Chapter 6

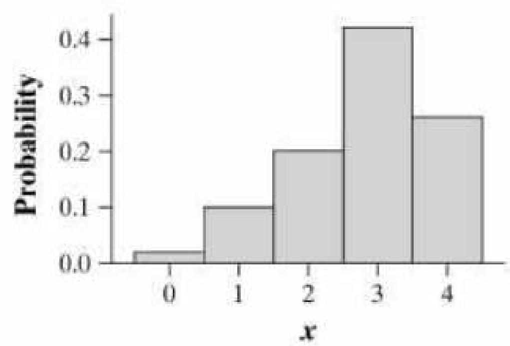
Section 6.1

Check Your Understanding, page 344:

1. We are looking for the probability that the student gets either an A or a B. This probability is $0.42 + 0.26 = 0.68$.

2. We are looking for $P(X < 2) = 0.02 + 0.10 = 0.12$.

3. This histogram is left skewed. This means that higher grades are more likely, but that there are a few lower grades.



Check Your Understanding, page 349:

1. $\mu_X = 0(0.3) + 1(0.4) + 2(0.2) + 3(0.1) = 1.1$. The long-run average, over many Friday mornings, will be about 1.1 cars sold.

2. $\sigma_X^2 = (0-1.1)^2(0.3) + (1-1.1)^2(0.4) + (2-1.1)^2(0.2) + (3-1.1)^2(0.1) = 0.89$. So $\sigma_X = 0.943$. On average, the number of cars sold on a randomly selected Friday will differ from the mean (1.1) by 0.943 cars sold.

Exercises, page 353:

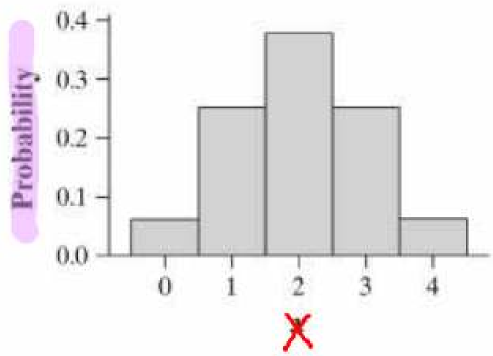
6.1 (a) If you toss a coin 4 times, the sample space is {HHHH, HHHT, HHTH, HTHH, THHH, HHTT, HTHT, HTTH, THHT, THTH, TTHH, HTTT, THTT, TTHT, TTTH, TTTT}. All of these outcomes are equally likely and so have probability $\frac{1}{16} = 0.0625$. To find the probability of X taking on any specific number, count the number of outcomes with exactly this number of heads and multiply by 0.0625. For example, there are 4 ways to get exactly 3 heads so $P(X = 3) = 4(0.0625) = 0.25$. This leads to the following distribution:

a)

X ~~Value~~	0	1	2	3	4
Probability	0.0625	0.25	0.375	0.25	0.0625
	1	4	6	4	1

X = # of heads you get

(b) The histogram shows that this distribution is symmetric with a center at 2.



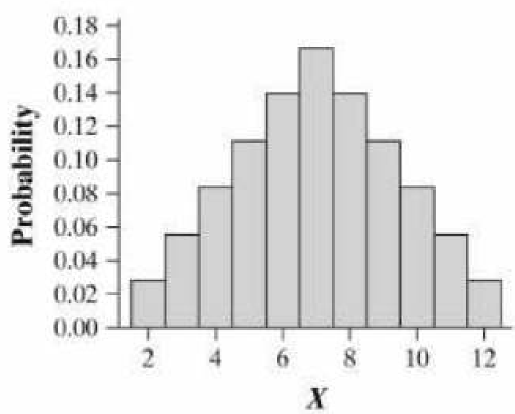
By Hand

(c) $P(X \le 3) = 1 - P(X = 4) = 1 - 0.0625 = 0.9375$. There is a 93.75% chance that you will get three or fewer heads on 4 tosses of a fair coin.

6.2 (a) If we roll two 6-sided dice, the sample space is {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}. All of these are equally likely, so the probability of any one outcome is $\frac{1}{36}$. To find the probability of X taking on any specific number, count the number of outcomes where the sum of the dice is exactly this number and multiply by 0.0278. For example, there are 4 ways to get a sum of 5 (the outcomes (1,4), (2,3), (3,2), and (4,1)) so $P(X = 3) = 4\left(\frac{1}{36}\right) = \frac{4}{36}$. This leads to the following distribution:

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

(b) The histogram shows that the distribution is symmetric about a center of 7.

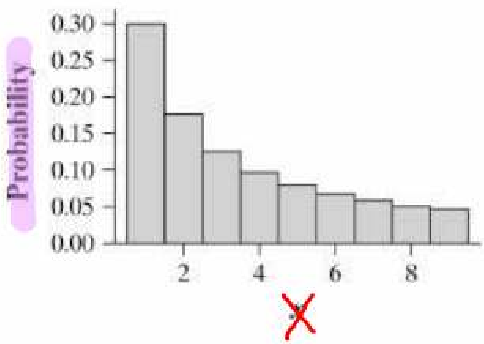


(c) $P(T \ge 5) = 1 - P(T \le 4) = 1 - \left(\frac{1}{36} + \frac{2}{36} + \frac{3}{36}\right) = 1 - \frac{6}{36} = 1 - \frac{1}{6} = \frac{5}{6}$. This means that about five-sixths of the time, when you roll a pair of 6-sided dice, you will have a sum of 5 or more.

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

6.4 (a) "Plays with at most two toys" is the event $\{X \leq 2\}$ or $\{X < 3\}$. $P(X \leq 2) = (X = 0) + P(X = 1) + P(X = 2) = 0.03 + 0.16 + 0.30 = 0.49$. (b) The event $\{X > 3\}$ is "the child plays with more than three toys." $P(X>3) = P(X=4) + P(X=5) = 0.17 + 0.11 = 0.28$. $P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.49 = 0.51$.

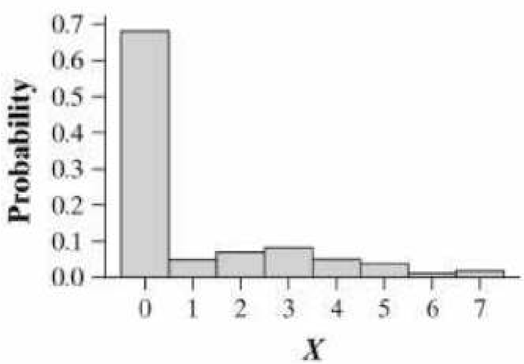
6.5 (a) All of the probabilities are between 0 and 1 and they sum to 1 so this is a legitimate probability distribution. (b) This is a right skewed distribution with the largest amount of probability on the digit 1.

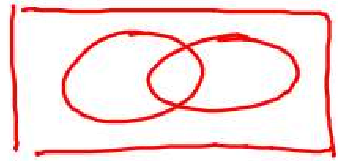


By Hand

(c) The event $\{X \geq 6\}$ is the event that "the first digit in a randomly chosen record is a 6 or higher." $P(X \geq 6) = P(X=6) + P(X=7) + P(X=8) + P(X=9) = 0.067 + 0.058 + 0.051 + 0.046 = 0.222$. (d) The event that "the first digit is at most 5" is the event $\{X \leq 5\}$.
$P(X \leq 5) = 1 - P(X \geq 6) = 1 - 0.222 = 0.778$.

6.6 (a) All of the probabilities are between 0 and 1 and they sum to 1 so this is a legitimate probability distribution. (b) Most people did not work out in the last 7 days. The distribution is skewed to the right with a peak at 0 days.





(c) The event {Y<7} is the event that the person "did not work out all 7 days." $P(Y<7)=1-P(Y=7)=1-0.02=0.98$. (d) The event "worked out at least once" is the event $\{Y \geq 1\}$. $P(Y \geq 1)=1-P(Y=0)=1-0.68=0.32$.

6.7 (a) The outcomes that make up the event A are {7, 8, 9}. $P(A)=P(X=7)+P(X=8)+P(X=9)=0.058+0.051+0.046=0.155$. (b) The outcomes that make up the event B are {1, 3, 5, 7, 9}. $P(B)=P(X=1)+P(X=3)+P(X=5)+P(X=7)+P(X=9)$ $=0.301+0.125+0.079+0.058+0.046=0.609$.
(c) The outcomes that make up the event "A or B" are {1, 3, 5, 7, 8, 9}. $P(A \text{ or } B)=P(A \cup B)=P(A)+P(B)-P(A \cap B)=0.155+0.609-(0.058+0.046)=0.66$. This is not the same as $P(A)+P(B)$ because A and B are not mutually exclusive.

6.8 (a) The outcomes that make up the event A are {1, 2, 3, 4, 5, 6, 7}. From exercise 6.6d, $P(X \geq 1)=0.32$. (b) The outcomes that make up the event B are {0, 1, 2, 3, 4}. $P(B)=0.68+0.05+0.07+0.08+0.05=0.93$. (c) The event "$A$ and B" has the outcomes {1, 2, 3, 4}. So $P(A \text{ and } B)=0.05+0.07+0.08+0.05=0.25$. Note that $P(A)P(B)=0.32(0.93)=0.2976$. These last two probabilities are not equal because A and B are not independent.

6.9 (a) The payoff is either \$0, with a probability of 0.75, or \$3, with a probability of 0.25.

Value	\$0	\$3
Probability	0.75	0.25

(b) For each \$1 bet, the mean payoff is μ_X = (\$0)(0.75) + (\$3)(0.25) = \$0.75. This means that for every \$1 the player bets, he only gets \$0.75 back. In other words, he loses \$0.25 on each bet, on average.

6.10 The company earns \$300, with a probability of 0.9998, and earns \$-199,700 with probability 0.0002.

Value	\$300	\$-199,700
Probability	0.9998	0.0002

(b) The expected value of Y is $\mu_Y=(\$300)(0.9998)+(\$-199,700)(0.0002)=\$260$. This means that, on average, the company gains \$260 per policy.

6.11 $\mu_X=0(0.1)+1(0.2)+2(0.3)+3(0.3)+4(0.1)=2.1$. On average, undergraduates make 2.1 nonword errors per 250-word essay.

6.12 $\mu_X=0(0.03)+1(0.16)+2(0.30)+3(0.23)+4(0.17)+5(0.11)=2.68$. On average, when children are given 5 toys to play with, they play with 2.68.
Do Not Round

6.13 (a) The mean of the random variable Y is located at 5 because this distribution is symmetric and 5 is located at the center. (b) The expected value of X, the first digit following Benford's law, is $\mu_X=1(0.301)+2(0.176)+3(0.125)+4(0.097)+5(0.079)+6(0.067)+7(0.058)+8(0.051)+9(0.046)$ $=3.441$. The average of first digits following Benford's law is 3.441. To detect a fake expense report, compute the sample mean of the first digits and see if it is near 5 or near 3.441. (c) Under the equally likely assumption, $P(Y>6)=0.111+0.111+0.111=0.333$. Under Benford's law $P(X>6)=0.058+0.051+0.046=0.155$. So on a fake expense report, we would expect 33% of numbers to start with digits higher than 6 whereas on a real expense report, only about 15% would start with digits higher than 6. So when looking at a suspect report, find the percentage of figures that start with numbers higher than 6. If that percentage is closer to 33% than to 15%, it is probably fake.

6.14 (a) and (b)

Death age	21	22	23	24	25	26/more
Profit	-\$99,750	-\$99,500	**-\$99,250**	**-\$99,000**	**-\$98,750**	**\$1,250**
Probability	0.00183	0.00186	0.00189	0.00191	0.00193	**0.99058**

(c) $\mu_X = (-\$99,750)(0.00183) + (-\$99,500)(0.00186) + (-\$99,250)(0.00189)$ $+(-\$99,000)(0.00191) + (-\$98,750)(0.00193) + (\$1,250)(0.99058) = \$303.35$. The company makes an average of \$303.35 per life insurance policy.

6.15 $\sigma_X^2 = (0-2.1)^2(0.1) + (1-2.1)^2(0.2) + (2-2.1)^2(0.3) + (3-2.1)^2(0.3) + (4-2.1)^2(0.1) = 1.29$. So $\sigma_X = \sqrt{1.29} = 1.1358$. This means that, on average, the number of nonword errors in a randomly selected essay will differ from the mean (2.1) by 1.14 words.

6.16 .215472

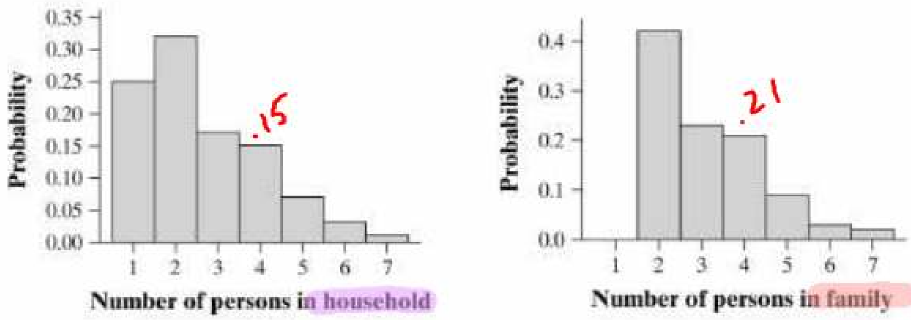
$\sigma_X^2 = (0-2.68)^2(0.03) + (1-2.68)^2(0.16) + (2-2.68)^2(0.3) + (3-2.68)^2(0.23) + (4-2.68)^2(0.17)$ $+(5-2.68)^2(0.11) = 1.7176$. So $\sigma_X = \sqrt{1.7176} = 1.3106$. This means that, on average, the number of toys a randomly selected child will play with will differ from the mean (2.68) by 1.31 toys.

6.17 (a) $\sigma_Y^2 = (1-5)^2(0.111) + (2-5)^2(0.111) + \ldots + (9-5)^2(0.111) = 6.667$. So $\sigma_Y = 2.58$.
(b) $\sigma_X^2 = (1-3.441)^2(0.301) + (2-3.441)^2(0.176) + \ldots + (9-3.441)^2(0.046) = 6.0605$. So $\sigma_X = 2.4618$. This would not be the best way to tell the difference between a fake and a real expense report because the standard deviations are not too different from one another.

6.18 (a) The mean μ_X of the company's "winnings" (premiums) and their "losses" (insurance claims) is about \$303.35. 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 from many thousands of 21-year-old men. In the long run, the insurance company can expect to make \$303.35 per insurance policy. The insurance company is relying on the Law of Large Numbers. (b) $\sigma_X^2 = (-\$99,750 - 303.35)^2(0.00183) + \ldots + (\$1,250 - 303.35)^2(0.97321) = \$94,236,826.64$. So $\sigma_X = \$9,707.57$.

6.19 (a) The probability histograms are shown below. The distribution of the number of rooms is roughly symmetric for owners (graph on the left) and skewed to the right for renters (graph on the right). The center is slightly over 6 units for owners and slightly over 4 for renters. Overall, renter-occupied units tend to have fewer rooms than owner-occupied units.

20a) Both skewed right; P(x=1)=0 for families by definition; family sizes tend to be larger...



(b) The means are: $\mu_X = 1(0.25)+2(0.32)+3(0.17)+4(0.15)+5(0.07)+6(0.03)+7(0.01) = 2.6$ people for a household and $\mu_Y = 1(0)+2(0.42)+3(0.23)+4(0.21)+5(0.09)+6(0.03)+7(0.02) = 3.14$ people for a family. The family distribution has a slightly larger mean than the household distribution, matching the observation in part (a) that family sizes tend to be larger than household sizes. (c) The standard deviations are: $\sigma_X^2 = (1-2.6)^2\times0.25 + (2-2.6)^2\times 0.32 + (3-2.6)^2\times0.17 + (4-2.6)^2\times0.15 + (5-2.6)^2\times0.07 + (6-2.6)^2\times0.03 + (7-2.6)^2\times0.01 = 2.02$, and $\sigma_X = \sqrt{2.02} = 1.421$ people for a household and $\sigma_Y^2 = (1-3.14)^2(0) + (2-3.14)^2(0.42) + (3-3.14)^2(0.23) + (4-3.14)^2(0.21) + (5-3.14)^2(0.09) + (6-3.14)^2(0.03) + (7-3.14)^2(0.02) = 1.5604$, and $\sigma_Y = \sqrt{1.5604} = 1.249$ people for a family. The standard deviation for households is only slightly larger, mainly due to the fact that a household can have only 1 person.

6.21 (a) $P(X>0.49)=0.51$. (b) $P(X\geq 0.49)=0.51$. *Note: (a) and (b) are the same because there is no area under the curve at any one particular point.* (c) $P(0.19\leq X<0.37 \text{ or } 0.84<X\leq 1.27) = P(0.19\leq X<0.37)+P(0.84<X\leq1.27)$. Note that $P(0.19\leq X<0.37)=0.37-0.19=0.18$. Note also that X cannot be bigger than 1, so $P(0.84<X\leq1.27)=P(0.84<X\leq1)=1-0.84=0.16$. Therefore $P(0.19\leq X<0.37 \text{ or } 0.84<X\leq1.27)=0.18+0.16=0.34$.

6.22 (a) $P(X\leq 0.4)=0.4$. (b) $P(X<0.4)=0.4$. *Note: (a) and (b) are the same because there is no area under the curve at any one particular point.* (c) $P(0.1<Y\leq0.15 \text{ or } 0.77\leq Y<0.88) = P(0.1<Y\leq0.15)+P(0.77\leq Y<0.88)$ $=(0.15-0.1)+(0.88-0.77)=0.05+0.11=0.16$.

6.23 *State*: What is the probability that a randomly chosen student scores a 9 or better on the ITBS? *Plan*: The score X of the randomly chosen student has the $N(6.8, 1.6)$ distribution. We want to find $P(X\geq 9)$. We'll standardize the scores and find the area shaded in the Normal curve.

6.32 No we cannot conclude that chess causes an increase in reading scores. We do not have a control group that did not participate in the chess program. This means that we have no comparison group. It may be that children of this age naturally improve their reading scores anyway and that the chess program had nothing to do with their improvement.

6.33 The equation of the linear regression model is: predicted post-test $= 17.897 + 0.78301(\text{pretest})$.

6.34 This line is appropriate. The residual plot does not show any patterns. The strength of the fit is only moderate with an $r^2 = 0.558$. The line does not do a great job of predicting any individual score as $s = 12.55$ so in using the line, we expect to be off by an average of 12.55 points. Given that the mean number of points is roughly 63, this is a lot to be off by.

Section 6.2

Check Your Understanding, page 362:

1. $\mu_Y = 500\mu_X = 500(1.1) = \$550$. $\sigma_Y = 500\sigma_X = 500(0.943) = \$471.5$
2. $\mu_T = \mu_Y - 75 = 550 - 75 = \$475$. $\sigma_T = \sigma_Y = \$471.5$.

Check Your Understanding, page 370:

1. $\mu_T = \mu_X + \mu_Y = 1.1 + 0.7 = 1.8$. On average, this dealership sells or leases 1.8 cars in the first hour of business on Fridays.
2. Assuming that X and Y are independent, $\sigma_T^2 = \sigma_X^2 + \sigma_Y^2 = (0.943)^2 + (0.64)^2 = 1.2988$ so $\sigma_T = \sqrt{1.2988} = 1.14$.
3. The total bonus is $B = 500X + 300Y$. This means that $\mu_B = 500\mu_X + 300\mu_Y = 500(1.1) + 300(0.7) = \$760$. To find the standard deviation, we must again assume that X and Y are independent and we must first calculate the variance. $\sigma_B^2 = (500)^2\sigma_X^2 + (300)^2\sigma_Y^2 = 250{,}000(0.943)^2 + 90{,}000(0.64)^2 = 259{,}176.25$. Therefore $\sigma_B = \sqrt{259{,}176.25} = \$509.09$.

Check Your Understanding, page 372:

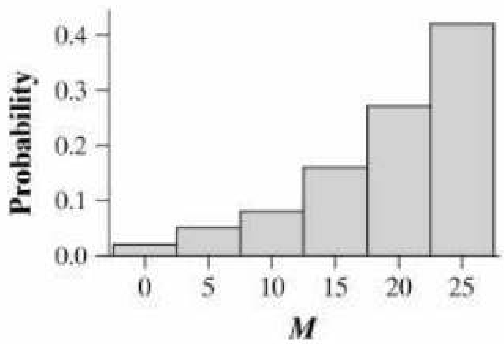
1. $\mu_D = \mu_X - \mu_Y = 1.1 - 0.7 = 0.4$. On average, this dealership sells 0.4 cars more than it leases during the first hour of business on Fridays.
2. $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 = (0.943)^2 + (0.64)^2 = 1.2998$ so $\sigma_D = \sqrt{1.2998} = 1.14$.
3. $B = 500X - 300Y$. This means that $\mu_B = 500\mu_X - 300\mu_Y = 500(1.1) - 300(0.7) = \$340$. To find the standard deviation, we must again assume that X and Y are independent and we must first calculate the variance. $\sigma_B^2 = (500)^2\sigma_X^2 + (300)^2\sigma_Y^2 = 250{,}000(0.943)^2 + 90{,}000(0.64)^2 = 259{,}176.25$. Therefore $\sigma_B = \sqrt{259{,}176.25} = \$509.09$.

Exercises, page 378:

6.35 The relationship between the length in centimeters and the length in inches is $Y = 2.54X$. So $\mu_Y = 2.54\mu_X = 2.54(1.2) = 3.048$ cm and $\sigma_Y = 2.54\sigma_X = 2.54(0.25) = 0.635$ cm.

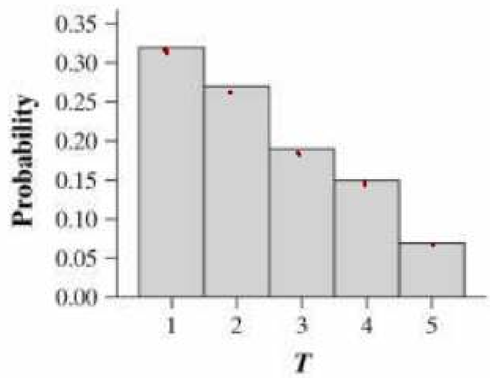
6.36 The relationship between heights in inches and heights in feet is $J = 12H$. So $\mu_J = 12\mu_H = 12(5.8) = 69.6$ in and $\sigma_J = 12\sigma_H = 12(0.24) = 2.88$ in.

6.37 (a) This graph is skewed to the left. Most of the time, the ferry makes \$20 or \$25.



(b) $\mu_M = 5\mu_X = 5(3.87) = \$19.35$. The ferry makes \$19.35 per trip on average. (c) $\sigma_M = 5\sigma_X = 5(1.29) = \$6.45$. The individual amounts made on the ferry trips will vary by about \$6.45 from the mean (\$19.35) on average.

6.38 (a) This distribution is skewed to the right. Ana is more likely to get 1 or 2 tickets than 4 or 5.



(b) $\mu_T = \frac{1}{10}\mu_X = \frac{1}{10}(23.8) = 2.38$ tickets. On average, Ana will receive 2.38 tickets on a single roll. (c) $\sigma_T = \frac{1}{10}\sigma_X = \frac{1}{10}(12.63) = 1.263$ tickets. In individual rolls, the number of tickets that Ana wins will vary by 1.263 from the mean (2.38) on average.

6.39 (a) The score on the test G is related to the number of questions X by the equation $G = 10X$. This means that $\mu_G = 10\mu_X = 10(7.6) = 76$. (b) $\sigma_G = 10\sigma_X = 10(1.32) = 13.2$. (c) Since the variance of G is the square of the standard deviation $\sigma_G^2 = (10\sigma_X)^2 = 100\sigma_X^2$. In other words, we need to multiply the variance of X by the square of the constant 10 to get the variance of G.

6.40 (a) The score on the test G is related to the number of questions X by the equation $G = 10X$. This means that $\text{median}_G = 10\text{median}_X = 10(8.5) = 85$. (b) $IQR_G = 10(IQR_X) = 10(1) = 10$. (c) The shape of

G's distribution will be the same as the shape of X's. Since the distance between the median and the minimum is much bigger than the distance between the median and the maximum, this distribution is skewed to the left.

6.41 (a) The mean of Y is 20 less than the mean of M. In other words $\mu_Y = \mu_M - 20$. Thus, the total mean profit per trip is \$20 less than the amount of money collected. (b) The standard deviation of Y is the same as the standard deviation of M. There is the same amount of variability in both the profit and the amount of money collected.

6.42 (a) $\mu_X = 0(0.999) + 500(0.001) = \$0.50$. $\sigma_X^2 = (0-0.50)^2(0.999) + (500-0.50)^2(0.001) = 249.75$ so $\sigma_X = \sqrt{249.75} = \$15.80$. (b) $\mu_W = \mu_X - 1 = 0.50 - 1 = -\$0.50$. $\sigma_W = \sigma_X = \$15.80$. On average, when playing this game, people will lose \$0.50. Individual outcomes will vary from this amount by \$15.80 on average.

6.43 $\mu_Y = 6\mu_X - 20 = 6(3.87) - 20 = \$3.22$. $\sigma_Y = 6\sigma_X = 6(1.29) = \$7.74$.

6.44 The mean and standard deviation of Y are $\mu_Y = 0.9\mu_X - 0.2 = 0.9(\$3) - 0.2 = \boxed{\$2.5 \text{ million}}$ and $\cancel{\sigma_Y = \sqrt{0.9^2\sigma_X^2} = \sqrt{0.9^2(6.35)} = \$2.2679 \text{ million}}$. $\sigma_Y = .9(2.52) = \$2.2679$ million

6.45 (a) $\mu_Y = \frac{9}{5}\mu_T + 32 = \frac{9}{5}(8.5) + 32 = 47.3$ degrees Fahrenheit. $\sigma_Y = \frac{9}{5}\sigma_T = \frac{9}{5}(2.25) = 4.05$ degrees Fahrenheit. (b) $P(Y < 40) = P\left(Z < \frac{40-47.3}{4.05}\right) = P(Z < -1.80) = 0.0359$.

6.46 (a) $\mu_Y = 28.35\mu_X - 273.01 = 28.35(9.70) - 273.01 = 1.985$ grams. $\sigma_Y = 28.35\sigma_X = 28.35(0.03) = 0.8505$ grams. (b) $P(Y \geq 3) = P\left(Z \geq \frac{3-1.985}{0.8505}\right) = P(Z \geq 1.19) = 0.1170$.

6.47 (a) The possible values of T are $\{1+2=3, 1+4=5, 2+2=4, 2+4=6, 5+2=7, 5+4=9\}$. Since X and Y are independent, $P(X = x \text{ and } Y = y) = P(X = x)P(Y = y)$. So $P(T=3) = P(X=1 \text{ and } Y=2) = P(X=1)P(Y=2) = 0.2(0.7) = 0.14$. The rest of the probabilities are found the same way. The probability distribution for T is

Value	3	4	5	6	7	9
Probability	0.14	0.35	0.06	0.15	0.21	0.09

(b) From the distribution of T we get $\mu_T = 3(0.14) + 4(0.35) + 5(0.06) + 6(0.15) + 7(0.21) + 9(0.09) = 5.3$. Now using the relationship between T, X and Y we get $\mu_T = \mu_X + \mu_Y = 2.7 + 2.6 = 5.3$ which is the same. (c) From the distribution of T we get $\sigma_T^2 = (3-5.3)^2(0.14) + (4-5.3)^2(0.35) + ... + (9-5.3)^2(0.09) = 3.25$. Now using the relationship between T, X and Y we get $\sigma_T^2 = \sigma_X^2 + \sigma_Y^2 = (1.55)^2 + (0.917)^2 = 3.24$ which is only different due to rounding error. Note that $\sigma_T = 1.80$ but $\sigma_X + \sigma_Y = 1.55 + 0.917 = 2.467$. These last two values are not the same.

55) $M_y - M_x = 2.001 - 2.000 = .001$ grams

$$\sigma_{y-x} = \sqrt{\sigma_y^2 + \sigma_x^2}$$

$$= \sqrt{(.001)^2 + (.002)^2}$$

$$= \sqrt{.000001 + .000004}$$

$$= \sqrt{.000005}$$

$= .0022$ grams $2.2E^{-3}$

$P\left(\frac{345-350}{3.7537} \le Z \le \frac{355-350}{3.7537}\right) = P(-1.332 \le Z \le 1.332) = 0.9086 - 0.0914 = 0.8172$ (Table A gives $0.9082 - 0.0918 = 0.8164$).

6.61 *State*: What is the probability that the difference between the NOX levels for two randomly selected cars is at least 0.80 g/mi? *Plan*: Since both cars have NOX levels that follow a $N(1.4, 0.3)$ distribution, the distribution of the difference $X - Y$ will be $N(0, \sqrt{0.3^2 + 0.3^2}) \approx N(0, 0.4243)$. Use this to find $P(X - Y > 0.8 \text{ or } X - Y < -0.8)$. *Do*:

$$P(X - Y > 0.8 \text{ or } X - Y < -0.8) = P(X - Y > 0.8) + P(X - Y < -0.8)$$

$$= P\left(Z > \frac{0.8-0}{0.4243}\right) + P\left(Z < \frac{-0.8-0}{0.4243}\right) = P(Z > 1.89) + P(Z < -1.89) = 0.0294 + 0.0294 = 0.0588.$$

Conclude: There is only about a 6% chance of randomly finding two cars with as big a difference in NOX levels as the attendant found.

6.62 *State*: What is the probability that the scores differ by 5 or more points? *Plan*: Let L and F denote the respective scores of Leona and Fred. The difference $L - F$ has a Normal distribution with mean $\mu_{L-F} = 24 - 24 = 0$ points and standard deviation $\sigma_{L-F} = \sqrt{2^2 + 2^2} \doteq 2.8284$ points. Use this to find $P(L - F < -5 \text{ or } L - F > 5)$. *Do*: $P(L - F < -5 \text{ or } L - F > 5) = P(L - F < -5) + P(L - F > 5)$

$$= P\left(Z < \frac{-5-0}{2.8284}\right) + P\left(Z > \frac{5-0}{2.8284}\right) = P(Z < -1.77) + P(Z > 1.77) = 0.0384 + 0.0384 = 0.0768.$$

Conclude: There is about an 8% chance that one of the two friends will have to buy the other one a pizza.

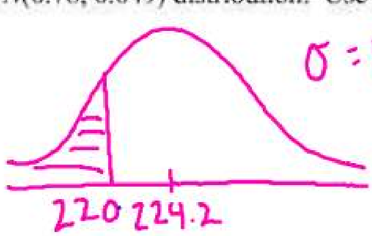
6.63 *State*: What is the probability that the total team swim time is less than 220 seconds. *Plan*: Let T stand for the total team swim time and X_1 be Wendy's time, X_2 be Jill's time, X_3 be Carmen's time, and X_4 be Latrice's time. Then $T = X_1 + X_2 + X_3 + X_4$. The mean and variance are $\mu_T = \mu_1 + \mu_2 + \mu_3 + \mu_4 = 55.2 + 58.0 + 56.3 + 54.7 = 224.2$ seconds and $\sigma_T^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \sigma_{X_3}^2 + \sigma_{X_4}^2 = (2.8)^2 + (3.0)^2 + (2.6)^2 + (2.7)^2 = 30.89$. This means that $\sigma_T = \sqrt{30.89} = 5.56$ seconds. Since each individual's time is approximately Normally distributed, that means that T is approximately Normally distributed. Putting this all together we have that T has a $N(224.2, 5.56)$ distribution. Use this to find $P(T < 220)$.

Do: $P(T < 220) = P\left(Z < \frac{220-224.2}{5.56}\right) = P(Z < -0.76) = 0.2236$. *Conclude*: There is approximately a 22% chance that the team's swim time will be less than 220 seconds. .2249 (normalcdf)

6.64 *State*: What is the probability that Ken uses more than 0.85 ounces of toothpaste if he brushes his teeth six times? *Plan*: Let T be the total amount of toothpaste and X_i be the amount used the ith time he brushes his teeth. $T = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$. Each of the X_i has a $N(0.13, 0.02)$ distribution and we will assume that the X_i are independent. So $\mu_T = \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6 = 6(0.13) = 0.78$ ounces and $\sigma_T^2 = \sum_{i=1}^{6} \sigma_{X_i}^2 = 6(0.02)^2 = 0.0024$. Since each of the X_i are Normally distributed, then so is T. Putting all of this together, T has a $N(0.78, 0.049)$ distribution. Use this to find $P(T > 0.85)$. *Do*:

$\sigma = 5.56$

normalcdf (0, 220, 224.2, 5.56)



probability of success is 0.20 for each trial. This is a binomial setting. Since X counts the number of successes, it is a binomial random variable with $n = 10$ and $p = 0.20$.

2. $P(X=3)=\binom{10}{3}(0.2)^3(0.8)^7=0.2013$. There is a 20% chance that Patti will answer exactly 3 questions correctly.

3. $P(X \geq 6)=1-P(X<6)=1-P(X \leq 5)$. Using binomcdf(10,0.2,5) we get 0.9936. So $P(X \geq 6)=1-0.9936=0.0064$. There is only a 0.64% chance that a student would pass so we would be quite surprised if Patti was able to pass.

Check Your Understanding, page 393:

1. $\mu_X = np = 10(0.20) = 2$. We would expect the average student to get 2 answers correct on such a quiz.

2. $\sigma_X = \sqrt{np(1-p)} = \sqrt{10(0.20)(0.80)} = 1.265$. We would expect individual students' scores to vary from a mean of 2 correct answers by an average of 1.265 correct answers.

3. Two standard deviations above the mean is $2+2(1.265)=4.53$. Since Patti can only score an integer number of questions correctly, we are asked to find $P(X \geq 5)=1-P(X \leq 4)$. Using binomcdf(10,0.2,4) we get 0.9672. So $P(X \geq 5)=1-0.9672=0.0328$. There is about a 3% chance that Patti will score more than two standard deviations above the mean.

Check Your Understanding, page 401:

1. Check the BITS: Binary? "Success" = roll doubles. "Failure" = don't roll doubles. Independent: Rolling dice is an independent process. Trials: We are counting trials up to and including the first success. Success? The probability of success is $\frac{1}{6}$ for each roll (there are 6 ways to get doubles out of a total of 36 possible rolls). This is a geometric setting. Since X is measuring the number of trials to get the first success, it is a geometric random variable with $p=\frac{1}{6}$.

2. $P(T=3)=\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)=0.1157$. There is about a 12% chance that you will get the first set of doubles on the third roll of the dice.

3.

$$P(T \leq 3)=P(T=1)+P(T=2)+P(T=3)=\frac{1}{6}+\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)+0.1157=0.1667+0.1389+0.1157=0.4213.$$

Exercises, page 403:

6.69 Binary? "Success" = seed germinates. "Failure" = seed does not germinate. Independent? Possibly although it is possible that if one seed does not germinate, it's more likely that others in the packet will not grow either. Number? 20 seeds. Success? The probability that each seed germinates is 85%, assuming the advertised percentage is true. Assuming independence does hold, this is a binomial setting and X would have a binomial distribution.

6.70 Binary? "Success" = name has more than 6 letters. "Failure" = name has 6 letters or less. Independent? Since we are selecting without replacement from a small number of students, the observations are not independent. Number? 4 names are drawn. Success? The probability that a student's name has more than 6 letters does not change from one draw to the next. This is a binomial setting and Y would have a binomial distribution.

6.71 Binary? "Success" = person is left-handed. "Failure" = person is right-handed. Independent? Since students are selected randomly, their handedness is independent. Number? There is not a fixed number of trials for this chance process since you continue until you find a left-handed student. Since the number of trials is not fixed, this is not a binomial setting and V is not a binomial random variable.

6.72 Binary? "Success" = person is left-handed. "Failure" = person is right-handed. Independent? Even though we are selecting students without replacement, 15 students is less than 10% of the total student body of most schools so the observations can be considered to be independent. Number? 15 students are selected. Success? The probability of left-handedness remains constant from one student to the next, about 0.10. This is a binomial setting and W has a binomial distribution with $n = 15$ and $p = 0.10$.

6.73 (a) A binomial distribution is *not* an appropriate choice for field goals made by the National Football League player, because given the different situations the kicker faces, his probability of success is likely to change from one attempt to another. (b) It would be reasonable to use a binomial distribution for free throws made by the NBA player because we have $n = 150$ attempts, presumably independent (or at least approximately so), with chance of success $p = 0.8$ each time.

6.74 (a) This is the binomial setting. We check the BINS. Binary? "Success" = reaching a live person. "Failure" = any other outcome. Independent? It is reasonable to believe that each call is independent of the others. Number? We have a fixed number of observations ($n = 15$). Success? Each randomly-dialed number has chance $p = 0.2$ of reaching a live person. (b) This is not a binomial setting because there are not a fixed number of attempts.

1

6.75 $P(X=4)=\binom{7}{4}(0.44)^4(0.56)^3=0.2304$. There is about a 23% chance that exactly 4 of the 7 people chosen have blood type O.

6.76 $P(Y=1)=\binom{10}{1}(0.05)(0.95)^9=0.3151$. There is about a 32% chance that exactly one of the 10 rhubarb plants will die before producing any rhubarb.

binomcdf (7, .44, 4)

2

6.77 Using technology, $P(X>4)=1-P(X\le 4)=1-0.8598=0.1402$. There is about a 14% chance that more than 4 people of the 7 chosen will have blood type O. This is not particularly surprising. GFY

6.78 Using technology, $P(Y\ge 3)=1-P(Y\le 2)=1-0.9885=0.0115$. There is only about a 1% chance that 3 or more of the plants would die before producing rhubarb. This would be surprising if it occurred.

6.79 (a) $P(X=17)=\binom{20}{17}(0.85)^{17}(0.15)^3=0.2428$. (b) Using technology, $P(X\le 12)=0.0059$. This would indeed be surprising and should make Judy suspicious.

1) binompdf (n, p, k)

2) binomcdf (n, p, k)

binompdf(15, .10, 3) binomcdf(15, .10, 3)

6.80 (a) $P(W=3)=\binom{15}{3}(0.10)^3(0.90)^{12}=0.1285$. (b) $P(W\geq 4)=1-P(W\leq 3)=1-0.9444=0.0556$.

There is about a 6% chance of finding 4 or more lefties in a sample of 15. This would be moderately surprising, but not completely unexpected.

6.81 (a) $\mu_X=np=15(0.20)=3$. You would expect to reach a live person in an average of 3 phone calls when making 15 calls. (b) $\sigma_X=\sqrt{np(1-p)}=\sqrt{15(0.20)(0.80)}=1.55$. In actual practice, you would expect the number of live persons you reach to vary from 3 per 15 calls by 1.55 on average.

6.82 (a) $\mu_X=np=12(0.20)=2.4$. You would expect to find an average of 2.4 people that the machine finds to be deceptive when testing 12 people actually telling the truth. (b) $\sigma_X=\sqrt{np(1-p)}=\sqrt{12(0.20)(0.80)}=1.39$. In actual practice, you would expect the number "deceivers" to vary from 2.4 by an average of 1.39.

6.83 (a) Y is also a binomial random variable. The only difference is that what we called a "failure" for X is a "success" for Y and vice versa. So the probability of success for Y is 0.80. Therefore $\mu_Y=np=15(0.80)=12$. Notice that $\mu_X=3$ and that $12+3=15$. In other words, if we reach an average of 3 live people in our 15 calls, we must not reach a live person in an average of 12 calls. (b) Notice that $\sigma_Y=\sqrt{np(1-p)}=\sqrt{15(0.80)(0.20)}=1.55$. This is the same thing as σ_X because we have just switched the definitions of p and $1-p$.

6.84 (a) Y is also a binomial random variable. The only difference is that what we called a "failure" for X is a "success" for Y and vice versa. So the probability of success for Y is 0.80.

$$P(Y\geq 10)=P(Y=10)+P(Y=11)+P(Y=12)=\binom{12}{10}(0.8)^{10}(0.2)^2+\binom{12}{11}(0.8)^{11}(0.2)^1+\binom{12}{12}(0.8)^{12}(0.2)^0$$

$=0.2835+0.2062+0.0687=0.5584$. Notice that $P(X\leq 2)=0.5584$ as well. If we find 2 or fewer lying, by definition we are saying that 10 or more are telling the truth. (b) $\mu_Y=np=12(0.8)=9.6$. Notice that $\mu_X=12(0.20)=2.4$ which is $12-\mu_Y$. Both σ_Y and σ_X are the same. $\sigma_Y=\sigma_X=\sqrt{12(0.8)(0.2)}=1.39$. The amount of variability around the number "failures" should be the same as the amount around the number of "successes."

6.85 (a) Check the BINS: Binary? "Success" = operates for an hour without failure. "Failure" = does not operate for an hour without failure. Independent? Whether one engine operates for an hour or not should not affect whether any other engine does or not. Number? We are looking at 350 engines. Success? Each engine has the same probability of success: 0.999. This is a binomial setting so, since X is counting the number of successes, it has a binomial distribution. (b) $\mu_X=np=350(0.999)=349.65$. If we were to test 350 engines over and over again, we would expect that, on average, 349.65 of them would operate for an hour without failure. $\sigma_X=\sqrt{np(1-p)}=\sqrt{350(0.999)(0.001)}=0.591$. In individual tests we would expect to find the number of engines that operate for an hour without failure to vary from 349.65 by an average of 0.591. (c) $P(X\leq 348)=1-P(X\geq 349)=1-P(X=349)-P(X=350)$

91a) ✓2 Outcomes < Visit Auction Site
Do Not Visit Auction Site

✓P(success) constant = .50

✓Observations Independent - Random Sample

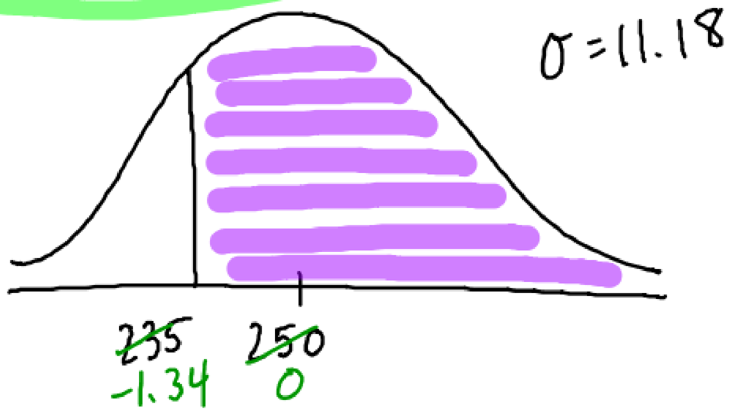
✓Fixed # Observations = 500

91b) $np = (500)(.5) = 250 > 10$ ✓

$n(1-p) = (500)(.5) = 250 > 10$ ✓

c) $\mu = np = 250$

$\sigma = \sqrt{n(p)(1-p)} = \sqrt{(500)(.5)(.5)} = 11.18$



normalcdf$(235, 1000, 250, 11.18) \approx .8982$

$P\left(z > \frac{235-250}{11.18} > -1.34\right) \approx .9099$

$1 - \text{binomcdf}(500, .5, 234) \approx .9172$

$\sigma_x = \sqrt{np(1-p)} = \sqrt{500(0.5)(0.5)} = 11.18$. So, using the Normal approximation,

$P(X \geq 235) = P\left(Z \geq \dfrac{235-250}{11.18}\right) = P(Z \geq -1.34) = 0.9099.$

6.92 (a) Check the BINS: Binary? "Success" = Person identify themselves as black. "Failure" = person does not identify themselves as black. Independent? The sample was random, so one person's identification shouldn't affect anyone else's and even though we are sampling without replacement, the sample size (1500) is far less than 10% of all American adults. Number? 1500 people were sampled. Success? The probability of success for any one individual is 0.12. This is a binomial setting so, since X measures the number of successes, X is approximately a binomial random variable. (b) The Normal approximation is quite safe: $np = 180$ and $n(1-p) = 1320$ are both more than 10. (c) First we need the mean and standard deviation for the binomial distribution. $\mu_X = np = 1500(0.12) = 180$.

$\sigma_x = \sqrt{1500(0.12)(0.88)} = 12.5857$. So, using the Normal approximation,

$P(165 \leq X \leq 195) = P\left(\dfrac{165-180}{12.5857} \leq Z \leq \dfrac{195-180}{12.5857}\right) = P(-1.19 \leq Z \leq 1.19) = 0.7660.$

6.93 Let X be the number of 1's and 2's; then X has a binomial distribution with $n = 90$ and $p = 0.477$ (in the absence of fraud). Using the calculator or software, we find $P(X \leq 29)$ = binomcdf(90, 0.477, 29) $\doteq$ 0.0021. Using the Normal approximation (the conditions are satisfied), we find a mean of 42.93 and standard deviation of $\sigma = \sqrt{90(0.477)(0.523)} = 4.7384$. Therefore,

$P(X \leq 29) = P\left(Z \leq \dfrac{29-42.93}{4.7384}\right) = P(Z \leq -2.94) = 0.0016$. Either way, the probability is quite small, so we have reason to be suspicious.

6.94 Let X be the number of hits out of 500 times at bat. Then X has a binomial distribution with $n = 500$ and $p = 0.260$. Using the Normal approximation (the conditions are satisfied), we find a mean of 130 and a standard deviation of $\sigma = \sqrt{500(0.26)(0.74)} = 9.808$. Therefore

$P(X \geq 150) = P\left(Z \geq \dfrac{150-130}{9.808}\right) = P(Z \geq 2.04) = 0.0207$. This could happen but we would only expect it to happen for about 2% of typical baseball players.

6.95 (a) Check the BITS: Binary? "Success" = get a card you don't already have. "Failure" = get a card you already have. Independent? Assuming the cards are put into the boxes randomly at the factory, what card you get in a particular box should be independent of what card you get in any other box. Trials? We are not counting trials until the first success. This is not a geometric setting. (b) Check the BITS: Binary? "Success" = Lola wins some money. "Failure" = Lola does not win money. Independent? The outcomes of the games are independent of each other. Trials? We are counting the number of games until she wins one. Success? The probability of success on any given game is 0.259. This is a geometric setting. Let Y be the number of games Lola plays up to and including her first win. This is a geometric random variable.

6.96 (a) Check the BITS: Binary? "Success" = get an ace. "Failure" = do not get an ace. Independent? The trials are not independent of one another because we are not replacing the previous cards drawn back into the deck. This is not a geometric setting. (b) Check the BITS: Binary? "Success" = get a bulls-eye. "Failure" = do not get a bulls-eye. Independent? Different shots should be independent of each other. Whether he makes one shot does not affect whether he makes any other shots. Trials? We are continuing until he gets his first bulls-eye. Success? The probability of success is 0.10 for all shots. This is a geometric setting. Let X be the number of shots Lawrence makes up to and including his first bulls-eye. This is a geometric random variable.

See Next Pages

6.97 (a) $P(X=3)=(0.8)^2(0.2)=0.128$. (b) Using technology, [illegible]$(X>10)=1-P(X\le 10)=1-0.[illegible]926=0.1074$.

6.9[illegible] [illegible] $P($[illegible]$=5)=\left(\frac{5}{6}\right)^4$ [illegible] [illegible]04. (b[illegible] [illegible]sin[illegible] [illegible]echnology, [illegible] [illegible]

6.99 (a) The expected value is $\mu_X=\frac{1}{p}=\frac{1}{\frac{1}{38}}=38$. We would expect it to take 38 games to win if one wins in 1 out of 38 games. (b) Using technology, $P(X\le 3)=0.0769$. While this is not a usual occurrence, it would happen about 8% of the time, so it is not completely surprising.

6.100 (a) The expected value is $\mu_X=\frac{1}{p}=\frac{1}{0.097}=10.31$. We would expect to examine about 10.31 invoices in order to find the first 8 or 9. (b) Using technology, $P(X\ge 40)=1-P(X\le 39)=1-0.9813=0.0187$. The likelihood of this happening is reasonably rare. We may begin to worry that the invoice amounts are a fraud.

6.101 b

6.102 c

6.103 b

6.104 c

6.105 c

6.106 This is an experiment; students were randomly (at the time they visited the site) assigned to a treatment. The explanatory variable is the login box (genuine or not), and the response variable is the student's action (logging in or not).

6.107 (a) [INSERT S6_107] $P(\text{smoke})=0.086+0.0986+0.0874=0.272$. 27.2% of British men smoke. (b) $P(\text{routine and manual}\mid\text{smoke})=\frac{P(\text{routine and manual and smoke})}{P(\text{smoke})}=\frac{0.0874}{0.272}=0.321$. About 32% of the smokers are routine and manual laborers.

97a) $P(X=n) = \text{geometpdf}(p, n)$

$P(X=3) = \text{geometpdf}\left(\frac{1}{5}, 3\right) = .128$

$P(X=3) = P(\text{No})P(\text{No})P(\text{Yes}) = \left(\frac{4}{5}\right)^2\left(\frac{1}{5}\right) = .128$

b) $P(X>n) = 1 - \text{geometcdf}(p, n)$

$P(X>10) = 1 - \text{geometcdf}\left(\frac{1}{5}, 10\right) = .1074$

$P(X>10) = P(\text{No})P(\text{No})P(\text{No})\ldots = \left(\frac{4}{5}\right)^{10} = .1074$

98a) $P(X=n) = \text{geometpdf}(p, n)$

$P(X=5) = \text{geometpdf}(\frac{1}{6}, 5) = .0804$

$P(X=5) = P(\text{No})P(\text{No})P(\text{No})P(\text{No})P(\text{yes}) = (\frac{5}{6})^4(\frac{1}{6}) = .0804$

b) $P(X \le n) = \text{geometcdf}(p, n)$

$P(X \le 8) = \text{geometcdf}(\frac{1}{6}, 8) = .7674$

99a) Expected Value $(\mu) = \frac{1}{p}$

$\mu = \frac{1}{\frac{1}{38}} = 38$ } Expect to play 38 games to win

b) $P(X \leq n) = \text{geometcdf}(p, n)$

$P(X \leq 3) = \text{geometcdf}(1/38, 3) = .0769$

Happen about 8% of time so not completely a surprise