Chapter 3 Solutions

3.1. Jamie and his friends do not constitute a random sample, and so would not be representative of the population of all "young people."

3.2. The anecdote describes a single unusual event. We would like data on deaths and injuries for occupants wearing/not wearing restraints for many accidents.

3.3. This is an observational study: No treatment was assigned to the subjects; we merely observed cell phone usage (and presence/absence of cancer). The explanatory variable is cell phone usage, and the response variable is whether or not a subject has brain cancer.

3.4. **(a)** No treatment is imposed on the subjects (children); they (or their parents) choose how much TV they watch. The explanatory variable is hours watching TV, and the response variable is "later aggressive behavior." **(b)** An adolescent who watches a lot of television probably is more likely to spend less time doing homework, playing sports, or having social interactions with peers. He or she may also have less contact with or guidance from his/her parents.

3.5. This is an experiment: Each subject is (presumably randomly) assigned to a group, each with its own treatment (computer animation or reading the textbook). The explanatory variable is the teaching method, and the response variable is the change in each student's test score.

3.6. **(a)** In this observational study, we do not know if the reduction in heart attack risk was due to hormones or other factors. An experiment would have involved taking a (very large) group of women, splitting them randomly into two groups; one would take hormones, and the other would not. The random allocation would allow us to conclude that, if we observed a difference in heart attack risk in either group, it could be attributed to the treatment. **(b)** Possible characteristics might include socio-economic factors; for example, women who take hormones might be wealthier and therefore have better access to health care. There might also be biological factors; perhaps women who are more sensitive to menopause (and are therefore more likely to use hormones to relieve its symptoms) are also at less risk for heart attacks.

3.7. **(a)** In an observational study, we might take a sample and classify each subject as a wine drinker or a beer drinker. For an experiment, we would assign each subject to drink either wine or beer. In either case, we would then observe the health of the subjects over time. **(b)** Wine drinkers might be wealthier, better educated, have white-collar jobs, and have different dietary habits ("beer and pretzels" vs. "wine and cheese").

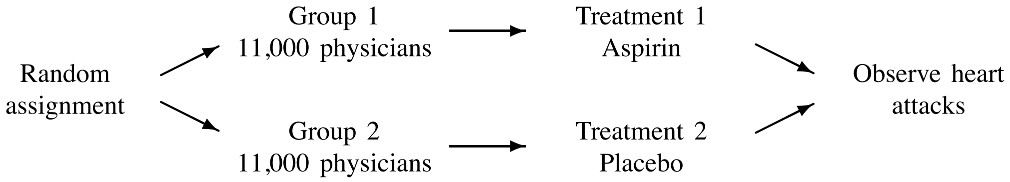
3.8. This number is subject to change, as is the NCES web site. At the time of this writing, a search at the NCES web site for the phrase "undergraduates work part-time" yielded a document (dated February 1998) which included the following statement: "In 1992–93, 72 percent of the undergraduates in this analysis worked while enrolled, . . . an average of 31 hours per week and 88 percent of the months they were enrolled." Note that this was one of the top 10 results, but not the first result. A search at `www.google.com` for the same phrase also yielded many matches, at sites other than NCES, but this same document was conveniently at the top of the list.

3.9. Experimental units: Pine tree seedlings. Factor: Amount of light. Treatments: Full light, or shaded to 5% of normal. Response variable: Dry weight at end of study.

3.10. Experimental units: Middle schools. Factors: Physical activity program, and nutrition program. Treatments (four): Activity intervention, nutrition intervention, both interventions, and neither intervention. Response variables: Physical activity and lunchtime consumption of fat.

3.11. Subjects: Adults (or registered voters) from selected households. Factors: Level of identification, and offer of survey results. Treatments (six): Interviewer's name with results, interviewer's name without results, university name with results, university name without results, both names with results, both names without results. Response variable: Whether or not the interview is completed.

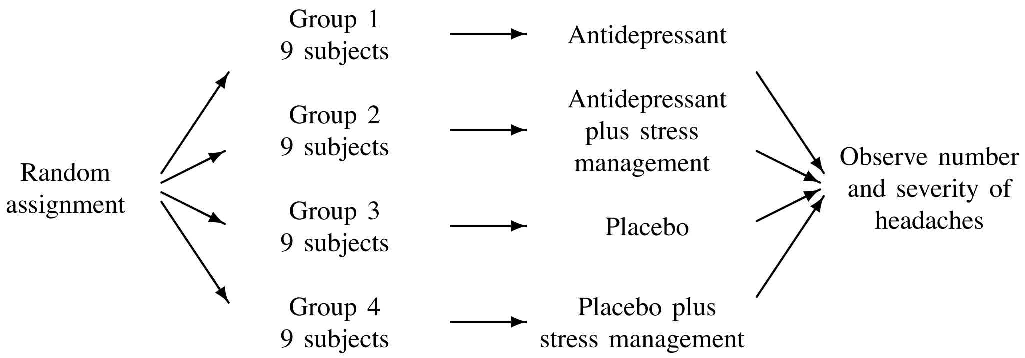
3.12. **(a)** The subjects are the physicians, the factor is medication (aspirin or placebo), and the response variable is health, specifically whether the subjects have heart attacks. **(b)** Below. **(c)** The difference in the number of heart attacks between the two groups was so great that it would rarely occur by chance if aspirin had no effect.



3.13. Assign nine subjects to each treatment. A diagram is below; if we assign labels 01 through 36, then line 130 gives:

Group 1		Group 2		Group 3	
05 Chen	32 Vaughn	31 Valasco	02 Asihiro	35 Willis	11 Fleming
16 Imrani	04 Bikalis	18 Kaplan	36 Zhang	21 Marsden	15 Hruska
17 James	25 Padilla	07 Duncan	23 O'Brian	26 Plochman	12 George
20 Maldonado	29 Trujillo	13 Han	27 Rosen	08 Durr	14 Howard
19 Liang		33 Wei		10 Farouk	

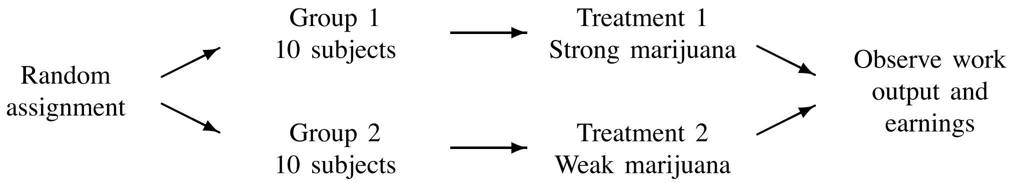
The other nine subjects are in Group 4. See note on page 50 about using Table B.



3.14. (a) A diagram is shown below. **(b)** Label the subjects from 01 through 20. From line 131, we choose

05, 19, 04, 20, 16, 18, 07, 13, 02, and 08;

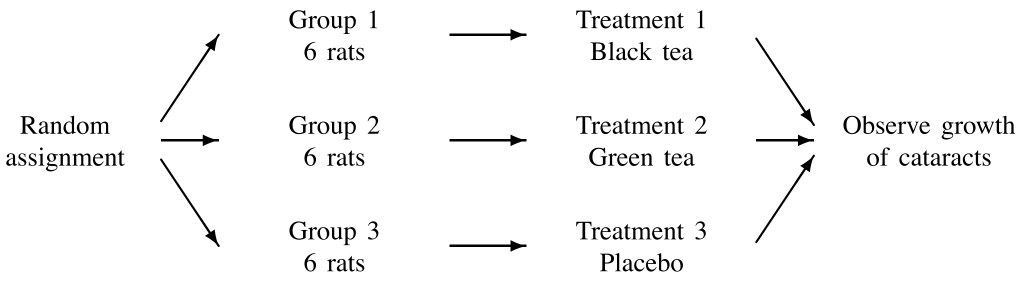
that is, Decker, Travers, Chen, Ullmann, Quinones, Thompson, Fluharty, Lucero, Afifi, and Gerson for one group, and the rest for the other. See note on page 50 about using Table B.



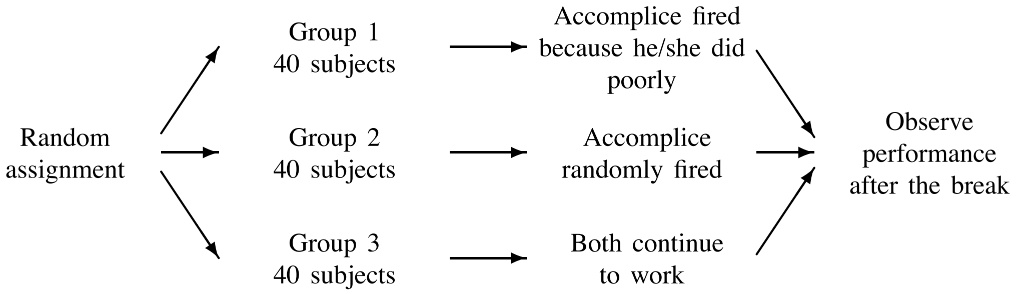
3.15. **(a)** Diagram below. **(b)** If we assign labels 01, . . . , 18 and begin on line 142, then we select:

02, 08, 17, 10, 05, and 09 for Group 1;
06, 16, 01, 07, 18, and 15 for Group 2.

The remaining rats are assigned to the placebo group. See note on page 50 about using Table B.



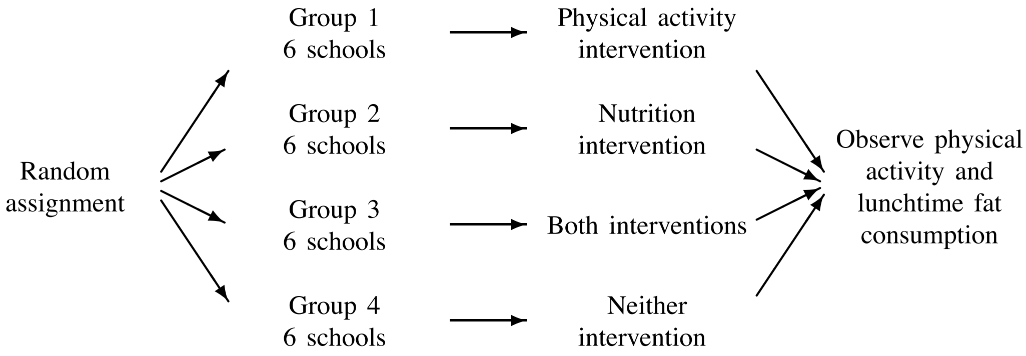
3.16. **(a)** Diagram below. **(b)** Using line 123 from Table B, the first four subjects are 102, 063, 035, and 090. See note on page 50 about using Table B.



3.17. Diagram below. Starting at line 105, we choose:

07, 19, 14, 17, 13, 15 for Group 1,
08, 21, 20, 11, 24, 09 for Group 2,
06, 23, 16, 18, 12, 04 for Group 3,

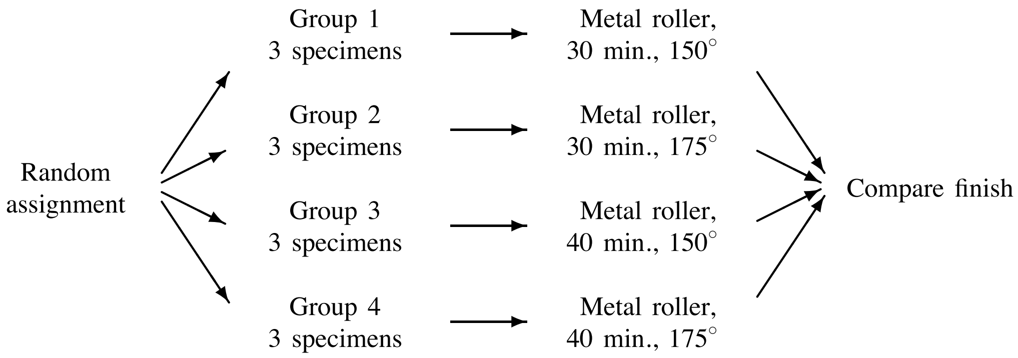
and the rest for Group 4. See note on page 50 about using Table B.



3.18. (a) The table below shows the 16 treatments—four levels for each of the two factors. **(b)** A diagram is not shown here (it is quite large). Five subjects are randomly assigned to each treatment; they read the ad for that treatment, and we record their attractiveness ratings for the ad. Using line 133, the first five subjects are 45, 74, 04, 18, and 07.

		Factor B **Fraction of shoes on sale**			
		25%	50%	75%	100%
Factor A **Discount level**	20%	Treatment 1	Treatment 2	Treatment 3	Treatment 4
	40%	Treatment 5	Treatment 6	Treatment 7	Treatment 8
	60%	Treatment 9	Treatment 10	Treatment 11	Treatment 12
	80%	Treatment 13	Treatment 14	Treatment 15	Treatment 16

3.19. (a) There are three factors (roller type, dyeing cycle time, and temperature), each with two levels, for a total of $2^3 = 8$ treatments. The experiment therefore requires 24 fabric specimens. **(b)** In the interest of space, <u>only the top half of the diagram</u> is shown below. The other half consists of Groups 5 to 8, for which the treatments have natural bristle rollers instead of metal rollers.



3.20. Population = 1 to **40** , Select a sample of size **20** , then click Reset and Sample .

3.21. (a) Population = 1 to **150** , Select a sample of size **25** , click Reset and Sample .

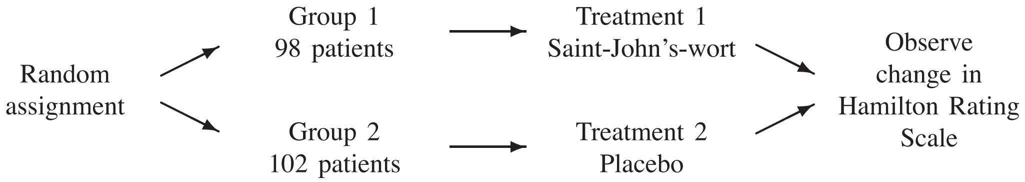
(b) Without resetting, click Sample again.

3.22. For a range of discounts, the attractiveness of the sale decreases slightly as the percentage of goods on sale increases. (The decrease is so small that it might not be significant.) With precise discounts, on the other hand, mean attractiveness increases with the percentage on sale. Range discounts are more attractive when only 25% of goods are marked down, while the precise discount is more attractive if 75% or 100% of goods are discounted.

3.23. The first design is an observational study. It is flawed because the women observed chose whether or not to take bee pollen; one might reasonably expect that people who choose to take bee pollen have other dietary or health habits that would differ from those who do not. The second design is an experiment; because the treatment is randomly assigned, the effect of other habits would be "diluted" because they would be more-or-less equally split between the two groups. Therefore, any difference in colon health between the two groups could be attributed to the treatment (bee pollen or not).

3.24. "Randomized" means that patients were randomly assigned to receive either the standard morphine treatment or CR morphine tablets. "Double blind" means that the treatment assigned to a patient was unknown to both the patient and those responsible for assessing the effectiveness of that treatment. (It is not clear how the treatment was hidden from the patients, because they would know when they received the morphine.) "Comparative" means researchers compared the effectiveness of two treatments, rather than simply trying to assess the effectiveness of one treatment—that is, researchers did not simply change over to CR morphine and try to judge if it was better than the standard treatment had been in the past.

3.25. (a) "Randomized" means that patients were randomly assigned to receive either Saint-John's-wort or a placebo. "Double blind" means that the treatment assigned to a patient was unknown to both the patient and those responsible for assessing the effectiveness of that treatment. "Placebo-controlled" means that some of the subjects were given placebos. Even though these possess no medical properties, some subjects may show improvement or benefits just as a result of participating in the experiment; the placebos allow those doing the study to observe this effect. **(b)** Diagram below.



3.26. The mean Monday return for the first three weeks of the month was both different from zero and higher than the mean for the last two Mondays. However, the difference from zero was small enough that it might have occurred purely by chance (and so it gives no reason to suspect that the first three Mondays tend to produce negative returns). On the other hand, the difference between the first three Mondays and the last two Mondays was so large that it would rarely occur by chance, leading us to conclude that the last two Mondays really do (for whatever reason) tend to yield lower returns than the first three Mondays.

3.27. As described, there are two factors: ZIP code (three levels: none, 5-digit, 9-digit) and the day on which the letter is mailed (three levels: Monday, Thursday, or Saturday) for a total of 9 treatments. To control lurking variables, aside from mailing all letters to the same address, all letters should be the same size, and either printed in the same handwriting or typed. The design should also specify how many letters will be in each treatment group. Also, the letters should be sent randomly over many weeks.

3.28. Results will vary, but probability computations reveal that more than 97% of samples will have 7 to 13 fast-reacting subjects (and 99.6% of samples have 8 to 14 fast-reacting subjects). Additionally, if students average their 10 samples, nearly all students (more than 99%) should find that the average number of fast-reacting subjects is between 8.5 and 11.5.

Note: *X, the number of fast-reacting subjects in the sample, has a hypergeometric distribution with parameters* $N = 40$, $r = 20$, $n = 20$, *so that* $P(7 \le X \le 13) \doteq 0.974$. *The theoretical average number of fast-reacting subjects is 10.*

3.29. Each player will be put through the sequence (100 yards, four times) twice—once with oxygen and once without, and we will observe the difference in their times on the final run. (If oxygen speeds recovery, we would expect that the oxygen-boosted time will be lower.) Randomly assign half of the players to use oxygen on the first trial, while the rest use it on the second trial. Trials should be on different days to allow ample time for full recovery.

If we label the players 01 through 20 and begin on line 170, we choose 07, 13, 14, 02, 12, 20, 06, 08, 16, 03 to be in the oxygen-first group. See note on page 50 about using Table B.

3.30. The sketches requested in the problem are not shown here; random assignments will vary among students. **(a)** Label the circles 1 to 6, then randomly select three (using Table B, or simply by rolling a die) to receive the extra CO_2. Observe the growth in all six regions, and compare the mean growth within the three treated circles with the mean growth in the other three (control) circles. **(b)** Select pairs of circles in each of three different areas of the forest. For each pair, randomly select one circle to receive the extra CO_2 (using Table B or by flipping a coin). For each pair, compute the difference in growth (treated minus control).

3.31. **(a)** Randomly assign half the girls to get high-calcium punch, and the other half get low-calcium punch. The response variable is not clearly described in this exercise; the best we can say is "observe how the calcium is processed." **(b)** Randomly select half of the girls to receive high-calcium punch first, while the other half gets low-calcium punch first, then for each subject, compute the difference in the response variable for each level. This is a better design because it deals with person-to-person variation; the differences in responses for 60 individuals gives more precise results than the difference in the average responses for two groups of 30 subjects. **(c)** The first five subjects are 16, 34, 59, 44, and 21. In the CR design, the first group receives high-calcium punch all summer; in the matched pairs design, they receive high-calcium punch for the first part of the summer, and then low-calcium punch in the second half.

3.32. **(a)** Ordered by increasing weight, the five blocks are

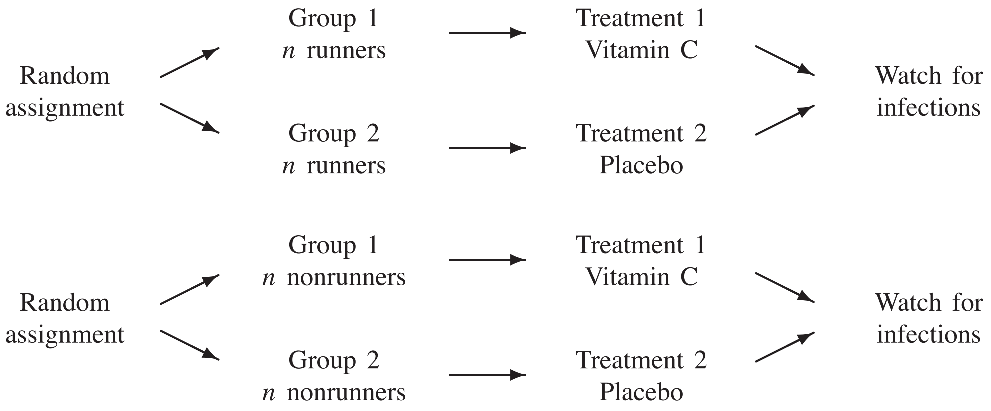
(1)	Williams	22	Festinger	24	Hernandez	25	Moses	25
(2)	Santiago	27	Kendall	28	Mann	28	Smith	29
(3)	Brunk	30	Obrach	30	Rodriguez	30	Loren	32
(4)	Jackson	33	Stall	33	Brown	34	Dixon	34
(5)	Birnbaum	35	Tran	35	Nevesky	39	Wilansky	42

(b) The exact randomization will vary with the starting line in Table B. Different methods are possible; perhaps the simplest is to number from 1 to 4 within each block, then assign the members of block 1 to a weight-loss treatment, then assign block 2, and so on. For example, starting on line 133, we assign 4–Moses to treatment A, 1–Williams to B, and

3–Hernandez to C (so that 2–Festinger gets treatment D), then carry on for block 2, and so forth (either continuing on the same line or starting over somewhere else).

3.33. The simplest design would be a completely randomized design, assigning half of the women to take strontium renelate and half to take the placebo. A better design would block according to the medical center (and country); that is, randomly select for the strontium renelate group half the women from country A, half of those from country B, and so on. This blocking would take care of any differences in the level of medical care from one country to the next.

3.34. **(a)** This is a block design. **(b)** The diagram might be similar to the one below (which assumes equal numbers of subjects in each group). **(c)** The results observed in this study would rarely have occurred by chance if vitamin C were ineffective.



3.35. **(a)** False. Such regularity holds only in the long run. If it were true, you could look at the first 39 digits and know whether or not the 40th was a 0. **(b)** True. All pairs of digits (there are 100, from 00 to 99) are equally likely. **(c)** False. Four random digits have chance 1/10000 to be 0000, so this sequence will occasionally occur. 0000 is no more or less random than 1234 or 2718, or any other four-digit sequence.

3.36. The population is (all) local businesses. The sample is the 73 businesses that return the questionnaire, *or* the 150 businesses selected. The nonresponse rate is $51.3\% = \frac{77}{150}$.

Note: *The definition of "sample" makes it somewhat unclear whether the sample includes all the businesses selected, or only those which responded. My inclination is toward the latter (the smaller group), which is consistent with the idea that the sample is "a part of the population that we actually examine."*

3.37. **(a)** The population is adult residents of the U.S. The sample size is 1002. **(b)** Perhaps people are more inclined to respond with the first or last option they hear. Rotating the order of the options would cancel out any effect on the response of such inclinations.

3.38. Exact descriptions of the populations may vary. **(a)** Adult residents of the U.S. **(b)** All pieces of wood in the lot. **(c)** All U.S. households.

3.39. **(a)** The population is (something like) adult residents of the U.S. **(b)** The nonresponse rate is $\frac{669}{1800} \doteq 37.2\%$. **(c)** This question will likely have response bias; specifically, many people will give an inaccurate count of how many movies they have seen in the past year.

3.40. Numbering from 01 to 33 alphabetically (down the columns), we enter Table B at line 117 and choose
16=Fairington, 32=Waterford Court, 18=Fowler, 06=Briarwood, 23=Mayfair Village.
See note on page 50 about using Table B.

3.41. Numbering from 01 to 32 alphabetically (down the columns), we enter Table B at line 139 and choose
04=Bell, 10=Fernandez, 17=Johnson, 19=Molina, 12=Gandhi,
32=Zhao, 13=Garcia, 18=Kim, 24=Prince, 28=Samuels.
See note on page 50 about using Table B.

3.42. Population = 1 to **33**, Select a sample of size **5**, then click Reset and Sample.

3.43. With the applet: Population = 1 to **371**, Select a sample of size **25**, then click Reset and Sample. With Table B, line 149 gives the codes labeled 353, 246, 027, 038, and 207.

3.44. One could use the labels already assigned to the blocks, but that would mean skipping a lot of four-digit combinations that do not correspond to any block. An alternative would be to drop the second digit and use labels 100,101,102,. . . ,105; 200,. . . ,211; 300,. . . ,325. But by far the simplest approach is to assign labels 01, . . . , 44 (in numerical order by the four-digit numbers already assigned), enter the table at line 125, and select:
21 (#3002), 37 (#3018), 18 (#2011), 44 (#3025), and 23 (#3004).
See note on page 50 about using Table B.

3.45. If one always begins at the same place, then the results could not really be called random.

3.46. The sample will vary with the starting line in Table B. The simplest method is to use the last digit of the numbers assigned to the blocks in Group 1 (that is, assign the labels 0–5), then choose one of those blocks; use the last two digits of the blocks in Group 2 (00–11) and choose two of those, and finally use the last two digits of the blocks in Group 3 (00–25) and choose three of them.

3.47. **(a)** If we choose one of the first 45 students, and then every 45th name after that, we will have a total of $\frac{9000}{45} = 200$ names. **(b)** Label the first 45 names 01–45. Beginning at line 145, the first number we find is 19, so we choose names 19, 64, 109, ...

3.48. Considering the 9000 students of Exercise 3.47, each student is equally likely; specifically, each name has chance 1/45 of being selected. To see this, note that each of the first 45 has chance 1/45 because one is chosen at random. But each student in the second 45 is chosen exactly when the corresponding student in the first 45 is, so each of the second

45 also has chance 1/45. And so on.

This is not an SRS because the only possible samples have exactly one name from the first 45, one name from the second 45, and so on; that is, there are only 45 possible samples. An SRS could contain *any* 200 of the 9000 students in the population.

3.49. **(a)** This is a stratified random sample. **(b)** Label from 01 through 25; beginning at line 111, we choose
12 (559), 04 (209), 11 (805), 19 (562), 02 (707),
06 (925), 08 (650), 25 (619), 17 (626), and 14 (661).

3.50. Assign labels 01–36 for the Climax 1 group, 01–72 for the Climax 2 group, and so on. Then beginning at line 162, choose
34, 14, 15, 36 from the Climax 1 group, and (continuing on in Table B) choose
23, 36, 21, 11, 55, 27, 14 from the Climax 2 group,
28, 31, 09 from the Climax 3 group, and
03, 41, 37, 16 from the mature secondary group.
See note on page 50 about using Table B.

3.51. Label the students 01, . . . , 30 and use Table B. Then label the faculty 0, . . . , 9 and use the table again. (You could also label the faculty from 01 to 10, but that would needlessly require two-digit labels.)

Note: *Students often try some fallacious method of choosing both samples simultaneously. We simply want to choose two separate SRSs: one from the students and one from the faculty. See note on page 50 about using Table B.*

3.52. Each student has a 10% chance: 3 out of 30 over-21 students, and 2 of 20 under-21 students. This is not an SRS because not every group of 5 students can be chosen; the only possible samples are those with 3 older and 2 younger students.

3.53. Label the 500 midsize accounts from 001 to 500, and the 4400 small accounts from 0001 to 4400. On line 115, we first encounter numbers 417, 494, 322, 247, and 097 for the midsize group, then 3698, 1452, 2605, 2480, and 3716 for the small group. See note on page 50 about using Table B.

3.54. The higher no-answer was probably the second period—more families are likely to be gone for vacations, and so on. Nonresponse of this type might underrepresent those who are more affluent (and are able to travel). In general, high nonresponse rates always make results less reliable, because we do not know what information we are missing.

3.55. **(a)** This design would omit households without telephones or with unlisted numbers. Such households would likely be made up of poor individuals (who cannot afford a phone), those who choose not to have phones, and those who do not wish to have their phone numbers published. **(b)** Those with unlisted numbers would be included in the sampling frame when a random-digit dialer is used.

3.56. **(a)** There were 14,484 responses. (Note that we have no guarantee that these came from 14,484 distinct people; some may have voted more than once.) **(b)** This voluntary response sample collects only the opinions of those who visit this site and feel strongly enough to respond.

3.57. **(a)** This will almost certainly produce a positive response, because it draws the dubious conclusion that cell phones *cause* brain cancer. Some people who drive cars, or eat carrots, or vote Republican develop brain cancer, too. Do we conclude that these activities should come with warning labels, also? **(b)** The phrasing of this question will tend to make people respond in favor of national health insurance: It lists two benefits of such a system, and no arguments from the other side of the issue. **(c)** This sentence is so convoluted and complicated that it is almost unreadable, and is also vague (what sort of 'economic incentives'? How much would this cost?). A better phrasing might be, "Would you be willing to pay more for the products you buy if the extra cost were used to conserve resources by encouraging recycling?" That is still vague, but less so, and is written in plain English.

3.58. The first wording brought the higher numbers in favor of a tax cut; "new government programs" has considerably less appeal than the list of specific programs given in the second wording.

3.59. Children from larger families will be overrepresented in such a sample. Student explanations of why will vary; a simple illustration can aid in understanding this effect. Suppose that there are 100 families with children; 60 families have one child, and the other 40 have three. Then there are a total of 180 children (an average of 1.8 children per family), and *two-thirds* of those children come from families with three children. Therefore, if we had a class (a sample) chosen from these 180 children, only one-third of the class would answer "one" to the teacher's question, and the rest would say "three." This would give an average of about 2.3 children per family.

3.61. Responses to public opinion polls can be affected by things like the wording of the question, as was the case here: Both statements address the question of how to distribute wealth in a society, but subtle (and not-so-subtle) slants in the wording suggest that the public holds conflicting opinions on the subjects.

3.62. 621 is a statistic (related to the sample of 2000 phone numbers); 35% is a parameter (related to the population of all residential phone numbers).

3.63. 72% is a statistic (related to the sample of 663 registered voters); 56% is a parameter (related to the population of all registered voters).

3.64. Both 40.2% and 31.7% are statistics (related, respectively, to the samples of small-class and large-class black students).

3.65. Both 283 and 311 pushes per minute are statistics (related to one sample: the subjects with placebo, and the same subjects with caffeine).

3.66. **(a)** High bias, high variability (many are low, wide scatter). **(b)** Low bias, low variability, (close to parameter, little scatter). **(c)** Low bias, high variability (neither too low nor too high, wide scatter). **(d)** High bias, low variability (too high, little scatter).

Note: *Make sure that students understand that "high bias" means that the values are far from the parameter,* not *that they are too high.*

3.67. **(a)** The sample size for Hispanics was smaller. Smaller sample sizes give less information about the population, and therefore lead to larger margins of error (with the same confidence level). **(b)** The sample size was so small, and the margin of error so large, that the results could not be viewed as an accurate reflection of the population of Cubans.

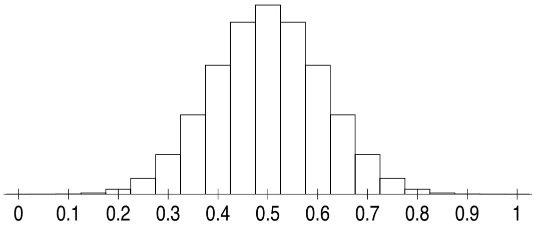
3.68. No: With sufficiently large populations ("at least 100 times larger than the sample"), the variability (and margin of error) depends on the sample size.

3.69. **(a)** Because the smallest population is still more than 100 times the sample size, the variability will be (approximately) the same for all states. **(b)** Yes, it will change—the sample size would vary from 500 in Wyoming to 35,000 in California, so the margin of error would be smaller in larger states.

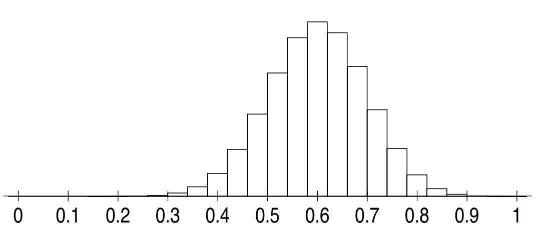
3.70. **(a)** The population is Ontario residents; the sample is the 61,239 people interviewed. **(b)** The sample size is very large, so if there were large numbers of both sexes in the sample—this is a safe assumption because we are told this is a "random sample"—these two numbers should be fairly accurate reflections of the values for the whole population.

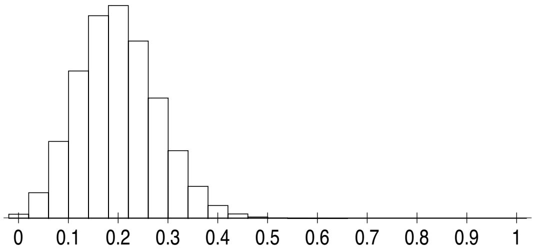
3.71. See also the solution to Exercise 1.146. **(a)** Answers will vary, but almost all means will be between \$25,000 and \$80,000 (see Figure 3.14 in the text). **(b)** For samples of size 25, means range from \$25,000 to \$80,000; for samples of size 100, the range is \$35,000 to \$65,000 (although very few means exceeded \$60,000). This illustrates that sampling variability decreases as sample size increases. **(c)** Based on the histogram, the mean is between \$45,000 and \$50,000. (The actual mean of the 71,076 incomes is \$46,050.)

3.72. **(a)** Answers will vary. If, for example, eight heads are observed, then $\hat{p} = \frac{8}{20} = 0.4 = 40\%$. **(b)** Note that all the leaves in the stemplot should be either 0 or 5, since all possible $\hat{p}$-values end in 0 or 5. For comparison, here is a histogram of the sampling distribution (assuming p really is 0.5). An individual student's stemplot will probably only roughly approximate this distribution, but pooled efforts should be fairly close.

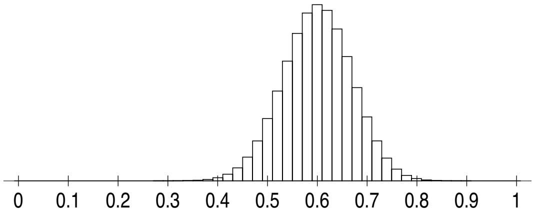


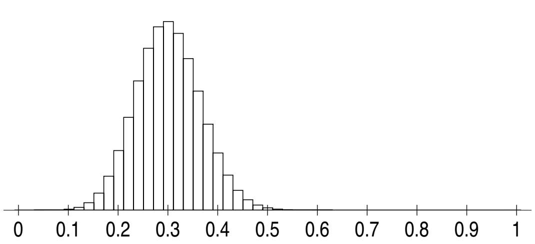
3.73. (a) The histogram should be centered at about 0.6 (with quite a bit of spread). For reference, the theoretical histogram is shown below on the left; student results should have a similar appearance. **(b)** The histogram should be centered at about 0.2 (with quite a bit of spread). The theoretical histogram is shown below on the right.

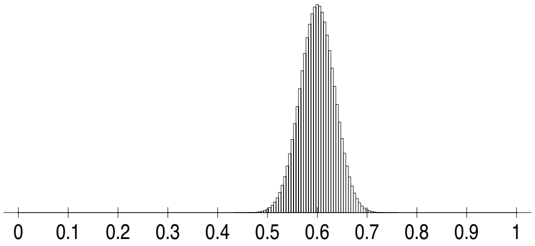




3.74. (a) The histogram of this theoretical sampling distribution is shown (on the right) for reference. **(b)** This theoretical sampling distribution is shown below on the left. Students should observe that their two stemplots have clearly different centers (near 0.6 and 0.3, respectively), but similar spreads. **(c)** The theoretical sampling distribution is below on the right. Compared to the distribution of (a), this has the same center, but is about half as wide; that is, the spread is about half as much when the sample size is multiplied by 4. (The vertical scale of this graph is not the same as the other two; it should be about twice as tall as it is, since it is only about half as wide.)

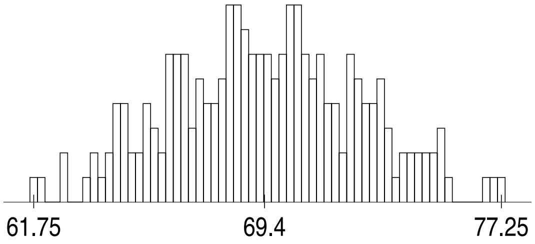


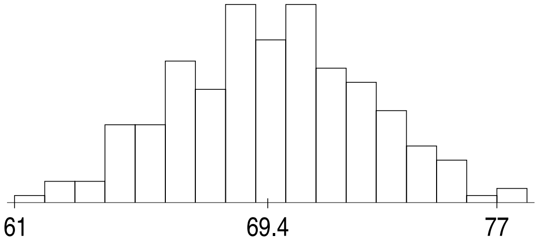




3.75. (a) The scores will vary depending on the starting row. Note that the smallest possible mean is 61.75 (from the sample 58, 62, 62, 65) and the largest is 77.25 (from 73, 74, 80, 82). **(b)** Answers will vary; shown below are two views of the (exact) sampling distribution. The first shows all possible values of the experiment (so the first rectangle is for 61.75, the next is for 62.00, etc.); the other shows values grouped from 61 to 61.75, 62 to 62.75, etc. (which makes the histogram less bumpy). The tallest rectangle in the first picture is 8 units; in the second, the tallest is 28 units.

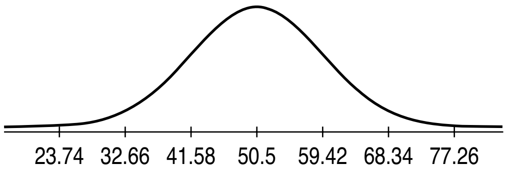
Note: *These histograms were found by considering all* $\binom{10}{4} = 210$ *of the possible* *samples. It happens that half (105) of those samples yield a mean smaller than 69.4 and half yield a greater mean.*

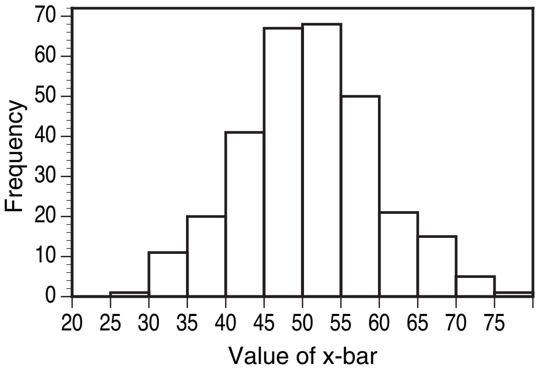




3.76. Student results will vary greatly, and ten values of $\bar{x}$ will give little indication of the appearance of the sampling distribution. In fact, the sampling distribution of $\bar{x}$ is approximately normal with a mean of 50.5 and a standard deviation of about 8.92; this approximating normal distribution is shown on the right (above). Therefore, nearly every sample of size 10 would yield a mean between 23 and 78.

The shape of the sampling distribution becomes more apparent if the results of many students are pooled. Below on the right is an example based on 300 sample means, which might arise from pooling all the results in a class of 30.

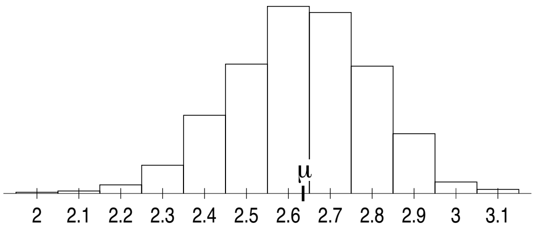




Note: *Because the values in these samples are not independent (there can be no repeats), a stronger version of the central limit theorem is needed to determine that the sampling distribution is approximately normal. Confirming the standard deviation given above is a reasonably difficult exercise even for a mathematics major.*

3.77. (a) Below is the population stemplot (which gives the same information as a histogram). The (population) mean GPA is $\mu \doteq 2.6352$, and the standard deviation is $\sigma \doteq 0.7794$. [Technically, we should take $\sigma \doteq 0.7777$, which comes from dividing by n rather than $n-1$, but few (if any) students would know this, and it has little effect on the results.] **(b)** & **(c)** Results will vary; these histograms are not shown. Not every sample of size 20 could be viewed as "generally representative of the population," but most should bear at least some resemblance to the population distribution.

```
0 | 134
0 | 567889
1 | 0011233444
1 | 5566667888888888999999
2 | 000000000111111111222222222333333333444444444
2 | 55555555555556666666677777777777777888888888888889999999
3 | 0000000000000011111111112222222223333333333333333444444444
3 | 556666666677777788889
4 | 0000
```



3.78. **(a)** Shown for reference is a histogram of the approximate sampling distribution of $\bar{x}$. This distribution is difficult to find exactly, but based on 1000 simulated samples, it is approximately normal with mean 2.6352 (the same as μ) and standard deviation $s_{\bar{x}} \doteq 0.167$. (Therefore, $\bar{x}$ will almost always be between 2.13 and 3.14.) **(b)** Results may vary, but most students should see no strong suggestion of bias. **(c)** Student means and standard deviations will vary, but for most (if not all) students, their values should meet the expectations (close to $\mu \doteq 2.6352$ and less than $\sigma \doteq 0.78$).

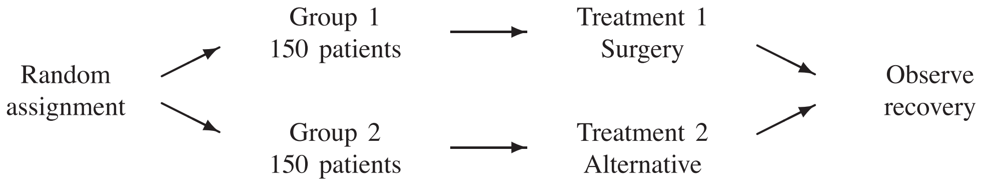
Note: *Observe that the distribution of $\bar{x}$ is slightly left-skewed, but less than the population distribution. Also note that $s_{\bar{x}}$, the standard deviation of the sampling distribution, is smaller than $\sigma/\sqrt{20} \doteq 0.174$, since we are sampling without replacement.*

3.79. **(a)** The simplest approach is to label from 00001 through 14959, and then take five digits at a time from the table. A few clever students might think of some ways to make this process more efficient, such as taking the first random digit chosen as "0" if it is even and "1" if odd. (This way, fewer numbers need to be ignored.) **(b)** Using labels 00001–14959, we choose 03638, 07871, and 12193. Students who try an alternate approach may have a different sample.

3.80. **(a)** Possible response variables: Whether or not a subject has a job within some period of time, number of hours worked during some period, length of time before subject became employed. For the design, randomly assign about one-third of the group (3,355 subjects) to each treatment, and observe the chosen response variables after a suitable amount of time. **(b)** The simplest approach is to label from 00001 through 10065, and then take five digits at a time from the table. (This means we have to skip about 90% of the five-digit sets, as only those beginning with 0 [and a few beginning with 1] are useful.) With this approach, we choose 00850, 02182, and 00681 (the last of these is on line 172). More efficient labellings are possible and will lead to different samples.

3.81. **(a)** A matched pairs design (two halves of the same board would have similar properties). **(b)** A sample survey (with a stratified sample: smokers and nonsmokers). **(c)** A block design (blocked by gender).

3.82. **(a)** In a serious case, when the patient has little chance of surviving, a doctor might choose not to recommend surgery; it might be seen as an unnecessary measure, bringing expense and a hospital stay with little benefit to the patient. **(b)** Diagram below.



3.83. This is an experiment, because each subject is (randomly, we assume) assigned to a treatment. The explanatory variable is the price history seen by the subject (steady prices or fluctuating prices), and the response variable is the price the subject expects to pay.

3.84. **(a)** A sample survey: We want to gather information about a population (U.S. residents) based on a sample. **(b)** An experiment: We want to establish a cause-and-effect relationship between teaching method and amount learned. **(c)** An observational study: There is no particular population from which we will sample; we simply observe "your teachers," much like an animal behavioral specialist might study animals in the wild.

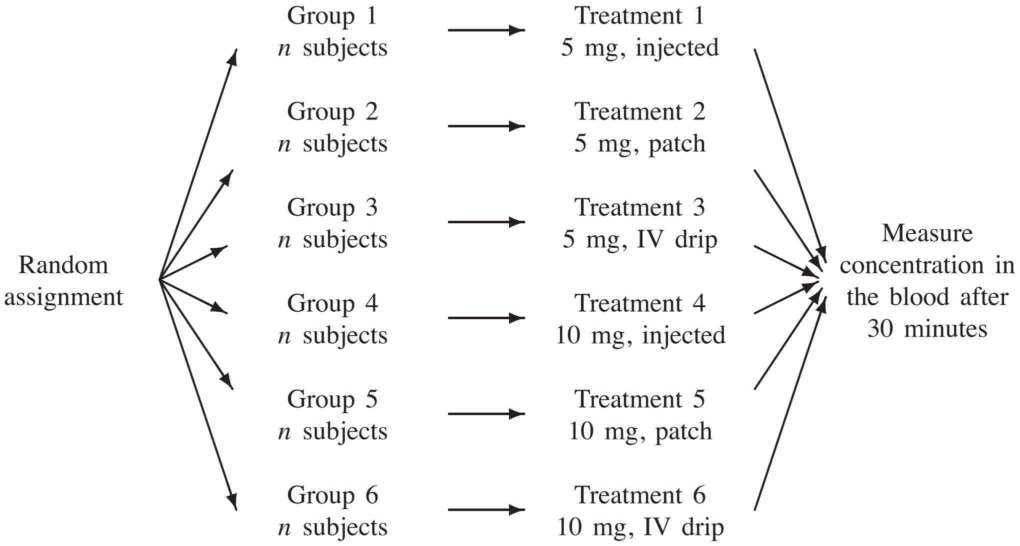
3.86. Each subject should taste both kinds of cheeseburger, in a randomly selected order, and then be asked about preference. Both burgers should have the same "fixings" (ketchup, mustard, etc.). Because some subjects might be able to identify the cheeseburgers by appearance, one might need to take additional steps (such as blindfolding or serving only the center part of the burger) in order to make this a true "blind" experiment.

3.87. The two factors are gear (three levels) and steepness of the course (number of levels not specified). Assuming there are at least three steepness levels—which seems like the smallest reasonable choice—that means at least nine treatments. Randomization should be used to determine the order in which the treatments are applied. Note that we must allow ample recovery time between trials, and it would be best to have the rider try each treatment several times.

3.89. **(a)** One possible population: all full-time undergraduate students in the fall term on a list provided by the registrar. **(b)** A stratified sample with 125 students from each year is one possibility. **(c)** Mailed (or e-mailed) questionnaires might have high nonresponse rates. Telephone interviews exclude those without phones and may mean repeated calling for those who are not home. Face-to-face interviews might be more costly than your funding will allow. There might also be some response bias: Some students might be hesitant about criticizing the faculty (while others might be far too eager to do so).

3.90. **(a)** The treatment combinations are shown in the table on the right, and the design is diagrammed below. **(b)** Larger samples give more information; in particular, with large samples, we reduce the variability in the observed mean concentrations, so that we can have more confidence that the differences we might observe are due to the treatment applied rather than random fluctuation.

		Factor B Administration method		
		Injection	Skin patch	IV drip
Factor A Dosage	5 mg	Treatment 1	Treatment 2	Treatment 3
	10 mg	Treatment 4	Treatment 5	Treatment 6



3.91. **(a)** The factors are storage method (three levels: fresh, room temperature for one month, refrigerated for one month) and preparation method (two levels: cooked immediately, or after one hour). There are therefore six treatments (summarized in the table on the right). The response variables are the tasters' color and flavor ratings. **(b)** Randomly allocate n potatoes to each of the six groups, then compare ratings. (Diagram not shown.) **(c)** For each taster, randomly choose the order in which the fries are tasted.

	Cooked immediately	Wait one hour
Fresh	Treatment 1	Treatment 2
Stored	Treatment 3	Treatment 4
Refrigerated	Treatment 5	Treatment 6

3.92. Use a block design: Separate men and women, and randomly allocate each gender among the six treatments.

3.93. Voluntary response samples are affected by strong feelings, promotional campaigns, and possibly multiple votes from one individual. Furthermore, online polls are also likely to overrepresent the opinions of the more technologically sophisticated (i.e., the "nerds"), which might account for tech hero Torvalds' advantage over figures like Mandela and Diana.

3.94. Parents who fail to return the consent form may be more likely to place less priority on education, and therefore may give their children less help with homework, and so forth. Including those children in the control group is likely to lower that group's score.

Note: *This is a generalization, to be sure: We are not saying that* every *such parent does not value education, only that the percentage of this group that highly values education will almost certainly be lower than that percentage of the parents who return the form.*

3.95. The latter method (CASI) will show a higher percentage of drug use, because respondents will generally be more comfortable (and more assured of anonymity) about revealing embarrassing or illegal behavior to a computer than to a person, so they will be more likely to be honest.

For answers to the EESEE Case Studies (exercises 96–99), see the instructor's version of EESEE.

Chapter 4 Solutions

4.1. Long trials of this experiment often approach 40% heads. One theory attributes this surprising result to a "bottle-cap effect" due to an unequal rim on the penny. We don't know. But a teaching assistant claims to have spent a profitable evening at a party betting on spinning coins after learning of the effect.

4.2. Note that in this experiment, factors other than the nickel's characteristics might affect the outcome. For example, if the surface used is not quite level, there will be a tendency for the penny to fall in the "downhill" direction.

4.3. The table on the right shows information from `www.mms.com` as of this writing. The exercise specified M&M's Milk Chocolate Candies, but based on these numbers, results will be similar for other popular varieties. Of course, answers will vary, but students who take reasonably large samples should get results close to the numbers in this table. (For example, samples of size 50 will almost always be within ±12%, while size 75 should give results within ±10%.)

M&M's variety	Blue %
Milk Chocolate	24%
Peanut	23%
Almond	20%
Peanut Butter	20%
Crispy	17%

4.4. In the long run, of a large number of hands of five cards, the fraction containing two pairs will be about 1/21. It does not mean that exactly one out of 21 hands contains two pairs; that would mean, for example, that if you've been dealt 20 hands without two pairs, that you could count on the next hand having two pairs.

4.5. **(a)** Most answers will be between 35% and 65%. **(b)** Based on 10,000 simulated trials—more than students are expected to do—there is about an 80% chance of having a longest run of four or more (i.e., either making or missing four shots in a row), a 54% chance of getting five or more, a 31% chance of getting six or more, and a 16% chance of getting seven or more. The average ("expected") longest run length is about six.

4.6. **(a)** – **(c)** Results will vary, but after n tosses, the distribution of the proportion $\hat{p}$ is approximately normal with mean 0.5 and standard deviation $1/(2\sqrt{n})$, while the distribution of the count of heads is approximately normal with mean $0.5n$ and standard deviation $\sqrt{n}/2$, so using the 68–95–99.7 rule, we have the results shown in the table on the right. Note that the range for $\hat{p}$ gets narrower, while the range for the count gets wider.

n	99.7% Range for $\hat{p}$	99.7% Range for count
40	0.5 ± 0.237	20 ± 9.5
120	0.5 ± 0.137	60 ± 16.4
240	0.5 ± 0.097	120 ± 23.2
480	0.5 ± 0.068	240 ± 32.9

4.7. The true probability (assuming perfectly fair dice) is $1 - \left(\frac{5}{6}\right)^4 \doteq 0.5177$, so students should conclude that the probability is "quite close to 0.5."

4.8. **(a)** The distribution of this proportion is roughly normal with mean 0.3 and standard deviation 0.1, so the 25 results can be quite variable. A stemplot for one set of 25 trials is shown. **(b)** With 200 tosses, the standard deviation drops to about 0.032, so the stemplot should be much less spread out. (Note the different "scales" in these two stemplots: The first ranges from 0.10 to 0.55, and the second from 0.235 to 0.355.)

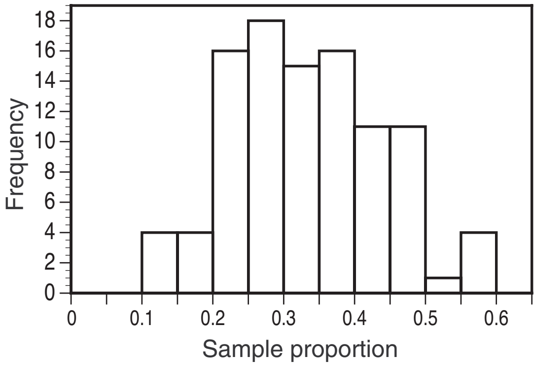
20 tosses

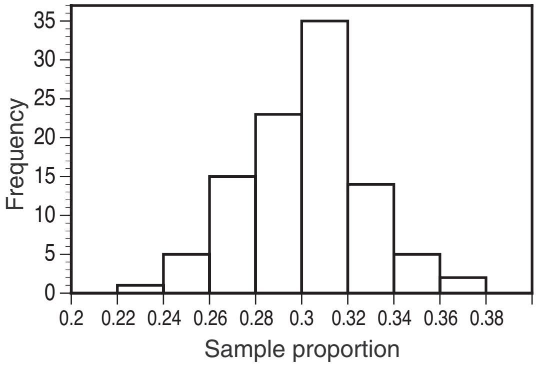
1	0
1	5
2	0
2	5555
3	0000000
3	5555
4	00000
4	5
5	
5	5

200 tosses

23	5
24	
25	
26	00
27	55
28	055
29	0055
30	05
31	0005
32	00
33	55
34	
35	055

4.9. **(a)** One such histogram is shown below on the left. It should appear approximately normal with center near 0.3 and standard deviation about 0.1, so it will typically range from about 0.1 to 0.5. **(b)** Now the histogram (sample below, right) should appear to be roughly normal with mean 0.3 and standard deviation about 0.026, spread from about 0.22 to 0.38. **(c)** Both distributions are roughly normal with center near 0.3, but as expected, the larger sample leads to much less variability.





4.10. **(a)** $S = \{\text{F, M}\}$ or {female, male}. **(b)** $S = \{6, 7, \ldots, 20\}$. **(c)** S = {All numbers between 2.5 and 6 l/min}. **(d)** S = {All numbers (or all whole numbers) between __ and __ bpm}. (Choices of upper and lower limits will vary.)

4.11. **(a)** S = {right, left}. **(b)** S = {All numbers between __ and __ cm}. (Choices of upper and lower limits will vary.) **(c)** S = {all numbers greater than or equal to 0}, or $S = \{0, 0.01, 0.02, 0.03, \ldots\}$. **(d)** S = {all numbers between 0 and 1440}. (There are 1440 minutes in one day, so this is the *largest* upper limit we could choose; many students will likely give a smaller upper limit.)

4.12. **(a)** The table on the right illustrates the 16 possible pair combinations in the sample space. **(b)** Add 1 to each pair-total on the right: $S = \{3, 4, 5, 6, 7, 8, 9\}$.

	⚀	⚁	⚂	⚃
⚀	⚀ ⚀	⚀ ⚁	⚀ ⚂	⚀ ⚃
⚁	⚁ ⚀	⚁ ⚁	⚁ ⚂	⚁ ⚃
⚂	⚂ ⚀	⚂ ⚁	⚂ ⚂	⚂ ⚃
⚃	⚃ ⚀	⚃ ⚁	⚃ ⚂	⚃ ⚃

4.13. **(a)** The given probabilities have sum 0.96, so $P(\text{type AB}) = 0.04$. **(b)** $P(\text{type O or B}) = 0.45 + 0.11 = 0.56$.

4.14. $P(\text{both are type O}) = (0.45)(0.35) = 0.1575$. $P(\text{both are the same type}) = (0.45)(0.35) + (0.40)(0.27) + (0.11)(0.26) + (0.04)(0.12) = 0.2989$.

4.15. **(a)** Legitimate. **(b)** Not legitimate: The total is more than 1. **(c)** Legitimate.

4.16. **(a)** The given probabilities have sum 0.89, so $P(\text{other language}) = 0.11$. **(b)** $P(\text{not English}) = 1 - 0.59 = 0.41$. (Or, add the other three probabilities.)

4.17. **(a)** The given probabilities have sum 0.72, so this probability must be 0.28. **(b)** $P(\text{at least a high school education}) = 1 - P(\text{has not finished HS}) = 1 - 0.12 = 0.88$. (Or, add the other three probabilities.)

4.18. **(a)** The given probabilities have sum 0.81, so $P(\text{other topic}) = 0.19$. **(b)** $P(\text{adult or scam}) = 0.145 + 0.142 = 0.287$.

Note: *An underlying assumption here is that each piece of spam falls into exactly one category.*

4.19. If $P(\text{Red Sox win}) = P(\text{Angels win}) = x$, then $P(\text{Athletics win}) = P(\text{White Sox win}) = \frac{1}{3}x$, and $0.6 + x + x + \frac{1}{3}x + \frac{1}{3}x = 1$. Therefore $\frac{8}{3}x = 0.4$, so $x = 0.15$. Thus the whole set of probabilities is

Team	Yankees	Red Sox	Angels	Athletics	White Sox
Probability	0.6	0.15	0.15	0.05	0.05

4.20. See also the solution to Exercise 4.12. As all faces are equally likely and the dice are independent, each of the 16 possible pairings is equally likely, so (for example) the probability of a total of 5 is 3/16, because 3 pairings add to 4 (and then we add 1). The complete set of probabilities is shown in the table.

Total	Probability
3	1/16
4	2/16
5	3/16
6	4/16
7	3/16
8	2/16
9	1/16

4.21. The probabilities of 2, 3, 4 and 5 are unchanged (1/6), so P(⚀ or ⚅) must still be 1/3. If P(⚅) = 0.2, then P(⚀) $= \frac{1}{3} - 0.2 = 0.1\overline{3}$ (or $\frac{2}{15}$).

Face	⚀	⚁	⚂	⚃	⚄	⚅
Probability	$0.1\overline{3}$	1/6	1/6	1/6	1/6	0.2

4.22. **(a)** It is legitimate because every person must fall into exactly one category, and the probabilities add to 1. **(b)** $P(A) = 0.125 = 0.000 + 0.003 + 0.060 + 0.062$. **(c)** B^c is the event "the person chosen is not white." $P(B^c) = 1 - P(B) = 1 - (0.060 + 0.691) = 0.249$. **(d)** $P(A^c \text{ and } B) = 0.691$ is the probability that a randomly chosen American is a non-Hispanic white.

4.23. For example, the probability for O-positive blood is $(0.45)(0.84) = 0.378$, and for O-negative it is $(0.45)(0.16) = 0.072$.

Blood type	O+	O–	A+	A–	B+	B–	AB+	AB–
Probability	0.3780	0.0720	0.3360	0.0640	0.0924	0.0176	0.0336	0.0064

4.24. We found in Exercise 4.22 that $P(A) = 0.125$ and that $P(B) = 0.751$. (We actually computed $P(B^c) = 0.249$.) Because $P(A)P(B) \doteq 0.094$ is not equal to $P(A \text{ and } B) = 0.060$ (from the table in Exercise 4.22), A and B are *not* independent.

4.25. **(a)** All are equally likely; the probability is 1/38. **(b)** Because 18 slots are red, the probability of a red is $P(\text{red}) = \frac{18}{38} \doteq 0.474$. **(c)** There are 12 winning slots, so $P(\text{win a column bet}) = \frac{12}{38} \doteq 0.316$.

4.26. **(a)** There are six arrangements of the digits 1, 2, and 3 (123, 132, 213, 231, 312, 321), so that $P(\text{win}) = \frac{6}{1000} = 0.006$. **(b)** With the digits 1, 1, and 2, there are only three distinct arrangements (112, 121, 211), so $P(\text{win}) = \frac{3}{1000} = 0.003$.

4.27. **(a)** There are $10^4 = 10{,}000$ possible PINs (0000 through 9999).* **(b)** The probability that a PIN has *no* 0s is 0.9^4 (because there are 9^4 PINs that can be made from the nine nonzero digits), so the probability of at least one 0 is $1 - 0.9^4 = 0.3439$.

*If we assume that PINs cannot have leading 0s, then there are only 9000 possible codes (1000–9999), and the probability of at least one 0 is $1 - \frac{9^4}{9000} = 0.271$.

4.28. $P(\text{none are O-negative}) = (1 - 0.07)^{10} \doteq 0.4840$, so $P(\text{at least one is O-negative}) \doteq 1 - 0.4840 = 0.5160$.

4.29. If we assume that each site is independent of the others (and that they can be considered as a random sample from the collection of sites referenced in scientific journals), then $P(\text{all seven are still good}) = 0.87^7 \doteq 0.3773$.

4.30. **(a)** About 0.33: $P(\text{no calls reach a live person}) \doteq (1 - 0.2)^5 = 0.8^5 = 0.32768$.
(b) About 0.66: $P(\text{no NY calls reach a live person}) \doteq 0.92^5 \doteq 0.65908$.

4.31. This computation would only be correct if the events "a randomly selected person is at least 75" and "a randomly selected person is a woman" were independent. This is likely not true; in particular, as women have a greater life expectancy than men, this fraction is probably greater than 3%.

4.32. As $P(\text{R}) = \frac{2}{6}$ and $P(\text{G}) = \frac{4}{6}$, and successive rolls are independent, the respective probabilities are $\left(\frac{2}{6}\right)^4\left(\frac{4}{6}\right) = \frac{2}{243} \doteq 0.00823$, $\left(\frac{2}{6}\right)^4\left(\frac{4}{6}\right)^2 = \frac{4}{729} \doteq 0.00549$, and $\left(\frac{2}{6}\right)^5\left(\frac{4}{6}\right) = \frac{2}{729} \doteq 0.00274$.

4.33. **(a)** $(0.65)^3 \doteq 0.2746$ (under the random walk theory). **(b)** 0.35 (because performance in separate years is independent). **(c)** $(0.65)^2 + (0.35)^2 = 0.545$.

4.34. For any event A, along with its complement A^c, we have $P(S) = P(A \text{ or } A^c)$ because "$A \text{ or } A^c$" includes all possible outcomes (that is, it is the entire sample space S). By Rule 2, $P(S) = 1$, and by Rule 3, $P(A \text{ or } A^c) = P(A) + P(A^c)$, because A and A^c are disjoint. Therefore, $P(A) + P(A^c) = 1$, from which Rule 4 follows.

4.35. Note that $A = (A \text{ and } B)$ or $(A \text{ and } B^c)$, and the events $(A \text{ and } B)$ and $(A \text{ and } B^c)$ are disjoint, so Rule 3 says that

$$P(A) = P\big((A \text{ and } B) \text{ or } (A \text{ and } B^c)\big) = P(A \text{ and } B) + P(A \text{ and } B^c).$$

If $P(A \text{ and } B) = P(A)P(B)$, then we have

$$P(A \text{ and } B^c) = P(A) - P(A)P(B) = P(A)(1 - P(B)),$$

which equals $P(A)P(B^c)$ by the complement rule.

4.36. **(a)** Hannah and Jacob's children can have alleles AA, BB, or AB, so they can have blood type A, B, or AB. (The table on the right shows the possible combinations.) **(b)** Either note that the four combinations in the table are equally likely, or compute:

	A	B
A	AA	AB
B	AB	BB

$$\begin{aligned} P(\text{type A}) &= P(\text{A from Hannah and A from Jacob}) = P(\text{A}_H)P(\text{A}_J) = 0.5^2 = 0.25 \\ P(\text{type B}) &= P(\text{B from Hannah and B from Jacob}) = P(\text{B}_H)P(\text{B}_J) = 0.5^2 = 0.25 \\ P(\text{type AB}) &= P(\text{A}_H)P(\text{B}_J) + P(\text{B}_H)P(\text{A}_J) = 2 \cdot 0.25 = 0.5. \end{aligned}$$

4.37. **(a)** Nancy and David's children can have alleles BB, BO, or OO, so they can have blood type B or O. (The table on the right shows the possible combinations.) **(b)** Either note that the four combinations in the table are equally likely, or compute $P(\text{type O}) = P(\text{O from Nancy and O from David}) = 0.5^2 = 0.25$, and $P(\text{type B}) = 1 - P(\text{type O}) = 0.75$.

	B	O
B	BB	BO
O	BO	OO

4.38. Any child of Jennifer and José has a 50% chance of being type A (alleles AA or AO), and each child inherits alleles independently of other children, so $P(\text{both are type A}) = 0.5^2 = 0.25$. For one child, we have $P(\text{type A}) = 0.5$ and $P(\text{type AB}) = P(\text{type B}) = 0.25$, so that $P(\text{both have same type}) = 0.5^2 + 0.25^2 + 0.25^2 = 0.375 = \frac{3}{8}$.

	A	O
A	AA	AO
B	AB	BO

4.39. **(a)** Any child of Jasmine and Joshua has an equal (1/4) chance of having blood type AB, A, B, or O (see the allele combinations in the table). Therefore, $P(\text{type O}) = 0.25$. **(b)** $P(\text{all three have type O}) = 0.25^3 = 0.015625 = \frac{1}{64}$. $P(\text{first has type O, next two do not}) = 0.25 \cdot 0.75^2 = 0.140625 = \frac{9}{64}$.

	A	O
B	AB	BO
O	AO	OO

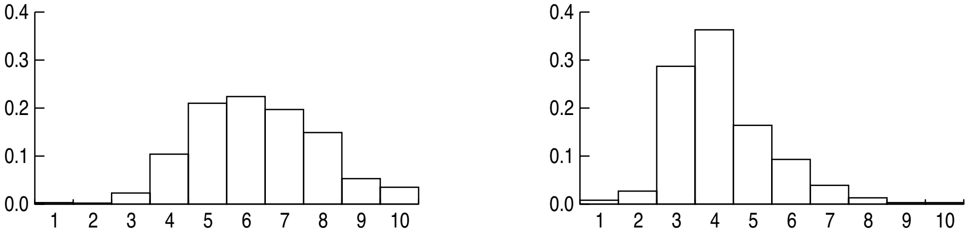
4.40. The RPE rating in (b) is discrete, because the description says that it "ranges in whole-number steps from 6 ... to 20." VO2 and maximum heart rate are continuous.

Note: *In practice, heart rate is given as a whole number, for example, "60 bpm" rather than "59.63 bpm." For this reason, the argument could be made that this is discrete. In fact, since we cannot measure anything with infinite precision, one could say that, in practice,* everything *is discrete. Be aware that this distinction can be somewhat unclear to students; you (the instructor) will need to decide how to handle this.*

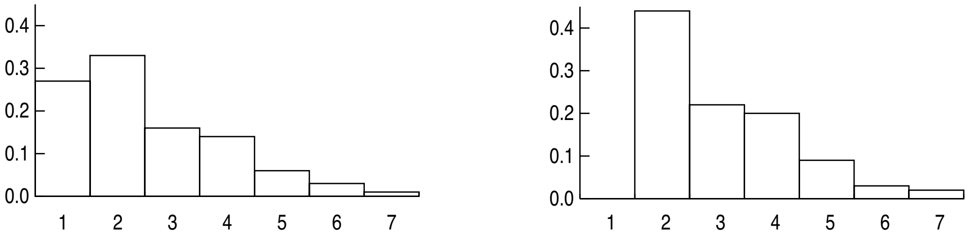
4.41. **(a)** "At least one nonword error" is the event "$X \geq 1$" (or "$X > 0$"). $P(X \geq 1) = 1 - P(X = 0) = 0.9$. **(b)** "$X \leq 2$" is "no more than two nonword errors," or "fewer than three nonword errors." $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.1 + 0.2 + 0.3 = 0.6$. $P(X < 2) = P(X = 0) + P(X = 1) = 0.1 + 0.2 = 0.3$.

4.42. About 0.017: $P(Y > 300) = P\left(\frac{Y-266}{16} > \frac{300-266}{16}\right) = P(Z > 2.125)$. Software gives 0.0168; Table A gives 0.0166 for −2.13 and 0.0170 for −2.12 (so the average is again 0.0168).

4.43. The two histograms are shown below. Rented housing typically has fewer rooms, and its distribution is considerably more skewed than the owned-housing distribution.



4.44. The two histograms are shown below. The most obvious difference is that a "family" must have at least two people. Otherwise, the family-size distribution has slightly larger probabilities for 2, 3, or 4, while for large family/household sizes, the differences between the distributions are small.



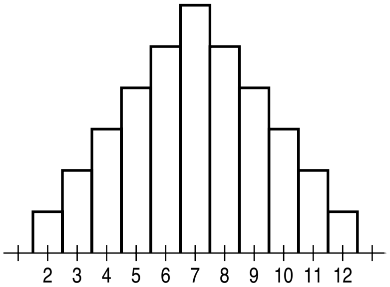
4.45. **(a)** "The unit has 5 or more rooms" is "$X \geq 5$" or "$X > 4$." $P(X \geq 5) = 0.210 + 0.224 + 0.197 + 0.149 + 0.053 + 0.035 = 0.868$. **(b)** "$X > 5$" is "the unit has more than 5 (or at least 6) rooms." $P(X > 5) = 0.658$. **(c)** With discrete distributions, one must pay attention to whether or not the endpoints should be included in the probability computation. (That is, pay attention to whether you have "greater than" or "greater than or equal to.")

4.46. **(a)** "More than one person lives in this household" is "$Y > 1$" or "$Y \geq 2$." $P(Y > 1) = 1 - P(Y = 1) = 1 - 0.27 = 0.73$. **(b)** $P(2 < Y \leq 4) = P(Y = 3 \text{ or } Y = 4) = 0.16 + 0.14 = 0.30$. **(c)** $P(Y \neq 2) = 1 - P(Y = 2) = 1 - 0.33 = 0.67$.

4.47. **(a)** $P(T = 2) = P(\text{your friend breaks the chain}) = 1 - 0.37 = 0.63$. $P(T = 3) = P(\text{your friend's friend breaks the chain}) = (0.37)(0.63) = 0.2331$. **(b)** $P(T \leq 4)$ is the probability that the message stops before reaching the fifth person—that is, one of the first three people to receive it breaks the chain. $P(T \leq 4) = 0.63 + (0.37)(0.63) + (0.37)(0.37)(0.63) \doteq 0.9493$.

4.48. **(a)** The pairs are given below. We must assume that we can distinguish between, e.g., "(1,2)" and "(2,1)"; otherwise the outcomes are not equally likely. **(b)** Each pair has probability 1/36. **(c)** The value of X is given below each pair. For the distribution (given below), we see (for example) that there are four pairs that add to 5, so $P(X = 5) = \frac{4}{36}$. Histogram below, right. **(d)** $P(7 \text{ or } 11) = \frac{6}{36} + \frac{2}{36} = \frac{8}{36} = \frac{2}{9}$. **(e)** $P(\text{not } 7) = 1 - \frac{6}{36} = \frac{5}{6}$.

(1,1) 2	(1,2) 3	(1,3) 4	(1,4) 5	(1,5) 6	(1,6) 7
(2,1) 3	(2,2) 4	(2,3) 5	(2,4) 6	(2,5) 7	(2,6) 8
(3,1) 4	(3,2) 5	(3,3) 6	(3,4) 7	(3,5) 8	(3,6) 9
(4,1) 5	(4,2) 6	(4,3) 7	(4,4) 8	(4,5) 9	(4,6) 10
(5,1) 6	(5,2) 7	(5,3) 8	(5,4) 9	(5,5) 10	(5,6) 11
(6,1) 7	(6,2) 8	(6,3) 9	(6,4) 10	(6,5) 11	(6,6) 12



Value of X	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

4.49. The possible values of Y are 1,2,3,. . . ,12, each with probability 1/12. Aside from drawing a diagram showing all the possible combinations, one can reason that the first (regular) die is equally likely to show any number from 1 through 6. Half of the time, the second roll shows 0, and the rest of the time it shows 6. Each possible outcome therefore has probability $\frac{1}{6} \cdot \frac{1}{2}$.

4.50. The table on the right shows the additional columns to add to the table given in the solution to Exercise 4.48. There are 48 possible (equally-likely) combinations.

(1,7) 8	(1,8) 9
(2,7) 9	(2,8) 10
(3,7) 10	(3,8) 11
(4,7) 11	(4,8) 12
(5,7) 12	(5,8) 13
(6,7) 13	(6,8) 14

Value of X	2	3	4	5	6	7	8	9	10	11	12	13	14
Probability	$\frac{1}{48}$	$\frac{2}{48}$	$\frac{3}{48}$	$\frac{4}{48}$	$\frac{5}{48}$	$\frac{6}{48}$	$\frac{6}{48}$	$\frac{6}{48}$	$\frac{5}{48}$	$\frac{4}{48}$	$\frac{3}{48}$	$\frac{2}{48}$	$\frac{1}{48}$

4.51. **(a)** W can be 0, 1, 2, or 3. **(b)** See the top two lines of the table below. **(c)** The distribution is given in the bottom two lines of the table. For example, $P(W = 0) = (0.73)(0.73)(0.73) \doteq 0.3890$, and in the same way, $P(W = 3) = 0.27^3 \doteq 0.1597$. For $P(W = 1)$, note that each of the three arrangements that give $W = 1$ have probability $(0.73)(0.73)(0.27) = 0.143883$, so $P(W = 1) = 3(0.143883) \doteq 0.4316$. Similarly, $P(W = 2) = 3(0.73)(0.27)(0.27) \doteq 0.1597$.

Arrangement	DDD	DDF	DFD	FDD	FFD	FDF	DFF	FFF
Probability	0.3890	0.1439	0.1439	0.1439	0.0532	0.0532	0.0532	0.0197
Value of W	0	1			2			3
Probability	0.3890	0.4316			0.1597			0.0197

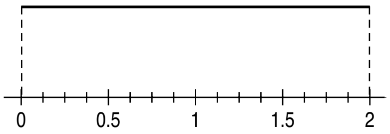
4.52. Let “S” mean that a student supports funding and “O” mean that the student opposes funding. **(a)** $P(\text{SSO}) = (0.6)(0.6)(0.4) = 0.144$. **(b)** See the top two lines of the table below. **(c)** The distribution is given in the bottom two lines of the table. For example, $P(X = 0) = (0.6)(0.6)(0.6) = 0.216$, and in the same way, $P(X = 3) = 0.4^3 = 0.064$. For $P(X = 1)$, note that each of the three arrangements that give $X = 1$ have probability 0.144, so $P(X = 1) = 3(0.144) = 0.432$. Similarly, $P(X = 2) = 3(0.6)(0.4)(0.4) = 0.288$. **(d)** A majority means $X \geq 2$; $P(X \geq 2) = 0.288 + 0.064 = 0.352$.

Arrangement	SSS	SSO	SOS	OSS	OOS	OSO	SOO	OOO
Probability	0.216	0.144	0.144	0.144	0.096	0.096	0.096	0.064
Value of X	0	1			2			3
Probability	0.216	0.432			0.288			0.064

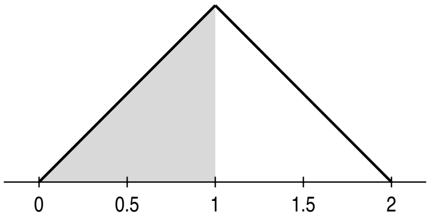
4.53. **(a)** $P(X < 0.5) = 0.5$. **(b)** $P(X \leq 0.5) = 0.5$. **(c)** For continuous random variables, “equal to” has no effect on the probability; that is, $P(X = c) = 0$ for any value of c.

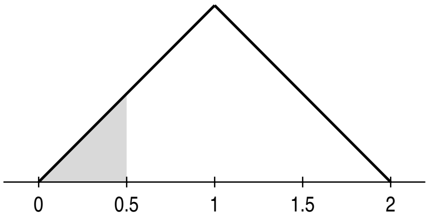
4.54. **(a)** $P(X \geq 0.27) = 0.73$. **(b)** $P(X = 0.27) = 0$. **(c)** $P(0.27 < X < 1.27) = P(0.27 < X < 1) = 0.73$. **(d)** $P(0.1 \leq X \leq 0.2 \text{ or } 0.8 \leq X \leq 0.9) = 0.1 + 0.1 = 0.2$. **(e)** $P(\text{not } [0.3 \leq X \leq 0.8]) = 1 - P(0.3 \leq X \leq 0.8) = 1 - 0.5 = 0.5$.

4.55. **(a)** The height should be $\frac{1}{2}$, since the area under the curve must be 1. The density curve is at the right. **(b)** $P(Y \leq 1) = \frac{1}{2}$. **(c)** $P(0.5 < Y < 1.3) = 0.4$. **(d)** $P(Y \geq 0.8) = 0.6$.



4.56. **(a)** The area of a triangle is $\frac{1}{2}bh = \frac{1}{2}(2)(1) = 1$. **(b)** $P(Y < 1) = 0.5$.
(c) $P(Y < 0.5) = 0.125$.





4.57. $P(8 \le \bar{x} \le 10) = P\left(\frac{8-9}{0.075} \le \frac{\bar{x}-9}{0.075} \le \frac{10-9}{0.075}\right) = P(-13.3 \le Z \le 13.3)$. This probability is essentially 1; $\bar{x}$ will almost certainly estimate μ within ±1 (in fact, it will almost certainly be much closer than this).

4.58. **(a)** $P(0.52 \le \hat{p} \le 0.60) = P\left(\frac{0.52-0.56}{0.019} \le \frac{\hat{p}-0.56}{0.019} \le \frac{0.60-0.56}{0.019}\right) = P(-2.11 \le Z \le 2.11) = 0.9826 - 0.0174 = 0.9652$. **(b)** $P(\hat{p} \ge 0.72) = P\left(\frac{\hat{p}-0.56}{0.019} \ge \frac{0.72-0.56}{0.019}\right) = P(Z \ge 8.42)$; this is basically 0.

4.59. The average grade is $\mu = (0)(0.01) + (1)(0.05) + (2)(0.30) + (3)(0.43) + (4)(0.21) = 2.78$.

4.60. The mean number of nonword errors is $(0)(0.1)+(1)(0.2)+(2)(0.3)+(3)(0.3)+(4)(0.1) = 2.1$, and the mean number of word errors is $(0)(0.4) + (1)(0.3) + (2)(0.2) + (3)(0.1) = 1$.

4.61. For owner-occupied units, the mean is

$$(1)(0.003) + (2)(0.002) + (3)(0.023) + (4)(0.104) + (5)(0.210) +$$
$$(6)(0.224) + (7)(0.197) + (8)(0.149) + (9)(0.053) + (10)(0.035) = 6.284 \text{ rooms.}$$

For rented units, the mean is

$$(1)(0.008) + (2)(0.027) + (3)(0.287) + (4)(0.363) + (5)(0.164) +$$
$$(6)(0.093) + (7)(0.039) + (8)(0.013) + (9)(0.003) + (10)(0.003) = 4.187 \text{ rooms.}$$

This agrees with the observation in Exercise 4.43 that rented housing typically has fewer rooms.

4.62. **(a)** The mean of Y is $\mu_Y = 1$—the obvious balance point of the triangle. **(b)** Both X_1 and X_2 have mean $\mu_{X_1} = \mu_{X_2} = 0.5$, and $\mu_Y = \mu_{X_1} + \mu_{X_2}$.

4.63. The owned-unit distribution seems to be more spread out, and the two standard deviations confirm this impression: $\sigma_o \doteq 1.6399$ and $\sigma_r \doteq 1.3077$ rooms. The full details of these computations are not shown, but for example the variance of the number of rooms in an owner-occupied unit looks like this:

$$\sigma_o^2 = (1 - \mu_o)^2(0.003) + (2 - \mu_o)^2(0.008) + \cdots + (10 - \mu_o)^2(0.035) \doteq 2.6893,$$

and in the same way, $\sigma_r^2 \doteq 1.7100$. Taking square roots completes the task.

4.64. Let N and W be nonword and word error counts. In Exercise 4.60, we found $\mu_N = 2.1$ errors and $\mu_W = 1$ error. The variances of these distributions are

$\sigma_N^2 = 1.29$ and $\sigma_W^2 = 1$, so the standard deviations are $\sigma_N = 1.1358$ and $\sigma_W = 1$ errors. The mean total error count is $\mu_N + \mu_W = 3.1$ errors for both cases. **(a)** If error counts are independent, $\sigma_{N+W}^2 = \sigma_N^2 + \sigma_W^2 = 2.29$ and $\sigma_{N+W} = 1.5133$ errors. (Note we add the *variances,* not the standard deviations.) **(b)** With $\rho = 0.5$, $\sigma_{N+W}^2 = \sigma_N^2 + \sigma_W^2 + 2\rho\sigma_N\sigma_W = 2.29 + 1.1358 = 3.4258$, and $\sigma_{N+W} = 1.8509$ errors.

Value	\$562.50	\$975	\$1690
Probability	0.25	0.5	0.25

4.65. **(a)** Over two days, the value of this stock changes by a factor of $1.3^2 = 1.69$, $0.75^2 = 0.5625$, or $1.3 \times 0.75 = 0.75 \times 1.3 = 0.975$. Therefore, the possible values and probabilities are as given in the table on the right. The stock gains value only 25% of the time. **(b)** The mean value is \$1050.625.

Note: *Each of the four possible pairs—fall/fall, fall/rise, rise/fall, and rise/rise—are equally likely.*

4.66. We are trying to estimate the probabilities in the distribution of first digits among the invoices. If we have inaccurate estimates, then small differences between the observed probabilities and those of Benford's law can be attributed to that inaccuracy; we would need to see very large discrepancies before we could conclude that the invoices were faked. We can get better (more accurate) estimates by taking more invoices; this would mean that it would be easier to detect a set of faked invoices.

4.67. **(a)** The mean for one coin is $\mu_1 = (0)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{2}\right) = 0.5$ and the variance is $\sigma_1^2 = (0 - 0.5)^2\left(\frac{1}{2}\right) + (1 - 0.5)^2\left(\frac{1}{2}\right) = 0.25$, so the standard deviation is $\sigma_1 = 0.5$. **(b)** Multiply μ_1 and σ_1^2 by 4: $\mu_4 = 4\mu_1 = 2$ and $\sigma_4^2 = 4\sigma_1^4 = 1$, so $\sigma_4 = 1$. **(c)** The computations (not shown here) are more tedious, but the results are the same. Note that because of the symmetry of the distribution, we do not need to compute the mean to see that $\mu_4 = 2$; this is the obvious balance point of the probability histogram in Figure 4.7.

4.68. If S and R are the results from the six- and eight-sided dice (respectively), then $\mu_S = 3.5$ and $\mu_E = 4.5$ by symmetry (this can be confirmed by computation). Therefore, $\mu_{S+E} = 8$.

For the variances,

$$\begin{aligned}\sigma_S^2 = (1-3.5)^2\left(\tfrac{1}{6}\right) + (2-3.5)^2\left(\tfrac{1}{6}\right) + (3-3.5)^2\left(\tfrac{1}{6}\right) + \\ (4-3.5)^2\left(\tfrac{1}{6}\right) + (5-3.5)^2\left(\tfrac{1}{6}\right) + (6-3.5)^2\left(\tfrac{1}{6}\right) \quad = \tfrac{35}{12} \doteq 2.9167,\end{aligned}$$

and by similar computations, $\sigma_E^2 = 5.25$. Therefore, $\sigma_{S+E}^2 = 8\frac{1}{6} \doteq 8.1667$, and the standard deviation is $\sigma_{S+E} \doteq 2.8577$.

4.69. With R as the rod length and B_1 and B_2 the bearing lengths, we have $\mu_{B_1+R+B_2} = 10 + 2 \cdot 2 = 14$ cm and $\sigma_{B_1+R+B_2} = \sqrt{0.005^2 + 2 \cdot 0.001^2} \doteq 0.005196$ cm.

4.70. **(a)** $d_1 = 2\sigma_R = 0.010$ cm and $d_2 = 2\sigma_B = 0.002$ cm. **(b)** The natural tolerance of the assembled parts is $2\sigma_{B_1+R+B_2} \doteq 0.01039$ cm.

4.71. **(a)** Not independent: Knowing the total X of the first two cards tells us something about the total Y for three cards. **(b)** Independent: Separate rolls of the dice should be independent.

4.72. **(a)** Independent: Weather conditions a year apart should be independent. **(b)** Not independent: Weather patterns tend to persist for several days; today's weather tells us something about tomorrow's. **(c)** Not independent: The two locations are very close together, and would likely have similar weather conditions.

4.73. **(a)** The total mean is $11 + 20 = 31$ seconds. **(b)** No: Changing the standard deviations does not affect the means. Reducing variation is useful, though, because it increases the consistency of the process; for example, we could more reliably predict how many units could be assembled during a shift. **(c)** The mean does not change, because correlation does not affect the total mean.

4.74. Divide the given values by 2.54: $\mu \doteq 69.6063$ in and $\sigma \doteq 2.8346$ in.

4.75. If the two times are independent, the total standard deviation is

$$\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{pos}} + \sigma^2_{\text{att}}} = \sqrt{2^2 + 8^2} = \sqrt{68} \doteq 8.2462 \text{ seconds.}$$

With correlation 0.3, we have

$$\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{pos}} + \sigma^2_{\text{att}} + 2\rho\sigma_{\text{pos}}\sigma_{\text{att}}} = \sqrt{68 + 2(0.3)(2)(8)} = \sqrt{77.6} \doteq 8.8091 \text{ seconds.}$$

With a positive correlation, if the first task takes longer, the second often does, too, while a shorter time for the first task often means the second time is shorter. Consequently, there is an increased chance of extreme (long or short) times, increasing the spread of the distribution.

4.76. **(a)** $\mu_Y = \frac{1}{2}(\mu_{X_1} + \mu_{X_2}) = 100$ m, and $\sigma_Y = \frac{1}{2}\sigma_{X_1+X_2} = \frac{1}{2}\sqrt{1.2^2 + 0.85^2} \doteq 0.7353$ m.

(b) $\mu_W = \frac{1}{3}\mu_{X_1} + \frac{2}{3}\mu_{X_2} = 100$ m, and $\sigma_W = \sqrt{\left(\frac{1}{3}\right)^2 \sigma^2_{X_1} + \left(\frac{2}{3}\right)^2 \sigma^2_{X_2}} \doteq 0.6936$ m.

4.77. With $\rho = 1$, we have

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\rho\sigma_X\sigma_Y = \sigma^2_X + \sigma^2_Y + 2\sigma_X\sigma_Y = (\sigma_X + \sigma_Y)^2,$$

and of course $\sigma_{X+Y} = \sqrt{(\sigma_X + \sigma_Y)^2} = \sigma_X + \sigma_Y$.

4.78. The mean of X is $(\mu - \sigma)(0.5) + (\mu + \sigma)(0.5) = \mu$, and the standard deviation is $\sqrt{(\mu - \sigma - \mu)^2(0.5) + (\mu + \sigma - \mu)^2(0.5)} = \sqrt{\sigma^2} = \sigma$.

4.79. 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.

4.80. The total loss T for 12 fires has mean $\mu_T = 12 \cdot \$250 = \$3000$, and standard deviation $\sigma_T = \sqrt{12 \cdot \$300^2} \doteq \$1039$. The average loss is $T/12$, so $\mu_{T/12} = \frac{1}{12}\mu_T = \$250$ and $\sigma_{T/12} = \frac{1}{12}\sigma_T = \$86.60$.

4.81. (a) Add up the given probabilities and subtract from 1; this gives P(man does not die in the next five years) $= 0.99749$. **(b)** The distribution of income (or loss) is given below. Multiplying each possible value by its probability gives the mean intake $\mu \doteq \$623.22$.

Age at death	21	22	23	24	25	Survives
Loss or income	−\$99,825	−\$99,650	−\$99,475	−\$99,300	−\$99,125	\$875
Probability	0.00039	0.00044	0.00051	0.00057	0.00060	0.99749

4.82. The mean μ of the company's "winnings" (premiums) and their "losses" (insurance claims) is positive. Even though the company will lose a large amount of money on a small number of policyholders who die, it will gain a small amount on the majority. The law of large numbers says that the average "winnings" minus "losses" should be close to μ, and overall the company will almost certainly show a profit.

4.83. With $R = 0.8W + 0.2Y$, we have $\mu_R = 0.8\mu_W + 0.2\mu_Y = 11.116\%$ and

$$\sigma_R = \sqrt{(0.8\sigma_W)^2 + (0.2\sigma_Y)^2 + 2\rho_{WY}(0.8\sigma_W)(0.2\sigma_Y)} \doteq 15.9291\%.$$

4.84. With $\rho_{WY} = 0$, the standard deviation drops to $\sqrt{(0.8\sigma_W)^2 + (0.2\sigma_Y)^2} \doteq 14.3131\%$. The mean is unaffected by the correlation.

4.85. With $R = 0.6W + 0.2X + 0.2Y$, we have $\mu_R = 0.6\mu_W + 0.2\mu_X + 0.2\mu_Y = 10.184\%$ and

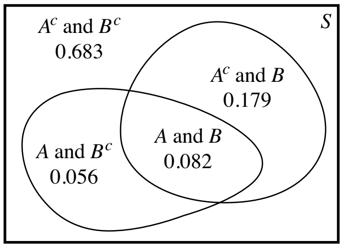
$$\begin{aligned}\sigma_R^2 &= (0.6\sigma_W)^2 + (0.2\sigma_X)^2 + (0.2\sigma_Y)^2 \\ &\quad + 2\rho_{WX}(0.6\sigma_W)(0.2\sigma_X) + 2\rho_{WY}(0.6\sigma_W)(0.2\sigma_Y) + 2\rho_{XY}(0.2\sigma_X)(0.2\sigma_Y) \\ &\doteq 152.3788\end{aligned}$$

so that $\sigma_R \doteq 12.3442\%$. The benefit of this diversification is reduced variability (σ_R is smaller than either σ_W and σ_Y), but the mean return is reduced because of the inclusion of the Bond Fund.

4.86. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 0.138 + 0.261 - 0.082 = 0.317$.

4.87. $P(A \mid B) = \dfrac{P(A \text{ and } B)}{P(B)} = \frac{0.082}{0.261} \doteq 0.3142$. If A and B were independent, then $P(A \mid B)$ would equal $P(A)$, and also $P(A \text{ and } B)$ would equal the product $P(A)P(B)$.

4.88. **(a)** $\{A \text{ and } B\}$: household is both prosperous and educated; $P(A \text{ and } B) = 0.082$ (given). **(b)** $\{A \text{ and } B^c\}$: household is prosperous but not educated; $P(A \text{ and } B^c) = P(A) - P(A \text{ and } B) = 0.056$. **(c)** $\{A^c \text{ and } B\}$: household is not prosperous but is educated; $P(A^c \text{ and } B) = P(B) - P(A \text{ and } B) = 0.179$. **(d)** $\{A^c \text{ and } B^c\}$: household is neither prosperous nor educated; $P(A^c \text{ and } B^c) = 0.683$ (so that the probabilities add to 1).



4.89. **(a)** "The vehicle is a car" $= A^c$; $P(A^c) = 1 - P(A) = 1 - 0.69 = 0.31$. **(b)** "The vehicle is an imported car" $= A^c \text{ and } B$. To find this probability, note that we have been given $P(B^c) = 0.78$ and $P(A \text{ and } B^c) = 0.55$. From this we can determine that $78\% - 55\% = 23\%$ of vehicles sold were domestic cars—that is, $P(A^c \text{ and } B^c) = 0.23$—so $P(A^c \text{ and } B) = P(A^c) - P(A^c \text{ and } B^c) = 0.31 - 0.23 = 0.08$.

	$P(A) = \mathbf{0.69}$	$P(A^c) = 0.31$
$P(B) = 0.22$	$P(A \text{ and } B) = 0.14$	$P(A^c \text{ and } B) = 0.08$
$P(B^c) = \mathbf{0.78}$	$P(A \text{ and } B^c) = \mathbf{0.55}$	$P(A^c \text{ and } B^c) = 0.23$

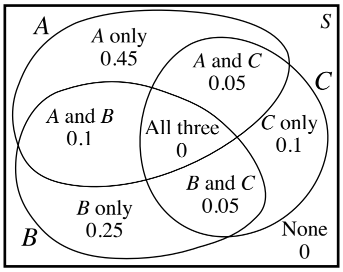
Note: *The table shown here summarizes all that we can determine from the given information (***bold***).*

4.90. Let A be the event "income $\geq$ \$1 million" and B be "income $\geq$ \$100,000." Then "A and B" is the same as A, so

$$P(A \mid B) = \frac{P(A)}{P(B)} = \frac{\frac{240{,}000}{129{,}075{,}000}}{\frac{10{,}855{,}000}{129{,}075{,}000}} = \frac{240{,}000}{10{,}855{,}000} \doteq 0.02211.$$

4.91. See also the solution to Exercise 4.89, especially the table of probabilities given there. **(a)** $P(A^c \mid B) = \frac{P(A^c \text{ and } B)}{P(B)} = \frac{0.08}{0.22} \doteq 0.3636$. **(b)** The events A^c and B are *not* independent; if they were, $P(A^c \mid B)$ would be the same as $P(A^c) = 0.31$.

4.92. To find the probabilities in this Venn diagram, begin with $P(A \text{ and } B \text{ and } C) = 0$ in the center of the diagram. Then each of the two-way intersections $P(A \text{ and } B)$, $P(A \text{ and } C)$, and $P(B \text{ and } C)$ go in the remainder of the overlapping areas; if $P(A \text{ and } B \text{ and } C)$ had been something other than 0, we would have subtracted this from each of the two-way intersection probabilities to find, for example, $P(A \text{ and } B \text{ and } C^c)$. Next, determine $P(A \text{ only})$ so that the total probability of the regions that make up the event A is 0.6. Finally, $P(\text{none}) = P(A^c \text{ and } B^c \text{ and } C^c) = 0$ because the total probability inside the three sets A, B, and C is 1.



4.93. We seek $P(\text{at least one offer}) = P(A \text{ or } B \text{ or } C)$; we can find this as $1 - P(\text{no offers}) = 1 - P(A^c \text{ and } B^c \text{ and } C^c)$. We see in the Venn diagram of Exercise 4.92 that this probability is 1.

4.94. This is $P(A \text{ and } B \text{ and } C^c)$. As was noted in Exercise 4.92, because $P(A \text{ and } B \text{ and } C) = 0$, this is the same as $P(A \text{ and } B) = 0.1$.

4.95. $P(B \mid C) = \dfrac{P(B \text{ and } C)}{P(C)} = \frac{0.05}{0.2} = 0.25$. $P(C \mid B) = \dfrac{P(B \text{ and } C)}{P(B)} = \frac{0.05}{0.4} = 0.125$.

4.96. Let W be the event "the person is a woman" and P be "the person earned a professional degree." **(a)** $P(W) = \frac{1119}{1944} \doteq 0.5756$. **(b)** $P(W \mid P) = \frac{39/1944}{83/1944} = \frac{39}{83} \doteq 0.4699$. **(c)** W and P are *not* independent; if they were, the two probabilities in (a) and (b) would be equal.

4.97. Let M be the event "the person is a man" and B be "the person earned a bachelor's degree." **(a)** $P(M) = \frac{825}{1944} \doteq 0.4244$. **(b)** $P(B \mid M) = \frac{559/1944}{825/1944} = \frac{559}{825} \doteq 0.6776$.
(c) $P(M \text{ and } B) = P(M)\,P(B \mid M) \doteq (0.4244)(0.6776) \doteq 0.2876$. This agrees with the directly computed probability: $P(M \text{ and } B) = \frac{559}{1944} \doteq 0.2876$.

4.98. Each unemployment rate is computed as shown on the right. (Alternatively, subtract the number employed from the number in the labor force, then divide that difference by the number in the labor force.) Because these rates (probabilities) are different, education level and being employed are not independent.

Did not finish HS	$1 - \frac{11{,}552}{12{,}623} \doteq 0.0848$
HS/no college	$1 - \frac{36{,}249}{38{,}210} \doteq 0.0513$
Some college	$1 - \frac{32{,}429}{33{,}928} \doteq 0.0442$
College graduate	$1 - \frac{39{,}250}{40{,}414} \doteq 0.0288$

4.99. (a) Add up the numbers in the first and second columns. We find that there are 186,210 thousand (that is, over 186 million) people aged 25 or older, of which 125,175 thousand are in the labor force, so $P(L) = \frac{125{,}175}{186{,}210} \doteq 0.6722$.
(b) $P(L \mid C) = \dfrac{P(L \text{ and } C)}{P(C)} = \frac{40{,}414}{51{,}568} \doteq 0.7837$. **(c)** L and C are *not* independent; if they were, the two probabilities in (a) and (b) would be equal.

4.100. For the first probability, add up the numbers in the third column. We find that there are 119,480 thousand (that is, over 119 million) employed people aged 25 or older. Therefore, $P(C \mid E) = \dfrac{P(C \text{ and } E)}{P(E)} = \frac{39{,}250}{119{,}480} \doteq 0.3285$.

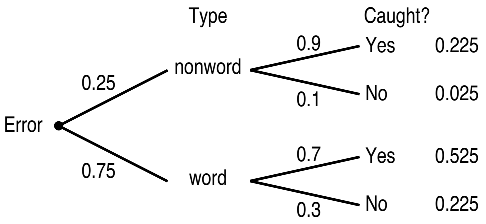
For the second probability, we use the total number of college graduates in the population: $P(E \mid C) = \dfrac{P(C \text{ and } E)}{P(C)} = \frac{39{,}250}{51{,}568} \doteq 0.7611$.

4.101. The population includes retired people who have left the labor force. Retired persons are more likely than other adults to have not completed high school; consequently, a relatively large number of retired persons fall in the "did not finish high school" category.

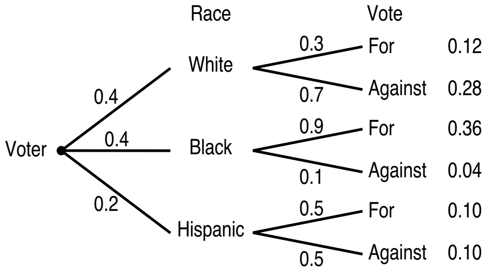
Note: *Details of this lurking variable can be found in the Current Population Survey annual report on "Educational Attainment in the United States." For 2003, this report says that among the 65-and-over population, about 29% have not completed high school, compared to about 15% of the under-65 group.*

4.102. By the multiplication rule, $P(E \text{ and } W) = P(E)\,P(W \mid E) = (0.15)(0.8) = 0.12$. Therefore, $P(E \mid W) = \dfrac{P(E \text{ and } W)}{P(W)} = \frac{0.12}{0.6} = 0.2$.

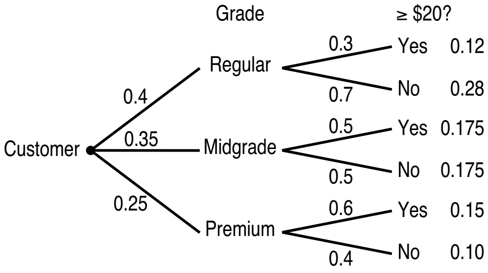
4.103. Tree diagram at right. The numbers on the right side of the tree are found by the multiplication rule; for example, P(“nonword error” and “caught”) $= P(N \text{ and } C) = P(N)\,P(C \mid N) = (0.25)(0.9) = 0.225$. A proofreader should catch about $0.225 + 0.525 = 0.75 = 75\%$ of all errors.



4.104. Tree diagram at right. The numbers on the right side of the tree are found by the multiplication rule; for example, P(“white” and “for”) $= P(W \text{ and } F) = P(W)\,P(F \mid W) = (0.4)(0.3) = 0.12$. The black candidate expects to get $12\% + 36\% + 10\% = 58\%$ of the vote.



4.105. Although this exercise does not call for a tree diagram, one is shown at right. The numbers on the right side of the tree are found by the multiplication rule; for example, P(“regular” and “$\geq$ \$20”) $= P(R \text{ and } T) = P(R)\,P(T \mid R) = (0.4)(0.3) = 0.12$. The probability that the next customer pays at least \$20 is $P(T) = 0.12 + 0.175 + 0.15 = 0.445$.



4.106. $P(B \mid F) = \dfrac{P(B \text{ and } F)}{P(F)} = \frac{0.36}{0.58} \doteq 0.6207$—about 62%.

4.107. $P(P \mid T) = \dfrac{P(P \text{ and } T)}{P(T)} = \frac{0.15}{0.445} \doteq 0.3371$—about 34%.

4.108. With B, M, and D representing the three kinds of degrees, and W meaning the degree recipient was a woman, we have been given

$$P(B) = 0.73, \quad P(M) = 0.21, \quad P(D) = 0.06,$$
$$P(W \mid B) = 0.48, \quad P(W \mid M) = 0.42, \quad P(W \mid D) = 0.29.$$

Therefore,we find $P(W) = P(W \text{ and } B) + P(W \text{ and } M) + P(W \text{ and } D) = P(B)\,P(W \mid B) + P(M)\,P(W \mid M) + P(D)\,P(W \mid D) = 0.456$, so

$$P(B \mid W) = \frac{P(B \text{ and } W)}{P(W)} = \frac{P(B)\,P(W \mid B)}{P(W)} = \frac{0.3504}{0.456} \doteq 0.7684.$$

4.109. **(a)** Her brother has type aa, and he got one allele from each parent. But neither parent is albino, so neither could be type aa. **(b)** The table on the right shows the possible combinations, each of which is equally likely, so $P(aa) = 0.25$, $P(Aa) = 0.5$, and $P(AA) = 0.25$. **(c)** Beth is either AA or Aa, and $P(AA \mid \text{not } aa) = \frac{0.25}{0.75} = \frac{1}{3}$ while $P(Aa \mid \text{not } aa) = \frac{0.50}{0.75} = \frac{2}{3}$.

	A	a
A	AA	Aa
a	Aa	aa

4.110. **(a)** If Beth is Aa, then the first table on the right gives the (equally likely) allele combinations for a child, so $P(\text{child is non-albino} \mid \text{Beth is } Aa) = \frac{1}{2}$. If Beth is AA, then as the second table shows, their child will definitely be type Aa (and non-albino), so $P(\text{child is non-albino} \mid \text{Beth is } AA) = 1$. **(b)** We have

$$\begin{aligned} P(\text{child is non-albino}) &= P(\text{child } Aa \text{ and Beth } Aa) + P(\text{child } Aa \text{ and Beth } AA) \\ &= P(\text{Beth } Aa)\, P(\text{child } Aa \mid \text{Beth } Aa) + P(\text{Beth } AA)\, P(\text{child } Aa \mid \text{Beth } AA) \\ &= \tfrac{2}{3} \cdot \tfrac{1}{2} + \tfrac{1}{3} \cdot 1 = \tfrac{2}{3}. \end{aligned}$$

Therefore, $P(\text{Beth is } Aa \mid \text{child is } Aa) = \frac{1/3}{2/3} = \frac{1}{2}$.

	A	a
a	Aa	aa
a	Aa	aa

A	A
Aa	Aa
Aa	Aa

4.111. Let T be the event "test is positive" and C be the event "Jason is a carrier." Since the given information says that the test is never positive for noncarriers, it clearly must be the case that $P(C \mid T) = 1$.

To confirm this, note that (if there is no human error), we have $P(T \text{ and } C^c) = 0$, and $P(T) = P(T \text{ and } C) + P(T \text{ and } C^c) = P(T \text{ and } C) = P(T)\, P(T \mid C) = (0.04)(0.9) = 0.036$. Therefore, $P(C \mid T) = \dfrac{P(C \text{ and } T)}{P(T)} = \frac{0.036}{0.036} = 1$.

4.112. Let C be the event that Julianne is a carrier, and let D be the event that Jason's and Julianne's child has the disease. We have been given $P(C) = \frac{2}{3}$, $P(D \mid C) = \frac{1}{4}$, and $P(D \mid C^c) = 0$. Therefore, $P(D^c) = P(C)\, P(D^c \mid C) + P(C^c)\, P(D^c \mid C^c) = \left(\frac{2}{3}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{3}\right)(1) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$, and $P(C \mid D^c) = \frac{1/2}{5/6} = \frac{3}{5}$.

4.113. Let C be the event "Toni is a carrier," T be the event "Toni tests positive," and D be "her son has DMD." We have $P(C) = \frac{2}{3}$, $P(T \mid C) = 0.7$, and $P(T \mid C^c) = 0.1$. Therefore, $P(T) = P(T \text{ and } C) + P(T \text{ and } C^c) = P(C)\, P(T \mid C) + P(C^c)\, P(T \mid C^c) = \left(\frac{2}{3}\right)(0.7) + \left(\frac{1}{3}\right)(0.1) = 0.5$, and

$$P(C \mid T) = \frac{P(T \text{ and } C)}{P(C)} = \frac{(2/3)(0.7)}{0.5} = \frac{14}{15} \doteq 0.9333.$$

4.114. $P(\text{A}) = P(\text{B}) = \cdots = P(\text{F}) = \frac{0.72}{6} = 0.12$ and $P(1) = P(2) = \cdots = P(8) = \frac{1-0.72}{8} = 0.035$.

4.115. **(a)** All probabilities are greater than or equal to 0, and their sum is 1. **(b)** Let R_1 be Taster 1's rating and R_2 be Taster 2's rating. Add the probabilities on the diagonal (upper left to lower right): $P(R_1 = R_2) = 0.03 + 0.08 + 0.25 + 0.20 + 0.06 = 0.62$. **(c)** $P(R_1 > 3) = 0.39$ (the sum of the ten numbers in the bottom two rows), and $P(R_2 > 3) = 0.39$ (the sum of the ten numbers in the right two rows). Note that because the matrix is symmetric (relative to the main diagonal), these probabilities agree.

4.116. **(a)** $\mu_X = (1)(0.1) + (1.5)(0.2) + (2)(0.4) + (4)(0.2) + (10)(0.1) = 3$ million dollars and $\sigma_X = \sqrt{(1-3)^2(0.1) + (1.5-3)^2(0.2) + (2-3)^2(0.4) + (4-3)^2(0.2) + (10-3)^2(0.1)} = \sqrt{6.35} \doteq 2.5199$ million dollars. **(b)** $\mu_Y = \mu_{0.9X-0.2} = 0.9\mu_X - 0.2 = 2.5$ million dollars, and $\sigma_Y = \sigma_{0.9X-0.2} = 0.9\sigma_X \doteq 2.2679$ million dollars.

4.117. **(a)** The probability of winning nothing is $1 - \left(\frac{1}{10{,}000} + \frac{1}{1{,}000} + \frac{1}{100} + \frac{1}{20}\right) = 0.9389$.
(b) The mean is $\mu = (\$1000)\left(\frac{1}{10{,}000}\right) + (\$200)\left(\frac{1}{1{,}000}\right) + (\$50)\left(\frac{1}{100}\right) + (\$10)\left(\frac{1}{20}\right) = \$1.30$.
(c) $\sigma^2 = (\$998.70)^2\left(\frac{1}{10{,}000}\right) + (\$198.70)^2\left(\frac{1}{1{,}000}\right) + (\$48.70)^2\left(\frac{1}{100}\right) + (\$8.70)^2\left(\frac{1}{20}\right) = 168.31$, so $\sigma \doteq \$12.9734$.

4.118. As $\sigma_{a+bX} = b\sigma_X$ and $\sigma_{c+dY} = d\sigma_Y$, we need $b = \frac{100}{108}$ and $d = \frac{100}{109}$. With these choices for b and d, we have $\mu_{a+bX} = a + b\mu_X \doteq a + 409.2593$, so $a \doteq 90.7407$, and $\mu_{c+dX} = c + d\mu_Y \doteq c + 519.2661$, so $c \doteq -19.2661$.

4.119. This is the probability of 19 (independent) losses, followed by a win; by the multiplication rule, this is $0.994^{19} \cdot 0.006 \doteq 0.005352$.

4.120. **(a)** $P(\text{win the jackpot}) = \left(\frac{1}{20}\right)\left(\frac{9}{20}\right)\left(\frac{1}{20}\right) = 0.001125$. **(b)** The other symbol can show up on the middle wheel, with probability $\left(\frac{1}{20}\right)\left(\frac{11}{20}\right)\left(\frac{1}{20}\right) = 0.001375$, or on either of the outside wheels, with probability $\left(\frac{19}{20}\right)\left(\frac{9}{20}\right)\left(\frac{1}{20}\right) = 0.021375$. Therefore, combining all three cases, we have $P(\text{exactly two bells}) = 0.001375 + 2 \cdot 0.021375 = 0.044125$.

4.121. Let R_1 be Taster 1's rating and R_2 be Taster 2's rating. $P(R_1 = 3) = 0.01 + 0.05 + 0.25 + 0.05 + 0.01 = 0.37$, so

$$P(R_2 > 3 \mid R_1 = 3) = \frac{P(R_2 > 3 \text{ and } R_1 = 3)}{P(R_1 = 3)} = \frac{0.05 + 0.01}{0.37} \doteq 0.1622.$$

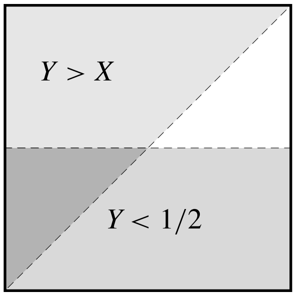
4.122. Let F be "adult is a full-time student," P be "adult is a part-time student," N be "adult is not a student," and A be "adult accesses Internet from someplace other than work or home."

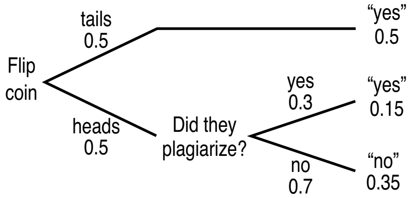
We were given $P(F) = 0.041$ and $P(P) = 0.029$, so that $P(N) = 1 - 0.041 - 0.029 = 0.93$. Also, $P(A \mid F) = 0.58$, $P(A \mid P) = 0.30$, and $P(A \mid N) = 0.21$. Therefore $P(A) = P(F)\,P(A \mid F) + P(P)\,P(A \mid P) + P(N)\,P(A \mid N) \doteq 0.22778$—about 22.8%.

4.123. Note first that $P(A) = \frac{1}{2}$ and $P(B) = \frac{2}{4} = \frac{1}{2}$. Now $P(B \text{ and } A) = P(\text{both coins are heads}) = 0.25$, so $P(B \mid A) = \dfrac{P(B \text{ and } A)}{P(A)} = \frac{0.25}{0.5} = 0.5 = P(B)$.

4.124. The event $\{Y < 1/2\}$ is the bottom half of the square, while $\{Y > X\}$ is the upper left triangle of the square. They overlap in a triangle with area 1/8, so

$$P(Y < \tfrac{1}{2} \mid Y > X) = \frac{P(Y < \frac{1}{2} \text{ and } Y > X)}{P(Y > X)} = \frac{1/8}{1/2} = \frac{1}{4}.$$





4.125. The response will be "no" with probability $0.35 = (0.5)(0.7)$. If the probability of plagiarism were 0.2, then $P(\text{student answers "no"}) = 0.4 = (0.5)(0.8)$. If 39% of students surveyed answered "no," then we estimate that $2 \cdot 39\% = 78\%$ have *not* plagiarized, so about 22% have plagiarized.

Chapter 5 Solutions

5.1. **(a)** This could be reasonably viewed as binomial with $n = 500$ and $p = 1/12$, because there is a fixed number (500) of independent trials with the same chance of success (1/12) on each try. **(b)** Not binomial: There is no fixed number of attempts (n). **(c)** Not binomial: There are no separate "trials" or "attempts" being observed here.

5.2. **(a)** No: There is no fixed number of observations. **(b)** A binomial distribution is reasonable here; a "large city" will have a population over 1000 (10 times as big as the sample). **(c)** In a "Pick 3" game, Joe's chance of winning the lottery is the same every week, so assuming that a year consists of 52 weeks (observations), this would be binomial.

5.3. **(a)** C, the number caught, is binomial with $n = 20$ and $p = 0.7$. M, the number missed, is binomial with $n = 20$ and $p = 0.3$. **(b)** Referring to Table C, we find $P(M \geq 9) = 0.0654 + 0.0308 + 0.0120 + 0.0039 + 0.0010 + 0.0002 = 0.1133$.

5.4. **(a)** X, the number of auction site visitors, is binomial with $n = 12$ and $p = 0.5$. **(b)** Referring to Table C, we find $P(X \geq 8) = 0.1208 + 0.0537 + 0.0161 + 0.0029 + 0.0002 = 0.1937$. (Software: 0.19385.)

5.5. **(a)** The mean of C is $(20)(0.7) = 14$ errors caught, and for M the mean is $(20)(0.3) = 6$ errors missed. **(b)** The standard deviation of C (or M) is $\sigma = \sqrt{(20)(0.7)(0.3)} \doteq 2.0494$ errors. **(c)** With $p = 0.9$, $\sigma = \sqrt{(20)(0.9)(0.1)} \doteq 1.3416$ errors; with $p = 0.99$, $\sigma \doteq 0.4450$ errors. σ decreases toward 0 as p approaches 1.

5.6. **(a)** The mean of X is $(12)(0.5) = 6$; the mean of $\hat{p}$ is 0.5. **(b)** The mean of X increases with n; it is 60 with $n = 120$, and 600 with $n = 1200$. The mean of $\hat{p}$ is 0.5 for any value of n.

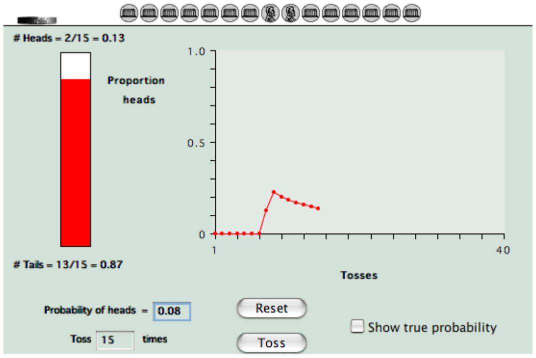
5.7. $m = 10$: From Table C, we see that $P(X \geq 10) = 0.0479$ and $P(X \geq 9) = 0.1133$. (Software gives $P(X \geq 10) = 0.04796$, but the conclusion is the same.)

5.8. **(a)** The population (the 75 members of the fraternity) is only three 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 $np = (500)(0.002) = 1$.

5.9. The count of 0s among n random digits has a binomial distribution with $p = 0.1$. **(a)** $P(\text{at least one } 0) = 1 - P(\text{no } 0) = 1 - (0.9)^5 \doteq 0.4095$. (Or take 0.5905 from Table C and subtract from 1.) **(b)** $\mu = (40)(0.1) = 4$.

5.10. 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 20 samples, most students should see 3 to 12 samples with exactly one bad record. That is, N, the number of samples with one bad record, has a binomial distribution with parameters $n = 20$ and $p = 0.3734$, and $P(3 \leq N \leq 12) = 0.9818$.


Heads = 2/15 = 0.13
Tails = 13/15 = 0.87
Probability of heads = 0.08
Toss 15 times
☐ Show true probability

5.11. **(a)** For Mark McGwire, $\mu_H = (509)(0.116) = 59.044$ home runs. **(b)** The exact answer (using software) is 0.0764. For the normal approximation, we compute $\sigma_H = \sqrt{(509)(0.116)(0.884)} \doteq 7.225$, so

$$P(H \geq 70) = P(\tfrac{H-59.044}{7.225} \geq \tfrac{70-59.044}{7.225}) = P(Z \geq 1.52) \doteq 0.0643.$$

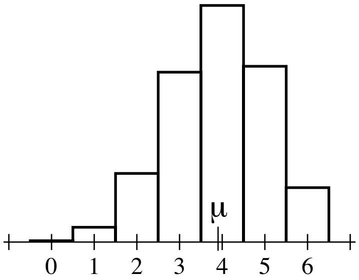
If we use the continuity correction,

$$P(H \geq 69.5) = P(\tfrac{H-59.044}{7.225} \geq \tfrac{69.5-59.044}{7.225}) = P(Z \geq 1.45) \doteq 0.0735.$$

(c) Regardless of the approach used to compute the probability, $P(H_2 \geq 73) < 0.000005$—basically 0.

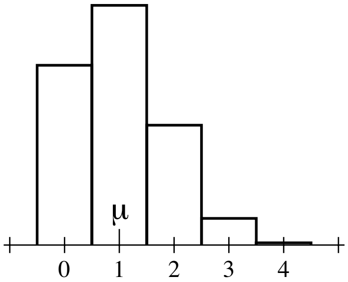
5.12. **(a)** $n = 6$ and $p = 0.65$. **(b)** The distribution is below; the histogram is on the right. **(c)** $\mu = np = 3.9$ years. **(d)** $\sigma = \sqrt{np(1-p)} \doteq 1.1683$ years; one standard deviation from μ means $P(3 \leq X \leq 5) = 0.8072$.

x	0	1	2	3	4	5	6
$P(X = x)$	.0018	.0205	.0951	.2355	.3280	.2437	.0754



5.13. **(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



5.14. For $\hat{p}$, $\mu = 0.49$ and $\sigma = \sqrt{p(1-p)/n} \doteq 0.01576$. As $\hat{p}$ is approximately normally distributed with this mean and standard deviation, we find

$$P(0.46 < \hat{p} < 0.52) \doteq P(-1.90 < Z < 1.90) = 0.9426$$

(Exact calculation gives 0.94565.)

5.15. 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 $P(0.29 < \hat{p} < 0.31) =$

$P(-0.69 < Z < 0.69) = 0.5098$. (Exact computation gives 0.50735.) **(b)** With $p = 0.06$, $\sigma \doteq 0.00747$, so $P(0.05 < \hat{p} < 0.07) = P(-1.34 < Z < 1.34) = 0.8198$. (Exact computation gives 0.81527.) **(c)** $P(-0.01 < \hat{p} - p < 0.01)$ increases to 1 as p gets closer to 0. (This is because σ also gets close to 0, so that $0.01/\sigma$ grows.)

5.16. When $n = 250$, the distribution of $\hat{p}$ is approximately normal with mean 0.49 and standard deviation 0.03162 (about twice that in Exercise 5.14). When $n = 4000$, the standard deviation drops to 0.00790 (half as big as in Exercise 5.14). Therefore,

$$\begin{aligned} n = 250: &\quad P(0.46 < \hat{p} < 0.52) \doteq P(-0.95 < Z < 0.95) = 0.6578 \\ n = 4000: &\quad P(0.46 < \hat{p} < 0.52) \doteq P(-3.80 < Z < 3.80) = 0.9998 \end{aligned}$$

Larger samples give a better probability that $\hat{p}$ will be close to the true proportion p. (Exact calculation of the first probability gives 0.68853, but this more accurate answer does not change our conclusion.)

5.17. **(a)** $\hat{p} = \frac{62}{100} = 0.62$. **(b)** We want $P(X \leq 62)$ or $P(\hat{p} \leq 0.62)$. 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 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.

			Continuity correction	
Exact prob.	Table normal	Software normal	Table normal	Software normal
0.1690	0.1446	0.1438	0.1685	0.1693

5.18. As $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$, we have $0.005^2 = (0.49)(0.51)/n$, so $n = 9996$.

5.19. **(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.

5.20. **(a)** $\mu = (1500)(0.7) = 1050$ and $\sigma = \sqrt{315} \doteq 17.7482$ students. **(b)** $P(X \geq 1000) \doteq P(Z \geq -2.82) = 0.9976$ (0.9978 with continuity correction; see the first line of the table on the right). **(c)** $P(X \geq 1201) \doteq P(Z \geq 8.51) < 0.00005$ (it's *very* small). **(d)** With $n = 1700$, $P(X \geq 1201) \doteq P(Z \geq 0.58) = 0.2810$. Other answers are shown in the second line of the table on the right.

		Continuity correction	
Table normal	Software normal	Table normal	Software normal
0.9976	0.9976	0.9978	0.9978
0.2810	0.2802	0.2877	0.2892

5.21. **(a)** X, the count of successes, has a binomial distribution with mean $\mu_X = np = (1000)(1/5) = 200$ and $\sigma_X = \sqrt{(1000)(0.2)(0.8)} \doteq 12.6491$ successes. **(b)** For $\hat{p}$, the mean is $\mu_{\hat{p}} = p = 0.2$ and $\sigma_{\hat{p}} = \sqrt{(0.2)(0.8)/1000} \doteq 0.01265$. **(c)** $P(\hat{p} > 0.24) \doteq P(Z > 3.16) = 0.0008$. (Exact computation gives 0.00111; using $P(X \geq 240)$ with the continuity correction gives 0.0009.) **(d)** From a standard normal distribution, $P(Z > 2.326) = 0.01$, so the subject must score 2.326 standard deviations above the mean: $\mu_{\hat{p}} + 2.326\sigma_{\hat{p}} = 0.2294$. This corresponds to 230 or more successes.

5.22. **(a)** M has a binomial distribution with $n = 30$ and $p = 0.7$, so $P(M = 20) = \binom{30}{20}(0.7)^{20}(0.3)^{10} \doteq 0.1416$. **(b)** $P(\text{1st woman is the 4th call}) = (0.7)^3(0.3) = 0.1029$.

Exact prob.	Table normal	Software normal	*Continuity correction* Table normal	Software normal
0.5038	0.4840	0.4834	0.5000	0.4996

5.23. **(a)** $p = \frac{23{,}772{,}494}{209{,}128{,}094} \doteq 0.1137$. **(b)** If B is the number of blacks, then B has a binomial distribution with parameters $n = 1500$ and $p \doteq 0.1137$, so the mean is $np \doteq 170.5$ blacks. **(c)** $P(B \le 170) \doteq P(Z < -0.04) = 0.4840$. See the table on the right for alternate answers.

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.*

5.24. **(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.

Table normal	Software normal	*Continuity correction* Table normal	Software normal
0.1251	0.1241	0.1492	0.1493
0.0336	0.0339	0.0401	0.0398

5.25. **(a)** $P(\hat{p} \le 0.70) = P(X \le 70)$ is on line 1. **(b)** $P(\hat{p} \le 0.70) = P(X \le 175)$ 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 $n = 100$, $\sigma = \sqrt{(0.75)(0.25)/100} \doteq 0.04330$; and with $n = 400$, $\sigma = \sqrt{(0.75)(0.25)/400} \doteq 0.02165$. **(d)** Yes: Regardless of p, n must be quadrupled to cut the standard deviation in half.

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

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

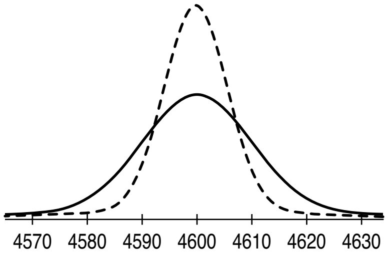
(c) $P(\text{first } ⚀ \text{ appears on toss 4}) = \left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)$.

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

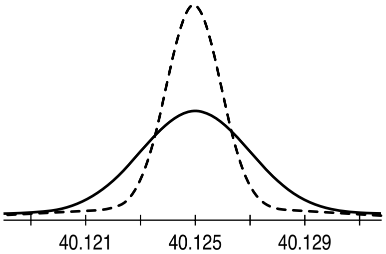
5.27. Y has possible values $1, 2, 3, \ldots$. $P(\text{first } ⚀ \text{ appears on toss } k) = \left(\frac{5}{6}\right)^{k-1}\left(\frac{1}{6}\right)$.

5.28. **(a)** $\sigma_{\bar{x}} = \sigma/\sqrt{3} = 10/\sqrt{3} \doteq 5.7735$ mg. **(b)** Solve $10/\sqrt{n} = 5$: $\sqrt{n} = 2$, so $n = 4$. The average of several measurements is more likely than a single measurement to be close to the mean.

5.29. Mean $\mu = 40.125$ mm and standard deviation $\sigma/\sqrt{4} = 0.001$ mm.



5.30. In Exercise 5.28, we found that $\sigma_{\bar{x}} \doteq 5.7735$ mg, so $\bar{x}$ has a $N(4600 \text{ mg}, 5.7735 \text{ mg})$ distribution. **(a)** On the right. The normal curve for $\bar{x}$ should be "taller and skinnier" (specifically, taller by a factor of $\sqrt{3}$, and skinnier by a factor of $1/\sqrt{3}$, although this detail will not be exact in a student's sketch). **(b)** The probability that an individual measurement misses the true mass by 10 mg—one standard deviation—is about 32% (by the 68–95–99.7 rule), or 0.3174 (using Table A). **(c)** 10 mg is $10/\sigma_{\bar{x}} \doteq 1.73$ standard deviations for the distribution of $\bar{x}$, so $P(\bar{x} < 4590 \text{ or } \bar{x} > 4610) = P(Z < -1.73 \text{ or } Z > 1.73) = 0.0836$ using Table A, or 0.0833 using software.



5.31. In Exercise 5.29, we found that $\sigma_{\bar{x}} = 0.001$ mm, so $\bar{x}$ has a $N(40.125 \text{ mm}, 0.001 \text{ mm})$ 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. **(b)** The probability that an individual axle diameter differs from the target by at least 0.004 mm—two standard deviations—is about 5% (by the 68–95–99.7 rule), or 0.0456 (using Table A). **(c)** 0.004 mm is four standard deviations for the distribution of $\bar{x}$, so this probability is $P(Z < -4 \text{ or } Z > 4) < 0.0001$.

5.32. **(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.

5.33. **(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).

5.34. **(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.

5.35. **(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, $P(\bar{x} > 0.6) \doteq P(Z > \frac{0.6-0.5}{0.09899}) = P(Z > 1.01) = 0.1562$.

5.36. **(a)** $\mu = (4)(0.21) + (3)(0.43) + (2)(0.30) + (1)(0.05) + (0)(0.01) = 2.78$ and $\sigma = \sqrt{0.7516} \doteq 0.8669$. **(b)** $\mu_{\bar{x}} = \mu = 2.78$ and $\sigma_{\bar{x}} = \sigma/\sqrt{50} \doteq 0.1226$. **(c)** $P(X \geq 3) = 0.21{+}0.43 = 0.64$, and $P(\bar{x} \geq 3) \doteq P(Z \geq \frac{3-2.78}{0.1226} = P(Z \geq 1.79) = 0.0367$ (software value: 0.0364).

5.37. 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 four measurements (assumed to be independent), then $\bar{x}$ has a $N(125, 10/\sqrt{4})$ or $N(125 \text{ mg/dl}, 5 \text{ mg/dl})$ distribution, and $P(\bar{x} > 140) = P(Z > 3) = 0.0013$.

5.38. **(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 $\sigma_{\bar{x}} = \sigma_X/\sqrt{365} \doteq \$0.8272$. The central limit theorem says that $\bar{x}$ is approximately normally distributed (with this mean and standard deviation). **(d)** Using this normal approximation, $P(\bar{x} > \$1) \doteq P(Z > 0.60) = 0.2743$ (software: 0.2728).

Note: *Joe comes out ahead if he wins at least once during the year. This probability is easily computed as* $1 - (0.999)^{365} \doteq 0.3059$. *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.*

5.39. The mean of four measurements has a $N(125 \text{ mg/dl}, 5 \text{ mg/dl})$ distribution, and $P(Z > 1.645) = 0.05$ if Z is $N(0, 1)$, so $L = 125 + 1.645 \cdot 5 = 133.225$ mg/dl.

5.40. $\bar{x}$ is approximately normal with mean 1.6 and standard deviation $1.2/\sqrt{200} \doteq 0.08485$ flaws/yd^2, so $P(\bar{x} > 2) \doteq P(Z > 4.71) = 0$ (essentially).

5.41. If W is total weight, and $\bar{x} = W/19$, then

$$P(W > 4000) = P(\bar{x} > 210.5263) \doteq P(Z > \tfrac{210.5263-190}{35/\sqrt{19}}) = P(Z > 2.56) = 0.0052$$

(software: 0.0053).

5.42. **(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.

5.43. 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.*

5.44. Let D be the shaft diameter, so that D has a $N(2.45\text{ cm}, 0.01\text{ cm})$ distribution.
(a) 99.38% fit into the hole: $P(D \le 2.475) = P(Z \le \frac{2.475-2.45}{0.01}) = P(Z \le 2.50) = 0.9938$.
(b) 96.16%: 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, $P(H - D \ge 0.025) = P(Z \ge \frac{0.025-0.05}{0.01414}) = P(Z \ge -1.77) \doteq 0.9616$ (software: 0.9615).

5.45. **(a)** The mean of five untreated specimens has a standard deviation of $2.3/\sqrt{5} \doteq 1.0286$ lbs, so $P(\bar{x}_u > 50) = P(Z > \frac{50-58}{1.0286}) = P(Z > -7.78)$, which is basically 1. **(b)** The mean of $\bar{x}_u - \bar{x}_t$ is $58 - 30 = 28$ lbs, and the standard deviation is $\sqrt{2.3^2/5 + 1.6^2/5} \doteq 1.2530$ lbs, so $P(\bar{x}_u - \bar{x}_t > 25) = P(Z > \frac{25-28}{1.2530}) = P(Z > -2.39) = 0.9916$ (software: 0.9917).

5.46. **(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 $1.5/\sqrt{30} \doteq 0.2739$. For *Enquirer* scores, $\bar{x}$ has mean 2.4 and standard deviation $1.6/\sqrt{30} \doteq 0.2921$. **(c)** $\bar{y} - \bar{x}$ has (approximately) a normal distribution with mean 2.4 and standard deviation $\sqrt{1.5^2/30 + 1.6^2/30} \doteq 0.4004$.
(d) $P(\bar{y} - \bar{x} \ge 1) = P(Z \ge \frac{1-2.4}{0.4004}) = P(Z \ge -3.50) \doteq 0.9998$.

5.47. **(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 $\sqrt{\sigma_Y^2/m + \sigma_X^2/n}$.

5.48. We have been given $\mu_X = 9\%$, $\sigma_X = 20\%$, $\mu_Y = 10\%$, $\sigma_Y = 17\%$, and $\rho = 0.6$.
(a) Linda's return $R = 0.8X + 0.2Y$ has mean $\mu_R = 0.8\mu_X + 0.2\mu_Y = 9.2\%$ and standard deviation $\sigma_R = \sqrt{(0.8\sigma_X)^2 + (0.2\sigma_Y)^2 + 2\rho(0.8\sigma_X)(0.2\sigma_Y)} \doteq 18.2439\%$. **(b)** $\overline{R}$, the average return over 20 years, has approximately a normal distribution with mean 9.2% and standard deviation $\sigma_R/\sqrt{20} \doteq 4.0795\%$, so $P(\overline{R} < 5\%) \doteq P(Z < -1.03) \doteq 0.1515$ (software gives 0.1516). **(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.*

5.49. 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. $P(H < 31.5 \text{ or } H > 32.5) = 1 - P(31.5 < H < 32.5) = 1 - P(-2.50 < Z < 2.50) = 1 - 0.9876 = 0.0124$.

5.50. The mean monthly fee for 500 households has approximately a $N(\$28, \$10/\sqrt{500}) = N(\$28, \$0.4472)$ distribution, so $P(\bar{x} > \$29) \doteq P(Z > 2.24) \doteq 0.0125$ (software: 0.0127).

5.51. **(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).

5.52. For each step of the random walk, the mean is $\mu = (1)(0.75) + (-1)(0.25) = 0.5$, the variance is $\sigma^2 = (1 - 0.5)^2(0.75) + (-1 - 0.5)^2(0.25) = 0.75$, and the standard deviation is $\sigma = \sqrt{0.75} \doteq 0.8660$. Therefore, $Y/500$ has approximately a $N(0.5, 0.03873)$ distribution, and $P(Y \geq 200) = P(\frac{Y}{500} \geq 0.4) \doteq P(Z \geq -2.58) = 0.9951$.

Note: *The number R of right-steps has a binomial distribution with $n = 500$ and $p = 0.75$. $Y \geq 200$ 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 0.99517.*

5.53. Out of six independent drivers, the number A who have an accident in the same year has a binomial distribution with $n = 6$ and $p = 0.2$, so $P(A \geq 3) = 0.0989$ (using Table C, software, or a calculator). Six roommates could not be considered independent, as they were not randomly chosen from the population of all drivers.

5.54. **(a)** Ramon's score is about the 64th percentile: $P(X \leq 1100) \doteq P(Z \leq \frac{1100-1026}{209}) = P(Z \leq 0.35) = 0.6368$ (software: 0.6384). Note that a continuity correction would be reasonable here, as SAT scores are reported as whole numbers; that is, we could compute $P(Z \leq \frac{1100.5-1026}{209}) = P(Z \leq 0.36) = 0.6406$ (software: 0.6393). **(b)** $\bar{x} = 1100$ is about the 99.9th percentile: $\bar{x}$ is approximately normal with mean 1026 and standard deviation $209/\sqrt{70} \doteq 24.9803$, so $P(\bar{x} \leq 1100) \doteq P(Z \leq 2.96) = 0.9985$. (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 Ramon's) might not be normal, but the central limit theorem says that the distribution of $\bar{x}$ will be close to normal.

5.55. **(a)** Out of ten independent vehicles, the number X with one person has a binomial distribution with $n = 10$ and $p = 0.7$, so $P(X \geq 6) = 0.8497$ (using Table C, software, or a calculator). **(b)** Y (the number of one-person cars in a sample of 100) has a binomial distribution with $n = 100$ and $p = 0.7$. Regardless of the approach used—normal approximation, or exact computation using software or a calculator—$P(Y \geq 51)$ is very close to 1 (the exact value is 0.99998).

5.56. 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 nine 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 $n = 9$ and $p = 0.056$. Then $P(N = 0) = (1 - 0.056)^9 \doteq 0.5953$.

5.57. 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 $n = 1200$ 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 $\sqrt{p(1-p)/1200} \doteq 0.01413$, so $P(\hat{p} \le \frac{680}{1200}) \doteq P(Z \le -2.50) = 0.0062$. Alternate answers shown on the right.

		Continuity correction	
Table normal	Software normal	Table normal	Software normal
0.0062	0.0062	0.0068	0.0067

5.58. **(a)** If R is the number of red-blossomed plants out of a sample of 8, then $P(R = 6) = 0.3115$, using a binomial distribution with $n = 8$ and $p = 0.75$. (For Table C, use $p = 0.25$ and find $P(X = 2)$, where $X = 8 - R$ is the number of flowers with nonred blossoms.) **(b)** With $n = 80$, the mean number of red-blossomed plants is $np = 60$. **(c)** If R_2 is the number of red-blossomed plants out of a sample of 80, then $P(R_2 \ge 50) \doteq P(Z \ge -2.58) = 0.9951$.

			Continuity correction	
Exact prob.	Table normal	Software normal	Table normal	Software normal
0.99542	0.9951	0.9951	0.9966	0.9966

5.59. If $\bar{x}$ is the average weight of 12 eggs, then $\bar{x}$ is $N(65\text{ g}, 5/\sqrt{12}\text{ g})$ or $N(65\text{ g}, 1.4434\text{ g})$, and $P(\frac{750}{12} < \bar{x} < \frac{825}{12}) \doteq P(-1.73 < Z < 2.60) = 0.9535$ (software: 0.9537).

5.60. **(a)** X has a binomial distribution with $n = 1555$ and $p = 0.2$. **(b)** $P(X \le 300) \doteq P(Z \le -0.7) = 0.2420$ (or see the table).

			Continuity correction	
Exact prob.	Table normal	Software normal	Table normal	Software normal
0.25395	0.2420	0.2428	0.2514	0.2528

Note: *Actually, X has a hypergeometric distribution, but the size of the population (all Internet users) is so much larger than the sample that the binomial distribution is an extremely good approximation.*

5.61. **(a)** No. Possible reasons: One could never have $X = 0$. There is no fixed number of "attempts" here. Solving $np = 1.5$ and $\sqrt{np(1-p)} = 0.75$ gives $p = 0.625$ and $n = 2.4$. **(b)** No: A count assumes only whole-number values, so it cannot be normally distributed. **(c)** $\bar{x}$ is approximately $N(1.5, 0.75/\sqrt{700})$ or $N(1.5, 0.02835)$. **(d)** P(more than 1075 in 700 cars) $= P(\bar{x} > \frac{1075}{700}) \doteq P(Z > 1.26) = 0.1038$ (software: 0.1039). We could also do a continuity correction for this question: $P(700\bar{x} > 1075.5) \doteq P(Z > 1.29) = 0.0985$ (software: 0.0994).

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

5.63. If $\hat{p}$ is the sample proportion who have been on a diet, then $\hat{p}$ has approximately a $N(0.70, 0.02804)$ distribution, so $P(\hat{p} \ge 0.75) \doteq P(Z \ge 1.78) = 0.0375$ (software: 0.0373). Alternatively, as $\hat{p} \ge 0.75$ is equivalent to 201 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
0.0329	0.0301	0.0298	0.0344	0.0347

5.64. (a) The machine that makes the caps and the machine that applies the torque are not the same. **(b)** T (torque) is $N(7, 0.9)$ and S (cap strength) is $N(10, 1.2)$, so $T - S$ is $N(7-10, \sqrt{0.9^2 + 1.2^2}) = N(-3 \text{ inch} \cdot \text{lb}, 1.5 \text{ inch} \cdot \text{lb})$. The probability that the cap breaks is $P(T > S) = P(T - S > 0) = P(Z > 2.00) = 0.0228$.

5.65. The center line is $\mu_{\bar{x}} = \mu = 4.22$ and the control limits are $\mu \pm 3\sigma/\sqrt{5} = 4.0496$ to 4.3904.

5.66. (a) $\bar{x}$ has a $N(32, 6/\sqrt{23}) = N(32, 1.2511)$ distribution, and $\bar{y}$ has a $N(29, 5/\sqrt{23})$ or $N(29, 1.0426)$ distribution, and **(b)** $\bar{y} - \bar{x}$ has a $N(29 - 32, \sqrt{5^2/23 + 6^2/23})$ or $N(-3, 1.6285)$ distribution. **(c)** $P(\bar{y} \geq \bar{x}) = P(\bar{y} - \bar{x} \geq 0) = P(Z \geq 1.84) = 0.0329$ (software: 0.0327).

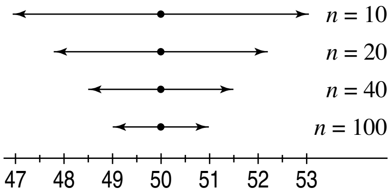
5.67. (a) $\hat{p}_F$ is approximately $N(0.82, 0.01718)$ and $\hat{p}_M$ is approximately $N(0.88, 0.01453)$. **(b)** When we subtract two independent normal random variables, the difference is normal. The new mean is the difference of the two means ($0.88-0.82 = 0.06$), and the new variance is the sum of the variances ($0.01453^2+0.01718^2 = 0.0005064$), so $\hat{p}_M-\hat{p}_F$ is approximately $N(0.06, 0.02250)$. **(c)** $P(\hat{p}_F > \hat{p}_M) = P(\hat{p}_M - \hat{p}_F < 0) \doteq P(Z < -2.67) = 0.0038$.

Chapter 6 Solutions

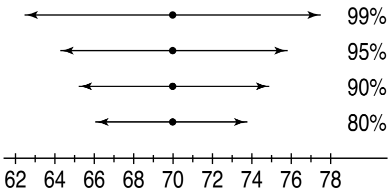
6.1. The 95% confidence interval is $76 \pm 12 = 64$ to 88. (The sample size is not needed.)

6.2. Greater than 12: A wider margin of error is needed in order to be more confident that the interval includes the true mean.

6.3. The margins of error are $1.96 \times 5/\sqrt{n}$, which yields 3.10, 2.19, 1.55, and 0.98. (And, of course, all intervals are centered at 50.) Interval width decreases with increasing sample size.



6.4. The margins of error are $z^* \times 15/\sqrt{25} = 3z^*$. With z^* equal to 1.282, 1.645, 1.960, and 2.576, this yields 3.846, 4.935, 5.880, and 7.728. (And, of course, all intervals are centered at 70.) Increasing confidence makes the interval wider.



6.5. The margin of error is 2.29 U/l, and the 95% confidence interval is $13.2 \pm (1.96)(6.5/\sqrt{31}) = 13.2 \pm 2.29 = 10.91$ to 15.49 U/l.

6.6. The 95% confidence interval is $33.4 \pm (1.96)(19.6/\sqrt{31}) = 33.4 \pm 6.90 = 26.50$ to 40.30 ng/ml.

6.7. The margin of error is 3.01 U/l, and the 99% confidence interval is $13.2 \pm (2.576)(6.5/\sqrt{31}) = 13.2 \pm 3.01 = 10.19$ to 16.21 U/l. This interval is wider than the 95% interval.

6.8. The 90% confidence interval is $33.4 \pm (1.645)(19.6/\sqrt{31}) = 33.4 \pm 5.79 = 27.61$ to 39.19 ng/ml. This interval is narrower than the 95% interval.

6.9. Scenario (B) has a smaller margin of error. Both samples would have the same value of z^* (1.96), but the value of σ would be smaller for (B), because we would have less variability in height for students in a single grade level.

6.10. The 95% confidence interval is $\$580 \pm (1.96)(\$90/\sqrt{10}) = \$580 \pm \$55.78 = \$524.22$ to \$635.78.

6.11. No: This is a range of values for the mean rent, not for individual rents.

Note: *To find a range to include 95% of all rents, we should take* $\mu \pm 2\sigma$*, where* μ *is the (unknown) mean rent for all apartments, and* σ *is the standard deviation for all apartments (assumed to be \$90 in Exercise 6.10). If* μ *were equal to \$580, for example, this range would be \$400 to \$760.*

6.12. If the distribution were roughly normal, the 68–95–99.7 rule says that 68% of all measurements should be in the range 13.8 to 53.0 ng/ml, 95% should be between -5.8 and 72.6 ng/ml, and 99.7% should be between -25.4 and 92.2 ng/ml. Because the measurements cannot be negative, this suggests that the distribution must be skewed to the right. The normal confidence interval should be fairly accurate nonetheless, because the sample mean $\bar{x}$ will still be roughly normal because of the central limit theorem.

6.13. **(a)** The 95% confidence interval is $80 \pm (1.96)(35/\sqrt{25}) = 80 \pm 13.72 = 66.28$ to 93.72 minutes. **(b)** No: This is a range of values for the mean study time, not for individual study times. (See also the solution to Exercise 6.11.)

6.14. **(a)** Divide both quantities by 60: $\bar{x} = \frac{80}{60} \doteq 1.3333$ hr and $\sigma = \frac{35}{60} \doteq 0.5833$ hr. **(b)** The 95% confidence interval is $1.3333 \pm (1.96)(0.5833/\sqrt{25}) = 1.3333 \pm 0.229 = 1.105$ to 1.562 hours. **(c)** Convert the interval from the previous exercise into hours: $\frac{66.28}{60} \doteq 1.105$ hr and $\frac{93.72}{60} \doteq 1.562$ hr.

6.15. **(a)** The standard deviation of $\bar{x}$ is $\sigma_{\bar{x}} = \sigma/\sqrt{200} \doteq 2.1920$ cal/day. **(b)** The probability is about 0.95 that $\bar{x}$ is within 4.3840 cal/day (two standard deviations) of the population mean μ. **(c)** About 95% of all samples will capture the true mean of calories consumed per day in the interval $\bar{x} \pm$ 4.3840 cal/day. Some students may use 1.96 rather than 2 and answer 4.2964. (This is the whole idea behind a confidence interval: Probability tells us that $\bar{x}$ is usually close to μ. That is equivalent to saying that μ is usually close to $\bar{x}$.)

6.16. **(a)** The standard deviation of $\bar{x}$ is $\sigma_{\bar{x}} = \sigma/\sqrt{20} \doteq 0.6485$ MPG. **(b)** The mean is $\bar{x} = 18.48$ MPG, and the 95% confidence interval is $18.48 \pm (1.96)(2.9/\sqrt{20}) = 18.48 \pm 1.27 = 17.21$ to 19.75 MPG.

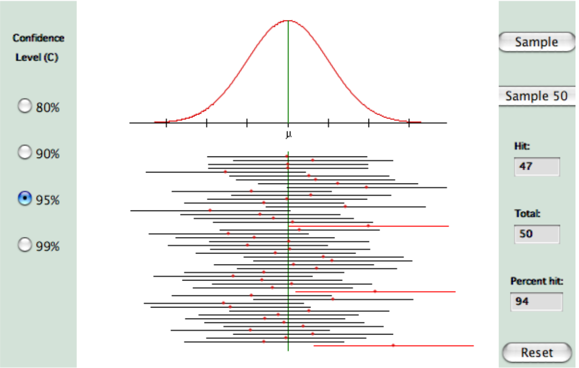
6.17. The mean is $\bar{x} = 30.8$ MPH, and the 95% confidence interval is $30.8 \pm (1.96)(10.3/\sqrt{20}) = 30.8 \pm 4.5142 = 26.2858$ to 35.3142 MPH.

6.18. The mean $\bar{x} = 30.8$ MPH converts to 49.28 KPH, and the margin of error 4.5142 MPH converts to 7.2227 KPH. The new interval is therefore 42.0573 to 56.5027 KPH.

6.19. The mean is $\bar{x} \doteq 35.0909$, and the 95% confidence interval is $35.0909 \pm (1.96)(11/\sqrt{44}) = 35.0909 \pm 3.2503 = 31.8406$ to 38.3412.

6.20. One sample screen is shown below (left), and a sample stemplot of results in shown on the right. The number of hits will vary, but the distribution should follow a binomial distribution with $n = 50$ and $p = 0.95$, so we expect the average number of hits to be about 47.5. We also find that about 99.7% of individual counts should be 43 or more, and the mean hit count for 30 samples should be approximately normal with mean 47.5 and standard deviation 0.2814—so almost all sample means should be between 46.66 and 48.34.

44	00
45	0000
46	00
47	00000000
48	000000
49	000
50	00000



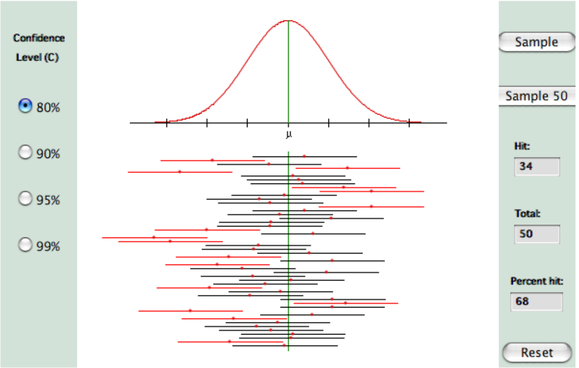
Confidence Level (C)

○ 80%
○ 90%
● 95%
○ 99%

Hit: 47

Total: 50

Percent hit: 94



Confidence Level (C)

● 80%
○ 90%
○ 95%
○ 99%

Hit: 34

Total: 50

Percent hit: 68

6.21. One sample screen is shown above (right), and a sample stemplot of results in shown on the right. The number of hits will vary, but the distribution should follow a binomial distribution with $n = 50$ and $p = 0.8$, so we expect the average number of hits to be about 40. We also find that about 99.7% of individual counts should be between 32 and 48, and the mean hit count for 30 samples should be approximately normal with mean 40 and standard deviation 0.5164—so almost all sample means should be between 38.45 and 41.55.

31	0
32	
33	
34	00
35	0
36	000
37	00
38	00
39	00
40	0000
41	0000
42	00000
43	0
44	00
45	0

6.22. $n = \left(\frac{1.96 \times \$9000}{\$400}\right)^2 = 1944.81$—take $n = 1945$.

6.23. It will be smaller, because we require less precision (and therefore need less information). Specifically, because the desired margin of error is twice as large, the sample size will be about one-fourth as large as before: $n = \left(\frac{1.96 \times \$9000}{\$800}\right)^2 \doteq 486.20$—take $n = 487$.

6.24. $n = \left(\frac{1.645 \times \$90}{\$20}\right)^2 \doteq 54.80$—take $n = 55$.

6.25. $n = \left(\frac{1.96 \times 6.5}{2.0}\right)^2 \doteq 40.58$—take $n = 41$.

6.26. If we start with a sample of size k and lose 20% of the sample, we will end with $0.8k$. Therefore, we need to *increase* the sample size by *25%*—that is, start with a sample of size $k = 1.25n$—so that we end with $(0.8)(1.25n) = n$. With $n = 41$, that means we should initially sample 52 subjects.

6.27. This is probably not a confidence interval; it is not intended to give an estimate of the mean income, but rather it gives the range of incomes earned by all (or most) telemarketers working for this company.

6.28. **(a)** The 98% confidence interval is $10.0023 \pm (2.326)(0.0002/\sqrt{5}) = 10.0023 \pm 0.0002 =$ 10.0021 to 10.0025 g. **(b)** $n = \left(\frac{(2.326)(0.0002)}{0.0001}\right)^2 \doteq 21.64$—take $n = 22$.

6.29. The number of hits has a binomial distribution with parameters $n = 4$ and $p = 0.95$, so the number of misses is binomial with $n = 4$ and $p = 0.05$. We can therefore use Table C to answer these questions. **(a)** The probability that all cover their means is $0.95^4 \doteq 0.8145$. (Or use Table C to find the probability of 0 misses.) **(b)** The probability that at least three cover their means is $0.95^4 + 4(0.05)(0.95^3) \doteq 0.9860$. (Or use Table C to find the probability of 0 or 1 misses.)

6.30. **(a)** The 95% confidence interval would be $0.80 \pm 0.04 = 0.76$ to 0.84. **(b)** The confidence interval methods used here require that the results come from an SRS of the population (which presumably is "the newspaper's readers"). This sample cannot be viewed as an SRS.

6.31. **(a)** The results will vary from one sample to the next; one sample will not perfectly represent the population. **(b)** The method used gives correct results 95% of the time. **(c)** The 95% confidence interval is $37\% \pm 3\% = 34\%$ to 40%. **(d)** Yes; some fans would change their responses based on what they are currently watching.

6.32. **(a)** Our hypothesis should be "some claim about the population." Whether or not it rains tomorrow is not such a statement. Put another way, hypothesis testing—at least as described in this text—does not deal with random outcomes, but rather with statements that are either true or false. Rain (or not) is a random outcome. **(b)** The standard deviation of the sample mean is $15/\sqrt{20}$. **(c)** The null hypothesis should be a statement about μ, not $\bar{x}$.

6.33. **(a)** H_0 and H_a have been switched: The null hypothesis should be a statement of "no change." **(b)** H_0 should concern μ, not $\bar{x}$. **(c)** A P-value of 0.95 is not significant by any reasonable standard.

6.34. **(a)** One possibility: If μ_1 is the mean change in bone density for a sample of rats eating a high-soy diet, and μ_2 is the mean change for a control group, we could test H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 \neq \mu_2$. (The alternative might be one-sided, for example, $\mu_1 > \mu_2$.) We could also state hypotheses which refer to a single mean μ, if we simply look at change for one group of rats (with no control group). **(b)** If μ is the mean rating for all students, we could test H_0: $\mu = 0$ vs. H_a: $\mu > 0$. **(c)** With μ as the mean score of all of this TA's students, we test H_0: $\mu = 75$ vs. H_a: $\mu > 75$.

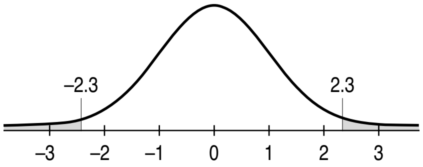
6.35. **(a)** If μ is the population mean proportion of food expenditures in restaurants (that is, this is the average of the fraction r/f for all local households, where f is total food expenditures, and r is restaurant food expenditures), then we test H_0: $\mu = 0.3$ vs. H_a: $\mu \neq 0.3$. **(b)** H_0: $\mu = 20$ seconds vs. H_a: $\mu < 20$ seconds. **(c)** If μ is the mean DXA reading for the phantom, we test H_0: $\mu = 1.3$ g/cm^2 vs. H_a: $\mu \neq 1.3$ g/cm^2.

Note: *For (a), the hypotheses could be stated in terms of a proportion p, but note that this kind of proportion is fundamentally different from proportions in binomial populations. The latter proportions represent the probability of success in independent trials; in this case, we are measuring a fraction of total food expenditures.*

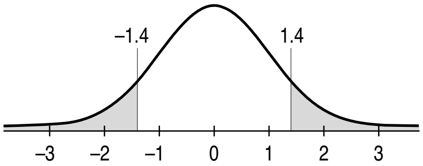
6.36. **(a)** H_0: $\mu_A = \mu_B$; H_a: $\mu_A > \mu_B$, where μ_A is the mean score for group A and μ_B is the group B mean. **(b)** H_0: $\rho = 0$; H_a: $\rho > 0$, where ρ is the (population) correlation between GPA and self-esteem. **(c)** H_0: $p_m = p_f$; H_a: $p_m < p_f$, where p_m is the proportion of males who name English as their favorite subject, and p_f is that proportion for females.

6.37. **(a)** H_0: $\mu = \$72{,}500$; H_a: $\mu > \$72{,}500$. **(b)** H_0: $\mu = 1.8$ hr; H_a: $\mu \neq 1.8$ hr.

6.38. $P(Z > 2.3) = 0.0107$, so the two-sided P-value is $2(0.0107) = 0.0214$.



6.39. $P(Z < -1.4) = 0.0808$, so the two-sided P-value is $2(0.0808) = 0.1616$.



6.40. **(a)** No, we do not reject H_0 at $\alpha = 0.05$. **(b)** No, we do not reject H_0 at $\alpha = 0.01$. **(c)** We have $P = 0.082$; to reject, we need $P < \alpha$.

6.41. **(a)** Yes, we reject H_0 at $\alpha = 0.05$. **(b)** No, we do not reject H_0 at $\alpha = 0.01$. **(c)** We have $P = 0.032$; to reject, we need $P < \alpha$.

6.42. **(a)** For H_a: $\mu > \mu_0$, the P-value is $P(Z > 1.6) = 0.0548$.
(b) For H_a: $\mu < \mu_0$, the P-value is $P(Z < 1.6) = 0.9452$.
(c) For H_a: $\mu \neq \mu_0$, the P-value is $2P(Z > 1.6) = 2(0.0548) = 0.1096$.

6.43. **(a)** For H_a: $\mu > \mu_0$, the P-value is $P(Z > -1.6) = 0.9452$.
(b) For H_a: $\mu < \mu_0$, the P-value is $P(Z < -1.6) = 0.0548$.
(c) For H_a: $\mu \neq \mu_0$, the P-value is $2P(Z < -1.6) = 2(0.0548) = 0.1096$.

6.44. Recall the statement from the text: "A level α two-sided significance test rejects a hypothesis H_0: $\mu = \mu_0$ exactly when the value μ_0 falls outside a level $1 - \alpha$ confidence interval for μ." **(a)** Yes, 30 is in the 95% confidence interval, because $P = 0.09$ means that we would not reject H_0 at $\alpha = 0.05$. **(b)** No, 30 is not in the 90% confidence interval, because we would reject H_0 at $\alpha = 0.10$.

6.45. Recall the statement from the text: "A level α two-sided significance test rejects a hypothesis H_0: $\mu = \mu_0$ exactly when the value μ_0 falls outside a level $1 - \alpha$ confidence interval for μ." **(a)** No, 30 is not in the 95% confidence interval, because $P = 0.04$ means that we would reject H_0 at $\alpha = 0.05$. **(b)** No, 30 is not in the 90% confidence interval, because we would also reject H_0 at $\alpha = 0.10$.

6.46. **(a)** If the alternative is two-sided, the answer is *yes*; see the quote from the text in the solutions to Exercises 6.44 and 6.45 for an explanation. If H_a is one-sided, then the answer depends on its direction: If H_a is $\mu > 68$, the answer is *no*. If H_a is $\mu < 68$, the answer is *yes*, because the given interval suggests $\bar{x} = 61$ and the standard error of the mean is about 2 (or less, if the interval were constructed using the t distribution rather than the normal distribution), so $\bar{x}$ is 3.5 standard errors from 68. **(b)** Regardless of the alternative, we would not reject H_0, because 62 falls well inside the confidence interval.

6.47. **(a)** Regardless of the alternative, we would not reject H_0, because 13 falls well inside the confidence interval. **(a)** If the alternative is two-sided, the answer is *yes*; see the quote from the text in the solutions to Exercises 6.44 and 6.45 for an explanation. If H_a is one-sided, then the answer depends on its direction: If H_a is $\mu < 10$, the answer is *no*. If H_a is $\mu > 10$, the answer is *yes*, because the given interval suggests $\bar{x} = 13.5$ and the standard error of the mean is about 0.9 (or less, if the interval were constructed using the t distribution rather than the normal distribution), so $\bar{x}$ is 3.8 standard errors from 10.

6.48. The study presumably examined malarial infection rates in two groups of subjects—one with bed nets, and one without. The observed difference between the two groups were so large that they would be unlikely to occur by chance if bed nets had no effect. Specifically, if the groups were the same, and we took many samples, the difference in malarial infections would be so large less than 0.1% of the time.

6.49. Even if the two groups (the health and safety class, and the statistics class) had the same level of alcohol awareness, there might be some difference in our sample due to chance. The difference observed was large enough that it would rarely arise by chance. The reason for this difference might be that health issues related to alcohol use are probably discussed in the health and safety class.

6.50. Even if scores had not changed over time, random fluctuation might cause the mean in 2003 to be different from the 2000 mean. However, in this case the difference was so great that it is unlikely to have occurred by chance; specifically, such a difference would arise less than 5% of the time if the actual mean had not changed. We therefore conclude that the mean did change from 2000 to 2003.

6.51. If the mean score for all Boston students were 287, then by chance we might expect to see a sample mean of 262. In fact, there is greater than a 5% chance of this happening, so we do not consider a mean of 262 to be strong enough evidence to conclude that Boston's mean is less than 287.

6.52. If μ is the mean north-south location, the hypotheses are H_0: $\mu = 100$ vs. H_a: $\mu \neq 100$. For testing these hypotheses, we find $z = \frac{99.74-100}{58/\sqrt{584}} \doteq -0.11$. This is not significant—$P = 2(0.4562) = 0.9124$—so we have no reason to doubt a uniform distribution based on this test.

6.53. If μ is the mean east-west location, the hypotheses are H_0: $\mu = 100$ vs. H_a: $\mu \neq 100$ (as in the previous exercise). For testing these hypotheses, we find $z = \frac{113.8-100}{58/\sqrt{584}} \doteq 5.75$. This is highly significant ($P < 0.0001$), we conclude that the trees are not uniformly spread from east to west.

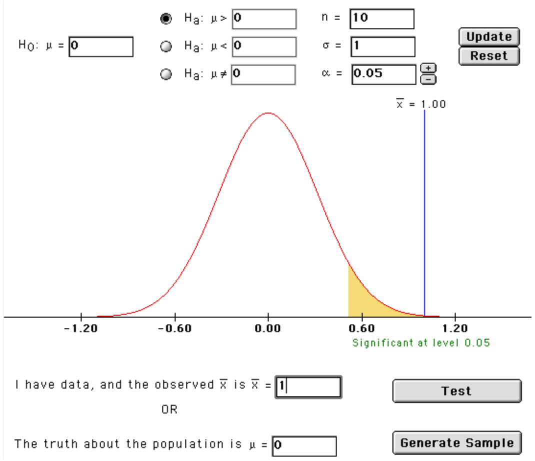
6.54. For testing these hypotheses, we find $z = \frac{10.2-8.9}{2.5/\sqrt{6}} \doteq 1.27$. This is not significant ($P = 0.1020$); there is not enough evidence to conclude that these sonnets were not written by our poet.

6.55. **(a)** $z = \frac{132.2-115}{30/\sqrt{25}} \doteq 2.87$, so the P-value is $P = P(Z > 2.87) = 0.0021$. This is strong evidence that the older students have a higher SSHA mean. **(b)** The important assumption is that this is an SRS from the population of older students. We also assume a normal distribution, but this is not crucial provided there are no outliers and little skewness.

6.56. **(a)** H_0: $\mu = 9.5$ mg/dl vs. H_a: $\mu \neq 9.5$ mg/dl. **(b)** $z = \frac{9.57-9.5}{0.4/\sqrt{160}} \doteq 2.21$, so the P-value is $P = 2P(Z > 2.21) = 0.0272$. This is pretty strong evidence that μ is different from (greater than) 9.5 mg/dl. **(c)** The 95% confidence interval is $9.57 \pm (1.96)(0.4/\sqrt{160}) = 9.57 \pm 0.062 = 9.508$ to 9.632.

6.57. **(a)** H_0: $\mu = 32$; H_a: $\mu > 32$. **(b)** $\bar{x} = 35.090-$, so $z = \frac{35.0909-32}{11/\sqrt{44}} \doteq 1.86$ and $P = 0.0314$. This is fairly good evidence that children in this district have a mean score higher than the national average—observations this extreme would occur in only about 3 out of 100 samples if H_0 were true.

6.58. A sample screen (for $\bar{x} = 1$) is shown below on the left. As one can judge from the shading under the normal curve, $\bar{x} = 0.5$ is not significant, but 0.6 is. (In fact, the cutoff is about 0.52, which is approximately $1.645/\sqrt{10}$.)



H_0: μ = 0

● H_a: $\mu >$ 0
○ H_a: $\mu <$ 0
○ H_a: $\mu \neq$ 0

n = 10
σ = 1
α = 0.05

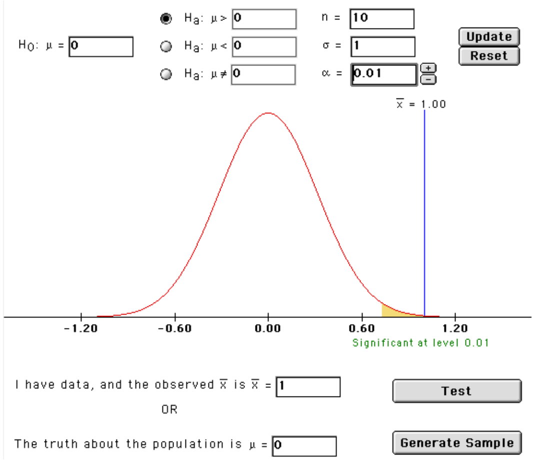
$\bar{x}$ = 1.00

Significant at level 0.05

I have data, and the observed $\bar{x}$ is $\bar{x}$ = 1

OR

The truth about the population is μ = 0



H_0: μ = 0

● H_a: $\mu >$ 0
○ H_a: $\mu <$ 0
○ H_a: $\mu \neq$ 0

n = 10
σ = 1
α = 0.01

$\bar{x}$ = 1.00

Significant at level 0.01

I have data, and the observed $\bar{x}$ is $\bar{x}$ = 1

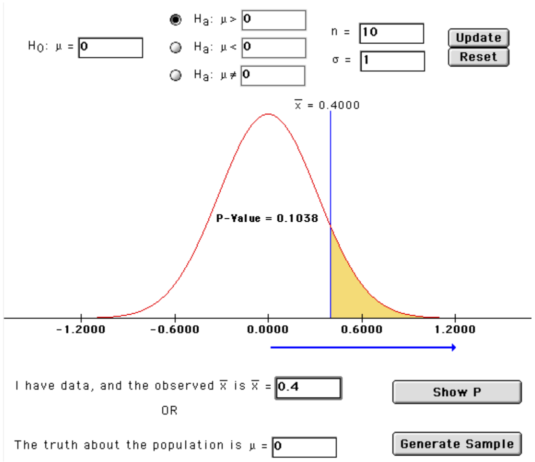
OR

The truth about the population is μ = 0

6.59. See the sample screen (for $\bar{x} = 1$) above on the right. As one can judge from the shading under the normal curve, $\bar{x} = 0.7$ is not significant, but 0.8 is. (In fact, the cutoff is about 0.7354, which is approximately $2.326/\sqrt{10}$.) Smaller α means that $\bar{x}$ must be farther away from μ_0 in order to reject H_0.

6.60. A sample screen (for $\bar{x} = 0.4$) is shown on the right. The P-values given by the applet are listed in the table below; as $\bar{x}$ moves farther away from μ_0, P decreases.

$\bar{x}$	P	$\bar{x}$	P
0.1	0.3745	0.6	0.0287
0.2	0.2643	0.7	0.0136
0.3	0.1711	0.8	0.0057
0.4	0.1038	0.9	0.0022
0.5	0.0571	1	0.0008



H_0: μ = 0

● H_a: $\mu >$ 0

○ H_a: $\mu <$ 0

○ H_a: $\mu \neq$ 0

n = 10

σ = 1

I have data, and the observed $\bar{x}$ is $\bar{x}$ = 0.4

OR

The truth about the population is μ = 0

6.61. Either compare with the critical values in Table D, or determine the P-value (0.0401). **(a)** Yes, because $z > 1.645$ (or because $P < 0.05$). **(b)** No, because $z < 2.326$ (or because $P > 0.01$).

6.62. **(a)** $z = \frac{0.4365-0.5}{0.2887/\sqrt{100}} \doteq -2.20$. For the other parts of this problem, either compare this with the critical values in Table D, or determine the P-value (0.0278). **(b)** This result is significant at the 5% level because $z < -1.960$ (or because $P < 0.05$). **(c)** It is not significant at 1% because $z \geq -2.576$ (or because $P > 0.01$).

6.63. When a test is significant at the 1% level, it means that if the null hypothesis were true, outcomes similar to those seen are expected to occur less than once in 100 repetitions of the experiment or sampling. "Significant at the 5% level" means we have observed something that occurs in less than 5 out of 100 repetitions (when H_0 is true). Something that occurs "less than once in 100 repetitions" also occurs "less than 5 times in 100 repetitions," so significance at the 1% level implies significance at the 5% level (or any higher level).

6.64. Something that occurs "less than 5 times in 100 repetitions" is not necessarily as rare as something that occurs "less than once in 100 repetitions," so a test that is significant at 5% is not necessarily significant at 1%.

6.65. Using Table D or software, we find that the 0.005 critical value test is 2.576, and the 0.0025 critical value is 2.807. Therefore, if $2.576 < |z| < 2.807$—that is, either $2.576 < z < 2.807$ or $-2.807 < z < -2.576$—then z would be significant at the 1% level but not at the 0.5% level.

6.66. As $-4.3 < -3.291$, the two-sided P-value is $P < 2(0.0005) = 0.001$.

6.67. As $0.22 < 0.674$, the one-sided P-value is $P > 0.25$.

6.68. **(a)** Reject H_0 if $z > 1.645$. **(b)** Reject H_0 if $|z| > 1.960$. **(c)** For tests at a fixed significance level (α), we reject H_0 when we observe values of our statistic that are so extreme (far from the mean, or other "center" of the sampling distribution) that they would rarely occur when H_0 is true. (Specifically, they occur with probability no greater than α.) For a two-sided alternative, we split the rejection region—this set of extreme values—into two pieces, while with a one-sided alternative, all the extreme values are in one piece, which is twice as large (in area) as either of the two pieces used for the two-sided test.

6.69. Because $1.282 < 1.34 < 1.645$, the P-value is between $2(0.05) = 0.10$ and $2(0.10) = 0.20$. From Table A, $P = 2(0.0901) = 0.1802$.

6.70. Because the alternative is two-sided, the answer is the same for $z = -1.34$: $-1.282 > -1.34 > -1.645$, so Table D says that $0.10 < P < 0.20$, and Table A gives $P = 2(0.0901) = 0.1802$.

6.71. **(a)** The 95% confidence interval is $104.1\overline{3} \pm (1.96)(9/\sqrt{12}) = 104.1\overline{3} \pm 5.0922 = 99.0411$ to 109.2256 pCi/L. **(b)** We test H_0: $\mu = 105$ pCi/L vs. H_a: $\mu \neq 105$ pCi/L. Because 105 is in the interval from (a), we do not have enough evidence to reject H_0.

6.72. Finding something to be "statistically significant" is not really useful unless the significance level is sufficiently small. While there is some freedom to decide what "sufficiently small" means, $\alpha = 0.5$ would lead your team to incorrectly rejecting H_0 half the time, so it is clearly a bad choice. (This approach would be essentially equivalent to flipping a coin to make your decision!)

6.73. The first test was barely significant at $\alpha = 0.05$, while the second was significant at any reasonable α.

6.74. We expect 50 tests to be statistically significant: Each of the 1000 tests has a 5% chance of being significant, so the number of significant tests has a binomial distribution with $n = 1000$ and $p = 0.05$, for which the mean is $np = 50$.

6.75. $P = 0.00001 = \frac{1}{100{,}000}$, so we would need $n = 100{,}000$ tests in order to expect one P-value of this size (assuming that all null hypotheses are true). That is why we reject H_0 when we see P-values such as this: It indicates that our results would rarely happen if H_0 were true.

6.76. The study may have rejected $\mu = \mu_0$ (or some other null hypothesis), but with such a large sample size, such a rejection might occur even if the actual mean (or other parameter) differs only slightly from μ_0. For example, there might be no practical importance to the difference between $\mu = 10$ and $\mu = 10.5$.

6.77. We expect more variation with small sample sizes, so even a large difference between $\bar{x}$ and μ_0 (or whatever measures are appropriate in our hypothesis test) might not turn out to be significant. If we were to repeat the test with a larger sample, the decrease in the standard error might give us a small enough P-value to reject H_0.

6.78. $P = 0.95$ means that we have no reason to doubt the null hypothesis. Specifically, results like those observed in the sample would occur 95% of the time when H_0 is true.

6.81. When many variables are examined, "significant" results will show up by chance, so we should not take it for granted that the variables identified are really indicative of future success. In order to decide if they are appropriate, we should track this year's trainees, and compare the success of those from urban/suburban backgrounds with the rest, and likewise compare those with a degree in a technical field with the rest.

6.82. **(a)** $z = \frac{483-480}{100/\sqrt{100}} = 0.3$, so $P = P(Z > 0.3) = 0.3821$. **(b)** $z = \frac{483-480}{100/\sqrt{1000}} \doteq 0.95$, so $P = P(Z > 0.95) = 0.1711$. **(c)** $z = \frac{483-480}{100/\sqrt{10000}} = 3$, so $P = P(Z > 3) = 0.0013$.

6.83. The interval is $483 \pm (2.576)(100/\sqrt{n})$. $n = 100$: 457.24 to 508.76. $n = 1000$: 474.85 to 491.15. $n = 10,000$: 480.42 to 485.58.

6.84. **(a)** $z = \frac{496.4-480}{100/\sqrt{100}} = 1.64$. This is not significant at $\alpha = 0.05$, because $z < 1.645$ (or $P = 0.0505$). **(b)** $z = \frac{496.5-480}{100/\sqrt{100}} = 1.65$. This *is* significant at $\alpha = 0.05$, because $z > 1.645$ (or $P = 0.0495$).

6.86. Using $\alpha/7 \doteq 0.007143$ as the cutoff, the second ($P = 0.003$), fourth ($P = 0.004$), and seventh ($P = 0.001$) are significant.

6.87. Using $\alpha/10 = 0.005$ as the cutoff, we reject the fourth ($P = 0.004$), fifth ($P = 0.001$), and ninth ($P = 0.002$).

6.88. With all 50 states (plus Puerto Rico) listed in the table, we have information about the entire population in question; no statistical procedures are needed (or meaningful).

6.90. The power of this study is far lower than what is generally desired—for example, it is well below the "80% standard" mentioned in the text. Twenty percent power for the specified effect means that, if the effect is present, we will only detect it 20% of the time. With such a small chance of detecting an important difference, the study should probably not be run (unless the sample size is increased to give sufficiently high power).

6.91. A larger sample gives more information and therefore gives a better chance of detecting a given alternative; that is, larger samples give more power.

6.92. The power for $\mu = -10$ is the same as the power for $\mu = 10$, because both alternatives are an equal distance from the null value of μ. (The symmetry of two-sided tests with the normal distribution means that we only need to consider the size of the difference, not the direction.)

6.93. The power for $\mu = 80$ will be higher than 0.5, because larger differences are easier to detect.

6.94. The applet (or other software) gives the power as about 0.061.

Note: *At the time these solutions were prepared, the* Power *applet would not allow sample sizes over 50. If that is still the case, students can be directed to use the online power calculators at* `http://calculators.stat.ucla.edu/powercalc`. *For this problem, choose "Normal: 1 sample."*

6.95. The applet (or other software) gives the power as about 0.99. (See the note in the previous solution.)

6.96. **(a)** $\alpha = P(\text{Type I error}) = P(\bar{x} > 43.12 \text{ when } \mu = 40)$

$= P(Z > \frac{43.12-40}{60/\sqrt{1000}}) = P(Z > 1.64) = 0.0505$ (software: 0.0500).

(b) $P(\text{Type II error when } \mu = 45) = P(\bar{x} \le 43.12 \text{ when } \mu = 45)$

$= P(Z \le \frac{43.12-45}{60/\sqrt{1000}}) = P(Z < -0.99) = 0.1611$ (software: 0.1609).

(c) $P(\text{Type II error when } \mu = 50) = P(\bar{x} < 43.12 \text{ when } \mu = 50)$

$= P(Z \le \frac{43.12-50}{60/\sqrt{1000}}) = P(Z < -3.63) = 0.0001$.

(d) The sample size ($n = 1000$) is so large that the mean will be very close to normal.

6.97. $z \ge 2.326$ is equivalent to $\bar{x} \ge 450 + 2.326(100/\sqrt{500}) \doteq 460.4$, so the power is

$$P(\text{reject } H_0 \text{ when } \mu = 462) = P(\bar{x} \ge 460.4 \text{ when } \mu = 462)$$
$$= P\left(Z \ge \frac{460.4-462}{100/\sqrt{500}}\right) \doteq P(Z \ge -0.36) = 0.6406.$$

(Software gives the power as 0.6394.) This is not too bad, but a bit less than the "80% power" standard.

6.98. **(a)** $P(\text{Type I error}) = P(X = 0 \text{ or } X = 1 \text{ when the distribution is } p_0) = 0.2$.

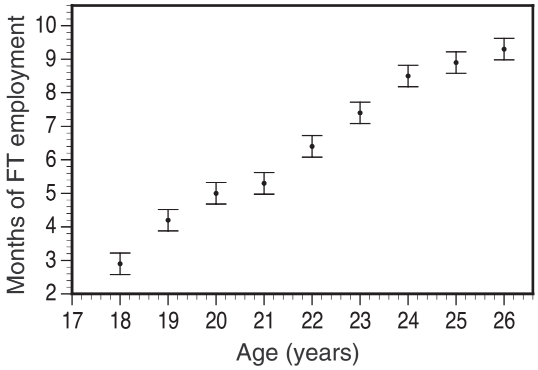
(b) $P(\text{Type II error}) = P(X > 1 \text{ when the distribution is } p_1) = 0.5$.

6.99. **(a)** H_0: The patient is ill (or "the patient should see a doctor"); H_a: The patient is healthy (or "the patient should not see a doctor"). A Type I error means a false negative—clearing a patient who should be referred to a doctor. A Type II error is a false positive—sending a healthy patient to the doctor. **(b)** One might wish to lower the probability of a false negative so that most ill patients are treated, especially for serious diseases that require fast treatment. On the other hand, if resources (for example, money or medical personnel) are limited, or for less serious health problems, lowering the probability of false positives might be desirable.

Note: *For (a), there is no clear choice for which should be the null hypothesis in this case. Because the subjects have no specific medical complaints,* H_0 *might be "the patient is healthy." This choice may also be affected by the factors considered in part (b).*

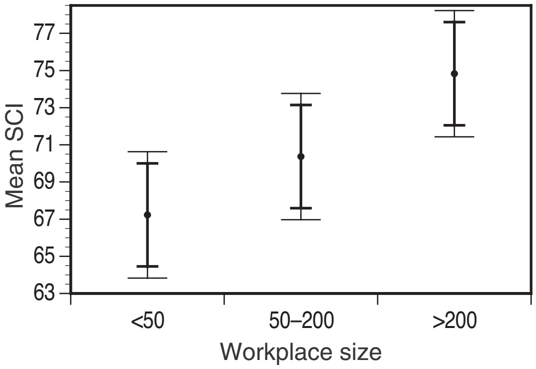
6.100. **(a)** Because all standard deviations and sample sizes are the same, the margin of error for all intervals is $(1.96)(4.5)/\sqrt{750} \doteq 0.3221$ months. The confidence intervals are listed in the table below. **(b)** Plot below. **(c)** The mean number of months of full-time employment shows a steady increase over time, at an average rate of about 0.8 months per year.

Age (years)	Average months worked
18	2.58 to 3.22
19	3.88 to 4.52
20	4.68 to 5.32
21	4.98 to 5.62
22	6.08 to 6.72
23	7.08 to 7.72
24	8.18 to 8.82
25	8.58 to 9.22
26	8.98 to 9.62



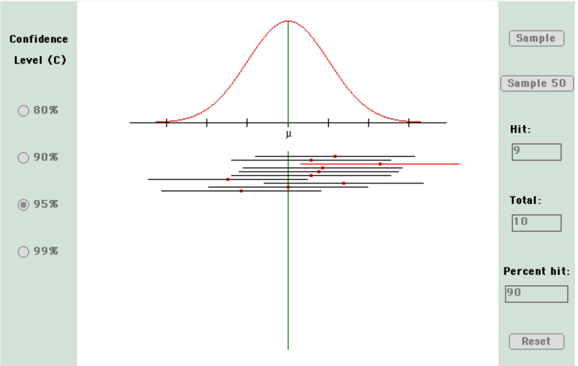
6.101. **(a)** Because all standard deviations and sample sizes are the same, the margin of error for all intervals is $(1.96)(19)/\sqrt{180} \doteq 2.7757$. The confidence intervals are listed in the table below. **(b)** The plot below shows the error bars for the confidence intervals of (a), and also for part (c). The limits for (a) are the thicker lines which do not extend as far above and below the mean. **(c)** With $z^* = 2.40$, the margin of error for all intervals is $(2.40)(19)/\sqrt{180} \doteq 3.3988$. The confidence intervals are listed in the table below and are shown in the plot (the thinner lines with the wider dashes). **(d)** When we use $z^* = 2.40$ to adjust for the fact that we are making three "simultaneous" confidence intervals, the margin of error is larger, so the intervals overlap more.

Workplace size	Mean SCI
< 50	64.45 to 70.01
50–200	67.59 to 73.15
> 200	72.05 to 77.61
< 50	63.83 to 70.63
50–200	66.97 to 73.77
> 200	71.43 to 78.23



6.102. Shown below is a sample screenshot from the applet and an example of what the resulting plot might look like. Most students (99.7% of them) should find that their final proportion is between 0.90 and 1; 90% will have a proportion between 0.925 and 0.975.

Note: *For each n (number of intervals), the number of "hits" would have a binomial distribution with* $p = 0.95$, *but these counts would not be independent; for example, if we knew there were 28 hits after 30 tries, we would know that there could be no more than 38 after 40 tries.*



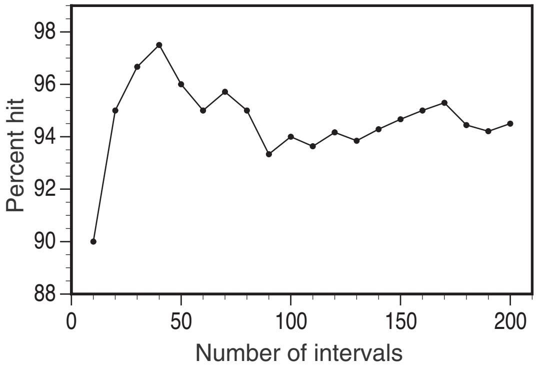
Confidence Level (C)

○ 80%
○ 90%
● 95%
○ 99%

Hit: 9

Total: 10

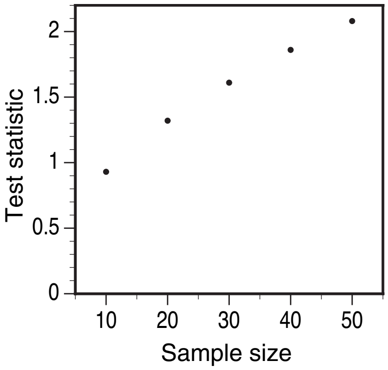
Percent hit: 90

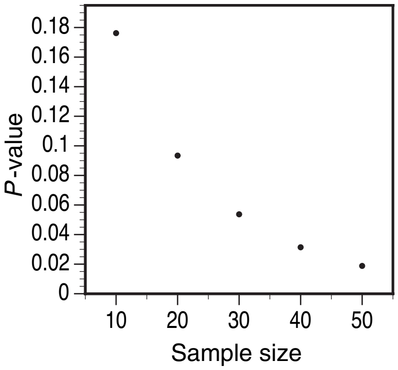


6.103. A sample screenshot and example plot are not shown but would be similar to those shown above for the previous exercise. Most students (99.4% of them) should find that their final proportion is between 0.84 and 0.96; 85% will have a proportion between 0.87 and 0.93.

6.104. For $n = 10$, $z = \frac{5-0}{17/\sqrt{10}} \doteq 0.93$, for which $P = 0.1762$. For the other sample sizes, the computations are similar; the resulting table and graphs are shown below. We see that sample size increases the value of the test statistic (assuming the mean is the same), which in turn decreases the size of the P-value.

n	z	P
10	0.93	0.1762
20	1.32	0.0934
30	1.61	0.0537
40	1.86	0.0314
50	2.08	0.0188





6.105. **(a)** $\bar{x} = 5.\overline{3}$ mg/dl, so $\bar{x} \pm 1.960\sigma/\sqrt{6}$ is 4.6132 to 6.0534 mg/dl. **(b)** To test H_0: $\mu = 4.8$ mg/dl vs. H_a: $\mu > 4.8$ mg/dl, we compute $z = \frac{\bar{x}-4.8}{0.9/\sqrt{6}} \doteq 1.45$ and $P \doteq 0.0735$. This is not strong enough to reject H_0.

Note: *The confidence interval in (a) would allow us to say without further computation that, against a two-sided alternative, we would have* $P > 0.05$. *Because we have a one-sided alternative, we could conclude from the confidence interval that* $P > 0.025$, *but that is not enough information to draw a conclusion.*

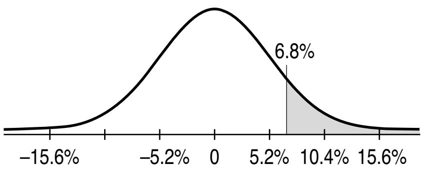
6.106. **(a)** The 95% confidence interval is $140 \pm (1.96)(8/\sqrt{16}) = 140 \pm 3.92 = 136.08$ to 143.92 mg/g. **(b)** Our hypotheses are H_0: $\mu = 135$ mg/g vs. H_a: $\mu > 135$ mg/g. The test statistic is $z = \frac{140-135}{8/\sqrt{16}} = 2.50$, so the P-value is $P = P(Z > 2.50) = 0.0062$. This is strong evidence against H_0; we conclude that the mean cellulose content is higher than 135 mg/g. **(c)** We must assume that the 16 cuttings in our sample are an SRS. Because our sample is not too large, the population should be normally distributed, or at least not extremely nonnormal.

6.107. **(a)** The stemplot is reasonably symmetric for such a small sample. **(b)** $\bar{x} = 30.6\ \mu\text{g/l}$; $30.6 \pm (1.96)(7/\sqrt{10})$ gives 26.2614 to 34.9386 μg/l. **(c)** H_0: $\mu = 25\ \mu\text{g/l}$; H_a: $\mu > 25\ \mu\text{g/l}$. $z = \frac{30.6-25}{7/\sqrt{10}} \doteq 2.53$, so $P = 0.0057$. (We knew from (b) that it had to be smaller than 0.025). This is fairly strong evidence against H_0; the beginners' mean threshold is higher than 25 μg/l.

2	04
2	5
3	0123
3	56
4	0

6.108. **(a)** The intended population is probably "the American public"; the population that was actually sampled was "citizens of Indianapolis (with listed phone numbers)." **(b)** Take $\bar{x} \pm 1.96s/\sqrt{201}$. Food stores: 15.22 to 22.12. Mass merchandisers: 27.77 to 36.99. Pharmacies: 43.68 to 53.52. **(c)** The confidence intervals do not overlap at all; in particular, the *lower* confidence limit of the rating for pharmacies is higher than the *upper* confidence limit for the other stores. This indicates that the pharmacies are *really* rated higher.

6.109. **(a)** Under H_0, $\bar{x}$ has a $N(0\%, 53\%/\sqrt{104}) \doteq N(0\%, 5.1971\%)$ distribution. **(b)** $z = \frac{6.8-0}{53/\sqrt{104}} \doteq 1.31$, so $P = P(Z > 1.31) = 0.0951$. **(c)** This is not significant at $\alpha = 0.05$. The study gives *some* evidence of increased compensation, but it is not very strong; similar results would happen almost 10% of the time just by chance.



6.110. No: "Significant at $\alpha = 0.05$" *does* mean that the null hypothesis is unlikely, but only in the sense that the evidence (from the sample) would not occur very often if H_0 were true. There is no probability associated with H_0 [unless one is a Bayesian statistician]; it is either true or it is not.

6.111. Yes. That's the heart of why we care about statistical significance. Significance tests allow us to discriminate between random differences ("chance variation") that might occur when the null hypothesis is true, and differences that are unlikely to occur when H_0 is true.

6.112. For each sample, find $\bar{x}$, then take $\bar{x} \pm 1.96(5/\sqrt{12}) = \bar{x} \pm 2.829$.

We "expect" to see that 95 of the 100 intervals will include 20 (the true value of μ); binomial computations show that (about 99% of the time) 90 or more of the 100 intervals will include 20.

6.113. For each sample, find $\bar{x}$, then compute $z = \frac{\bar{x}-20}{5/\sqrt{12}}$. Choose a significance level α and the appropriate cutoff point—for example, with $\alpha = 0.10$, reject H_0 if $|z| > 1.645$; with $\alpha = 0.05$, reject H_0 if $|z| > 1.96$.

If, for example, $\alpha = 0.05$, we "expect" to reject H_0 (that is, make the wrong decision) only 5 of the 100 times.

6.114. For each sample, find $\bar{x}$, then compute $z = \frac{\bar{x}-18}{5/\sqrt{12}}$. Choose a significance level α and the appropriate cutoff point (z^*)—for example, with $\alpha = 0.10$, reject H_0 if $|z| > 1.645$; with $\alpha = 0.05$, reject H_0 if $|z| > 1.96$.

Because the true mean is 20, $Z = \frac{\bar{x}-20}{5/\sqrt{12}}$ has a $N(0,1)$ distribution, so the probability that we will accept H_0 is $P\left(-z^* < \frac{\bar{x}-18}{5/\sqrt{12}} < z^*\right) = P(-z^* < Z + 1.3856 < z^*) = P(-1.3856 - z^* < Z < -1.3856 + z^*)$. If $\alpha = 0.10$ ($z^* = 1.645$), this probability is $P(-3.03 < Z < 0.26) = 0.6014$; if $\alpha = 0.05$ ($z^* = 1.96$), this probability is $P(-3.35 < Z < 0.57) = 0.7153$. For smaller α, the probability will be larger. Thus we "expect" to (wrongly) accept H_0 a majority of the time, and correctly reject H_0 about 40% of the time or less. (The probability of rejecting H_0 is essentially the power of the test against the alternative $\mu = 20$.)

6.116. The listing below shows the statements ordered from largest to smallest means (that is, with the highest level of agreement at the top) for both age groups. Lines are drawn in the list to show sets of similar complaints for both groups; for example, the top complaint is the same for both groups, and the next set is the same (although the order is different). Restauranteurs wishing to make their businesses attractive to older customers should probably pay the most attention to the top of this list. (Note that those statements which are in the bottom half of the list have ratings of just over 3 or less. On a 1-to-5 scale, "3" is a neutral response.)

Ages 50–64	Ages 65–79
I would rather be served ...	I would rather be served ...
I would rather pay the server ...	Tables are too close together.
Tables are too close together.	Print size is not large enough.
Print size is not large enough.	Background music is often too loud.
Most restaurants are too noisy.	I would rather pay the server ...
Background music is often too loud.	Most restaurants are too noisy.
Restaurants are too smoky.	Tables are too small.
Service is too slow.	Restaurants are too smoky.
Tables are too small.	Service is too slow.
Glare makes menus difficult to read.	Glare makes menus difficult to read.
Most restaurants are too dark.	It is difficult to hear the service staff.
It is difficult to hear the service staff.	Most restaurants are too dark.
Colors make menus difficult to read.	Colors make menus difficult to read.

In comparing the means, we use a two-sided alternative, because we have no prior expectation that the differences should be either positive or negative. As the text says, the z statistics are found by taking the differences between the means and dividing by 0.08; for example, the first is $z = \frac{2.75-2.93}{0.08} \doteq -2.25$, for which $P = 0.0244$. Those with P-values less than 0.05 are marked with an asterisk. Because we are performing 13 tests, a P-value below 0.05 should not be viewed as establishing significance. Using the Bonferroni approach (see Exercise 6.86), we would be willing to declare as significant any differences with $P < 0.05/13 \doteq 0.0038$; these are marked with two asterisks. Given the clustering of complaints noted in the list above, it is not too surprising that few differences are significant. Two of these three differences suggest that the older group may have more concerns about noise level and hearing.

	Statement	z	P	
Ambience:	Most restaurants are too dark.	−2.25	0.0244	*
	Most restaurants are too noisy.	−1.25	0.2113	
	Background music is often too loud.	−3.50	0.0005	**
	Restaurants are too smoky.	0.62	0.5353	
	Tables are too small.	−2.38	0.0173	*
	Tables are too close together.	−0.25	0.8026	
Menu design:	Print size is not large enough.	−1.13	0.2585	
	Glare makes menus difficult to read.	−2.50	0.0124	*
	Colors of menus make them difficult to read.	−2.38	0.0173	*
Service:	It is difficult to hear the service staff.	−4.38	0.0000	**
	I would rather be served than serve myself.	1.13	0.2585	
	I would rather pay the server than a cashier.	5.00	0.0000	**
	Service is too slow.	0.37	0.7114	