children would have a z-score of $z = \dfrac{21-12}{2.4495} = 3.67$. A value that is 3.67 standard deviations above the mean would be highly unlikely. Thus, we would agree with the statement, "Rodgers men produce boys."

4.68 If x has a normal distribution with mean μ and standard deviation σ, then the distribution of $z = \dfrac{x-\mu}{\sigma}$ is a normal with a mean of $\mu = 0$ and a standard deviation of $\sigma = 1$.

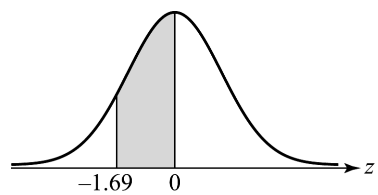
4.70 a. $P(-2.00 < z < 0) = .4772$
(from Table III, Appendix A)



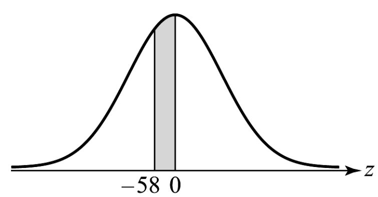
b. $P(-1.00 < z < 0) = .3413$
(from Table III, Appendix A)



c. $P(-1.69 < z < 0) = .4545$
(from Table III, Appendix A)

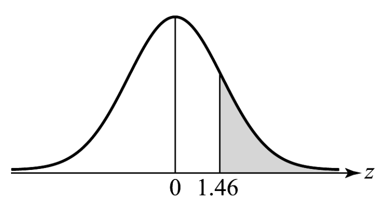


d. $P(-.58 < z < 0) = .2190$
(from Table IV, Appendix A)

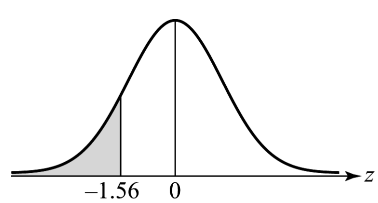


4.72 Using Table III, Appendix A:

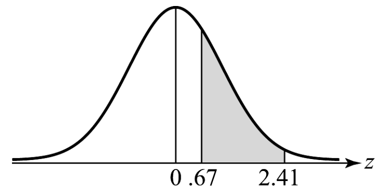
a. $P(z > 1.46) = .5 - P(0 < z \le 1.46)$
$= .5 - .4279 = .0721$



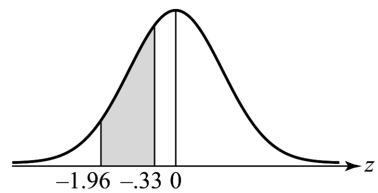
b. $P(z < -1.56) = .5 - P(-1.56 \le z < 0)$
$= .5 - .4406 = .0594$



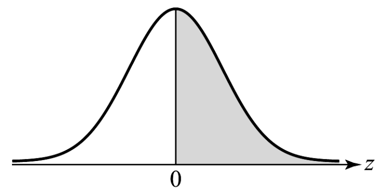
c. $P(.67 \le z \le 2.41)$
$= P(0 < z \le 2.41) - P(0 < z < .67)$
$= .4920 - .2486 = .2434$



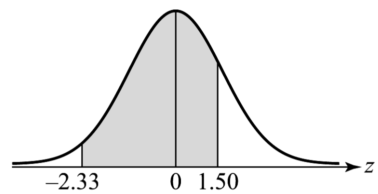
d. $P(-1.96 \le z < -.33)$
$= P(-1.96 \le z < 0) - P(-.33 \le z < 0)$
$= .4750 - .1293 = .3457$



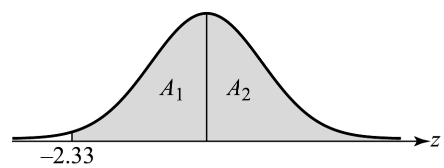
e. $P(z \ge 0) = .5$



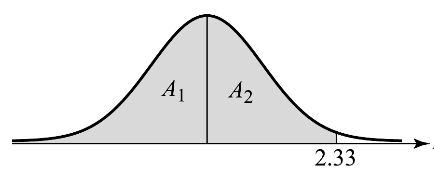
f. $P(-2.33 < z < 1.50)$
$= P(-2.33 < z < 0) + P(0 < z < 1.50)$
$= .4901 + .4332 = .9233$



g. $P(z \ge -2.33) = P(-2.33 \le z \le 0) + P(z \ge 0)$
$= .4901 + .5000$
$= .9901$



h. $P(z < 2.33) = P(z \le 0) + P(0 \le z \le 2.33)$
$= .5000 + .4901$
$= .9901$



4.74 Using the formula $z = \dfrac{x - \mu}{\sigma}$ with $\mu = 25$ and $\sigma = 5$:

a. $z = \dfrac{25 - 25}{5} = 0$

b. $z = \dfrac{30 - 25}{5} = 1$

c. $z = \dfrac{37.5 - 25}{5} = 2.5$

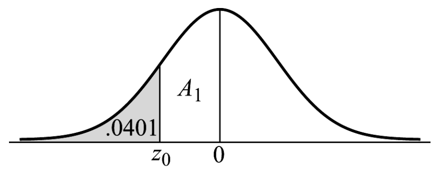
d. $z = \dfrac{10 - 25}{5} = -3$

e. $z = \dfrac{50-25}{5} = 5$

f. $z = \dfrac{32-25}{5} = 1.4$

4.76 Using Table III of Appendix A:

a. $P(z < z_o) = .0401$

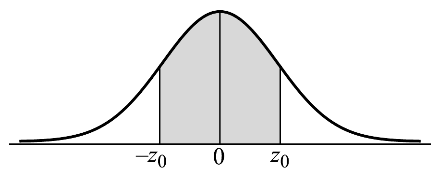


A_1 = .5000 − .0401 = .4591

Look up the area .4591 in the body of Table III;
$z_0 = -1.75$

(z_0 is negative since the graph shows z_0 is on the left side of 0.)

b. $P(-z_o \le z \le z_o) = .95$

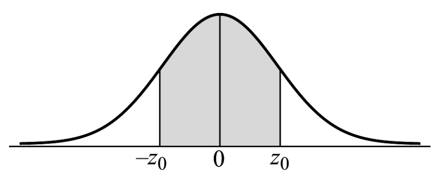


$P(-z_o \le z \le z_o) = 2P(0 \le z \le z_o)$
$2P(0 \le z \le z_o) = .95$

Therefore, $P(0 \le z \le z_o) = .4750$

Look up the area .4750 in the body of Table III; z_0 = 1.96

c. $P(-z_o \le z \le z_o) = .90$

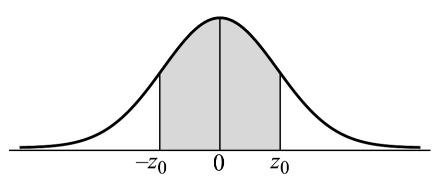


$P(-z_o \le z \le z_o) = 2P(0 \le z \le z_o)$
$2P(0 \le z \le z_o) = .90$

Therefore, $P(0 \le z \le z_o) = .45$

Look up the area .45 in the body of Table III; z_0 = 1.645 (.45 is half way between .4495 and .4505; therefore, we average the z-scores $\dfrac{1.64+1.65}{2} = 1.645$)

d. $P(-z_o \le z \le z_o) = .8740$

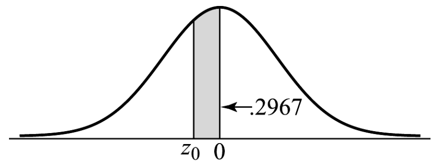


$P(-z_o \le z \le z_o) = 2P(0 \le z \le z_o)$
$2P(0 \le z \le z_o) = .8740$

Therefore, $P(0 \le z \le z_o) = .4370$

Look up the area .4370 in the body of Table III; z_0 = 1.53

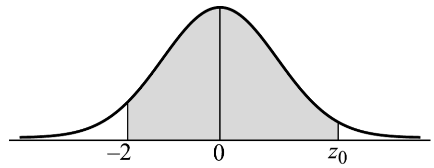
e. $P(-z_o \le z \le 0) = .2967$



$P(-z_o \le z \le 0) = P(0 \le z \le z_o)$

Look up the area .2967 in the body of Table III; z_0 = .83 and -z_0 = -.83

f. $P(-2 < z < z_o) = .9710$



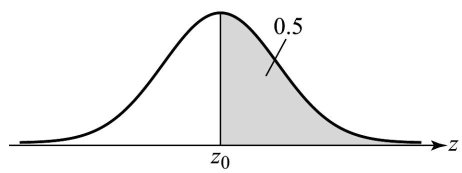
$P(-2 < z < z_o)$
$= P(-2 < z < 0) + P(0 < z < z_0) = .9710$

$P(0 < z < 2) + P(0 < z < z_0) = .9710$

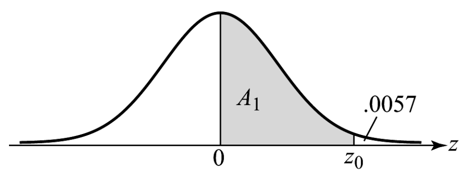
Thus, $P(0 < z < z_0) = .9710 - P(0 < z < 2) = .9710 - .4772 = .4938$

Look up the area .4938 in the body of Table III; z_0 = 2.50

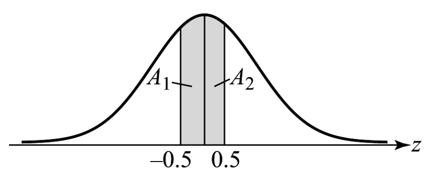
g. $P(z \ge z_0) = .5$
$z_0 = 0$



h. $P(z \ge z_0) = .0057$
A_1 = .5 − .0057 = .4943
Looking up the area .4943 in Table III
gives z_0 = 2.53.

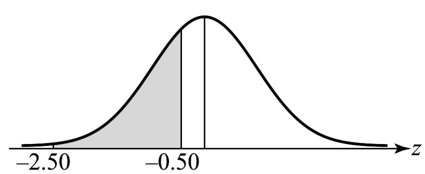


4.78 a. $P(10 \le x \le 12) = P\left(\frac{10-11}{2} \le z \le \frac{12-11}{2}\right)$



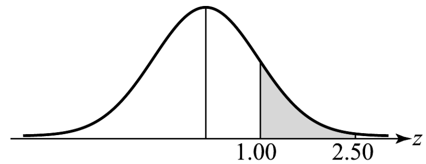
$= P(-0.50 \le z \le 0.50)$
$= A_1 + A_2$
$= .1915 + .1915 = .3830$

b. $P(6 \le x \le 10) = P\left(\frac{6-11}{2} \le z \le \frac{10-11}{2}\right)$



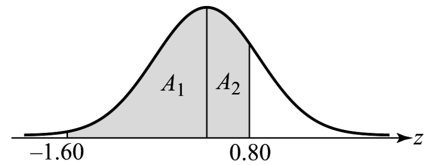
$= P(-2.50 \le z \le -0.50)$
$= P(-2.50 \le z \le 0) - P(-0.50 \le z \le 0)$
$= .4938 - .1915 = .3023$

c. $P(13 \le x \le 16) = P\left(\frac{13-11}{2} \le z \le \frac{16-11}{2}\right)$



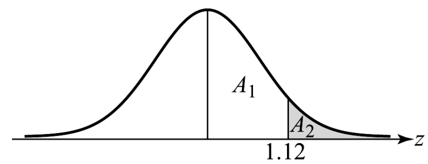
$= P(1.00 \le z \le 2.50)$
$= P(0 \le z \le 2.50) - P(0 \le z \le 1.00)$
$= .4938 - .3413 = .1525$

d. $P(7.8 \le x \le 12.6) = P\left(\frac{7.8-11}{2} \le z \le \frac{12.6-11}{2}\right)$



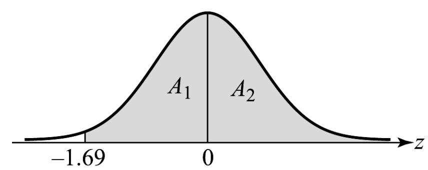
$= P(-1.60 \le z \le 0.80)$
$= A_1 + A_2$
$= .4452 + .2881 = .7333$

e. $P(x \ge 13.24) = P\left(z \ge \frac{13.24-11}{2}\right)$



$= P(z \ge 1.12)$
$= A_2 = .5 - A_1$
$= .5000 - .3868 = .1314$

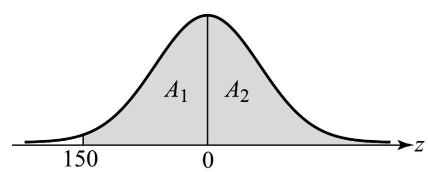
f. $P(x \ge 7.62) = P\left(z \ge \frac{7.62-11}{2}\right)$



$= P(z \ge -1.69)$
$= A_1 + A_2$
$= .4545 + .5000 = .9545$

4.80 The random variable x has a normal distribution with $\sigma = 25$.

We know $P(x > 150) = .90$. So, $A_1 + A_2 = .90$. Since $A_2 = .50$, $A_1 = .90 - .50 = .40$. Look up the area .40 in the body of Table III; (take the closest value) $z_0 = -1.28$.



To find μ, substitute all the values into the z-score formula:

$$z = \frac{x-\mu}{\sigma} \Rightarrow -1.28 = \frac{150-\mu}{25} \Rightarrow \mu = 150 + 25(1.28) = 182$$

4.82 a. $P(x > 120) = P\left(z > \frac{120-105.3}{8}\right) = P(z > 1.84)$

$= .5 - P(0 < z < 1.84) = .5 - .4671 = .0329$

(Using Table III, Appendix A)

b. $P(100 < x < 110) = P\left(\frac{100-105.3}{8} < z < \frac{110-105.3}{8}\right) = P(-.66 < z < .59)$

$= P(-.66 < z < 0) + P(0 < z < .59) = .2454 + .2224 = .4678$

(Using Table III, Appendix A)

c. $P(x < a) = .25 \Rightarrow P\left(z < \frac{a-105.3}{8}\right) = P(z < z_o) = .25$

$A_1 = .5 - .25 = .25$

Looking up the area .25 in Table III gives z_o = -.67.

$z_o = -.67 = \frac{a-105.3}{8} \Rightarrow a = 8(-.67) + 105.3 = 99.94$

4.84 a. Using Table III, Appendix A,

$P(x > 0) = P\left(z > \frac{0-5.26}{10}\right) = P(z > -.53) = .5000 + .2019 = .7019$

b. $P(5 < x < 15) = P\left(\frac{5-5.26}{10} < z < \frac{15-5.26}{10}\right) = P(-.03 < z < .97) = .0120 + .3340 = .3460$

c. $P(x < 1) = P\left(z < \frac{1-5.26}{10}\right) = P(z < -.43) = .5000 - .1664 = .3336$

d. $P(x < -25) = P\left(z < \frac{-25-5.26}{10}\right) = P(z < -3.03) = .5000 - .4988 = .0012$

Since the probability of seeing an average casino win percentage of -25% or smaller after 100 bets on black/red is so small (.0012), we would conclude that either the mean casino win percentage is not 5.26% but something smaller or the standard deviation of 10% is too small.

4.86 a. Let x = carapace length of green sea turtle. Then x has a normal distribution with $\mu = 55.7$ and $\sigma = 11.5$.

$P(x < 40) + P(x > 60) = P\left(z < \frac{40-55.7}{11.5}\right) + P\left(z > \frac{60-55.7}{11.5}\right)$

$= P(z < -1.37) + P(z > .37)$

$= .5 - .4147 + .5 - .1443 = .4410$

b. $P(x > L) = .10 \Rightarrow P\left(z > \frac{L-55.7}{11.5}\right) = P(z > z_o) = .10$

$A_1 = .5 - .10 = .40$

Looking up the area .40 in Table III gives z_o = 1.28.

$$z_o = 1.28 = \frac{L - 55.7}{11.5} \Rightarrow L = 1.28(11.5) + 55.7 = 70.42$$

4.88 a. If the player aims at the right goal post, he will score if the ball is less than 3 feet away from the goal post inside the goal (because the goalie is standing 12 feet from the goal post and can reach 9 feet). Using Table III, Appendix A,

$$P(0 < x < 3) = P\left(0 < z < \frac{3-0}{3}\right) = P(0 < z < 1) = .3413$$

b. If the player aims at the center of the goal, he will be aimed at the goalie. In order to score, the player must place the ball more than 9 feet away from the goalie.
Using Table III, Appendix A

$$P(x < -9) + P(x > 9) = P\left(z < \frac{-9-0}{3}\right) + P\left(z > \frac{9-0}{3}\right)$$
$$= P(z < -3) + P(z > 3) \approx .5 - .5 + .5 - .5 = 0$$

c. If the player aims halfway between the goal post and the goalie's reach, he will be aiming 1.5 feet from the goal post. Therefore, he will score if he hits from 1.5 feet to the left of where he is aiming to 1.5 feet to the right of where he is aiming. Using Table III, Appendix A,

$$P(-1.5 < x < 1.5) = P\left(\frac{-1.5-0}{3} < z < \frac{1.5-0}{3}\right)$$
$$= P(-.5 < z < .5) = .1915 + .1915 = .3830$$

4.90 a. Let x = rating of employee's performance. Then x has a normal distribution with $\mu = 50$ and $\sigma = 15$. The top 10% get "exemplary" ratings.

$$P(x > x_o) = .10 \Rightarrow P\left(z > \frac{x_o - 50}{15}\right) = P(z > z_o) = .10$$

$$A_1 = .5 - .10 = .40$$

Looking up the area .40 in Table III gives z_o = 1.28.

$$z_o = 1.28 = \frac{x_o - 50}{15} \Rightarrow x_o = 1.28(15) + 50 = 69.2$$

b. Only 30% of the employees will get ratings lower than "competent".

$$P(x < x_o) = .30 \Rightarrow P\left(z < \frac{x_o - 50}{15}\right) = P(z < z_o) = .30$$

$$A_1 = .5 - .30 = .20$$

Looking up the area .20 in Table III gives z_o = -.52. The value of z_o is negative because it is in the lower tail.

$$z_o = -.52 = \frac{x_o - 50}{15} \Rightarrow x_o = -.52(15) + 50 = 42.2$$

4.92 a. Using Table III, Appendix A,

$$P(40 < x < 50) = P\left(\frac{40-37.9}{12.4} < z < \frac{50-37.9}{12.4}\right) = P(.17 < z < .98)$$
$$= .3365 - .0675 = .2690.$$

b. Using Table III, Appendix A,

$$P(x < 30) = P\left(z < \frac{30-37.9}{12.4}\right) = P(z < -.64) = .5 - .2389 = .2611 .$$

c. We know that if $P(z_L < z < z_U) = .95$, then $P(z_L < z < 0) + P(0 < z < z_U) = .95$ and

$$P(z_L < z < 0) = P(0 < z < z_U) = .95 / 2 = .4750 .$$

Using Table III, Appendix A, z_U = 1.96 and z_L = -1.96.

$$P(x_L < x < x_U) = .95 \Rightarrow P\left(\frac{x_L - 37.9}{12.4} < z < \frac{x_U - 37.9}{12.4}\right) = .95$$

$$\Rightarrow \frac{x_L - 37.9}{12.4} = -1.96 \quad \text{and} \quad \frac{x_U - 37.9}{12.4} = 1.96$$

$$\Rightarrow x_L - 37.9 = -24.3 \text{ and } x_U - 37.9 = 24.3 \Rightarrow x_L = 13.6 \text{ and } x_U = 62.2$$

d. $P(z > z_0) = .10 \Rightarrow P(0 < z < z_0) = .4000$. Using Table III, Appendix A, z_0 = 1.28.

$$P(x > x_0) = .10 \Rightarrow \frac{x_0 - 37.9}{12.4} = 1.28 \Rightarrow x_0 - 37.9 = 15.9 \Rightarrow x_0 = 53.8 .$$

4.94 a. Let x = fill of container. Using Table III, Appendix A,

$$P(x < 10) = P\left(z < \frac{10-10}{.2}\right) = P(z < 0) = .5$$

b. Profit = Price – cost – reprocessing fee = \$230 – \$20(10.6) – \$10
= \$230 – \$212 – \$10 = \$8.

c. If the probability of underfill is approximately 0, then Profit = Price – Cost.

E(Profit) = E(Price − Cost) = \$230 – E(Cost) = \$230 − \$20$E(x)$ = \$230 − \$20(10.5)

= \$230 − \$210 = \$20.

4.96 Let x = load. From the problem, we know that the distribution of x is normal with a mean of 20.

$$P(10 < x < 30) = P\left(\frac{10-20}{\sigma} < z < \frac{30-20}{\sigma}\right) = P(-z_0 < z < z_0) = .95$$

First, we need to find z_0 such that $P(-z_0 < z < z_0) = .95$. Since we know that the center of the z distribution is 0, half of the area or .95/2 = .475 will be between $-z_0$ and 0 and half will be between 0 and z_0.

We look up .475 in the body of Table III, Appendix A to find z_0 = 1.96.
Thus,

$$\frac{30-20}{\sigma} = 1.96$$
$$\Rightarrow 1.96\sigma = 10 \Rightarrow \sigma = \frac{10}{1.96} = 5.102$$

4.98 Four methods for determining whether the sample data come from a normal population are:

1. Use either a histogram or a stem-and-leaf display for the data and note the shape of the graph. If the data are approximately normal, then the graph will be similar to the normal curve.

2. Compute the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, $\bar{x} \pm 3s$, and determine the percentage of measurements falling in each. If the data are approximately normal, the percentages will be approximately equal to 68%, 95%, and 100%, respectively.

3. Find the interquartile range, IQR, and the standard deviation, s, for the sample, then calculate the ratio IQR / s. If the data are approximately normal, then IQR / $s \approx 1.3$.

4. Construct a normal probability plot for the data. If the data are approximately normal, the points will fall (approximately) on a straight line.

4.100 In a normal probability plot, the observations in a data set are ordered from smallest to largest and then plotted against the expected z-scores of observations calculated under the assumption that the data come from a normal distribution. If the data are normally distributed, a linear or straight-line trend will result.

4.102 a. IQR = $Q_U - Q_L$ = 195 − 72 = 123

b. IQR/s = 123/95 = 1.295

c. Yes. Since IQR is approximately 1.3, this implies that the data are approximately normal.

4.104 a. Using MINITAB, the stem-and-leaf display of the data is:

Stem-and-Leaf Display: Data

```
Stem-and-leaf of Data  N  = 28
Leaf Unit = 0.10

 2   1  16
 3   2  1
 7   3  1235
 11  4  0356
(4)  5  0399
 13  6  03457
 8   7  34
 6   8  2446
 2   9  47
```

The data are somewhat mound-shaped, so the data could be normally distributed.

b. Using MINITAB, the descriptive statistics are:

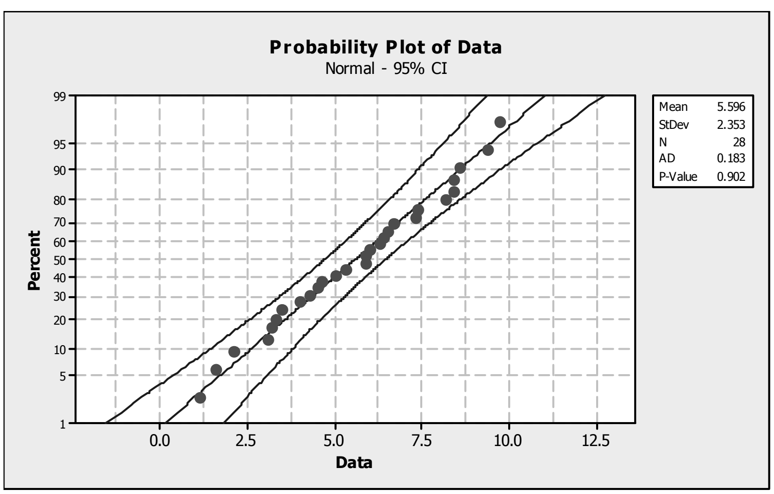
Descriptive Statistics: Data

```
Variable   N   Mean  StDev  Minimum     Q1  Median     Q3  Maximum
Data      28  5.596  2.353    1.100  3.625   5.900  7.375    9.700
```

From the printout, the standard deviation is $s = 2.353$.

c. From the printout, $Q_L = 3.625$ and $Q_U = 7.375$. The interquartile range is $IQR = Q_U - Q_L = 7.375 - 3.625 = 3.75$. If the data are approximately normal, then $IQR / s \approx 1.3$. For this data, $IQR / s = 3.75 / 2.353 = 1.397$. This is fairly close to 1.3, so the data could be normal.

d. Using MINITAB, the normal probability plot is:



Since the data are very close to a straight line, it indicates that the data could be normally distributed.

4.106 The histogram of the data is mound-shaped. It is somewhat skewed to the right, so it is not exactly symmetric. However, it is very close to a mound shaped distribution, so the engineers could use the normal probability distribution to model the behavior of shear strength for rock fractures.

4.108 a. We know that approximately 68% of the observations will fall within 1 standard deviation of the mean, approximately 95% will fall within 2 standard deviations of the mean, and approximately 100% of the observations will fall within 3 standard deviation of the mean. From the printout, the mean is 89.29 and the standard deviation is 3.18.

$\bar{x} \pm s \Rightarrow 89.29 \pm 3.18 \Rightarrow (86.11, 92.47)$. Of the 50 observations, 34 (or 34/50 = .68) fall between 86.11 and 92.47. This is close to what we would expect if the data were normally distributed.

$\bar{x} \pm 2s \Rightarrow 89.29 \pm 2(3.18) \Rightarrow 89.29 \pm 6.36 \Rightarrow (82.93, 95.65)$. Of the 50 observations, 48 (or 48/50 = .96) fall between 82.93 and 95.65. This is close to what we would expect if the data were normally distributed.

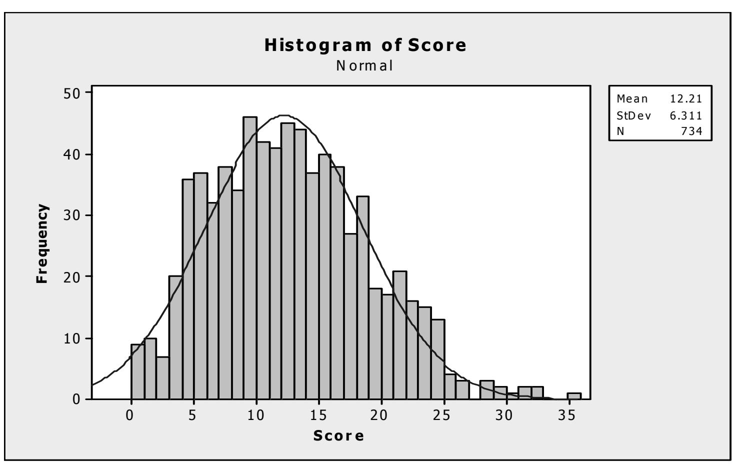
$\bar{x} \pm 3s \Rightarrow 89.29 \pm 3(3.18) \Rightarrow 89.29 \pm 9.54 \Rightarrow (79.75, 98.83)$. Of the 50 observations, 50 (or 50/50 = 1.00) fall between 79.75 and 98.83. This is close to what we would expect if the data were normally distributed.

The IQR = 4.84 and s = 3.18344. The ratio of the IQR and s is $\dfrac{IQR}{s} = \dfrac{4.84}{3.18344} = 1.52$. This is close to 1.3 that we would expect if the data were normally distributed.

Thus, there is evidence that the data are normally distributed.

b. If the data are normally distributed, the points will form a straight line when plotted using a normal probability plot. From the normal probability plot, the data points are close to a straight line. There is evidence that the data are normally distributed.

4.110 Using MINITAB, a histogram of the data with a normal curve drawn on the graph is:



From the graph, the data appear to be close to mound-shaped, so the data may be approximately normal.

4.112 **Distance**: To determine if the distribution of distances is approximately normal, we will run through the tests. Using MINITAB, the stem-and-leaf display is:

Stem-and-Leaf Display: Distance

```
Stem-and-leaf of Distance  N  = 40
Leaf Unit = 1.0


 1     28  3
 4     28  689
 10    29  011144
(11)   29  55556778889
 19    30  0000001112234
 6     30  59
 4     31  01
 2     31  68
```

From the stem-and-leaf display, the data look to be mound-shaped. The data may be normal.

Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Distance

```
Variable   N    Mean  StDev  Minimum      Q1  Median      Q3  Maximum
Distance  40  298.95   7.53   283.20  294.60  299.05  302.00   318.90
```

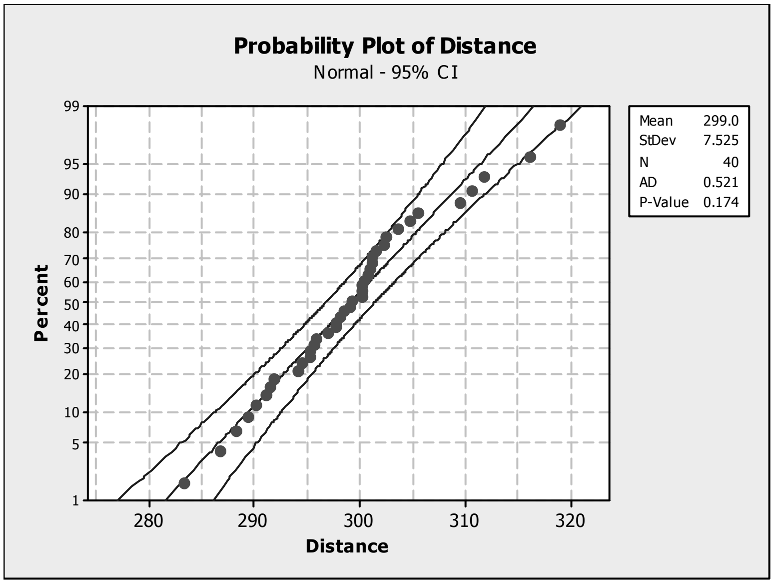
The interval $\bar{x} \pm s \Rightarrow 298.95 \pm 7.53 \Rightarrow (291.42,\ \ 306.48)$ contains 28 of the 40 observations. The proportion is 28 / 40 = .70. This is very close to the .68 from the Empirical Rule.

The interval $\bar{x} \pm 2s \Rightarrow 298.95 \pm 2(7.53) \Rightarrow 298.95 \pm 15.06 \Rightarrow (283.89,\ \ 314.01)$ contains 37 of the 40 observations. The proportion is 37 / 40 = .925. This is somewhat smaller than the .95 from the Empirical Rule.

The interval $\bar{x} \pm 3s \Rightarrow 298.95 \pm 3(7.53) \Rightarrow 298.95 \pm 22.59 \Rightarrow (276.36,\ \ 321.54)$ contains 40 of the 40 observations. The proportion is 40 / 40 = 1.00. This is very close to the .997 from the Empirical Rule. Thus, it appears that the data may be normal.

The lower quartile is Q_L = 294.60 and the upper quartile is Q_U = 302. The interquartile range is IQR = $Q_U - Q_L$ = 302 – 294.60 = 7.4. From the printout, s = 7.53. IQR / s = 7.4 / 7.53 = .983. This is somewhat less than the 1.3 that we would expect if the data were normal. Thus, there is evidence that the data may not be normal.

Using MINITAB, the normal probability plot is:



The data are very close to a straight line. Thus, it appears that the data may be normal.

From 3 of the 4 indicators, it appears that the distances come from an approximate normal distribution.

Accuracy: To determine if the distribution of accuracies is approximately normal, we will run through the tests. Using MINITAB, the stem-and-leaf display is:

Stem-and-Leaf Display: Accuracy

```
Stem-and-leaf of Accuracy  N  = 40
Leaf Unit = 1.0


 1   4  5
 1   4
 1   4
 2   5  0
 2   5
 3   5  4
 7   5  6777
 11  5  8999
 20  6  000001111
 20  6  2223333333
 10  6  4
 9   6  6667
 5   6  899
 2   7  0
 1   7  3
```

From the stem-and-leaf display, the data look to be skewed to the left. The data may not be normal.

Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Accuracy

```
Variable   N    Mean  StDev  Minimum      Q1  Median      Q3  Maximum
Accuracy  40  61.970  5.226   45.400  59.400  61.950  64.075   73.000
```

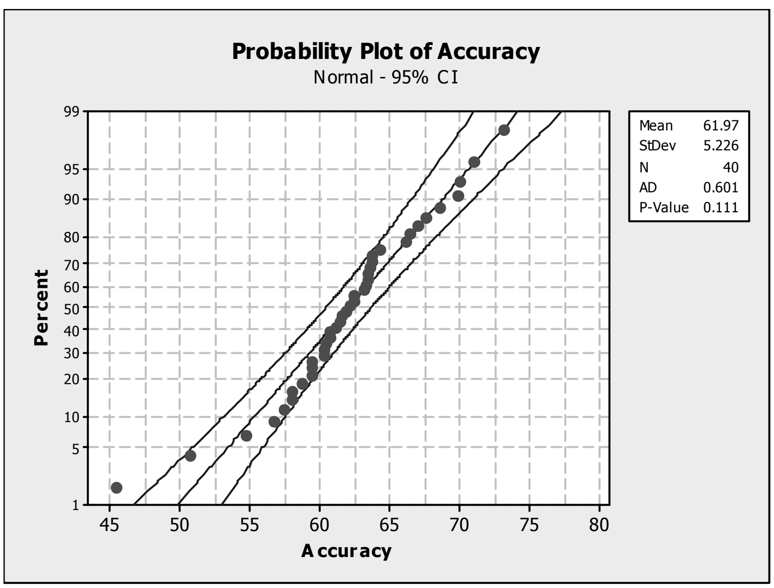
The interval $\bar{x} \pm s \Rightarrow 61.970 \pm 5.226 \Rightarrow (56.744, \;\; 67.196)$ contains 30 of the 40 observations. The proportion is 30 / 40 = .75. This is somewhat larger than the .68 from the Empirical Rule.

The interval $\bar{x} \pm 2s \Rightarrow 61.970 \pm 2(5.226) \Rightarrow 61.970 \pm 10.452 \Rightarrow (51.518, \;\; 72.422)$ contains 37 of the 40 observations. The proportion is 37 / 40 = .925. This is somewhat smaller than the .95 from the Empirical Rule.

The interval $\bar{x} \pm 3s \Rightarrow 61.970 \pm 3(5.226) \Rightarrow 61.970 \pm 15.678 \Rightarrow (46.292, \;\; 77.648)$ contains 39 of the 40 observations. The proportion is 39 / 40 = .975. This is somewhat smaller than the .997 from the Empirical Rule. Thus, it appears that the data may not be normal.

The lower quartile is Q_L = 59.400 and the upper quartile is Q_U = 64.075. The interquartile range is IQR = $Q_U - Q_L$ = 64.075 – 59.400 = 4.675. From the printout, s = 5.226. IQR / s = 4.675 / 5.226 = .895. This is less than the 1.3 that we would expect if the data were normal. Thus, there is evidence that the data may not be normal.

Using MINITAB, the normal probability plot is:



The data are not real close to a straight line. Thus, it appears that the data may not be normal.

From the 4 indicators, it appears that the accuracy values do not come from an approximate normal distribution.

Index: To determine if the distribution of driving performance index scores is approximately normal, we will run through the tests. Using MINITAB, the stem-and-leaf display is:

Stem-and-Leaf Display: ZSUM

```
Stem-and-leaf of ZSUM  N  = 40
Leaf Unit = 0.10

  1   1  1
  8   1  2233333
  18  1  4444445555
 (4)  1  7777
  18  1  8899
  14  2  0
  13  2  22222
  8   2  5
  7   2  77
  5   2  8
  4   3  1
  3   3  2
  2   3  45
```

From the stem-and-leaf display, the data look to be skewed to the right. The data may not be normal.

Using MINITAB, the descriptive statistics are:

Descriptive Statistics: ZSUM

```
Variable   N   Mean  StDev  Minimum     Q1  Median     Q3  Maximum
ZSUM      40  1.927  0.660    1.170  1.400   1.755  2.218    3.580
```

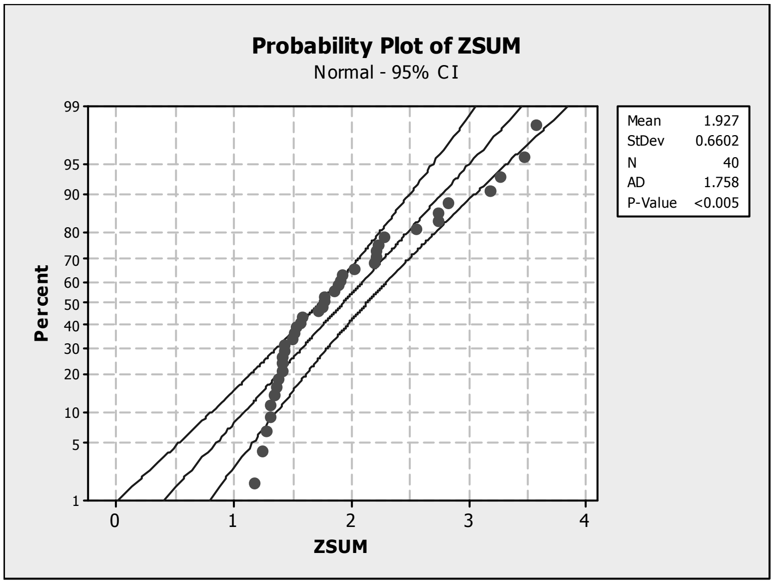
The interval $\bar{x} \pm s \Rightarrow 1.927 \pm .660 \Rightarrow (1.267,\ \ 2.587)$ contains 30 of the 40 observations. The proportion is 30 / 40 = .75. This is somewhat larger than the .68 from the Empirical Rule.

The interval $\bar{x} \pm 2s \Rightarrow 1.927 \pm 2(.660) \Rightarrow 1.927 \pm 1.320 \Rightarrow (.607,\ \ 3.247)$ contains 37 of the 40 observations. The proportion is 37 / 40 = .925. This is somewhat smaller than the .95 from the Empirical Rule.

The interval $\bar{x} \pm 3s \Rightarrow 1.927 \pm 3(.660) \Rightarrow 1.927 \pm 1.98 \Rightarrow (-.053,\ \ 3.907)$ contains 40 of the 40 observations. The proportion is 40 / 40 = 1.000. This is slightly larger than the .997 from the Empirical Rule. Thus, it appears that the data may not be normal.

The lower quartile is $Q_L = 1.4$ and the upper quartile is $Q_U = 2.218$. The interquartile range is IQR = $Q_U - Q_L = 2.218 - 1.4 = .818$. From the printout, $s = .66$. IQR / s = .818 / .66 = 1.24. This is fairly close to the 1.3 that we would expect if the data were normal. Thus, there is evidence that the data may be normal.

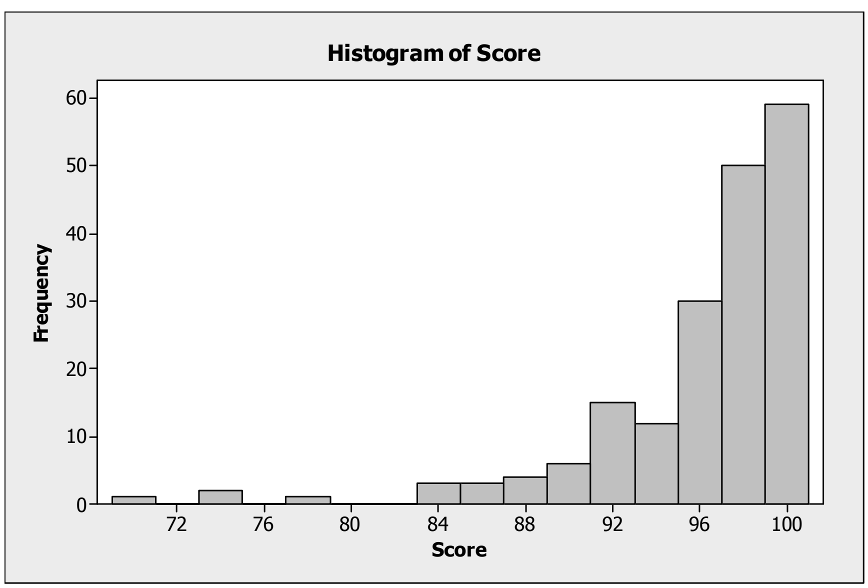
Using MINITAB, the normal probability plot is:



The data are not real close to a straight line. Thus, it appears that the data may not be normal.

From 3 of the 4 indicators, it appears that the driving performance index scores do not come from an approximate normal distribution.

4.114 To determine if the distribution of sanitation scores is approximately normal, we will run through the tests. Using MINITAB, the histogram of the data is:



From the histogram, the data look to be skewed to the left. The data may not be normal.

Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Score

```
Variable    N    Mean  StDev  Minimum      Q1  Median      Q3  Maximum
Score     186  95.699  4.963   69.000  94.000  97.000  99.000  100.000
```

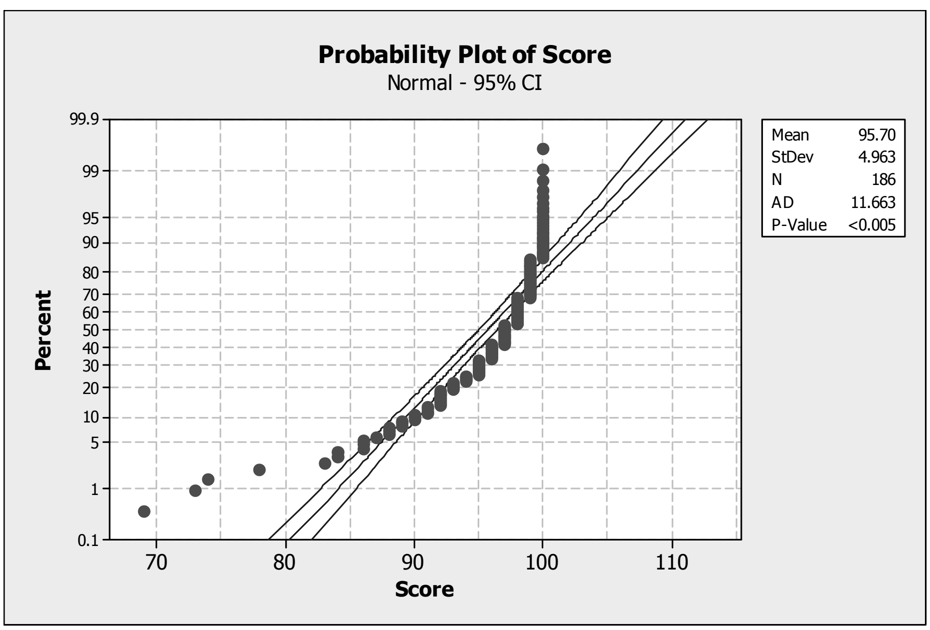
The interval $\bar{x} \pm s \Rightarrow 95.70 \pm 4.96 \Rightarrow (90.74,\ \ 100.66)$ contains 166 of the 186 observations. The proportion is 166 / 186 = .892. This is much larger than the .68 from the Empirical Rule.

The interval $\bar{x} \pm 2s \Rightarrow 95.70 \pm 2(4.96) \Rightarrow 95.70 \pm 9.92 \Rightarrow (85.78,\ \ 105.62)$ contains 179 of the 186 observations. The proportion is 179 / 186 = .962. This is somewhat larger than the .95 from the Empirical Rule.

The interval $\bar{x} \pm 3s \Rightarrow 95.70 \pm 3(4.96) \Rightarrow 95.70 \pm 14.88 \Rightarrow (80.82,\ \ 110.58)$ contains 182 of the 186 observations. The proportion is 182 / 186 = .978. This is somewhat smaller than the .997 from the Empirical Rule. Thus, it appears that the data may not be normal.

The lower quartile is $Q_L = 94$ and the upper quartile is $Q_U = 99$. The interquartile range is IQR = $Q_U - Q_L = 99 - 94 = 5$. From the printout, s = 4.963. IQR / s = 5 / 4.963 = 1.007. This is not particularly close to the 1.3 that we would expect if the data were normal. Thus, there is evidence that the data may not be normal.

Using MINITAB, the normal probability plot is:



The data are not real close to a straight line. Thus, it appears that the data may not be normal.

From the 4 indicators, it appears that the sanitation scores do not come from an approximate normal distribution.

4.116 a. From the normal probability plot, it is very unlikely that the data are normally distributed. If the data are normal, the normal probability plot will form a straight line. For this normal probability plot, the data do not form a straight line. Thus, the data are not normal.

b. Since the data points corresponding to the largest z-scores are spread out further than the points corresponding to the smallest z-scores, the data are skewed to the right.

4.118 The binomial probability distribution is a discrete distribution. The random variable can take on only a limited number of values. The normal distribution is a continuous distribution. The random variable can take on an infinite number of values. To get a better estimate of probabilities for the binomial probability distribution using the normal distribution, we use the continuity correction factor.

4.120 a. $\mu = np = 100(.01) = 1.0$, $\sigma = \sqrt{npq} = \sqrt{100(.01)(.99)} = .995$

$\mu \pm 3\sigma \Rightarrow 1 \pm 3(.995) \Rightarrow 1 \pm 2.985 \Rightarrow (-1.985,\ 3.985)$

Since this interval does not fall in the interval (0, n = 100), the normal approximation is not appropriate.

b. $\mu = np = 20(.6) = 12$, $\sigma = \sqrt{npq} = \sqrt{20(.6)(.4)} = 2.191$

$\mu \pm 3\sigma \Rightarrow 12 \pm 3(2.191) \Rightarrow 12 \pm 6.573 \Rightarrow (5.427,\ 18.573)$

Since this interval falls in the interval (0, n = 20), the normal approximation is appropriate.

c. $\mu = np = 10(.4) = 4$, $\sigma = \sqrt{npq} = \sqrt{10(.4)(.6)} = 1.549$

$\mu \pm 3\sigma \Rightarrow 4 \pm 3(1.549) \Rightarrow 4 \pm 4.647 \Rightarrow (-.647,\ 8.647)$

Since this interval does not fall within the interval (0, n = 10), the normal approximation is not appropriate.

d. $\mu = np = 1000(.05) = 50$, $\sigma = \sqrt{npq} = \sqrt{1000(.05)(.95)} = 6.892$

$\mu \pm 3\sigma \Rightarrow 50 \pm 3(6.892) \Rightarrow 50 \pm 20.676 \Rightarrow (29.324,\ 70.676)$

Since this interval falls within the interval (0, n = 1000), the normal approximation is appropriate.

e. $\mu = np = 100(.8) = 80$, $\sigma = \sqrt{npq} = \sqrt{100(.8)(.2)} = 4$

$\mu \pm 3\sigma \Rightarrow 80 \pm 3(4) \Rightarrow 80 \pm 12 \Rightarrow (68, 92)$

Since this interval falls within the interval (0, n = 100), the normal approximation is appropriate.

f. $\mu = np = 35(.7) = 24.5$, $\sigma = \sqrt{npq} = \sqrt{35(.7)(.3)} = 2.711$

$\mu \pm 3\sigma \Rightarrow 24.5 \pm 3(2.711) \Rightarrow 24.5 \pm 8.133 \Rightarrow (16.367, 32.633)$

Since this interval falls within the interval (0, n = 35), the normal approximation is appropriate.

4.122 $\mu = np = 1000(.5) = 500$, $\sigma = \sqrt{npq} = \sqrt{1000(.5)(.5)} = 15.811$

a. Using the normal approximation,

$$P(x > 500) \approx P\left(z > \frac{(500+.5)-500}{15.811}\right) = P(z > .03) = .5 - .0120 = .4880$$

(from Table III, Appendix A)

b. $$P(490 \le x < 500) \approx P\left(\frac{(490-.5)-500}{15.811} \le z < \frac{(500-.5)-500}{15.811}\right)$$

$$= P(-.66 \le z < -.03) = .2454 - .0120 = .2334$$

(from Table III, Appendix A)

c. $$P(x > 550) \approx P\left(z > \frac{(500+.5)-500}{15.811}\right) = P(z > 3.19) \approx .5 - .5 = 0$$

(from Table III, Appendix A)

4.124 a. For this exercise n = 500 and p = .5.

$\mu = np = 500(.5) = 250$ and $\sigma = \sqrt{npq} = \sqrt{500(.5)(.5)} = \sqrt{125} = 11.1803$

b. $z = \dfrac{x-\mu}{\sigma} = \dfrac{240-250}{11.1803} = -.89$

c. $z = \dfrac{x-\mu}{\sigma} = \dfrac{270-250}{11.1803} = 1.79$

d. $\mu \pm 3\sigma \Rightarrow 250 \pm 3(11.1803) \Rightarrow 250 \pm 33.5409 \Rightarrow (216.4591, 283.5409)$

Since the above is completely contained in the interval 0 to 500, the normal approximation is valid.

$$P(240 < x < 270) = P\left(\frac{(240+.5)-250}{11.1803} < z < \frac{(270-.5)-250}{11.1803}\right)$$
$$= P(-.85 < z < 1.74) = .3023 + .4591 = .7614$$

(Using Table III, Appendix A)

4.126 a. Let x = number of patients who experience serious post-laser vision problems in 100,000 trials. Then x is a binomial random variable with n = 100,000 and p =.01.

$$E(x) = \mu = np = 100,000(.01) = 1000.$$

b. $V(x) = \sigma^2 = npq = 100,000(.01)(.99) = 990$

c. $z = \dfrac{x-\mu}{\sigma} = \dfrac{950-1000}{\sqrt{990}} = \dfrac{-50}{31.4643} = -1.59$

d. $\mu \pm 3\sigma \Rightarrow 1000 \pm 3(31.4643) \Rightarrow 1000 \pm 94.3929 \Rightarrow (905.6071,\ \ 1,094.3929)$ Since the interval lies in the range 0 to 100,000, we can use the normal approximation to approximate the binomial probability.

$$P(x < 950) = P\left(z < \frac{(950-.5)-1000}{31.4643}\right) = P(z < -1.60) = .5 - .4452 = .0548$$

(Using Table III, Appendix A)

4.128 a. $\mu = E(x) = np = 1000(.32) = 320$. This is the same value that was found in Exercise 4.56 a.

b. $\sigma = \sqrt{npq} = \sqrt{1000(.32)(1-.32)} = \sqrt{217.6} = 14.751$ This is the same value that was found in Exercise 4.56 b.

c. $z = \dfrac{x-\mu}{\sigma} = \dfrac{200.5-320}{14.751} = -8.10$

d. $\mu \pm 3\sigma \Rightarrow 320 \pm 3(14.751) \Rightarrow 320 \pm 44.253 \Rightarrow (275.474,\ 364.253)$

Since the above is completely contained in the interval 0 to 1000, the normal approximation is valid.

$$P(x \le 200) \approx P\left(z \le \frac{200.5-320}{14.751}\right) = P(z \le -8.10) \approx .5 - .5 = 0$$

(Using Table III, Appendix A)

4.130 For this exercise, n = 100 and p = .4.

$$\mu = np = 100(.4) = 40 \quad \text{and} \quad \sigma = \sqrt{npq} = \sqrt{100(.4)(.6)} = \sqrt{24} = 4.899$$

$\mu \pm 3\sigma \Rightarrow 40 \pm 3(4.899) \Rightarrow 40 \pm 14.697 \Rightarrow (25.303, 54.697)$

Since the above is completely contained in the interval 0 to 100, the normal approximation is valid.

$$P(x < 50) = P\left(z < \frac{(50-.5)-40}{4.899}\right) = P(z < 1.94) = .5 + .4738 = .9738$$

(Using Table III, Appendix A)

4.132 a. Let x = number of abused women in a sample of 150. The random variable x is a binomial random variable with n = 150 and p = 1/3. Thus, for the normal approximation,

$$\mu = np = 150(1/3) = 50 \text{ and } \sigma = \sqrt{npq} = \sqrt{150(1/3)(2/3)} = 5.7735$$
$$\mu \pm 3\sigma \Rightarrow 50 \pm 3(5.7735) \Rightarrow 50 \pm 17.3205 \Rightarrow (32.6795, 67.3205)$$

Since this interval lies in the range from 0 to n = 150, the normal approximation is appropriate.

$$P(x > 75) \approx P\left(z > \frac{(75+.5)-50}{5.7735}\right) = P(z > 4.42) \approx .5 - .5 = 0$$

(Using Table III, Appendix A.)

b. $$P(x < 50) \approx P\left(z < \frac{(50-.5)-50}{5.7735}\right) = P(z < -.09) \approx .5 - .0359 = .4641$$

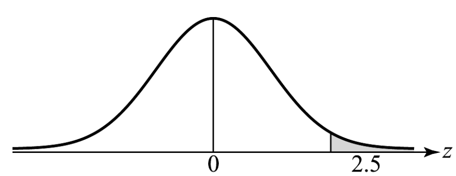
c. $$P(x < 30) \approx P\left(z < \frac{(30-.5)-50}{5.7735}\right) = P(z < -3.55) \approx .5 - .5 = 0$$

Since the probability of seeing fewer than 30 abused women in a sample of 150 is so small ($p \approx 0$), it would be very unlikely to see this event.

4.134 a. Let x equal the percentage of body fat in American men. The random variable x is a normal random variable with μ = 15 and σ = 2.

$$\begin{aligned} P(\text{Man is obese}) &= P(x \ge 20) \\ &\approx P\left(z \ge \frac{20-15}{2}\right) \\ &= P(z \ge 2.5) \\ &= .5000 - .4938 = .0062 \end{aligned}$$

(Using Table III in Appendix A.)

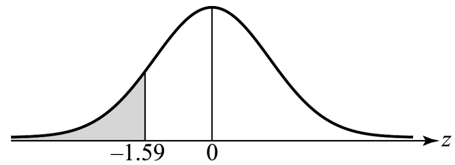


Let y equal the number of men in the U.S. Army who are obese in a sample of 10,000. The random variable y is a binomial random variable with n = 10,000 and p = .0062.

$$\mu \pm 3\sigma \Rightarrow np \pm 3\sqrt{npq} \Rightarrow 10{,}000(.0062) \pm 3\sqrt{10{,}000(.0062)(1-.0062)}$$
$$\Rightarrow 62 \pm 3(7.85) \Rightarrow (38.45, 85.55)$$

Since the interval does lie in the range 0 to 10,000, we can use the normal approximation to approximate the probability.

$$P(x<50) \approx P\left(z<\frac{(50-.5)-62}{7.85}\right)$$
$$\approx P(z<-1.59)$$
$$=.5000-.4441=.0559$$
(Using Table III in Appendix A.)



b. The probability of finding less than 50 obese Army men in a sample of 10,000 is .0559. Therefore, the probability of finding only 30 would even be smaller. Thus, it looks like the Army has successfully reduced the percentage of obese men since this did occur.

4.136 Let x = number of patients who wait more than 30 minutes. Then x is a binomial random variable with n = 150 and p = .5.

a. $\mu = np = 150(.5) = 75,\ \sigma = \sqrt{npq} = \sqrt{150(.5)(.5)} = 6.124$

$$P(x>75) \approx P\left(z>\frac{(75+.5)-75}{6.124}\right) = P(z>.08) = .5-.0319 = .4681$$

(from Table III, Appendix A)

b. $$P(x>85) \approx P\left(z>\frac{(85+.5)-75}{6.124}\right) = P(z>1.71) = .5-.4564 = .0436$$

(from Table III, Appendix A)

c. $$P(60<x<90) \approx P\left(\frac{(60+.5)-75}{6.124}<z<\frac{(90-.5)-75}{6.124}\right)$$
$$= P(-2.37<z<2.37) = .4911+.4911 = .9822$$

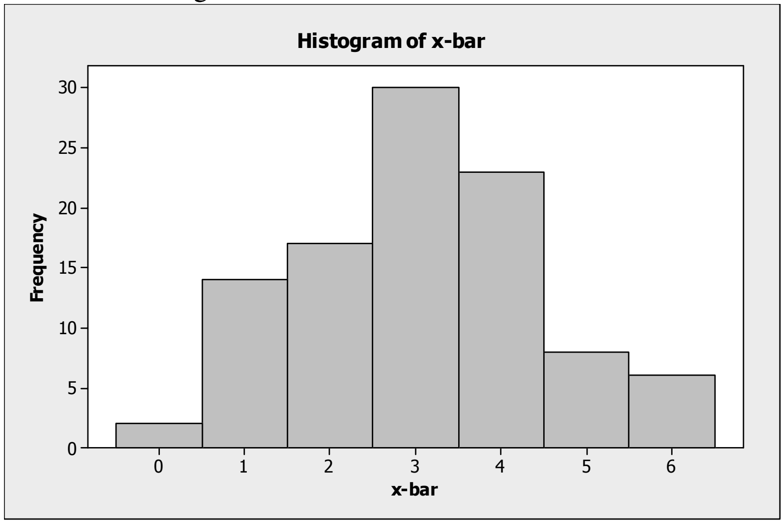
(from Table III, Appendix A)

4.138 The sampling distribution of a sample statistic calculated from a sample of n measurements is the probability distribution of the statistic.

4.140 Answers will vary. One hundred samples of size 2 were generated and the value of $\bar{x}$ computed for each. The first 10 samples along with the values of $\bar{x}$ are shown in the table:

Sample	Values		x-bar		Sample	Values		x-bar
1	2	4	3		6	2	0	1
2	4	4	4		7	2	0	1
3	2	4	3		8	0	4	2
4	0	2	1		9	4	6	5
5	4	6	5		10	0	4	2

Using MINITAB, the histogram of the 100 values of $\bar{x}$ is:



The shape of this histogram is very similar to that of the exact distribution in Exercise 4.139e. This histogram is not exactly the same because it is based on a sample size of only 100.

4.142 $E(x) = \mu = \sum xp(x) = 1(.2) + 2(.3) + 3(.2) + 4(.2) + 5(.1) = .2 + .6 + .8 + .5 = 2.7$

From Exercise 4.141, the sampling distribution of $\bar{x}$ is

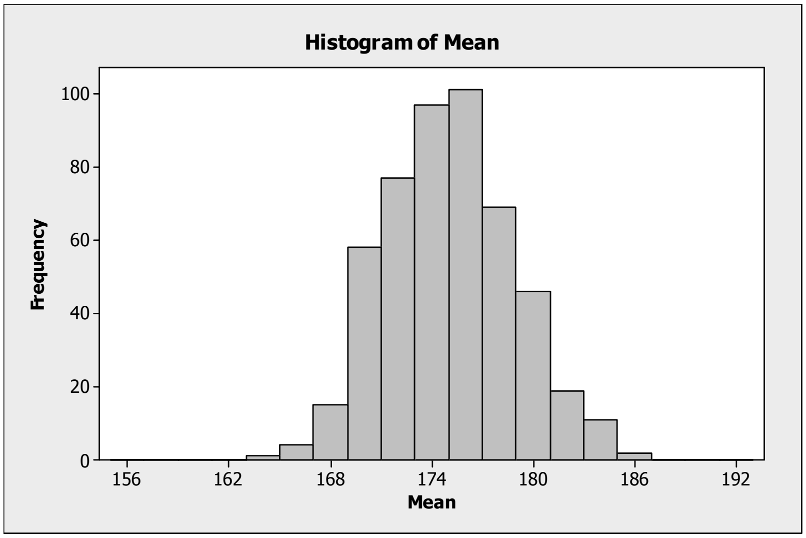
$\bar{x}$	1	1.5	2	2.5	3	3.5	4	4.5	5
$p(\bar{x})$	.04	.12	.17	.20	.20	.14	.08	.04	.01

$$E(\bar{x}) = \sum \bar{x}p(\bar{x}) = 1.0(.04) + 1.5(.12) + 2.0(.17) + 2.5(.20) + 3.0(.20) + 3.5(.14) + 4.0(.08) + 4.5(.04) + 5.0(.01)$$
$$= .04 + .18 + .34 + .50 + .60 + .49 + .32 + .18 + .05 = 2.7$$

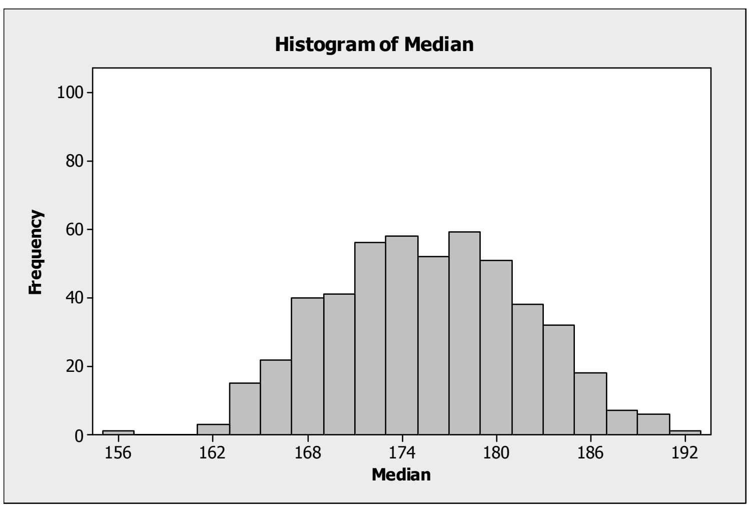
4.144 a. Answers will vary. MINITAB was used to generate 500 samples of size n = 15 observations from a uniform distribution over the interval from 150 to 200. The first 10 samples along with the sample means are shown in the table below:

Sample	Observations															Mean	Median
1	159	200	177	158	195	165	196	180	174	181	180	154	160	192	153	174.93	177
2	180	166	157	195	173	168	190	168	170	199	198	165	180	166	175	176.67	173
3	173	179	159	170	162	194	165	167	168	160	164	153	154	154	165	165.80	165
4	200	155	166	152	164	165	190	176	165	197	164	173	187	152	164	171.33	165
5	190	196	154	183	170	172	200	158	150	187	184	191	182	180	188	179.00	183
6	194	185	186	190	180	178	183	196	193	170	178	197	173	196	196	186.33	186
7	166	196	156	151	151	168	158	185	160	199	166	185	159	161	184	169.67	166
8	154	164	188	158	167	153	174	188	185	153	161	188	198	173	192	173.07	173
9	177	152	161	156	177	198	185	161	167	156	157	189	192	168	175	171.40	168
10	192	187	176	161	200	184	154	151	185	163	176	155	155	191	171	173.40	176

Using MINITAB, the histogram of the 500 values of $\bar{x}$ is:



b. The sample medians were computed for each of the samples. The medians of the first 10 samples are shown in the table in part a. Using MINITAB, the histogram of the 500 values of the median is:



The graph of the sample medians is flatter and more spread out than the graph of the sample means.

4.146 The symbol $\mu_{\bar{x}}$ represents the mean of the sampling distribution of $\bar{x}$. The symbol $\sigma_{\bar{x}}$ represents the standard deviation of the sampling distribution of $\bar{x}$.

4.148 The standard deviation of the sampling distribution of $\bar{x}$, $\sigma_{\bar{x}}$, is the standard deviation of the population from which the sample is selected divided by $\sqrt{n}$.

4.150 The sampling distribution is approximately normal only if the sample size is sufficiently large.

4.152 a. $\mu_{\bar{x}} = \mu = 10$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3}{\sqrt{25}} = 0.6$

b. $\mu_{\bar{x}} = \mu = 100$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{25}{\sqrt{25}} = 5$

c. $\mu_{\bar{x}} = \mu = 20$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{40}{\sqrt{25}} = 8$

d. $\mu_{\bar{x}} = \mu = 10$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{100}{\sqrt{25}} = 20$

4.154 a. $\mu_{\bar{x}} = \mu = 20$, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{16}{\sqrt{64}} = 2$

b. By the Central Limit Theorem, the distribution of $\bar{x}$ is approximately normal. In order for the Central Limit Theorem to apply, n must be sufficiently large. For this problem, $n = 64$ is sufficiently large.

c. $z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \dfrac{16-20}{2} = -2.00$

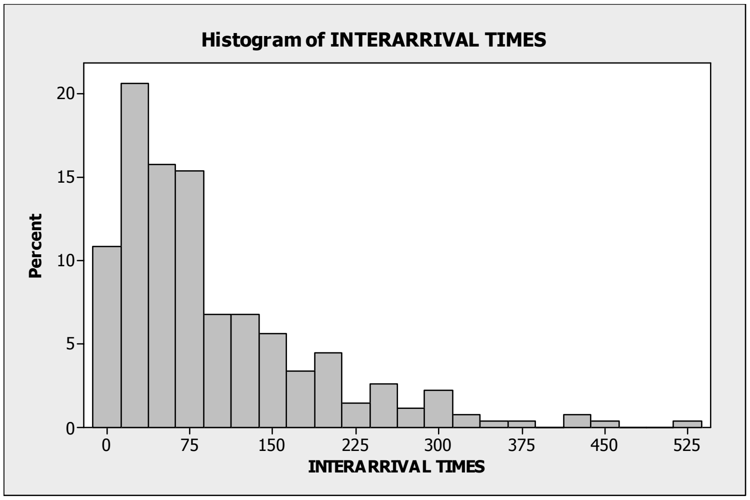
d. $z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \dfrac{23-20}{2} = 1.50$

e. $P(\bar{x} < 16) = P\left(z < \dfrac{16-20}{2}\right) = P(z < -2) = .5 - .4772 = .0228$

f. $P(\bar{x} > 23) = P\left(z > \dfrac{23-20}{2}\right) = P(z > 1.50) = .5 - .4332 = .0668$

g. $P(16 < \bar{x} < 23) = P\left(\dfrac{16-20}{2} < z < \dfrac{23-20}{2}\right) = P(-2 < z < 1.5) = .4772 + .4332 = .9104$

4.156 a. From Exercise 2.43, the histogram of the data is:



The distribution of the interarrival times is skewed to the right.

b. Using MINITAB, the descriptive statistics are:

Descriptive Statistics: INTTIME

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
INTTIME	267	95.52	91.54	1.86	30.59	70.88	133.34	513.52

The mean is 95.52 and the standard deviation is 91.54.

c. The sampling distribution of $\bar{x}$ should be approximately normal with a mean of $\mu_{\bar{x}} = \mu = 95.52$ and a standard deviation of $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{91.54}{\sqrt{40}} = 14.47$.

d. $P(\bar{x} < 90) = P\left(z < \dfrac{90 - 95.52}{91.54/\sqrt{40}} \right) = P(z < -.38) = .5 - .1480 = .3520$

e. Answers can vary. Using SAS, 40 randomly selected interarrival times were:

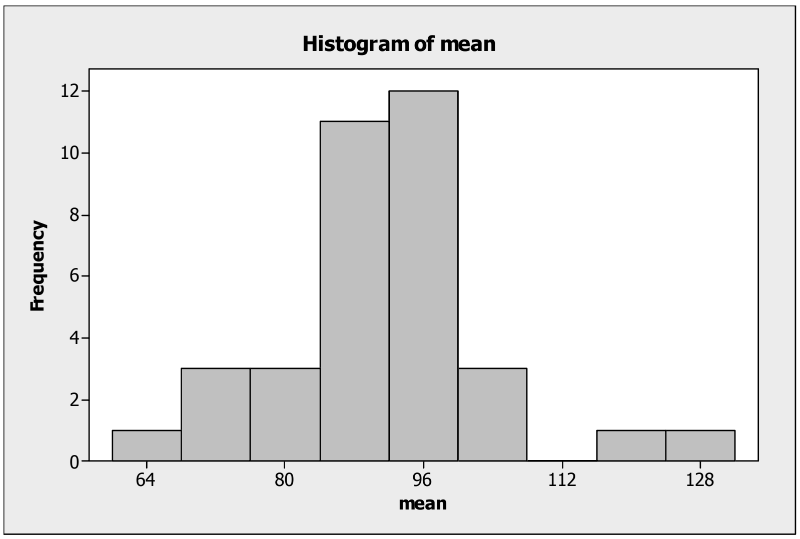
17.698	153.772	20.229	20.740	49.822
38.303	7.995	148.289	267.534	13.411
63.037	4.626	105.527	163.349	5.016
47.203	38.791	36.421	120.624	41.905
88.855	2.954	17.886	428.507	82.760
84.203	54.291	289.669	11.062	303.903
259.861	126.448	53.104	32.193	37.720
256.385	185.598	30.720	40.760	72.648

For this sample, $\bar{x} = 95.595$.

f. Answers can vary. Thirty-five samples of size 40 were selected and the value of $\bar{x}$ was computed for each. The thirty-five values of $\bar{x}$ were:

95.595	97.476	99.125	70.176	83.850
92.095	125.785	88.988	85.435	93.682
87.741	103.584	96.868	92.104	93.697
88.209	87.617	117.301	84.088	69.962
105.779	96.428	94.949	102.164	87.467
78.566	86.630	92.719	84.370	73.767
95.066	67.268	86.906	84.404	77.111

Using MINITAB, the histogram of these means is:



This histogram is somewhat mound-shaped.

g. Answers will vary. Using MINITAB, the descriptive statistics for these $\bar{x}$-values are:

Descriptive Statistics: mean

```
Variable   N   Mean  StDev  Minimum     Q1  Median     Q3  Maximum
mean      35  90.48  12.15    67.27  84.37   88.99  96.43   125.79
```

The mean of the $\bar{x}$-values is 90.48. This is somewhat close to the value of $\mu_{\bar{x}} = 95.52$.

The standard deviation of the $\bar{x}$-values is 12.15. Again, this is somewhat close to the value of $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{91.54}{\sqrt{40}} = 14.47$. If more than 35 samples of size 40 were selected, the mean of the $\bar{x}$-values will get closer to 95.52 and the standard deviation will get closer to 14.47.

4.158 a. Let $\bar{x}$ = sample mean FNE score. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with

$$\mu_{\bar{x}} = \mu = 18 \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{45}} = .7453.$$

$$P(\bar{x} > 17.5) = P\left(z > \frac{17.5-18}{.7453}\right) = P(z > -.67) = .5 + .2486 = .7486$$

(Using Table III, Appendix A)

b. $$P(18 < \bar{x} < 18.5) = P\left(\frac{18-18}{.7453} < z < \frac{18.5-18}{.7453}\right) = P(0 < z < .67) = .2486$$

(Using Table III, Appendix A)

c. $$P(\bar{x} < 18.5) = P\left(z < \frac{18.5-18}{.7453}\right) = P(z < .67) = .5 + .2486 = .7486$$

(Using Table III, Appendix A)

4.160 Let $\bar{x}$ = sample mean shell length. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with $\mu_{\bar{x}} = \mu = 50$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{76}} = 1.147$.

$$P(\bar{x} > 55.5) = P\left(z > \frac{55.5-50}{1.147}\right) = P(z > 4.79) \approx .5 - .5 = 0 \quad \text{(using Table III, Appendix A)}$$

4.162 a. Let $\bar{x}$ = sample mean amount of uranium. $E(x) = \mu = \frac{c+d}{2} = \frac{1+3}{2} = 2$. Thus, $E(\bar{x}) = \mu_{\bar{x}} = \mu = 2$.

b. $V(x) = \sigma^2 = \frac{1}{3}$. Thus, $V(\bar{x}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{1/3}{60} = \frac{1}{180}$ and $\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}^2} = \sqrt{\frac{1}{180}} = .0745$.

c. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with $\mu_{\bar{x}} = \mu = 2$ and $\sigma_{\bar{x}} = .0745$.

d. $$P(1.5 < \bar{x} < 2.5) = P\left(\frac{1.5-2}{.0745} < z < \frac{2.5-2}{.0745}\right) = P(-6.71 < z < 6.71) \approx .5 + .5 = 1$$

(using Table III, Appendix A)

e. $P(\bar{x} > 2.2) = P\left(z > \frac{2.2-2}{.0745}\right) = P(z > 2.68) = .5 - .4963 = .0037$

(using Table III, Appendix A)

4.164 a. $\mu_{\bar{x}} = \mu = 2.78$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.15}{\sqrt{100}} = .015$

b. $P(2.78 < \bar{x} < 2.80) = P\left(\frac{2.78-2.78}{.015} < z < \frac{2.80-2.78}{.015}\right) = P(0 < z < 1.33) = .4082$

(Using Table III, Appendix A)

c. $P(\bar{x} > 2.80) = P\left(z > \frac{2.80-2.78}{.015}\right) = P(z > 1.33) = .5 - .4082 = .0918$

(Using Table III, Appendix A)

d. $\mu_{\bar{x}} = \mu = 2.78$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.15}{\sqrt{200}} = .0106$

The mean of the sampling distribution of $\bar{x}$ remains the same, but the standard deviation would decrease.

$$P(2.78 < \bar{x} < 2.80) = P\left(\frac{2.78-2.78}{.0106} < z < \frac{2.80-2.78}{.0106}\right) = P(0 < z < 1.89) = .4706$$

(Using Table III, Appendix A)

The probability is larger than when the sample size is 100.

$$P(\bar{x} > 2.80) = P\left(z > \frac{2.80-2.78}{.0106}\right) = P(z > 1.89) = .5 - .4706 = .0294$$

(Using Table III, Appendix A)

The probability is smaller than when the sample size is 100.

4.166 Let $\bar{x}$ = sample mean WR score. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with $\mu_{\bar{x}} = \mu = 40$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{100}} = .5$.

$P(\bar{x} \geq 42) = P\left(z \geq \frac{42-40}{.5}\right) = P(z \geq 4) \approx .5 - .5 = 0$ (using Table III, Appendix A)

Since the probability of seeing a mean WR score of 42 or higher is so small if the sample had been selected from the population of convicted drug dealers, we would conclude that the sample was *not* selected from the population of convicted drug dealers.

4.168 Let $\bar{x}$ = sample mean attitude score (KAE-A). By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with

$$\mu_{\bar{x}} = \mu = 11.92 \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.95}{\sqrt{100}} = .295.$$

If $\bar{x}$ is from KAE-A, then $z = \dfrac{6.5 - 11.92}{.295} = -18.37$

If $\bar{x}$ = sample mean knowledge score (KAE-GK). By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with

$$\mu_{\bar{x}} = \mu = 6.35 \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.12}{\sqrt{100}} = .212.$$

If $\bar{x}$ is from KAE-GK, then $z = \dfrac{6.5 - 6.35}{.212} = .71$

The score of $\bar{x}$ = 6.5 is much more likely to have come from the KAE-GK distribution because the z-score associated with 6.5 is much closer to 0. A z-score of –18.37 indicates that it would be almost impossible that the value came from the KAE-A distribution.

4.170 a. The length of time that an exercise physiologist's program takes to elevate her client's heart rate to 140 beats per minute is measured on an interval and thus, is continuous.

b. The number of crimes committed on a college campus per year is a whole number such as 0, 1, 2, etc. This variable is discrete.

c. The number of square feet of vacant office space in a large city is a measurement of area and is measured on an interval. Thus, this variable is continuous.

d. The number of voters who favor a new tax proposal is a whole number such as 0, 1, 2, etc. This variable is discrete.

4.172 We could obtain a simulated sampling distribution of a sample statistic by taking random samples from a single population, computing $\bar{x}$ for each sample, and then finding a histogram of these $\bar{x}$'s.

4.174 The statement "The sampling distribution of $\bar{x}$ is normally distributed regardless of the size of the sample *n*" is false. If the original population being sampled from is normal, then the sampling distribution of $\bar{x}$ is normally distributed regardless of the size of the sample *n*. However, if the original population being sampled from is not normal, then the sampling distribution of $\bar{x}$ is normally distributed only if the size of the sample *n* is sufficiently large.

4.176 a. $\mu = \sum xp(x) = 10(.2) + 12(.3) + 18(.1) + 20(.4) = 15.4$

$\sigma^2 = \sum (x-\mu)^2 p(x)$

$= (10-15.4)^2(.2) + (12-15.4)^2(.3) + (18-15.4)^2(.1) + (20-15.4)^2(.4) = 18.44$

$\sigma = \sqrt{18.44} \approx 4.294$

b. $P(x < 15) = p(10) + p(12) = .2 + .3 = .5$

c. $\mu \pm 2\sigma \Rightarrow 15.4 \pm 2(4.294) \Rightarrow 15.4 \pm 8.588 \Rightarrow (6.812,\ 23.988)$

d. $P(6.812 < x < 23.988) = .2 + .3 + .1 + .4 = 1.0$

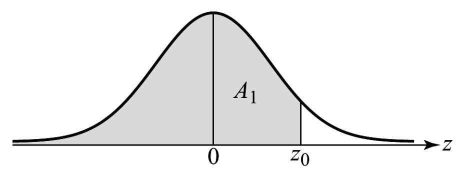
4.178 Using Table III, Appendix A:

a. $P(x \ge x_0) = .5$. Find x_0.

$$P(x \ge x_0) = P\left(z \ge \frac{x_0 - 40}{6}\right) = P(z \ge z_0) = .5 \Rightarrow z_0 = 0$$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow 0 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 40$$

b. $P(x \le x_0) = .9911$. Find x_0.



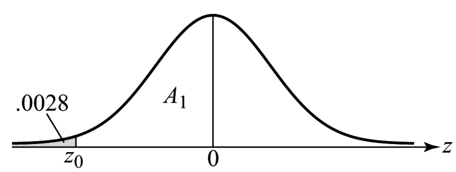
$$P(x \le x_0) = P\left(z \le \frac{x_0 - 40}{6}\right) = P(z \le z_0) = .9911$$

$A_1 = .9911 - .5 = .4911$

Looking up area .4911, $z_0 = 2.37$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow 2.37 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 54.22$$

c. $P(x \le x_0) = .0028$. Find x_0.



$$P(x \le x_0) = P\left(z \le \frac{x_0 - 40}{6}\right) = P(z \le z_0) = .0028$$

$A_1 = .5 - .0028 = .4972$

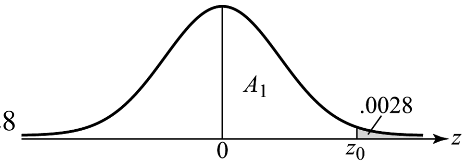
Looking up area .4972, $z_0 = 2.77$

Since z_0 is to the left of 0, $z_0 = -2.77$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow -2.77 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 23.38$$

d. $P(x \geq x_0) = .0228$. Find x_0.

$$P(x \geq x_0) = P\left(z \geq \frac{x_0 - 40}{6}\right) = P(z \geq z_0) = .0228$$

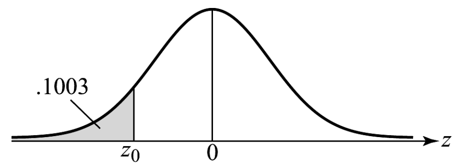


$A_1 = .5 - .0228 = .4772$
Looking up area .4772, $z_0 = 2.0$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow 2 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 52$$

e. $P(x \leq x_0) = .1003$. Find x_0.

$$P(x \leq x_0) = P\left(z \leq \frac{x_0 - 60}{8}\right) = P(z \leq z_0) = .1003$$

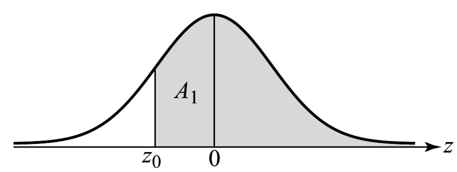


$A_1 = .5 - .1003 = .3997$
Looking up area .3997, $z = 1.28$.
Since z_0 is to the left of 0, $z_0 = -1.28$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow -1.28 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 32.32$$

f. $P(x \geq x_0) = .7995$. Find x_0.

$$P(x \geq x_0) = P\left(z \geq \frac{x_0 - 60}{8}\right) = P(z \geq z_0) = .7995$$



$A_1 = .7995 - .5 = .2995$
Looking up area .2995, $z = .84$.
Since z_0 is to the left of 0, $z_0 = -.84$

$$z_0 = \frac{x_0 - 40}{6} \Rightarrow -.84 = \frac{x_0 - 40}{6} \Rightarrow x_0 = 34.96$$

4.180 By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal.

$$\mu_{\bar{x}} = \mu = 19.6, \quad \sigma_{\bar{x}} = \frac{3.2}{\sqrt{68}} = .388$$

a. $P(\bar{x} \leq 19.6) = P\left(z \leq \dfrac{19.6 - 19.6}{.388}\right) = P(z \leq 0) = .5$ (using Table III, Appendix A)

b. $P(\bar{x} \leq 19) = P\left(z \leq \dfrac{19 - 19.6}{.388}\right) = P(z \leq -1.55) = .5 - .4394 = .0606$

(using Table III, Appendix A)

c. $P(\bar{x} \geq 20.1) = P\left(z \geq \dfrac{20.1-19.6}{.388}\right) = P(z \geq 1.29) = .5 - .4015 = .0985$

(using Table III, Appendix A)

d. $P(19.2 \leq \bar{x} \leq 20.6) = P\left(\dfrac{19.2-19.6}{.388} < z < \dfrac{20.6-19.6}{.388}\right)$

$= P(-1.03 \leq z \leq 2.58) = .3485 + .4951 = .8436$

(using Table III, Appendix A)

4.182 a. From the problem, x is a binomial random variable with $n = 3$ and $p = .6$.

$$P(x=0) = \binom{3}{0}(.6)^0(.4)^{3-1} = \frac{3!}{0!3!}(.6)^0(.4)^3 = .064$$

b. $P(x \geq 1) = 1 - P(x=0) = 1 - .064 = .936$

c. $\mu = E(x) = np = 3(.6) = 1.8$

$\sigma = \sqrt{npq} = \sqrt{3(.6)(.4)} = .8485$

In samples of 3 parents, on the average, 1.8 condone spanking.

4.184 The random variable x would have a binomial distribution:

1. n identical trials. A sample of married women was taken and would be considered small compared to the entire population of married women, so the trials would be very close to identical.

2. There are only two possible outcomes for each trial. In this experiment, there are only two possible outcomes: the woman would marry the same man (S) or she would not (F).

3. The probability of Success stays the same from trial to trial. Here, $P(S) = .50$ and $P(F) = .50$. In reality, these probabilities would not be exactly the same from trial to trial, but rounded off to 4 decimal places, they would be the same.

4. The trials are independent. In this experiment, the trials would not be exactly independent because we would be sampling without replacement from a finite population. However, if the sample size is fairly small compared to the population size, the trials will be essentially independent.

5. The binomial random variable x would be the number of successes in n trials. For this experiment, x = the number of women out of those sampled who would marry the same man again.

   Thus, x would possess (approximately) a binomial distribution.

4.186 a. $\mu = np = 200(.5) = 100$

b. $\sigma = \sqrt{npq} = \sqrt{200(.5)(.5)} = \sqrt{50} = 7.071$

c. $z = \dfrac{x-\mu}{\sigma} = \dfrac{110-100}{7.071} = 1.41$

d. $P(x \le 110) \approx P\left(z \le \dfrac{(110+.5)-100}{7.071}\right) = P(z \le 1.48) = .5 + .4306 = .9306$

(Using Table III, Appendix A)

4.188 a. $\mu_{\bar{x}}$ is the mean of the sampling distribution of $\bar{x}$. $\mu_{\bar{x}} = \mu = 97,300$

b. $\sigma_{\bar{x}}$ is the standard deviation of the sampling distribution of $\bar{x}$.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{30,000}{\sqrt{50}} = 4,242.6407$$

c. By the Central Limit Theorem (n = 50), the sampling distribution of $\bar{x}$ is approximately normal.

d. $z = \dfrac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} = \dfrac{89,500-97,300}{4,242.6407} = -1.84$

e. $P(\bar{x} > 89,500) = P(z > -1.84) = .5 + .4671 = .9671$ (Using Table III, Appendix A)

4.190 a. Let x = number of beach trees damaged by fungi in 20 trials. Then x is a binomial random variable with n = 20 and p = .25.

$$\begin{aligned} P(x<10) &= P(x=0) + P(x=1) + \cdots + P(x=9) \\ &= \binom{20}{0}.25^0.75^{20} + \binom{20}{1}.25^1.75^{19} + \binom{20}{2}.25^2.75^{18} + \cdots + \binom{20}{9}.25^9.75^{11} \\ &= .0032 + .0211 + .0669 + .1339 + .1897 + .2023 + .1686 + .1124 + .0609 + .0271 \\ &= .9861 \end{aligned}$$

b. $$\begin{aligned} P(x>15) &= P(x=16) + P(x=17) + \cdots + P(x=20) \\ &= \binom{20}{16}.25^{16}.75^4 + \binom{20}{17}.25^{17}.75^3 + \binom{20}{18}.25^{18}.75^2 + \cdots + \binom{20}{20}.25^{20}.75^0 \\ &= .000000356 + .000000027 + .000000001 + 0 + 0 = .000000384 \end{aligned}$$

c. $E(x) = \mu = np = 20(.25) = 5$

4.192 a. Let x = number of democratic regimes that allow a free press in 50 trials. For this problem, $p = .8$.

$$\mu = E(x) = np = 50(.8) = 40$$

We would expect 40 democratic regimes out of the 50 to have a free press.

$$\sigma = \sqrt{np(1-p)} = \sqrt{50(.8)(.2)} = 2.828$$

We would expect most observations to fall within 2 standard deviations of the mean:

$$\mu \pm 2\sigma \Rightarrow 40 \pm 2(2.828) \Rightarrow 40 \pm 5.656 \Rightarrow (34.344,\ 45.656)$$

We would expect to see anywhere between 35 and 45 democratic regimes to have a free press out of a sample of 50.

b. Let x = number of non-democratic regimes that allow a free press in 50 trials. For this problem, $p = .1$.

$$\mu = E(x) = np = 50(.1) = 5$$

We would expect 5 non-democratic regimes out of the 50 to have a free press.

$$\sigma = \sqrt{np(1-p)} = \sqrt{50(.1)(.9)} = 2.121$$

We would expect most observations to fall within 2 standard deviations of the mean:

$$\mu \pm 2\sigma \Rightarrow 5 \pm 2(2.121) \Rightarrow 5 \pm 4.242 \Rightarrow (0.758,\ 9.242)$$

We would expect to see anywhere between 1 to 9 non-democratic regimes to have a free press out of a sample of 50.

4.194 Let x_1 = score on the blue exam. Then x_1 is approximately normal with $\mu_1 = 53\%$ and $\sigma_1 = 15\%$.

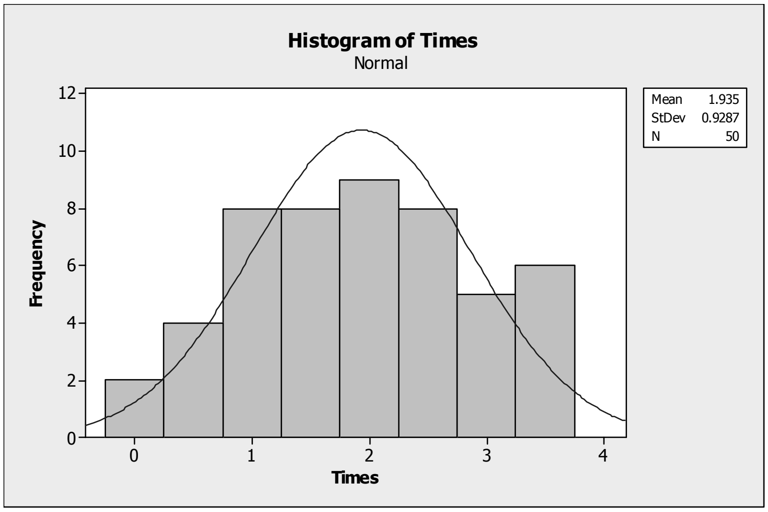
$$P(x_1 < 20\%) = P\left(z < \frac{20-53}{15}\right) = P(z < -2.20) = .5 - .4861 = .0139$$

Let x_2 = score on the red exam. Then x_2 is approximately normal with $\mu_2 = 39\%$ and $\sigma_2 = 12\%$.

$$P(x_2 < 20\%) = P\left(z < \frac{20-39}{12}\right) = P(z < -1.58) = .5 - .4429 = .0571$$

Since the probability of scoring below 20% on the red exam is greater than the probability of scoring below 20% on the blue exam, it is more likely that a student will score below 20% on the red exam.

4.196 First, we will graph the data. Using MINITAB, a histogram of the data with a normal distribution superimposed on the graph is:



The data are somewhat mound shaped. The data could be normally distributed.

The descriptive statistics of the data are:

Descriptive Statistics: Times

```
Variable   N   Mean  StDev  Minimum     Q1  Median     Q3  Maximum
Times     50  1.935  0.929   0.0100  1.218   1.835  2.645    3.500
```

We know if the data are normally distributed, then about 68% of the observations will fall within 1 standard deviations of the mean, about 95% of the observations will fall within 2 standard deviations of the mean, and about all of the observations will fall within 3 standard deviations of the mean.

$\bar{x} \pm s \Rightarrow 1.935 \pm .929 \Rightarrow (1.006,\ 2.864)$. 33 of the 50 or 33/50 = .66 of the observations are within 1 standard deviation of the mean. This is close to the .68 if the data are normal.

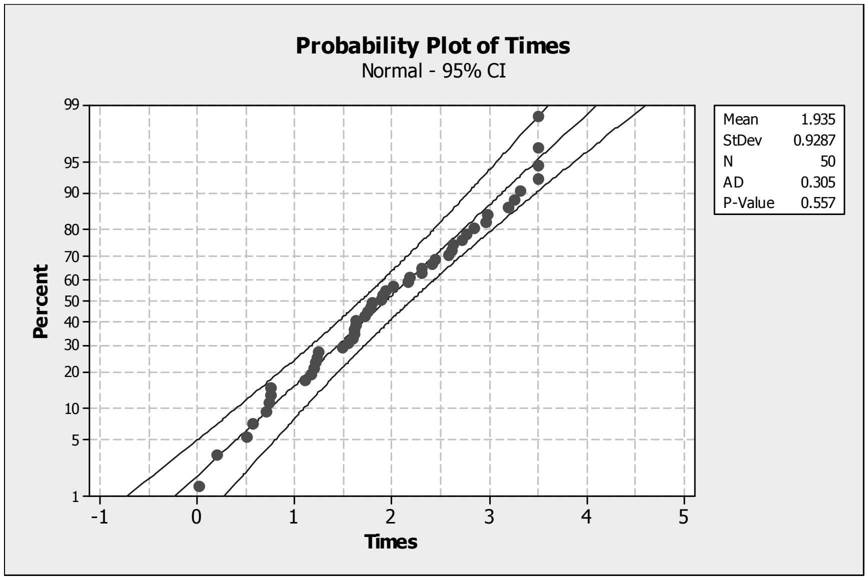
$\bar{x} \pm 2s \Rightarrow 1.935 \pm 2(.929) \Rightarrow 1.935 \pm 1.858 \Rightarrow (.077,\ 3.793)$. 49 of the 50 or 49/50 = .98 of the observations are within 2 standard deviations of the mean. This is higher than the .95 if the data are normal.

$\bar{x} \pm 3s \Rightarrow 1.935 \pm 3(.929) \Rightarrow 1.935 \pm 2.787 \Rightarrow (-.852,\ 4.722)$. 50 of the 50 or 59/50 = 1.00 of the observations are within 3 standard deviations of the mean. This is equal to the 1.00 if the data are normal.

The data appear to not be exactly normal.

$\text{IQR} = 3.5 - 1.218 = 2.282$ and s = .929. $\frac{IQR}{s} = \frac{2.282}{.929} = 2.456$. This is much larger than the 1.3 expected is the data are normal. This indicates that the data are probably not normal.

Finally, the probability plot of the data using MINITAB is:



If the data are normal, then the normal probability plot should be a straight line. On the right side of the graph, the plot of the data is almost a vertical line. There is some indication that the data are not normally distributed.

Thus, it appears that the data may not be normal.

4.198 By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with $\mu_{\bar{x}} = \mu = -2$ and $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{.3}{\sqrt{42}} = .0463$.

a. $$P(\bar{x} > -2.05) = P\left(z > \frac{-2.05-(-2)}{.0463}\right) = P(z > -1.08) = .5 + .3599 = .8599$$

(Using Table III, Appendix A.)

b. $$P(-2.20 < \bar{x} < -2.10) = P\left(\frac{-2.20-(-2)}{.0463} < z < \frac{-2.10-(-2)}{.0463}\right) = P(-4.32 < z < -2.16)$$
$$= .5 - .4846 = .0154$$

(Using Table III, Appendix A.)

4.200 From Table III, Appendix A, and $\sigma = .4$:

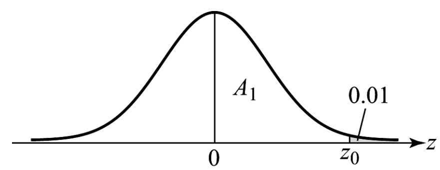
$$P(x > 6) = .01$$

$$P(x > 6) = P\left(z > \frac{6-\mu}{.4}\right) = P(z > z_0) = .01$$

$$A_1 = .5 - .01 = .4900$$

From Table III, $z_0 = 2.33$

$$2.33 = \frac{6-\mu}{.4} \Rightarrow \mu = 6 - 2.33(.4) = 5.068$$



4.202 Let x = number of disasters in 25 trials. If NASA's assessment is correct, then x is a binomial random variable with $n = 25$ and $p = 1 / 60{,}000 = .00001667$. If the Air Force's assessment is correct, then x is a binomial random variable with $n = 25$ and $p = 1 / 35 = .02857$.

If NASA's assessment is correct, then the probability of no disasters in 25 missions would be:

$$P(x = 0) = \binom{25}{0}(1/60{,}000)^0(59{,}999/60{,}000)^{25} = .9996$$

Thus, the probability of at least one disaster would be

$$P(x \geq 1) = 1 - P(x = 0) = 1 - .9996 = .0004$$

If the Air Force's assessment is correct, then the probability of no disasters in 25 missions would be:

$$P(x = 0) = \binom{25}{0}(1/35)^0(34/35)^{25} = .4845$$

Thus, the probability of at least one disaster would be

$$P(x \geq 1) = 1 - P(x = 0) = 1 - .4845 = .5155$$

One disaster actually did occur. If NASA's assessment was correct, it would be almost impossible for at least one disaster to occur in 25 trials. If the Air Force's assessment was correct, one disaster in 25 trials would not be an unusual event. Thus, the Air Force's assessment appears to be appropriate.

4.204 Answers will vary. We are to assume that the fecal bacteria concentrations of water specimens follow an approximate normal distribution. Now, suppose that the distribution of the fecal bacteria concentration at a beach is normal with a true mean of 360 with a standard deviation of 40. If only a single sample was selected, then the probability of getting an observation at the 400 level or higher would be:

$$P(x \geq 400) = P\left(z \geq \frac{400-360}{40}\right) = P(z \geq 1) = .5 - .3413 = .1587$$

(Using Table III, Appendix A)

Thus, even if the water is safe, the beach would be closed approximately 15.87% of the time.

On the other hand, if the mean was 440 and the standard deviation was still 40, then the probability of getting a single observation less than the 400 level would also be .1587. Thus, the beach would remain open approximately 15.78% of the time when it should be closed.

Now, suppose we took a random sample of 64 water specimens. The sampling distribution of $\bar{x}$ is approximately normal by the Central Limit Theorem with $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{64}} = 5.$

If $\mu = 360$, $P(\bar{x} \leq 400) = P\left(z \leq \frac{400-440}{5}\right) = P(z \leq -8) \approx .5 - .5 = 0$. Thus, the beach would never be shut down if the water was actually safe if we took samples of size 64.

If $\mu = 440$, $P(\bar{x} \geq 400) = P\left(z \geq \frac{400-360}{5}\right) = P(z \geq 8) \approx .5 - .5 = 0$. Thus, the beach would never be left open if the water was actually unsafe if we took samples of size 64.

The single sample standard can lead to unsafe decisions or inconvenient decisions, but is much easier to collect than samples of size 64.