Chapter 6

Discussion Problem Solutions

D1. The distribution in Display 6.6 is more symmetric than that in Display 6.5, which is skewed toward the larger values. Display 6.6 has a mean that is about double that of Display 6.5 (about 3.5 versus 1.75) and has a spread that is larger than the spread of Display 6.5.

D2. While three or more cars in a household is not unusual, selecting two such households in a sample of 2 is much less probable. For example, the probability of selecting a single house that requires three parking spaces is 0.201, but the probability of selecting two households that require three parking spaces each is $0.201 \cdot 0.201 = 0.0404$. In general, an event happening twice in a row is less likely than the event happening just once.

D3. Using the probabilities given in Display 6.5,

$$\begin{aligned}
P(\textit{two vehicles in a duplex}) &= P(\textit{no vehicle in the 1st household} \text{ and} \\
&\qquad \textit{two vehicles in the 2nd household} \\
&\qquad \text{or } \textit{one vehicle in the 1st household} \text{ and} \\
&\qquad \textit{one vehicle in the 2nd household} \\
&\qquad \text{or } \textit{two vehicles in the 1st household} \text{ and} \\
&\qquad \textit{no vehicle in the 2nd household}) \\
&= P(\textit{no vehicle in 1}) \cdot P(\textit{two vehicles in 2}) + \\
&\quad\ P(\textit{one vehicle in 1}) \cdot P(\textit{one vehicle in 2}) + \\
&\quad\ P(\textit{two vehicles in 1}) \cdot P(\textit{no vehicle in 2}) \\
&= 0.087 \cdot 0.381 + 0.331 \cdot 0.331 + 0.381 \cdot 0.087 \\
&= 0.175855
\end{aligned}$$

which rounds to the value 0.176 given in Display 6.6.

D4. a.

Value, x	Frequency, f	$x \cdot f$
5	12	60
6	23	138
8	15	120
Total	**50**	**318**

$$\bar{x} = \frac{\sum x \cdot f}{n} = \frac{318}{50} = 6.36$$

Some students will want to divide by 3 because there are three rows, but it is the total frequency that gives the sample size: $12+23+15=50$. An alternative formula is

$$\bar{x}=\frac{\sum x\cdot f}{\sum f}$$

b.

Value, x	Proportion, $\frac{f}{n}$
5	0.24
6	0.46
8	0.30

c. Using the values from part b. above,

$$\bar{x}=\sum x\cdot\left(\frac{f}{n}\right)=5(0.24)+6(0.46)+8(0.30)=6.36$$

d. You can think of $\frac{f_i}{n}$ as being the probability p_i that the random variable takes on the specific value x_i. . So, $\sum x\cdot\left(\frac{f}{n}\right)=\sum x\cdot p$ which is the formula for expected value.

D5. The lesson here is that the expected value of a probability distribution of outcomes need not be one of the outcomes itself. The average number of children per household can be 2.2 children even though it is impossible for one such household to have 2.2 children.

D6. The first distribution in 6.2, the sum of two dice, is symmetric about its middle or mean, increasing up to its mean and then decreasing. The second distribution, the larger of two dice, is strictly increasing and skewed to the left. Let *S* denote the sum of two rolled dice and *L* the larger of two rolled dice. Then,

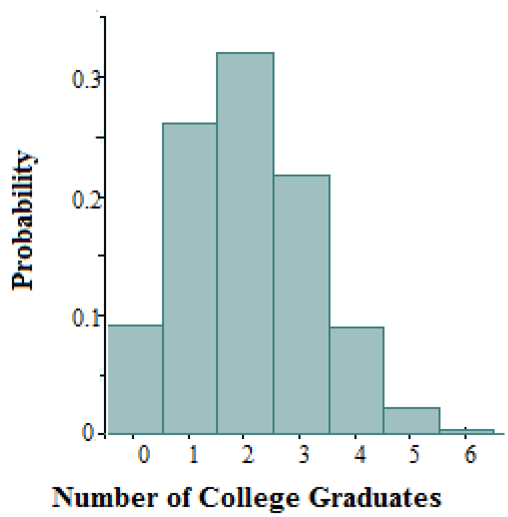
$$\begin{aligned}E(S)&=2\cdot\left(\frac{1}{36}\right)+3\cdot\left(\frac{2}{36}\right)+4\cdot\left(\frac{3}{36}\right)+5\cdot\left(\frac{4}{36}\right)+6\cdot\left(\frac{5}{36}\right)+7\cdot\left(\frac{6}{36}\right)\\&\quad+8\cdot\left(\frac{5}{36}\right)+9\cdot\left(\frac{4}{36}\right)+10\cdot\left(\frac{3}{36}\right)+11\cdot\left(\frac{2}{36}\right)+12\cdot\left(\frac{1}{36}\right)\\&=\frac{252}{36}=7\\E(L)&=1\cdot\left(\frac{1}{36}\right)+2\cdot\left(\frac{3}{36}\right)+3\cdot\left(\frac{5}{36}\right)+4\cdot\left(\frac{7}{36}\right)+5\cdot\left(\frac{9}{36}\right)+6\cdot\left(\frac{11}{36}\right)\\&=\frac{161}{36}\approx 4.47.\end{aligned}$$

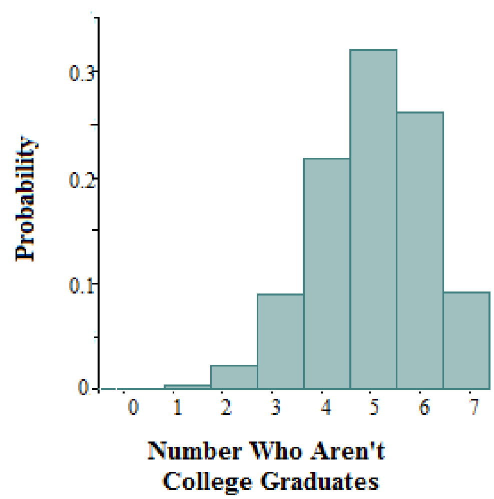
When rolling two dice, the expected sum of the dice is seven. The larger value when two dice are rolled will, on average, be between 4 and 5, around 4.47.

D7. A reasonable estimate of the center of each interval is its midpoint. For example, the center of the interval 1-25 is $\frac{1+25}{2}=13.$ The midpoints of 0, 1-25, 26-50, 51-75, and 76-100 are 0, 13, 38, 63, and 88 respectively. Using these values with the given frequencies, the expected value is $\mu_X = 0(0.834)+13(0.149)+38(0.009)+63(0.004)+88(0.004) = 2.883$ and the variance is $\sigma_X^2 = \sum (x-\mu_X)^2 \cdot p$ $= (0-2.883)^2(0.834)+(13-2.883)^2(0.149)+(38-2.883)^2(0.009)$ $+(63-2.883)^2(0.004)+(88-2.883)^2(0.004)$ $= 76.717311$ Taking the square root gives a standard deviation of 8.759. In summary, a judge can expect 2.883% of cases where they believe the compensatory damages are disproportionately high although typically this number will differ from the mean by 8.759% more or less. Note that 2.883% less 8.759% is a negative number, which is not a realistic percent of cases so the percent of such cases would typically range from 0 to 2.883% plus 8.759%.
D8. The total number of hours you expect to spend is $10+15=25$ hours. The variance of the total is $2^2+3^2=13,$ so the standard deviation is $\sqrt{13}\approx 3.61$ hours. However, using the rule for adding variances is not reasonable in this case because the two times undoubtedly are not independent. With only so many hours in the week, if you work a lot, you must spend less time studying, and vice versa.
D9. Say you have a machine that fills soda bottles. Your target is 16 fluid ounces, but the machine is not perfect, so there is some variability in the actual amount put into the bottle. The actual amount in the bottle can be considered a random variable. Let's say that you can expect to be off by up to half a fluid ounce in either direction. Now imagine that you pick two bottles at random and pour the contents of both bottles into a pitcher. Since you could be up to half an ounce off in either direction on an individual bottle, the pitcher could be off by as much as an ounce in either direction. There is more variability in the sum than in either of the original random variables. Now say you have selected one of the bottles and want to pour 6 ounces out into a drinking glass. Of course you can't measure perfectly, so the amount actually poured out is a second random variable. For ease of discussion, assume you can expect to be off by as much as half an ounce. That is, the original amount could be off by half an ounce either way, as could the amount poured out. Thus, the remaining amount (the difference) could be off by 1 ounce in either direction, just like the sum in the first part. In short, whether you are summing up random variables or subtracting two random variables, you are increasing the variability.

D10. $P(number\ of\ college\ grads = k) = P(number\ not\ college\ grads = 7 - k)$. In other words, the number of people in the sample who are not college grads is 7 minus the number of college grads in the sample. Based the table given in Display 6.34, the distribution of non-graduates is shown below along with graphs of both distributions.

Number of College Graduates	Number Who Aren't College Graduates	Probability
0	7	0.091
1	6	0.260
2	5	0.319
3	4	0.217
4	3	0.089
5	2	0.022
6	1	0.003
7	0	0.000





The graph of the distribution of the number of non-college graduates is a reflection of the distribution of the number of college graduates.

D11. Using the general Multiplication Rule, the probability that none will be chocoholics is $\frac{8}{12}\cdot\frac{7}{11}\cdot\frac{6}{10}\cdot\frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}\cdot\frac{2}{6} \approx 0.010101$. Using the binomial probability formula (inappropriately), where $p = \frac{4}{12} = \frac{1}{3}$, the probability is

$P(X = 0) = \binom{7}{0}\left(\frac{1}{3}\right)^0\left(1 - \frac{1}{3}\right)^7 \approx 0.05853$. Quite a difference!

D12. *Shape:* For small values of p (close to 0) the binomial distributions are highly skewed toward the larger values of x. As p increases to 0.5, the distributions get more symmetric, with perfect symmetry achieved at $p = 0.5$. As p increases from 0.5 to 1, the distributions get increasingly skewed toward the smaller values of x. As n increases, the skewness decreases for values of p different from 0.5 and the distributions tend to look more and more normal. For
$p = 0.5$, the binomial distribution is symmetric and mound-shaped for any sample size. *Center:* Because the mean is located at np, it increases with both n and p. *Spread:* The standard deviation increases with n. For a fixed n, the standard deviation is largest around $p = 0.5$ and gets smaller as p gets closer to 0 or 1.

D13. If X denote the number of MP3 players that fail in the first month, X has a binomial distribution with $n = 10$ and $p = 0.08$. Because the seller sells all 10 players for \$100 each, the gain, G, is $G = 100(10) - 200X$. Then

$$E(G) = 100(10) - 200E(X) = 100(10) - 200(10)(0.08)$$
$$= 1000 - 160 = \$840.$$

D14. If you randomly select three adults and record the number of college graduates among the three, the result will be either 0, 1, 2, or 3. Suppose you conduct such a sample and the result is 1. You repeat the experiment with another random sample of three adults and observe 2 college graduates. Thus far, your average observation is $(1+2)/2 = 1.5$ college graduates, which is not a whole number of people. Saying the expected number of college graduates is 0.87 means that if you were to continue this sampling process and compute the average observation, this average would approach 0.87.

Practice Problem Solutions

P1. Define X to be the smaller of the two numbers when rolling two six-sided dice. Then the probability distribution of the random variable X is given by this table:

x	p	x	p
1	$\frac{11}{36}$	4	$\frac{5}{36}$
2	$\frac{9}{36}$	5	$\frac{3}{36}$
3	$\frac{7}{36}$	6	$\frac{1}{36}$

P2. The four possible outcomes are:

None are spam
The first selected message is the only spam
The second selected message is the only spam
The first and second selected messages are spam

The corresponding probabilities are:

P(None are Spam) = 0.3 • 0.3 = 0.09
P(*The first selected message is the only spam*) = 0.7 • 0.3 = 0.21
P(*The second selected message is the only spam*) = 0.3 • 0.7 = 0.21
P(*The first and second selected messages are spam*) = 0.7 • 0.7 = 0.49

Define *X* to be the number of selected messages that are spam. Then the probability distribution of the random variable *X* is given by this table:

X	**P(X)**
0	0.09
1	0.42
2	0.49

P3. a. There are 10 possible samples:

BCM, BCA, BCS, BMA, BMS, BAS, CMA, CMS, CAS, MAS

b. Each sample has a probability of 1/10 of being selected.
c. Define *X* to be the number of committee members who live in the dorms. Then the probability distribution of the random variable *X* is given by this table:

x	p	x	p
0	$\frac{1}{10}$	2	$\frac{3}{10}$
1	$\frac{6}{10}$	3	$\frac{0}{10}$

P4. 0(0.09) + 1(0.42) + 2(0.49) = 1.4

P5. a. $\mu = \sum x \cdot P(x)$ = 0(0.524) + 1(0.201) + 2(0.179) + 3(0.070) + 4(0.026) = 0.873

b. On average, a family in the US has 0.873 children.

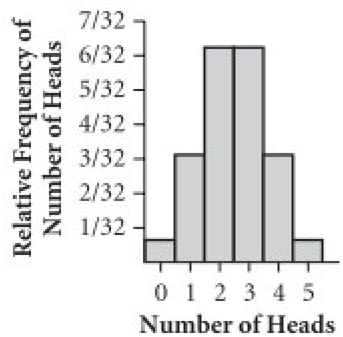
c. The expected number of children when 10 families are chosen at random is 10(0.873) = 8.73.

P6. The possible outcomes when flipping a coin five times are:

5 Heads	4 Heads	3 Heads	2 Heads	1 Head	0 Heads
HHHHH	*HHHHT*	*HHHTT*	*HHTTT*	*HTTTT*	*TTTTT*
	HHHTH	*HHTHT*	*HTHTT*	*THTTT*	
	HHTHH	*HTHHT*	*THHTT*	*TTHTT*	
	HTHHH	*THHHT*	*HTTHT*	*TTTHT*	
	THHHH	*HHTTH*	*THTHT*	*TTTTH*	
		HTHTH	*TTHHT*		
		THHTH	*HTTTH*		
		HTTHH	*THTTH*		
		THTHH	*TTHTH*		
		TTHHH	*TTTHH*		

Each of these outcomes is equally likely with probability 1/32. Define *X* to be the number of heads when flipping a coin five times. Then the probability distribution of the random variable *X* is given by this table and graph:

x	*p*	*x*	*p*
0	$\frac{1}{32}$	3	$\frac{10}{32}$
1	$\frac{5}{32}$	4	$\frac{5}{32}$
2	$\frac{10}{32}$	5	$\frac{1}{32}$



The expected value is 2.5.

P7. a.

Sum of Two Tetrahedral Dice	Probability
2	$\frac{1}{16}$
3	$\frac{2}{16}$
4	$\frac{3}{16}$
5	$\frac{4}{16}$
6	$\frac{3}{16}$
7	$\frac{2}{16}$
8	$\frac{1}{16}$
Total	$\mathbf{\frac{16}{16} = 1}$

The probability that the sum is 3 is $\frac{2}{16}$.

b. The expected value of this distribution is

$$2\left(\tfrac{1}{16}\right)+3\left(\tfrac{2}{16}\right)+4\left(\tfrac{3}{16}\right)+5\left(\tfrac{4}{16}\right)+6\left(\tfrac{3}{16}\right)+7\left(\tfrac{2}{16}\right)+8\left(\tfrac{1}{16}\right)=5.0.$$

P8. a. $\mu_X = 600 \cdot \frac{1}{500} + 0 \cdot \frac{499}{500} = \$1.20$ **b.** $\mu_X = 1000 \cdot \frac{1}{500} + 400 \cdot \frac{2}{500} + 0 \cdot \frac{497}{500} = \$3.60$

P9.

Outcome	Payout	Probability
No burglary	$0	0.9768
Burglary	$5000	0.0232

The expected payout per policy is 0(0.9768) + 5000(0.0232) = $116, so you should charge $116 to break even.

P10.

x	p	$x \cdot p$
0	0.45	0
1	0.09	0.09
2	0	0
3	0.04	0.12
4	0.15	0.60
5	0.27	1.35
Sum	**1**	**2.16**

Observe that

$$\mu = \sum x \cdot P(x) = 2.16, \qquad \sigma = \sqrt{\sum (x-\mu)^2 \cdot P(x)} \approx 2.221.$$

The mean score on the problem was 2.16, but it would be typical for a student to get a score 2.221 points higher or lower than that. (But not less than 0.)

P11. Answers will vary depending on the choice of center. Here we use the midpoint of the first two age groups, and 35 for the last group.

Age	Proportion
19.5	0.442
26	0.314
35	0.244

Observe that

$$\mu = \sum x \cdot P(x) = 25.323, \qquad \sigma = \sqrt{\sum (x-\mu)^2 \cdot P(x)} \approx 6.163.$$

Using the midpoint as the center for the first two intervals and 35 for the last interval, the mean age of students at UTSA is 25.232, but it would be typical for a student to be 6.163 years older or younger than that.

P12. a. 1,691,290 tickets out of 8,054,375, or 0.210, are winners. This is about equal to $1/4.76 \approx 0.210$.

b. On average, you will lose 70¢ for each ticket purchased. The variability in the outcomes is quite large: $23.69, which indicates there is a small chance of winning a very big prize.

P13. a. The probabilities do not add to one. So, it must be the case that the probability of not being able to rent the cleaner on a typical Saturday is 1 – (0.4 + 0.3 + 0.1) = 0.2. When this happens, the rental income is $0. The expected rental income for the store is then:

$$\mu_x = \sum x \cdot P(x) = 30 \bullet 0.4 + 50 \bullet 0.3 + 60 \bullet 0.1 + 0 \bullet 0.2 = \$33$$

b.

$$\begin{aligned}\sigma_x &= \sqrt{\sum (x-\mu_x)^2 \cdot P(x)} \\ &= \sqrt{(30-33)^2 \cdot 0.4 + (50-33)^2 \cdot 0.3 + (60-33)^2 \cdot 0.1 + (0-33)^2 \cdot 0.2} \approx \$19.52\end{aligned}$$

The possible deviations from the mean (in absolute value) are $3, $17, $27, and $33. No deviation is very close to $19.52, but as typical deviation, it seems reasonable.

P14. a. Observe that

$$\mu_x = \sum x \cdot P(x) = 0(0.524) + 1(0.201) + 2(0.179) + 3(0.070) + 4(0.026) = 0.873$$

$$\begin{aligned}\sigma_x &= \sqrt{\sum (x-\mu_x)^2 \cdot P(x)} \\ &= \sqrt{\begin{array}{l}(0-0.873)^2 (0.524) + (1-0.873)^2 (0.201) + (2-0.873)^2 (0.179) + (3-0.873)^2 (0.070) \\ \qquad + (4-0.873)^2 (0.026)\end{array}} \\ &= 1.096\end{aligned}$$

On average, a randomly selected family will have 0.873 children, give or take 1.096 children. (The fact that the standard deviation is larger than the mean is an indication that the distribution is skewed right.)

b. On average, 30 randomly selected families will have 30(0.873) = 26.19 children, give or take 6.003 children.

P15. a. The expected yearly savings is 52($35) = $1,820. The standard deviation for yearly savings is $\sqrt{52}(\$15) = \$108.17$.

b. The expected yearly savings in this case is 52($35 + $10) = $2,340. The standard deviation for yearly savings in this case is the same as in (a) because adding a constant to a random variable does not change its standard deviation.

P16. a. The expected value and standard deviation of this distribution are

$$\mu = 2\left(\tfrac{1}{16}\right)+3\left(\tfrac{2}{16}\right)+4\left(\tfrac{3}{16}\right)+5\left(\tfrac{4}{16}\right)+6\left(\tfrac{3}{16}\right)+7\left(\tfrac{2}{16}\right)+8\left(\tfrac{1}{16}\right)=5.0$$

$$\sigma^2 = (2-5)^2\left(\tfrac{1}{16}\right)+(3-5)^2\left(\tfrac{2}{16}\right)+(4-5)^2\left(\tfrac{3}{16}\right)+(5-5)^2\left(\tfrac{4}{16}\right)$$
$$+(6-5)^2\left(\tfrac{3}{16}\right)+(7-5)^2\left(\tfrac{2}{16}\right)+(8-5)^2\left(\tfrac{1}{16}\right)=2.5.$$

b. The expected value and standard deviation of this distribution are

$$\mu = 1\left(\tfrac{1}{4}\right)+2\left(\tfrac{1}{4}\right)+3\left(\tfrac{1}{4}\right)+4\left(\tfrac{1}{4}\right)=2.5$$

$$\sigma^2 = (1-2.5)^2\left(\tfrac{1}{4}\right)+(2-2.5)^2\left(\tfrac{1}{4}\right)+(3-2.5)^2\left(\tfrac{1}{4}\right)+(4-2.5)^2\left(\tfrac{1}{4}\right)=1.25.$$

c. The mean roll is 2.5. So, the sum of the means of two such rolls is 5. This coincides with the mean of the sum from (a).

d. This is similar to (c). The variance of each roll is 1.25, so that the sum for two rolls is 2.5. This coincides with the variance from (a).

P17. a. For a duplex, the mean and standard deviation are

$$\mu = 0(0.008)+1(0.058)+2(0.176)+3(0.287)+4(0.278)$$
$$+5(0.153)+6(0.040)=3.388$$

$$\sigma^2 = (0-3.388)^2(0.008)+(1-3.388)^2(0.058)+(2-3.388)^2(0.176)+(3-3.388)^2(0.287)$$
$$+(4-3.388)^2(0.278)+(5-3.388)^2(0.153)+(6-3.388)^2(0.040)=1.579$$

For a single household, the mean and standard deviation are

$$\mu = 0(0.087)+1(0.331)+2(0.381)+3(0.201)=1.696$$

$$\sigma^2 = (0-1.696)^2(0.087)+(1-1.696)^2(0.331)+(2-1.696)^2(0.381)+(3-1.696)^2(0.201)$$
$$=0.78758$$

b. This is similar to P16 c and d. Indeed, from part (a), the mean number of vehicles for a single household is 1.696, so that for two such households, the mean is 1.6962=3.392, which coincides with the mean for a duplex. Similarly, the variance for each household is 0.78758, so that for two households the variance is 2*0.78758=1.575, which also coincides with the results from part (a) for a duplex.

P18. Let X = number of heads from 8 flips of a coin.

a. $P(X=3)=\binom{8}{3}(0.5)^8=0.21875$

b. Since 25% of 8 is 2 heads, we compute $P(X=2)=\binom{8}{2}(0.5)^8=0.109375$.

c. $P(X=7 \text{ or } X=8)=0.03125+0.00391\approx 0.03516$

P19. a. $\binom{5}{1}\left(\frac{6}{36}\right)^1\left(\frac{30}{36}\right)^4 \approx 0.4019$

b. $\binom{5}{3}\left(\frac{6}{36}\right)^3\left(\frac{30}{36}\right)^2 \approx 0.0322$

c. $P(\text{at least one } 7) = 1 - P(\text{no } 7) = 1 - (\frac{30}{36})^5 \approx 0.5981$

d. $P(\text{at most one } 7) = P(\text{no } 7) + P(\text{one } 7) = \binom{5}{0}\left(\frac{6}{36}\right)^0\left(\frac{30}{36}\right)^5 + \binom{5}{1}\left(\frac{6}{36}\right)^1\left(\frac{30}{36}\right)^4 \approx 0.8038$.

P20. a. The probability of touching either side of the forehead is the same, namely 0.5.

b. Since such trials are assumed to be independent, the probability that she touches the mark every time is $\binom{12}{12}(0.5)^{12}(0.5)^0 \approx 0.0002441$, or about a 0.02% chance.

c. Happy isn't picking the side to touch at random; she is favoring the side with the mark so she must understand where the mark is on her forehead.

P21. Let X = number with blood type A and p = probability of having type A blood.

a. This is a binomial experiment with n = 10 and p = 0.40. As such, we have

$$P(X = 0) + P(X = 1) = 0.04636.$$

b. Yes, since $P(X \geq 2) = 0.95364$.

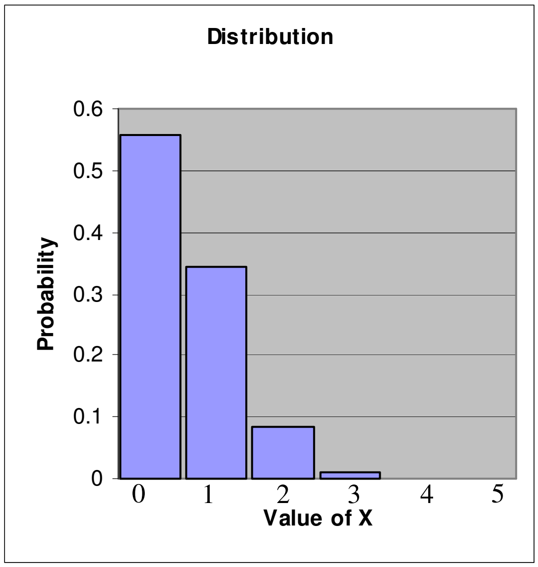
P22. Let X = number of dropouts. Choosing a sample of 5 people at random in the given age group, we have a binomial random variable with p = 0.11.

a. The distribution is as follows:

X	**Probability**
0	$\binom{5}{0}(0.11)^0(0.89)^5 \approx 0.5584$
1	$\binom{5}{1}(0.11)^1(0.89)^4 \approx 0.3451$
2	$\binom{5}{2}(0.11)^2(0.89)^3 \approx 0.0853$

3	$\binom{5}{3}(0.11)^3(0.89)^2 \approx 0.0105$
4	$\binom{5}{4}(0.11)^4(0.89)^1 \approx 0.0006515$
5	$\binom{5}{5}(0.11)^5(0.89)^0 \approx 0.00001605$

b. The histogram is as follows. It is strongly skewed right, towards larger values



c. $\mu = 5(0.11) = 0.55$, $\sigma^2 = 5(0.11)(0.89) = 0.4895$, so that $\sigma = 0.6996$

d. No. If 11% is the correct percentage, the probability that none are dropouts is quite high, 0.5584.

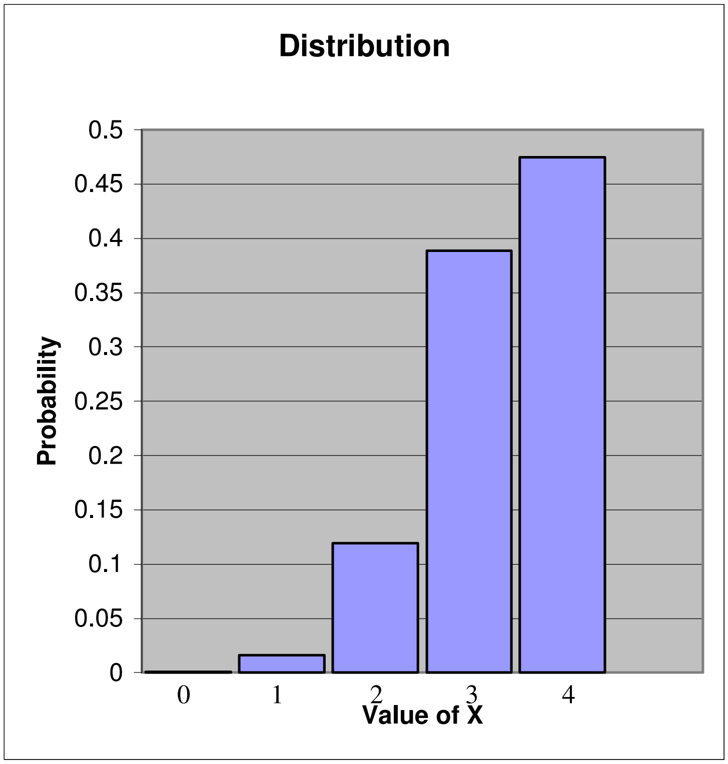
P23. Let X = number of drivers wearing seatbelts. Choosing a sample of 4 people at random, we have a binomial random variable with $p = 0.84$.

a. The distribution is as follows:

X	**Probability**
0	$\binom{4}{0}(0.16)^4(0.84)^0 \approx 0.000655$
1	$\binom{4}{1}(0.16)^3(0.84)^1 \approx 0.0138$
2	$\binom{4}{2}(0.16)^2(0.84)^2 \approx 0.1084$

3	$\binom{4}{3}(0.16)^1(0.84)^3 \approx 0.3793$
4	$\binom{4}{4}(0.16)^0(0.84)^4 \approx 0.4979$

The histogram is as follows. It is skewed left, towards the smaller values.



b. Observe that $\mu = 4(0.84) = 3.36$, So, on average, 3.36 of the 4 randomly selected drivers will be wearing seatbelts.

c. The standard deviation is $\sigma = \sqrt{4(0.16)(0.84)} \approx 0.7332$, so it would be typical to find 3 or 4 drivers wearing seatbelts, which are the values within 0.7332 of the mean.

d. Definitely; the probability that no drivers in a random sample of 4 drivers were wearing seatbelts is only 0.000655 if NHTSA is correct that 84% of drivers use seatbelts regularly.

Exercise Solutions

E1. a. The outcomes are pairs of selections of distinct concertos. Assuming each digit in a pair refers to the number of the concerto played, the 10 outcomes are as follows:

12, 13, 14, 15, 23, 24, 25, 34, 35, 45

Let X = number of minutes played in pair of concertos. The values of X corresponding to each outcome are as follows:

Pair	Value of X	Pair	Value of X
12	68	24	63
13	73	25	70
14	71	34	68
15	78	35	75
23	65	45	73

The distribution for X is as follows:

X	Probability
63	$\frac{1}{10}$
65	$\frac{1}{10}$
68	$\frac{2}{10}$
70	$\frac{1}{10}$
71	$\frac{1}{10}$
73	$\frac{2}{10}$
75	$\frac{1}{10}$
78	$\frac{1}{10}$

The expected value of X is:

$$\mu = 63\left(\tfrac{1}{10}\right) + 65\left(\tfrac{1}{10}\right) + 68\left(\tfrac{2}{10}\right) + 70\left(\tfrac{1}{10}\right) + 71\left(\tfrac{1}{10}\right) + 73\left(\tfrac{2}{10}\right) + 75\left(\tfrac{1}{10}\right) + 78\left(\tfrac{1}{10}\right) = 70.4 \text{ min.}$$

So, on average, 2 concertos played back to back will provide 70.4 minutes of music.

b. $P(X \le 74) = 1 - [P(X = 75) + P(X = 78)] = 0.8$.

E2. a. The fifteen possible samples of size two are

1 and 2; 1 and 3; 1 and 4; 1 and 5; 1 and 6; 2 and 3;
2 and 4; 2 and 5; 2 and 6; 3 and 4; 3 and 5; 3 and 6;
4 and 5; 4 and 6; 5 and 6

Assume computers 1, 2, and 3 are the defective monitors. The probability would be the same no matter which three are assigned as the defective monitors.

Define X to be the number of defective monitors in the sample. Then the probability distribution of the random variable X is given by this table:

x	p
0	$\frac{3}{15}$
1	$\frac{9}{15}$
2	$\frac{3}{15}$

Note that

$$\mathrm{E}(X) = \sum x \cdot P(x) = 0(3/15) + 1(9/15) + 2(3/15) = 1$$

$$\sigma_x = \sqrt{\sum (x - E(x))^2 \cdot P(x)} = \sqrt{(0-1)^2 \cdot \tfrac{3}{15} + (1-1)^2 \cdot \tfrac{9}{15} + (2-1)^2 \cdot \tfrac{3}{15}} \approx 0.632$$

b. The probability of a complete inspection being triggered by this set of six monitors is $P(X = 1) + P(X = 2) = 12/15$.

E3. a. Let N = no seat belt and S = wears seat belt.
If we choose a sample of three people, the 8 possible outcomes are:

NNN, SNN, NSN, NNS, SSN, SNS, NSS, SSS

Let X = number of drivers who DO NOT wear a seat belt.
This is a binomial random variable with $p = 0.84$. The distribution of X is as follows:

X	**Probability**
0	$\binom{3}{0}(0.16)^0(0.84)^3 \approx 0.593$
1	$\binom{3}{1}(0.16)^1(0.84)^2 \approx 0.339$
2	$\binom{3}{2}(0.16)^2(0.84)^1 \approx 0.065$
3	$\binom{3}{3}(0.16)^3(0.84)^0 \approx 0.004$

b. No, the probability that none of the three were wearing seat belts is 0.593, which is quite likely if 16% of pickup truck occupants wear seat belts.

E4. a. Let Y = number of drivers who DO wear a seat belt.
Assuming that the probability of not wearing a seat belt is the same as in E3, the distribution for Y is very similar to that of X. Indeed, observe that

Y	*X* (from E3)	Probability
0	3	0.004
1	2	0.065
2	1	0.339
3	0	0.593

b. Yes, because the probability that none of the three people chosen was wearing a seat belt is 0.004, which is unlikely if 84% of people are assumed to wear a seat belt.

E5. a. Let X = number of speeders who get another speeding ticket within the year.
Let Y = yes, they get another ticket and N = no, they do not get another ticket.
If choosing a random sample of three drivers from this population, the possible outcomes are:

YYY, NYY, YNY, YYN, NNY, NYN, YNN, NNN

The distribution for X is as follows:

X	Probability
0	$\binom{3}{0}(0.11)^0(0.89)^3 \approx 0.705$
1	$\binom{3}{1}(0.11)^1(0.89)^2 \approx 0.261$
2	$\binom{3}{2}(0.11)^2(0.89)^1 \approx 0.032$
3	$\binom{3}{3}(0.11)^3(0.89)^0 \approx 0.001$

b. You cannot tell because you need to know something about the people who were speeding and might have gotten a ticket but didn't. If more than 11% of those people got at least one speeding ticket in the next year, then you would have some evidence that tickets deter speeding.

E6. Let S represent a patient whose lung cancer was caused by smoking and let N represent a patient whose lung cancer was not caused by smoking. The possible outcomes for samples of three patients are

Outcome	Probability
NNN	0.13 • 0.13 • 0.13 = 0.002
NNS	0.13 • 0.13 • 0.87 = 0.015
NSN	0.13 • 0.87 • 0.13 = 0.015
SNN	0.87 • 0.13 • 0.13 = 0.015
NSS	0.13 • 0.87 • 0.87 = 0.098
SNS	0.87 • 0.13 • 0.87 = 0.098
SSN	0.87 • 0.87 • 0.13 = 0.098
SSS	0.87 • 0.87 • 0.87 = 0.659

Where, for example, NNS means that for the first two patients selected, the cancer was not caused by smoking and for the third patient selected, the cancer was caused by smoking.

Define X to be the number of patients with cancer caused by smoking. Then the probability distribution of the random variable X is given by this table:

X	Probability
0	0.002
1	0.045
2	0.294
3	0.659

b. Yes, because the likelihood of this occurring assuming that 87% of people's lung cancer was caused by smoking is very small.

E7. a. Possible outcomes for number of TVs in Units 1 and 2, respectively, in a duplex:

Outcome (unit 1, unit 2)	Probability
0,0	$0.012 \cdot 0.012 = 0.000144$
0,1	$0.012 \cdot 0.274 = 0.003288$
0,2	$0.012 \cdot 0.359 = 0.004308$
0,3	$0.012 \cdot 0.218 = 0.002616$
0,4	$0.012 \cdot 0.095 = 0.001140$
0,5	$0.012 \cdot 0.042 = 0.000504$
1,0	$0.274 \cdot 0.012 = 0.003288$
1,1	$0.274 \cdot 0.274 = 0.075076$
1,2	$0.274 \cdot 0.359 = 0.098366$
1,3	$0.274 \cdot 0.218 = 0.059732$
1,4	$0.274 \cdot 0.095 = 0.026030$
1,5	$0.274 \cdot 0.042 = 0.011508$
2,0	$0.359 \cdot 0.012 = 0.004308$
2,1	$0.359 \cdot 0.274 = 0.098366$

Outcome (unit 1, unit 2)	Probability
2,2	$0.359 \cdot 0.359 = 0.128881$
2,3	$0.359 \cdot 0.218 = 0.078262$
2,4	$0.359 \cdot 0.095 = 0.034105$
2,5	$0.359 \cdot 0.042 = 0.015078$
3,0	$0.218 \cdot 0.012 = 0.002616$
3,1	$0.218 \cdot 0.274 = 0.059732$
3,2	$0.218 \cdot 0.359 = 0.078262$
3,3	$0.218 \cdot 0.218 = 0.047524$
3,4	$0.218 \cdot 0.095 = 0.027010$
3,5	$0.218 \cdot 0.042 = 0.009156$
4,0	$0.095 \cdot 0.012 = 0.001140$
4,1	$0.095 \cdot 0.274 = 0.026030$
4,2	$0.095 \cdot 0.359 = 0.034105$
4,3	$0.095 \cdot 0.218 = 0.027010$
4,4	$0.095 \cdot 0.095 = 0.009025$
4,5	$0.095 \cdot 0.042 = 0.003990$
5,0	$0.042 \cdot 0.012 = 0.000504$
5,1	$0.042 \cdot 0.274 = 0.011508$
5,2	$0.042 \cdot 0.359 = 0.015078$
5,3	$0.042 \cdot 0.218 = 0.009156$
5,4	$0.042 \cdot 0.095 = 0.003990$
5,5	$0.042 \cdot 0.042 = 0.001764$

Define X to be the number of color television sets per duplex. Then the probability distribution of the random variable X is given by this table

x	p
0	0.000144
1	0.006576
2	0.083692
3	0.201964
4	0.250625
5	0.209592

x	p
6	0.138750
7	0.084176
8	0.027337
9	0.007980
10	0.001764

b. $E(X) = 0 \bullet 0.000144 + 1 \bullet 0.006576 + 2 \bullet 0.083692 + 3 \bullet 0.201964 + 4 \bullet 0.250625$
$+ 5 \bullet 0.209592 + 6 \bullet 0.138750 + 7 \bullet 0.084176 + 8 \bullet 0.027337 + 9 \bullet 0.007980$
$+ 10 \bullet 0.001764 = 4.5602$.

The expected value is 4.5602 color televisions per duplex.

c. The expected value of number of color televisions per household is:

$$\mu = 0(0.012) + 1(0.274) + 2(0.359) + 3(0.218) + 4(0.095) + 5(0.042) = 2.236 .$$

On average, a household has 2.236 color televisions.

E8. a. The events "number of vehicles from House I" and "number of vehicles from House 2" are independent and have the same distribution. As such, we have:

P(x in I and y in II) = P(x in I) * P(y in II).

Applying this principle yields the following probability calculations:

P(2 vehicles) = P(0 in I)*P(2 in II) + P(1 in I)*P(1 in II) + P(2 in I)*P(0 in II)
= (0.087)(0.381) + (0.331)(0.331) + (0.381)(0.087) = 0.176

P(3 vehicles) = P(0 in I)*P(3 in II) + P(1 in I)*P(2 in II)
+ P(2 in I)*P(1 in II) + P(3 in I)*P(0 in II)
= 2(0.087)(0.201) + 2(0.331)(0.381) = 0.287

P(4 vehicles) = P(1 in I)*P(3 in II) + P(2 in I)*P(2 in II) + P(3 in I)*P(1 in II)
= 2(0.331)(0.201) + (0.381)(0.381) = 0.278

P(5 vehicles) = P(2 in I)*P(3 in II) + P(3 in I)*P(2 in II)
= 2(0.381)(0.201) = 0.153

P(6 vehicles) = P(3 in I)*P(3 in II)
= (0.201)(0.201) = 0.040

b. The expected value is:

0(0.008) + 1(0.058) + 2(0.176) + 3(0.287) + 4(0.278)
+5(0.153) + 6(0.040) = 3.388

So, on average, there are 3.388 cars per duplex.

c. This is exactly twice the expectation of the distribution of the number of vehicles for a single household.

E9. 13.7% of the houses, or about 69 houses, would be expected to have 3 vehicles. Each of these would need to park one vehicle on the street. 5.8% of the houses, or about 29 of them would have four vehicles, two of which would need to be parked on the street. This means about 127 vehicles would be parked on the street. With three spaces per house, the 29 houses with four vehicles would each have one vehicle on the street.

E10. In single households, 0.137 + 0.058 = 0.195, or 19.5% of households will have more than two vehicles and must park at least one vehicle on the street. According to the simulation shown in Display 6.3, 0.160 + 0.064 + 0.010 + 0.002 = 0.236, or 24% of duplexes will have more than four vehicles and must park at least one vehicle on the street. **Alternatively,** for every 500 single households there are, on average, 872.5 vehicles. About 68.5 of the households will need to park one vehicle on the street, and about 29 will need to park two vehicles on the street. This gives a proportion of 126.5/872.5 or about 14.5% of all vehicles for these 500 households will be parked on the street. For every 500 duplexes, according to the estimation in Display 6.3, there are, on average, 1765 vehicles. About 80 of the duplexes will have five vehicles and must park one vehicle on the street, about 32 of the duplexes will need to park two vehicles on the street, about 5 of the duplexes will need to park three vehicles on the street, and one duplex will need to park four vehicles on the street, for a total of 163 vehicles on the street. So 163/1765, or about 9.2% of the vehicles for these 500 duplexes will be parked on the street.
E11. a. \$134.50 profit if there is no burglary; 5000 – 134.50, or \$4865.50 loss if there is a burglary. **b.** \$134.50 + \$5000 / 1000 = \$139.50. At the \$134.50 rate the expected gain of the company is 0. Therefore, they need to add \$5.00 per customer to the rate charged. **c.** Examples include whether there is a burglar alarm; whether there are good locks on the doors; whether the house contains expensive items; what proportion of the time someone is at home.
E12. \$0(0.9798) + \$14,939(0.0202) = \$301.7678
E13. a. μ_X = \$0.185 **b.** 1,000,000(0.50 – 0.185), or \$315,000. Alternatively, the state would take in \$500,000 for 1,000,000 tickets and would pay out prizes worth \$185,000 for a profit of \$315,000. **c.** 37% of the income was returned in prizes.
E14. a. \$148,861,425 **b.** The value of all of the nonfood prizes is \$5,781,425, and $\frac{1}{10}$ of this, or \$578,142.50, is expected to be paid out. The value of the 196,000,000 cards with food prizes is \$143,080,000, and 1/5 of this, or \$28,616,000, is expected to be paid out. Burger King expected to pay out a total of \$29,194,142.50, but almost all of this is in food. **c.** The expected value of a card can be computed as $\frac{29,194,142.50}{196,011,489}$, or about \$0.149.

E15. a. Buy service contract: Expected cost is $250
Don't buy service contract: Expected cost is
$150(1)(0.30) + $150(2)(0.15) + $150(3)(0.10) + $150(4)(0.05) = $165

b. The advantage of buying the service plan is that you if the computer happens to be a lemon and require at least two repairs, you will have saved money, given that unlimited repairs are free. The advantage of not buying the service contract is that you expect to pay less than if you were to buy the service contract.

E16. a. The expected costs per month are:

AT&T: 49.99 + 100(0.1)(0.45) = $54.49
Verizon: 49.00 + 50(0.3)(0.45) + 100(0.2)(0.45) + 200(0.1)(0.45) = $73.75
Voicestream: 39.99 + 50(0.3)(0.25) + 100(0.2)(0.25) + 200(0.1)(0.25) = $53.74
Sprint PCS: 49.99 + 50(0.3)(0.30) + 100(0.2)(0.30) + 200(0.1)(0.30) = $66.49

b. Voicestream is still the cheapest (but not by a lot), followed by AT&T, Sprint PCS, and Verizon.

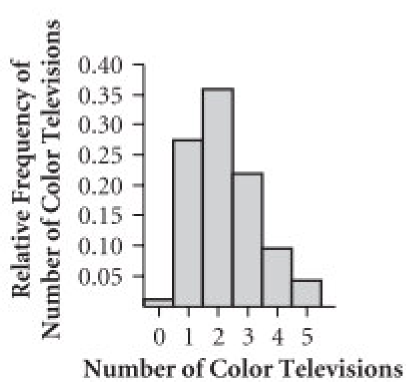
E17. a. The expected number of vehicles is 200 • 1.696 = 339.2.

b. The standard deviation can be calculated by $\sqrt{200} \cdot 0.88745926 = 12.551$. It would not be unusual to be off by 12 or so vehicles.

E18. a. $\mu_X = \sum x \cdot P(x) = 0 \bullet 0.012 + 1 \bullet 0.274 + 2 \bullet 0.359 + 3 \bullet 0.218 + 4 \bullet 0.095 + 5 \bullet 0.042 = 2.236$

$\sigma_X = 1.115$

The distribution is as follows:



b. For 30 households, the expected number of color televisions would be 30(2.236) = 67.08 and the standard deviation would be $\sqrt{30}(1.115) = 6.107$.

E19. a. The expected total weekly tips is

$$200(0.1)+300(0.3)+400(0.4)+500(0.2)=\$370.$$

The standard deviation for total weekly tips is

$$\sigma^2=(200-370)^2(0.1)+(300-370)^2(0.3)+(400-370)^2(0.4)+(500-370)^2(0.2)=\$8100$$

$$\sigma=\$90$$

b. The expected weekly salary is \$60 + \$370 = \$430. The standard deviation for the weekly salary remains \$90 since adding a constant to a random variable does not change its standard deviation.

c. Assuming that you keep 80% of the tips, the expected weekly salary would be \$60 + 0.80(\$370)= \$356. The corresponding standard deviation is 0.80(s.d. in (a)) = \$72.

E20. Let X = number of televisions you sell in a week.

a. The expected weekly commission is $\$25\mu_X$, where

$$\mu_X=0(0.10)+5(0.40)+10(0.25)+15(0.15)+20(0.10)=8.75.$$

So, the expected weekly commission is \$218.75.

The standard deviation for weekly commission is $5\sigma_X$, where

$$\sigma_X^2=(0-8.75)^2(0.10)+(5-8.75)^2(0.40)+(10-8.75)^2(0.25)$$
$$+(15-8.75)^2(0.15)+(20-8.75)^2(0.10)=32.1875.$$

So, the standard deviation is $5\sqrt{32.1875}\approx\$28.37$.

b. The expected total weekly salary is 150 + 218.75 = \$368.75. The standard deviation is the same as in (a) since adding a constant to a random variable does not change its standard deviation.

c. Arguing as in E19c, we have

$$\mu=\$368.75(0.7)=\$258.13$$
$$\sigma=0.70(28.37)=\$19.86$$

E21. a. Begin with possible combinations of working times:

Main Pump	Backup Pump	Total Working Time	Probability
1	1	2	$0.1\cdot0.2=0.02$
1	2	3	$0.1\cdot0.8=0.08$
2	1	3	$0.3\cdot0.2=0.06$
2	2	4	$0.3\cdot0.8=0.24$
3	1	4	$0.6\cdot0.2=0.12$
3	2	5	$0.6\cdot0.8=0.48$

Define X to be the number of months the two pumps together work. Then the probability distribution of the random variable X is given by this table:

x	p	x	p
2	0.02	4	0.36
3	0.14	5	0.48

b. $\mu_x = 4.3$ months, $\sigma_x = 0.781$

c. $\mu_{\text{Main}} = 2.5$ months, $\mu_{\text{Backup}} = 1.8$ months. The sum of these is 4.3 months.

E22. Begin with possible combinations of gains:

Game A	Game B	Total	Probability
−1	−2	−3	$0.4 \cdot 0.7 = 0.28$
−1	0	−1	$0.4 \cdot 0.2 = 0.08$
−1	2	1	$0.4 \cdot 0.1 = 0.04$
0	−2	−2	$0.3 \cdot 0.7 = 0.21$
0	0	0	$0.3 \cdot 0.2 = 0.06$
0	2	2	$0.3 \cdot 0.1 = 0.03$
1	−2	−1	$0.2 \cdot 0.7 = 0.14$
1	0	1	$0.2 \cdot 0.2 = 0.04$
1	2	3	$0.2 \cdot 0.1 = 0.02$
2	−2	0	$0.1 \cdot 0.7 = 0.07$
2	0	2	$0.1 \cdot 0.2 = 0.02$
2	2	4	$0.1 \cdot 0.1 = 0.01$

Define X to be the total gain from playing the two games. Then the probability distribution of the random variable X is given by this table:

x	p	x	p
−3	0.28	1	0.08
−2	0.21	2	0.05
−1	0.22	3	0.02
0	0.13	4	0.01

b. $\mu_x = -1.2$ dollars, $\sigma_x \approx 1.661$ dollars.

c. $\mu_{Game\ A} = 0$ dollars, $\mu_{Game\ B} = -1.2$ dollars. The sum of these is −1.2 dollars.

E23. a.

Difference	Probability
−3	$\frac{1}{24}$
−2	$\frac{2}{24}$
−1	$\frac{3}{24}$
0	$\frac{4}{24}$
1	$\frac{4}{24}$

Difference	Probability
2	$\frac{4}{24}$
3	$\frac{3}{24}$
4	$\frac{2}{24}$
5	$\frac{1}{24}$

b. Using the formula $\mu = \sum x \cdot P(x)$, we find this distribution has mean μ = 1. (It is also apparent by symmetry that the mean is 1.) The mean of the distribution of the outcomes of the six-sided die is 3.5, and the mean of the distribution of the outcomes of the four-sided die is 2.5. Thus, the mean of the sampling distribution of the difference is equal to $\mu_1 - \mu_2 = 3.5 - 2.5 = 1$.

Using the formula $\sigma^2 = \sum (x-\mu)^2 \cdot P(x)$, we find this distribution has variance σ^2 = 4.167. The variance of the distribution of the outcomes of the six-sided die is 2.917, and the variance of the distribution of the outcomes of the four-sided die is 1.25. Thus, the variance of the sampling distribution of the difference is indeed equal to 2.917 + 1.25 = 4.167.

E24. a. The possible outcomes for the larger number are: 1, 2, 3, 4, 5, 6
Each of these has a probability of 1/6 of occurring.
Also, $\mu = \sum x \cdot P(x) = 3.5$ and $\sigma^2 = \sum (x-\mu)^2 \cdot P(x) = 2.917$.

b. The possible outcomes for the smaller number are: 0, 1, 2, 3, 4, 5
Each of these also has a probability of 1/6 of occurring.
Also, $\mu = \sum x \cdot P(x) = 2.5$ and $\sigma^2 = \sum (x-\mu)^2 \cdot P(x) = 2.917$.

c.

Difference	Probability
1	1

The mean is μ = 1(1) = 1. The variance is $\sigma^2 = (1-1)^2 \bullet (1) = 0$.

d. Yes: μ = 3.5 – 2.5 = 1.
e. No: $0 \neq 2.917 + 2.917$. The two values are not independent.

E25. a. The expected total yearly earnings is 52($356) = $18,512.
The standard deviation for yearly earnings is $\sqrt{52}(\$72) = \$519.20$.

b. Note that 18,512 + 2.87(519.20) = 20,000. So, a yearly salary of $20,000 is nearly 3 standard deviations from the mean. So, it would be very unusual to earn more than $20,000 in one year.

E26. a. Let X = number of customers to whom you sell a TV.
For n = 1, the distribution of X is

X	**Probability**
0	0.9
1	0.1

The commission earned is given by the random variable $Y = 25X$. Its distribution is:

Y	**Probability**
0	0.9
25	0.1

b. Using Y from (a), and assuming a weekly base pay of \$150, the expected weekly earnings are

$$\mu_Y = \$150 + \left[0(0.9) + 25(0.1)\right] = \$152.50$$

with a standard deviation of

$$\sigma_Y^2 = \sigma_X^2 = (0-0.1)^2(0.9) + (1-0.1)^2(0.1) = \$0.09$$
$$\sigma_Y = \$0.30$$

c. Now, assuming you've interacted with 100 customers, all according to the same distribution as in (a), we have a binomial random variable with n = 100 and p = 0.1. The mean and standard deviation in such case are

$$\mu = 100(0.1) = 10, \qquad \sigma = \sqrt{100(0.1)(0.9)} = 3.$$

In such case, the expected weekly earnings are 150 + 25(10) = \$400 with a standard deviation of \$3.

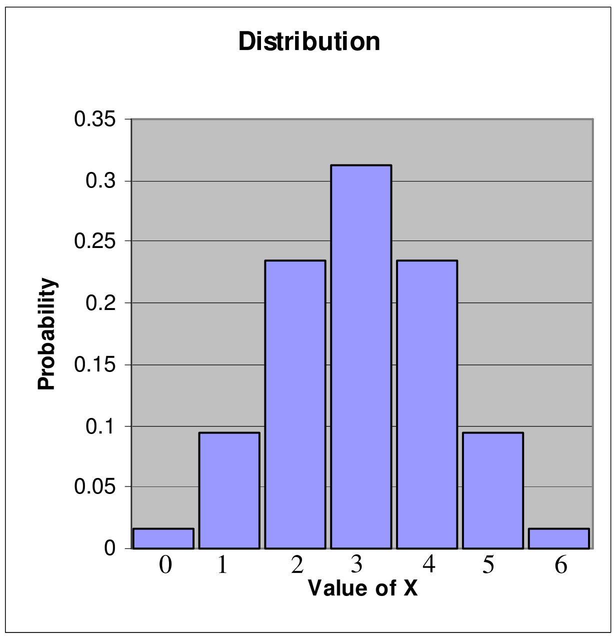
d. It would be very unusual to earn less than \$300 since such earnings would be more than 30 standard deviations below the mean earnings.

E27. Let X = number who correctly identify gourmet coffee.
The situation described is binomial with n = 6 and p = 0.5 (since equally likely to tell gourmet coffee from ordinary coffee).
The distribution of X is as follows:

X	**Probability**
0	$\binom{6}{0}(0.5)^6 \approx 1(0.015625) = 0.015625$
1	$\binom{6}{1}(0.5)^6 \approx 6(0.015625) = 0.09375$

2	$\binom{6}{2}(0.5)^6 \approx 15(0.015625) = 0.234375$
3	$\binom{6}{3}(0.5)^6 \approx 20(0.015625) = 0.3125$
4	$\binom{6}{4}(0.5)^6 \approx 15(0.015625) = 0.234375$
5	$\binom{6}{5}(0.5)^6 \approx 6(0.015625) = 0.09375$
6	$\binom{6}{6}(0.5)^6 \approx 1(0.015625) = 0.015625$

The histogram is as follows. It is symmetric and mound shaped.



b. On average, the expected number of people who correctly identify the gourmet coffee is np = 6(0.5) = 3.
c. While you would get 3 people on average who correctly identify the gourmet coffee, you could well get $\sqrt{np(1-p)} = \sqrt{6(0.5)(0.5)} = 1.225$ people more or fewer.

d. Yes, if the people can't tell the difference, the probability is only 0.015625 that all 6 will choose correctly.

E28. a.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - \binom{25}{0}(0.132)^0(0.868)^{25} + \binom{25}{1}(0.132)^1(0.868)^{24} \approx 0.8606$$

b. $\mu_X = 25(0.132) = 3.3$, and $\sigma_X = \sqrt{25(0.132)(0.868)} \approx 1.692$

E29. Let X be the number of people in the sample of 20 residents that do not have health insurance.

a. $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$

$$1-\left(\binom{20}{0}(0.15)^0(0.85)^{20}+\binom{20}{1}(0.15)^1(0.85)^{19}+\binom{20}{2}(0.15)^2(0.85)^{18}\right)=0.5951$$

b. $E(X) = np = 20 \bullet 0.15 = 3$ people, and $\sigma_X = \sqrt{np(1-p)} = \sqrt{20 \cdot 0.15 \cdot 0.85} = 1.5969$ people.

E30. a. Because the median income has half of the incomes below it and half above it, the chance of a randomly selected income exceeding the median is 0.5.

b. We have a binomial distribution with $n = 5$ and $p = 0.5$. If X denotes the number of households with income exceeding the median, the probability distribution is the following:

Number of households whose income exceeds $50,740	Probability
0	0.03125
1	0.15625
2	0.31250
3	0.31250
4	0.15625
5	0.03125

The probability of four or more households having incomes exceeding the median is 0.1875, the sum of the bottom two entries in the table. (In answering specific probability questions of this type, it is good to have students get into the habit of looking at the entire distribution.)

c. i. If X denotes the number of households with incomes below the median in a random sample of 16, then X has a binomial distribution with $n = 16$ and $p = 0.5$. Thus, $E(X) = np = 16(0.5) = 8$.

ii. The standard deviation of X is given by $\sigma_X = \sqrt{np(1-p)} = \sqrt{16(0.5)(0.5)} = 2$

iii. The exact binomial calculation gives $P(X \geq 10) \approx 0.227$.

d. This would be a very rare result, so you might suspect that the sample was not selected randomly. The sample selection method seems biased toward households with larger incomes.

E31. To get 60% correct, you have to get 3 or more on the 5-question quiz and 12 or more on the 20-question quiz. Looking at the graphs in the display, the probability of getting 3 or more correct on a 5-question quiz is 0.5. On a 20-question quiz, the probability of getting 12 or more correct is about 0.25. It would be much better to have a 5-question quiz.

E32. The standard deviation would be significantly reduced for n = 20 versus n = 5.

E33. a. Because of the \$15 spent for the tickets, in order to gain \$10 or more, you must win \$25 or more. The only way you can do this is to win on 3 or more of the tickets:

$$P(\textit{3 or more wins}) = P(X \geq 3) = 1 - P(X \leq 2) \approx 1 - 0.9429 = 0.0571.$$

b. $E(\textit{earnings}) = 15(0.06)(10) - 15(1) = -\$6.00$

c. $\sigma_W = \sqrt{np(1-p)} = \sqrt{15 \cdot 0.06 \cdot 0.94} \approx 0.9198$, where W is the number of winning tickets. The earnings $E = 10W - 15$, so the $\mu_E = 10\,\mu_W \approx 10 \cdot 0.9198$, or \$9.20.

E34. a. If the firm's gain is denoted by G and the number of wells that produce oil is denoted by X, then G = 1,000,000X – 600,000 because it costs \$60,000 to drill each of 10 wells. Then,

$$E(G) = 1{,}000{,}000 \cdot E(X) - 600{,}000 = 1{,}000{,}000(10)(0.1) - 600{,}000 = \$400{,}000$$

b. $\sigma_G = 1{,}000{,}000\sigma_X = 1{,}000{,}000\sqrt{10(0.1)(0.9)} \approx 948{,}683.30$

c. For the firm to lose money, all 10 wells must be dry. The probability of this is $(0.9)^{10} \approx 0.3487$.

d. For the firm to make \$1.5 million or more, 3 or more wells must produce oil. The probability of this is 1 – P(*0 or 1 or 2 produces*) < 1 – 0.9298 = 0.0702.

E35. a. $P(\textit{at least one alarm sounds}) = 1 - P(\textit{all alarms fail}) = 1 - (0.3)^3 = 0.973$

b. $P(X \geq 1) = 1 - P(X = 0) = 1 - (0.3)^6 = 0.99927$. No, the two probabilities are nearly the same.

c. Solving $1 - (0.3)^n = 0.99$ gives $n \approx 3.825$, which rounds up to n = 4. This equation can be solved by trial and error, by graphing, or by logarithms.

E36. a. $P(\textit{at least one works correctly}) = 1 - P(\textit{none work correctly}) = 1 - 0.08^2 = 0.9936$.

b. $1 - 0.08^3 = 0.999488$

c. Solving $1 - 0.08^n = 0.99999$ gives $n \approx 4.5583$, which rounds up to n = 5. This equation can be solved by trial and error, by graphing, or by logarithms.

E37. a. If X denotes the number of defective DVDs in the sample, then

$$P(\textit{X is at least 1}) = P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - p)^n,$$

where p is the probability of observing a defective DVD. This probability, $P(X \geq 1)$, is given as 0.5. With p = 0.1, the equation becomes $1 - (0.9)^n = 0.5$ or $n \approx 6.5788$, which rounds up to $n = 7$.

b. With p = 0.04, the equation is $1 - (0.96)^n = 0.5$ or $n \approx 16.980 \approx 17$.

E38. a. If 5% of the population is tagged, the probability of capturing at least one tagged fish in a sample of size n is $1 - 0.95^n$. Solving $1 - 0.95^n = 0.80$ gives $n \approx 31.377$, so 32 fish would need to be captured.

b. $1 - 0.98^n = 0.80$ gives 79.664, meaning that 80 fish would need to be captured.

E39. a. Observe that

$$\begin{aligned}
\mu &= 0(1-p)+1(p) = p \\
\sigma^2 &= (0-p)^2(1-p)+(1-p)^2(p) \\
&= p^2(1-p)+p(1-p)^2 \\
&= p(1-p)\big(p+(1-p)\big) = p(1-p) \\
\sigma &= \sqrt{p(1-p)}
\end{aligned}$$

b. Observe that

$$\begin{aligned}
\sigma^2_{X_1+X_2+\cdots+X_n} &= \sigma^2_{X_1}+\sigma^2_{X_2}+\cdots+\sigma^2_{X_n} \\
&= p(1-p)+p(1-p)+\cdots+p(1-p) \\
&= np(1-p)
\end{aligned}$$

So, $\sigma_{X_1+X_2+\cdots+X_n} = \sqrt{np(1-p)}$.

E40. a. The following table shows the distribution for n = 2 and probability of success p.

k	$P(X = k)$
0	$\binom{2}{0}p^0q^2 = q^2$
1	$\binom{2}{1}p^1q^1 = 2pq$
2	$\binom{2}{2}p^2q^0 = p^2$

The sum of the probabilities is $p^2 + 2pq + q^2 = (p + q)^2 = 1^2 = 1$. (recall that $p + q = 1$).

b. This table shows the distribution for n = 3.

k	$P(X = k)$
0	$\binom{3}{0} p^0 q^3 = q^3$
1	$\binom{3}{1} p^1 q^2 = 3pq^2$
2	$\binom{3}{2} p^2 q^1 = 3p^2 q$
3	$\binom{3}{3} p^3 q^0 = p^3$

The sum of the probabilities is $p^3 + 3p^2q + 3pq^2 + q^3 = (p + q)^3 = 1^3 = 1$.

Note: In general, using the binomial theorem yields

$$\begin{aligned}(q+p)^n &= \binom{n}{n} q^n p^0 + \binom{n}{n-1} q^{n-1} p^1 \\ &\quad + \binom{n}{n-2} q^{n-2} p^2 + \cdots \\ &\quad + \binom{n}{2} q^2 p^{n-2} + \binom{n}{1} q^1 p^{n-1} + \binom{n}{0} q^0 p^n \\ &= P(X=0) + P(X=1) \\ &\quad + P(X=2) + \cdots + P(X=n-2) \\ &\quad + P(X=n-1) + P(X=n)\end{aligned}$$

So the terms in the expansion are the probabilities of a binomial distribution with n trials and probability of success p. We have shown at left that

$$P(X=0) + P(X=1) + \cdots + P(X=n-1) + P(X=n) = (q+p)^n = ((1-p)+p)^n = 1^n = 1$$

So the sum of the probabilities is 1.

E41. a. There are (12)(12) = 144 possible outcomes. Fortunately, we do not have to list them because there are only three ways to get a sum of 3 or less: 1, 1; or 2, 1; or 1, 2. Thus the probability is 3/144.

b. Each of the outcomes 1, 2, 3, 4, . . . , 12 has a $\frac{1}{12}$ chance of occurring. By symmetry, the mean is 6.5. The standard deviation is 3.452.

c. The mean of the sampling distribution of the sum of two rolls of a 12-sided die is

$$\mu_{sum} = 2(6.5) = 13$$

and the standard deviation is

$$\sigma_{sum} = \sqrt{2 \cdot 3.452^2} = 4.882.$$

E42. The outcome of the roll of a single die, *X*, has mean 3.5 and variance 2.917. The mean of two rolls, written as $\frac{1}{2}(X_1 + X_2)$, has mean (1/2)(3.5 + 3.5) = 3.5 and variance (1/4)(2.917 + 2.917) = 2.917/2 = 1.4585. Thus, the mean of the sampling distribution of the difference between your average and your partner's average is

$$\mu_{you-partner} = \mu_{you} - \mu_{partner} = 3.5 - 3.5 = 0.$$

The standard error is

$$\sigma_{you-partner} = \sqrt{\sigma^2_{you} + \sigma^2_{partner}} = \sqrt{1.4585 + 1.4585} = 1.708$$

The distribution will be mound-shaped and symmetric about the mean of 0.

E43. a. You are looking for the distribution of the minimum lifelength of the two pumps, since the system stops when the first pump stops. The combinations can be displayed in a two-way table. Since the pumps operate independently, the probabilities can be calculated by multiplying the individual probabilities.

		Pump 1		
		1 month	2 months	3 months
	1 month	0.01	0.03	0.06
Pump 2	2 months	0.03	0.09	0.18
	3 months	0.06	0.18	0.36

Define *X* to be the minimum lifelength of Pumps 1 and 2. Then the probability distribution of the random variable *X* is given by this table:

x	*p*
1 month	0.19
2 months	0.45
3 months	0.36

b. The lifelength of System II is the maximum lifelength of the two pumps. Using the same table, replacing 'Pump 1' and 'Pump 2' with 'Pump 3' and 'Pump 4,' respectively:

Define *X* to be the maximum lifelength of Pumps 3 and 4. Then the probability distribution of the random variable *X* is given by this table:

x	*p*
1 month	0.01
2 months	0.15
3 months	0.84

c. The expected lifelength of System I is 2.17 months and the expected lifelength for System II is 2.83 months. You should recommend System II because it has a larger expected lifelength.

E44. **a.** In order to get the number in 300,000 who suffer from each type of accident, multiply the table values by 300,000. This yields:

Event	Number in 300,000 who suffer
Stairs	1,143,000
Bicycle	489,000
Jewelry	84,000
Lightning Strike	750

b. The expected cost to the insurance company for an injury associated with stairs is:

$$100{,}000(0.0381) = \$3{,}810.$$

So, a customer should pay $3,810 for such a policy to ensure the insurance company breaks even, on average.

c. Arguing as in (b), the expected cost is 10,000(0.000025) = $0.25.

d. Solving the equation $x(0.0028) = 50$ for x yields $x = \$17{,}857.14$.

E45. The chance that at least one donation out of ten will be type B is

$$1-(0.9)^{10} \approx 1-0.349 \approx 0.651$$

You are assuming that the probability that the first donation checked is type B is 0.1, that the probability that any subsequent donation is type B is 0.1, and that the probability that a donation is type B is independent of the type of the donations previously checked.

E46. Using the binomial model for the number of students that apply out of n = 120,000, with p = 0.1, gives a mean of 12,000 and a standard deviation of about 104. But, the binomial model might not be a good model here, because these are not 120,000 independent decisions with p = 0.1 for the "success" of each one. Students tend to apply where their friends apply and certain high schools encourage groups of similarly qualified students to apply. Also, the probability of applying to any college is not the same across all graduates because, for example, some do not apply to any college.

E47. **a.** Let X = number of women who have the BS degree. Then, X is binomial with $n = 20$ and $p = 0.33$. $P(X > 10) = 0.0350$; assuming the women are selected randomly and independently and that the probability is 0.33 for each one.

b. Let X = number of men who have the BS degree. Then, X is binomial with $n = 20$ and $p = 0.26$. $P(X > 10) = 0.0132$; assuming the men are selected randomly and independently and that the probability is 0.26 for each one.

c. Assuming that the proportion of all people with bachelor's degrees is about 29.5% (which is the average of the probabilities used in (a) and (b).
Let X = number of people with BS degree. Then, X is binomial with $n = 40$ and p =0.295. $P(X > 20) = 0.0019$.

E48. Let X = number of people with health insurance. Then,

$$E(X) = np = 500(0.85) = 425.$$

So, the expected cost is 40(425) + 20 (75) = $18,500.

E49. a. On a TI-83 Plus or TI-84 Plus calculator, enter the command **binompdf(130,0.9)**, press $\boxed{\text{STO}\rightarrow}$ $\boxed{\text{2nd}}$ **[L1]** $\boxed{\text{ENTER}}$, which stores the probabilities of the different numbers of no shows in list **L1.** The first *P*(*X*) in the list below, 0.84793, is a cumulative probability.

Number of Passengers Who Show	Payout x	Probability p
≤120	0	0.84793
121	100	0.06399
122	200	0.04248
123	300	0.02487
124	400	0.01263
125	500	0.00546
126	600	0.00195
127	700	0.00055
128	800	0.00012
129	900	0.00002
130	1000	0.00000

b. $E(payout) = \sum x \cdot P(x) \approx \$31.807$.

E50. a. The nine possible totals are 300, 450, 900, 325, 475, 925, 700, 850, and 1300. The standard deviation is 313.803, and the variance is 98,472.22. (You must compute all variances in this question by dividing by *n,* not *n* – 1, because you are dealing with a population in which the true mean is known.)

b. The variance of List A is 33,472.22 and of List B is 65,000. The variance of List A plus the variance of List B exactly equals the variance of the nine totals computed in (a).

c. The standard deviation of List A is 182.95 and of List B is 254.95. The standard deviation of List A plus the standard deviation of List B (437.9) does *not* equal the standard deviation of the nine totals computed in part a!

d. Yes: 341.67 + 350 = 691.67.

e. No: 225 + 250 doesn’t equal 700, the median of the totals. Sums and differences of random variables are often used as summary statistics in part because their means and variances are easily computed, especially when the variables are independent.

E51. a.

GAME B	
Value of Prize	**Probability**
$0.69	0.25
$0.00	0.75

GAME C	
Value of Prize	**Probability**
$1.44	0.125
$0.00	0.875

GAME D	
Value of Prize	**Probability**
$1.99	0.0625
$0.00	0.9375

The expected winnings for each game are:

Game A: $0.55(0.50) = $0.275
Game B: $0.69(0.25) = $0.1725
Game C: $1.44(0.125) = $0.18
Game D: $1.99(0.0625) = $0.124375

Game A yields the highest expected earnings.

b. Some may not want the drink or might just want to try for the bigger prize.

E52. There are 12 outcomes. These outcomes with their probabilities are listed below:

Number hours tutoring	Number hours dance lessons	Combined earnings	Probability
1	0	$12	(0.3)(0.4) = 0.12
1	1	$4	(0.3)(0.3) = 0.09
1	2	- $4	(0.3)(0.3) = 0.09
2	0	$24	(0.3)(0.4) = 0.12
2	1	$16	(0.3)(0.3) = 0.09
2	2	$8	(0.3)(0.3) = 0.09
3	0	$36	(0.2)(0.4) = 0.08
3	1	$28	(0.2)(0.3) = 0.06
3	2	$20	(0.2)(0.3) = 0.06
4	0	$48	(0.2)(0.4) = 0.08
4	1	$40	(0.2)(0.3) = 0.06
4	2	$32	(0.2)(0.3) = 0.06

So, the expected weekly earnings are obtained by multiplying the rows in the rightmost two columns and summing. This yields $20.40.

To obtain the associated variance, subtract 20.4 from each value in the third column, square it, multiply by the probability in the rightmost column, and sum all of these values. This yields 218.4. The standard deviation is then $\sqrt{218.4} \approx \$14.78$.
E53. The probability of rolling an even number in one roll of a six-sided die is 3/6 or 0.5. Let X = number of even numbers rolled from seven rolls of a die. $$P(X=2) = \binom{7}{2}(0.5)^7 = 0.1641$$ $$P(X \geq 4) = \binom{7}{4}(0.5)^7 + \binom{7}{5}(0.5)^7 + \binom{7}{6}(0.5)^7 + \binom{7}{7}(0.5)^7 = 0.5$$
E54. a. $P(even) = 0.5$, so we are looking at $P(X = 3)$, where X is the number of even two digit numbers. $P(X=3) = \binom{5}{3}(0.5)^5 = 0.3125$. **b.** $P(sum \geq 9) = 54/90 = 0.55$. We want $P(X = 1)$, where X is the number of two-digit numbers whose digits sum to a number greater than or equal to 9. $$P(X=1) = \binom{5}{1}(0.55)^1(0.45)^4 \approx 0.11277$$ **c.** $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{5}{0}(0.55)^0(0.45)^5 = 0.981547$.

Concept Review Solutions

C1. B. To have a total of 2 children, either (first woman has 0 AND second has 2) or (first woman has 1 AND second has 1) or (first woman has 2 AND second has 0). So, the probability is 0.18(0.35) + 0.17(0.17) + 0.35(0.18) = 0.1549.
C2. B. The expected winnings are \$0.50. The variance is $$(5-0.50)^2(0.1) + (0-0.50)^2(0.9) = 2.25,$$ so that the standard deviation is \$1.50.
C3. B. Let X = number of year June lives and Y = number of years Ray lives. Observe that $$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 = 3^2 + 4^2 = 25,$$ and so $\sigma = 5$ years.
C4. D. There are $\binom{5}{2} = 10$ ways to select which 2 days he goes jogging, and then for each pattern (such as jog, no jog, no jog, jog, no jog) the probability is $\left(\frac{2}{6}\right)^2\left(\frac{4}{6}\right)^3$, so the net result is $10\left(\frac{2}{6}\right)^2\left(\frac{4}{6}\right)^3 \approx 0.329$.

C5. D. Dawna's mean walking time is 170 minutes, and we are given that her standard deviation is 10 minutes. Since Jeanne completes her run in half the time as Dawna's routine, her mean and standard deviation are each half those of Dawna; that is, a mean of 85 minutes with a standard deviation of 5 minutes. Since she leaves at 1:30pm, her average finishing time is 2:55pm with a standard deviation of 5 minutes.

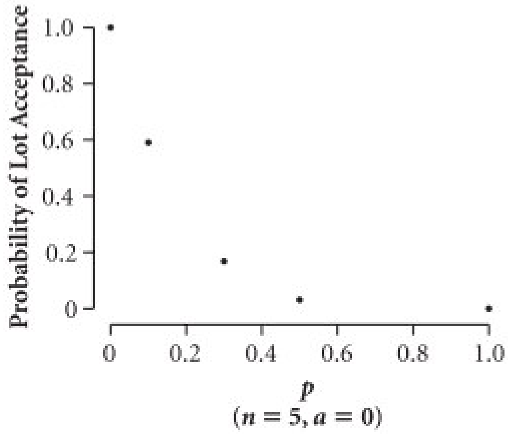
C6. D. The mean is $np = 100(0.35) = 35$, and the standard deviation is

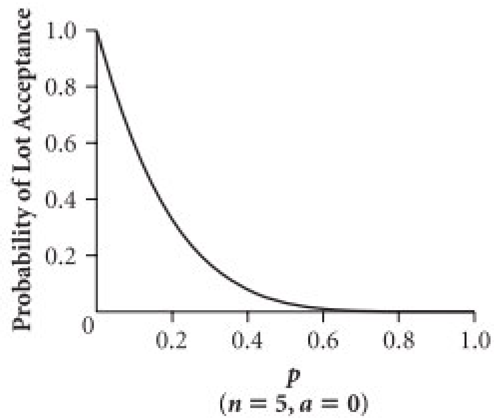
$$\sqrt{np(1-p)} = \sqrt{100(0.35)(0.65)} \approx 4.7697\,.$$

C7. **a.** The lot is accepted when there are no defective items in a random sample of size 5. On the TI-83+ and TI-84+, the command **binompdf(5,p,0)** gives the probability that there will be no defective items for each p. Alternatively, you can compute the probability that all five items will be good using $(1-p)^5$. The probabilities are given in the following table:

p	*P(no defectives)*
0	1
0.1	0.59049
0.3	0.16807
0.5	0.03125
1	0

b. The operating characteristic curve for the points in the table is shown in the first plot below. The probability of acceptance is high when the proportion of defective items in the lot is low and drops off rapidly as the proportion of defectives increases.

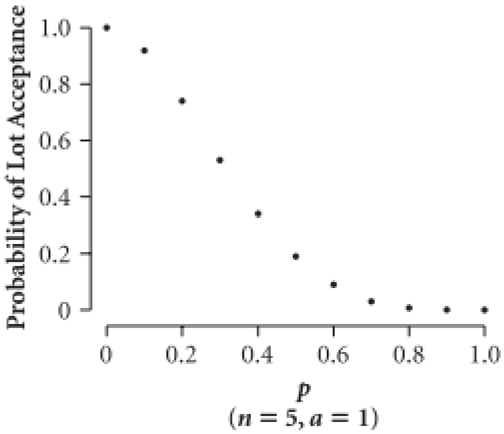


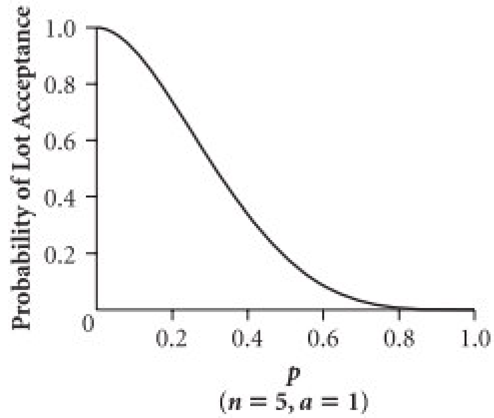


The second plot is the actual operating characteristic curve, which is a smooth continuous curve for all values of p, $0 \le p \le 1$. You can graph this curve on the TI-83+ or TI-84+ calculator by defining $Y_1 =$ **binompdf(5,X,0)** in the Y= menu and graphing Y_1 in an appropriate window. Here, X stands for the proportion of defective items, p.

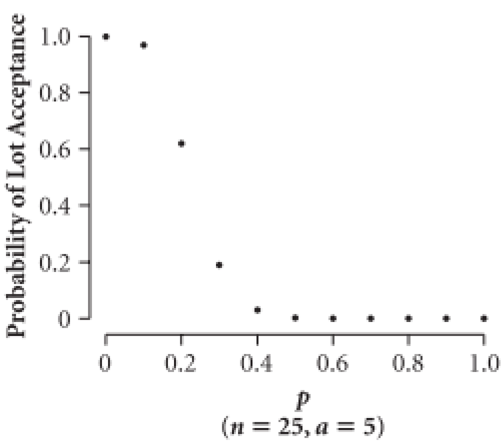
c. To compute the probability of 1 or fewer defective items for a sample size of 5, use the calculator command **binomcdf(5,p,1)**. To compute the probability of 5 or fewer defective items for a sample of size 25, use the command **binomcdf(25,p,5)**.

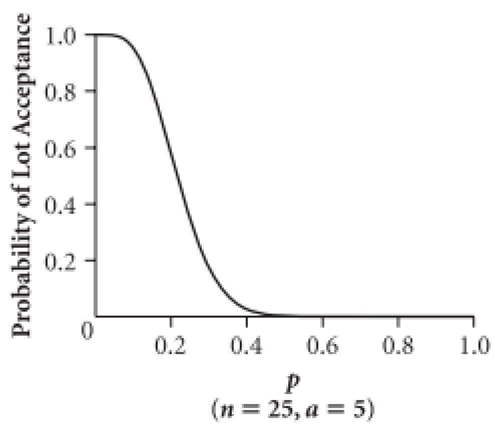
The operating characteristic curves for the two plans are shown in the plots below. Again, these will be smooth continuous curves for all values of p, $0 \le p \le 1$. On a calculator, the equation for the graph when n = 5 and a = 1 is $Y_1 =$ **binomcdf**(5,X,1).





On a calculator, the equation for the graph when $n = 25$ and $a = 5$ is
$Y_1 =$ **binomcdf**(25,X,5).





d. You would want a high probability of acceptance for p between 0 and 0.1. Both plans have high acceptance probabilities over this interval, but they are slightly higher for the plan ($n = 25$, $a = 5$).

e. At 30% defectives, the plan ($n = 5$, $a = 1$) has an acceptance probability around 0.5, which is quite high. (You would be accepting many bad lots.) The plan ($n = 25$, $a = 5$) has an acceptance probability of less than 0.20 when p is 0.30, and drops rapidly as p increases from there, which is what you are looking for.

C8. Let X be the event that a woman is chosen. Then, the complement of X, denoted $\tilde{X}$, is the event that a man is chosen. We are given that

$$P(X) = 0.57, \quad P(\tilde{X}) = 0.43.$$

Assume choices are independent.

a. 0.43

b. $P(X)P(\tilde{X}) = (0.57)(0.43) = 0.2451$

c. $P(X)P(X)P(\tilde{X}) = (0.57)^2(0.43) = 0.1397$

d. $\underbrace{P(X)\cdots P(X)}_{n-1 \text{ times}} P(\tilde{X}) = (0.57)^{n-1}(0.43)$

e.

n	**Probability of first male chosen at *n*th stage**
1	0.43
2	0.2451
3	0.1397
4	0.0796
5	0.0454
6	0.0259
7	0.0147
8	0.0084

The histogram is as follows. It is skewed right.