CHAPTER 3

Section 3.1

1.

$\mathcal{S}$:	*FFF*	*SFF*	*FSF*	*FFS*	*FSS*	*SFS*	*SSF*	*SSS*
X:	0	1	1	1	2	2	2	3

3. Examples include: M = the difference between the large and the smaller outcome with possible values 0, 1, 2, 3, 4, or 5; $T = 1$ if the sum of the two resulting numbers is even and $T = 0$ otherwise, a Bernoulli random variable. See the back of the book for other examples.

5. No. In the experiment in which a coin is tossed repeatedly until a H results, let $Y = 1$ if the experiment terminates with at most 5 tosses and $Y = 0$ otherwise. The sample space is infinite, yet Y has only two possible values. See the back of the book for another example.

7.

a. Possible values of X are 0, 1, 2, …, 12; discrete.

b. With n = # on the list, values of Y are 0, 1, 2, … , N; discrete.

c. Possible values of U are 1, 2, 3, 4, … ; discrete.

d. Possible values of X are $(0, \infty)$ if we assume that a rattlesnake can be arbitrarily short or long; not discrete.

e. With c = amount earned in royalties per book sold, possible values of Z are 0, c, $2c$, $3c$, … , $10{,}000c$; discrete.

f. Since 0 is the smallest possible pH and 14 is the largest possible pH, possible values of Y are [0, 14]; not discrete.

g. With m and M denoting the minimum and maximum possible tension, respectively, possible values of X are $[m, M]$; not discrete.

h. The number of possible tries is 1, 2, 3, …; each try involves 3 coins, so possible values of X are 3, 6, 9, 12, 15, …; discrete.

9.

a. Returns to 0 can occur only after an even number of tosses, so possible X values are 2, 4, 6, 8, …. Because the values of X are enumerable, X is discrete.

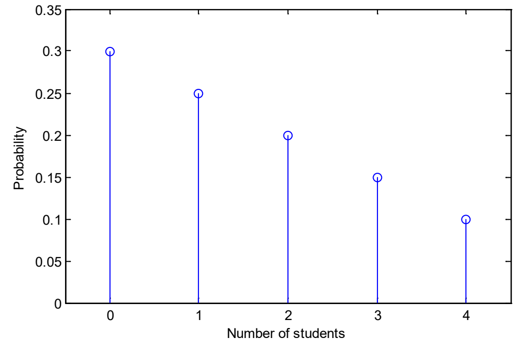
b. Now a return to 0 is possible after any number of tosses greater than 1, so possible values are 2, 3, 4, 5, …. Again, X is discrete.

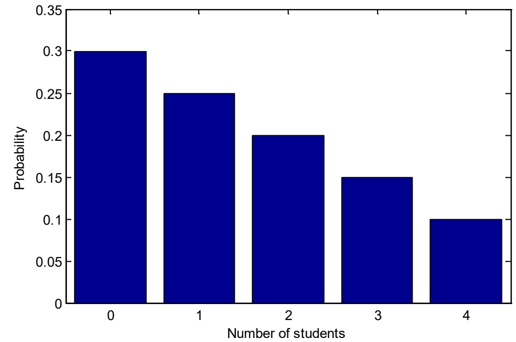
Section 3.2

11.

a. The sum of the probabilities must be 1, so $p(4) = 1 - [p(0) + p(1) + p(2) + p(3)] = 1 - .90 = .10$.

b.





c. $P(X \geq 2) = p(2) + p(3) + p(4) = .20 + .15 + .10 = .45$. $P(X > 2) = p(3) + p(4) = .15 + .10 = .25$.

d. This was just a joke — professors should always show up for office hours!

13.

a. $P(X \leq 3) = p(0) + p(1) + p(2) + p(3) = .10+.15+.20+.25 = .70$.

b. $P(X < 3) = P(X \leq 2) = p(0) + p(1) + p(2) = .45$.

c. $P(X \geq 3) = p(3) + p(4) + p(5) + p(6) = .55$.

d. $P(2 \leq X \leq 5) = p(2) + p(3) + p(4) + p(5) = .71$.

e. The number of lines not in use is $6 - X$, and $P(2 \leq 6 - X \leq 4) = P(-4 \leq -X \leq -2) =$
$P(2 \leq X \leq 4) = p(2) + p(3) + p(4) = .65$.

f. $P(6 - X \geq 4) = P(X \leq 2) = .10 + .15 + .20 = .45$.

15.

a. (1,2) (1,3) (1,4) (1,5) (2,3) (2,4) (2,5) (3,4) (3,5) (4,5)

b. X can only take on the values 0, 1, 2. $p(0) = P(X = 0) = P(\{(3,4)\ (3,5)\ (4,5)\}) = 3/10 = .3$;
$p(2) = P(X = 2) = P(\{(1,2)\}) = 1/10 = .1$; $p(1) = P(X = 1) = 1 - [p(0) + p(2)] = .60$; and otherwise $p(x) = 0$.

c. $F(0) = P(X \leq 0) = P(X = 0) = .30$;
$F(1) = P(X \leq 1) = P(X = 0 \text{ or } 1) = .30 + .60 = .90$;
$F(2) = P(X \leq 2) = 1$.
Therefore, the complete cdf of X is

$$F(x) = \begin{cases} 0 & x < 0 \\ .30 & 0 \le x < 1 \\ .90 & 1 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

17.

a. $p(2) = P(Y = 2) = P(\text{first 2 batteries are acceptable}) = P(AA) = (.9)(.9) = .81$.

b. $p(3) = P(Y = 3) = P(UAA \text{ or } AUA) = (.1)(.9)^2 + (.1)(.9)^2 = 2[(.1)(.9)^2] = .162$.

c. The fifth battery must be an A, and exactly one of the first four must also be an A. Thus, $p(5) = P(AUUUA \text{ or } UAUUA \text{ or } UUAUA \text{ or } UUUAA) = 4[(.1)^3(.9)^2] = .00324$.

d. $p(y) = P(\text{the } y^{\text{th}} \text{ is an } A \text{ and so is exactly one of the first } y - 1) = (y - 1)(.1)^{y-2}(.9)^2$, for $y = 2, 3, 4, 5, \ldots$.

19.

a. First, $1 + 1/x > 1$ for all $x = 1, \ldots, 9$, so $\log(1 + 1/x) > 0$. Next, check that the probabilities sum to 1: $\sum_{x=1}^{9} \log_{10}(1 + 1/x) = \sum_{x=1}^{9} \log_{10}\left(\frac{x+1}{x}\right) = \log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) + \cdots + \log_{10}\left(\frac{10}{9}\right)$; using properties of logs, this equals $\log_{10}\left(\frac{2}{1} \times \frac{3}{2} \times \cdots \times \frac{10}{9}\right) = \log_{10}(10) = 1$.

b. Using the formula $p(x) = \log_{10}(1 + 1/x)$ gives the following values: $p(1) = .301$, $p(2) = .176$, $p(3) = .125$, $p(4) = .097$, $p(5) = .079$, $p(6) = .067$, $p(7) = .058$, $p(8) = .051$, $p(9) = .046$. The distribution specified by Benford's Law is <u>not</u> uniform on these nine digits; rather, lower digits (such as 1 and 2) are much more likely to be the lead digit of a number than higher digits (such as 8 and 9).

c. The jumps in $F(x)$ occur at 0, … , 8. We display the cumulative probabilities here: $F(1) = .301$, $F(2) = .477$, $F(3) = .602$, $F(4) = .699$, $F(5) = .778$, $F(6) = .845$, $F(7) = .903$, $F(8) = .954$, $F(9) = 1$. So, $F(x) = 0$ for $x < 1$; $F(x) = .301$ for $1 \le x < 2$; $F(x) = .477$ for $2 \le x < 3$; etc.

d. $P(X \le 3) = F(3) = .602$; $P(X \ge 5) = 1 - P(X < 5) = 1 - P(X \le 4) = 1 - F(4) = 1 - .699 = .301$.

21. The jumps in $F(x)$ occur at $x = 0, 1, 2, 3, 4, 5$, and 6, so we first calculate $F(x)$ at each of these values:
$F(0) = P(X \le 0) = P(X = 0) = p(0) = .10$, $F(1) = P(X \le 1) = p(0) + p(1) = .25$,
$F(2) = P(X \le 2) = p(0) + p(1) + p(2) = .45$, $F(3) = .70$, $F(4) = .90$, $F(5) = .96$, and $F(6) = 1$.
The complete cdf of X is

$$F(x) = \begin{cases} .00 & x < 0 \\ .10 & 0 \le x < 1 \\ .25 & 1 \le x < 2 \\ .45 & 2 \le x < 3 \\ .70 & 3 \le x < 4 \\ .90 & 4 \le x < 5 \\ .96 & 5 \le x < 6 \\ 1.00 & 6 \le x \end{cases}$$

Then **a.** $P(X \le 3) = F(3) = .70$, **b.** $P(X < 3) = P(X \le 2) = F(2) = .45$, **c.** $P(X \ge 3) = 1 - P(X \le 2) = 1 - F(2) = 1 - .45 = .55$, **d.** $P(2 \le X \le 5) = F(5) - F(1) = .96 - .25 = .71$.

23.

a. Possible X values are those values at which $F(x)$ jumps, and the probability of any particular value is the size of the jump at that value. Thus we have:

x	1	3	4	6	12
$p(x)$	.30	.10	.05	.15	.40

b. $P(3 \le X \le 6) = F(6) - F(3-) = .60 - .30 = .30$; $P(4 \le X) = 1 - P(X < 4) = 1 - F(4-) = 1 - .40 = .60$.

25.

a. Possible X values are 1, 2, 3, …

$p(1) = P(X = 1) = P(\text{return home after just one visit}) = \frac{1}{3}$;

$p(2) = P(X = 2) = P(\text{visit a second friend, and then return home}) = \frac{2}{3} \cdot \frac{1}{3}$;

$p(3) = P(X = 3) = P(\text{three friend visits, and then return home}) = \left(\frac{2}{3}\right)^2 \cdot \frac{1}{3}$;

and in general, $p(x) = \left(\frac{2}{3}\right)^{x-1} \cdot \frac{1}{3}$ for $x = 1, 2, 3, \ldots$

b. The number of straight line segments is $Y = 1 + X$ (since the last segment traversed returns Alvie to 0). Borrow the answer from **a**, and $p(y) = \left(\frac{2}{3}\right)^{y-2} \cdot \frac{1}{3}$ for $y = 2, 3, \ldots$.

c. Possible Z values are 0, 1, 2, 3 , …. In what follows, notice that Alvie can't visit two female friends in a row or two male friends in a row.

$p(0) = P(\text{male first and then home}) = \frac{1}{2}\frac{1}{3} = \frac{1}{6}$;

$p(1) = P(\text{exactly one visit to a female}) = P(F \text{ first, then } 0) + P(F, M, 0) + P(M, F, 0) + P(M, F, M, 0)$ $= \frac{1}{2}\frac{1}{3} + \frac{1}{2}\frac{2}{3}\frac{1}{3} + \frac{1}{2}\frac{2}{3}\frac{1}{3} + \frac{1}{2}\frac{2}{3}\frac{2}{3}\frac{1}{3} = \frac{25}{54}$; for the event $Z = 2$, two additional visits occur, and the probability of those is $\frac{2}{3}\frac{2}{3} = \frac{4}{9}$, so $p(2) = \frac{4}{9}p(1) = \frac{4}{9} \cdot \frac{25}{54}$; similarly, $p(3) = \frac{4}{9}p(2) = \left(\frac{4}{9}\right)^2 \cdot \frac{25}{54}$; and so on. Therefore, $p(0) = \frac{1}{6}$ and $p(z) = \left(\frac{4}{9}\right)^{z-1} \cdot \frac{25}{54}$ for $z = 1, 2, 3, \ldots$.

27. If $x_1 < x_2$, then $F(x_2) = P(X \le x_2) = P(\{X \le x_1\} \cup \{x_1 < X \le x_2\}) = P(X \le x_1) + P(x_1 < X \le x_2)$. Since all probabilities are non-negative, $P(X \le x_1) + P(x_1 < X \le x_2) \ge P(X \le x_1) + 0 = P(X \le x_1) = F(x_1)$. That is, $x_1 < x_2$ implies $F(x_1) \le F(x_2)$, QED.

Looking at the proof above, $F(x_1) = F(x_2)$ iff $P(x_1 < X \le x_2) = 0$.

Section 3.3

29.

a. $E(Y) = \sum_{y=0}^{3} y \cdot p(y) = 0(.60) + 1(.25) + 2(.10) + 3(.05) = .60$ moving violations.

b. Use the Law of the Unconscious Statistician:

$E(100Y^2) = \sum_{y=0}^{3} 100y^2 \cdot p(y) = 0(.60) + 100(.25) + 400(.10) + 900(.05) = \$110$.

31.

a. $E(X) = (13.5)(.2) + (15.9)(.5) + (19.1)(.3) = 16.38$; $E(X^2) = (13.5)^2(.2) + (15.9)^2(.5) + (19.1)^2(.3) = 272.298$. Put these together, and $V(X) = E(X^2) - [E(X)]^2 = 272.298 - (16.38)^2 = 3.9936$.

b. Use the linearity/rescaling property of expected value: $E(17X + 180) = 17\mu + 180 = 17(16.38) + 180 =$ \$458.46. Alternatively, you can figure out the price for each of the three freezer types and take the weighted average.

c. Use the linearity/rescaling property of standard deviation: $\text{SD}(17X + 180) = |17|\sigma = 17\sqrt{3.9936} =$ \$33.97. Alternatively, you can figure out the price for each of the three freezer types and determine the standard deviation of that distribution.

d. We cannot use the linearity/rescaling properties for $E(X - .01X^2)$, since this isn't a linear function of X. However, since we've already found both $E(X)$ and $E(X^2)$, we may as well use them: the expected actual capacity of a freezer is $E(X - .01X^2) = E(X) - .01E(X^2) = 16.38 - .01(272.298) = 13.657$ cubic feet. Alternatively, you can figure out the actual capacity for each of the three freezer types and take the weighted average.

33. Yes, the expectation is finite. $E(X) = \sum_{x=1}^{\infty} x \cdot p(x) = \sum_{x=1}^{\infty} x \cdot \frac{c}{x^3} = c\sum_{x=1}^{\infty}\frac{1}{x^2}$; it is a well-known result from the theory of infinite series that $\sum_{x=1}^{\infty}\frac{1}{x^2} < \infty$, so $E(X)$ is finite.

35. You have to be careful here: if \$0 damage is incurred, then there's no deductible for the insured driver to pay! Here's one approach: let $h(X)$ = the amount paid by the insurance company on an accident claim, which is \$0 for a "no damage" event and \$500 less than actual damages ($X - 500$) otherwise. The pmf of $h(X)$ looks like this:

x	0	1000	5000	10000
$h(x)$	0	500	4500	9500
$p(x)$	.8	.1	.08	.02

Based on the pmf, the average payout across these types of accidents is $E(h(X)) = 0(.8) + 500(.1) + 4500(.08) + 9500(.02) =$ \$600. If the insurance company charged \$600 per client, they'd break even (a bad idea!). To have an expected profit of \$100 — that is, to have a *mean* profit of \$100 per client — they should charge \$600 + \$100 = \$700.

37. Determine the expected payout from rolling the die: using the Law of the Unconscious Statistician,

$$E[h(X)] = E[h(X)] = E\left[\frac{350}{X}\right] = \sum_{x=1}^{6}\left(\frac{350}{x}\right)\cdot p(x) = \sum_{x=1}^{6}\left(\frac{350}{x}\right)\cdot\frac{1}{6} = \frac{350}{6}\sum_{x=1}^{6}\frac{1}{x} = \$142.92.$$

Since \$142.92 > \$100, you expect to win more if you gamble.

Notes: (1) Even though $E[X] = 3.5$ (the average of 1, 2, 3, 4, 5, 6), the mean of $350/X$ is <u>not</u> 350/3.5.
(2) In general, if $h(x)$ is concave up then $E[h(X)] > h(E(X))$, while the opposite is true if $h(x)$ is concave down.

39.

a. For a single number, $E(X) = –1(37/38) + (35)(1/38) = –2/38 = –\$1/19$, or about –5.26 cents.
For a square, $E(X) = –1(34/38) + (8)(4/38) = –2/38 = –\$1/19$, or about –5.26 cents.

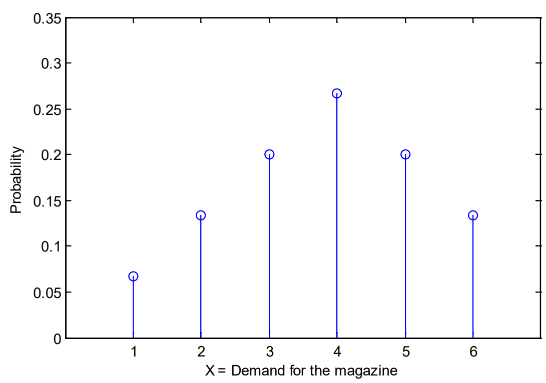
b. The expected return for a \$1 wager on roulette is the same no matter how you bet. (Specifically, in the long run players lose a little more than 5 cents for every \$1 wagered. This is how the payoffs are designed.) This seems to contradict our intuitive notion that betting on 1 of 38 numbers is riskier than betting on a color (18 of 38). The real lesson is that expected value does not, in general, correspond to the riskiness of a wager.

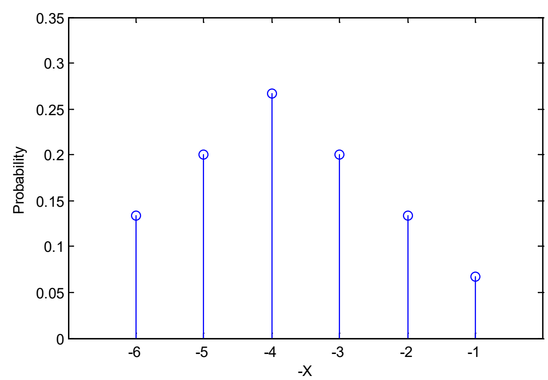
c. From Example 2.24, the standard deviation from a \$1 wager on a color is roughly \$1.
For a single number, $V(X) = (–1 – [–1/19])^2(37/38) + (35 – [–1/19])^2(1/38) = 33.208$, so $SD(X) \approx \$5.76$.
For a square, $V(X) = (–1 – [–1/19])^2(34/38) + (8 – [–1/19])^2(4/38) = 7.6288$, so $SD(X) \approx \$2.76$.

d. The standard deviations, in increasing order, are \$1 (color) < \$2.76 (square) < \$5.76 (single number). So, unlike expected value, standard deviation does correspond to our natural sense of which bets are riskier and which are safer. Specifically, low-risk/low-reward bets (such as a color) have smaller standard deviation than high-risk/high-reward bets (such as a single number).

41.

a. The line graph of the pmf of $–X$ is just the line graph of the pmf of X reflected about zero, so both have the same degree of spread about their respective means, suggesting $V(–X) = V(X)$.





b. With $a = –1$ and $b = 0$, $V(–1X + 0) = (–1)^2V(X) = V(X)$.

43.

a. By linearity of expectation, $E[X(X–1)] = E(X^2 – X) = E(X^2) – E(X) \Rightarrow$
$E(X^2) = E[X(X–1)] + E(X) = 27.5 + 5 = 32.5$.

b. $V(X) = E(X^2) – [E(X)]^2 = 32.5 – (5)^2 = 7.5$.

c. Substituting **a** into **b**, $V(X) = E[X(X–1)] + E(X) – [E(X)]^2$.

45.

a. See the table below.

k	2	3	4	5	10
$1/k^2$	.25	.11	.06	.04	.01

b. From the table in Exercise 13, $\mu = 2.64$ and $\sigma^2 = 2.3704 \Rightarrow \sigma = 1.54$.
For $k = 2$, $P(|X - \mu| \geq 2\sigma) = P(|X - 2.64| \geq 2(1.54)) = P(X \geq 2.64 + 2(1.54)$ or $X \leq 2.64 - 2(1.54)) = P(X \geq 5.72$ or $X \leq -.44) = P(X = 6) = .04$. Chebyshev's bound of .25 is much too conservative.

For $k = 3$, 4, 5, or 10, $P(|X - \mu| \geq k\sigma)$ turns out to be zero, whereas Chebyshev's bound is positive. Again, this points to the conservative nature of the bound $1/k^2$.

c. X is $\pm d$ with probability 1/2 each. So, $E(X) = d(1/2) + (-d)(1/2) = 0$; $V(X) = E(X^2) - \mu^2 = E(X^2) = d^2(1/2) + (-d)^2(1/2) = d^2$; and $SD(X) = d$.
$P(|X - \mu| < \sigma) = P(|X - 0| < d) = P(-d < X < d) = 0$, since $X = \pm d$. Here, Chebyshev's inequality gives no useful information: for $k = 1$ it states that at least $1 - 1/1^2 = 0$ percent of a distribution lies within one standard deviation of the mean.

d. $\mu = 0$ and $\sigma = 1/3$, so $P(|X - \mu| \geq 3\sigma) = P(|X| \geq 1) = P(X = -1 \text{ or } +1) = \frac{1}{18} + \frac{1}{18} = \frac{1}{9}$, identical to Chebyshev's upper bound of $1/k^2 = 1/3^2 = 1/9$.

e. There are many. For example, let $p(-1) = p(1) = \frac{1}{50}$ and $p(0) = \frac{24}{25}$.

Section 3.4

47. $M_X(t) = \sum_{x=1}^{\infty} e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx} (.5)^x = \sum_{x=1}^{\infty} (.5e^t)^x = (.5e^t)[1 + (.5e^t) + (.5e^t)^2 + \cdots] = \frac{.5e^t}{1 - .5e^t} = \frac{e^t}{2 - e^t}$, provided $|.5e^t| < 1$. (This condition is equivalent to $t < \ln(2)$ or $(-\infty, \ln(2))$ for the domain of M_X, which contains an open interval around 0.) From this, $E(X) = M_X'(0) = \left.\frac{2e^t}{(2 - e^t)^2}\right|_{t=0} = 2$.

Next, $E(X^2) = M_X''(0) = \left.\frac{2e^t(2 + e^t)}{(2 - e^t)^3}\right|_{t=0} = 6$, from which $V(X) = 6 - 2^2 = 2$ and $SD(X) = \sqrt{2}$.

49.

a. $M_X(t) = \sum_{x=9}^{12} e^{tx} p(x) = .01e^{9t} + .05e^{10t} + .16e^{11t} + .78e^{12t}$.

b. $M_X'(t) = .01 \cdot 9e^{9t} + .05 \cdot 10e^{10t} + .16 \cdot 11e^{11t} + .78 \cdot 12e^{12t} = .09e^{9t} + .50e^{10t} + 1.76e^{11t} + 9.36e^{12t} \Rightarrow$ by Theorem (2.21), $E(X) = M_X'(0) = .09 + .50 + 1.76 + 9.36 = 11.71$.
Next, $M_X''(t) = .09 \cdot 9e^{9t} + .50 \cdot 10e^{10t} + 1.76 \cdot 11e^{11t} + 9.36 \cdot 12e^{12t} =$
$M_X''(t) = .81e^{9t} + 5e^{10t} + 19.36e^{11t} + 112.32e^{12t} \Rightarrow E(X^2) = M_X''(0) = .81 + 5 + 19.36 + 112.32 = 137.49$.
Hence, by the variance shortcut formula, $V(X) = E(X^2) - [E(X)]^2 = 137.49 - (11.71)^2 = 0.3659$, which gives $SD(X) = \sqrt{0.3659} = 0.605$.

51. By the uniqueness property of mgfs, we easily identify this as the discrete distribution with $p(0) = .2$, $p(1) = .3$, and $p(3) = .5$. From the mgf or the pmf, we find $E(X) = 1.8$ and $E(X^2) = 4.8$, from which $V(X) = 1.56$.

53. All mgfs must abide the property $M_X(0) = E[e^{0X}] = E[1] = 1$, but $g(0) = 0$.

55. The point of this exercise is that the natural log of the mgf is often a faster way to obtain the mean and variance of a rv, especially if the mgf has the form e^{function}.

a. $M_X(t) = e^{5t+2t^2} \Rightarrow M'_X(t) = (4t+5)e^{5t+2t^2} \Rightarrow E(X) = M'_X(0) = 5$;

$M''_X(t) = (4)e^{5t+2t^2} + (4t+5)^2 e^{5t+2t^2} \Rightarrow E(X^2) = M''_X(0) = 4 + 5^2 = 29 \Rightarrow V(X) = 29 - 5^2 = 4$.

b. $R_X(t) = \ln[M_X(t)] = 2t^2 + 5t$. From Exercise 54,
$R'_X(t) = 4t + 5 \Rightarrow E(X) = R'_X(0) = 5$ and $R''_X(t) = 4 \Rightarrow V(X) = R''_X(0) = 4$.

57.

$$M_X(t) = \sum e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx}(1-p)^{x-1} p = pe^t \sum_{x=1}^{\infty} (qe^t)^{x-1} \quad \text{with } q = 1-p$$

$$= pe^t[1 + qe^t + (qe^t)^2 + \cdots] = pe^t\left[\frac{1}{1-qe^t}\right] \quad \text{provided } |qe^t| < 1$$

$$= \frac{pe^t}{1-(1-p)e^t}$$

Compare the two mgfs: they are an exact match, with $p = .75$. Therefore, by the uniqueness property of mgfs, the pmf of Y must be $p(y) = (.25)^{y-1}(.75)$ for $y = 1, 2, 3, \ldots$.

59. Use the last proposition with $a = 1/2$ and $b = -5/2$: $M_Y(t) = e^{-5/2t} M_X(t/2) = e^{-5/2t} e^{5(t/2)+2(t/2)^2} = e^{t^2/2}$.
From this, $M'_Y(t) = te^{t^2/2} \Rightarrow E(Y) = M'_Y(0) = 0$ and $M''_Y(t) = 1 \cdot e^{t^2/2} + t \cdot te^{t^2/2} \Rightarrow E(Y^2) = M''_Y(0) = 1 \Rightarrow V(Y) = 1 - 0^2 = 1$.

61. If $Y = 10 - X$, $E(Y) = E(10 - X) = 10 - E(X)$ and $\text{SD}(Y) = \text{SD}(10 - X) = |-1|\text{SD}(X) = \text{SD}(X)$. Also, $Y - \mu_Y = (10 - X) - (10 - \mu_X) = -(X - \mu_X)$. Thus, the skewness coefficient of Y is
$E[(Y - \mu_Y)^3]/\sigma_Y^3 = E[(-(X - \mu_X))^3]/\sigma_X^3 = E[(-1)^3(X - \mu_X)^3]/\sigma_X^3 = -E[(X - \mu_X)^3]/\sigma_X^3 = -c$.

Section 3.5

63.

a. $b(3;8,.6) = \binom{8}{3}(.6)^3(.4)^5 = .124$.

b. $b(5;8,.6) = \binom{8}{5}(.6)^5(.4)^3 = .279$.

c. $P(3 \le X \le 5) = b(3;8,.6) + b(4;8,.6) + b(5;8,.6) = .635$.

d. $P(1 \le X) = 1 - P(X = 0) = 1 - b(0;12, .9) = 1 - \binom{12}{0}(.1)^0(.9)^{12} = 1 - (.9)^{12} = .718.$

65. $X \sim \text{Bin}(25, .05)$

a. $P(X \le 2) = B(2;25,.05) = .873.$

b. $P(X \ge 5) = 1 - P(X \le 4) = 1 - B(4;25,.05) = .1 - .993 = .007.$

c. $P(1 \le X \le 4) = P(X \le 4) - P(X \le 0) = .993 - .277 = .716.$

d. $P(X = 0) = P(X \le 0) = .277.$

e. $E(X) = np = (25)(.05) = 1.25$, $\text{SD}(X) = \sqrt{np(1-p)} = \sqrt{25(.05)(.95)} = 1.09.$

67. Let X be the number of drivers that come to a complete stop, so $X \sim \text{Bin}(20, .25)$.

a. $P(X \le 6) = B(6;20,.25) = .786.$

b. $P(X = 6) = b(6;20,.25) = .169.$

c. $P(X \ge 6) = 1 - P(X \le 5) = 1 - B(5;20,.25) = .382.$

69. Let "success" = has at least one citation and define X = number of individuals with at least one citation. Then $X \sim \text{Bin}(n = 15, p = .4)$.

a. If at least 10 have no citations (failure), then at most 5 have had at least one citation (success): $P(X \le 5) = B(5;15,.40) = .403.$

b. Half of 15 is 7.5, so less than half means 7 or fewer: $P(X \le 7) = B(7;15,.40) = .787.$

c. $P(5 \le X \le 10) = P(X \le 10) - P(X \le 4) = .991 - .217 = .774.$

71. Let "success" correspond to a telephone that is submitted for service while under warranty and must be replaced. Then $p = P(\text{success}) = P(\text{replaced} \mid \text{submitted})\cdot P(\text{submitted}) = (.40)(.20) = .08$. Thus X, the number among the company's 10 phones that must be replaced, has a binomial distribution with $n = 10$ and $p = .08$. Therefore, $P(X = 2) = \binom{10}{2}(.08)^2(.92)^8 = .1478$.

73. Let X = the number of flashlights that work, and let event B = {battery has acceptable voltage}.
Then $P(\text{flashlight works}) = P(\text{both batteries work}) = P(B)P(B) = (.9)(.9) = .81$. We have assumed here that the batteries' voltage levels are independent.
Finally, $X \sim \text{Bin}(10, .81)$, so $P(X \ge 9) = P(X = 9) + P(X = 10) = .285 + .122 = .407.$

75. Let X = the number of bits that are <u>not</u> switched during transmission ("successes"), so $X \sim \text{Bin}(3, .94)$.

a. A triplet is decoded <u>incorrectly</u> if 2 or 3 of the bits are switched during transmission — equivalently, if the number of preserved bits is less than half (one or none). So,

$P(\text{decoded incorrectly}) = P(X = 0 \text{ or } 1) = \binom{3}{0}(.94)^0(.06)^3 + \binom{3}{1}(.94)^1(.06)^2 = .0104.$

b. While this type of repeating system triples the number of bits required to transmit a message, it reduces the likelihood of a transmitted bit being wrongly decoded by a factor of about 6 (from 6% to ~1%).

c. Similar to part **a**, let $X \sim \text{Bin}(5, .94)$ model the number of correctly decoded bits in a 5-bit message. A 5-bit message is decoded <u>incorrectly</u> if less than half the bits are preserved, and "less than half" of 5 is 0, 1, or 2:

$P(X = 0, 1, \text{ or } 2) = \binom{5}{0}(.94)^0(.06)^5 + \binom{5}{1}(.94)^1(.06)^4 + \binom{5}{2}(.94)^2(.06)^3 = .00197.$

d. With no repetition, the expected number of bit errors is simply 25,000(.06) = 1500. According to part **a**, the probability drops to .0104 with the use of triplets, and the resulting expected number of bit errors falls to 25,000(.0104) = 260.

77. In this example, $X \sim \text{Bin}(25, p)$ with p unknown.

a. $P(\text{rejecting claim when } p = .8) = P(X \le 15 \text{ when } p = .8) = B(15; 25, .8) = .017.$

b. $P(\underline{\text{not}} \text{ rejecting claim when p} = .7) = P(X > 15 \text{ when } p = .7) = 1 - P(X \le 15 \text{ when } p = .7) =$
$= 1 - B(15; 25, .7) = 1 - .189 = .811.$
For $p = .6$, this probability is $= 1 - B(15; 25, .6) = 1 - .575 = .425.$

c. The probability of rejecting the claim when $p = .8$ becomes $B(14; 25, .8) = .006$, smaller than in **a** above. However, the probabilities of **b** above increase to .902 and .586, respectively. So, by changing 15 to 14, we're making it less likely that we will reject the claim when it's true (p really is $\ge .8$), but more likely that we'll "fail" to reject the claim when it's false (p really is $< .8$).

79. If topic A is chosen, then $n = 2$. When $n = 2$, $P(\text{at least half received}) = P(X \ge 1) = 1 - P(X = 0) =$

$1 - \binom{2}{0}(.9)^0(.1)^2 = .99.$

If topic B is chosen, then $n = 4$. When $n = 4$, $P(\text{at least half received}) = P(X \ge 2) = 1 - P(X \le 1) =$

$1 - \left[\binom{4}{0}(.9)^0(.1)^4 + \binom{4}{1}(.9)^1(.1)^3\right] = .9963.$

Thus topic B should be chosen if $p = .9$.
However, if $p = .5$, then the probabilities are .75 for A and .6875 for B (using the same method as above), so now A should be chosen.

81.

a. Although there are three payment methods, we are only concerned with S = uses a debit card and F = does not use a debit card. Thus we can use the binomial distribution. So, if X = the number of customers who use a debit card, $X \sim \text{Bin}(n = 100, p = .2)$. From this,
$E(X) = np = 100(.2) = 20$, and $V(X) = npq = 100(.2)(1–.2) = 16$.

b. With S = doesn't pay with cash, $n = 100$ and $p = .7$, so $\mu = np = 100(.7) = 70$ and $\sigma^2 = 21$.

83.

a. $np(1 - p) = 0$ if either $p = 0$ (whence every trial is a failure, so there is no variability in X) or if $p = 1$ (whence every trial is a success and again there is no variability in X).

b. $\frac{d}{dp}\left[np(1-p)\right] = n[(1)(1 - p) + p(-1)] = n[1 - 2p] = 0 \Rightarrow p = .5$, which is easily seen to correspond to a maximum value of $V(X)$.

85. When $n = 20$ and $p = .5$, $\mu = 10$ and $\sigma = 2.236$, so $2\sigma = 4.472$ and $3\sigma = 6.708$.
The inequality $|X - 10| \geq 4.472$ is satisfied if either $X \leq 5$ or $X \geq 15$, or
$P(|X - \mu| \geq 2\sigma) = P(X \leq 5 \text{ or } X \geq 15) = .021 + .021 = .042$. The inequality $|X - 10| \geq 6.708$ is satisfied if either $X \leq 3$ or $X \geq 17$, so $P(|X - \mu| \geq 3\sigma) = P(X \leq 3 \text{ or } X \geq 17) = .001 + .001 = .002$.

In the case $p = .75$, $\mu = 15$ and $\sigma = 1.937$, so $2\sigma = 3.874$ and $3\sigma = 5.811$.
$P(|X - 15| \geq 3.874) = P(X \leq 11 \text{ or } X \geq 19) = .041 + .024 = .065$, whereas
$P(|X - 15| \geq 5.811) = P(X \leq 9) = .004$.

All these probabilities are considerably less than the upper bounds given by Chebyshev:
for $k = 2$, Chebyshev's bound is $1/2^2 = .25$; for $k = 3$, the bound is $1/3^2 = .11$.

87. $R_X(t) = \ln[(pe^t + q)^n] = n\ln(pe^t + q) \Rightarrow R_X'(t) = n \cdot \frac{1}{pe^t + q} \cdot pe^t = \frac{npe^t}{pe^t + q} \Rightarrow E(X) = R_X'(0) = \frac{np}{p+q} = np$.

Next, $R_X''(t) = \frac{(pe^t + q)(npe^t) - (npe^t)(pe^t)}{(pe^t + q)^2} \Rightarrow$

$V(t) = R_X''(0) = \frac{(p+q)(np) - (np)(p)}{(p+q)^2} = \frac{1np - np^2}{1^2} = np - np^2 = np(1-p)$.

Section 3.6

89. All these solutions are found using the cumulative Poisson table, $P(x; \mu) = P(x; 5)$.

a. $P(X \leq 8) = P(8; 5) = .932$.

b. $P(X = 8) = P(8; 5) - P(7; 5) = .065$.

c. $P(X \geq 9) = 1 - P(X \leq 8) = .068$.

d. $P(5 \leq X \leq 8) = P(8; 5) - P(4; 5) = .492$.

e. $P(5 < X < 8) = P(7; 5) - P(5; 5) = .867 - .616 = .251$.

91. Let $X \sim \text{Poisson}(\mu = 20)$.

a. $P(X \le 10) = P(10; 20) = .011$.

b. $P(X > 20) = 1 - P(20; 20) = 1 - .559 = .441$.

c. $P(10 \le X \le 20) = P(20; 20) - P(9; 20) = .559 - .005 = .554$;
$P(10 < X < 20) = P(19; 20) - P(10; 20) = .470 - .011 = .459$.

d. $E(X) = \mu = 20$, so $\sigma = \sqrt{20} = 4.472$. Therefore, $P(\mu - 2\sigma < X < \mu + 2\sigma) =$
$P(20 - 8.944 < X < 20 + 8.944) = P(11.056 < X < 28.944) = P(X \le 28) - P(X \le 11) =$
$P(28; 20) - P(11; 20) = .966 - .021 = .945$.

93. The exact distribution of X = "number of implants that fail out of 200" is binomial with $n = 200$ and $p = .017$; we can approximate this distribution by the Poisson distribution with $\mu = np = 3.4$.

a. $P(X = 3) = \dfrac{e^{-3.4}3.4^3}{3!} = .219$.

b. $P(X \le 3) = \displaystyle\sum_{x=0}^{3} \frac{e^{-3.4}3.4^x}{x!} = .558$.

95. Let X = the number of pages with typos. The exact distribution of X is Bin(400, .005), which we can approximate by a Poisson distribution with $\mu = np = 400(.005) = 2$. Based on this model,
$P(X = 1) = \dfrac{e^{-2}2^1}{1!} = .271$ and $P(X \le 3) = P(3; 2) = .857$.

97.

a. $\mu = 8$ when $t = 1$, so $P(X = 6) = \dfrac{e^{-8}8^6}{6!} = .122$; $P(X \ge 6) = 1 - P(5; 8) = .809$; and
$P(X \ge 10) = 1 - P(9; 8) = .283$.

b. $t = 90$ min $= 1.5$ hours, so $\mu = 12$; thus the expected number of arrivals is 12 and the standard deviation is $\sigma = \sqrt{12} = 3.464$.

c. $t = 2.5$ hours implies that $\mu = 20$. So, $P(X \ge 20) = 1 - P(19; 20) = .530$ and
$P(X \le 10) = P(10; 20) = .011$.

99.

a. For a two-hour period the parameter of the distribution is $\mu = (4)(2) = 8$, so $P(X = 10) = \dfrac{e^{-8}8^{10}}{10!} = .099$.

b. For a 30-minute period, $\mu = (4)(.5) = 2$, so $P(X = 0) = \dfrac{e^{-2}2^0}{2!} = .135$.

c. $E(X) = (4)(.5) = 2$ calls.

101. $1/\lambda = .5$, so $\lambda = 2$.

a. $\mu = \lambda t = (2)(2) = 4$.

b. $P(X > 5) = 1 - P(X \le 5) = 1 - .785 = .215$.

c. Solve for t, given $\lambda = 2$: $.1 = P(X = 0) = e^{-2t} \Rightarrow t = -\ln(.1)/2 = 1.15$ years.

103.

a. For a quarter-acre (.25 acre) plot, the mean parameter is $\mu = (80)(.25) = 20$, so $P(X \le 16) = P(16; 20) = .221$.

b. The expected number of trees is $\lambda\cdot$(area) = 80 trees/acre (85,000 acres) = 6,800,000 trees.

c. The area of the circle is $\pi r^2 = \pi(.1)^2 = .01\pi = .031416$ square miles, which is equivalent to .031416(640) = 20.106 acres. Thus X has a Poisson distribution with parameter $\mu = \lambda(20.106) = 80(20.106) = 1608.5$. That is, the pmf of X is the function $p(x; 1608.5)$.

105.

a. With probability 1, no events occur in the interval $(0, t + \Delta t)$ if and only if no events occur in either of the two intervals $(0, t)$ and $(t, t + \Delta t)$. The assumptions of independence and Poisson distributions then imply $P_0(t + \Delta t) = P_0(t) \cdot P(\text{no events in } (t, t + \Delta t)) = P_0(t) \cdot [1 - \lambda \Delta t + o(\Delta t)]$.

b. Rearranging part **a**, $\dfrac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) - P_0(t) \cdot \dfrac{o(\Delta t)}{\Delta t}$. The limit of the left-hand side as $\Delta t \to 0$ is the derivative of $P_0(t)$ with respect to t. On the right-hand side, by definition $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$. Therefore, $P_0'(t) = -\lambda P_0(t)$.

c. $P_0'(t) = \dfrac{d}{dt}(e^{-\lambda t}) = -\lambda e^{-\lambda t} = -\lambda P_0(t)$, as desired.

d.

$$\frac{d}{dt}\left(\frac{e^{-\lambda t}(\lambda t)^k}{k!}\right) = \frac{\lambda^k}{k!}\frac{d}{dt}(t^k e^{-\lambda t}) = \frac{\lambda^k}{k!}\left[kt^{k-1}e^{-\lambda t} - \lambda t^k e^{-\lambda t}\right] = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!} - \frac{\lambda^{k+1} t^k e^{-\lambda t}}{k!}$$
$$= \lambda \frac{e^{-\lambda t}(\lambda t)^{k-1}}{(k-1)!} - \lambda \frac{e^{-\lambda t}(\lambda t)^k}{k!} = \lambda P_{k-1}(t) - \lambda P_k(t), \text{ as desired.}$$

107. The binomial mgf can be re-written as $(pe^t + (1-p))^n = (1 + p(e^t - 1))^n = \left(1 + \dfrac{np(e^t - 1)}{n}\right)^n = \left(1 + \dfrac{a_n}{n}\right)^n$, where $a_n = np(e^t - 1)$. By assumption, $np \to \mu$, so $a_n \to \mu(e^t - 1)$. Applying the calculus theorem, as $n \to \infty$ $\left(1 + \dfrac{a_n}{n}\right)^n \to e^{\lim a_n} = e^{\mu(e^t - 1)}$, which is indeed the Poisson mgf.

Section 3.7

109. According to the problem description, X is hypergeometric with $n = 6$, $N = 12$, and $M = 7$.

a. $P(X = 5) = \dfrac{\binom{7}{5}\binom{5}{1}}{\binom{12}{6}} = \dfrac{105}{924} = .114$.

b. $P(X \le 4) = 1 - P(X > 4) = 1 - [P(X = 5) + P(X = 6)] = 1 - \left[\dfrac{\binom{7}{5}\binom{5}{1}}{\binom{12}{6}} + \dfrac{\binom{7}{6}\binom{5}{0}}{\binom{12}{6}}\right] =$

$1 - [.114 + .007] = 1 - .121 = .879$.

c. $E(X) = n \cdot \dfrac{M}{N} = 6 \cdot \dfrac{7}{12} = 3.5$; $V(X) = \left(\dfrac{12-6}{12-1}\right)6\left(\dfrac{7}{12}\right)\left(1-\dfrac{7}{12}\right) = 0.795$; $\sigma = 0.892$. So,
$P(X > \mu + \sigma) = P(X > 3.5 + 0.892) = P(X > 4.392) = P(X = 5 \text{ or } 6) = .121$ (from part **b**).

d. We can approximate the hypergeometric distribution with the binomial if the population size and the number of successes are large. Here, $n = 15$ and $M/N = 40/400 = .1$, so
$h(x;15, 40, 400) \approx b(x;15, .10)$. Using this approximation, $P(X \le 5) \approx B(5; 15, .10) = .998$ from the binomial tables. (This agrees with the exact answer to 3 decimal places.)

111.

a. Possible values of X are 5, 6, 7, 8, 9, 10. (In order to have less than 5 of the granite, there would have to be more than 10 of the basaltic). X is hypergeometric, with $n = 15$, $N = 20$, and $M = 10$. So, the pmf of X is

$$p(x) = h(x; 15, 10, 20) = \frac{\binom{10}{x}\binom{10}{15-x}}{\binom{20}{15}}.$$

The pmf is also provided in table form below.

x	5	6	7	8	9	10
$p(x)$	.0163	.1354	.3483	.3483	.1354	.0163

b. P(all 10 of one kind or the other) $= P(X = 5) + P(X = 10) = .0163 + .0163 = .0326$.

c. $\mu = n \cdot \dfrac{M}{N} = 15 \cdot \dfrac{10}{20} = 7.5$; $V(X) = \left(\dfrac{20-15}{20-1}\right)15\left(\dfrac{10}{20}\right)\left(1-\dfrac{10}{20}\right) = .9868$; $\sigma = .9934$.

$\mu \pm \sigma = 7.5 \pm .9934 = (6.5066, 8.4934)$, so we want $P(6.5066 < X < 8.4934)$. That equals
$P(X = 7) + P(X = 8) = .3483 + .3483 = .6966$.

113.

a. The successes here are the top $M = 10$ pairs, and a sample of $n = 10$ pairs is drawn from among the $N = 20$. The probability is therefore $h(x; 10, 10, 20) = \dfrac{\binom{10}{x}\binom{10}{10-x}}{\binom{20}{10}}$.

b. Let X = the number among the top 5 who play east-west. (Now, $M = 5$.)
Then P(all of top 5 play the same direction) = $P(X = 5) + P(X = 0) = h(5; 10, 5, 20) + h(5; 10, 5, 20) =$

$$\frac{\binom{5}{5}\binom{15}{5}}{\binom{20}{10}} + \frac{\binom{5}{0}\binom{15}{10}}{\binom{20}{10}} = .033 .$$

c. Generalizing from earlier parts, we now have $N = 2n$; $M = n$. The probability distribution of X is hypergeometric: $p(x) = h(x; n, n, 2n) = \dfrac{\binom{n}{x}\binom{n}{n-x}}{\binom{2n}{n}}$ for $x = 0, 1, \ldots, n$. Also,

$E(X) = n \cdot \dfrac{n}{2n} = \dfrac{1}{2}n$ and $V(X) = \left(\dfrac{2n-n}{2n-1}\right) \cdot n \cdot \dfrac{n}{2n} \cdot \left(1 - \dfrac{n}{2n}\right) = \dfrac{n^2}{4(2n-1)}$.

115.

a. Let Y = the number of children the couple has until they have two girls. Then Y follows a negative binomial distribution with $r = 2$ (two girls/successes) and $p = .5$ (note: p here denotes P(girl)). The probability the family has exactly x <u>male</u> children is the probability they have $x + 2$ total children:
$P(Y = x + 2) = nb(x + 2; 2, .5) = \binom{x+2-1}{2-1}(.5)^2(1-.5)^{(x+2)-2} = (x + 1)(.5)^{x+2}$.

b. $P(Y = 4) = nb(4; 2, .5) = \binom{4-1}{2-1}(.5)^2(1-.5)^{4-2} = .1875$.

c. $P(Y \le 4) = \sum_{y=2}^{4} nb(y; 2, .5) = .25 + .25 + .1875 = .6875$.

d. The expected number of children is $E(Y) = \dfrac{r}{p} = \dfrac{2}{.5} = 4$. Since exactly two of these children are girls, the expected number of male children is $E(Y - 2) = E(Y) - 2 = 4 - 2 = 2$.

117. From Exercise 115**d**, the expected number of male children in each family is 2. Hence, the expected total number of male children across the three families is just 2 + 2 + 2 = 6, the sum of the expected number of males born to each family.

119. Let X = the number of ICs, among the 4 randomly selected ICs, that are defective. Then X has a hypergeometric distribution with n = 4 (sample size), M = 5 (# of "successes"/defectives in the population), and N = 20 (population size).

a. $P(X = 0) = h(0; 4, 5, 20) = \dfrac{\binom{5}{0}\binom{15}{4}}{\binom{20}{4}} = .2817.$

b. $P(X \le 1) = P(X = 0, 1) = .2817 + \dfrac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}} = .7513.$

c. Logically, if there are <u>fewer</u> defective ICs in the shipment of 20, then we ought to be <u>more</u> likely to accept the shipment. Replace M = 5 with M = 3 in the above calculations, and you will find that $P(X = 0) = .4912$ and $P(X \le 1) = .9123$.

121. Let X = the number of students the kinesiology professor needs to ask until she gets her 40 volunteers. With the specified assumptions, X has a negative binomial distribution with r = 40 and p = .25.

a. $E(X) = \dfrac{r}{p} = \dfrac{40}{.25} = 160$ students, and SD(X) = $\sqrt{\dfrac{r(1-p)}{p^2}} = \sqrt{\dfrac{40(.75)}{(.25)^2}} = \sqrt{480} = 21.9$ students.

b. $P(160 - 21.9 < X < 160 + 21.9) = P(138.1 < X < 181.9) = P(139 \le X \le 181) =$
$\sum_{x=139}^{181} \binom{x-1}{40-1} (.25)^{40}(.75)^{x-40} = .6756$ using software.

123. By definition, $M_X(t) = \sum_{x \in D} e^{tx} p(x) = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r}$. To match up with the provided formula, substitute $k = x - r$, which is equivalent to $x = r + k$:

$$M_X(t) = \sum_{k=r-r}^{\infty} e^{t(r+k)} \binom{r+k-1}{r-1} p^r (1-p)^k = e^{rt} p^r \sum_{k=0}^{\infty} e^{tk} \binom{r+k-1}{r-1} (1-p)^k = (pe^t)^r \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} [(1-p)e^t]^k$$

Now apply Newton's formula:

$$M_X(t) = (pe^t)^r (1-(1-p)e^t)^{-r} = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r.$$

For the derivatives, let $g(t) = \dfrac{pe^t}{1-(1-p)e^t}$, so $M_X(t) = [g(t)]^r$. Applying the quotient rule,

$$g'(t) = \frac{[1-(1-p)e^t]pe^t + pe^t(1-p)e^t}{(1-(1-p)e^t)^2} = \frac{pe^t}{(1-(1-p)e^t)^2} \text{ and } g''(t) = \frac{pe^t(1+(1-p)e^t)}{(1-(1-p)e^t)^3}. \text{ So,}$$

$$M'_X(t) = r[g(t)]^{r-1} g'(t) \Rightarrow E(X) = M'_X(0) = r[g(0)]^{r-1} g'(0) = r[1]^{r-1} \frac{p}{p^2} = \frac{r}{p} \text{ and}$$

$$M''_X(t) = r(r-1)[g(t)]^{r-2}[g'(t)]^2 + r[g(t)]^{r-1} g''(t) \Rightarrow$$

$$E(X^2) = M''_X(0) = r(r-1)[g(0)]^{r-2}[g'(0)]^2 + r[g(0)]^{r-1} g''(0) = r(r-1)[1]^{r-2}\left[\frac{1}{p}\right]^2 + r[1]^{r-1}\frac{2-p}{p^2}$$

$$= \frac{r^2 + r(1-p)}{p^2} \Rightarrow$$

$$V(X) = \frac{r^2 + r(1-p)}{p^2} - \left(\frac{r}{p}\right)^2 = \frac{r(1-p)}{p^2}$$

Section 3.8

125. Using the built-in commands of R, our "program" can actually be extremely short. The program below simulates 10,000 values from the pmf in Exercise 29.

```
Y=sample(c(0,1,2,3),10000,TRUE,c(.60,.25,.10,.05))
mean(Y)
sd(Y)
```

One execution of the program returned $\hat{\mu} = \bar{x} = 0.5968$ and $\hat{\sigma} = s = 0.8548$. From Exercise 29, the actual mean and sd are $\mu = 0.60$ and $\sigma = \sqrt{0.74} \approx 0.86$, so these estimated values are pretty close.

127. There are numerous ways to approach this problem. In our program below, we count the number of times the event $A = \{X \le 25\}$ occurs in 10,000 simulations. Five initial tests are administered, with a positive test (denoted 1) occurring with probability .75 and a negative test (denoted 0) with probability .25. The while statement looks at the most recent five tests and determines if the sum of the indicators (1's and 0's) is 5, since that would signify five consecutive positive test results. If not, then another test is administered, and the result of that test is appended to the previous results. At the end, a vector `X` of 0's and 1's terminating in 11111 exits the while loop; if its length is ≤ 25, then the number of tests was ≤ 25, and the count for the event A is increased by 1.

```
A<-0
for(i in 1:10000){
     k<-1
     X<-sample(c(0,1),5,
               TRUE,c(.25,.75))
     while(sum(X[k:(k+4)])<5){
          k<-k+1
          test<-sample(c(0,1),1,
               TRUE,c(.25,.75))
          X<-c(X,test)
     }
     if(length(X)<=25){
          A<-A+1
     }
}
```

An execution of the program gave A=9090, so $\hat{P}(X \le 25) = \hat{P}(A) = \frac{9090}{10000} = .9090$.

129.

a. Modify the programs from Exercise 101 of Chapter 2. (See Chapter 2 solutions for an explanation of the code.)

```
X<-NULL
for(i in 1:10000){
    u<-runif(10); v<-runif(13)
    w<-runif(13); x<-runif(5)
    right<-sum(u<1/2)+sum(v<1/3)
             +sum(w<1/4)+sum(x<1/5)
    X[i]<-right
}
```

The program returns a 10,000-by-1 vector of simulated values of X. Executing this program gave a vector with `mean(X)` = 13.5888 and `sd(X)` = 2.9381.

b. The command `mean(X>=mean(X)+sd(X))` computes one standard deviation above the mean, then determines which entries of `X` are at least one standard deviation above the mean (so the object in parentheses is a vector of 0's and 1's). Finally, taking the mean of that vector returns the *proportion* of entries that are 1's; using `sum(…)/10000` would do the same thing. R returned .1562, so the estimated probability is $\hat{P}(X \geq \mu+\sigma) \approx \hat{P}(X \geq \bar{x}+s)$ = .1562.

131. The cdf of Benford's law actually has a simple form:
$F(1) = p(1) = \log_{10}(2)$; $F(2) = p(1) + p(2) = \log_{10}\left(\frac{2}{1}\right)+\log_{10}\left(\frac{3}{2}\right)=\log_{10}\left(\frac{2}{1}\cdot\frac{3}{2}\right)=\log_{10}(3)$; and using the same property of logs, $F(x) = \log_{10}(x + 1)$ for x = 1, 2, 3, 4, 5, 6, 7, 8, 9. Using the inverse cdf method, we should assign a value x so long as $F_{x-1} \leq u < F_x$; i.e., $\log_{10}(x) \leq u < \log_{10}(x + 1)$. But that's equivalent to the inequalities $x \leq 10^u < x + 1$, which implies that x is the greatest integer less than or equal to 10^u.
Thus, we may implement the inverse cdf method here not with a series of if-else statements, but with a simple call to the "floor" function.

```
u<-runif(10000)
X<-floor(10^u)
```

For one execution of this (very simple) program in R, the sample mean and variance were `mean(X)` = 3.4152 and `var(X)` = 5.97.

133.

a. The function below returns a 10000-by-1 vector containing simulated values of the profit the airline receives, based on the number of tickets sold and the loss for over-booking.

```
airline<-function(t){
Profit<-NULL
for(i in 1:10000){
    X<-rbinom(1,t,.85)
    B<-max(X-120,0)
    Profit[i]<-250*t-500*B
}
return(Profit)
}
```

b. Execute the program above for input t = 140, 141, … 150, and determine the sample mean profit for each simulation. R returned the following estimates of expected profit:

t	140	141	142	143	…	150
Avg. Profit	\$34,406	\$34,458	\$34,468	\$34,431	…	\$33,701

Average profit increased monotonically from t = 140 to t = 142 and decreased monotonically from t = 142 to t = 150 (not shown). Thus, the airline should sell 142 tickets in order to maximize their expected profit.

135. The program below first simulates the number of chips a player earns, stored as `chips`; then, `chips` is passed as an argument in the second simulation to determine how many values from the "Plinko distribution" from Exercise 40 should be generated. The output vector `W` contains 10,000 simulations of the contestant's total winnings.

```
plinkowin<-function(){
W<-NULL
ch<-c(1,2,3,4,5)
pc<-c(.03,.15,.35,.34,.13)
x<-c(0,100,500,1000,10000)
px<-c(.39,.03,.11,.24,.23)
for(i in 1:10000){
    chips<-sample(ch,1,TRUE,pc)
    wins<-sample(x,chips,TRUE,px)
    W[i]<-sum(wins)
}
return(W)
}
```

Using the preceding R program gave the following estimates.

a. $P(W > 11{,}000) \approx \hat{P}(W > 11{,}000)$ = `mean(W>11000)` = .2991.

b. $E(W) \approx \overline{w}$ = `mean(W)` = \$8,696.

c. SD(X) ≈ s = `sd(X)` = \$7,811.

d. Delete the lines creating the pmf of C, and replace the line of code creating the count chips as follows:
`chips<-1+rbinom(1,4,.5)`
Running the new program gave the following estimates: $P(W > 11{,}000) \approx .2342$, $E(X) \approx \$7{,}767$, and SD($X$) ≈ \$7,571.

137.

a. The program below generates a random sample of size $n = 150$ from the ticket pmf described in the problem. Those 150 requests are stored in `tickets`, and then the total number of requested tickets is stored in `T`.

```
numboxes<-NULL
for(i in 1:10000){
    coupon<-rep(0,10)
    while(all(coupon>0)==0){
        new<-sample(1:10,1)
        coupon[new]<-coupon[new]+1
    }
    numboxes[i]<-sum(coupon)
}
```

b. Using the program above, $P(T \le 410) \approx$ `mean(T<=410)` = .9196 for one run. The estimated standard error of that estimate is $\sqrt{\frac{(.9196)(1-.9196)}{10{,}000}}$ = .00272. Hence, a 95% confidence interval for the true probability (see Chapter 5) is .9196 ± 1.96(.00272) = (.9143, .9249).

Supplementary Exercises

139.

a. We'll find $p(1)$ and $p(4)$ first, since they're easiest, then $p(2)$. We can then find $p(3)$ by subtracting the others from 1.

$p(1) = P$(exactly one suit) = P(all ♠) + P(all ♥) + P(all ♦) + P(all ♣) =

4 · P(all ♠) = $4 \cdot \frac{\binom{13}{5}\binom{39}{0}}{\binom{52}{5}} = .00198$, since there are 13 ♠s and 39 other cards.

$p(4)$ = 4 · P(2 ♠, 1 ♥, 1 ♦, 1 ♣) = $4 \cdot \frac{\binom{13}{2}\binom{13}{1}\binom{13}{1}\binom{13}{1}}{\binom{52}{5}} = .26375$.

$p(2) = P$(all ♥s and ♠s, with ≥ one of each) + … + P(all ♦s and ♣s with ≥ one of each) =

$\binom{4}{2}$ · P(all ♥s and ♠s, with ≥ one of each) =

6 · [P(1 ♥ and 4 ♠) + P(2 ♥ and 3 ♠) + P(3 ♥ and 2 ♠) + P(4 ♥ and 1 ♠)] =

$$6 \cdot \left[2 \cdot \frac{\binom{13}{4}\binom{13}{1}}{\binom{52}{5}} + 2 \cdot \frac{\binom{13}{3}\binom{13}{2}}{\binom{52}{5}} \right] = 6\left[\frac{18{,}590 + 44{,}616}{2{,}598{,}960}\right] = .14592.$$

Finally, $p(3) = 1 - [p(1) + p(2) + p(4)] = .58835$.

b. $\mu = \sum_{x=1}^{4} x \cdot p(x) = 3.114$; $\sigma^2 = \left[\sum_{x=1}^{4} x^2 \cdot p(x)\right] - (3.114)^2 = .405 \Rightarrow \sigma = .636$.

141.

a. From the description, $X \sim$ Bin(15, .75). So, the pmf of X is $b(x; 15, .75) = \binom{15}{x}(.75)^x(.25)^{15-x}$.

b. $P(X > 10) = 1 - P(X \le 10) = 1 - B(10;15, .75) = 1 - .314 = .686$.

c. $P(6 \le X \le 10) = B(10; 15, .75) - B(5; 15, .75) = .314 - .001 = .313$.

d. $\mu = (15)(.75) = 11.75$, $\sigma^2 = (15)(.75)(.25) = 2.81 \Rightarrow \sigma = 1.68$.

e. Requests can all be met if and only if $X \le 10$ <u>and</u> $15 - X \le 8$, i.e. iff $7 \le X \le 10$. So, $P(\text{all requests met}) = P(7 \le X \le 10) = B(10; 15, .75) - B(6; 15, .75) = .310$.

143.

a. Let X = the number of bits transmitted until the third error occurs. Then $X \sim$ NB($r = 3$, $p = .05$). Thus $P(X = 50) = nb(50; 3, .05) = \binom{50-1}{3-1}(.05)^3(.95)^{47} = .013$.

b. Using the mean of the geometric distribution, the average number of bits up to and including the first error is $1/p = 1/.05 = 20$ bits. Hence, the average number <u>before</u> the first error is $20 - 1 = 19$.

c. Now let X = the number of bit errors in a 32-bit word. Then $X \sim$ Bin($n = 32$, $p = .05$), and $P(X = 2) = b(2; 32, .05) = \binom{32}{2}(.05)^2(.95)^{30} = .266$.

d. The rv X has a Bin(10000, .05) distribution. Since n is large and p is small, this can be approximated by a Poisson distribution, with $\mu = 10000(.05) = 500$.

145.

a. $X \sim$ Bin($n = 500$, $p = .005$). Since n is large and p is small, X can be approximated by a Poisson distribution with $\mu = np = 2.5$. The approximate pmf of X is $p(x; 2.5) = \dfrac{e^{-2.5}2.5^x}{x!}$.

b. $P(X = 5) = \dfrac{e^{-2.5}2.5^5}{5!} = .0668$.

c. $P(X \ge 5) = 1 - P(X \le 4) = 1 - P(4; 2.5) = 1 - .8912 = .1088$.

147. Let Y denote the number of tests carried out.
For $n = 3$, possible Y values are 1 and 4. $P(Y = 1) = P(\text{no one has the disease}) = (.9)^3 = .729$ and $P(Y = 4) = 1 - .729 = .271$, so $E(Y) = (1)(.729) + (4)(.271) = 1.813$, as contrasted with the 3 tests necessary without group testing.
For $n = 5$, possible values of Y are 1 and 6. $P(Y = 1) = P(\text{no one has the disease}) = (.9)^5 = .5905$, so $P(Y = 6) = 1 - .5905 = .4095$ and $E(Y) = (1)(.5905) + (6)(.4095) = 3.0475$, less than the 5 tests necessary without group testing.

149. $p(2) = P(X = 2) = P(SS) = p^2$, and $p(3) = P(FSS) = (1 - p)p^2$.

For $x \geq 4$, consider the first $x - 3$ trials and the last 3 trials separately. To have $X = x$, it must be the case that the *last* three trials were FSS, and that two-successes-in-a-row was not already seen in the first $x - 3$ tries.

The probability of the first event is simply $(1 - p)p^2$.
The second event occurs if two-in-a-row hadn't occurred after 2 or 3 or … or $x - 3$ tries. The probability of this second event equals $1 - [p(2) + p(3) + \ldots + p(x - 3)]$. (For $x = 4$, the probability in brackets is empty; for $x = 5$, it's $p(2)$; for $x = 6$, it's $p(2) + p(3)$; and so on.)

Finally, since trials are independent, $P(X = x) = (1 - [p(2) + \ldots + p(x - 3)]) \cdot (1 - p)p^2$.

For $p = .9$, the pmf of X up to $x = 8$ is shown below.

x	2	3	4	5	6	7	8
$p(x)$	.81	.081	.081	.0154	.0088	.0023	.0010

So, $P(X \leq 8) = p(2) + \ldots + p(8) = .9995$.

151.

a. Let event A = seed carries single spikelets, and event B = seed produces ears with single spikelets. Then $P(A \cap B) = P(A) \cdot P(B \mid A) = (.40)(.29) = .116$.
Next, let X = the number of seeds out of the 10 selected that meet the condition $A \cap B$. Then $X \sim \text{Bin}(10, .116)$. So, $P(X = 5) = \binom{10}{5}(.116)^5(.884)^5 = .002857$.

b. For any one seed, the event of interest is B = seed produces ears with single spikelets. Using the law of total probability, $P(B) = P(A \cap B) + P(A' \cap B) = (.40)(.29) + (.60)(.26) = .272$.
Next, let Y = the number out of the 10 seeds that meet condition B. Then $Y \sim \text{Bin}(10, .272)$.
$P(Y = 5) = \binom{10}{5}(.272)^5(1 - .272)^5 = .0767$, while
$P(Y \leq 5) = \sum_{y=0}^{5}\binom{10}{y}(.272)^y(1 - .272)^{10-y} = .041813 + \ldots + .076719 = .97024$.

153.

a. Using the Poisson model with $\mu = 2(1) = 2$, $P(X = 0) = p(0; 2)$ or $\frac{e^{-2}2^0}{0!} = .135$.

b. Let S = an operator receives no requests. Then the number of operators that receive no requests follows a Bin($n = 5$, $p = .135$) distribution. So, P(exactly 4 S's in 5 trials) = $b(4; 5, .135)$ = $\binom{5}{4}(.135)^4(.865)^1 = .00144$.

c. For any non-negative integer x, P(all operators receive exactly x requests) = P(first operator receives x) · … · P(fifth operator receives x) = $[p(x; 2)]^5 = \left[\frac{e^{-2}2^x}{x!}\right]^5 = \frac{e^{-10}2^{5x}}{(x!)^5}$.
Then, P(all receive the same number) = P(all receive 0 requests) + P(all receive 1 request) + P(all receive 2 requests) + … = $\sum_{x=0}^{\infty}\frac{e^{-10}2^{5x}}{(x!)^5}$.

155. The number of magazine copies sold is X so long as X is no more than five; otherwise, all five copies are sold. So, mathematically, the number sold is min(X, 5), and

$E[\min(x, 5)] = \sum_{x=0}^{\infty}\min(x,5)p(x;4) = 0p(0; 4) + 1p(1; 4) + 2p(2; 4) + 3p(3; 4) + 4p(4; 4) + \sum_{x=5}^{\infty}5p(x;4) =$

$1.735 + 5\sum_{x=5}^{\infty}p(x;4) = 1.735 + 5\left[1-\sum_{x=0}^{4}p(x;4)\right] = 1.735 + 5[1 - P(4; 4)] = 3.59$.

157.

a. No, since the probability of a "success" is not the same for all tests.

b. There are four ways exactly three could have positive results. Let D represent those with the disease and D' represent those without the disease. Adding up the probabilities associated with the four combinations yields 0.0273.

Combination		**Probability**
D	D'	
0	3	$\left[\binom{5}{0}(.2)^0(.8)^5\right]\cdot\left[\binom{5}{3}(.9)^3(.1)^2\right]$ =(.32768)(.0729) = .02389
1	2	$\left[\binom{5}{1}(.2)^1(.8)^4\right]\cdot\left[\binom{5}{2}(.9)^2(.1)^3\right]$ =(.4096)(.0081) = .00332
2	1	$\left[\binom{5}{2}(.2)^2(.8)^3\right]\cdot\left[\binom{5}{1}(.9)^1(.1)^4\right]$ =(.2048)(.00045) = .00009216
3	0	$\left[\binom{5}{3}(.2)^3(.8)^2\right]\cdot\left[\binom{5}{0}(.9)^0(.1)^5\right]$ =(.0512)(.00001) = .000000512

159.

a. Notice that $p(x; \mu_1, \mu_2) = .5\, p(x; \mu_1) + .5\, p(x; \mu_2)$, where both terms $p(x; \mu_i)$ are Poisson pmfs. Since both pmfs are ≥ 0, so is $p(x; \mu_1, \mu_2)$. That verifies the first requirement.

Next, $\sum_{x=0}^{\infty} p(x;\mu_1,\mu_2) = .5\sum_{x=0}^{\infty} p(x;\mu_1) + .5\sum_{x=0}^{\infty} p(x;\mu_2) = .5 + .5 = 1$, so the second requirement for a pmf is met. Therefore, $p(x; \mu_1, \mu_2)$ is a valid pmf.

b. $E(X) = \sum_{x=0}^{\infty} x\cdot p(x;\mu_1,\mu_2) = \sum_{x=0}^{\infty} x[.5p(x;\mu_1) + .5p(x;\mu_2)] = .5\sum_{x=0}^{\infty} x\cdot p(x;\mu_1) + .5\sum_{x=0}^{\infty} x\cdot p(x;\mu_2) = .5E(X_1) +$ $.5E(X_2)$, where $X_i \sim$ Poisson(μ_i). Therefore, $E(X) = .5\mu_1 + .5\mu_2$.

c. This requires using the variance shortcut. Using the same method as in **b,**

$E(X^2) = .5\sum_{x=0}^{\infty} x^2\cdot p(x;\mu_1) + .5\sum_{x=0}^{\infty} x^2\cdot p(x;\mu_2) = .5E(X_1^2) + .5E(X_2^2)$. For any Poisson rv,

$E(X^2) = V(X) + [E(X)]^2 = \mu + \mu^2$, so $E(X^2) = .5(\mu_1 + \mu_1^2) + .5(\mu_2 + \mu_2^2)$.

Finally, $V(X) = .5(\mu_1 + \mu_1^2) + .5(\mu_2 + \mu_2^2) - [.5\mu_1 + .5\mu_2]^2$, which can be simplified to equal $.5\mu_1 + .5\mu_2 + .25(\mu_1 - \mu_2)^2$.

d. Simply replace the weights .5 and .5 with .6 and .4, so $p(x; \mu_1, \mu_2) = .6\, p(x; \mu_1) + .4\, p(x; \mu_2)$.

161. Since $X \sim$ Poisson(μ) for some unknown μ, $P(X = 1) = \dfrac{e^{-\mu}\mu^1}{1!} = \mu e^{-\mu}$ and $P(X = 2) = \dfrac{e^{-\mu}\mu^2}{2!} = .5\mu^2 e^{-\mu}$.

Thus, $P(X = 1) = 4P(X = 2)$ implies $\mu e^{-\mu} = 4\cdot .5\mu^2 e^{-\mu} \Rightarrow 1 = 2\mu \Rightarrow \mu = .5$.

163. Start with the hint: $V(X) = \sum_{\text{all } x}(x-\mu)^2 p(x) \geq \sum_{x:\,|x-\mu|\geq k\sigma}(x-\mu)^2 p(x)$.

Next, for any x value satisfying $|x-\mu| \geq k\sigma$, $(x-\mu)^2 \geq (k\sigma)^2$, so we have

$V(X) \geq \sum_{x:\,|x-\mu|\geq k\sigma}(k\sigma)^2 p(x) = k^2\sigma^2 \sum_{x:\,|x-\mu|\geq k\sigma} p(x) \Rightarrow 1 \geq k^2 \sum_{x:\,|x-\mu|\geq k\sigma} p(x)$.

Third, for any set A, $\sum_A p(x) = P(X \in A)$ by definition. Therefore,

$1 \geq k^2 \cdot P(|X-\mu| \geq k\sigma)$, or $P(|X-\mu| \geq k\sigma) \leq 1/k^2$, QED.

165. If choices are independent and have the same probability p, then $X \sim$ Bin(25, p), where p denotes the probability that customer wants the fancy coffee maker. Using the rescaling properties of mean and standard deviation,

$E(20X + 750) = 20E(X) + 750 = 20 \cdot 25p + 750 = 500p + 750$, and

$SD(20X + 750) = 20SD(X) = 20\cdot\sqrt{25p(1-p)} = 100\sqrt{p(1-p)}$.

Customers' choices might not be independent because, for example, the store may have a limited supply of one or both coffee makers, customers might affect each other's choices, the sales employees may influence customers, …

167.

$P(X = 0) = P(\text{none are late}) = (.6)^5 = .07776$; $P(X = 1) = P(\text{one single is late}) = 2(.4)(.6)^4 = .10368$.
$P(X = 2) = P(\text{both singles are late or one couple is late}) = (.4)^2(.6)^3 + 3(.4)(.6)^4 = .19008$.
$P(X = 3) = P(\text{one single and one couple is late}) = 2(.4)3(.4)(.6)^3 = .20736$. Continuing in this manner,
$P(X = 4) = .17280$,
$P(X = 5) = .13824$,
$P(X = 6) = .06912$,
$P(X = 7) = .03072$, and
$P(X = 8) = (.4)^5 = .01024$.

169.

a. The expected value of X is given by

$$\begin{aligned}\sum_{x=0}^{N}[x\cdot p(x)] &= 0p(0)+1p(1)+2p(2)+\cdots+Np(N)\\ &= p(1)+2p(2)+\cdots+Np(N)\\ &=[p(1)+p(2)+\cdots+p(N)]+[1p(2)+2p(3)+\cdots+(N-1)p(N)]\\ &=[p(1)+p(2)+\cdots+p(N)]+[p(2)+p(3)+\cdots+p(N)]\\ &\quad+[1p(3)+2p(4)+\cdots+(N-2)p(N)]\\ &=\cdots\end{aligned}$$

Continuing in this fashion, we ultimately have

$$\begin{aligned}\sum_{x=0}^{N}[x\cdot p(x)] = p(1)+p(2)+p(3)+\cdots+p(N)&\\ +p(2)+p(3)+\cdots+p(N)&\\ +p(3)+\cdots+p(N)&\\ \vdots&\\ +p(N)&\end{aligned}$$

b. The first row of the above expression is $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X \leq 0) = 1 - F(0)$. Similarly, the second row is $P(X \geq 2) = 1 - F(1)$; the third row is $1 - F(2)$; and finally the bottom row is $p(N) = P(X = N) = 1 - P(X \leq N - 1)$ because the range of X is 0 to $N = 1 - F(N - 1)$. Equating the original summation with the sum of these terms, we have

$$\sum_{x=0}^{N}[x\cdot p(x)] = [1-F(0)]+[1-F(1)]+\cdots+[1-F(N-1)] = \sum_{x=0}^{N-1}[1-F(x)].$$

c. The relationship in part **b** holds for any positive integer N. Hence, we may take the limit as $N \to \infty$ of both sides:

$$\sum_{x=0}^{N}[x\cdot p(x)] = \sum_{x=0}^{N-1}[1-F(x)] \text{ for all } N \Rightarrow \lim_{N\to\infty}\sum_{x=0}^{N}[x\cdot p(x)] = \lim_{N\to\infty}\sum_{x=0}^{N-1}[1-F(x)] \Rightarrow$$

$\sum_{x=0}^{\infty}[x\cdot p(x)] = \sum_{x=0}^{\infty}[1-F(x)]$. Both expressions represent the expected value of a discrete rv X whose range is $\{0, 1, 2, \ldots\}$.

This formula holds even if the range of X is finite: suppose the maximum value of X is N. Then $F(N) = P(X \leq N) = 1$, so $1 - F(N) = 0$; similarly, $F(x) = 1$ for $x > N$ and so $1 - F(x) = 0$ for all $x > N$. Thus, all the terms past $x = N - 1$ drop out of the right-hand side. Meanwhile, $p(x) = 0$ for all $x > N$ on the left-hand side, and so the equation with ∞ as the upper limit of each sum returns to the answer to part **b**.

d. The pmf of X is $p(x) = (1 - p)^{x-1}p$ for $x = 1, 2, 3, \ldots$. Using the formula for a geometric series, the cdf at any positive integer x is
$F(x) = p(1) + p(2) + \ldots + p(x) = (1 - p)^{1-1}p + (1 - p)^{2-1}p + (1 - p)^{x-1}p = p[1 + (1 - p) + \ldots + (1 - p)^{x-1}]$
$= p\dfrac{1-(1-p)^x}{1-(1-p)} = p\dfrac{1-(1-p)^x}{p} = 1-(1-p)^x$. This is even valid for $x = 0$: $F(0) = P(X \le 0) = 0$ and $1 - (1 - p)^x = 1 - (1 - p)^0 = 1 - 1 = 0$.

Now apply the alternative expression we have derived for the expected value. Again using the formula for a geometric series,

$$E(X) = \sum_{x=0}^{\infty}[1 - F(x)] = \sum_{x=0}^{\infty}[1-(1-(1-p)^x)] = \sum_{x=0}^{\infty}[(1-p)^x]$$
$$= 1 + (1-p) + (1-p)^2 + \cdots$$
$$= \frac{1}{1-(1-p)} = \frac{1}{p}$$