Chapter 5

Exercises for Section 5.2

1. Probabilities must lie between 0 and 1 and add to 1. The value 1.10 is clearly incorrect for a probability. If it is changed to 0.11 then the set of probabilities add to 1.

2. $\text{pr}(X = 1) = \frac{176}{200} = 0.88$, $\text{pr}(X = 2) = \frac{22}{200} = 0.11$, etc. Placing these into a table, we get the following.

x	1	2	3
pr(x)	0.88	0.11	0.01

3. We have, using independence, $\text{pr}(BBB) = \text{pr}(B) \times \text{pr}(B) \times \text{pr}(B) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$. Every other outcome has the same probability. (This is the same probability distribution as in Example 4.4.6(b).) $X = 0$ corresponds to one outcome (BBB) and thus has probability $\frac{1}{8}$. $X = 1$ corresponds to three outcomes (GBB, BGB, and BBG) and thus has probability $\frac{3}{8}$, and so on. We obtain the following.

x	0	1	2	3
pr(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

4. $\text{pr}(4 < X < 8) = \text{pr}(X = 5) + \text{pr}(X = 6) + \text{pr}(X = 7) = 0.844$ or $\text{pr}(4 < X < 8) = \text{pr}(X \leq 7) - \text{pr}(x \leq 4) = 0.882 - 0.038 = 0.844$

5. **(a)** The cumulative probabilities are given below.

x	3	4	5	6	7	8	9	10	11	12	13
$\text{pr}(X = x)$	.07	.01	.09	.01	.16	.25	.20	.03	.02	.11	.05
$\text{pr}(X \leq x)$	.07	.08	.17	.18	.34	.59	.79	.82	.84	.95	1.0

(b) $\text{pr}(X \leq 5) = \text{pr}(3) + \text{pr}(4) + \text{pr}(5) = .07 + .01 + .09 = 0.17$. We also read the value directly off the cumulative probabilities row.

(c) $\text{pr}(X > 9) = \text{pr}(10) + \text{pr}(11) + \text{pr}(12) + \text{pr}(13) = .03 + .02 + .11 + .05 = 0.21$, or, using the cumulative probabilities row, $\text{pr}(X > 9) = 1 - \text{pr}(x \leq 9) = 1 - 0.79 = 0.21$.

(d) $\text{pr}(X \geq 9) = \text{pr}(9) + \ldots + \text{pr}(13) = 0.41$, or, using the cumulative probabilities row, $\text{pr}(X \geq 9) = 1 - \text{pr}(x \leq 8) = 1 - 0.59 = 0.41$.

(e) $\text{pr}(X < 12) = \text{pr}(3) + \ldots + \text{pr}(11) = 0.84$, or using the cumulative probabilities table, $\text{pr}(X < 12) = \text{pr}(X \leq 11) = 0.84$.

(f) $\text{pr}(5 \leq X \leq 9) = \text{pr}(5) + \ldots + \text{pr}(9) = 0.71$, or using the cumulative probabilities table, $\text{pr}(5 \leq X \leq 9) = \text{pr}(X \leq 9) - \text{pr}(X \leq 4) = 0.79 - 0.08 = 0.41$.

(g) $\text{pr}(4 < X < 11) = \text{pr}(5) + \ldots + \text{pr}(10) = 0.74$, or using the cumulative probabilities table, $\text{pr}(4 < X < 11) = \text{pr}(X \leq 10) - \text{pr}(X \leq 4) = 0.82 - 0.08 = 0.74$.

(h) and **(i)** The probabilities for the values of x listed in the table add to 1 so these take up all the possible values. Everything outside this has probability 0. Thus $\text{pr}(X = 14) = 0$ and $\text{pr}(X < 3) = 0$.

6. Arguing exactly as in the Case Study, $0.9^3 = 0.729$. As $\text{pr}(X \le 6) = 0.469$ and $\text{pr}(X \le 7) = 0.522$, she should try seven times.

Exercises for Section 5.3

1. **(a)** 0.2668. **(b)** $1 - \text{pr}(X \le 3) = 0.3504$. **(c)** $1 - \text{pr}(X \le 6) = 0.01059$.
(d) $\text{pr}(X \le 6) = 0.9894$. **(e)** 0.1493. **(f)** $\text{pr}(X \le 7) - \text{pr}(X \le 3) = 0.3488$.
(g) $\text{pr}(X \le 7) - \text{pr}(X \le 3) = 0.3488$. **(h)** 0 as X only takes values $0, 1, \ldots, 10$.
(i) $\text{pr}(X \le 10) = 1$.

2. **(a)** Here the 20% refers to a conceptual population rather an actual one, so that the coin-tossing model is a candidate. The urn model does not apply, as the 10 chosen cars are not a simple random sample of all the cars parking. You need to have the probability of overstaying remaining constant at 0.2. Here $n = 10$ and $p = 0.2$. The arrival of each car is like a binomial trial. You can use Binomial$(n = 10, p = 0.2)$.

(b) Urn model. The 50 cars watched must be a simple random sample of the cars. Here $N = 10,000$; M, the total number of overstayers, is unknown; and $n = 50$. You can approximate by Binomial$(n = 50, p)$, where $p = \frac{M}{N}$, as $\frac{n}{N} < 0.1$. You may be able to use $p = 0.2$ from part (a).

(c) Urn model. The 50 people dialed must be a simple random sample of subscribers. Here $N = 7400$, $M = 2730$, and $n = 50$. You can approximate by Binomial$(n = 50, p = \frac{M}{N} = 0.3689)$ as $\frac{n}{N} < 0.1$.

(d) Urn model. Here N and M are large unknown numbers, $\frac{M}{N} = 0.45$, and $n = 100$. You can approximate by Binomial$(n = 100, p = 0.45)$ as you can expect $\frac{n}{N} < 0.1$.

(e) Urn model. Here $N = 100$, $M = 12$ and $n = 7$. You can approximate by Binomial$(n = 100, p = \frac{M}{N} = 0.12)$ as $\frac{n}{N} = \frac{7}{100} < 0.1$.

(f) Urn model. Here N and M are large unknown numbers, $\frac{M}{N} = 0.64$, and $n = 50$. You can approximate by Binomial$(n = 50, p = 0.64)$ as you can expect $\frac{n}{N} < 0.1$.

(g) Urn model. Here $N = 188$, $M = 99$, and $n = 30$. You cannot use the Binomial as $\frac{n}{N} > 0.1$.

(h) Coin-tossing model. Here $n = 10$ and $p = \frac{1}{6}$. Can use Binomial$(n = 10, p = \frac{1}{6})$

(i) Urn model. Here $N = 52$, $M = 4$, and $n = 7$. You need random shuffling of the pack before dealing. You cannot use the Binomial as $\frac{n}{N} > 0.1$.

(j) Urn model. Here N and M are large unknown numbers, $\frac{M}{N} = 0.1$ and $n = 30$. You can approximate by Binomial$(n = 30, p = 0.1)$ as you can expect $\frac{n}{N} < 0.1$.
Either model is a candidate. For an urn model, the 30% must refer to the actual population sampled with M and N unknown and $n = 20$. For the coin-tossing model, you must have p constant. Since you can expect $\frac{n}{N} < 0.1$, either model will lead to Binomial$(n = 20, p = 0.3)$.

(k) Either model is a candidate, both leading to Binomial$(n = 20, p = 0.3)$. For example, urn model applies if we think of the 20 bearings being sampled from the existing population of bearings of which 30% will function for a year of continuous use. If, however, we think in terms of the bearings being random items from a continuous manufacturing process producing items with the "30% will function" being stable (or constant) over time, the coin-tossing model applies.

(l) Like (a), either model is a candidate. The urn model, however, will apply only if the 50 patients can be regarded as a simple random sample of patients, which is probably not the case. The coin-tossing model is dubious as p may not be constant.

Exercises for Section 5.4.1

1. $\mathrm{E}(X) = \sum x\,\mathrm{pr}(x) = 2 \times 0.2 + 3 \times 0.1 + 5 \times 0.3 + 7 \times 0.4 = 5.0.$

2. $\mathrm{E}(X) = \sum x\,\mathrm{pr}(x) = 0 \times 0.49 + 1 \times 0.42 + 2 \times 0.09 = 0.6 \quad (= n\,p).$

Exercises for Section 5.4.2

1. $\mathrm{E}(X-\mu)^2\mathrm{pr}(x) = (2-5)^2 \times 0.2 + (3-5)^2 \times 0.1 + (5-5)^2 \times 0.3 + (7-5)^2 \times 0.4 = 3.8.$
$\mathrm{sd}(X) = \sqrt{3.8} = 1.9494.$

2. $\mathrm{E}(X-\mu)^2\mathrm{pr}(x) = (0-1.25)^2 \times \frac{1}{8} + (1-1.25)^2 \times \frac{5}{8} + (2-1.25)^2 \times \frac{1}{8} + (3-1.25)^2 \times \frac{1}{8}$
$= 0.6875, \qquad \mathrm{sd}(X) = \sqrt{0.6875} = 0.8292.$

3. $\mathrm{E}(X-\mu)^2\mathrm{pr}(x) = (0-0.6)^2 \times 0.49 + (1-0.6)^2 \times 0.42 + (2-0.6)^2 \times 0.09 = 0.42.$
$\mathrm{sd}(X) = \sqrt{0.42} = 0.6481.$

Exercises for Section 5.4.3

(a) $\mathrm{E}(2X) = 2\,\mathrm{E}(X) = 6, \quad \mathrm{sd}(2X) = 2\,\mathrm{sd}(X) = 4.$

(b) $\mathrm{E}(4+X) = 4 + \mathrm{E}(X) = 7, \quad \mathrm{sd}(4+X) = \mathrm{sd}(X) = 2.$

(c) As for (b), since $\mathrm{E}(X+4) = \mathrm{E}(4+X)$ and $\mathrm{sd}(X+4) = \mathrm{sd}(4+X)$.

(d) $\mathrm{E}(3X+2) = 3\,\mathrm{E}(X) + 2 = 11, \quad \mathrm{sd}(3X+2) = 3\,\mathrm{sd}(X) = 6.$

(e) $\mathrm{E}(4+5X) = 4 + 5\,\mathrm{E}(X) = 19, \quad \mathrm{sd}(4+5X) = 5\,\mathrm{sd}(X) = 10.$

(f) $\mathrm{E}(-5X) = -5\,\mathrm{E}(X) = -15, \quad \mathrm{sd}(-5X) = 5\,\mathrm{sd}(X) = 10.$

(g) $\mathrm{E}(-5X+4) = -5\,\mathrm{E}(X) + 4 = -11, \quad \mathrm{sd}(-5X+4) = 5\,\mathrm{sd}(X) = 10.$

(h) As for (g).

(i) $\mathrm{E}(-7X-9) = -7\,\mathrm{E}(X) - 9 = -30, \quad \mathrm{sd}(-7X-9) = 7\,\mathrm{sd}(X) = 14.$

Review Exercises 5

1. In the following "N/A" is used when neither model is applicable.

- **(a)** Coin-tossing model. $X_1 \sim \text{Binomial}(n = 20, p = 0.2)$.
- **(b)** Urn model with $N = 1000$, $M = 100$, and $n = 20$. You can use the Binomial approximation, $X_2 \sim \text{Binomial}(n = 20, p = 0.1)$ as $\frac{n}{N} < 0.1$.
- **(c)** N/A.
- **(d)** Coin-tossing model. $X_4 \sim \text{Binomial}(n = 120, p = 0.6)$.
- **(e)** Urn model with $N = 120$, $M = 70$ and $n = 10$. You can use the Binomial approximation, $X_5 \sim \text{Binomial}(n = 10, p = \frac{7}{12})$, as $\frac{n}{N} < 0.1$.
- **(f)** Urn model with $N = 20$, $M = 9$ and $n = 15$. You cannot use a Binomial approximation as $\frac{n}{N} > 0.1$.
- **(g)** Coin-tossing model. $X_7 \sim \text{Binomial}(n = 12, p = \frac{1}{6})$.
- **(h)** N/A. (Not the number of "heads" in a fixed number of "tosses".)
- **(i)** Coin-tossing model. $X_9 \sim \text{Binomial}(n = 12, p = \frac{1}{36})$.
- **(j)** Urn model with $N = 98$, $M = 44$, and $n = 7$. You can use the Binomial approximation, $X_{10} \sim \text{Binomial}(n = 7, p = \frac{44}{98})$, as $\frac{n}{N} < 0.1$.

2.
- **(a)** No. He does not guess all the questions so that p is not constant.
- **(b)** Yes. His little brother will guess them all so that $p = 0.2$.
- **(c)** Yes, provided the probability of having an error free page is constant.
- **(d)** Probably no. You can expect the probability of an error-free page to depend on the number of mathematical symbols and numbers on the page.
- **(e)** No, as the number of trials is not fixed in advance.
- **(f)** No, as what happens in successive months (trials) will almost certainly not be independent.

3.
- **(a)** $p = \frac{\text{sampled area}}{\text{population area}} = \frac{20\times100\times100}{2000\times2000} = \frac{1}{20}$.
- **(b)** The 420 animals may be regarded as 420 independent Binomial trials each with probability of success, where success means "found in the sample area"and failure means "found outside the sample area." The four assumptions are satisfied because the animals (trials) are independent.
  $X \sim \text{Binomial}(n = 400, p = \frac{1}{20})$.
- **(c)** For a single plot $p = \frac{100\times100}{2000\times2000} = \frac{1}{400}$ and $W \sim \text{Binomial}(n = 420, p = \frac{1}{400})$.
- **(d)** Any two of the following.
  - **(i)** Animals tend to exhibit social tendencies and so are not generally independent.
  - **(ii)** Animals do not move randomly but usually have well-defined territories or "home ranges."
  - **(iii)** The presence of observers may disturb the animals so that they move out of the area.

(iv) Some animals may be missed. Deer, for example, are very hard to spot.

4. **(a)** The ten fish are a simple random sample of fish. Fish do not lose their tags. All the fish stay alive and no new fish are born.

(b) Binomial($n = 10, p$) where $p = \frac{M}{N} = \frac{50}{1000} = 0.05$.

(c) $\text{pr}(X \geq 1) = 1 - \text{pr}(X = 0) = 0.4013$.

5. **(a)** $L \sim \text{Binomial}(n = 160, p = \frac{120}{80{,}000})$.

(b) $\text{E}(L) = np = 0.24$ (about one every four years).

(c) Each ship has the same probability of being lost and ships are lost or not lost independently. This is probably not true, but it may be a reasonable approximation.

6. Let X = number of women from the 30 who become pregnant in the first year. Then $X \sim \text{Binomial}(n = 30, p = 0.11)$.

(a) $\text{pr}(X = 0) = 0.0303$.

(b) $\text{pr}(X \leq 2) = 0.3442$.

(c) pr(*No pregnancies in 2 years*)

= pr(*No pregnancies in 1st year* and *no pregnancies in 2nd year)*

= pr(*No pregnancies in 1st year*) × pr(*No pregnancies in 2nd year)*

= 0.0303×0.0303 (Assuming independence and constant probs)

= 0.0009.

7. **(a)** Binomial($n = 12, p = 0.18$).

(b) pr(*No failures*) = $\text{pr}(X = 0) = 0.09242$.

8. The drugs will be discredited if less than 7 of the 12 patients recover. Let X = the number that recover and assume that $X \sim \text{Binomial}(n = 12, p = 0.5)$. Then $\text{pr}(X \leq 6) = 0.6128$. (What assumptions have been made?)

9. If X = number of rods that perform satisfactorily, you assume that $X \sim \text{Binomial}(n = 10, p = 0.80)$. Then $\text{pr}(X \leq 4) = 0.006369$.

10. **(a)** If X = number of sixes when 6 dice are rolled, then $X \sim \text{Binomial}(n = 6, p = \frac{1}{6})$. $\text{pr}(X \geq 1) = 1 - \text{pr}(X = 0) = 0.6651$.

(b) If Y = number of sixes when 12 dice are rolled, then $Y \sim \text{Binomial}(n = 12, p = \frac{1}{6})$. $\text{pr}(Y \geq 2) = 1 - \text{pr}(Y \leq 1) = 0.6187$.

(c) If Z = number of sixes when 18 dice are rolled, then $Z \sim \text{Binomial}(n = 18, p = \frac{1}{6})$. $\text{pr}(Z \geq 3) = 1 - \text{pr}(Z \leq 2) = 0.5974$.

11. **(a)** The value -0.39 is clearly in error (probabilities cannot be negative). We replace it with $1 - (0.23 + 0.18 + 0.17 + 0.13) = 0.29$.

(b) $\text{pr}(X \geq 1) = 0.18 + 0.17 + 0.13 = 0.48$.

(c) $\text{pr}(X \leq 0) = 0.23 + 0.29 = 0.52$.

(d) $E(X) = \sum x\,\text{pr}(x) = (-3)\times 0.23 + 0\times 0.29 + 1\times 0.18 + 3\times 0.17 + 8\times 0.13 = 1.04$,
$E[(X-\mu)^2] = \sum (x-\mu)^2 \text{pr}(x) = (-3-1.04)^2 \times 0.23 + (0-1.04)^2 \times 0.29.$
$+(1-1.04)^2\times 0.18 + (3-1.04)^2\times 0.17 + (8-1.04)^2 \times 0.13 = 11.0184.$
$\text{sd}(X) = \sqrt{11.0184} = 3.32.$

12. **(a)** Let B = "black face shows uppermost" $= \{2,5\}$.
Let E = "even numbered face" $= \{2,4,6\}$. Then $\text{pr}(B \text{ or } E) = \frac{4}{6} = \frac{2}{3}$.

(b) $\text{pr}(X=-10) = \text{pr}(\{5\}) = \frac{1}{6}$; $\text{pr}(X=-4) = \text{pr}(\{2\}) = \frac{1}{6}$;
$\text{pr}(X=0) = \text{pr}(\{1\}) = \frac{1}{6}$; $\text{pr}(X=3) = \text{pr}(\{3\}) = \frac{1}{6}$;
$\text{pr}(X=4) = \text{pr}(\{4\}) = \frac{1}{6}$; $\text{pr}(X=6) = \text{pr}(\{6\}) = \frac{1}{6}$.

(c) Expected amount in dollars won:
$E(X) = \sum x\,\text{pr}(x) = -10\times\frac{1}{6} - 4\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 6\times\frac{1}{6} = -\frac{1}{6}$. No, since one expects to lose.

13. **(a)** In the table, $a(b) = a \times 10^b$.

x	0	1.61	7.85	25.90	34.29	224.13	558.00
pr(x)	0.9333	6.511(−2)	1.190(−3)	2.381(−4)	9.524(−5)	1.720(−5)	1.190(−5)

(b) $E(X) = \sum x\text{pr}(x)$ is most quickly calculated by noting that here $\sum x\ \text{pr}(x) = \sum \frac{x\times n(x)}{3{,}780{,}000}$, where $n(x)$ is the number of vouchers with value x. Thus
$E(X) = \frac{1.61\times 246130 + 7.85\times 4500 + ... + 558.00\times 45}{3{,}780{,}000} = 0.1341$, or 13.41 cents.

(c) We calculate $E[(X-\mu)^2] = \sum (x-\mu)^2\text{pr}(x) = \frac{\sum (x-\mu)^2 \times n(x)}{3{,}780{,}000}$
$= \frac{(1.61-0.1341)^2\times 246130 + (7.85-0.1341)^2\times 4500 + ... + (558.00-0.1341)^2 \times 45}{3{,}780{,}000}$ and find the square root of the answer to obtain $\text{sd}(X) = 2.2509$.

(d) The expected redeemable value in cents is calculated by (since we only incur a postage cost if we get a voucher)
$E(X) - (\textit{Postage Cost}) \times \text{pr}(\textit{getting a voucher})$
$= 13.41 - 40 \times 0.0677 = 10.74 \approx 11$ cents. Thus the expected cost of a box of Almond Delight is \$1.84 – \$0.11 = \$1.73. At \$1.60, the alternative brand is better value.

14. Eventually you must get a prize or nothing. Dividing the probabilities by 0.9 you get

Prize	\$ 0	\$3	\$7	\$11	\$ 21	\$ 2100
Prob.	0.9333	0.0370	0.0111	0.011	7.407×10^{-3}	3.703×10^{-6}

$E(X) = \$0\times 0.933 + \$3\times 0.037 + ... + \$2100 \times 3.7\times 10^{-6} = 0.4742.$
$\text{sd}(X) = \sqrt{(0-0.4774)^2\times 0.933 + ... + (2100-0.4774)^2 \times 3.7\times 10^{-6}}$
$= 4.6473.$

15. Let X = number of defective rivets in the sample.

(a) $X \sim \text{Binomial}(n=8, p=0.01)$. Then $\text{pr}(X\geq 2) = 1 - \text{pr}(X\leq 1) = 0.00269$.

(b) $Y \sim \text{Binomial}(n=8, p=0.02)$. Then $\text{pr}(X \leq 1) = 0.9897$.

***16.** **(a)** $X \sim \text{Binomial}(n, p = \frac{1}{12{,}000{,}000})$. The following four assumptions support the Binomial: (i) the total number of trials or number of couples is fixed; (ii) a given couple either fits or doesn't fit the description; (iii) the chance of any couple matching the description is constant, and (iv) independence is assumed.

(b) $$\text{pr}(X \geq 2 \,|\, X \geq 1) = \frac{\text{pr}(X \geq 2 \text{ and } X \geq 1)}{\text{pr}(X \geq 1)} = \frac{\text{pr}(X \geq 2)}{\text{pr}(X \geq 1)} = \frac{1 - \text{pr}(X \leq 1)}{1 - \text{pr}(X = 0)}.$$

(c) $n = 10^6$, pr $= 0.04109$; $n = 4 \times 10^6$, pr $= 0.1574$; $n = 10^7$, pr $= 0.3595$.

17. **(a)** Let X = number of attempts made.
From the case study:

x	1	2	3	4
$\text{pr}(X = x)$	0.1	0.09	0.081	0.729

(b) $\text{E}(X) = \sum x\,\text{pr}(x) = 1 \times 0.1 + 2 \times 0.09 + 3 \times 0.081 + 4 \times 0.729 = 3.439$.
$\text{sd}(X) = \sqrt{\sum (x - \mu)^2 \text{pr}(x)} = \sqrt{(1 - 3.439)^2 \times 0.1 + ... + (4 - 3.439)^2 \times 0.729}$
$= \sqrt{1.0263} = 1.0131$.

(c) $\text{E}(Cost) = \$7,000 \times \text{E}(X) = \$7,000 \times 3.439 = \$24,073$.

(d) $\text{pr}(\textit{Still childless}) = (\frac{9}{10})^4 = 0.6561$. [Alternatively, use $\text{pr}(Y = 0)$ where $Y \sim \text{Binomial}(n = 4, p = 0.1)$.]

(e) and **(f)** The frequencies are computed in the following table.

Group	Starting number	Yes 1st	No	Yes 2nd	No	Yes 3rd	No	Yes 4th
1	30,000	6,000	24,000	4,800	19,200	3,840	15,360	3,072
2	30,000	3,000	27,000	2,700	24,300	2,430	21,870	2,187
3	40,000	400	39,600	396	39,204	392	38,812	388
Total	100,000	9,400	90,600	7,896	82,704	6,662	76,042	5,647

Here "Yes" means the number who were successful, "No" means the number unsuccessful and "1st" means first attempt, and so on.

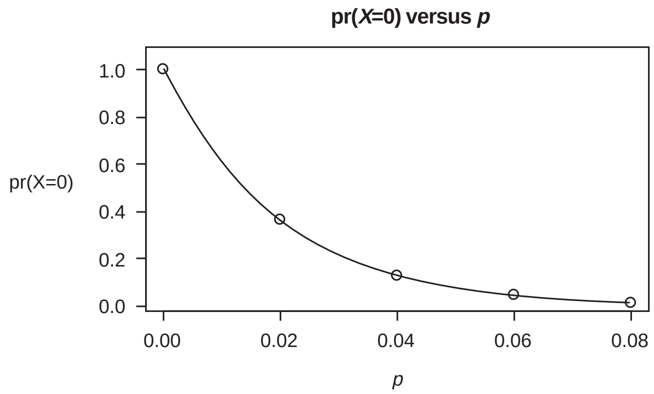
(g) Proportion who succeed on first attempt $= \frac{9{,}400}{100{,}000} = 0.094$.
Proportion who succeed on second attempt $= \frac{7{,}896}{90{,}600} = 0.0872$.
Proportion who succeed on third attempt $= \frac{6{,}662}{82{,}704} = 0.0806$.
Proportion who succeed on fourth attempt $= \frac{5{,}647}{76{,}042} = 0.0743$.

18. **(a)** You can use either an urn model with $\frac{n}{N} < 0.1$ and $p = \frac{M}{N}$ the proportion of defectives, or a coin tossing model with p the probability of getting a defective item. Either way you have $X \sim \text{Binomial}(n = 50, p)$.

(b) Implicit in our answers to (a) are the assumptions that the manufacturing process is stable or is in control, that no time drifts occur in the process which may otherwise lead to deterioration resulting from the wearing of machine parts or the loss of accuracy in machine settings. This implies that the probability of a defective item produced remains constant. You also assume that the sampling is taken with a sufficient time interval between selections to preserve the independence condition.

(c) Using $X \sim \text{Binomial}(n = 50, p)$, for $p = 0, 0.02, 0.04, 0.06, 0.08$ you obtain:

p	0	0.02	0.04	0.06	0.08
$\text{pr}(X = 0)$	1	0.3642	0.1299	0.0453	0.0155



(d) $\text{pr}(X = 0) = (1 - p)^{50}$. If $\text{pr}(X = 0) = 0.1$, then $1 - p = 0.1^{1/50} = 0.9550$. Hence $p = 0.045$.

19. Let X = number of incorrect identifications out of 91 calves. Then $X \sim \text{Binomial}(n = 91, p)$ and you can compute $\text{pr}(X \leq 1)$ for different p, namely:

p	0	0.02	0.04	0.06	0.08
$\text{pr}(X \leq 1)$	1	0.4545	0.1167	0.0244	0.0045

By calculating a couple more points between 0.04 and 0.06, and plotting a graph, you can estimate that the value of p giving $\text{pr}(X \leq 1) = 0.1$ is 0.042.

20. **(a)** Let X = number delivered in a day. Then $X \sim \text{Binomial}(n = 15, p = 0.95)$, and $\text{pr}(X \leq 11) = 0.00547$.

(b) Independence of the delivery of each of the 15 letters.

(c) If N.Z. Post's claim was true, we would be highly unlikely to get so few letters delivered the next day. 99.5% of the time $(1 - .005)$, we would get more than 11 letters delivered the next day. The data does not support N.Z. Post's claim.

21. If X is the number of mutations, then $X \sim \text{Binomial}(n = 20,000, p = \frac{1}{10{,}000})$.

(a) $\text{pr}(X = 0) = 0.1353$.

(b) $\text{pr}(X \geq 1) = 1 - \text{pr}(X = 0) = 0.8647$.

(c) $\text{pr}(X \leq 3) = 0.8571$.

22. If X is the number of cases for 150,000 people, then $X \sim \text{Binomial}(n = 150{,}000, p = \frac{3.1}{100{,}000})$. $E(X) = np = 4.65$, i.e., 4 or 5 claims per year. We also have, $\text{pr}(X > 6) = 1 - \text{pr}(X \leq 6) = 0.1886$.

23. **(a)** $X \sim \text{Binomial}(n = 10, p = \frac{1}{50})$. The Binomial distribution is appropriate since the person either wins or does not win a bottle, sampling is with replacement so that the draws are (supposedly) independent, and the probability of winning is the same for each draw, namely, $\frac{1}{50}$.

(b) **(i)** $\text{pr}(X = 0) = 0.81707$, $\text{pr}(X \leq 2) = 0.999136$. **(ii)** $\text{pr}(X \geq 3) = 1 - \text{pr}(X \leq 2) = 0.000864$.

(c) Since the probability in (b)(ii) is so small, the fact that one person has won three bottles would lead us to suspect that the names were not properly stirred.

(d) By b(ii), $\text{pr}(E_i) = \text{pr}(X \geq 3) \approx 0.00086 \neq 0$. If the event E_i has occurred for at least 3 people, then the event E_j $(j \neq i)$ cannot occur for the other people (since there are only 10 bottles to be won), so $\text{pr}(E_j) = 0$. This establishes that the outcome of E_i depends on the occurrence or otherwise of the other outcomes. Hence the E_i 's are not independent.

(e) Let Y = number of people who win 3 or more bottles. Assuming independence of the E_i's and regarding E_i as a binomial trial you have $Y \sim \text{Binomial}(n = 50, p = 0.000864)$. Hence $\text{pr}(Y \geq 1) = 1 - \text{pr}(Y = 0) = 0.0423$.

(f) Although the last probability in (e) is somewhat larger (slightly over 4 chances in 100), our opinion in (c) is not changed.

24. (a) The forty weeks of data is given below. Each row corresponds to a week, 1 corresponds to a day with rain and 0 corresponds to a day without rain.

Week 1 :	1	0	0	1	1	0	0
Week 2 :	1	0	0	0	0	0	0
Week 3 :	0	0	1	0	0	1	0
Week 4 :	0	1	1	0	0	1	1
Week 5 :	0	1	0	0	0	0	0
Week 6 :	0	0	1	0	1	0	0
Week 7 :	0	1	1	0	1	0	0
Week 8 :	1	1	0	0	0	1	0
Week 9 :	1	0	0	0	0	0	0
Week 10:	1	0	0	0	0	0	0
Week 11:	1	0	0	1	1	1	1
Week 12:	0	1	1	1	1	0	1
Week 13:	0	0	0	0	1	0	1
Week 14:	1	1	0	1	0	1	0
Week 15:	1	0	0	1	0	0	0
Week 16:	0	0	0	0	1	1	1
Week 17:	0	0	1	0	0	1	1
Week 18:	1	1	1	0	0	0	1
Week 19:	1	0	0	1	0	0	0
Week 20:	0	0	0	0	0	0	0
Week 21:	1	0	0	0	1	0	0
Week 22:	1	0	0	0	1	1	0
Week 23:	0	1	1	0	0	0	0
Week 24:	1	1	0	1	0	0	1
Week 25:	1	1	0	1	0	0	1
Week 26:	0	1	0	0	0	1	0
Week 27:	0	1	1	1	1	0	0
Week 28:	0	0	1	0	1	0	1
Week 29:	1	0	1	1	0	0	1
Week 30:	1	0	1	0	0	0	0
Week 31:	0	0	1	0	0	0	0
Week 32:	0	1	0	1	0	1	1
Week 33:	1	0	1	0	1	1	1
Week 34:	1	0	0	1	0	1	0
Week 35:	0	1	0	1	1	0	0
Week 36:	1	0	0	0	0	1	1
Week 37:	1	1	0	0	0	1	1
Week 38:	1	0	1	0	0	1	0
Week 39:	1	1	0	0	1	1	1
Week 40:	1	0	1	1	0	1	1

In our simulated weeks given above, we obtain the following (your's will be somewhat different):

- There were **26** weeks with more dry days than wet days.
- There were **14** weeks with more wet days than dry days.
- The longest run of wet days was **4** days (weeks 11, 12 and 27).
- The longest run of dry days was **7** days (week 20).
- The proportion of wet days was $\frac{117}{280} = 0.4178$.

(b) The number of wet days for each of the forty weeks is as follows.

3 1 2 4 1 2 3 3 1 1 5 5 2 4 2 3 3 4 2 0 2 3 2 4 4 2 4 3 4 2 1 4 5 3 3 3 4 3 5 5.

Each simulated day has a probability 0.4 of being wet, days are generated independently and we count the number of wet days ("heads") in a fixed number of "tosses", namely 7, so the distribution is Binomial($n = 7, p = 0.4$).
Forty random observations from a Binomial($n = 7, p = 0.4$) distribution were generated and are given below

0 5 3 3 3 2 1 4 2 4 5 3 2 2 3 4 3 3 2 2 4 3 2 1 4 3 6 3 2 4 5 3 1 3 5 2 1 4 5 2.

(c) Two hundred Binomial($n = 7, p = 0.4$) random numbers were generated; the numbers and the relative-frequency table are given below.

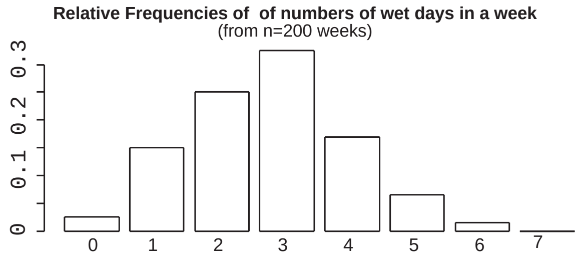
0	4	4	3	2	3	3	4	3	3	4	3	4	3	6	1	4	2	4	3	4	1	3	2	4
2	3	3	2	2	4	2	3	2	4	2	2	2	4	3	3	6	3	4	3	4	4	4	2	1
2	4	4	2	3	4	1	3	1	4	3	3	2	2	4	6	2	4	2	5	2	1	2	4	1
3	3	1	4	2	1	4	3	3	3	3	3	3	2	5	2	1	4	1	3	4	4	3	3	4
4	3	5	3	2	3	3	3	4	1	3	2	6	4	3	4	2	2	4	4	2	3	1	2	2
4	2	2	2	3	5	3	5	2	2	4	2	3	2	2	3	3	5	1	3	2	3	1	2	1
2	4	5	4	2	3	3	2	4	3	2	5	5	5	5	2	2	4	0	4	2	4	3	6	4
1	4	4	3	2	3	2	1	4	5	4	3	1	4	1	2	3	4	1	1	4	3	2	4	4

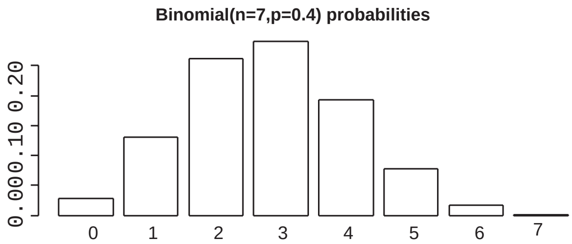
Days wet per week	0	1	2	3	4	5	6
Frequency	5	30	50	65	34	13	3
Rel. frequency	0.025	0.150	0.25	0.325	0.17	0.065	0.015

The probabilities for the Binomial($n = 7, p = 0.4$) distribution are shown in the following table. You will see that they are reasonably similar to the relative frequencies above. [The following probabilities have been rounded to 3 decimal places.]

x	0	1	2	3	4	5	6	7
pr(x)	0.028	0.131	0.261	0.290	0.194	0.077	0.017	0.002

A bar graph of the relative frequency of wet days in each of 200 weeks and a bar graph of the Binomial($n = 7, p = 0.4$) probabilities are given below.





The bar graphs are quite similar in shape. (The most distinct differences come in the bars relating to 2 and 4.). If we had used, say, 10,000 Binomial random numbers instead of only 200, there would be no visible difference.

(d) A relative frequency table of the longest run of dry days in each of 200 weeks is given below.

Longest dry	1	2	3	4	5	6	7
Frequency	5	16	8	6	1	3	1
Rel. Frequency	0.125	0.4	0.2	0.15	0.25	0.075	0.025

(e) Using the table in (d) the expected value of the longest run of dry days can be estimated by

$$1 \times 0.125 + 2 \times 0.4 + 3 \times 0.2 + 4 \times 0.15 + 5 \times 0.025 + 6 \times 0.075 + 7 \times 0.025 = 2.875.$$