Chapter 5 Solutions

5.1. **(a)** $n = 1500$ (the sample size). **(b)** The "Yes" count seems like the most reasonable choice, but either count is defensible. **(c)** $X = 525$ (or $X = 975$). **(d)** $\hat{p} = \frac{525}{1500} = 0.35$ (or $\hat{p} = \frac{975}{1500} = 0.65$).

5.2. $n = 200$ (the sample size), $\hat{p} = 40\% = 0.40$, and $X = n\hat{p} = 80$.

5.3. We have 20 independent trials, each with probability of success (heads) equal to 0.5, so X has a $B(20, 0.5)$ distribution.

5.4. Assuming no multiple births (twins, triplets, quadruplets), we have four independent trials, each with probability of success (type O blood) equal to 0.25, so the number of children with type O blood has a $B(4, 0.25)$ distribution.

5.5. **(a)** For a $B(4, 0.3)$ distribution, $P(X = 0) = 0.2401$ and $P(X \geq 3) = 0.0837$. **(b)** For a $B(4, 0.7)$ distribution, $P(X = 4) = 0.2401$ and $P(X \leq 1) = 0.0837$. **(c)** The number of "failures" in a $B(4, 0.3)$ distribution has a $B(4, 0.7)$ distribution. With 4 trials, 0 successes is equivalent to 4 failures, and 3 or more successes is equivalent to 1 or fewer failures.

5.6. **(a)** For a $B(100, 0.5)$ distribution, $\mu_{\hat{p}} = p = 0.5$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \frac{1}{20} = 0.05$. **(b)** No; the mean and standard deviation of the sample count are both 100 times bigger. (That is, $\hat{p} = X/100$, so $\mu_{\hat{p}} = \mu_X/100$ and $\sigma_{\hat{p}} = \sigma_X/100$.)

5.7. **(a)** $\hat{p}$ has approximately a Normal distribution with mean 0.5 and standard deviation 0.025, so $P(0.4 < \hat{p} < 0.6) = P(-2 < Z < 2) \doteq 0.9544$.
(b) $P(0.45 < \hat{p} < 0.55) = P(-1 < Z < 1) \doteq 0.6826$.

Note: *The answers given are from Table A or software. The 68–95–99.7 rule would give the answers 0.95 and 0.68—quite reasonable answers, especially since this is an approximation anyway. For comparison, the exact answers are* $P(0.4 < \hat{p} < 0.6) \doteq 0.9431$ *or* $P(0.4 \leq \hat{p} \leq 0.6) \doteq 0.9648$, *and* $P(0.45 < \hat{p} < 0.55) \doteq 0.6318$ *or* $P(0.45 \leq \hat{p} \leq 0.55) \doteq 0.7287$. *(Notice that the "correct" answer depends on our understanding of "between.")*

5.8. **(a)** $P(X \geq 3) = \binom{4}{3}0.52^3 0.48 + \binom{4}{4}0.52^4 \doteq 0.3431$. **(b)** If the coin were fair, $P(X \geq 3) = \binom{4}{3}0.5^3 0.5 + \binom{4}{4}0.5^4 = 0.3125$.

5.9. **(a)** Separate flips are independent (coins have no "memory," so they do not try to compensate for a lack of tails). **(b)** Separate flips are independent (coins have no "memory," so they do not get on a "streak" of heads). **(c)** $\hat{p}$ can vary from one set of observed data to another; it is not a parameter.

5.10. (a) X is a *count;* $\hat{p}$ is a proportion. **(b)** The given formula is the *standard deviation* for a binomial *proportion*. The variance for a binomial count is $np(1-p)$. **(c)** The rule of thumb in the text is that np and $n(1-p)$ should both be at least 10. If p is close to 0 (or close to 1), $n = 1000$ might not satisfy this rule of thumb. (See also the solution to Exercise 5.18.)

5.11. (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.

5.12. (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) a $B(n, p)$ 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.

5.13. (a) C, the number caught, is $B(10, 0.7)$. M, the number missed, is $B(10, 0.3)$. **(b)** Referring to Table C, we find $P(M \geq 4) = 0.2001 + 0.1029 + 0.0368 + 0.0090 + 0.0014 + 0.0001 = 0.3503$ (software: 0.3504).

5.14. (a) X, the number of auction site visitors, is $B(15, 0.5)$. **(b)** Symmetry tells us that this must be 0.5, which is confirmed in Table C: $P(X \geq 8) = 0.1964 + 0.1527 + 0.0916 + 0.0417 + 0.0139 + 0.0032 + 0.0005 = 0.5$.

5.15. (a) The mean of C is $(10)(0.7) = 7$ errors caught, and for M the mean is $(10)(0.3) = 3$ errors missed. **(b)** The standard deviation of C (or M) is $\sigma = \sqrt{(10)(0.7)(0.3)} \doteq 1.4491$ errors. **(c)** With $p = 0.9$, $\sigma = \sqrt{(10)(0.9)(0.1)} \doteq 0.9487$ errors; with $p = 0.99$, $\sigma \doteq 0.3146$ errors. σ decreases toward 0 as p approaches 1.

5.16. (a) The mean of X is $(15)(0.5) = 7.5$; the mean of $\hat{p}$ is 0.5. **(b)** The mean of X increases with n; it is 75 with $n = 150$, and 750 with $n = 1500$. The mean of $\hat{p}$ is 0.5 for any value of n.

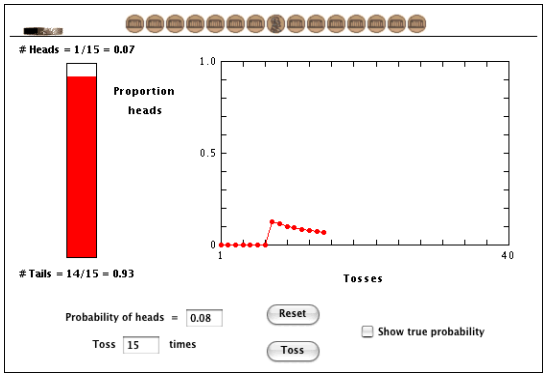
5.17. $m = 6$: $P(X \geq 6) = 0.0473$ and $P(X \geq 5) = 0.1503$.

5.18. (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 np and $n(1-p)$ to be at least 10; we have $np = (1000)(0.002) = 2$.

5.19. The count of 5s among n random digits has a binomial distribution with $p = 0.1$. **(a)** $P(\text{at least one 5}) = 1 - P(\text{no 5}) = 1 - (0.9)^5 \doteq 0.4095$. (Or take 0.5905 from Table C and subtract from 1.) **(b)** $\mu = (40)(0.1) = 4$.

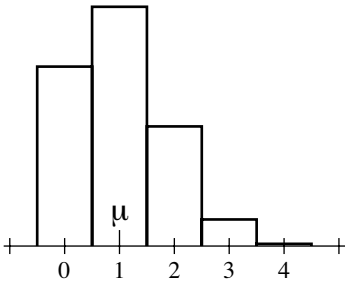
5.20. One sample of 15 flips is shown on the right. Results will vary quite a bit; Table C shows that 99.5% of the time, there will be 4 or fewer bad records in a sample of 15.

Out of 25 samples, most students should see 2 to 12 samples with no bad records. That is, N, the number of samples with no bad records, has a $B(25, 0.2683)$ distribution, and $P(2 \le N \le 12) = 0.9894$.



5.21. **(a)** $n = 4$ and $p = 1/4 = 0.25$. **(b)** The distribution is below; the histogram is on the right. **(c)** $\mu = np = 1$.

x	0	1	2	3	4
$P(X = x)$	.3164	.4219	.2109	.0469	.0039



5.22. For $\hat{p}$, $\mu = 0.49$ and $\sigma = \sqrt{p(1-p)/n} \doteq 0.01568$. As $\hat{p}$ is approximately Normally distributed with this mean and standard deviation, we find:

$$P(0.46 < \hat{p} < 0.52) \doteq P(-1.91 < Z < 1.91) \doteq 0.9438$$

(Software computation of the Normal probability gives 0.9442. Using a binomial distribution, we can also find $P(468 \le X \le 428) \doteq 0.9444$.)

5.23. Recall that $\hat{p}$ is approximately Normally distributed with mean $\mu = p$ and standard deviation $\sqrt{p(1-p)/n}$. **(a)** With $p = 0.30$, $\sigma \doteq 0.01441$, so $P(0.28 < \hat{p} < 0.32) = P(-1.39 < Z < 1.39) \doteq 0.8354$. (Software computation of the Normal probability gives 0.8348. Using a binomial distribution, we can also find $P(284 \le X \le 323) \doteq 0.8301$.) **(b)** With $p = 0.06$, $\sigma \doteq 0.00747$, so $P(0.04 < \hat{p} < 0.08) = P(-2.68 < Z < 2.68) = 0.9926$. (Using a binomial distribution, we can also find $P(41 \le X \le 80) \doteq 0.9918$.) **(c)** $P(-0.02 < \hat{p} - p < 0.02)$ increases to 1 as p gets closer to 0. (This is because σ also gets close to 0, so that $0.02/\sigma$ grows.)

5.24. When $n = 300$, the distribution of $\hat{p}$ is approximately Normal with mean 0.49 and standard deviation 0.02886 (nearly twice that in Exercise 5.22). When $n = 5000$, the standard deviation drops to 0.00707 (less than half as big as in Exercise 5.22). Therefore:

$$n = 300: \quad P(0.46 < \hat{p} < 0.52) \doteq P(-1.04 < Z < 1.04) \doteq 0.7016$$
$$n = 5000: \quad P(0.46 < \hat{p} < 0.52) \doteq P(-4.24 < Z < 4.24) \doteq 1$$

Larger samples give a better probability that $\hat{p}$ will be close to the true proportion p. (Software computation of the first Normal probability gives 0.7014; using a binomial distribution, we can also find $P(138 \le X \le 156) \doteq 0.7275$. These more accurate answers do not change our conclusion.)

5.25. **(a)** $\hat{p} = \frac{140}{200} = 0.7$. **(b)** We want $P(X \geq 140)$ or $P(\hat{p} \geq 0.7)$. The first can be found exactly (using a binomial distribution), or we can compute either using a Normal approximation (with or without the continuity correction). All possible answers are shown on the right. **(c)** The sample results are higher 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.

			Continuity correction	
Exact prob.	Table Normal	Software Normal	Table Normal	Software Normal
0.2049	0.1841	0.1835	0.2033	0.2041

5.26. As $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$, we have $0.004^2 = (0.49)(0.51)/n$, so $n = 15{,}618.75$—round up to 15,619.

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

5.28. **(a)** $\mu = (1200)(0.75) = 900$ and $\sigma = \sqrt{225} = 15$ students. **(b)** $P(X \geq 951) \doteq P(Z \geq 3.4) = 0.0003$. **(c)** With $n = 1300$, $P(X \geq 951) \doteq P(Z \geq -1.54) = 0.9382$. Other answers are shown in the table on the right.

		Continuity correction	
Table Normal	Software Normal	Table Normal	Software Normal
0.9382	0.9379	0.9418	0.9417

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

5.30. **(a)** M has a binomial distribution with $n = 30$ and $p = 0.6$, so $P(M = 20) = \binom{30}{20}(0.6)^{20}(0.4)^{10} \doteq 0.1152$. **(b)** $P(\text{1st woman is the 4th call}) = (0.6)^3(0.4) = 0.0864$.

5.31. **(a)** $p = \frac{23{,}772{,}494}{209{,}128{,}094} \doteq 0.1137$. **(b)** If B is the number of blacks, then B has (approximately) a binomial distribution with parameters $n = 1200$ and $p \doteq 0.1137$, so the mean is $np \doteq 136.4$ blacks. **(c)** $P(B \leq 100) \doteq P(Z < -3.31) = 0.0005$.

Note: *In (b), the population is at least 20 times as large as the sample, so our "rule of thumb" for using a binomial distribution is satisfied. In fact, the mean would be the same even if we could not use a binomial distribution, but we need to have a binomial distribution for part (c), so that we can approximate it with a Normal distribution—which we can safely do, because both np and n(1 − p) are much greater than 10.*

5.32. **(a)** $\binom{n}{n} = \frac{n!}{n!\,0!} = 1$. The only way to distribute n successes among n observations is for all observations to be successes. **(b)** $\binom{n}{n-1} = \frac{n!}{(n-1)!\,1!} = \frac{n \cdot (n-1)!}{(n-1)!} = n$. To distribute $n-1$ successes among n observations, the one failure must be either observation 1, 2, 3, $\ldots$, $n-1$, or n. **(c)** $\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n!}{(n-k)!\,[n-(n-k)]!} = \binom{n}{n-k}$. Distributing k successes is equivalent to distributing $n-k$ failures.

5.33. **(a)** $P(\hat{p} \le 0.8) = P(X \le 80)$ is on line 1. **(b)** $P(\hat{p} \le 0.8) = P(X \le 200)$ is on line 2. **(c)** For a test with 400 questions, the standard deviation of $\hat{p}$ would be half as big as the standard deviation of $\hat{p}$ for a test with 100 questions: With $n = 100$, $\sigma = \sqrt{(0.85)(0.15)/100} \doteq 0.03571$; and with $n = 400$, $\sigma = \sqrt{(0.85)(0.15)/400} \doteq 0.01785$. **(d)** Yes: Regardless of p, n must be quadrupled to cut the standard deviation in half.

		Continuity correction	
Table Normal	Software Normal	Table Normal	Software Normal
0.0808	0.0807	0.1038	0.1038
0.0136	0.0134	0.0166	0.0168

5.34. **(a)** $P(\text{first } \boxdot \text{ appears on toss 2}) = \left(\frac{5}{6}\right)\left(\frac{1}{6}\right) = \frac{5}{36}$.

(b) $P(\text{first } \boxdot \text{ appears on toss 3}) = \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) = \frac{25}{216}$.

(c) $P(\text{first } \boxdot \text{ appears on toss 4}) = \left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)$.

$P(\text{first } \boxdot \text{ appears on toss 5}) = \left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)$.

5.35. Y has possible values $1, 2, 3, \ldots$. $P(\text{first } \boxdot \text{ appears on toss } k) = \left(\frac{5}{6}\right)^{k-1}\left(\frac{1}{6}\right)$.

5.36. With $\mu = 200$, $\sigma = 10$, and $n = 25$, we have mean $\mu_{\bar{x}} = \mu = 200$ and standard deviation $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 2$.

5.37. When $n = 100$, the mean is $\mu_{\bar{x}} = \mu = 200$ (unchanged), and the standard deviation is $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 1$. Increasing n does not change $\mu_{\bar{x}}$ but decreases $\sigma_{\bar{x}}$ (the variability of the sampling distribution).

5.38. When $n = 100$, $\sigma_{\bar{x}} = \sigma/\sqrt{100} = 10/10 = 1$. The sampling distribution of $\bar{x}$ is approximately $N(200, 1)$, so about 95% of the time, $\bar{x}$ is between 198 and 202.

5.39. When $n = 400$, $\sigma_{\bar{x}} = \sigma/\sqrt{400} = 10/20 = 0.5$. The sampling distribution of $\bar{x}$ is approximately $N(200, 0.5)$, so about 95% of the time, $\bar{x}$ is between 199 and 201.

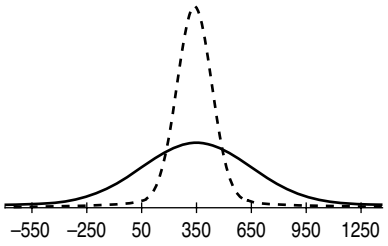
5.40. With $\sigma/\sqrt{70} \doteq 0.1195$, we have $P(\bar{x} < 1.1) \doteq P(Z < 0.84) \doteq 0.7995$. (If we use the rounded value $\sigma/\sqrt{70} \doteq 0.12$ given in the text, then $P(\bar{x} < 1.1) \doteq P(Z < 0.83) \doteq 0.7967$.)

5.41. **(a)** The given formula is the *standard deviation* of the mean. The variance of the sample mean is $10^2/20$. **(b)** Standard deviation decreases with increasing sample size. **(c)** $\mu_{\bar{x}}$ always equals μ, regardless of the sample size.

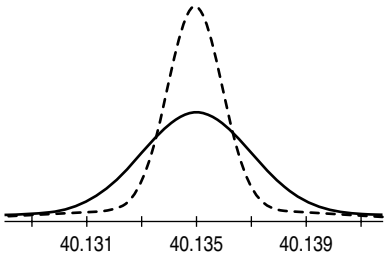
5.42. **(a)** The standard deviation is $\sigma/\sqrt{10} \doteq 94.8683$ seconds. **(b)** In order to have $\sigma/\sqrt{n} = 30$ seconds, we need a sample of size $n = 100$.

5.43. Mean $\mu = 40.135$ mm and standard deviation $\sigma/\sqrt{4} = 0.0015$ mm.

5.44. **(a)** For this exercise, bear in mind that the actual distribution for a single song length is definitely *not* Normal; in particular, a Normal distribution with mean 350 seconds and standard deviation 300 seconds extends well below 0 seconds. The Normal curve for $\bar{x}$ should be taller by a factor of $\sqrt{10}$ and skinnier by a factor of $1/\sqrt{10}$. **(b)** Using a $N(350, 300)$ distribution, $1 - P(331 < X < 369) \doteq 1 - P(-0.06 < Z < 0.06) \doteq 0.9522$. **(c)** Using a $N(350, 94.8683)$ distribution, $1 - P(331 < X < 369) \doteq 1 - P(-0.20 < Z < 0.20) \doteq 0.8415$.



5.45. In Exercise 5.43, we found that $\sigma_{\bar{x}} = 0.0015$ mm, so $\bar{x}$ has a $N(40.135 \text{ mm}, 0.0015 \text{ mm})$ distribution. **(a)** On the right. The Normal curve for $\bar{x}$ should be taller by a factor of 2 and skinnier by a factor of 0.5 (twice as tall, half as wide). **(b)** The probability that an individual axle diameter differs from the target by at least 0.006 mm—two standard deviations—is about 5% (by the 68–95–99.7 rule), or 0.0456 (using Table A). **(c)** 0.006 mm is four standard deviations for the distribution of $\bar{x}$, so this probability is $P(Z < -4 \text{ or } Z > 4) < 0.0001$.



5.46. **(a)** If T is the total number of lightning strikes (in one year), then $\mu_T = 10 \times 6 = 60$ strikes and $\sigma_T = \sqrt{10} \times 2.4 \doteq 7.5895$ strikes. **(b)** If $\bar{x}$ is the mean number of strikes per square kilometer (that is, $\bar{x} = T/10$), then $\mu_{\bar{x}} = 6$ strikes/km^2 and $\sigma_{\bar{x}} = 2.4/\sqrt{10} \doteq 0.7589$ strikes/km^2.

5.47. **(a)** $\bar{x}$ is not systematically higher than or lower than μ; that is, it has no particular tendency to underestimate or overestimate μ. **(b)** With large samples, $\bar{x}$ is more likely to be close to μ because with a larger sample comes more information (and therefore less uncertainty).

5.48. **(a)** $P(X \geq 23) \doteq P(Z \geq \frac{23 - 20.8}{4.8}) = P(Z \geq 0.46) = 0.3428$ (with software: 0.3234). Because ACT scores are reported as whole numbers, we might instead compute $P(X \geq 22.5) \doteq P(Z \geq 0.35) = 0.3632$ (software: 0.3616). **(b)** $\mu_{\bar{x}} = 20.8$ and $\sigma_{\bar{x}} = \sigma/\sqrt{25} = 0.96$. **(c)** $P(\bar{x} \geq 23) \doteq P(Z \geq \frac{23 - 20.8}{0.96}) = P(Z \geq 2.29) = 0.0110$. (In this case, it is not appropriate to find $P(\bar{x} \geq 22.5)$, unless $\bar{x}$ is rounded to the nearest whole number.) **(d)** Because individual scores are only roughly Normal, the answer to (a) is approximate. The answer to (c) is also approximate but should be more accurate because $\bar{x}$ should have a distribution that is closer to Normal.

5.49. **(a)** $\mu_{\bar{x}} = 0.5$ and $\sigma_{\bar{x}} = \sigma/\sqrt{50} = 0.7/\sqrt{50} \doteq 0.09899$. **(b)** Because this distribution is only approximately Normal, it would be quite reasonable to use the 68–95–99.7 rule to give a rough estimate: 0.6 is about one standard deviation above the mean, so the probability should be about 0.16 (half of the 32% that falls outside ± 1 standard deviation). Alternatively, $P(\bar{x} > 0.6) \doteq P(Z > \frac{0.6-0.5}{0.09899}) = P(Z > 1.01) = 0.1562$.

5.50. **(a)** $\mu = (4)(0.31) + (3)(0.40) + (2)(0.20) + (1)(0.04) + (0)(0.05) = 2.88$ and $\sigma = \sqrt{1.1056} \doteq 1.0515$. **(b)** $\mu_{\bar{x}} = \mu = 2.88$ and $\sigma_{\bar{x}} = \sigma/\sqrt{50} \doteq 0.1487$. **(c)** $P(X \geq 3) = 0.31 + 0.40 = 0.71$, and $P(\bar{x} \geq 3) \doteq P(Z \geq \frac{3-2.88}{0.1487} = P(Z \geq 0.81) = 0.2090$ (software value: 0.2098).

5.51. Let X be Sheila's measured glucose level. **(a)** $P(X > 140) = P(Z > 1.5) = 0.0668$. **(b)** If $\bar{x}$ is the mean of three measurements (assumed to be independent), then $\bar{x}$ has a $N(125, 10/\sqrt{3})$ or $N(125 \text{ mg/dl}, 5.7735 \text{ mg/dl})$ distribution, and $P(\bar{x} > 140) = P(Z > 2.60) = 0.0047$.

5.52. **(a)** $\mu_X = (\$500)(0.001) = \$0.50$ and $\sigma_X = \sqrt{249.75} \doteq \$15.8035$. **(b)** In the long run, Joe makes about 50 cents for each \$1 ticket. **(c)** If $\bar{x}$ is Joe's average payoff over a year, then $\mu_{\bar{x}} = \mu = \$0.50$ and $\sigma_{\bar{x}} = \sigma_X/\sqrt{104} \doteq \$1.5497$. The central limit theorem says that $\bar{x}$ is approximately Normally distributed (with this mean and standard deviation). **(d)** Using this Normal approximation, $P(\bar{x} > \$1) \doteq P(Z > 0.32) = 0.3745$ (software: 0.3735).

Note: *Joe comes out ahead if he wins at least once during the year. This probability is easily computed as* $1 - (0.999)^{104} \doteq 0.0988$. *The distribution of* $\bar{x}$ *is different enough from a Normal distribution so that answers given by the approximation are not as accurate in this case as they are in many others.*

5.53. The mean of three measurements has a $N(125 \text{ mg/dl}, 5.7735 \text{ mg/dl})$ distribution, and $P(Z > 1.645) = 0.05$ if Z is $N(0, 1)$, so $L = 125 + 1.645 \cdot 5.7735 \doteq 134.5$ mg/dl.

5.54. $\bar{x}$ is approximately Normal with mean 1.5 and standard deviation $1.3/\sqrt{200} \doteq 0.09192$ flaws/yd^2, so $P(\bar{x} > 2) \doteq P(Z > 5.44) = 0$ (essentially).

5.55. If W is total weight, and $\bar{x} = W/25$, then:

$$P(W > 5200) = P(\bar{x} > 208) \doteq P(Z > \tfrac{208-190}{35/\sqrt{25}}) = P(Z > 2.57) = 0.0051$$

5.56. **(a)** Although the probability of having to pay for a total loss for one or more of the 12 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 they collect 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 10,000 policies will be approximately Normally distributed with mean \$250 and standard deviation $\sigma/\sqrt{10{,}000} = \$1000/100 = \$10$. Since \$275 is 2.5 standard deviations above the mean, the probability of seeing an average loss over \$275 is about 0.0062.

5.57. Over 45 years, $\bar{x}$ (the mean return) is approximately Normal with $\mu_{\bar{x}} = 9.2\%$ and $\sigma_{\bar{x}} = 20.6\%/\sqrt{45} \doteq 3.0709\%$, so $P(\bar{x} > 15\%) = P(Z > 1.89) = 0.0294$ and $P(\bar{x} < 5\%) = P(Z < -1.37) = 0.0853$.

Note: *We have to assume that returns in separate years are independent.*

5.58. Let D be the shaft diameter so that D has a $N(2.45 \text{ cm}, 0.01 \text{ cm})$ distribution. **(a)** 99.53% fit into the hole: $P(D \le 2.476) = P(Z \le \frac{2.476-2.45}{0.01}) = P(Z \le 2.60) = 0.9953$. **(b)** 96.71%: If H is the hole diameter, then $H - D$ has a Normal distribution with mean $2.5 - 2.45 = 0.05$ cm and standard deviation $\sqrt{0.01^2 + 0.01^2} \doteq 0.01414$ cm. Therefore, $P(H - D \ge 0.024) = P(Z \ge \frac{0.024-0.05}{0.01414}) = P(Z \ge -1.84) \doteq 0.9671$ (software: 0.9670).

5.59. (a) The mean of six untreated specimens has a standard deviation of $2.2/\sqrt{6} \doteq 0.8981$ lbs, so $P(\bar{x}_u > 50) = P(Z > \frac{50-57}{0.8981}) = P(Z > -7.79)$, which is basically 1. **(b)** The mean of $\bar{x}_u - \bar{x}_t$ is $57 - 30 = 27$ lbs, and the standard deviation is $\sqrt{2.2^2/6 + 1.6^2/6} \doteq 1.1106$ lbs, so $P(\bar{x}_u - \bar{x}_t > 25) = P(Z > \frac{25-27}{1.1106}) = P(Z > -1.80) \doteq 0.9641$.

5.60. (a) The central limit theorem says that the sample means will be roughly Normal. Note that the distribution of individual scores cannot have extreme outliers because all scores are between 1 and 7. **(b)** For *Journal* scores, $\bar{y}$ has mean 4.8 and standard deviation $1.5/\sqrt{28} \doteq 0.2835$. For *Enquirer* scores, $\bar{x}$ has mean 2.4 and standard deviation $1.6/\sqrt{28} \doteq 0.3024$. **(c)** $\bar{y} - \bar{x}$ has (approximately) a Normal distribution with mean 2.4 and standard deviation $\sqrt{1.5^2/28 + 1.6^2/28} \doteq 0.4145$. **(d)** $P(\bar{y} - \bar{x} \ge 1) = P(Z \ge \frac{1-2.4}{0.4145}) = P(Z \ge -3.38) \doteq 0.9996$.

5.61. (a) $\bar{y}$ 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 $\sqrt{\sigma_Y^2/m + \sigma_X^2/n}$.

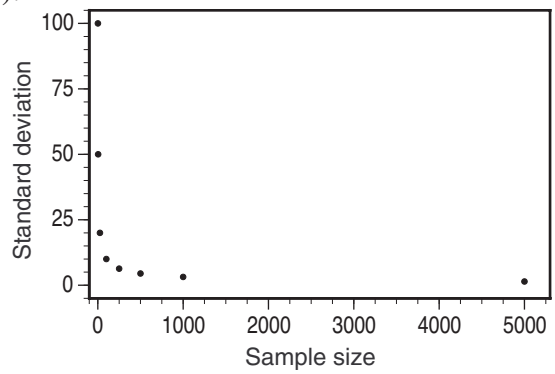
5.62. We have been given $\mu_X = 9\%$, $\sigma_X = 19\%$, $\mu_Y = 11\%$, $\sigma_Y = 17\%$, and $\rho = 0.6$. **(a)** Linda's return $R = 0.7X + 0.3Y$ has mean $\mu_R = 0.7\mu_X + 0.3\mu_Y = 9.6\%$ and standard deviation $\sigma_R = \sqrt{(0.7\sigma_X)^2 + (0.3\sigma_Y)^2 + 2\rho(0.7\sigma_X)(0.3\sigma_Y)} \doteq 16.8611\%$. **(b)** $\overline{R}$, the average return over 20 years, has approximately a Normal distribution with mean 9.6% and standard deviation $\sigma_R/\sqrt{20} \doteq 3.7703\%$, so $P(\overline{R} < 5\%) \doteq P(Z < -1.22) \doteq 0.1112$. **(c)** After a 12% gain in the first year, Linda would have \$1120; with a 6% gain in the second year, her portfolio would be worth \$1187.20. By contrast, two years with a 9% return would make her portfolio worth \$1188.10.

Note: *As the text suggests, the appropriate average for this situation is (a variation on) the* geometric mean, *computed as* $\sqrt{(1.12)(1.06)} - 1 \doteq 8.9587\%$. *Generally, if the sequence of annual returns is* $r_1, r_2, \ldots, r_k$ *(expressed as decimals), the mean return is* $\sqrt[k]{(1+r_1)(1+r_2)\cdots(1+r_k)} - 1$. *It can be shown that the geometric mean is always smaller than the arithmetic mean, unless all the returns are the same.*

5.63. The total height H of the four rows has a Normal distribution with mean $4 \times 8 = 32$ inches and standard deviation $0.1\sqrt{4} = 0.2$ inch. $P(H < 31.5 \text{ or } H > 32.5) = 1 - P(31.5 < H < 32.5) = 1 - P(-2.50 < Z < 2.50) = 1 - 0.9876 = 0.0124$.

5.64. **(a)** The table of standard deviations is given below. **(b)** The graph is below on the right; it is shown as a scatterplot, but in this situation it would be reasonable to "connect the dots" because the relationship between standard deviation and sample size holds for all n. **(c)** As n increases, the standard deviation decreases—at first quite rapidly, then more slowly (a demonstration of the law of diminishing returns).

n	$\sigma/\sqrt{n}$
1	100
4	50
25	20
100	10
250	6.32
500	4.47
1000	3.16
5000	1.41



5.65. Out of five independent drivers, the number A who have an accident in the same year has a binomial distribution with $n = 5$ and $p = 0.2$, so $P(A \geq 3) = 0.0579$ (using Table C, software, or a calculator). Five roommates could not be considered independent, as they were not randomly chosen from the population of all drivers.

5.66. **(a)** Julie's score is about the 65th percentile: $P(X \leq 1110) \doteq P(Z \leq \frac{1110 - 1026}{209}) = P(Z \leq 0.40) = 0.6554$ (software: 0.6561). Note that a continuity correction would be reasonable here, as SAT scores are reported as whole numbers; that is, we could compute $P(X \leq 1110.5)$. However, this makes little difference: Table A gives the same answer, while software changes slightly to 0.6570. **(b)** $\bar{x} = 1110$ is about the 99.99th percentile: $\bar{x}$ is approximately Normal with mean 1026 and standard deviation $209/\sqrt{80} \doteq 23.3669$, so $P(\bar{x} \leq 1110) \doteq P(Z \leq 3.59) = 0.9998$. (The continuity correction would not be appropriate in this case.) **(c)** The first answer is less accurate: The distribution of an individual's score (like Julie's) might not be Normal, but the central limit theorem says that the distribution of $\bar{x}$ will be close to Normal.

5.67. **(a)** Out of 12 independent vehicles, the number X with one person has a binomial distribution with $n = 12$ and $p = 0.7$, so $P(X \geq 7) = 0.8822$ (using Table C, software, or a calculator). **(b)** Y (the number of one-person cars in a sample of 80) has a binomial distribution with $n = 80$ and $p = 0.7$. Regardless of the approach used—Normal approximation, or exact computation using software or a calculator—$P(Y \geq 41) \doteq 0.9999$.

5.68. This would not be surprising: Assuming that all the authors are independent (for example, none were written by siblings or married couples), we can view the 12 names as being a random sample so that the number N of occurrences of the ten most common names would have a binomial distribution with $n = 12$ and $p = 0.056$. Then $P(N = 0) = (1 - 0.056)^{12} \doteq 0.5008$.

5.69. The probability that the first digit is 1, 2, or 3 is $0.301 + 0.176 + 0.125 = 0.602$, so the number of invoices for amounts beginning with these digits should have a binomial distribution with $n = 1000$ and $p = 0.602$. More usefully, the proportion $\hat{p}$ of such invoices should have approximately a Normal distribution with mean $p = 0.602$ and standard deviation $\sqrt{p(1-p)/1000} \doteq 0.01548$, so $P(\hat{p} \le \frac{560}{1000}) \doteq P(Z \le -2.71) = 0.0034$. Alternate answers shown on the right.

		Continuity correction	
Table Normal	Software Normal	Table Normal	Software Normal
0.0034	0.0033	0.0037	0.0037

5.70. **(a)** If R is the number of red-blossomed plants out of a sample of 12, then $P(R = 9) = 0.2581$, using a binomial distribution with $n = 12$ and $p = 0.75$. (For Table C, use $p = 0.25$ and find $P(X = 3)$, where $X = 12 - R$ is the number of flowers with nonred blossoms.) **(b)** With $n = 120$, the mean number of red-blossomed plants is $np = 90$. **(c)** If R_2 is the number of red-blossomed plants out of a sample of 120, then $P(R_2 \ge 80) \doteq P(Z \ge -2.11) = 0.9826$. (Other possible answers are given in the table on the right.)

			Continuity correction	
Exact prob.	Table Normal	Software Normal	Table Normal	Software Normal
0.9845	0.9826	0.9825	0.9864	0.9866

5.71. If $\bar{x}$ is the average weight of 12 eggs, then $\bar{x}$ has a $N(65 \text{ g}, 5/\sqrt{12} \text{ g}) = N(65 \text{ g}, 1.4434 \text{ g})$ distribution, and $P(\frac{755}{12} < \bar{x} < \frac{830}{12}) \doteq P(-1.44 < Z < 2.89) = 0.9231$ (software: 0.9236).

5.72. The probability that an airman completes a tour of duty without being on a lost aircraft is $0.95^{30} \doteq 0.2146$.

5.73. If $\hat{p}$ is the sample proportion who have been on a diet, then $\hat{p}$ has approximately a $N(0.70, 0.02739)$ distribution, so $P(\hat{p} \ge 0.75) \doteq P(Z \ge 1.83) = 0.0336$ (software: 0.0339). Alternatively, as $\hat{p} \ge 0.75$ is equivalent to 210 or more dieters in the sample, we can compute this probability using the binomial distribution; these answers are shown in the table.

			Continuity correction	
Exact prob.	Table Normal	Software Normal	Table Normal	Software Normal
0.0375	0.0336	0.0339	0.0392	0.0392

5.74. **(a)** The machine that makes the caps and the machine that applies the torque are not the same. **(b)** T (torque) is $N(7.0, 0.9)$ and S (cap strength) is $N(10.1, 1.2)$, so $T - S$ is $N(7 - 10.1, \sqrt{0.9^2 + 1.2^2}) = N(-3.1 \text{ inch} \cdot \text{lb}, 1.5 \text{ inch} \cdot \text{lb})$. The probability that the cap breaks is $P(T > S) = P(T - S > 0) = P(Z > 2.07) = 0.0192$ (software: 0.0194).

5.75. The center line is $\mu_{\bar{x}} = \mu = 4.25$ and the control limits are $\mu \pm 3\sigma/\sqrt{5} = 4.0689$ to 4.4311.

5.76. **(a)** $\bar{x}$ has a $N(32, 6/\sqrt{25}) = N(32, 1.2)$ distribution, and $\bar{y}$ has a $N(29, 5/\sqrt{25}) = N(29, 1)$ distribution. **(b)** $\bar{y} - \bar{x}$ has a $N(29 - 32, \sqrt{5^2/25 + 6^2/25}) \doteq N(-3, 1.5620)$ distribution. **(c)** $P(\bar{y} \ge \bar{x}) = P(\bar{y} - \bar{x} \ge 0) = P(Z \ge 1.92) = 0.0274$.

5.77. **(a)** $\hat{p}_F$ is approximately $N(0.82, 0.01921)$ and $\hat{p}_M$ is approximately $N(0.88, 0.01625)$. **(b)** When we subtract two independent Normal random variables, the difference is Normal. The new mean is the difference of the two means ($0.88 - 0.82 = 0.06$), and the new variance is the sum of the variances ($0.01921^2 + 0.01625^2 = 0.000633$), so $\hat{p}_M - \hat{p}_F$ is approximately $N(0.06, 0.02516)$. **(c)** $P(\hat{p}_F > \hat{p}_M) = P(\hat{p}_M - \hat{p}_F < 0) \doteq P(Z < -2.38) = 0.0087$ (software: 0.0085).

5.78. **(a)** Yes; this rule works for any random variables X and Y. **(b)** No; this rule requires that X and Y be independent. The incomes of two married people are certainly not independent, as they are likely to be similar in many characteristics that affect income (for example, educational background).

5.79. For each step of the random walk, the mean is $\mu = (1)(0.6) + (-1)(0.4) = 0.2$, the variance is $\sigma^2 = (1 - 0.2)^2(0.6) + (-1 - 0.2)^2(0.4) = 0.96$, and the standard deviation is $\sigma = \sqrt{0.96} \doteq 0.9798$. Therefore, $Y/500$ has approximately a $N(0.2, 0.04382)$ distribution, and $P(Y \geq 200) = P(\frac{Y}{500} \geq 0.4) \doteq P(Z \geq 4.56) \doteq 0$.

Note: *The number R of right-steps has a binomial distribution with $n = 500$ and $p = 0.6$. $Y \geq 200$ is equivalent to taking at least 350 right-steps, so we can also compute this probability as $P(R \geq 350)$, for which software gives the exact value 0.00000215....*