7.1 Population - All people who signed a stop smoking card

Parameter - Proportion of people who signed card and actually quit smoking

Sample - Random sample of 1000 people who signed card

Statistic - $\hat{p} = .21$

7.2 Population - All individuals in the US

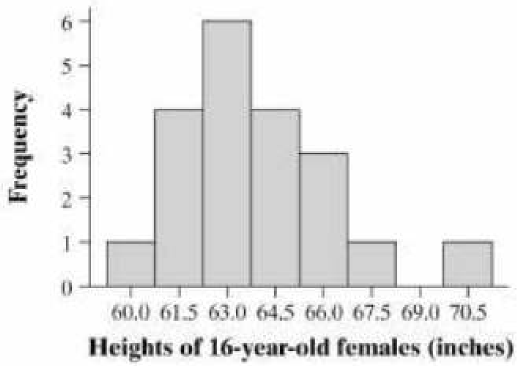
Parameter - Proportion of US population unemployed

Sample - Random sample of individuals from 55,000 US households

Statistic - $\hat{p} = .10$

	Parameter	Statistic
7.5	$\mu = 2.5003$	$\bar{x} = 2.5009$
7.6	$p = .41$	$\hat{p} = .33$
7.7	$p = .52$	$\hat{p} = .48$
7.8	$\mu = 63''$	$\bar{x} = 64.5''$

(b) Answers will vary. An example histogram is given below.



7.13 (a) The approximate sampling distribution is skewed to the right with a center at $9\left({}^{\circ}\mathrm{F}\right)^2$. The values vary from about 2 to 27.5 $\left({}^{\circ}\mathrm{F}\right)^2$. (b) A sample variance of 25 is quite large compared with what we would expect, since only one out of 500 SRSs had a variance that high. It suggests that the manufacturer's claim is false and that the thermostat actually has more variability than claimed.

7.14 (a) The approximate sampling distribution is reasonably symmetric and centered at $45.5^{\circ}\mathrm{F}$. The values vary from about 39 to $50^{\circ}\mathrm{F}$. (b) A sample minimum of $40^{\circ}\mathrm{F}$ is quite low compared with what we would expect. This suggests that the manufacturer's claim is false.

Aprox

7.15 (a) The population is the 12,000 students; the population distribution (Normal with mean 7.11 minutes and standard deviation 0.74 minutes) Aprox describes the time it takes randomly selected individuals to run a mile. (b) The sampling distribution (Normal with mean of 7.11 minutes and standard deviation of 0.074 minutes) describes the average mile-time for 100 randomly selected students. This is different from the population distribution in that it has a smaller standard deviation and it describes the mean of 100 mile times rather than individual mile times.

7.16 (a) The population is the 4000 beads in the container. 1000 of the beads are white and 3000 are red. (b) The distribution of the sample proportion is approximately Normal with mean 0.75 and standard deviation 0.06. The sample proportion is a numerical variable and so its distribution could be shown using a histogram or dotplot. The color of the individual beads is a categorical variable and so its distribution would be best shown with a bar graph.

7.17 (a) Since the smallest number of total tax returns (i.e., the smallest population) is still more than 10 times the sample size, the variability of the sample proportion will be (approximately) the same for all states. (b) Yes. It will change—the sample taken from Wyoming will be about the same size, but the sample in, for example, California will be considerably larger, and therefore the variability of the sample proportion will be smaller.

7.18 (a) A larger sample does not reduce the bias of a poll result. If the sampling technique results in bias, simply increasing the sample size will not reduce the bias. (b) A larger sample will reduce the variability of the result. More people means more information which means less variability.

7.19 (a) Graph (c) shows an unbiased estimator because the mean of the distribution is very close to the population parameter. (b) The graph in (b) shows the statistic that does the best job at estimating the parameter. Although it is biased, the bias is small and the statistic has very little variability.

Exercises, page 439:

7.27 (a) We would not be surprised to find 8 (32%) orange candies. From the graph in figure 7.11 there were a fair number of simulations in which there were 8 or fewer orange candies. On the other hand, there were only a couple of simulations where there were 5 (20%) or fewer so if this occurred, we should be surprised. (b) It is more surprising to get 32% orange candies in a sample of 50 than it is in a sample of 25. Comparing the graphs in figures 7.11 and 7.12, there were a fair number of simulations in 7.11 (sample size 25) with 32% or less, but very few in 7.12 (sample size 50) with 32% or less.

7.28 (a) We would be surprised to find 32% orange candies in this case. Very few of the simulations with sample size 25 had 32% or more orange candies. However, we would not be surprised to find 20% orange candies. This is very near the center of the distribution. (b) We would be surprised to find 32% orange candies in either case since neither simulation had many samples with 32% or more orange candies. However, it is even rarer when the sample size is 50.

7.29 (a) The mean of the sampling distribution is the same as the population proportion so $\mu_{\hat{p}} = p = 0.45$. (b) The standard deviation of the sampling distribution is

$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.45(0.55)}{25}} = 0.0995$. In this case the 10% condition is met because it is very likely true that there are ~~more than~~ 250 candies. (c) The sampling distribution is approximately Normal because $np = 25(0.45) = 11.25$ and $n(1-p) = 25(0.55) = 13.75$ are both at least 10. (d) If the sample size were 50 rather than 25, the sampling distribution would still be approximately Normal with mean 0.45, but the standard deviation would be $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.45(0.55)}{50}} = 0.0704$.

7.30 (a) The mean of the sampling distribution is the same as the population proportion so $\mu_{\hat{p}} = p = 0.15$. (b) The standard deviation of the sampling distribution is

$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.15(0.85)}{25}} = 0.0714$. In this case the 10% condition is met because it is very likely true that there are more than 250 candies. (c) The sampling distribution is not approximately Normal because $np = 25(0.15) = 3.75$ is less than 10. Note that $n(1-p) = 25(0.85) = 21.25$ is at least 10 but for the Normal approximation to be correct, both numbers must be at least 10. (d) If the sample size were 75 rather than 25, the sampling distribution would now be approximately Normal with mean 0.15 and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.15(0.85)}{75}} = 0.0412$, since $np = 75(0.15) = 11.25$ and $n(1-p) = 75(0.85) = 63.75$ are both at least 10.

7.31 (a) The 10% condition is not met here. Out of the population of 76 passengers, 10 people were screened (13%). This means that they sampled more than 10% of the population. (b) No. The Normal condition is also not met since the total sample size was 10. Necessarily, both np and $n(1-p)$ will be less than 10, violating the condition for Normality.

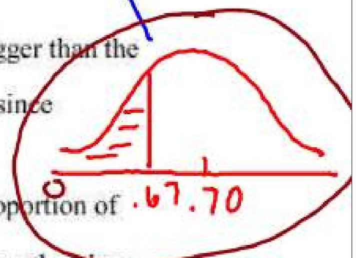
7.32 (a) The 10% condition is met here. We are drawing a sample of 7 out of 100 tiles. This is less than 10% of the population. (b) The Normal condition is not met here since the total sample size was 7. Necessarily, both np and $n(1-p)$ will be less than 10, violating the condition for Normality.

7.33 The Normal condition is not met here. $np = 15(0.3) = 4.5 < 10$. *Challenge*: Let X be the number of Hispanic workers in the sample. X has an approximate binomial distribution with $n = 15$ and $p = 0.3$.

$$P(X \le 3) = \binom{15}{0}(0.3)^0(1-0.3)^{15} + \cdots + \binom{15}{3}(0.3)^3(1-0.3)^{12} = 0.2969.$$

7.34 The 10% condition is not met here. The sample of 50 is more than 10% of the population (which is of size 316).

7.35 (a) The mean of the sampling distribution is the same as the population proportion so it is $\mu_{\hat{p}} = p = 0.70$. (b) The standard deviation of the sampling distribution is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.7(0.3)}{1012}} = 0.0144$. The population (all U.S. adults) is clearly at least 10 times as large as the sample (the 1012 surveyed adults) so the 10% condition is met. (c) The sampling distribution is approximately Normal since $np = 1012(0.70) = 708.4$ and $n(1-p) = 1012(0.30) = 303.6$ are both at least 10. (d) $P(\hat{p} \le 0.67) = P(z \le -2.08) = 0.0188$. This is a fairly unusual result if 70% of the population actually drink the cereal milk.



7.36 (a) The mean is the same as the population proportion so it is $\mu_{\hat{p}} = p = 0.4$. (b) The standard deviation is $\sigma_{\hat{p}} = \sqrt{\frac{0.4(0.6)}{1785}} = 0.0116$. Since the population is clearly at least 10 times bigger than the sample, the 10% condition is met. (c) The sampling distribution is approximately Normal since $np = 1785(0.4) = 714$ and $n(1-p) = 1785(0.6) = 1071$ are both at least 10. (d) $P(\hat{p} \ge 0.44) = P\left(z \ge \frac{0.44 - 0.4}{0.0116}\right) = P(z \ge 3.45) = 0.0003$. It is very unlikely, if the true proportion of people who attend church or synagogue is 0.40, that 44% or more will answer the poll saying that they attend church or synagogue.

7.37 Since the standard deviation is found by deviding by $\sqrt{n}$, using $4n$ for the sample size halves the standard deviation $\left(\sqrt{4n} = 2\sqrt{n}\right)$; we would need to sample $1012(4) = 4048$ adults.

7.38 Since the standard deviation is found by deviding by $\sqrt{n}$, using $9n$ for the sample size halves the standard deviation $\left(\sqrt{9n} = 3\sqrt{n}\right)$; we would need to sample $9(1785) = 16{,}065$ adults.

7.39 *State*: We want to find the probability that $\hat{p}$ is at least 0.75. In symbols, that's $P(\hat{p} \ge 0.75)$. *Plan*: We have an SRS of size 267 drawn from a population in which the proportion $p = 0.70$ of college women have been on a diet within the past 12 months. This means that $\mu_{\hat{p}} = 0.70$ and since the population clearly contains more than $267(10) = 2670$ college women, $\sigma_{\hat{p}} = \sqrt{\frac{0.7(0.3)}{267}} = 0.0280$. We also check the Normal condition: $np = 267(0.7) = 186.9$ and $n(1-p) = 267(0.3) = 80.1$ are both at least

40a) Sampling Distribution Normal?

$np \geq 10$? $n(1-p) \geq 10$?

$(500)(.14) \geq 10$? $(500)(.86) \geq 10$?

$70 \geq 10$ ✓ $430 \geq 10$ ✓

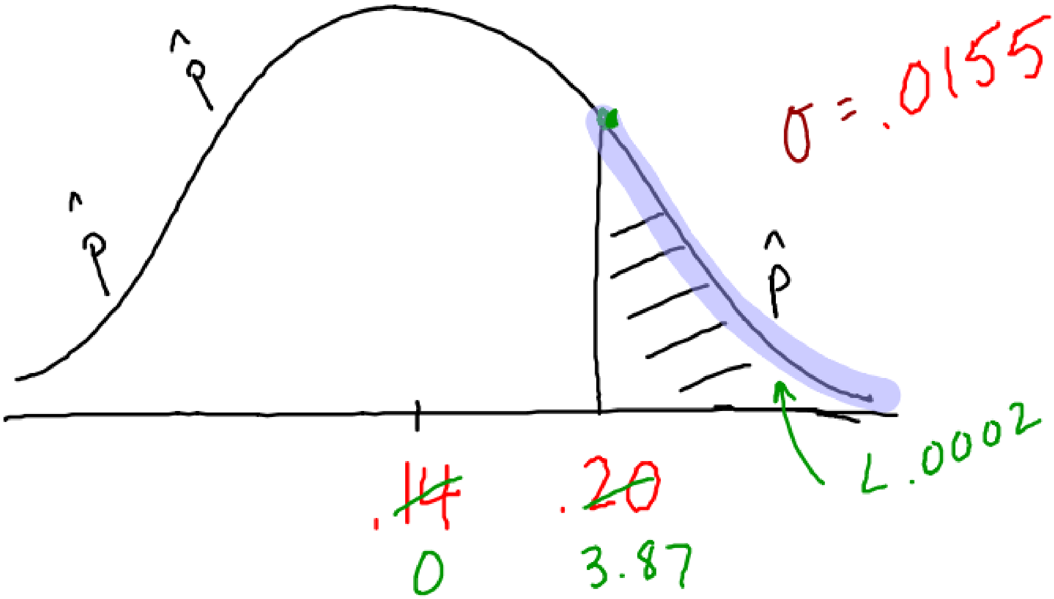
b) $N > 10n$?

$N > 10(500)$

$N > 5000$ registered Harley owners ✓

c) $\mu_{\hat{p}} = p = .14$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(.14)(.86)}{500}} = \underline{.0155}$$



normalcdf (.20, 1, .14, .0155) = .00005 ✓

★ Assuming 14% of all motorcycle owners have a Harley, the probability that 20% of this sample own a Harley is .00005 - <u>very</u> unlikely sample results are lower than the national percentage, but the sample was so small that such a difference could arise by chance even if the true campus proportion is the same.

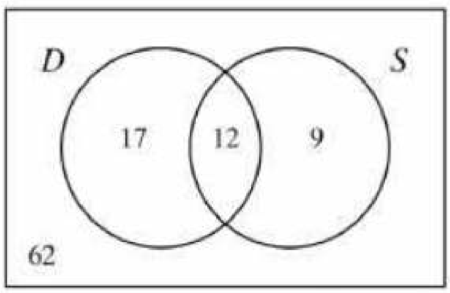
7.43 b.

7.44 c.

7.45 b.

7.46 b.

7.47



62% neither download nor share music files.

7.48 (a) Assign numbers 01-14 to the animals (01 to the desert tortoise, 02 to the Olive Ridley sea turtle,…, 14 to the San Francisco garter snake). Starting at line 111 in Table D, read pairs of numbers until you get three different numbers between 01 and 14. These numbers represent the animals chosen.
(b) Using Table D, the animals chosen are 12, 04, and 11 which represent the blunt-nosed leopard lizard, the flat-tailed horned lizard and the Coachella Valley fringe-toed lizard.

Section 7.3

Check Your Understanding, page 448:

1. $P(X > 270) = P\left(z > \frac{270-266}{16}\right) = P(z > 0.25) = 0.4013$

2. The mean of the sampling distribution of $\bar{x}$ is the same as the mean of the distribution of X so $\mu_{\bar{x}} = \mu_X = 266$ days.

3. First, we check the 10% condition. We are taking a sample of 6 pregnant women. There are clearly more than $10(6) = 60$ pregnant women so this condition is met. Therefore, the standard deviation of the sampling distribution is $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{16}{\sqrt{6}} = 6.532$ days.

4. $P(\bar{x} > 270) = P\left(z > \frac{270-266}{6.532}\right) = P(z > 0.61) = 0.2709$

#Songs > 10(10) > 100 ✓

Exercises, page 454:

7.49 The mean is $\mu_{\bar{x}} = \mu_X = 225$ seconds, and the standard deviation is $\frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{10}} \approx 18.794$ secs

$\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{60}{\sqrt{10}} = 18.974$ seconds. These results do not depend on the shape of the distribution of the individual play times.

7.50 The mean is $\mu_{\bar{x}} = \mu_X = 40.125$ mm, and the standard deviation is $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{0.002}{\sqrt{4}} = 0.001$ mm. These results do not depend on the shape of the distribution of the individual axele diameters.

7.51 If we want $\sigma_{\bar{x}} = 30$, then we need to solve the following equation for n:

$30 = \frac{60}{\sqrt{n}} \rightarrow \sqrt{n} = \frac{60}{30} = 2 \rightarrow n = 4$. So we need a sample of size 4.

7.52 If we want $\sigma_{\bar{x}} = 0.0005$, then we need to solve the following equation for n:

$0.0005 = \frac{0.002}{\sqrt{n}} \rightarrow \sqrt{n} = \frac{0.002}{0.0005} = 4 \rightarrow n = 16$. So we need a sample of size 16.

7.53 (a) The sampling distribution of $\bar{x}$ is aprox Normal with $\mu_{\bar{x}} = \mu_X = 188$ mg/dl and $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{41}{\sqrt{100}} = 4.1$ mg/dl.

Next Page

(b) $P(185 \le \bar{x} \le 191) = P\left(\frac{185-188}{4.1} \le z \le \frac{191-188}{4.1}\right) = P(-0.73 \le z \le 0.73) = 0.7673 - 0.2327 = 0.5346$

(c) In this case $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{41}{\sqrt{1000}} = 1.30$ mg/dl. So now the probability becomes

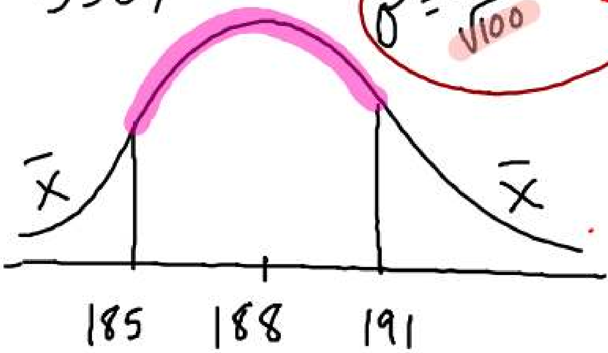
$P(185 \le \bar{x} \le 191) = P\left(\frac{185-188}{1.30} \le z \le \frac{191-188}{1.30}\right) = P(-2.31 \le z \le 2.31) = 0.9896 - 0.0104 = 0.9792$. The larger sample is better since it is more likely to produce a sample mean within 3 mg/dl of the population mean.

7.54 (a) The sampling distribution of $\bar{x}$ is Normal with $\mu_{\bar{x}} = \mu_X = 55{,}000$ miles and $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{4500}{\sqrt{8}} = 1591$ miles. (b) $P(\bar{x} < 51{,}800) = P\left(z < \frac{51{,}800 - 55{,}000}{1591}\right) = P(z < -2.01) = 0.0222$ Getting a sample mean this low would be a surprising result if the company's claim was true. Thus, I would doubt the company's claim.

7.55 (a) Let X denote the amount of cola in a bottle. $P(X < 295) = P\left(z < \frac{295-298}{3}\right) = P(z < -1) =$ 0.1587. (b) If $\bar{x}$ is the mean contents of six bottles (assumed to be independent), then $\bar{x}$ has a Normal distribution with $\mu_{\bar{x}} = \mu_X = 298$ ml and $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{3}{\sqrt{6}} = 1.2247$ ml (10% condition OK since there are more than 60 bottles in the population). $P(\bar{x} < 295) = P\left(z < \frac{295-298}{1.2247}\right) = P(z < -2.45) = 0.0071$.

53b)

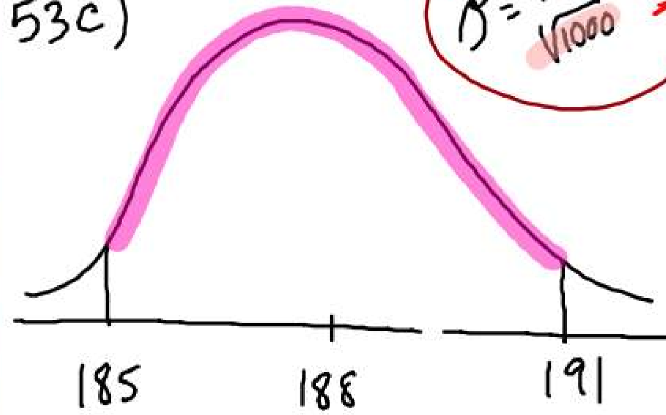
$$\sigma = \frac{41}{\sqrt{100}} = 4.1$$



normalcdf (185, 191, 188, 4.1) ≈ .5357

53c)

$$\sigma = \frac{41}{\sqrt{1000}} = 1.3$$

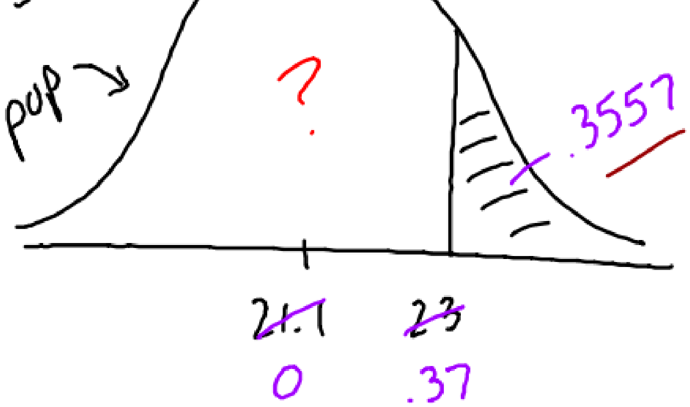


normalcdf (185, 191, 188, 1.3) ≈ .9790

Greater probability that a sample mean is within 3 points of population mean

56a)

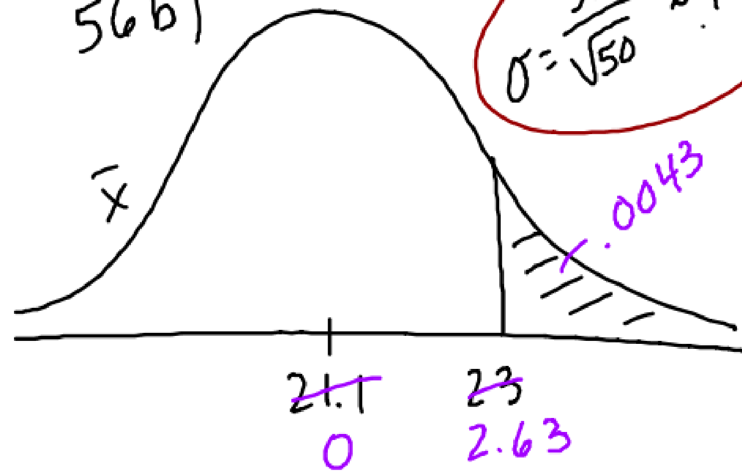
$\sigma = 5.1$



normalcdf(23, 100, 21.1, 5.1) ≈ .3547

56b)

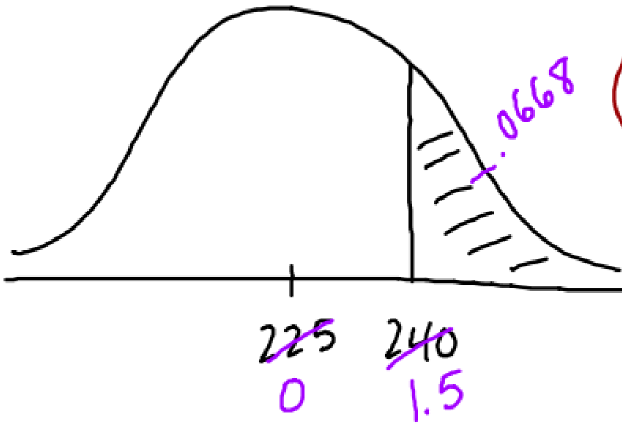
$\sigma = \frac{5.1}{\sqrt{50}} \approx .7212$



normalcdf(23, 102, 21.1, .7212) ≈ .0042

159 a) Samp dist not Normal

b)



$$\sigma = \frac{60}{\sqrt{36}} = 10 \text{ secs}$$

if $N > 360$ ✓

Normalcdf (240, 1001, 225, 10) ≈ .0668