5.15. Applet, answers will vary. **(a)** and **(b)** The mean should be "close" to 50.5. **(c)** The histogram theoretically should be a Normal distribution centered at 50.5.

5.16. (a) The larger sample size should have smaller variability. **(b)** Answers will vary; generally, the sample of size 10 will produce larger variability in the sampling distribution. **(c)** Theoretically, as the sample size increases, the variability of the sampling distribution decreases.

5.17. Population: all adults in the U.S. Statistic: mean of 37 hours and 6 min. Likely values: answers will vary, any amount of time someone could spend during a month using mobile apps (4.5, 68, 0, etc).

5.18. $\mu_{\bar{x}} = \mu = 44.$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{64}} = 2.$

5.19. $\mu_{\bar{x}} = \mu = 44.$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{576}} = 0.667.$ When the sample size increases, the mean of the sampling distribution remains the same, but the standard deviation of the sampling distribution decreases.

5.20. With $n = 64$, $\mu_{\bar{x}} = \mu = 82.$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{24}{\sqrt{64}} = 3.$ So, $\bar{x} \sim N(82, 3)$. About 95% of the time, $\bar{x}$ should be between 76 and 88 (two standard deviations either side of the mean).

5.21. With $n = 2304$, $\mu_{\bar{x}} = \mu = 82.$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{24}{\sqrt{2304}} = 0.5.$ So, $\bar{x} \sim N(82, 0.5)$. About 95% of the time, $\bar{x}$ should be between 81 and 83 (two standard deviations either side of the mean). With the larger sample size, the standard deviation decreased.

5.22. Applet, answers will vary. **(a)** Shown below. **(b)** 4.983 and 2.912, both are fairly close to the true values. **(c)** The shape is definitely not Normal, it is uniform.

5.23. Applet, answers will vary. **(a)** Each sample size has $\mu_{\bar{x}} = 1$. For n = 2, $\sigma_{\bar{x}} = 1/\sqrt{2} = 0.7071$. For n = 10, $\sigma_{\bar{x}} = 1/\sqrt{10} = 0.3162$. For n = 25, $\sigma_{\bar{x}} = 1/\sqrt{25} = 0.2$. **(b)** Shown below is one realization for n = 25. **(c)** The mean and standard deviations should be fairly close to the values calculated in part (a). As the sample size increases, the sampling distribution for the sample mean gets less right-skewed and closer to Normal. **(d)** Answers will vary; generally, n = 40 is "large enough" for most skewed distributions like an exponential random variable.

5.24. Using $\bar{x} \sim N(15, 2.12)$. $P(\bar{x} < 15) = P\left(Z < \frac{15-15}{2.12} = 0\right) = 0.5000$. This is close to the exact answer 0.5188.

5.25. **(a)** The standard deviation for n = 10 will be $\sigma_{\bar{x}} = 10/\sqrt{10}$. **(b)** Standard deviation *decreases* with increasing sample size. **(c)** $\mu_{\bar{x}}$ always equals μ and does not depend on the sample size n. **(d)** The population size N does not matter, as long as it is relatively large compared with n.

5.26. **(a)** The population mean μ is considered fixed. It does not have a distribution. The distribution of *sample* mean will be approximately Normal with large n. **(b)** The distribution of the observed values will "look like" the population distribution, which doesn't have to be Normal. The distribution of the sample mean becomes approximately Normal for large n. **(c)** For large n, $\bar{x}$ will be within $2\sigma/\sqrt{n}$ of μ about 95% of the time, not within 2σ of μ.

5.27. **(a)** μ = 120.5. **(b)** Answers will vary. **(c)** Answers will vary. **(d)** The center of the histogram should theoretically be close to μ, but random sampling does not guarantee this. With larger n, we would expect the mean of this sampling distribution to get closer to μ.

5.28. (a) $\sigma_{\bar{x}} = 1.24/\sqrt{120} = 0.113$. **(b)** 95% of the time, we'll expect the sample mean to be within $6.78 \pm 2(0.113)$ hours or between 6.554 and 7.006 hours. **(c)** $z = \frac{6.9 - 6.78}{0.113} = 1.06$. Using Table A, the probability is 0.8554.

5.29. (a) Larger. **(b)** We need $\sigma_{\bar{x}} \leq 0.08/2 = 0.04$. **(c)** We need $1.24/\sqrt{n} \leq 0.04$. Using algebra, we find $n \geq (1.24/0.04)^2 = 775$. So $n = 775$.

5.30. (a) The standard deviation will be $\sigma_{\bar{x}} = \frac{3.25}{\sqrt{25}} = 0.65$. **(b)** We need *n* large enough so that $\frac{3.25}{\sqrt{n}} \leq 0.50$. Using algebra to solve for *n*, we find $n \geq \left(\frac{3.25}{0.50}\right)^2 = 42.25$, so $n = 43$.

5.31. The mean of the sample means will still be 250 ml. The standard deviation of the sample means will be $\sigma_{\bar{x}} = 0.4/\sqrt{5} = 0.1789$ml.

5.32. (a) $n = 40$ is generally considered "large enough" for the sample mean to be approximately Normal even for skewed distributions. **(b)** Shown below. The standard deviation is $\sigma_{\bar{x}} = 3.25/\sqrt{50} = 0.4596$. **(c)** We need $P(\bar{x} < 2.20 \text{ or } \bar{x} > 2.50)$. By symmetry, this is $2P(\bar{x} < 2.20) = 2P\left(Z < \frac{2.20 - 2.35}{0.4596} = -0.33\right)$. Using Table A, the desired probability is 2(0.3707) = 0.7414.

5.33. (a) Graph shown. **(b)** To be more than 1 ml away from the target value means the volume is less than 249 or more than 251. Using symmetry, $P = 2P(X < 249) = 2P(Z < -2.5) = 2(0.0062) = 0.0124$. **(c)** $P = 2P(X < 249) = 2P(Z < -5.59)$, which is essentially 0.

5.34. (a) With n = 80 in the SRS, the sample mean will have $\mu_{\bar{x}} = 338$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{380}{\sqrt{80}} = 42.4853$. **(b)** $z = \frac{350-338}{42.4853} = 0.28$. $P(Z > 0.28) = 0.3897$. **(c)** The mean of total number of friends in the sample is 80(338) = 27,040, and the standard deviation of the total is $\sqrt{80*380^2} = 3398.823$. **(d)** $P(T > 28{,}000) = P\left(Z > \frac{28000-27040}{3398.823} = 0.28\right) = 0.3897$.

5.35. (a) $\bar{x}$ is not systematically higher than or lower than μ. **(b)** With large samples, $\bar{x}$ is more likely to be close to μ.

5.36. (a) $\mu = 4(0.18) + 3(0.31) + 2(0.26) + 1(0.13) + 0(0.12) = 2.3$ and $\sigma^2 = (4-2.3)^2(0.18) + (3-2.3)^2(0.31) + (2-2.3)^2(0.26) + (1-2.3)^2(0.13) + (0-2.3)^2(0.12) = 1.55$, $\sigma = \sqrt{1.55} = 1.245$. **(b)** $\mu_{\bar{x}} = 2.3$, $\sigma_{\bar{x}} = \sigma/\sqrt{25} = 0.249$. **(c)** $P(X \geq 3) = 0.31 + 0.18 = 0.49$. **(d)** $P(\bar{x} \geq 3) = P\left(Z \geq \frac{3-2.3}{0.249} = 2.81\right) = 0.0025$.

5.37. (a) $\mu_{\bar{x}} = 0.4$. $\sigma_{\bar{x}} = 0.9/\sqrt{100} = 0.09$. **(b)** $z = \frac{0.5-0.4}{0.09} = 1.11$. From Table A, $P(Z > 1.11) = 0.1335$. **(c)** Yes, n = 100 is a large enough sample to be able to use the central limit theorem.

5.38. (a) Although the probability of having to pay for a total loss for one or more of the 100 policies is very small, if this were to happen, it would be financially disastrous. On the other hand, for thousands of policies, the law of large numbers says that the average claim on many policies will be close to the mean, so the insurance company can be assured that the premiums it collects will (almost certainly) cover the claims. **(b)** The central limit theorem says that, in spite of the skewness of the population distribution, the average loss among 50,000 policies will be approximately Normally distributed with mean \$500 and standard deviation $\sigma/\sqrt{50{,}000}$ = \$44.72. $P(\bar{x} > 600) = P\left(Z > \frac{600-500}{44.72}\right) = P(Z > 2.24) = 0.0125$.

5.39. If W is total weight, and $\bar{x} = W/25$, then $P(W > 5200) = P(\bar{x} > 208) = P\left(Z > \frac{208-190}{35/\sqrt{25}}\right) = P(Z > 2.57) = 0.0051$.

5.40. (a) The distribution of the mean for $n = 24$ rowers with IDNA is $N(58,\ 11/\sqrt{24} = 2.245)$. $P(\bar{x} > 63) = P\left(Z > \frac{63-58}{2.245}\right) = P(Z > 2.23) = 0.0129$. **(b)** The distribution of the mean for $n = 24$ rowers with Normal iron status is $N(69,\ 18/\sqrt{24} = 3.674)$. $P(\bar{x} < 63) = P\left(Z < \frac{63-69}{3.674}\right) =$ $P(Z < -1.63) = 0.0516$ (software gives 0.0512). **(c)** The distribution of the mean difference (rowers with IDNA – rowers with Normal iron status) is $N(-11,\ \sqrt{2.245^2 + 3.674^2} = 4.306)$. $P(\bar{x}_{IDNA} - \bar{x}_{Normal} > 0) =$ $P\left(Z > \frac{0-(-11)}{4.306}\right) = P(Z > 2.55) = 0.0054$.

5.41. (a) Assuming both samples are "large," $\bar{y}$ has a $N(\mu_Y, \sigma_Y/\sqrt{m})$ distribution, and $\bar{x}$ has a $N(\mu_X, \sigma_X/\sqrt{n})$ distribution. **(b)** $\bar{y} - \bar{x}$ has a Normal distribution with mean $\mu_Y - \mu_X$ and standard deviation $\sqrt{\sigma_{\bar{y}}^2 + \sigma_{\bar{x}}^2}$.

5.57. (a) Separate flips are independent (coins have no "memory," so they do not get on a "streak" of heads). **(b)** The coin is fair. The probabilities are still $P(H) = P(T) = 0.5$. **(c)** The parameters for a binomial distribution are n and p. $\hat{p}$ is a sample statistic. **(d)** This is best modeled with a Poisson distribution.

5.58. (a) X is a count. $\hat{p}$ is a proportion. **(b)** The given formula is the standard deviation of the sample proportion. The variance for a binomial count is $np(1 - p)$. **(c)** Accuracy depends not only on n but also on p. If p is small, using the approximation may not be valid. **(d)** The population size should be at least 20 times as large as the sample size.

5.59. (a) A $B(200, p)$ distribution seems reasonable for this setting (even though we do not know what p is). **(b)** This setting is not binomial; there is no fixed value of n. **(c)** A $B(500, 1/12)$ distribution seems appropriate for this setting. **(d)** This is not binomial because separate cards are not independent.

5.60. (a) This is not binomial; X is not a count of successes. **(b)** A $B(20, p)$ distribution seems reasonable, where p (unknown) is the probability of a defective pair. **(c)** This should be (at least approximately) the $B(n, p)$ distribution, where n is the number of students in our sample, and p is the probability that a randomly chosen student eats at least five servings of fruits and vegetables per day. **(d)** This is not binomial. There are more than two possible values for the number of days that you skip a class during a school year.

5.61. (a) The distribution of those who say they have stolen something is $B(10, 0.2)$. The distribution of those who do not say they have stolen something is $B(10, 0.8)$. **(b)** X is the number who say they have stolen something. $P(X \geq 4) = 1 - P(X \leq 3) = 0.1209$.

5.62. (a) $B(50, 0.33)$. **(b)** Using software, $P(X \geq 18) = 1 - P(X \leq 17) = 0.6240$.

5.63. (a) $\mu = 10(0.2) = 2$ will say they have stolen. $\mu = 10(0.8) = 8$ will say they have not stolen. **(b)** $\sigma = \sqrt{10(0.2)(1-0.2)} = 1.265$. **(c)** If $p = 0.1$, $\sigma = \sqrt{10(0.1)(1-0.1)} = 0.949$. If $p = 0.01$, $\sigma = \sqrt{10(0.01)(1-0.01)} = 0.315$. As p gets smaller, the standard deviation becomes smaller.

5.64. (a) $\mu_X = 15(0.75) = 11.25$. The mean of $\hat{p}$ is $p = 0.75$. **(b)** For $n = 150$, $\mu_X = 150(0.75) = 112.5$. For $n = 1500$ $\mu_X = 1500(0.75) = 1125$. The mean of $\hat{p}$ is always $p = 0.75$. As the sample size increases, the expected count of successes increases, and the mean proportion of successes stays the same.

5.65. (a) $P(X \geq 12) = 0.0833$ and $P(X \geq 13) = 0.0206$, so 13 is the smallest value of m. **(b)** Using 75%, $P(X \leq 13) = 0.9198$. **(c)** The probability will decrease. When the sample size increases, the mean will increase as well.

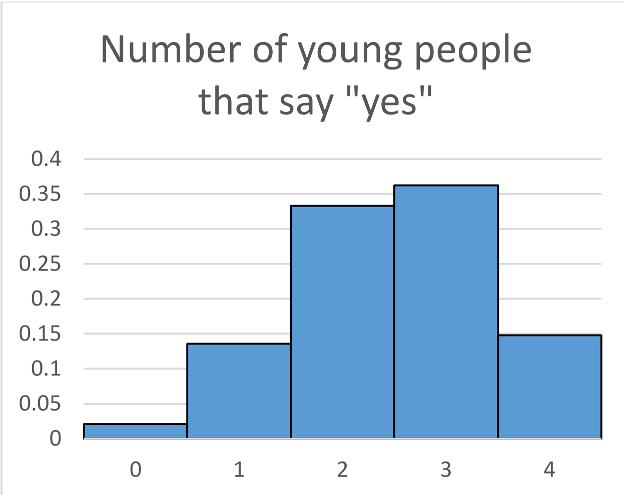
5.66. (a) The population (the 75 members of the fraternity) is only 2.5 times the size of the sample. Our rule of thumb says that this ratio should be at least 20. **(b)** Our rule of thumb for the Normal approximation calls for both np and $n(1 - p)$ to be at least 10; we have $np = (1000)(0.002) = 2$

5.67. The probability that a digit is greater than 5 is $P(6 \text{ or } 7 \text{ or } 8 \text{ or } 9) = 0.1 + 0.1 + 0.1 + 0.1 = 0.4$ and 0.6 that the digit is not greater than 5. **(a)** P(at least one digit greater than 5) = 1 – (no digits greater than 5) = $1 - (0.6)^6 = 0.9533$. **(b)** $\mu = (40)(0.4) = 16$.

5.68. Answers will vary. **(a)** Students should repeat until they have accumulated 25 samples of 15 tosses and use those results to answer the questions. **(b)** $B(25, 0.2863)$. **(c)** If we repeated part **(a)** many times, $\hat{p}$ is distributed $N(0.2863, 0.0904)$. So, on average, the proportion of samples with 0 bad records will be 0.2863 but with a 0.0904 standard deviation. Thus, each group of 25 could have a proportion somewhat different than the 0.2863 that we expect.

5.69. (a) $n = 4$, $p = 0.62$. **(b)** See table and graph below. **(c)** $\mu = 4(0.62) = 2.48$.

x	0	1	2	3	4
$P(x)$	0.0209	0.1361	0.3330	0.3623	0.1478



5.70. With n = 1020 and p = 0.48, both np and $n(1-p)$ are close to 500 (further exact calculations are unnecessary), so use of the Normal approximation is appropriate. We have $\mu_{\hat{p}} = 0.48,$ and

$$\sigma_{\hat{p}} = \sqrt{\frac{(0.48)(1-0.48)}{1020}} = 0.0156. \; P(0.44 < \hat{p} < 0.52) = P\left(\frac{0.44-0.48}{0.0156} < z < \frac{0.44-0.52}{0.0156}\right) =$$

$P(-2.56 < z < 2.56) = 0.9948 - 0.0052 = 0.9896$. There is almost a 99% probability of obtaining a sample proportion between 44% and 52%, but it is not guaranteed.

5.71. (a) Because 500(0.62) = 310 and 500(0.38) = 190, the approximate distribution is $p\ hat \sim N(0.62, \surd\{(0.62)(0.38) \div 500\} = 0.0217)$. $P(0.59 < \hat{p} < 0.65) =$

$P\left(\frac{0.59-0.62}{0.0217} < z < \frac{0.65-0.62}{0.0217}\right) = P(-1.38 < z < 1.38) = 0.9162 - 0.0838 = 0.8324$. **(b)** If p = 0.9, the distribution of $\hat{p}$ is approximately N(0.9, 0.0134) because 500(0.9) = 450 and 500(0.1) = 50.

$P(0.87 < \hat{p} < 0.93) = P\left(\frac{0.87-0.9}{0.0134} < z < \frac{0.93-0.9}{0.0134}\right) = P(-2.24 < z < 2.24) = 0.9875 - 0.0125 = 0.9750$.

(c) As p gets closer to 1, the probability of being within ±0.03. of p increases because the standard deviation decreases.

5.72. The mean of the sampling distribution is still p = 0.48 regardless of sample size. If n = 300, the standard deviation of the sample proportion becomes $\sigma_{\hat{p}} = \sqrt{\frac{0.48(1-0.48)}{300}} = 0.0288;$ when n = 5000, we would have $\sigma_{\hat{p}} = 0.0071$. When n = 300, $P(0.44 < \hat{p} < 0.52) = P(-1.39 < Z < 1.39) = 0.9177 - 0.0823 = 0.8354$. When n = 5000, $P(0.49 < \hat{p} < 0.57)$ is essentially 1. We see clearly that larger samples are more likely to have results that are close to the truth for the population.

5.73. (a) The mean is $\mu = p = 0.69$, and the standard deviation is $\sigma = \sqrt{p(1-p)/n} = 0.0008444$. **(b)** $\mu \pm 2\sigma$ gives the range 68.83% to 69.17%. **(c)** This range is considerably narrower than the historical range. In fact, 67% and 70% correspond to $z = -23.7$ and $z = 11.8$, suggesting that the observed percentages do not come from a N(0.69, 0.0008444) distribution; that is, the population proportion has changed over time.

5.74. (a) The sample proportion who preferred blended instruction was $\hat{p} = \frac{311}{400} = 0.7775$. (b) If p = 0.85, $\sigma_{\hat{p}} = \sqrt{\frac{0.85(1-0.85)}{400}} = 0.0178$. **(c)** About 95% of the time, we would expect $\hat{p}$ to be within the interval $p \pm 2\sigma_{\hat{p}}$, or 0.8144 to 0.8856. **(d)** The sample from our college found a proportion that was not in this interval; we should believe that our students prefer blended instruction less than students nationally.

5.75. (a) $\hat{p} = 56/200 = 0.28.$ (b) We need $P(\hat{p} \geq 0.28)$. $\hat{p}$ is approximately $N(0.28, 0.0317)$ because 200(0.28) = 56 and 200(0.72) = 144. $P(\hat{p} \geq 0.28) = P\left(z \geq \frac{0.28-0.28}{0.0317}\right) = P(Z \geq 0) = 0.5$. **(c)** Answers may vary. The key is that there is a 50% chance of obtaining a sample proportion 0.28 or larger if the proportion of binge drinkers at your campus is really 0.28, so it is likely that percent of binge drinkers on your campus is similar to the national average.

5.76. We need $\sigma = \sqrt{\frac{0.48(1-0.48)}{n}} = 0.005.$ This means $n = \left(\frac{1}{0.005}\right)^2 (0.48)(1-0.48) = 9984.$

5.77. (a) $p = 1/4 = 0.25$. **(b)** $P(X \geq 10) = 0.0139$. **(c)** $\mu = np = 5$ and $\sigma = \sqrt{np(1-p)} = \sqrt{3.75} = 1.9365$. **(d)** No. The trials would not be independent because the subject may alter his/her guessing strategy based on this information.

5.78. (a) $\mu = 1200(0.83) = 996$ and $\sigma = \sqrt{np(1-p)} = 13.0123$ students. **(b)** $P(X \geq 800) = P(Z \geq -15.06) = 1$ (essentially). **(c)** $P(X \geq 1000) = P(Z \geq 0.31) = 0.3783$. **(d)** With $n = 1150$, $\mu = 1150(0.83) = 954.5$ and $\sigma = \sqrt{1150(0.83)(0.17)} = 12.7383$. $P(X \geq 1000) = P(Z \geq 3.57) < 0.0002$.

5.79. (a) X has a $B(900, 1/5)$ distribution, with mean $\mu = np = 900(1/5) = 180$ and $\sigma = \sqrt{(900)(0.2)(0.8)} = 12$ successes. **(b)** For $\hat{p}$, the mean is $\mu_{\hat{p}} = p = 0.2$ and $\sigma_{\hat{p}} = \sqrt{(0.2)(0.8)/900} = 0.01333$. **(c)** $P(\hat{p} > 0.24) = P(Z > 3) = 0.0013$. **(d)** From a standard Normal distribution, $P(Z > 2.326) = 0.01$, so the subject must score 2.326 standard deviations above the mean: $\mu_{\hat{p}} + 2.326\sigma_{\hat{p}}$ = 0.20. + 2.326(0.01333) = 0.2310. This corresponds to 208 or more successes (correct guesses) in 900 attempts.

5.80. (a) $\binom{n}{n} = \frac{n!}{n!(n-n)!} = \frac{n!}{n!(0)!} = 1$. The number of ways to distribute n successes among n trials is 1. **(b)** $\binom{n}{n-1} = \frac{n!}{(n-1)!(n-(n-1))!} = \frac{n!}{(n-1)!(1)!} = \frac{n\cdot(n-1)!}{(n-1)!} = n$. The number of ways to distribute $n-1$ successes (and one failure) among n trials is n. **(c)** $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!(n)!} = \binom{n}{n-k}$. The number of ways to distribute k successes (and $n-k$ failures) among n trials is the same as distributing $n-k$ successes and k failures.

5.81. Poisson with $\mu = 2.768$. **(a)** $P(X = 0) = 0.0628$. **(b)** $P(X \geq 3) = 1 - P(X \leq 2) = 0.5229$.