Sampling Distributions

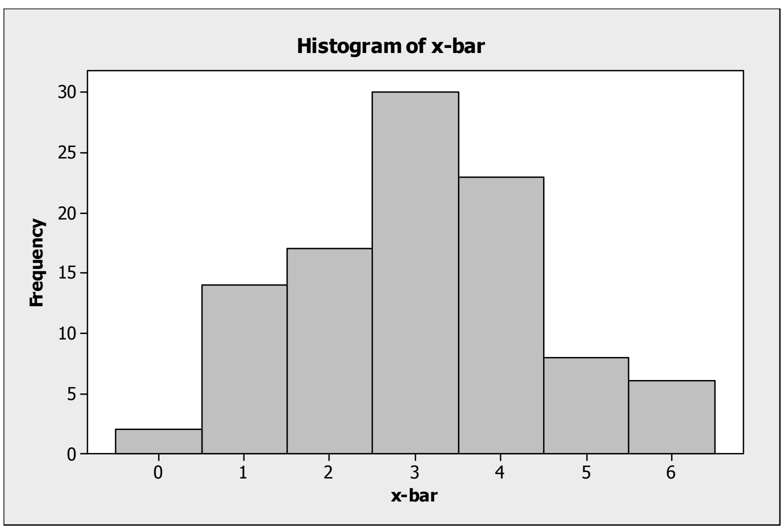
Chapter 6

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

6.4 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 6.3e. This histogram is not exactly the same because it is based on a sample size of only 100.

6.6 $E(x) = \mu = \sum xp(x) = 1(.2) + 2(.3) + 3(.2) + 4(.2) + 5(.1) = .2 + .6 + .8 + .5 = 2.7$

From Exercise 6.5, the sampling distribution of $\bar{x}$ is

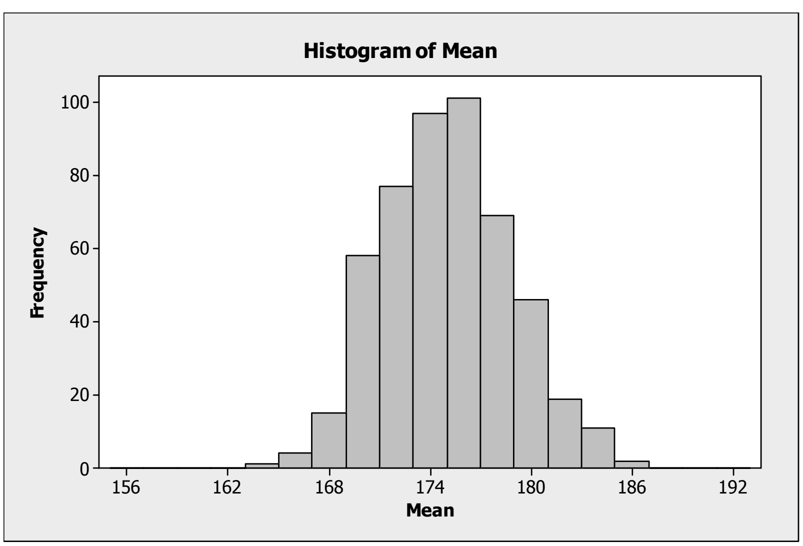
$\bar{x}$	1	1.5	2	2.5	3	3.5	4	4.5	5
$p(\bar{x})$	.04	.12	.17	.20	.20	.14	.08	.04	.01

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

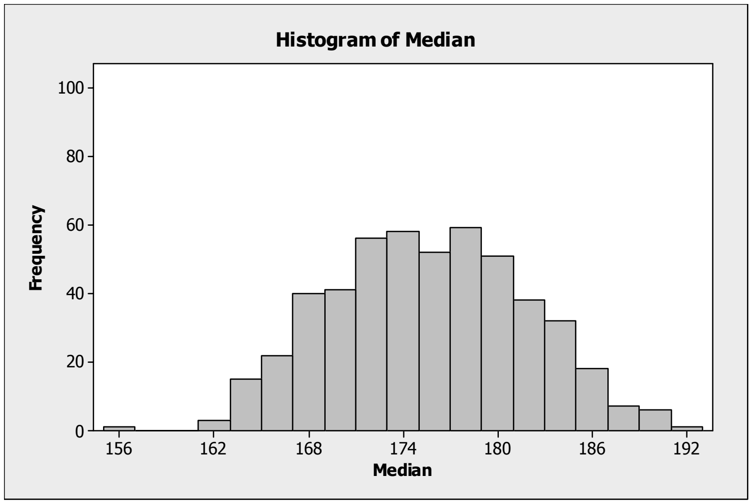
6.8 a. Answers will vary. MINITAB was used to generate 500 samples of size $n = 15$ observations from a uniform distribution over the interval from 150 to 200. The first 10 samples along with the sample means are shown in the table below:

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.

6.10 A point estimator of a population parameter is a rule or formula that tells us how to use the sample data to calculate a single number that can be used as an *estimate* of the population parameter.

6.12 The MVUE is the minimum variance unbiased estimator. The MUVE for a parameter is an unbiased estimator of the parameter that has the minimum variance of all unbiased estimators.

6.14 a. $\mu = \sum xp(x) = 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right) + 4\left(\frac{1}{3}\right) = \frac{5}{3} = 1.667$

$$\sigma^2 = \sum (x-\mu)^2 p(x) = \left(0-\frac{5}{3}\right)^2\left(\frac{1}{3}\right) + \left(1-\frac{5}{3}\right)^2\left(\frac{1}{3}\right) + \left(4-\frac{5}{3}\right)^2\left(\frac{1}{3}\right) = \frac{78}{27} = 2.889$$

b.

Sample	$\bar{x}$	Probability
0, 0	0	1/9
0, 1	0.5	1/9
0, 4	2	1/9
1, 0	0.5	1/9
1, 1	1	1/9
1, 4	2.5	1/9
4, 0	2	1/9
4, 1	2.5	1/9
4, 4	4	1/9

$\bar{x}$	**Probability**
0	1/9
0.5	2/9
1	1/9
2	2/9
2.5	2/9
4	1/9

c. $$E(\bar{x}) = \sum \bar{x}p(\bar{x}) = 0\left(\frac{1}{9}\right) + 0.5\left(\frac{2}{9}\right) + 1\left(\frac{1}{9}\right) + 2\left(\frac{2}{9}\right) + 2.5\left(\frac{2}{9}\right) + 4\left(\frac{1}{9}\right) = \frac{15}{9} = \frac{5}{3} = 1.667$$

Since $E(\bar{x}) = \mu$, $\bar{x}$ is an unbiased estimator for μ.

d. Recall that $s^2 = \dfrac{\sum x^2 - \dfrac{\left(\sum x\right)^2}{n}}{n-1}$

For the first sample, $s^2 = \dfrac{0^2 + 0^2 - \dfrac{(0+0)^2}{2}}{2-1} = 0$.

For the second sample, $s^2 = \dfrac{1^2 + 0^2 - \dfrac{(1+0)^2}{2}}{2-1} = \dfrac{1 - \dfrac{(1)^2}{2}}{2-1} = \dfrac{1}{2}$

The rest of the values are shown in the table below.

Sample	s^2	**Probability**
0, 0	0	1/9
0, 1	0.5	1/9
0, 4	8	1/9
1, 0	0.5	1/9
1, 1	0	1/9
1, 4	4.5	1/9
4, 0	8	1/9
4, 1	4.5	1/9
4, 4	0	1/9

The sampling distribution of s^2 is:

s^2	Probability
0	3/9
0.5	2/9
4.5	2/9
8	2/9

e. $$E(s^2)=\sum s^2 p(s^2)=0\left(\frac{3}{9}\right)+0\left(\frac{2}{9}\right)+4.5\left(\frac{2}{9}\right)+8\left(\frac{2}{9}\right)=\frac{26}{9}=2.889$$

Since $E(s^2)=\sigma^2$, s^2 is an unbiased estimator for σ^2.

6.16 a. $$\mu=\sum xp(x)=0\left(\frac{1}{3}\right)+1\left(\frac{1}{3}\right)+2\left(\frac{1}{3}\right)=1$$

b.

Sample		Probability	Sample		Probability
0, 0, 0	0	1/27	1, 1, 2	4/3	1/27
0, 0, 1	1/3	1/27	1, 2, 0	1	1/27
0, 0, 2	2/3	1/27	1, 2, 1	4/3	1/27
0, 1, 0	1/3	1/27	1, 2, 2	5/3	1/27
0, 1, 1	2/3	1/27	2, 0, 0	2/3	1/27
0, 1, 2	1	1/27	2, 0, 1	1	1/27
0, 2, 0	2/3	1/27	2, 0, 2	4/3	1/27
0, 2, 1	1	1/27	2, 1, 0	1	1/27
0, 2, 2	4/3	1/27	2, 1, 1	4/3	1/27
1, 0, 0	1/3	1/27	2, 1, 2	5/3	1/27
1, 0, 1	2/3	1/27	2, 2, 0	4/3	1/27
1, 0, 2	1	1/27	2, 2, 1	5/3	1/27
1, 1, 0	2/3	1/27	2, 2, 2	2	1/27
1, 1, 1	1	1/27			

From the above table, the sampling distribution of the sample mean would be:

$\bar{x}$	Probability
0	1/27
1/3	3/27
2/3	6/27
1	7/27
4/3	6/27
5/3	3/27
2	1/27

c.

Sample	m	Probability	Sample	m	Probability
0, 0, 0	0	1/27	1, 1, 2	1	1/27
0, 0, 1	0	1/27	1, 2, 0	1	1/27
0, 0, 2	0	1/27	1, 2, 1	1	1/27
0, 1, 0	0	1/27	1, 2, 2	2	1/27
0, 1, 1	1	1/27	2, 0, 0	0	1/27
0, 1, 2	1	1/27	2, 0, 1	1	1/27
0, 2, 0	0	1/27	2, 0, 2	2	1/27
0, 2, 1	1	1/27	2, 1, 0	1	1/27
0, 2, 2	2	1/27	2, 1, 1	1	1/27
1, 0, 0	0	1/27	2, 1, 2	2	1/27
1, 0, 1	1	1/27	2, 2, 0	2	1/27
1, 0, 2	1	1/27	2, 2, 1	2	1/27
1, 1, 0	1	1/27	2, 2, 2	2	1/27
1, 1, 1	1	1/27			

From the above table, the sampling distribution of the sample median would be:

m	Probability
0	7/27
1	13/27
2	7/27

d. $$E(\bar{x})=\sum \bar{x}p(\bar{x})=0\left(\frac{1}{27}\right)+\frac{1}{3}\left(\frac{3}{27}\right)+\frac{2}{3}\left(\frac{6}{27}\right)+1\left(\frac{7}{27}\right)+\frac{4}{3}\left(\frac{6}{27}\right)+\frac{5}{3}\left(\frac{3}{27}\right)+2\left(\frac{1}{27}\right)=1$$

Since $E(\bar{x})=\mu$, $\bar{x}$ is an unbiased estimator for μ.

$$E(m)=\sum mp(m)=0\left(\frac{7}{27}\right)+1\left(\frac{13}{27}\right)+2\left(\frac{7}{27}\right)=1$$

Since $E(m) = \mu$, m is an unbiased estimator for μ.

e. $$\sigma_{\bar{x}}^2=\sum(\bar{x}-\mu)^2 p(\bar{x})=(0-1)^2\left(\frac{1}{27}\right)+\left(\frac{1}{3}-1\right)^2\left(\frac{3}{27}\right)+\left(\frac{2}{3}-1\right)^2\left(\frac{6}{27}\right)+(1-1)^2\left(\frac{7}{27}\right)$$
$$+\left(\frac{4}{3}-1\right)^2\left(\frac{6}{27}\right)+\left(\frac{5}{3}-1\right)^2\left(\frac{3}{27}\right)+(2-1)^2\left(\frac{1}{27}\right)=\frac{2}{9}=.2222$$

$$\sigma_m^2=\sum(m-1)^2 p(m)=(0-1)^2\left(\frac{7}{27}\right)+(1-1)^2\left(\frac{13}{27}\right)+(2-1)^2\left(\frac{7}{27}\right)=\frac{14}{27}=.5185$$

f. Since both the sample mean and median are unbiased estimators and the variance is smaller for the sample mean, it would be the preferred estimator of μ.

6.18 a. The mean of the random variable x is:

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

From Exercise 6.5, the sampling distribution of $\bar{x}$ is:

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

The mean of the sampling distribution of $\bar{x}$ is:

$$E(\bar{x}) = \sum \bar{x}p(\bar{x}) = 1(.04) + 1.5(.12) + 2(.17) + 2.5(.20) + 3(.20) + 3.5(.14)$$
$$+ 4(.08) + 4.5(.04) + 5(.01) = 2.7$$

Since $E(\bar{x}) = E(x) = \mu$, $\bar{x}$ is an unbiased estimator of μ.

b. The variance of the sampling distribution of $\bar{x}$ is:

$$\sigma_{\bar{x}}^2 = \sum (\bar{x} - \mu)^2 p(\bar{x}) = (1 - 2.7)^2(.04) + (1.5 - 2.7)^2(.12) + (2 - 2.7)^2(.17)$$
$$+ (2.5 - 2.7)^2(.20) + (3 - 2.7)^2(.20) + (3.5 - 2.7)^2(.14)$$
$$+ (4 - 2.7)^2(.08) + (4.5 - 2.7)^2(.04) + (5 - 2.7)^2(.01) = .805$$

c. $\mu \pm 2\sigma_{\bar{x}} \Rightarrow 2.7 \pm 2\sqrt{.805} \Rightarrow 2.7 \pm 1.794 \Rightarrow (.906, 4.494)$

$P(.906 \le \bar{x} \le 4.494) = .04 + .12 + .17 + .2 + .2 + .14 + .08 = .95$

6.20 The mean of the random variable x is:

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

From Exercise 6.7, the sampling distribution of the sample median is:

m	1	1.5	2	2.5	3	3.5	4	4.5	5
p(m)	.04	.12	.17	.20	.20	.14	.08	.04	.01

The mean of the sampling distribution of the sample median *m* is:

$$E(m) = \sum mp(m) = 1(.04) + 1.5(.12) + 2(.17) + 2.5(.20) + 3(.20) + 3.5(.14) + 4(.08) = + 4.5(.04) + 5(.01) = 2.7$$

Since $E(m) = \mu$, *m* is an unbiased estimator of μ.

6.22 The mean of the sampling distribution of $\bar{x}$, $\mu_{\bar{x}}$, is the same as the mean of the population from which the sample is selected.

6.24 Another name given to the standard deviation of $\bar{x}$ is the <u>standard error</u>.

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

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

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

6.32 For this population and sample size,

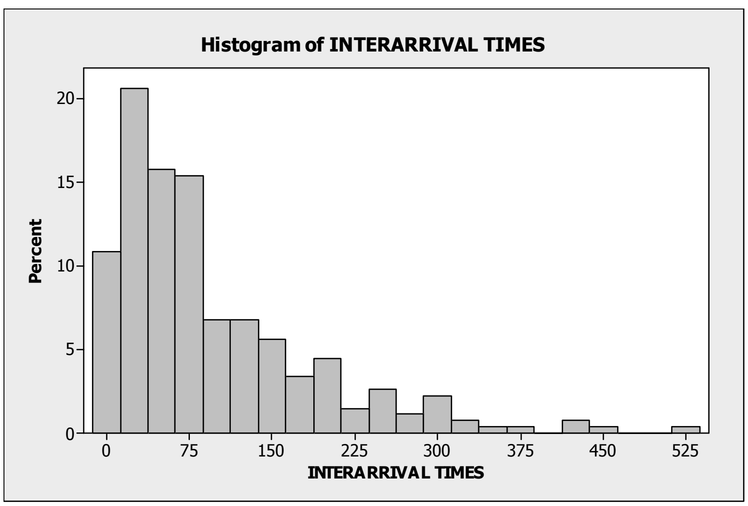
$$E(\bar{x}) = \mu_{\bar{x}} = \mu = 100, \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{900}} = \frac{1}{3}$$

a. Approximately 95% of the time, $\bar{x}$ will be within two standard deviations of the mean, i.e., $\mu \pm 2\sigma \Rightarrow 100 \pm 2\left(\frac{1}{3}\right) \Rightarrow 100 \pm \frac{2}{3} \Rightarrow (99.33, 100.67)$. Almost all of the time, the sample mean will be within three standard deviations of the mean, i.e., $\mu \pm 3\sigma \Rightarrow 100 \pm 3\left(\frac{1}{3}\right) \Rightarrow 100 \pm 1 \Rightarrow (99, 101)$.

b. No more than three standard deviations, i.e., $3\left(\frac{1}{3}\right) = 1$

c. No, the previous answer only depended on the standard deviation of the sampling distribution of the sample mean, not the mean itself.

6.34 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 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{91.54}{\sqrt{40}} = 14.47$.

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

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

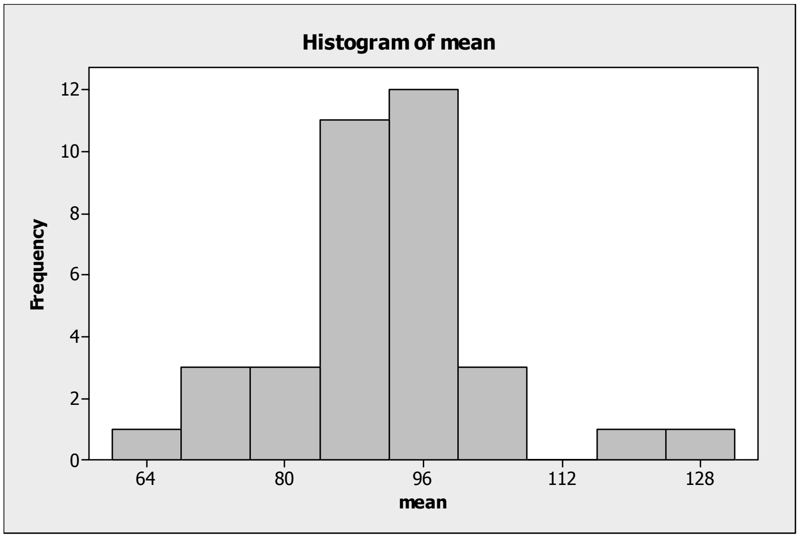
17.698	153.772	20.229	20.740	49.822
38.303	7.995	148.289	267.534	13.411
63.037	4.626	105.527	163.349	5.016
47.203	38.791	36.421	120.624	41.905
88.855	2.954	17.886	428.507	82.760
84.203	54.291	289.669	11.062	303.903
259.861	126.448	53.104	32.193	37.720
256.385	185.598	30.720	40.760	72.648

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.

6.36 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 IV, 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 IV, 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 IV, Appendix A)

6.38 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 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{76}} = 1.147$.

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

6.40 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. $V(x) = \sigma^2 = \frac{(d-c)^2}{12} = \frac{(3-1)^2}{12} = \frac{1}{3}$. 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 IV, 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 IV, Appendix A)

6.42 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 IV, 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 IV, 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 IV, 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 IV, Appendix A)

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

6.44 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 IV, 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.

6.46 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 $z = \frac{6.5 - 11.92}{.295} = -18.37$

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.

6.48 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.

6.50 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.

6.52 a. As the sample size increases, the standard error will decrease. This property is important because we know that the larger the sample size, the less variable our estimator will be. Thus, as n increases, our estimator will tend to be closer to the parameter we are trying to estimate.

b. This would indicate that the statistic would not be a very good estimator of the parameter. If the standard error is not a function of the sample size, then a statistic based on one observation would be as good an estimator as a statistic based on 1000 observations.

c. $\bar{x}$ would be preferred over A as an estimator for the population mean. The standard error of $\bar{x}$ is smaller than the standard error of A.

d. The standard error of $\bar{x}$ is $\dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{64}} = 1.25$ and the standard error of A is $\dfrac{10}{\sqrt[3]{64}} = 2.5$.

If the sample size is sufficiently large, the Central Limit Theorem says the distribution of $\bar{x}$ is approximately normal. Using the Empirical Rule, approximately 68% of all the values of $\bar{x}$ will fall between $\mu - 1.25$ and $\mu + 1.25$. Approximately 95% of all the values of $\bar{x}$ will fall between $\mu - 2.50$ and $\mu + 2.50$. Approximately all of the values of will fall between $\mu - 3.75$ and $\mu + 3.75$.

Using the Empirical Rule, approximately 68% of all the values of A will fall between $\mu - 2.50$ and $\mu + 2.50$. Approximately 95% of all the values of A will fall between $\mu - 5.00$ and $\mu + 5.00$. Approximately all of the values of A will fall between $\mu - 7.50$ and $\mu + 7.50$.

6.54 a. First we must compute μ and σ. The probability distribution for x is:

x	$p(x)$
1	.3
2	.2
3	.2
4	.3

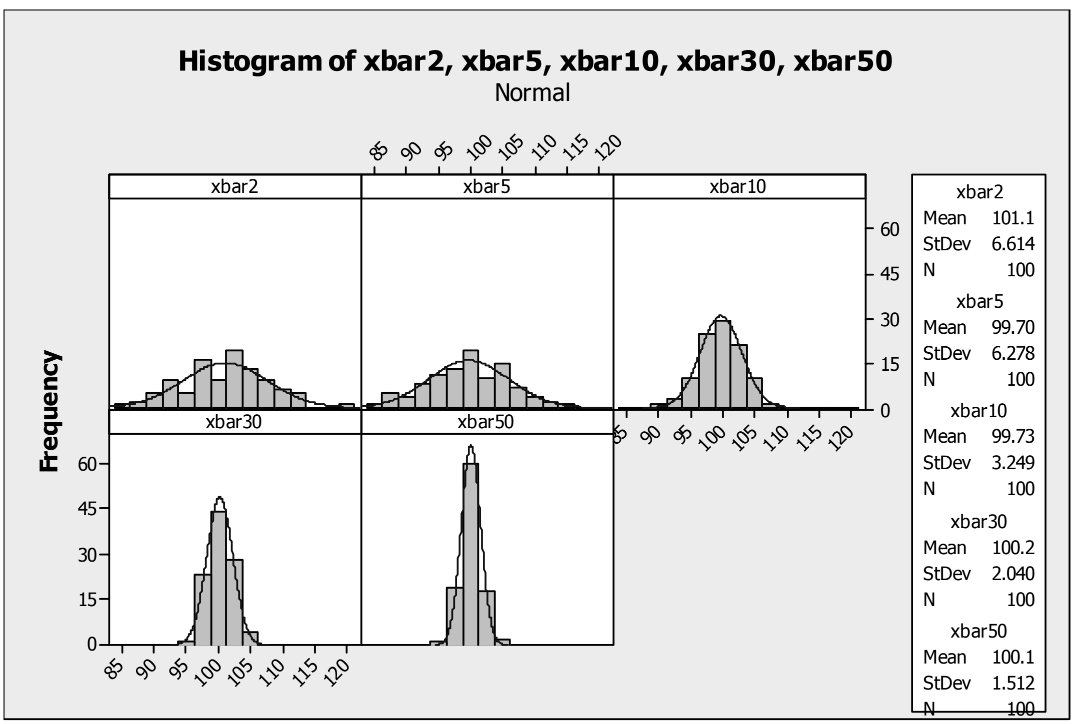
$$\mu = E(x) = \sum xp(x) = 1(.3) + 2(.2) + 3(.2) + 4(.3) = 2.5$$

$$\sigma^2 = E\sum(x-\mu)^2 = \sum(x-\mu)^2 p(x) = (1-2.5)^2(.3) + (2-2.5)^2(.2) + (3-2.5)^2(.2)$$
$$(4-2.5)^2(.3) + = 1.45$$

$$\mu_{\bar{x}} = \mu = 2.5\,,\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{1.45}}{\sqrt{40}} = .1904$$

b. By the Central Limit Theorem, the distribution of is approximately normal. The sample size, $n = 40$, is sufficiently large. Yes, the answer depends on the sample size.

6.56 Answers will vary. One hundred samples of size $n = 2$ were selected from a normal distribution with a mean of 100 and a standard deviation of 10. The process was repeated for samples of size $n = 5$, $n = 10$, $n = 30$, and $n = 50$. For each sample, the value of $\bar{x}$ was computed. Using MINITAB, the histograms for each set of 100 $\bar{x}$'s were constructed:



The sampling distribution of $\bar{x}$ is normal regardless of the sample size because the population we sampled from was normal. Notice that as the sample size n increases, the variances of the sampling distributions decrease.

6.58 a. Tossing a coin two times can result in:

2 heads (2 ones)
2 tails (2 zeros)
1 head, 1 tail (1 one, 1 zero)

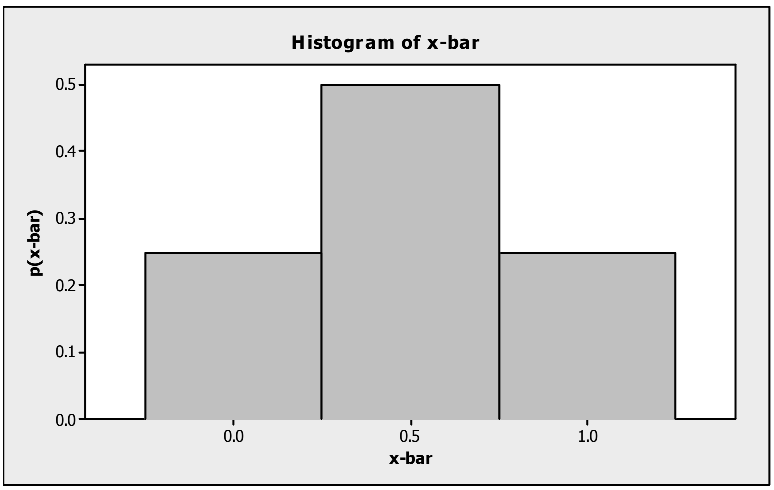
b. $\bar{x}_{2 \text{ heads}} = 1$; $\bar{x}_{2 \text{ tails}} = 0$; $\bar{x}_{1\text{H},1\text{T}} = \frac{1}{2}$

c. There are four possible combinations for one coin tossed two times, as shown below:

Coin Tosses	$\bar{x}$
H, H	1
H, T	1/2
T, H	1/2
T, T	0

$\bar{x}$	$P(\bar{x})$
0	1/4
1/2	1/2
1	1/4

d. The sampling distribution of $\bar{x}$ is given in the histogram shown.



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

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

6.62 a. $\mu_{\bar{x}} = \mu = 5.1$; $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6.1}{\sqrt{150}} = .4981$

b. Because the sample size is large, n = 150, the Central Limit Theorem says that the sampling distribution of $\bar{x}$ is approximately normal.

c. $P(\bar{x} > 5.5) = P\left(z > \frac{5.5 - 5.1}{.4981}\right) = P(z > .80) = .5 - .2881 = .2119$

(using Table IV, Appendix A)

d. $P(4 < \bar{x} < 5) = P\left(\frac{4 - 5.1}{.4981} < z < \frac{5 - 5.1}{.4981}\right) = P(-2.21 < z < -.20) = .4864 - .0793 = .4071$

(using Table IV, Appendix A)

6.64 a. The mean, μ, diameter of the bearings is unknown with a standard deviation of $\sigma = .001$ inch. Assuming that the distribution of the diameters of the bearings is normal, the sampling distribution of the sample mean is also normal. The mean and variance of the distribution are:

$$\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.001}{\sqrt{25}} = .0002$$

Having the sample mean fall within .0001 inch of μ implies

$$|\bar{x} - \mu| \leq .0001 \quad \text{or} \quad -.0001 \leq \bar{x} - \mu \leq .0001$$

$P(-.0001 \leq \bar{x} - \mu \leq .0001) =$

$$P(-.0001 \leq \bar{x} - \mu \leq .0001) = P\left(\frac{-.0001}{.0002} \leq z \leq \frac{.0001}{.0002}\right) = P(-.50 \leq z \leq .50)$$

$$= 2P(0 \leq z \leq .50) = 2(.1915) = .3830$$

(using Table IV, Appendix A)

b. The approximation is unlikely to be accurate. In order for the Central Limit Theorem to apply, the sample size must be sufficiently large. For a very skewed distribution, n = 25 is not sufficiently large, and thus, the Central Limit Theorem will not apply.

6.66 a. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal with $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{344}} = .3235$.

b. If $\mu = 18.5$, $P(\bar{x} > 19.1) = P\left(z > \frac{19.1 - 18.5}{.3235}\right) = P(z > 1.85) = .5 - .4678 = .0322$

(Using Table IV, Appendix A)

c. If $\mu = 19.5$, $P(\bar{x} > 19.1) = P\left(z > \frac{19.1 - 19.5}{.3235}\right) = P(z > -1.24) = .5 + .3925 = .8925$

(Using Table IV, Appendix A)

d. $P(\bar{x} > 19.1) = P\left(z > \frac{19.1-\mu}{.3235}\right) = .5$

We know that $P(z > 0) = .5$. Thus, $\frac{19.1-\mu}{.3235} = 0 \Rightarrow \mu = 19.1$

e. $P(\bar{x} > 19.1) = P\left(z > \frac{19.1-\mu}{.3235}\right) = .2$

Thus, μ must be less than 19.1. If μ = 19.1, then $P(P(\bar{x} > 19.1) = .5$.
Since $P(\bar{x} > \mu) < .5$, then $\mu < 19.1$.

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

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

6.70 a. By the Central Limit Theorem, the sampling distribution of is approximately normal, regardless of the shape of the distribution of the verbal IQ scores. The mean is $\mu_{\bar{x}} = \mu = 107$, and the standard deviation is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{84}} = 1.637$. This does not depend on the shape of the distribution of verbal IQ scores.

b. $P(\bar{x} \geq 110) = P\left(z \geq \frac{110-107}{1.637}\right) = P(z \geq 1.83) = .5 - .464 = .0336$
(using Table IV, Appendix A)

c. No. If the mean and standard deviation for the nondelinquent juveniles were the same as those for all juveniles, it would be very unlikely (probability = .0336) to observe a sample mean of 110 or higher.

6.72 a. If x is an exponential random variable, then $\mu = E(x) = \theta = 60$. The standard deviation of x is $\sigma = \theta = 60$.

Then, $E(\bar{x}) = \mu_{\bar{x}} = \mu = 60$; $V(\bar{x}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{60^2}{100} = 36$

b. Because the sample size is fairly large, the Central Limit Theorem says that the sampling distribution of $\bar{x}$ is approximately normal.

c. $P(\bar{x} \le 30) = P\left(z \le \frac{30-60}{\sqrt{36}}\right) = P(z \le -5.0) \approx .5 - .5 = 0$

(Using Table IV, Appendix A)

6.74 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:

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

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

If $\mu = 440$, $P(\bar{x} \ge 400) = P\left(z \ge \frac{400-360}{5}\right) = P(z \ge 8) \approx .5 - .5 = 0$. 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.