Solutions to Exercise 17A

1 $\quad n = 5;\quad X = \text{number of boys};\quad p = \frac{1}{2};\quad q = \frac{1}{2}$

1a $\quad P(X = 5)$

$= {}^nC_x\, p^x q^{n-x} \quad \text{where } x = 5$

$= {}^5C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^0$

$= 1 \times \left(\frac{1}{2}\right)^5 \times 1$

$= \dfrac{1}{32}$

1b $\quad P(X = 3)$

$= {}^nC_x\, p^x q^{n-x} \quad \text{where } x = 3$

$= {}^5C_3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^2$

$= \dfrac{5}{16}$

1c $\quad P(X = 4)$

$= {}^nC_x\, p^x q^{n-x} \quad \text{where } x = 4$

$= {}^5C_4 \times \left(\frac{1}{2}\right)^4 \times \left(\frac{1}{2}\right)^1$

$= \dfrac{5}{32}$

1d $P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{5}\mathrm{C}_{0} \times \left(\frac{1}{2}\right)^{0} \times \left(\frac{1}{2}\right)^{5}$

$= \frac{31}{32}$

2 $n = 6$; $X =$ number of boundaries hit; $p = \frac{1}{5}$; $q = \frac{4}{5}$

$P(X = 2)$

$= {}^{n}\mathrm{C}_{x}\, p^{x} q^{n-x}$ where $x = 2$

$= {}^{6}\mathrm{C}_{2} \times \left(\frac{1}{5}\right)^{2} \times \left(\frac{4}{5}\right)^{4}$

$= \frac{768}{3125}$

3a Because $\frac{1300}{2000} \doteqdot \frac{1299}{1999} \doteqdot \frac{1298}{1998}$ ... and so on; i.e., the probability remains constant as the numbers are quite large and the ratios shown above almost remain constant.

3b $P(X = 12)$

$\doteqdot {}^{12}\mathrm{C}_{12} \times \left(\frac{1300}{2000}\right)^{12} \times \left(\frac{700}{2000}\right)^{0}$

$= 1 \times \left(\frac{13}{20}\right)^{12} \times 1$

$= \left(\frac{13}{20}\right)^{12}$

$= (0.65)^{12}$

Exact probability $= \frac{1300! \times 1988!}{2000! \times 1288!}$

4 $n = 12$; $X =$ number of times 5 appears on uppermost face of die; $p = \frac{1}{6}$; $q = \frac{5}{6}$

4a $P(X = 3)$

$= {}^{12}C_3 \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^9$

4b $P(X = 8)$

$= {}^{12}C_8 \times \left(\frac{1}{6}\right)^8 \times \left(\frac{5}{6}\right)^4$

4c $P(X \geq 10)$

$= P(X = 10) + P(X = 11) + P(X = 12)$

$= {}^{12}C_{10} \times \left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^2 + {}^{12}C_{11} \times \left(\frac{1}{6}\right)^{11} \times \left(\frac{5}{6}\right)^1 + {}^{12}C_{12} \times \left(\frac{1}{6}\right)^{12} \times \left(\frac{5}{6}\right)^0$

$= {}^{12}C_{10} \times \left(\frac{1}{6}\right)^{10} \times \left(\frac{5}{6}\right)^2 + {}^{12}C_{11} \times \left(\frac{1}{6}\right)^{11} \times \left(\frac{5}{6}\right)^1 + \left(\frac{1}{6}\right)^{12}$

5 $n = 6$; $N =$ number of times 3 is shown on uppermost face of die; $p = \frac{1}{6}$; $q = \frac{5}{6}$

5a $P(N = 2)$

$= {}^{6}C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^4$

$\doteqdot 0.2009$

5b $P(N < 2)$

$= P(N = 0) + P(N = 1)$

$= {}^{6}C_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^6 + {}^{6}C_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^5$

$\doteqdot 0.7368$

5c $P(N \geq 2)$

$= 1 - P(N < 2)$

$\doteqdot 1 - 0.7368$

$= 0.2632$

6 $n = 20$; $X =$ number of times archer hits bulls-eye; $p = \frac{9}{10}$; $q = \frac{1}{10}$

6a $P(X \geq 18)$

$= P(X = 18) + P(X = 19) + P(X = 20)$

$$= {}^{20}\mathrm{C}_{18} \times \left(\frac{9}{10}\right)^{18} \times \left(\frac{1}{10}\right)^{2} + {}^{20}\mathrm{C}_{19} \times \left(\frac{9}{10}\right)^{19} \times \left(\frac{1}{10}\right)^{1} + {}^{20}\mathrm{C}_{20} \times \left(\frac{9}{10}\right)^{20} \times \left(\frac{1}{10}\right)^{0}$$

$$= {}^{20}\mathrm{C}_{18} \times \left(\frac{9}{10}\right)^{18} \times \left(\frac{1}{10}\right)^{2} + {}^{20}\mathrm{C}_{19} \times \left(\frac{9}{10}\right)^{19} \times \left(\frac{1}{10}\right)^{1} + \left(\frac{9}{10}\right)^{20} \times \left(\frac{1}{10}\right)^{0}$$

6b $P(\text{misses at least once})$

$= 1 - P(\text{no misses})$

$= 1 - P(X = 20)$

$$= 1 - {}^{20}\mathrm{C}_{20} \times \left(\frac{9}{10}\right)^{20} \times \left(\frac{1}{10}\right)^{0}$$

$$= 1 - \left(\frac{9}{10}\right)^{20}$$

7 $n = 10$; $X =$ number of defective bulbs; $p = 0.09$; $q = 0.91$

7a $P(X \leq 2)$

$= P(X = 0) + P(X = 1) + P(X = 2)$

$$= {}^{10}\mathrm{C}_{0} \times (0.09)^0 \times (0.91)^{10} + {}^{10}\mathrm{C}_{1} \times (0.09)^1 \times (0.91)^9 + {}^{10}\mathrm{C}_{2} \times (0.09)^2 \times (0.91)^8$$

$$= (0.91)^{10} + {}^{10}\mathrm{C}_{1} \times (0.09)^1 \times (0.91)^9 + {}^{10}\mathrm{C}_{2} \times (0.09)^2 \times (0.91)^8$$

7b $P(X \geq 2)$

$= 1 - \big(P(X = 0) + P(X = 1)\big)$

$= 1 - ({}^{10}C_0 \times (0.09)^0 \times (0.91)^{10} + {}^{10}C_1 \times (0.09)^1 \times (0.91)^9)$

$= 1 - (0.91)^{10} - {}^{10}C_1 \times (0.09)^1 \times (0.91)^9$

8a $S = \{\text{HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH,}$
$\text{THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT}\}$

$n(S) = 16 = 2^4$

8b Let X be the event when Janice wins.

$X = \{\text{HHTT, HTHT, HTTH, THHT, THTH, TTHH}\}$; $n(X) = 6$

$P(X = 6)$

$= \frac{n(X)}{n(S)}$

$= \frac{6}{16}$

$= \frac{3}{8}$

8c If we are to order the letters of the word HHTT, we have four places and two choices for each place.

So the number of arrangements $= {}^4C_2 = 6$. This answer agrees with part b.

8d The number of ways of choosing the two coins from four that are to be heads up is ${}^4C_2$.

9 $n = 5$; $X =$ number who support WTP party policies; $p = 0.55$; $q = 0.45$

Majority means 3 or more people.

$P(X \geq 3)$

$= P(X = 3) + P(X = 4) + P(X = 5)$

$= {}^5C_3 \times (0.55)^3 \times (0.45)^2 + {}^5C_4 \times (0.55)^4 \times (0.45)^1 + {}^5C_5 \times (0.55)^5 \times (0.45)^0$

$= 0.593\ 126\ \ldots$

$\doteqdot 0.593$

10 $n = 31$; $X =$ number of days a small earthquake occurs; $p = 0.95$; $q = 0.05$

$P(X = 28)$

$= {}^{31}\mathrm{C}_{28} \times (0.95)^{28} \times (0.05)^{3}$

11 $n = 10$; $X =$ number of times a jackpot prize is won; $p = 0.012$; $q = 0.988$;

11a i $P(X = 1)$

$= {}^{10}\mathrm{C}_{1} \times (0.012)^{1} \times (0.988)^{9}$

$= 0.107\ 644\ 91\ \ldots$

$\doteqdot 0.107\ 64$

11a ii $P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}\mathrm{C}_{0} \times (0.012)^{0} \times (0.988)^{10}$

$= 1 - (0.988)^{10}$

$= 0.113\ 723\ 06\ \ldots$

$\approx 0.113\ 72$

11b For the jackpot to reach \$200 000, the prize must be won on the 20th draw.

$n = 20$; $x = 20$; $p = 0.012$; $q = 0.988$

$P(X = 20)$

$= {}^{20}\mathrm{C}_{0} \times (0.012)^{0} \times (0.988)^{20}$

$= 0.785\ 485\ 486\ \ldots$

$\doteqdot 0.785\ 49$

12a $n = ?$; $X =$ number of times a six is rolled; $p = \frac{1}{6}$; $q = \frac{5}{6}$

Since $P(\text{rolling at least one six}) = 1 - P(\text{rolling no sixes})$

$P(X \geq 1) = 1 - P(X = 0) > 0.95$

$P(X = 0) < 0.05$

$^nC_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^n < 0.05$

$1 \times 1 \times \left(\frac{5}{6}\right)^n < 0.05$

$\left(\frac{5}{6}\right)^n < 0.05$

$\log_e \left(\frac{5}{6}\right)^n < \log_e 0.05$

$n \log_e \left(\frac{5}{6}\right) < \log_e 0.05$

$n > \dfrac{\log_e 0.05}{\log_e \left(\frac{5}{6}\right)}$

$n > 16.431\ldots$

The die must be rolled 17 times.

12b $n = ?$; $X =$ number of times a tail is tossed; $p = \frac{1}{2}$; $q = \frac{1}{2}$

$P(\text{tossing at least one tail}) > 0.99$

Since $P(\text{tossing at least one tail}) = 1 - P(\text{tossing no tails})$

$P(X \geq 1) = 1 - P(X = 0) > 0.99$

$P(X = 0) < 0.01$

$^nC_0 \times \left(\frac{1}{2}\right)^0 \times \left(\frac{1}{2}\right)^n < 0.01$

$1 \times 1 \times \left(\frac{1}{2}\right)^n < 0.01$

$\left(\frac{1}{2}\right)^n < 0.01$

$\log_e \left(\frac{1}{2}\right)^n < \log_e 0.01$

$n \log_e \left(\frac{1}{2}\right) < \log_e 0.01$

$n > \dfrac{\log_e 0.01}{\log_e \left(\frac{1}{2}\right)}$

$n > 6.643\ 856\ldots$

The coin must be tossed 7 times.

13a i For the probability that a family has three boys:

$p = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$

For the probability a family does not have three boys:

$q = 1 - \frac{1}{8}$

$= \frac{7}{8}$

The probability is same for each family.

$n = 5$; $X =$ number of families with three boys; $p = \frac{1}{8}$; $q = \frac{7}{8}$

So, for the probability that at least one family has three boys:

$P(\text{at least one family has three boys}) = 1 - P(\text{no family has three boys})$

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^5\mathrm{C}_0 \times \left(\frac{1}{8}\right)^0 \times \left(\frac{7}{8}\right)^5$

$= 1 - 0.512\ 90\ldots$

$= 0.487\ 091\ \ldots$

$\doteqdot 0.487$

13a ii For a family to have more boys than girls, then each family must have 2 boys and one girl, or 3 boys.

$n = 3$; $X =$ number of boys in a family; $p = \frac{1}{2}$; $q = \frac{1}{2}$

$P(X \geq 2)$

$= P(X = 2) + P(X = 3)$

$= {}^3C_2 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right) + {}^3C_3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^0$

$= 3 \times \left(\frac{1}{2}\right)^3 + 1 \times \left(\frac{1}{2}\right)^3 \times 1$

$= \dfrac{1}{2}$

So, for the probability that each of the five families has more boys than girls:

$n = 5$; $N =$ number of families with at least two boys; $p = \frac{1}{2}$; $q = \frac{1}{2}$

$P(N = 5)$

$= {}^5C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^0$

$= \dfrac{1}{32}$

$= 0.03125$

$\doteqdot 0.031$

13b Following assumptions are made:

1. Each birth is independent.
2. There are only two genders. (biologically true genders)
3. The probability of a child being a boy or a girl is equal.
4. Each birth in each family is an independent Bernoulli trial.

14a The argument is invalid. Normally, mathematics books are grouped together, so that once the shelf is chosen, one would expect all or none of the books to be mathematics books, thus the five stages are not independent events. The result would be true if the books were each chosen at random from the library.

14b The argument is invalid. People in a particular neighbourhood tend to vote more similarly than the population at large, so the four events are not independent. This method also oversamples small streets, which may introduce an additional bias.

15a $P(\text{rains on a winter day}) = \frac{18}{30} = \frac{3}{5}$

So, $p = \dfrac{3}{5}$; $\quad q = \dfrac{2}{5}$

$P(\text{first two days are fine and next three days are wet})$

$= \left(\frac{2}{5}\right) \times \left(\frac{2}{5}\right) \times \left(\frac{3}{5}\right) \times \left(\frac{3}{5}\right) \times \left(\frac{3}{5}\right)$

$= \dfrac{108}{3125}$

$= 0.034\ 56$

$\doteqdot 0.0346$

15b $n = 5$; $X =$ number of days that it rains; $p = \frac{3}{5}$; $\quad q = \frac{2}{5}$

$P(X \geq 3)$

$= P(X = 3) + P(X = 4) + P(X = 5)$

$= {}^5C_3 \times \left(\frac{3}{5}\right)^3 \times \left(\frac{2}{5}\right)^2 + {}^5C_4 \times \left(\frac{3}{5}\right)^4 \times \left(\frac{2}{5}\right)^1 + {}^5C_5 \times \left(\frac{3}{5}\right)^5 \times \left(\frac{2}{5}\right)^0$

$= 0.682\ 56$

$\doteqdot 0.6827$

16a $n = 4$; $X_s =$ number of serves in; $\quad p_s = \frac{8}{10} = \frac{4}{5}$; $\quad q_s = \frac{1}{5}$

$P(X = 4)$

$= {}^4C_4 \times \left(\frac{4}{5}\right)^4 \times \left(\frac{1}{5}\right)^0$

$= 0.409\ 600$

16b $n = 4; X_\mathrm{a} = \text{number of aces served};\quad p_\mathrm{a} = \frac{1}{15};\quad q_\mathrm{a} = \frac{14}{15}$

$P(X \geq 3)$

$= P(X = 3) + P(X = 4)$

$= {}^4C_3 \times \left(\frac{1}{15}\right)^3 \times \left(\frac{14}{15}\right) + {}^4C_4 \times \left(\frac{1}{15}\right)^4 \times \left(\frac{14}{15}\right)^0$

$= 0.001\,125\,92 \ldots$

$\fallingdotseq 0.001\,126$

16c $n = 4;\quad x_a = 3;\quad p_\mathrm{a} = \frac{1}{15};\quad q_\mathrm{a} = \frac{14}{15};\quad p_\mathrm{s} = \frac{4}{5};\quad q_\mathrm{s} = \frac{1}{5}$

$P(\text{exactly three aces and other serve is in})$

$= P(X_a = 3 \text{ and } X_s = 1)$

$= {}^4C_3 \times \left(\frac{1}{15}\right)^3 \times \left(\frac{14}{15}\right) \times \left(\frac{4}{5}\right)$

$= 0.000\,884\,93 \ldots$

$\fallingdotseq 0.000\,885$

17a $X_1 = \text{number of 1955 model cars that start};\ n_1 = 6;\quad p_1 = 0.65;\ q_1 = 0.35$

$X_2 = \text{number of 1962 model cars that start};\ n_2 = 4;\quad p_2 = 0.8;\ q_2 = 0.2$

$P(X_1 = 3 \text{ and } X_2 = 1)$

$= {}^6C_3 \times (0.65)^3 \times (0.35)^3 \times {}^4C_1 \times (0.8)^1 \times (0.2)^3$

$\fallingdotseq 0.0060$

17b Four 1955 models:

$P(X_1 = 4 \text{ and } X_2 = 0)$

$= {}^6C_4 \times (0.65)^4 \times (0.35)^2 \times {}^4C_0 \times (0.8)^0 \times (0.2)^4$

$= 0.000\,524\,8 \ldots$

Four 1962 models:

$P(X_1 = 0 \text{ and } X_2 = 4)$

$= {}^6C_0 \times (0.65)^0 \times (0.35)^6 \times {}^4C_4 \times (0.8)^4 \times (0.2)^0$

$= (0.35)^6 \times (0.8)^4$

$= 0.000\,752\,953 \ldots$

Two of each model:

$P(X_1 = 2 \text{ and } X_2 = 2)$

$= {}^6C_2 \times (0.65)^2 \times (0.35)^4 \times {}^4C_2 \times (0.8)^2 \times (0.2)^2$

$= 0.014\,607\,684 \ldots$

Three 1955 models and one 1962 model:

$P(X_1 = 3 \text{ and } X_2 = 1)$

$= {}^6C_3 \times (0.65)^3 \times (0.35)^3 \times {}^4C_1 \times (0.8)^1 \times (0.2)^3$

$= 0.006\,028\,568 \ldots$

One 1955 model and three 1962 models:

$P(X_1 = 1 \text{ and } X_2 = 3)$

$= {}^6C_1 \times (0.65)^1 \times (0.35)^5 \times {}^4C_3 \times (0.8)^3 \times (0.2)^1$

$= 0.008\,390\,054 \ldots$

Hence, probability of exactly four cars starting

$= 0.030\,304\,059 \ldots$

$\doteqdot 0.0303$

18a X_{GD} = number of Golden Delicious apples to be discarded; $p_{\text{GD}} = \frac{1}{50}$;

X_{RD} = number of Red Delicious apples to be discarded; $p_{\text{RD}} = \frac{1}{100}$

$\frac{4}{5}$ of apples are Red Delicious; $\frac{1}{5}$ of apples are Golden Delicious

$P(\text{selected apple is discarded})$

$= \left(\frac{4}{5}\right) \times \left(\frac{1}{100}\right) + \left(\frac{1}{5}\right) \times \left(\frac{1}{50}\right)$

$= \frac{3}{250}$

18b $n = 10$; $X =$ number of apples to be discarded; $p = \frac{3}{250}$; $q = \frac{247}{250}$

18b i $P(X = 10)$

$= {}^{10}C_{10} \times \left(\frac{3}{250}\right)^{10} \times \left(\frac{247}{250}\right)^{0}$

$= \left(\frac{3}{250}\right)^{10}$

18b ii $P(X = 5)$

$= {}^{10}C_{5} \times \left(\frac{3}{250}\right)^{5} \times \left(\frac{247}{250}\right)^{5}$

18b iii $P(X < 2)$

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

$= {}^{10}C_{0} \times \left(\frac{3}{250}\right)^{0} \times \left(\frac{247}{250}\right)^{10} + {}^{10}C_{1} \times \left(\frac{3}{250}\right)^{1} \times \left(\frac{247}{250}\right)^{9}$

$= \left(\frac{247}{250}\right)^{10} + 10\left(\frac{3}{250}\right)^{1}\left(\frac{247}{250}\right)^{9}$

19a Probability of selecting a bag $= \frac{1}{2}$; $p_{\text{bag 1,red}} = \frac{3}{8}$; $p_{\text{bag 2,red}} = \frac{4}{8}$

$P(\text{selecting a bag and selecting a red ball})$

$= P(\text{selecting bag 1 and selecting red ball or selecting bag 2 and selecting red ball})$

$= \left(\frac{1}{2}\right) \times \left(\frac{3}{8}\right) + \left(\frac{1}{2}\right) \times \left(\frac{4}{8}\right)$

$= \frac{7}{16}$

19b $n = 8$; $X =$ number of red balls drawn; $p = \frac{7}{16}$; $q = \frac{9}{16}$

19b i $P(X = 3)$

$= {}^8C_3 \times \left(\frac{7}{16}\right)^3 \times \left(\frac{9}{16}\right)^5$

19b ii $P(X \geq 3)$

$= 1 - P(X < 3)$

$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$

$= 1 - \left[{}^8C_0 \times \left(\frac{7}{16}\right)^0 \times \left(\frac{9}{16}\right)^8 + {}^8C_1 \times \left(\frac{7}{16}\right)^1 \times \left(\frac{9}{16}\right)^7 + {}^8C_2 \times \left(\frac{7}{16}\right)^2 \times \left(\frac{9}{16}\right)^6\right]$

$= 1 - \left(\frac{9}{16}\right)^8 - 8\left(\frac{7}{16}\right)^1\left(\frac{9}{16}\right)^7 - {}^8C_2\left(\frac{7}{16}\right)^2\left(\frac{9}{16}\right)^6$

20a Probability that the number showing is even for a die: $p_{even} = \frac{1}{2}$

More even faces than odd faces for six dice in one throw means rolling 4, 5 or 6 even faces.

$n = 6$; $X =$ number of even faces on one throw of six dice; $p = \frac{1}{2}$; $q = \frac{1}{2}$

$P(X \geq 4)$

$= P(X = 4) + P(X = 5) + P(X = 6)$

$= {}^6C_4 \times \left(\frac{1}{2}\right)^4 \times \left(\frac{1}{2}\right)^2 + {}^6C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right) + {}^6C_6 \times \left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^0$

$= 15\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 + 6\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^6$

$= 22 \times \left(\frac{1}{2}\right)^6$

$= 0.343\ 75$

Number of times this happens in 100 throws

$= 100 \times 0.343\ 75$

$= 34.375$

$\doteqdot 34$

20b Probability that a head is tossed with a coin: $p_{\text{head}} = \frac{1}{2}$

More heads than tails in one toss of eight coins means tossing 5, 6, 7 or 8 heads.

$n = 8; X =$ number of heads on one toss of eight coins; $p = \dfrac{1}{2}; q = \dfrac{1}{2}$

$P(X \geq 5)$

$= P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8)$

$= {}^8C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^3 + {}^8C_6 \times \left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^2 + {}^8C_7 \times \left(\frac{1}{2}\right)^7 \times \left(\frac{1}{2}\right)^1 + {}^8C_8 \times \left(\frac{1}{2}\right)^8 \times \left(\frac{1}{2}\right)^0$

$= 56\left(\frac{1}{2}\right)^8 + 28\left(\frac{1}{2}\right)^8 + 8\left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right)^8$

$= 93 \times \left(\frac{1}{2}\right)^8$

$= 0.363\,281\,25$

Number of times this happens in 60 throws

$= 60 \times 0.363\,281\,25$

$= 21.796\,875$

$\doteqdot 22$

21a $n = 5; X =$ number of games where 19 is drawn; $p = 0.2;\ q = 0.8$

$P(X = 2)$

$= {}^8C_5 \times (0.2)^2 \times (0.8)^3$

$= 0.2048$

21b $n = 5;\ X =$ number of games where 19 is drawn; $p = 0.2;\ q = 0.8$

$P(X \geq 2)$

$= 1 - P(X < 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - [{}^5C_0 \times (0.2)^0 \times (0.8)^5 + {}^5C_1 \times (0.2)^1 \times (0.8)^4]$

$= 0.262\,72$

21c i Suppose we select the ball numbered n in the first draw. Now, we have $(n-1)$ choices for the second draw, $(n-2)$ for the third and $(n-3)$ for the fourth. Each time we draw a ball, the total number of balls reduce by 1. So, the probability that all 4 numbers selected are less than or equal to n is:

$P(X \leq n)$

$$= \frac{n}{20} \times \frac{n-1}{19} \times \frac{n-2}{18} \times \frac{n-3}{17}$$

$$= \frac{n(n-1)(n-2)(n-3)}{20 \times 19 \times 18 \times 17}$$

Alternatively, using combinatorics, the probability is:

$$\frac{\text{Number of ways to select 4 from } n}{\text{Number of ways to select 4 from 20}}$$

$$= {}^{n}C_4 \div {}^{20}C_4$$

$$= \frac{n!}{(n-4)!\,4!} \div \frac{20!}{16!\,4!}$$

$$= \frac{n!}{(n-4)!\,4!} \times \frac{16!\,4!}{20!}$$

$$= \frac{n(n-1)(n-2)(n-3)}{20 \times 19 \times 18 \times 17}$$

21c ii The number of ways to choose a ball labelled n and three other balls labelled with any number up to $n-1$ is $1 \times {}^{n-1}C_3$. Dividing by the total number of unrestricted combinations gives the result.

Hence, the probability that n is the largest of the numbers drawn

$$= \frac{{}^{n-1}C_3}{{}^{20}C_4}$$

22a $(a+b+c)^3$

$= (a+b+c)(a+b+c)(a+b+c)$

$= (a^2+ab+ac+ba+b^2+bc+ca+cb+c^2)(a+b+c)$

$= (a^2+b^2+c^2+2ab+2ac+2bc)(a+b+c)$

$= a^3+a^2b+a^2c+b^2a+b^3+b^2c+c^2a+c^2b+c^3+2a^2b+2ab^2+2abc$
$+2a^2c+2abc+2ac^2+2abc+2b^2c+2bc^2$

$= a^3+b^3+c^3+3a^2b+3ab^2+3b^2c+3bc^2+3ac^2+3a^2c+\ 6abc$

22b Let a correspond with Hawthorn, b with Collingwood and c with Sydney.

$p_{\text{Hawthorn}} = 0.65 = a;\quad p_{\text{Collingwood}} = 0.24 = b;\quad p_{\text{Sydney}} = 0.11 = c$

22b i The coefficient of the abc term is 6, demonstrating there are six ways to rearrange this outcome abc amongst the three supporters.

P(one supporter of each team is selected)

$= 6abc$

$= 6 \times 0.65 \times 0.24 \times 0.11$

$= 0.102\ 96$

22b ii This outcome corresponds to ab^2 or b^2c.

P(exactly two supporters of Collingwood are selected)

$= 3ab^2 + 3b^2c$

$= 3 \times 0.65 \times (0.24)^2 + 3 \times (0.24)^2 \times 0.11$

$= 0.131\ 328$

$\doteqdot 0.131\ 33$

22b iii $P(\text{at least two supporters of the same team are selected})$

$= 1 - P(\text{one supporter of each team is selected})$

$= 1 - 6abc$

$= 1 - 0.102\ 96$

$= 0.897\ 04$

Solutions to Exercise 17B

1a Rain is a seasonal phenomenon and cannot be considered as a strictly independent event. If it rains one day it is more likely to rain the next day because rainy days tend to come in groups.

1b Yes, this can be modelled as a binomial random variable.

$X =$ number of throws where the result was less than 5

1c Cases may repeat and there is no certainty if the game will ever end or not. It cannot be predicted how many trials will be required. The stages are not independent because if she wins, then the game stops.

1d Yes, this can be modelled as a binomial random variable.

$X =$ number of heads turning up in 20 trials

1e Yes, this can be modelled as a binomial random variable.

Strictly, the pens are not replaced, so the probability changes as each pen is removed and tested. If the population of pens is large, then p is almost constant with each selection, and it could be modelled with a binomial distribution.

$X =$ number of defective pens in the batch of 20

1f No, there are not two outcomes at each stage. The pupil just goes to school and measures the travel time. It could be modified to 'arrives on time' or 'takes less than 20 minutes', but the events may still not be independent.

1g Yes, this can be modelled as a binomial random variable. Note that while the experiment is different at each stage, the probabilities at each stage are independent and have the same probability 0.01 of success.

$X =$ selecting a number that matches n

2a

Number of heads x	0	1	2	3	4	5	6	Total
Number of ways	1	6	15	20	15	6	1	64
Probability p	0.016	0.094	0.234	0.313	0.234	0.094	0.016	1
xp	0	0.094	0.469	0.938	0.938	0.469	0.094	3
x^2p	0	0.094	0.938	2.813	3.75	2.344	0.563	10.5

2b Mode = 3 heads

2c Expected value $= \mu$

$$= \sum xp$$

$$= 3$$

Variance $= \sum x^2p - \mu^2$

$$= 10.5 - 3^2$$

$$= 10.5 - 9$$

$$= 1.5$$

2d $\mathrm{E}(X) = np$

$$= 6 \times \frac{1}{2}$$

$$= 3$$

$\mathrm{Var}(X) = npq$

$$= 6 \times \frac{1}{2} \times \frac{1}{2}$$

$$= 1.5$$

We get the same results.

2e The distribution is symmetric, thus the centre of the distribution is exactly the midpoint.

3a $n = 5$; $X =$ number of heads that are face up when a coin is tossed five times;

$p = \frac{1}{2}$; $q = \frac{1}{2}$

x	0	1	2	3	4	5
$P(X = x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{5}{32}$	$\frac{1}{32}$

Mode $= 2$ or 3

$\mu = np$

$= 5 \times \frac{1}{2}$

$= \frac{5}{2}$

So, mean $= \frac{5}{2}$

$\sigma = \sqrt{npq}$

$= \sqrt{5 \times \frac{1}{2} \times \frac{1}{2}}$

$= \sqrt{\frac{5}{4}}$

$= \frac{\sqrt{5}}{2}$

So, standard deviation $= \frac{\sqrt{5}}{2}$

3b $n = 5$; $X =$ number of times 5 or 6 occurs when a die is thrown five times;

$p = \frac{1}{3}$; $q = \frac{2}{3}$

x	0	1	2	3	4	5
$p(X = x)$	$\frac{32}{243}$	$\frac{80}{243}$	$\frac{80}{243}$	$\frac{40}{243}$	$\frac{10}{243}$	$\frac{1}{243}$

Mode $= 1$ or 2

$\mu = np$

$= 5 \times \frac{1}{3}$

$= \frac{5}{3}$

So, mean $= \frac{5}{3}$

$\sigma = \sqrt{npq}$

$= \sqrt{5 \times \frac{1}{3} \times \frac{2}{3}}$

$= \sqrt{\frac{10}{9}}$

$= \frac{\sqrt{10}}{3}$

So, standard deviation $= \frac{\sqrt{10}}{3}$

3c $n = 5$; $X =$ number of court cards when five cards are drawn;

$p = \frac{12}{52}$; $q = \frac{40}{52}$

x	0	1	2	3	4	5
$p(X = x)$	0.269	0.404	0.242	0.073	0.011	0.001

Mode $= 1$

$\mu = np$

$= 5 \times \frac{12}{52}$

$= 1.153\,84\ldots$

$\doteqdot 1.154$

So, mean $\doteqdot 1.154$

$\sigma = \sqrt{npq}$

$$= \sqrt{5 \times \frac{12}{52} \times \frac{40}{52}}$$

$$= 0.942\ 11 \ldots$$

$$\doteqdot 0.942$$

So, standard deviation $\doteqdot 0.942$

4a $n = 24$; $X =$ number of times Larry wins; $p = \frac{8}{20} = \frac{2}{5}$; $q = \frac{12}{20} = \frac{3}{5}$

$$\mu = np$$

$$= \frac{48}{5}$$

$$= 9.6$$

$$\sigma = \sqrt{npq}$$

$$= \sqrt{\frac{144}{25}}$$

$$= \frac{12}{5}$$

$$= 2.4$$

The expected value is 9.6 wins and the standard deviation is 2.4.

4b $x = 6$; $\mu = 9.6$; $\sigma = 2.4$

$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{6 - 9.6}{2.4}$$

$$= -1.5$$

Larry's result is 1.5 standard deviations below the mean.

5a $n = 6$; $X =$ number of sixes from throw of six dice; $p = \frac{1}{6}$; $q = \frac{5}{6}$

$P(X \geq 2)$

$= 1 - P(X < 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - \left[{}^{6}C_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^6 + {}^{6}C_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^5\right]$

$= 1 - \left(\frac{5}{6}\right)^6 - 6 \times \frac{5^5}{6^6}$

$= 0.263\ 22 \ldots$

$\doteqdot 26\%$

$\mathrm{E}(X) = np$

$= 6 \times \frac{1}{6}$

$= 1$

$\sigma = \sqrt{npq}$

$= \sqrt{6 \times \frac{1}{6} \times \frac{5}{6}}$

$= \sqrt{\frac{5}{6}}$

$= 0.912\ 87 \ldots$

$\doteqdot 0.91$

5b i $n = 12$; $X =$ number of sixes from throw of 12 dice; $p = \frac{1}{6}$; $q = \frac{5}{6}$

$P(X \geq 2)$

$= 1 - P(X < 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - \left[{}^{12}C_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^{12} + {}^{12}C_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^{11}\right]$

$= 1 - \left(\frac{5}{6}\right)^{12} - 12 \times \frac{5^{11}}{6^{12}}$

$= 0.618\ 66 \ldots$

$\doteqdot 62\%$

$\mathrm{E}(X) = np$

$= 12 \times \frac{1}{6}$

$= 2$

$\sigma = \sqrt{npq}$

$= \sqrt{12 \times \frac{1}{6} \times \frac{5}{6}}$

$= \sqrt{\frac{5}{3}}$

$= 1.290\ 99 \ldots$

$\doteqdot 1.29$

5b ii $n = 24$; $X =$ number of sixes from throw of 24 dice; $p = \frac{1}{6}$; $q = \frac{5}{6}$

$P(X \geq 2)$

$= 1 - P(X < 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - \left[{}^{24}\mathrm{C}_0 \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^{24} + {}^{24}\mathrm{C}_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^{23}\right]$

$= 1 - \left(\frac{5}{6}\right)^{24} - 24 \times \frac{5^{23}}{6^{24}}$

$= 0.927\ 04 \ldots$

$\doteqdot 93\%$

$\mathrm{E}(X) = np$

$= 24 \times \frac{1}{6}$

$= 4$

$\sigma = \sqrt{npq}$

$= \sqrt{24 \times \frac{1}{6} \times \frac{5}{6}}$

$= \sqrt{\frac{10}{3}}$

$= 1.825\,74\ldots$

$\doteqdot 1.83$

6a From graph, distribution has longer 'tail' on the right. So, the distribution is right skewed.

6b $n = 48; p = 0.25; q = 0.75$

$\mu = np$

$= 48 \times 0.25$

$= 12$

$\sigma = \sqrt{npq}$

$= \sqrt{48 \times 0.25 \times 0.75}$

$= \sqrt{9}$

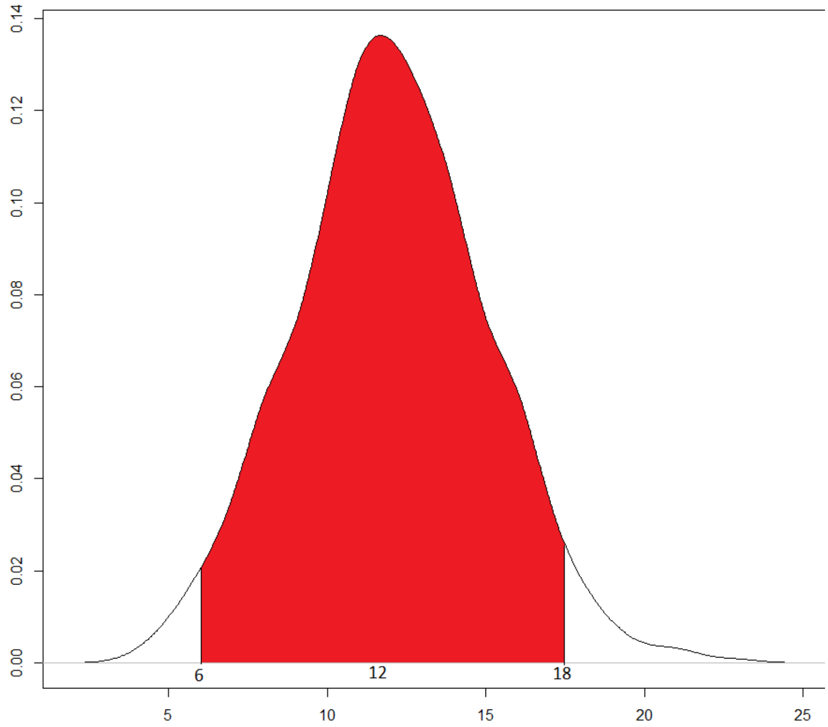
$= 3$

6c Mode $= 12$

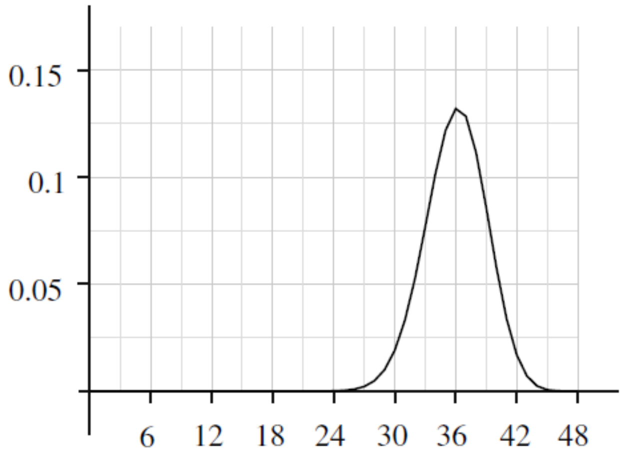
(Most likely outcome is value along horizontal axis of graph that gives the highest vertical value.)

From graph, $P(X = 12) \doteqdot 0.13$

6d Shade the region bounded by [6, 18] on the horizontal axis.



6e This is the graph for $p = 0.25$ reflected horizontally in $x = 24$.



7 Spreadsheet investigation

7e $P(X \geq 60) = 0.0284439\ldots \approx 0.028$

7f $P(30 \leq X \leq 50) = 0.864357\ldots \approx 0.864$

7g Mode = 50 heads

7h $n = 100;\quad p = 0.5;\quad q = 0.5;\quad \mu = 50;\quad \sigma = 5$

$P\big((50 - i) \leq X \leq (50 + i)\big) = 0.5$; 68% of data lies within 1 standard deviation of the mean. So, $i < 5$; Starting from $i = 1$ to $i = 4$; we find that for $i = 3$; the probability $p(X = x) > 0.5$. Hence, $i = 3$, and the interval is $[47, 53]$.

8 $n = 48$; $X =$ number of questions a person gets correct; $p = \frac{1}{4}$; $q = \frac{3}{4}$

8a
$$\begin{aligned} \mathrm{E}(X) &= np \\ &= 48 \times \frac{1}{4} \\ &= 12 \end{aligned}$$

Assuming all the students simply guess the answer or randomly tick an option, a student would be expected to get 12 correct answers.

8b
$$\begin{aligned} \sigma &= \sqrt{npq} \\ &= \sqrt{48 \times \frac{1}{4} \times \frac{3}{4}} \\ &= \sqrt{\frac{144}{16}} \\ &= \sqrt{9} \\ &= 3 \end{aligned}$$

8c $x = 24$
$$\begin{aligned} z &= \frac{x - \mu}{\sigma} \\ &= \frac{24 - 12}{3} \\ &= 4 \end{aligned}$$

Fayola's score is 4 standard deviations away from the mean.

8d $n = 100$; Y = number of questions a person gets correct; $p = \frac{1}{5}$; $q = \frac{4}{5}$

$$\begin{aligned}\mathrm{E}(Y) &= np \\ &= 100 \times \frac{1}{5} \\ &= 20\end{aligned}$$

$$\begin{aligned}\sigma &= \sqrt{npq} \\ &= \sqrt{100 \times \frac{1}{5} \times \frac{4}{5}} \\ &= \sqrt{16} \\ &= 4\end{aligned}$$

8e $x = 40$

$$\begin{aligned}z &= \frac{x-\mu}{\sigma} \\ &= \frac{40-20}{4} \\ &= 5\end{aligned}$$

This time she is five standard deviations above the mean, which is even more unusual than her previous result.

8f $75\% = 0.75$; $60\% = 0.60$

$x_1 = 0.75 \times 48 = 36$

$x_2 = 0.60 \times 100 = 60$

$$\begin{aligned}z_1 &= \frac{x_1-\mu_1}{\sigma_1} \\ &= \frac{36-12}{3} \\ &= 8\end{aligned}$$

$$\begin{aligned}z_2 &= \frac{x_2-\mu_2}{\sigma_2} \\ &= \frac{60-20}{4}\end{aligned}$$

$= 10$

Idette's score of 75% in the first test is 8 standard deviations above the mean and her score of 60% in the second test is 10 standard deviations above the mean. As $z_2 > z_1$; her score on the second test is farther away from the mean than her first score. Hence, her second result is more unusual than her first result. Note, however, that both results are almost impossible to achieve just by guessing.

9 $X =$ number of patients who show improvement using the drug; $p = 0.7; q = 0.3;$

$n_\text{A} = 50;\ x_\text{A} = 45;\ n_\text{B} = 90;\ \ x_\text{B} = 74$

$$\begin{aligned}\mu_\text{A} &= n_\text{A}p\\ &= 50 \times 0.7\\ &= 35\end{aligned}$$

$$\begin{aligned}\mu_\text{B} &= n_\text{B}p\\ &= 90 \times 0.7\\ &= 63\end{aligned}$$

$$\begin{aligned}\sigma_\text{A} &= \sqrt{n_\text{A}pq}\\ &= \sqrt{50 \times 0.7 \times 0.3}\\ &= \sqrt{10.5}\\ &= 3.240\,37\ldots\\ &\doteqdot 3.2\end{aligned}$$

$$\begin{aligned}\sigma_\text{B} &= \sqrt{n_\text{B}pq}\\ &= \sqrt{90 \times 0.7 \times 0.3}\\ &= \sqrt{18.9}\\ &= 4.347\,41\\ &\doteqdot 4.3\end{aligned}$$

$$\begin{aligned}z_\text{A} &= \frac{x_\text{A} - \mu_\text{A}}{\sigma_\text{A}}\\ &= \frac{45 - 35}{3.240\,37\ldots}\end{aligned}$$

$= 3.086\ 06\ \ldots$

$\doteqdot 3.1$

$z_{\mathrm{B}} = \dfrac{x_{\mathrm{B}} - \mu_{\mathrm{B}}}{\sigma_{\mathrm{B}}}$

$= \dfrac{74 - 63}{4.347\ 41\ \ldots}$

$= 2.530\ 24\ \ldots$

$\doteqdot 2.5$

Team A's results are 3.1 standard deviations above the mean, compared to Team B's which are only 2.5 standard deviations above the mean. Hence, Team A's changes to the drug show stronger evidence for improvement.

10 $n = 100$; $X =$ number of people who voted for the WTP; $p = 0.15$; $q = 0.85$

10a $\mu = np$

$= 100 \times 0.15$

$= 15$

$\sigma = \sqrt{npq}$

$= \sqrt{100 \times 0.15 \times 0.85}$

$= \sqrt{12.75}$

$= 3.570\ 71\ \ldots$

$\doteqdot 3.57$

10b $\sigma \div 2 \doteqdot 1.79$

$x_1 = \mu - \dfrac{\sigma}{2} \doteqdot 15 - 1.79 = 13.21$

$x_2 = \mu + \dfrac{\sigma}{2} \doteqdot 15 + 1.79 = 16.79$

There are 14, 15 or 16 people voting for WTP within half a standard deviation.

$P(14 \le X \le 16)$

$= P(X = 14) + P(X = 15) + P(X = 16)$

$= {}^{100}\mathrm{C}_{14}\,(0.15)^{14}(0.85)^{86} + {}^{100}\mathrm{C}_{15}\,(0.15)^{15}(0.85)^{85} + {}^{100}\mathrm{C}_{16}\,(0.15)^{16}(0.85)^{84}$

$= 0.325\ 03\ \ldots$

$\doteqdot 32.5\%$

11a $\sigma^2 = npq$

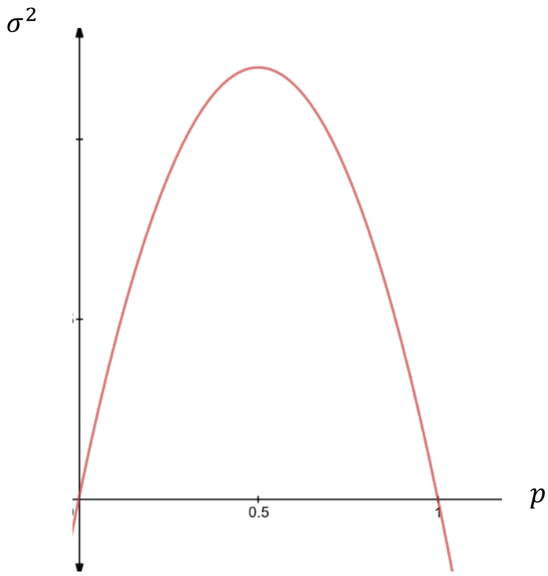
Since $q = 1 - p$,

$\sigma^2 = np(1-p)$

$= np - np^2$

$= -np^2 + np$

Sketch graph of $\sigma^2 = -np^2 + np$ shown below.



11b The graph of σ^2 is a parabola, symmetric in its axis of symmetry $p = \frac{1}{2}$.

11c As $0 \le p \le 1$ and $0 \le q \le 1$,

$npq < np$ and $npq < nq$ for $n \ge 1$.

Now $\sigma^2 = npq$,

hence $\sigma^2 < np$ and $\sigma^2 < nq$

11d From graph, maximum occurs (at vertex) at $p = \frac{1}{2}$.

Hence,

$$\sigma^2 = np(1-p)$$
$$= \frac{n}{2}\left(1-\frac{1}{2}\right)$$
$$= \frac{n}{4}$$

11e $\sigma^2 = np(1-p)$

As $\lim_{p\to 0^+} np(1-p) = n(0^+)(1-0^+) = 0$

And $\lim_{p\to 1^-} np(1-p) = n(1^-)(1-1^-) = 0$

Hence, as $p \to 0^+$ or $p \to 1^-$; $\sigma \to 0$

12a $n = 16;\ p = \frac{1}{2};\ q = \frac{1}{2}$

$$\sigma = \sqrt{npq}$$
$$= \sqrt{16 \times \frac{1}{2} \times \frac{1}{2}}$$
$$= \sqrt{4}$$
$$= 2 \text{ units}$$

There are 4 columns in the interval of one standard deviation or less from the mean (2 columns on each side of the mean).

12b $n = 36;\ p = \frac{1}{2};\ q = \frac{1}{2}$

$$\sigma = \sqrt{npq}$$
$$= \sqrt{36 \times \frac{1}{2} \times \frac{1}{2}}$$
$$= \sqrt{9}$$

$= 3$ units

There are 6 columns in the interval of one standard deviation or less from the mean.

12c $n = 64;\ p = \frac{1}{2};\ q = \frac{1}{2}$

$\sigma = \sqrt{npq}$

$= \sqrt{64 \times \frac{1}{2} \times \frac{1}{2}}$

$= \sqrt{16}$

$= 4$ units

There are 8 columns in the interval of one standard deviation or less from the mean.

13 Using the pattern observed in Q12,

$\sigma = \sqrt{npq} = \frac{1}{2}\sqrt{n}$, since $p = q = \frac{1}{2}$.

Thus an interval σ either side of the mean covers a width of $\sqrt{n}$.

Similarly if $p = \frac{1}{4}, q = \frac{3}{4}$ and $\sigma = \frac{1}{4}\sqrt{3n}$.

Thus an interval σ either side of the mean covers a width of $\frac{1}{2}\sqrt{3n}$.

The ratio of these widths is:

$\frac{\sqrt{n}}{\frac{1}{2}\sqrt{3n}}$

$= \frac{\sqrt{n}}{\frac{1}{2}\sqrt{3} \times \sqrt{n}}$

$= \frac{2}{\sqrt{3}}$

or $2 : \sqrt{3}$

14a $X =$ number of successful trials

Probability of obtaining at least one success

$= P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{n}C_{0}\, p^{0} q^{n}$

$= 1 - 1 \times 1 \times (1 - p)^{n}$ (since $q = 1 - p$)

$= 1 - (1 - p)^{n}$

14b $P(X \geq 1) > 0.95$

$1 - (1 - p)^{n} > 0.95$

$(1 - p)^{n} < 0.05$

$\log(1 - p)^{n} < \log(0.05)$

$n \times \log(1 - p) < \log(0.05)$

$$n > \frac{\log(0.05)}{\log(1 - p)}$$

Note, the inequality sign is reversed in the last step since $\log(1 - p)$ is negative as $1 - p < 1$.

So, the number of trials required to ensure that the probability of obtaining a success is at least 95% is:

$$n = \frac{\log(0.05)}{\log(1 - p)}$$

14c Using

$$n = \frac{\log(0.05)}{\log(1 - p)}$$

and rounding up to the nearest whole number produces the following table.

p	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05
n	2	3	4	5	6	9	14	29	59

14d If an experiment testing a certain result is repeated enough times, it is expected that the hypothesis will be upheld eventually. If 99 times it fails and once it succeeds, then only publishing the success gives a skewed picture of the truth.

15 $X =$ first trial producing a success

15a i $P(X = x)$ is the probability of exactly $(x - 1)$ failures and one success, in that order. We can multiply the probabilities of these independent events.

$$P(X = x) = q^{x-1} \times p$$

$$= pq^{x-1}$$

15a ii The mean is:

$$\mu = \sum x \times P(X = x) \qquad \text{(for } x \geq 1\text{)}$$

$$= 1 \times p + 2 \times pq + 3 \times pq^2 + 4 \times pq^3 + \cdots$$

15a iii Subtracting carefully in columns:

$$\mu - qu = p + 2pq + 3pq^2 + 4pq^3 + \cdots$$

$$-pq - 2pq^2 - 3pq^3 - \cdots$$

$$= p + pq + pq^2 + pq^3 + \cdots$$

Using the formula for the infinite sum of a geometric progression with $a = p$ and $r = q$ gives:

$$\mu - qu = \frac{p}{1 - q}$$

$$\mu - qu = \frac{p}{p} \qquad \text{(since } p = 1 - q\text{)}$$

$$\mu - qu = 1$$

$$\mu(1 - q) = 1$$

$$\mu = \frac{1}{1 - q}$$

$$\mu = \frac{1}{p}$$

15b The probability of being chosen is:

$$p = \frac{1}{20}$$

So the mean waiting time is:

$$\mu = \frac{1}{p}$$

$$= 20 \text{ time periods}$$

$$= 5 \times 20 \text{ minutes}$$

$$= 100 \text{ minutes}$$

Solutions to Exercise 17C

Note: the standard normal probability values are found using the table in Exercise 17C. Hence the answers for some questions may vary slightly depending on the accuracy used (table or other technology).

1a $n = 20;\ \ p = 0.3;\ \ q = 0.7$

Symbolic form is $B(20, 0.3)$

1b $P(X = 9, 10 \text{ or } 11)$

$= P(X = 9) + P(X = 10) + P(X = 11)$

$= {}^{20}\text{C}_9\,(0.3)^9(0.7)^{11} + {}^{20}\text{C}_{10}\,(0.3)^{10}(0.7)^{10} + {}^{20}\text{C}_{11}\,(0.3)^{11}(0.7)^9$

$= 0.108\,193\ldots$

$\doteqdot 0.1082$

$= 10.82\%$

1c $n = 20;\quad p = 0.3;\quad q = 0.7$

$np = 20 \times 0.3 = 6;\quad \text{so } np > 5$

$nq = 20 \times 0.7 = 14;\ \text{so } nq > 5$

1d $\mu = np = 6$

$\sigma^2 = npq = 6 \times 0.7 = 4.2$

$B(20, 0.3) \doteqdot N(6, 4.2)$

$P(8.5 \le X \le 11.5)$

$\doteqdot P\left(\dfrac{8.5 - 6}{\sqrt{4.2}} \le Z \le \dfrac{11.5 - 6}{\sqrt{4.2}}\right)$

$\doteqdot P(1.22 \le Z \le 2.68)$

$= P(Z \le 2.68) - P(Z \le 1.22)$

$= 0.9963 - 0.8888$ (using standard normal probability table in textbook)

$= 0.1075$

$= 10.75\%$

1e Percentage error using normal approximation

$= \dfrac{0.1082 - 0.1075}{0.1082} \times 100\%$

$= 0.646\,95\ldots\%$

$\doteqdot 0.65\%$

So percentage error is about 0.65%.

The result is very accurate.

2a $n = 50;\ \ p = 0.5;\ \ q = 0.5$

Symbolic form is $B(50, 0.5)$

$P(18 \le X \le 20)$

$= P(X = 18) + P(X = 19) + P(X = 20)$

$= {}^{50}\mathrm{C}_{18}\,(0.5)^{18}(0.5)^{32} + {}^{50}\mathrm{C}_{19}\,(0.5)^{19}(0.5)^{31} + {}^{50}\mathrm{C}_{20}\,(0.5)^{20}(0.5)^{30}$

$= {}^{50}\mathrm{C}_{18}\,(0.5)^{50} + {}^{50}\mathrm{C}_{19}\,(0.5)^{50} + {}^{50}\mathrm{C}_{20}\,(0.5)^{50}$

$= 0.084\,899\ldots$

$\doteqdot 0.0849$

$= 8.49\%$

$np = 50 \times 0.5 = 25$ so $np > 5$

$nq = 50 \times 0.5 = 25$ so $nq > 5$

$\mu = np = 25$

$\sigma^2 = npq = 25 \times 0.5 = 12.5$

$B(50, 0.5) \doteqdot N(25, 12.5)$

$P(18 \le X \le 20)$

$\doteqdot P\left(\dfrac{17.5 - 25}{\sqrt{12.5}} \le Z \le \dfrac{20.5 - 25}{\sqrt{12.5}}\right)$

$\doteqdot P(-2.12 \le Z \le -1.27)$

$= P(1.27 \le Z \le 2.12)$

$= P(Z \le 2.12) - P(Z \le 1.27)$

$= 0.9830 - 0.8980$

$= 0.0850$

$= 8.50\%$

Percentage error using normal approximation

$= \dfrac{0.0849 - 0.0850}{0.0849} \times 100\%$

$= -0.117\,785\ldots\%$

$\doteqdot -0.1\%$

So percentage error is about 0.1%.

2b $n = 20;\ p = 0.4;\ \ q = 0.6$

Symbolic form is $B(20, 0.4)$

$P(8 \le X \le 9)$

$= P(X = 8) + P(X = 9)$

$= {}^{20}\mathrm{C}_8\,(0.4)^8(0.6)^{12} + {}^{20}\mathrm{C}_9\,(0.4)^9(0.6)^{11}$

$= 0.339\,44\ldots$

$\doteqdot 0.3394$

$= 33.94\%$

$np = 20 \times 0.4 = 8$ so $np > 5$

$nq = 20 \times 0.6 = 12$ so $nq > 5$

$\mu = np = 8$

$\sigma^2 = npq = 8 \times 0.6 = 4.8$

$B(20, 0.4) \doteqdot N(8, 4.8)$

$P(8 \le X \le 9)$

$\doteqdot P\left(\frac{7.5-8}{\sqrt{4.8}} \le Z \le \frac{9.5-8}{\sqrt{4.8}}\right)$

$\doteqdot P(-0.23 \le Z \le 0.68)$

$= P(Z \le 0.68) - [1 - P(Z \le 0.23)]$

$= 0.7517 - (1 - 0.5910)$

$= 0.3427$

$= 34.27\%$

Percentage error using normal approximation

$= \frac{0.3394 - 0.3427}{0.3394} \times 100\%$

$= -0.9723 \ldots \%$

$\doteqdot -1\%$

So percentage error is about 1%.

2c $n = 30;\ p = 0.3;\ \ q = 0.7$

Symbolic form is $B(30, 0.3)$

$P(5 \le X \le 7)$

$= P(X = 5) + P(X = 6) + P(X = 7)$

$= {}^{30}\mathrm{C}_5\,(0.3)^5(0.7)^{25} + {}^{30}\mathrm{C}_6\,(0.3)^6(0.7)^{24} + {}^{30}\mathrm{C}_7\,(0.3)^7(0.7)^{23}$

$= 0.251\,221 \ldots$

$\doteqdot 0.2512$

$= 25.12\%$

$np = 30 \times 0.3 = 9$ so $np > 5$

$nq = 30 \times 0.7 = 21$ so $nq > 5$

$\mu = np = 9$

$\sigma^2 = npq = 9 \times 0.7 = 6.3$

$B(30, 0.3) \doteqdot N(9, 6.3)$

$P(5 \leq X \leq 7)$

$\doteqdot P\left(\frac{4.5-9}{\sqrt{6.3}} \leq Z \leq \frac{7.5-9}{\sqrt{6.3}}\right)$

$\doteqdot P(-1.79 \leq Z \leq -0.60)$

$= P(0.60 \leq Z \leq 1.79)$

$= P(Z \leq 1.79) - P(Z \leq 0.60)$

$= 0.9633 - 0.7257$

$= 0.2376$

$= 23.76\%$

Percentage error using normal approximation

$= \frac{0.2512 - 0.2376}{0.2512} \times 100\%$

$= 5.414 \ldots \%$

$\doteqdot 5\%$

So percentage error is about 5%.

2d $n = 40;\ p = 0.2;\ \ q = 0.8$

Symbolic form is $B(40, 0.2)$

$P(9 \leq X \leq 12)$

$= P(X = 9) + P(X = 10) + P(X = 11) + P(X = 12)$

$= {}^{40}\mathrm{C}_9\,(0.2)^9(0.7)^{31} + {}^{40}\mathrm{C}_{10}\,(0.2)^{10}(0.7)^{30} + {}^{40}\mathrm{C}_{11}\,(0.2)^{11}(0.7)^{29} + {}^{40}\mathrm{C}_{12}\,(0.2)^{12}(0.7)^{28}$

$= 0.363\,631 \ldots$

$\doteqdot 0.3636$

$= 36.36\%$

$np = 40 \times 0.2 = 8$ so $np > 5$

$nq = 40 \times 0.8 = 32$ so $nq > 5$

$\mu = np = 8$

$\sigma^2 = npq = 8 \times 0.8 = 6.4$

$B(40, 0.2) \doteqdot N(8, 6.4)$

$P(9 \leq X \leq 12)$

$\doteqdot P\left(\frac{8.5 - 8}{\sqrt{6.4}} \leq Z \leq \frac{12.5 - 8}{\sqrt{6.4}}\right)$

$\doteqdot P(0.20 \leq Z \leq 1.78)$

$= P(Z \leq 1.78) - P(Z \leq 0.20)$

$= 0.9625 - 0.5793$

$= 0.3832$

$= 38.32\%$

Percentage error using normal approximation

$= \frac{0.3636 - 0.3832}{0.3636} \times 100\%$

$= -5.3905 \ldots \%$

$\doteqdot -5\%$

So percentage error is about 5%.

2e $n = 22;\ p = 0.6;\ \ q = 0.4$

Symbolic form is $B(22, 0.6)$

$P(13 \leq X \leq 15)$

$= P(X = 13) + P(X = 14) + P(X = 15)$

$= {}^{22}\mathrm{C}_{13}\,(0.6)^{13}(0.4)^9 + {}^{22}\mathrm{C}_{14}\,(0.6)^{14}(0.4)^8 + {}^{22}\mathrm{C}_{15}\,(0.6)^{15}(0.4)^7$

$= 0.465\,907 \ldots$

$\doteqdot 0.4659$

$= 46.59\%$

$np = 22 \times 0.6 = 13.2$ so $np > 5$

$nq = 22 \times 0.4 = 8.8$ so $nq > 5$

$\mu = np = 13.2$

$\sigma^2 = npq = 13.2 \times 0.4 = 5.28$

$B(22, 0.6) \doteqdot N(13.2, 5.28)$

$P(13 \leq X \leq 15)$

$\doteqdot P\left(\frac{12.5 - 13.2}{\sqrt{5.28}} \leq Z \leq \frac{15.5 - 13.2}{\sqrt{5.28}}\right)$

$\doteqdot P(-0.30 \leq Z \leq 1.00)$

$= P(Z \leq 1.00) - [1 - P(Z \leq 0.30)]$

$= 0.8413 - (1 - 0.6179)$

$= 0.4592$

$= 45.92\%$

Percentage error using normal approximation

$= \frac{0.4659 - 0.4592}{0.4659} \times 100\%$

$= 1.438 \ldots \%$

$\doteqdot 1.4\%$

So percentage error is about 1%.

2f $n = 80;\ p = 0.1;\ \ q = 0.9$

Symbolic form is $B(80, 0.1)$

$P(10 \leq X \leq 13)$

$= P(X = 10) + P(X = 11) + P(X = 12) + P(X = 13)$

$= {}^{80}C_{10}\,(0.1)^{10}(0.9)^{70} + {}^{80}C_{11}\,(0.1)^{11}(0.9)^{69} + {}^{80}C_{12}\,(0.1)^{12}(0.9)^{68} + {}^{80}C_{13}\,(0.1)^{13}(0.9)^{67}$

$= 0.249\,801 \ldots$

$\doteqdot 0.2498$

$= 24.98\%$

$np = 80 \times 0.1 = 8$ so $np > 5$

$nq = 80 \times 0.9 = 72$ so $nq > 5$

$\mu = np = 8$

$\sigma^2 = npq = 8 \times 0.9 = 7.2$

$B(80, 0.1) \doteqdot N(8, 7.2)$

$P(10 \leq X \leq 13)$

$\doteqdot P\left(\frac{9.5 - 8}{\sqrt{7.2}} \leq Z \leq \frac{13.5 - 8}{\sqrt{7.2}}\right)$

$\doteqdot P(0.56 \leq Z \leq 2.05)$

$= P(Z \leq 2.05) - P(Z \leq 0.56)$

$= 0.9798 - 0.7123$

$= 0.2675$

$= 26.75\%$

Percentage error using normal approximation

$= \frac{0.2498 - 0.2675}{0.2498} \times 100\%$

$= -7.0856 \ldots \%$

$\doteqdot -7\%$

So percentage error is about 7%.

2g $n = 500;\ p = 0.25;\ q = 0.75$

Symbolic form is $B(500, 0.25)$

$P(100 \leq X \leq 103)$

$= P(X = 100) + P(X = 101) + P(X = 102) + P(X = 103)$

$$= {}^{500}\mathrm{C}_{100}\,(0.25)^{100}(0.75)^{400} + {}^{500}\mathrm{C}_{101}\,(0.25)^{101}(0.75)^{399} + {}^{500}\mathrm{C}_{102}\,(0.25)^{102}(0.75)^{398} + {}^{500}\mathrm{C}_{103}\,(0.25)^{103}(0.75)^{397}$$

$= 0.008\,413\ldots$

$\doteqdot 0.0084$

$= 0.84\%$

$np = 500 \times 0.25 = 125$ so $np > 5$

$nq = 500 \times 0.75 = 375$ so $nq > 5$

$\mu = np = 125$

$\sigma^2 = npq = 125 \times 0.75 = 93.75$

$B(500, 0.25) \doteqdot N(125, 93.75)$

$P(100 \le X \le 103)$

$\doteqdot P\left(\dfrac{99.5 - 125}{\sqrt{93.75}} \le Z \le \dfrac{103.5 - 125}{\sqrt{93.75}}\right)$

$\doteqdot P(-2.63 \le Z \le -2.22)$

$= P(2.22 \le Z \le 2.63)$

$= P(Z \le 2.63) - P(Z \le 2.22)$

$= 0.9957 - 0.9868$

$= 0.0089$

$= 0.89\%$

Percentage error using normal approximation

$= \dfrac{0.0084 - 0.0089}{0.0084} \times 100\%$

$= -5.952\,38\ldots\%$

$\doteqdot -6\%$

So percentage error is about 6%.

2h $n = 200;\ p = 0.9;\ \ q = 0.1$

Symbolic form is $B(200, 0.9)$

$P(170 \le X \le 172)$

$= P(X = 170) + P(X = 171) + P(X = 172)$

$= {}^{200}\mathrm{C}_{170}\,(0.9)^{170}(0.1)^{30} + {}^{200}\mathrm{C}_{171}\,(0.9)^{171}(0.1)^{29} + {}^{200}\mathrm{C}_{172}\,(0.9)^{172}(0.1)^{28}$

$= 0.033\,920\,\ldots$

$\doteqdot 0.0339$

$= 3.39\%$

$np = 200 \times 0.9 = 180$ so $np > 5$

$nq = 200 \times 0.1 = 20$ so $nq > 5$

$\mu = np = 180$

$\sigma^2 = npq = 180 \times 0.1 = 18$

$B(200, 0.9) \doteqdot N(180, 18)$

$P(170 \le X \le 172)$

$\doteqdot P\left(\dfrac{169.5 - 180}{\sqrt{18}} \le Z \le \dfrac{172.5 - 180}{\sqrt{18}}\right)$

$\doteqdot P(-2.47 \le Z \le -1.77)$

$= P(1.77 \le Z \le 2.47)$

$= P(Z \le 2.47) - P(Z \le 1.77)$

$= 0.9932 - 0.9616$

$= 0.0316$

$= 3.16\%$

Percentage error using normal approximation

$= \dfrac{0.0339 - 0.0316}{0.0339} \times 100\%$

$= 6.784\,66\ldots\%$

$\doteqdot 7\%$

So percentage error is about 7%.

3a There are only two possible outcomes – pink or blue. Also, for each stage of the experiment, the total number of counters remain the same and the probability of choosing a blue or a pink counter remains the same. Before each stage, the counters are stirred properly. Hence, at each stage, the probability of selecting a counter is independent of its predecessor. So, each stage of the process is a Bernoulli trial.

3b There are n stages. Each stage is independent, and each stage has the same probability of success.

3c Yes. If the counter is not returned, the stages of the experiment will not be independent. With the large number of counters, however, the probability will not change much, and we could approximate the experiment as binomial.

3d $n = 20; X =$ number of pink counters selected

$$p_{\text{pink}} = \frac{600}{600 + 400} = 0.6$$

$$\begin{aligned}\mu &= np \\ &= 20 \times 0.6 \\ &= 12\end{aligned}$$

$$\begin{aligned}\sigma &= \sqrt{npq} \\ &= \sqrt{12 \times 0.4} \\ &= \sqrt{4.8} \\ &= 2.190\,89\ldots \\ &\doteqdot 2.19\end{aligned}$$

3e $P(X = 14)$

$= {}^{20}\mathrm{C}_{14}\,(0.6)^{14}(0.4)^{6}$

$= 0.124\,411\ldots$

$\doteqdot 0.12$

3f i Considering the diagram of the histogram and the overlayed normal distribution curve, we can observe that the curve is a continuous function and at point 14, it is just a value and not an area whereas the histogram, being a discrete figure, has a width of 1 unit. So, the area under the histogram denoting the probability is $p \times 1 = p$. To find the probability for 14 using the normal curve, we need to use the same width as that of the histogram. Hence, we use the interval 13.5 to 14.5 so that the average is 14, width is 1 and we get the approximately correct value.

3f ii $P(13.5 < X < 14.5)$

$\doteqdot P\left(\frac{13.5 - 12}{2.19} < Z < \frac{14.5 - 12}{2.19}\right)$

$\doteqdot P(0.68 < Z < 1.14)$

$= P(Z < 1.14) - P(Z < 0.68)$

$= 0.8729 - 0.7517$

$= 0.1212$

$\doteqdot 0.12$

To two significant figures, both values are 0.12 or 12%.

4a $N = 3000$; $G = 1320$; $X =$ number of girls in sample

$p_{\text{girl}} = \frac{1320}{3000} = 0.44$

4b $n = 15$; $p = 0.44$; $q = 0.56$

$\mu = np$

$= 15 \times 0.44$

$= 6.6$

$\sigma = \sqrt{npq}$

$= \sqrt{6.6 \times 0.56}$

$= \sqrt{3.696}$

$= 1.922\ 49 \ldots$

$\doteqdot 1.92$

4c $P(X = 9 \text{ or } 10)$

$= P(X = 9) + P(X = 10)$

$= {}^{15}\mathrm{C}_9\,(0.44)^9(0.56)^6 + {}^{15}\mathrm{C}_{10}\,(0.44)^{10}(0.56)^5$

$= 0.140\ 393 \ldots$

$\doteqdot 0.14$

$= 14\%$

4d $np = 15 \times 0.44 = 6.6 > 5$

$nq = 15 \times 0.56 = 8.4 > 5$

Hence, normal approximation can be used.

4e $P(8.5 \le X \le 10.5)$

$\doteqdot P\left(\dfrac{8.5 - 6.6}{\sqrt{3.696}} \le Z \le \dfrac{10.5 - 6.6}{\sqrt{3.696}}\right)$

$\doteqdot P(0.99 \le Z \le 2.03)$

$= P(Z \le 2.03) - P(Z \le 0.99)$

$= 0.9788 - 0.8389$

$= 0.1399$

$\doteqdot 0.14$

$= 14\%$

4f Percentage error using normal approximation

$= \dfrac{0.1404 - 0.1399}{0.1404} \times 100\%$

$= 0.356 \ldots \%$

$\doteqdot 0.36\%$

So percentage error is less than 1%.

5 $np > 5$; $n(1-p) > 5$ or $nq > 5$

$n > \dfrac{5}{p}$ and $n > \dfrac{5}{q}$

5a $p = 0.5$; $q = 1 - p = 0.5$

$n > \dfrac{5}{p}$ and $n > \dfrac{5}{q}$

$n > \dfrac{5}{0.5}$ and $n > \dfrac{5}{0.5}$

$n > 10$ and $n > 10$

Hence, $n > 10$.

5b $p = 0.25$; $q = 1 - p = 0.75$

$n > \dfrac{5}{p}$ and $n > \dfrac{5}{q}$

$n > \dfrac{5}{0.25}$ and $n > \dfrac{5}{0.75}$

$n > 20$ and $n > 6.66 \ldots$

Hence, $n > 20$.

5c $p = 0.125$; $q = 1 - p = 0.875$

$n > \dfrac{5}{p}$ and $n > \dfrac{5}{q}$

$n > \dfrac{5}{0.125}$ and $n > \dfrac{5}{0.875}$

$n > 40$ and $n > 5.71 \ldots$

Hence, $n > 40$.

5d $p = 0.01; \quad q = 1 - p = 0.99$

$n > \frac{5}{p}$ and $n > \frac{5}{q}$

$n > \frac{5}{0.01}$ and $n > \frac{5}{0.99}$

$n > 500$ and $n > 5.05 \ldots$

Hence, $n > 500$.

5e $p = 0.75; \quad q = 1 - p = 0.25$

$n > \frac{5}{p}$ and $n > \frac{5}{q}$

$n > \frac{5}{0.75}$ and $n > \frac{5}{0.25}$

$n > 6.66 \ldots$ and $n > 20$

Hence, $n > 20$.

5f $p = 0.875; \quad q = 1 - p = 0.125$

$n > \frac{5}{p}$ and $n > \frac{5}{q}$

$n > \frac{5}{0.875}$ and $n > \frac{5}{0.125}$

$n > 5.71 \ldots$ and $n > 40$

Hence, $n > 40$.

5g $p = 0.9; \quad q = 1 - p = 0.1$

$n > \frac{5}{p}$ and $n > \frac{5}{q}$

$n > \frac{5}{0.9}$ and $n > \frac{5}{0.1}$

$n > 5.55\ldots$ and $n > 50$

Hence, $n > 50$.

5h $p = 0.55;\ \ q = 1 - p = 0.45$

$n > \frac{5}{p}$ and $n > \frac{5}{q}$

$n > \frac{5}{0.55}$ and $n > \frac{5}{0.45}$

$n > 9.09\ldots$ and $n > 11.11\ldots$

Hence, $n > 11$.

6a As the number of trials is very large, $n = 854$, the normal distribution will approximate the binomial distribution to a greater accuracy. Also, it is easier to compute the probabilities using normal distribution than binomial distribution. Such high computations are beyond the scope of simple calculators.

6b It mostly depends on the method of sampling used. For being a representative sample, it must have people from all over the world evenly distributed and there should not be any sampling bias involved to ensure randomness.

It is hard to get a representative sample of the whole world, because different ethnic groups will have different tendencies to colour blindness.

6c $n = 854;\ X =$ number of people in the sample who are colour blind;

$p = 0.08;\ \ q = 0.92$

7% of 854

$= 0.07\ \times 854$

$= 59.78$

9% of 854

$= 0.09\ \times 854$

$= 76.86$

So $60 < X < 76$

$\mu = np$

$= 854 \times 0.08$

$= 68.32$

$\sigma = \sqrt{npq}$

$= \sqrt{68.32 \times 0.92}$

$= \sqrt{62.8544}$

$= 7.928\,07 \ldots$

$\doteqdot 7.93$

$P(60 < X < 76)$

$\doteqdot P\left(\dfrac{60 - 68.32}{\sqrt{62.8544}} < Z < \dfrac{76 - 68.32}{\sqrt{62.8544}}\right)$

$\doteqdot P(-1.05 < Z < 0.97)$

$= P(Z < 0.97) - [1 - P(Z < 1.05)]$

$= 0.8340 - (1 - 0.8531)$

$= 0.6871$

$\doteqdot 68.7\%$

6d $P(76 < X < 76.5)$

$\doteqdot P\left(\dfrac{76 - 68.32}{\sqrt{62.8544}} < Z < \dfrac{76.5 - 68.32}{\sqrt{62.8544}}\right)$

$\doteqdot P(0.97 < Z < 1.03)$

$= P(Z < 1.03) - P(Z < 0.97)$

$= 0.8485 - 0.8340$

$= 0.0145$

$\doteqdot 1.5\%$

7 $n = 15$; X = number of offspring with red flowers; $p = 0.25$; $q = 0.75$

7a $P(X = 0)$

$= {}^{15}C_0\ (0.25)^0(0.75)^{15}$

$= (0.75)^{15}$

$= 0.013\ 36\ \ldots$

$\doteqdot 1\%$

7b $P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - (0.75)^{15}$

$= 0.986\ 636\ \ldots$

$\doteqdot 99\%$

7c 20% of $15 = 0.2 \times 15 = 3$

$\mu = np$

$= 15 \times 0.25$

$= 3.75$

$\sigma = \sqrt{npq}$

$= \sqrt{3.75 \times 0.75}$

$= \sqrt{2.8125}$

$= 1.677\ 05\ \ldots$

$\doteqdot 1.68$

Using a normal approximation:

$P(X > 3)$

$\doteqdot P\left(Z > \dfrac{2.5 - 3.75}{\sqrt{2.8125}}\right)$

$\doteqdot P(Z > -0.75)$

$= 1 - [1 - P(Z < 0.75)]$

$= 1 - (1 - 0.7734)$

$= 0.7734$

$\doteqdot 77\%$

8 $n = 20$; $X =$ number of eighteen-year-olds without a driver's licence;

$p = 0.45$; $q = 0.55$

$\mu = np$

$= 20 \times 0.45$

$= 9$

$\sigma = \sqrt{npq}$

$= \sqrt{9 \times 0.55}$

$= \sqrt{4.95}$

$= 2.224\,85\ldots$

$\doteqdot 2.225$

Looking for $P(X > 10) = P(X \geq 11)$.

Using a normal approximation:

$P(X \geq 10.5)$

$\doteqdot P\left(Z \geq \dfrac{10.5 - 9}{\sqrt{4.95}}\right)$

$\doteqdot P(Z \geq 0.67)$

$= 1 - 0.7486$

$= 0.2514$

$\doteqdot 25\%$

9 $n = 30$; $X =$ number of people who prefer Country music;

$p = 0.6$; $q = 0.4$

$\mu = np$

$= 30 \times 0.6$

$= 18$

$\sigma = \sqrt{npq}$

$= \sqrt{18 \times 0.4}$

$= \sqrt{7.2}$

$= 2.683\ 28 \ldots$

$\doteqdot 2.68$

Looking for $P(X > 20)$.

Using a normal approximation:

$P(X \geq 21)$

$\doteqdot P\left(Z \geq \dfrac{20.5 - 18}{\sqrt{7.2}}\right)$

$\doteqdot P(Z \geq 0.93)$

$= \ 1 - 0.8238$

$= 0.1762$

$\doteqdot 17.6\%$

The underlying Bernoulli distribution is not applied with replacement, because the same person will not be in the park twice at the same gathering. If the population of Nashville is large; it should be reasonable to neglect this fact. It is also assumed that the visitors to the park are a random cross-section of Nashville. Groups with similar tastes may arrive together.

10a Spreadsheet

10b There are still 100 trials, but the basic Bernoulli trial has changed. It could be that an extremely biased coin is tossed, or a card labelled 1 is selected (with replacement) from a pack of cards labelled 1-10.

10c The graphs are bell-shaped curves. Smaller probabilities give a curve centred to the left (skewed to the right), and larger probabilities give a curve centred to the right (skewed to the left). Probabilities further from 0.5 give a narrower curve (distribution).

11 Spreadsheet investigation

12a i $n = 10$; $X =$ number of punnets rejected; $p = 0.05$; $q = 0.95$

$P(\text{batch will be rejected})$

$= P(\text{at least one punnet is rejected})$

$= P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}C_0\,(0.05)^0(0.95)^{10}$

$= 1 - (0.95)^{10}$

$= 0.401\,26\ldots$

$\doteqdot 0.4$

12a ii For $p = 0.1$:

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}C_0\,(0.1)^0(0.9)^{10}$

$= 1 - (0.9)^{10}$

$= 0.651\,32\ldots$

$\doteqdot 0.65$

For $p = 0.2$:

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}C_0\,(0.2)^0(0.8)^{10}$

$= 1 - (0.8)^{10}$

$= 0.892\,62\ldots$

$\doteqdot 0.89$

For $p = 0.3$:

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}C_0\,(0.3)^0(0.7)^{10}$

$= 1 - (0.7)^{10}$

$= 0.971\ 75 \ldots$

$\doteqdot 0.97$

For $p = 0.4$:

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{10}C_0\ (0.4)^0(0.6)^{10}$

$= 1 - (0.6)^{10}$

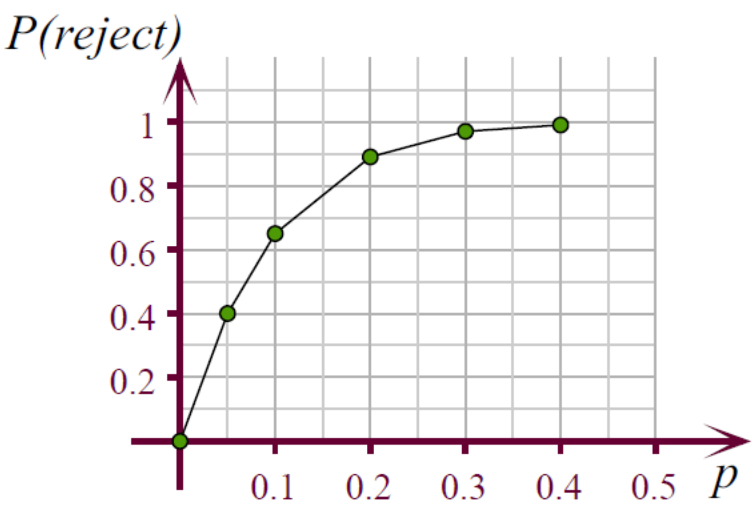
$= 0.993\ 95 \ldots$

$\doteqdot 0.99$

This can be summarised in the following table.

p	0	0.05	0.1	0.2	0.3	0.4
P(reject)	0	0.4	0.65	0.89	0.97	0.99

12a iii



12b $n = 15$; $X =$ number of punnets rejected

For $p = 0.05$; $q = 0.95$:

$P(\text{batch will be rejected})$

$= P(\text{at least two punnets are rejected})$

$= P(X \geq 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - {}^{15}C_0\,(0.05)^0(0.95)^{15} - {}^{15}C_1\,(0.05)^1(0.95)^{14}$

$= 1 - (0.95)^{15} - 15(0.05)(0.95)^{14}$

$= 0.170\,95\ldots$

$\doteqdot 0.17$

For $p = 0.1; q = 0.9$:

$= P(X \geq 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - {}^{15}C_0\,(0.1)^0(0.9)^{15} - {}^{15}C_1\,(0.1)^1(0.9)^{14}$

$= 1 - (0.9)^{15} - 15(0.1)(0.9)^{14}$

$= 0.450\,95\ldots$

$\doteqdot 0.45$

For $p = 0.2$; $q = 0.8$:

$= P(X \geq 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - {}^{15}C_0\,(0.2)^0(0.8)^{15} - {}^{15}C_1\,(0.2)^1(0.8)^{14}$

$= 1 - (0.8)^{15} - 15(0.2)(0.8)^{14}$

$= 0.832\,87\ldots$

$\doteqdot 0.83$

For $p = 0.3; q = 0.7$:

$= P(X \geq 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - {}^{15}C_0\,(0.3)^0(0.7)^{15} - {}^{15}C_1\,(0.3)^1(0.7)^{14}$

$= 1 - (0.7)^{15} - 15(0.3)(0.7)^{14}$

$= 0.964\,73\ldots$

$\doteqdot 0.96$

For $p = 0.4$; $q = 0.6$:

$= P(X \geq 2)$

$= 1 - [P(X = 0) + P(X = 1)]$

$= 1 - {}^{15}C_0\,(0.4)^0(0.6)^{15} - {}^{15}C_1\,(0.4)^1(0.6)^{14}$

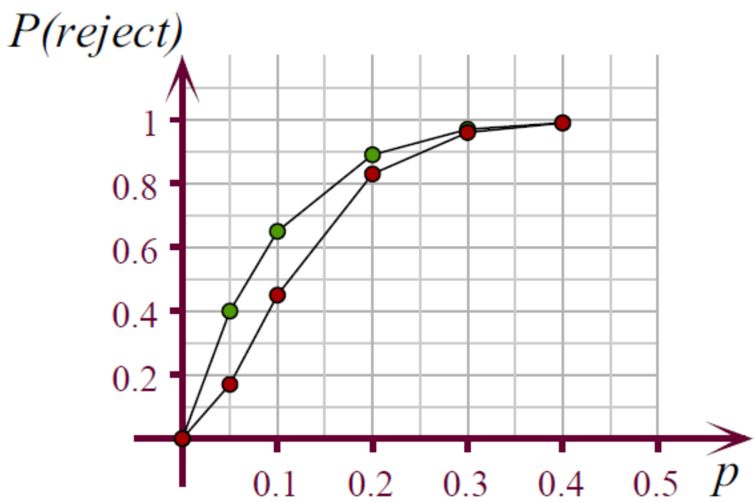
$= 1 - (0.6)^{15} - 15(0.4)(0.6)^{14}$

$= 0.994\,82\ldots$

$\doteqdot 0.99$

This can be summarised in the following table.

p	0	0.05	0.1	0.2	0.3	0.4
$P(\text{reject})$	0	0.17	0.45	0.83	0.96	0.99



12c The second method is more forgiving if there are a few punnets that need to be rejected. Both methods are strongly likely to reject the batch if p is high, indeed the curves approach one another closely by the time p reaches 20%. Which method to apply depends on other considerations, such as how forgiving the customers are, or whether distributers and shops are happy to throw out defective stock before it reaches the shelves.

Solutions to Exercise 17D

1a

$$\hat{p} = \frac{n(H)}{n(S)} = \frac{2}{5}$$

1b

$$\hat{p} = \frac{n(\text{Spades})}{n(\text{S})} = \frac{4}{10} = \frac{2}{5}$$

1c

$$\hat{p} = \frac{n(P)}{n(S)} = \frac{9}{12} = \frac{3}{4}$$

2a $n = 5$; $X = \text{number of heads}$; $p = \frac{1}{2}; q = \frac{1}{2}$

$P(X = x)$

$= {}^5\mathrm{C}_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x}$

$= {}^5\mathrm{C}_x \left(\frac{1}{2}\right)^5$

$= {}^5\mathrm{C}_x \times \frac{1}{32}$

$P(X = 0) = {}^5\mathrm{C}_0 \left(\frac{1}{2}\right)^5 = 1 \times \frac{1}{32} = \frac{1}{32}$

$P(X = 1) = {}^5\mathrm{C}_1 \left(\frac{1}{2}\right)^5 = 5 \times \frac{1}{32} = \frac{5}{32}$

$P(X = 2) = {}^5\mathrm{C}_2 \left(\frac{1}{2}\right)^5 = 10 \times \frac{1}{32} = \frac{10}{32}$

$P(X = 3) = {}^5\mathrm{C}_3 \left(\frac{1}{2}\right)^5 = 10 \times \frac{1}{32} = \frac{10}{32}$

$P(X = 4) = {}^5\mathrm{C}_4 \left(\frac{1}{2}\right)^5 = 5 \times \frac{1}{32} = \frac{5}{32}$

$$P(X = 5) = {}^5C_5\left(\frac{1}{2}\right)^5 = 1 \times \frac{1}{32} = \frac{1}{32}$$

The results are summarised in the following table.

x	0	1	2	3	4	5
$P(X = x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

2b The results from part a are summarised in the following table.

$\hat{p}$	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	1
$P(\hat{p})$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

2c $\mu_{\hat{p}}$

$$= 0 \times \frac{1}{32} + \frac{1}{5} \times \frac{5}{32} + \frac{2}{5} \times \frac{10}{32} + \frac{3}{5} \times \frac{10}{32} + \frac{4}{5} \times \frac{5}{32} + 1 \times \frac{1}{32}$$

$$= \frac{16}{32}$$

$$= 0.5$$

2d It is the probability p in each Bernoulli trial, that is, it is the probability of a coin landing heads.

3a $n = 5$; $X =$ number of yes answers in a week; 20 weeks

$\hat{p}$	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	1
f_r	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{2}{10}$	$\frac{6}{20}$	$\frac{7}{20}$

3b $\mu_{\hat{p}}$

$$= 0 \times \frac{1}{20} + \frac{1}{5} \times \frac{1}{20} + \frac{2}{5} \times \frac{3}{20} + \frac{3}{5} \times \frac{2}{20} + \frac{4}{5} \times \frac{6}{20} + \frac{5}{5} \times \frac{7}{20}$$

$= \dfrac{72}{100}$

$= 0.72$

3c The mean is an estimate of the probability chance that a shopper chosen at random lives in the suburb.

4a 75% of 10 $= 0.75 \times 10 = 7.5$

For more than 75% of the coins to show heads, this would be 8, 9 or 10 coins.

4b $n = 10$; $X =$ number of times the coin lands heads; $p = 0.5$; $q = 0.5$

$P(X > 7.5)$

$= P(X = 8) + P(X = 9) + P(X = 10)$

$= {}^{10}C_8 \times (0.5)^8 \times (0.5)^2 + {}^{10}C_9 \times (0.5)^9 \times (0.5)^1 + {}^{10}C_{10} \times (0.5)^{10} \times (0.5)^0$

$= 45 \times (0.5)^{10} + 10 \times (0.5)^{10} + 1 \times (0.5)^{10}$

$= 0.054\,687\ldots$

$\doteqdot 5.47\%$

5a $n = 50$; $X =$ number of sixes rolled; $p = \frac{1}{6}$; $q = \frac{5}{6}$

9% of 50 $= 0.09 \times 50 = 4.5$

$P(X < 4.5)$

$= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$= {}^{50}C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{50} + {}^{50}C_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{49} + {}^{50}C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{48} + {}^{50}C_3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{47} + {}^{50}C_4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^{46}$

$= 0.064\,31\ldots$

$\doteqdot 6.4\%$

5b For sample proportion:

$$\mu_{\hat{p}} = \frac{1}{6} = 0.166\,66\ldots \doteqdot 0.17$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{\frac{1}{6} \times \frac{5}{6}}{50}}$$

$$= \sqrt{\frac{1}{360}}$$

$$= \frac{1}{\sqrt{360}}$$

$$= 0.052\,70\ldots$$

$$P(\hat{p} < 0.09)$$

$$\doteqdot P\left(Z < \frac{0.09 - 0.166\,66\ldots}{0.052\,70\ldots}\right)$$

$$\doteqdot P(Z < -1.45)$$

$$= P(Z > 1.45)$$

$$= 1 - P(Z < 1.45)$$

$$= 1 - 0.9265$$

$$= 0.0735$$

$$\doteqdot 7.4\%$$

6a Number of students who bought the lunch from canteen $= 12$

Population proportion

$$= \frac{12}{32}$$

$$= \frac{3}{8}$$

$$= 0.375$$

6b Only one of the five students (Dakarai) buys lunch regularly.

Sample proportion

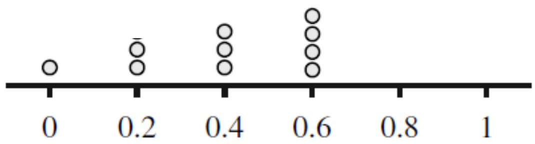
$= \dfrac{1}{5}$

$= 0.2$

6c

$\hat{p}$	0	0.2	0.4	0.6	0.8	1.0
Tally	|	||	|||	||||		
Frequency	1	2	3	4	0	0

6d



7 $n = 500$; $X =$ number of people who voted independent; $p = 0.2; q = 0.8$

22% of $500 = 0.22 \times 500 = 110$

$\mu = np$

$= 500 \times 0.2$

$= 100$

$\sigma = \sqrt{npq}$

$= \sqrt{100 \times 0.8}$

$= \sqrt{80}$

$\doteqdot 8.94$

$P(X > 110)$

$= P(X \geq 111)$

$\doteqdot P\left(Z \geq \dfrac{110.5 - 100}{\sqrt{80}}\right)$

$\doteqdot P(Z \geq 1.17)$

$= 1 - 0.8790$

$= 0.1210$

$\doteqdot 12\%$

8 $n = 80$; $X =$ number of hearts selected; $p = \frac{13}{52} = 0.25; q = 0.75$

$\mu = np$

$= 80 \times 0.25$

$= 20$

$\sigma = \sqrt{npq}$

$= \sqrt{20 \times 0.75}$

$= \sqrt{15}$

$\doteqdot 3.87$

20% of $80 = 16$

30% of $80 = 24$

8a $P(16 \leq X \leq 24)$

$\doteqdot P\left(\frac{15.5 - 20}{\sqrt{15}} \leq Z \leq \frac{24.5 - 20}{\sqrt{15}}\right)$

$\doteqdot P(-1.16 \leq Z \leq 1.16)$

$= P(Z \leq 1.16) - [1 - P(Z \leq 1.16)]$

$= 0.8770 - (1 - 0.8770)$

$= 0.7540$

$\doteqdot 75\%$

8b For sample proportion:

$\mu_{\hat{p}} = 0.25$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

$= \sqrt{\frac{0.25 \times 0.75}{80}}$

$= \sqrt{\frac{3}{1280}}$

$= 0.048\ 41 \ldots$

$\doteqdot 0.0484$

$P(0.20 \le \hat{p} \le 0.30)$

$\doteqdot P\left(\frac{0.20 - 0.25}{0.048\ 41 \ldots} \le Z \le \frac{0.30 - 0.25}{0.048\ 41 \ldots}\right)$

$\doteqdot P(-1.03 \le Z \le 1.03)$

$= P(Z < 1.03) - P(Z < -1.03)$

$= P(Z < 1.03) - [1 - P(Z < 1.03)]$

$= 0.8485 - (1 - 0.8485)$

$= 0.697$

$\doteqdot 70\%$

9 $n = 300$; $X =$ number of seeds that germinate; $p = \mu_{\hat{p}} = 0.7$; $q = 0.3$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

$= \sqrt{\frac{0.7 \times 0.3}{300}}$

$= \sqrt{\frac{7}{10\ 000}}$

$= \frac{\sqrt{7}}{100}$

$= 0.026\ 45\ \ldots$

$\doteqdot 0.0265$

9a $\quad$ 65% of $300 = 0.65 \times 300 = 195$

$P(X > 195)$

$= P(\hat{p} > 0.65)$

$\doteqdot P\left(Z > \dfrac{0.65 - 0.7}{0.026\ 45\ \ldots}\right)$

$\doteqdot P(Z > -1.89)$

$= P(Z < 1.89)$

$= 0.9706$

$\doteqdot 97\%$

9b $\quad$ 75% of $300 = 0.75 \times 300 = 225$

$P(195 < X < 225)$

$= P(0.65 < \hat{p} < 0.75)$

$\doteqdot P\left(\dfrac{0.65 - 0.7}{0.026\ 45\ \ldots} < Z < \dfrac{0.75 - 0.7}{0.026\ 45\ \ldots}\right)$

$\doteqdot P(-1.89 < Z < 1.89)$

$= P(Z < 1.89) - P(Z < -1.89)$

$= P(Z < 1.89) - [1 - P(Z < 1.89)]$

$= 0.9706 - (1 - 0.9706)$

$= 0.9706 - 0.0294$

$= 0.9412$

$\doteqdot 94\%$

10a $\quad$ $n = 100$; $X =$ number of people who have painful reaction to medication;

$p = 0.05$; $q = 0.95$

10a i $P(X = 0)$

$= {}^{100}\text{C}_0\ (0.05)^0(0.95)^{100}$

$= (0.95)^{100}$

$= 0.005\ 92\ \ldots$

$\doteqdot\ 0.6\%$

10a ii 2% of $100 = 0.2 \times 100 = 2$

$P(X < 2)$

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

$= (0.95)^{100} +\ {}^{100}\text{C}_1\ (0.05)^1(0.95)^{99}$

$=\ 0.037\ 08\ \ldots$

$\doteqdot\ 3.7\%$

10b $n = 1000; X =$ number of people who have painful reaction to medication;

For sample proportion:

$\mu_{\hat{p}} = p = 0.05;\ q = 0.95$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{0.05 \times 0.95}{1000}}$$

$= \sqrt{0.000\ 047\ 5}$

$= 0.006\ 89\ \ldots$

$\doteqdot 0.0069$

$P(X < 30)$

$= P(\hat{p} < 0.03)$

$$\doteqdot P\left(Z < \frac{0.03 - 0.05}{\sqrt{0.000\ 047\ 5}}\right)$$

$\doteqdot P(Z < -2.90)$

$= P(Z > 2.90)$

$= 1 - P(Z < 2.90)$

$= 1 - 0.9981$

$= 0.0019$

$\doteqdot 0.2\%$

10b ii The result of this test is significantly different from the previous claim that 5% of patients will have a reaction. They should check whether the sample was random – perhaps it consisted of patients more resistant to the side effects of the medication. They should also check whether there have been any changes to the medication to reduce patient reactions. It is also possibly just chance that this result occurred, but the likelihood of this is small.

11 $n = 20$; $X =$ number of pink counters selected; $p = 0.6$; $q = 0.4$

11a $\mu_{\hat{p}} = p = 0.6$

$${\sigma_{\hat{p}}}^2 = \frac{pq}{n}$$

$$= \frac{0.6 \times 0.4}{20}$$

$$= 0.012$$

$\sigma_{\hat{p}} = \sqrt{0.012}$

$= 0.109\,54\ldots$

$\doteqdot 0.1095$

11b $\hat{p}$ is the distribution of the binomial random variable divided by the number of trials. It has the similar properties to a binomial random variable and hence is not a continuous distribution.

11c $P(\hat{p} \le 0.4)$

$$\doteqdot P\left(Z < \frac{0.4 - 0.6}{\sqrt{0.012}}\right)$$

$\doteqdot P(Z < -1.83)$

$= P(Z > 1.83)$

$= 1 - P(Z < 1.83)$

$= 1 - 0.9664$

$= 0.0336$

$\doteqdot 3.4\%$

11d $p_{\text{exact}} = 5.7\% = 0.057$

Percentage error

$= \dfrac{0.057 - 0.034}{0.057} \times 100\%$

$= 40.350 \ldots \%$

$\doteqdot 40\%$

This is not a good estimate. The sample is too small and we are not using any continuity correction.

11e $P(\hat{p} \leq 0.75)$

$\doteqdot P\left(Z \leq \dfrac{0.75 - 0.6}{\sqrt{0.012}}\right)$

$\doteqdot P(Z \leq 1.37)$

$= 0.9147$

$\doteqdot 91.5\%$

11f $p_{\text{exact}} = 95\% = 0.95$

Percentage error

$= \dfrac{0.95 - 0.9147}{0.95} \times 100\%$

$= 3.715\,78 \ldots \%$

$\doteqdot 3.7\%$

The curve is flatter at the top end and varies less with $\hat{p}$. Percentage difference is also exaggerated by small values, such as at the left end of the curve.

12a X = number of times a coin lands heads; $p = 0.5$

$\mu_{\hat{p}} = p = 0.5;\ q = 0.5$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{0.5 \times 0.5}{n}}$$

$$= \frac{0.5}{\sqrt{n}}$$

$P(\hat{p} \leq 0.52)$

$$\doteqdot P\left(Z \leq \frac{0.52 - 0.5}{\frac{0.5}{\sqrt{n}}}\right)$$

$$= P\left(Z \leq \frac{0.02\sqrt{n}}{0.5}\right)$$

$$= P\left(Z \leq 0.04\sqrt{n}\right)$$

When $n = 1000$,

$P(\hat{p} \leq 0.52)$

$\doteqdot P\left(Z \leq 0.04\sqrt{1000}\right)$

$\doteqdot P(Z \leq 1.26)$

$= 0.8962$

When $n = 500$,

$P(\hat{p} \leq 0.52)$

$\doteqdot P\left(Z \leq 0.04\sqrt{500}\right)$

$\doteqdot P(Z \leq 0.89)$

$= 0.8133$

When $n = 100$,

$P(\hat{p} \leq 0.52)$

$\doteqdot P\left(Z \leq 0.04\sqrt{100}\right)$

$= P(Z \leq 0.40)$

$= 0.6554$

When $n = 50$,

$P(\hat{p} \leq 0.52)$

$\doteqdot P\left(Z \leq 0.04\sqrt{50}\right)$

$\doteqdot P(Z \leq 0.28)$

$= 0.6103$

When $n = 25$,

$P(\hat{p} \leq 0.52)$

$\doteqdot P\left(Z \leq 0.04\sqrt{25}\right)$

$= P(Z \leq 0.20)$

$= 0.5793$

The results are summarised in the following table.

n	1000	500	100	50	25
exact	0.9026	0.8262	0.6914	0.6641	0.6550
approx.	0.8962	0.8133	0.6554	0.6103	0.5793

12b When $n = 1000$,

percentage error

$= \dfrac{0.9026 - 0.8962}{0.9026} \times 100\%$

$= 0.709\,06\ldots\%$

$\doteqdot 0.7\%$

When $n = 500$,

percentage error

$= \frac{0.8262 - 0.8133}{0.8262} \times 100\%$

$= 1.561\,36\ldots\%$

$\doteqdot 1.6\%$

When $n = 100$,

percentage error

$= \frac{0.6914 - 0.6554}{0.6914} \times 100\%$

$= 5.206\,82\ldots\%$

$\doteqdot 5.2\%$

When $n = 50$,

percentage error

$= \frac{0.6641 - 0.6103}{0.6641} \times 100\%$

$= 8.101\,18\ldots\%$

$\doteqdot 8.1\%$

When $n = 25$,

percentage error

$= \frac{0.6550 - 0.5793}{0.6550} \times 100\%$

$= 11.557\,25\ldots\%$

The results are summarised in the following table.

n	1000	500	100	50	25
exact	0.9026	0.8262	0.6914	0.6641	0.6550
approx.	0.8962	0.8133	0.6554	0.6103	0.5793
% error	0.7	1.6	5.2	8.1	11.6

12c Using the above table, we can see that as n increases, the accuracy of our approximation improves. This is because for very large trials, a binomial distribution is almost similar to the normal distribution.

13 $n = 50$; $X =$ number of people choosing the branded pineapple; $p = 0.5; q = 0.5$

For sample proportion:

$\mu_{\hat{p}} = p = 0.5$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{0.5 \times 0.5}{50}}$$

$$= \sqrt{0.005}$$

$$= 0.070\,71 \ldots$$

$$\doteqdot 0.071$$

60% of $50 = 0.6 \times 50 = 30$

$P(X > 30)$

$= P(\hat{p} > 0.6)$

$\doteqdot P\left(Z > \dfrac{0.6 - 0.5}{\sqrt{0.005}}\right)$

$\doteqdot P(Z > 1.41)$

$= 1 - P(Z < 1.41)$

$= 1 - 0.9207$

$= 0.0793$

$\doteqdot 8\%$

It appears that people strongly prefer the branded version, even though the two versions are identical. There may be an expectation that the branded version is superior, or they may prefer the packaging.

14 $n = 100$; $X =$ number of people who respond positively to a drug;

$p = 0.3; q = 0.7$

For sample proportion:

$\mu_{\hat{p}} = p = 0.3$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{0.3 \times 0.7}{100}}$$

$$= \sqrt{0.0021}$$

$$= 0.045\,82\ldots$$

$$\doteqdot 0.046$$

$P(\hat{p} \geq 0.4)$

$$\doteqdot P\left(Z \geq \frac{0.4 - 0.3}{\sqrt{0.0021}}\right)$$

$$\doteqdot P(Z \geq 2.18)$$

$$= 1 - P(Z \leq 2.18)$$

$$= 1 - 0.9854$$

$$= 0.0146$$

$$\doteqdot 1.5\%$$

The probability that this occurred simply by chance is very low.

15a $X =$ number of people on a college campus that are living at home; $p = 0.7$; $q = 0.3$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

15a i $\sigma_{\hat{p}} < 4\%$

$\sigma_{\hat{p}} < 0.04$

$\sqrt{\frac{0.7 \times 0.3}{n}} < 0.04$

$\frac{0.21}{n} < (0.04)^2$

$\frac{n}{0.21} > \frac{1}{(0.04)^2}$

$n > \frac{0.21}{(0.04)^2}$

$n > 131.25$

$n \geq 132$

So sample needs to be at least 132 residents.

15a ii $\sigma_{\hat{p}} < 3\%$

$\sigma_{\hat{p}} < 0.03$

$\sqrt{\frac{0.7 \times 0.3}{n}} < 0.03$

$n > \frac{0.21}{(0.03)^2}$

$n > 233.33 \ldots$

$n \geq 234$

So sample needs to be at least 234 residents.

15a iii $\sigma_{\hat{p}} < 2\%$

$\sigma_{\hat{p}} < 0.02$

$\sqrt{\frac{0.7 \times 0.3}{n}} < 0.02$

$n > \frac{0.21}{(0.02)^2}$

$n > 525$

$n \geq 526$

So sample needs to be more than 525 residents or at least 526 residents.

15a iv $\sigma_{\hat{p}} < 1\%$

$\sigma_{\hat{p}} < 0.01$

$\sqrt{\dfrac{0.7 \times 0.3}{n}} < 0.01$

$n > \dfrac{0.21}{(0.01)^2}$

$n > 2100$

$n \geq 2101$

So sample needs to be more than 2100 residents or at least 2101 residents.

15a v $\sigma_{\hat{p}} < k\%$

$\sigma_{\hat{p}} < \dfrac{k}{100}$

$\sqrt{\dfrac{0.7 \times 0.3}{n}} < \dfrac{k}{100}$

$\dfrac{0.21}{n} < \left(\dfrac{k}{100}\right)^2$

$n > \dfrac{0.21}{\left(\dfrac{k}{100}\right)^2}$

$n > \dfrac{2100}{k^2}$

So sample needs to be more than $\frac{2100}{k^2}$ residents.

15b $X =$ number of people on a college campus that are living at home; $p = 0.8$; $q = 0.2$

$\sigma_{\hat{p}} = \sqrt{\dfrac{pq}{n}}$

15b i $\sigma_{\hat{p}} < 4\%$

$\sigma_{\hat{p}} < 0.04$

$\sqrt{\frac{0.8 \times 0.2}{n}} < 0.04$

$\frac{0.16}{n} < (0.04)^2$

$\frac{n}{0.16} > \frac{1}{(0.04)^2}$

$n > \frac{0.16}{(0.04)^2}$

$n > 100$

$n \geq 101$

So sample needs to be more than 100 residents or at least 101 residents.

15b ii $\sigma_{\hat{p}} < 3\%$

$\sigma_{\hat{p}} < 0.03$

$\sqrt{\frac{0.8 \times 0.2}{n}} < 0.03$

$n > \frac{0.16}{(0.03)^2}$

$n > 177.77\ldots$

$n \geq 178$

So sample needs to be at least 178 residents.

15b iii $\sigma_{\hat{p}} < 2\%$

$\sigma_{\hat{p}} < 0.02$

$\sqrt{\frac{0.8 \times 0.2}{n}} < 0.02$

$n > \frac{0.16}{(0.02)^2}$

$n > 400$

$n \geq 401$

So sample needs to be more than 400 residents or at least 401 residents.

15b iv $\sigma_{\hat{p}} < 1\%$

$\sigma_{\hat{p}} < 0.01$

$\sqrt{\dfrac{0.8 \times 0.2}{n}} < 0.01$

$n > \dfrac{0.16}{(0.01)^2}$

$n > 1600$

$n \geq 1601$

So sample needs to be more than 1600 residents or at least 1601 residents.

15b v $\sigma_{\hat{p}} < k\%$

$\sigma_{\hat{p}} < \dfrac{k}{100}$

$\sqrt{\dfrac{0.8 \times 0.2}{n}} < \dfrac{k}{100}$

$\dfrac{0.16}{n} < \left(\dfrac{k}{100}\right)^2$

$n > \dfrac{0.16}{\left(\dfrac{k}{100}\right)^2}$

$n > \dfrac{1600}{k^2}$

So sample needs to be more than $\frac{1600}{k^2}$ residents.

15c $X =$ number of people on a college campus that are living at home; $p = 0.5$; $q = 0.5$

$\sigma_{\hat{p}} = \sqrt{\dfrac{pq}{n}}$

15c i $\sigma_{\hat{p}} < 4\%$

$\sigma_{\hat{p}} < 0.04$

$\sqrt{\dfrac{0.5 \times 0.5}{n}} < 0.04$

$\dfrac{0.25}{n} < (0.04)^2$

$\dfrac{n}{0.25} > \dfrac{1}{(0.04)^2}$

$n > \dfrac{0.25}{(0.04)^2}$

$n > 156.25$

$n \geq 157$

So sample needs to be at least 157 residents.

15c ii $\sigma_{\hat{p}} < 3\%$

$\sigma_{\hat{p}} < 0.03$

$\sqrt{\dfrac{0.5 \times 0.5}{n}} < 0.03$

$n > \dfrac{0.25}{(0.03)^2}$

$n > 277.77 \ldots$

$n \geq 278$

So sample needs to be at least 278 residents.

15c iii $\sigma_{\hat{p}} < 2\%$

$\sigma_{\hat{p}} < 0.02$

$\sqrt{\dfrac{0.5 \times 0.5}{n}} < 0.02$

$n > \dfrac{0.25}{(0.02)^2}$

$n > 625$

$n \geq 626$

So sample needs to be more than 625 residents or at least 626 residents.

15c iv $\sigma_{\hat{p}} < 1\%$

$\sigma_{\hat{p}} < 0.01$

$\sqrt{\frac{0.5 \times 0.5}{n}} < 0.01$

$n > \frac{0.25}{(0.01)^2}$

$n > 2500$

$n \geq 2501$

So sample needs to be more than 2500 residents or at least 2501 residents.

15c v $\sigma_{\hat{p}} < k\%$

$\sigma_{\hat{p}} < \frac{k}{100}$

$\sqrt{\frac{0.5 \times 0.5}{n}} < \frac{k}{100}$

$\frac{0.25}{n} < \left(\frac{k}{100}\right)^2$

$n > \frac{0.25}{\left(\frac{k}{100}\right)^2}$

$n > \frac{2500}{k^2}$

So sample needs to be more than $\frac{2500}{k^2}$ residents.

16a i $n = 10$; $X =$ number of heads from toss of 10 coins; $p = 0.5; q = 0.5$

$P(X = 6)$

$= {}^{10}C_6\ (0.5)^6(0.5)^4$

$= 0.205\ 078\ \ldots$

$\doteqdot 0.205$

16a ii $n = 40$; $p = 0.205$

$\mathrm{E}(X = 6)$

$= np$

$= 40 \times 0.205$

$= 8.2$

$\doteqdot 8$

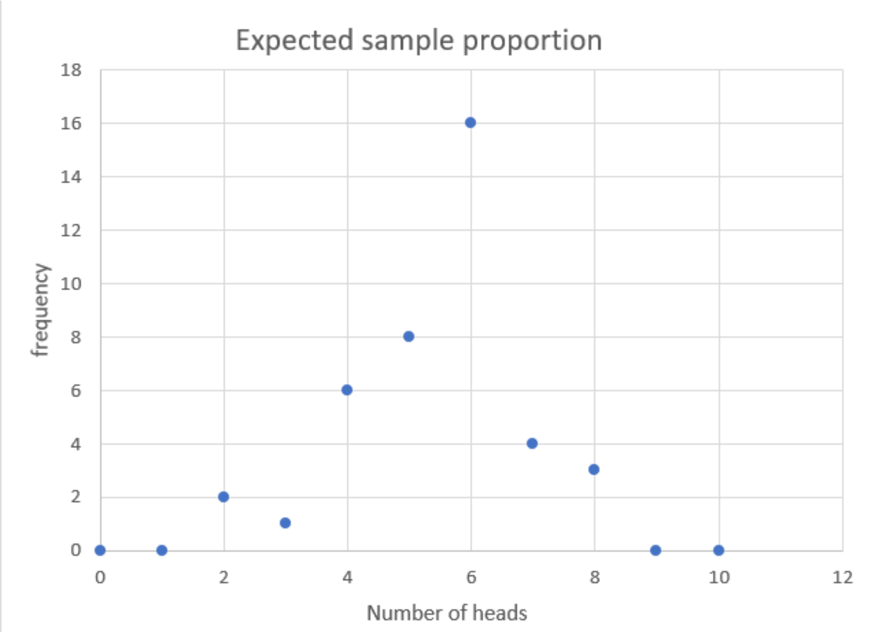
From 40 trials, you would expect to get exactly 6 heads about eight times.

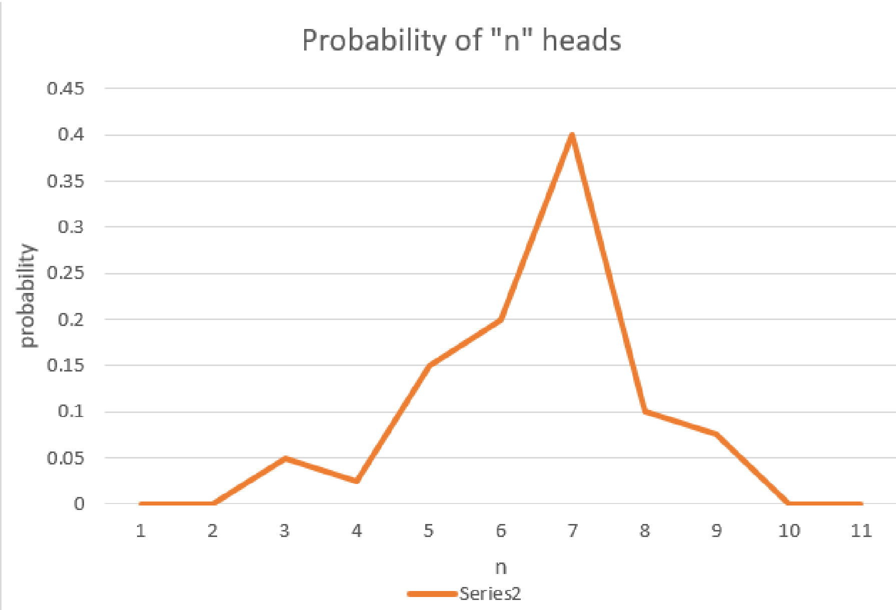
16b i Answer will vary. Sample spreadsheet simulation shown below.

<table>
<tr><th>Number of heads</th><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th></tr>
<tr><td>Proportion of heads</td><td>0</td><td>0.1</td><td>0.2</td><td>0.3</td><td>0.4</td><td>0.5</td><td>0.6</td><td>0.7</td><td>0.8</td><td>0.9</td><td>1</td></tr>
<tr><td>Expected frequency</td><td>0</td><td>0.4</td><td>1.8</td><td>4.7</td><td>8.2</td><td>9.8</td><td>8.2</td><td>4.7</td><td>1.8</td><td>0.4</td><td>0</td></tr>
<tr><td>Tally</td><td>|</td><td></td><td>|</td><td>||||</td><td>||||</td><td>|||||||||</td><td>||||||||</td><td>|||||||||</td><td>||</td><td>|</td><td>|</td></tr>
<tr><td>Frequency</td><td>1</td><td>0</td><td>1</td><td>4</td><td>4</td><td>9</td><td>8</td><td>9</td><td>2</td><td>1</td><td>1</td></tr>
<tr><td>probability</td><td>0.025</td><td>0</td><td>0.025</td><td>0.1</td><td>0.1</td><td>0.225</td><td>0.2</td><td>0.225</td><td>0.05</td><td>0.025</td><td>0.025</td></tr>
</table>

16b ii Answers will vary but for the sample simulation it was observed that the expected frequency is 8.

16b iii Charts for sample simulation.





17a Sum is at least 9 for the cases $6 + 3$, $6 + 4$, $6 + 5$, $6 + 6$, $5 + 4$, $5 + 5$, $5 + 6$, $4 + 5$, $4 + 6$, $3 + 6$. That is, for 10 cases out of the 36 possible outcomes of throwing two dice.

$$P(\text{sum is at least } 9) = \frac{10}{36} = \frac{5}{18}$$

17b $n = 20$; $X =$ number of successes where sum of numbers is at least 9;

$p = \frac{5}{18}; q = \frac{13}{18}$

$P(0 \le X \le 4)$

$= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= {}^{20}\mathrm{C}_0 \left(\frac{5}{18}\right)^0 \left(\frac{13}{18}\right)^{20} + {}^{20}\mathrm{C}_1 \left(\frac{5}{18}\right)^1 \left(\frac{13}{18}\right)^{19} + {}^{20}\mathrm{C}_2 \left(\frac{5}{18}\right)^2 \left(\frac{13}{18}\right)^{18} + {}^{20}\mathrm{C}_3 \left(\frac{5}{18}\right)^3 \left(\frac{13}{18}\right)^{17} + {}^{20}\mathrm{C}_4 \left(\frac{5}{18}\right)^4 \left(\frac{13}{18}\right)^{16}$$

$= 0.309\,61 \ldots$

$\doteqdot 0.3096$

17c We approximate the binomial by treating X as a normal random variable with:

$\mu = np$

$= 20 \times \frac{5}{18}$

$= \frac{50}{9}$

$\sigma = \sqrt{npq}$

$= \sqrt{\frac{50}{9} \times \frac{13}{18}}$

$= \sqrt{\frac{325}{81}}$

$P(0 \le X \le 4)$

$\doteqdot P\left(\frac{0-\mu}{\sigma} \le Z \le \frac{4-\mu}{\sigma}\right)$

$\doteqdot P(-2.77 \le Z \le -0.78)$

$= P(0.78 \le Z \le 2.77)$

$= P(Z \le 2.77) - P(Z \le 0.78)$

$= 0.9972 - 0.7823$

$= 0.2149$

17d $P(-0.5 \le X \le 4.5)$

$\doteqdot P\left(\frac{-0.5-\mu}{\sigma} \le Z \le \frac{4.5-\mu}{\sigma}\right)$

$\doteqdot P(-3.02 \le Z \le -0.53)$

$= P(0.53 \le Z \le 3.02)$

$= P(Z \le 3.02) - P(Z \le 0.53)$

$= 0.9987 - 0.7019$

$= 0.2968$

17e For sample proportion:

$\mu_{\hat{p}} = p = \frac{5}{18}$

$\sigma_{\hat{p}} = \sqrt{\frac{\frac{5}{18} \times \frac{13}{18}}{20}}$

$= \sqrt{\frac{13}{1296}}$

$P(0 \le \hat{p} \le 0.2)$

$\doteqdot P\left(\frac{0-\mu}{\sigma} \le Z \le \frac{0.2-\mu}{\sigma}\right)$

$\doteqdot P(-2.77 \le Z \le -0.78)$

$= P(0.78 \le Z \le 2.77)$

$= P(Z \le 2.77) - P(Z \le 0.78)$

$= 0.9972 - 0.7823$

$= 0.2149$

This agrees with the result to part c. The sample proportion distribution is just the binomial stretched vertically by a factor of n and compressed horizontally by a factor $\frac{1}{n}$, thus the corresponding areas will be the same. After the distribution has been converted to standard normal, the calculation is identical.

17f $P\left(0 - \frac{1}{40} \le \hat{p} \le 0.2 + \frac{1}{40}\right)$

$= P(-0.025 \le \hat{p} \le 0.225)$

$\doteqdot P\left(\frac{-0.025 - \mu}{\sigma} \le Z \le \frac{0.225 - \mu}{\sigma}\right)$

$\doteqdot P(-3.02 \le Z \le -0.53)$

$= P(0.53 \le Z \le 3.02)$

$= P(Z \le 3.02) - P(Z \le 0.53)$

$= 0.9987 - 0.7019$

$= 0.2968$

This result agrees with part d. The factor $\frac{1}{40}$ corresponds to half an interval on the histogram and thus applies the same continuity correction as part d.

18a i $P(\mu - 1.96\sigma \le X \le \mu + 1.96\sigma)$

$\doteqdot P\left(\frac{\mu - 1.96\sigma - \mu}{\sigma} \le Z \le \frac{\mu + 1.96\sigma - \mu}{\sigma}\right)$

$= P\left(\frac{-1.96\sigma}{\sigma} \le Z \le \frac{1.96\sigma}{\sigma}\right)$

$= P(-1.96 \le Z \le 1.96)$

$= P(Z \le 1.96) - P(Z \le -1.96)$

$= P(Z \le 1.96) - [1 - P(Z \le 1.96)]$

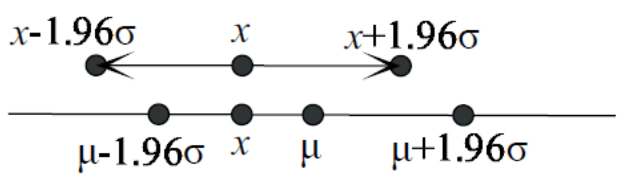
$= 0.9750 - (1 - 0.9750)$

$= 0.9750 - 0.0250$

$= 0.95$

$= 95\%$

18a ii



18b i $n = 100$; $p \doteqdot \hat{p} = 67\% = 0.67; q = 1 - p = 0.33$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$= \sqrt{\frac{0.67 \times 0.33}{100}}$$

$$= \sqrt{0.002\ 211}$$

$$\doteqdot 0.047$$

18b ii Margin of error

$= 1.96\sigma$

$= 1.96 \times \sqrt{0.002\ 211}$

$= 0.092\ 16 \ldots$

$\doteqdot 0.09$

$= 9\%$

Confidence interval

$= [\hat{p} - 1.96\sigma, \hat{p} + 1.96\sigma]$

$= [67\% - 9\%, 67\% + 9\%]$

$= [58\%, 76\%]$

18b iii $p \doteqdot 67\% = 0.67; q = 0.33$; margin of error to be reduced to 1%

Margin of error $= 1.96\sigma$ where $\sigma = \sqrt{\frac{pq}{n}}$

$$1.96 \times \sqrt{\frac{0.67 \times 0.33}{n}} \leq 0.01$$

$$\frac{0.67 \times 0.33}{n} \leq \left(\frac{0.01}{1.96}\right)^2$$

$$\frac{n}{0.67 \times 0.33} \geq \left(\frac{1.96}{0.01}\right)^2$$

$$n \geq \left(\frac{1.96}{0.01}\right)^2 \times 0.67 \times 0.33$$

$$n \geq 8493.77 \ldots$$

$$n > 8493$$

Sample size to be more than 8493 or at least 8494.

Solutions to Chapter review

1 $n = 4; X =$ number of times marksman hits target; $p = \frac{5}{6}$; $q = \frac{1}{6}$

1a $P(X = 3)$

$= {}^4C_3 \left(\frac{5}{6}\right)^3 \left(\frac{1}{6}\right)$

$= 0.385\ 80 \ldots$

$\doteqdot 0.39$

1b Exactly two misses means exactly two hits in the four shots.

$P(X = 2)$

$= {}^4C_2 \left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^2$

$= 0.115\ 74 \ldots$

$\doteqdot 0.12$

2 $n = 15$; $X =$ number of people who think Tasmania is the most beautiful state in Australia;

$p = \frac{5}{6}$; $q = \frac{1}{6}$

$P(X \geq 13)$

$= P(X = 13) + P(X = 14) + P(X = 15)$

$= {}^{15}C_{13} \left(\frac{5}{6}\right)^{13} \left(\frac{1}{6}\right)^2 + {}^{15}C_{14} \left(\frac{5}{6}\right)^{14} \left(\frac{1}{6}\right)^1 + {}^{15}C_{15} \left(\frac{5}{6}\right)^{15} \left(\frac{1}{6}\right)^0$

$= {}^{15}C_{13} \left(\frac{5}{6}\right)^{13} \left(\frac{1}{6}\right)^2 + {}^{15}C_{14} \left(\frac{5}{6}\right)^{14} \left(\frac{1}{6}\right)^1 + \left(\frac{5}{6}\right)^{15}$

$= 0.532\ 22 \ldots$

$\doteqdot 0.53$

3 $n = 10$; X = number of questions answered correctly; $p = \frac{1}{5}$; $q = \frac{4}{5}$

$P(X = 7)$

$= {}^{10}\mathrm{C}_7 \left(\frac{1}{5}\right)^7 \left(\frac{4}{5}\right)^3$

$= 0.000\ 786\ 432 \ldots$

$\doteqdot 0.000\ 786$

4 $n = ?$; X = number of hearts drawn from pack of cards; $p = \frac{1}{4}$; $q = \frac{3}{4}$

Need $P(X \geq 1) > 0.95$.

$P(X \geq 1)$

$= 1 - P(X = 0)$

$= 1 - {}^{n}\mathrm{C}_0 \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^n$

$= 1 - \left(\frac{3}{4}\right)^n$

So $1 - \left(\frac{3}{4}\right)^n > 0.95$

$1 - (0.75)^n > 0.95$

$(0.75)^n < 0.05$

$\log(0.75)^n < \log(0.05)$

$n \log(0.75) < \log(0.05)$

$n > \frac{\log(0.05)}{\log(0.75)}$ (as $\log(0.75)$ is negative)

$n > 10.413\ 34 \ldots$

$n \geq 11$

The experiment needs to be performed 11 times.

5a $n(S) = 8;\quad n(A) = 1$

$$p = \frac{n(A)}{n(S)}$$

$$= \frac{1}{8}$$

5b $X =$ number of eights that occur when six eight-sided dice are thrown;

$$p = \frac{1}{8};\quad q = \frac{7}{8}$$

$$P(X = 0)$$

$$= {}^6\mathrm{C}_0\left(\frac{1}{8}\right)^0\left(\frac{7}{8}\right)^6$$

$$= 1 \times 1 \times \left(\frac{7}{8}\right)^6$$

$$= 0.448\,79\ldots$$

$$\doteqdot 0.4488$$

$$P(X = 1)$$

$$= {}^6\mathrm{C}_1\left(\frac{1}{8}\right)^1\left(\frac{7}{8}\right)^5$$

$$= 0.384\,68\ldots$$

$$\doteqdot 0.3847$$

$$P(X = 2)$$

$$= {}^6\mathrm{C}_2\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^4$$

$$= 0.137\,38\ldots$$

$$\doteqdot 0.1374$$

$$P(X = 3)$$

$$= {}^6\mathrm{C}_3\left(\frac{1}{8}\right)^3\left(\frac{7}{8}\right)^3$$

$= 0.026\ 16\ \ldots$

$\doteqdot 0.0262$

$P(X = 4)$

$= {}^6C_4 \left(\frac{1}{8}\right)^4 \left(\frac{7}{8}\right)^2$

$= 0.002\ 80\ \ldots$

$\doteqdot 0.0028$

$P(X = 5)$

$= {}^6C_5 \left(\frac{1}{8}\right)^5 \left(\frac{7}{8}\right)^1$

$= 0.000\ 16\ \ldots$

$\doteqdot 0.0002$

$P(X = 6)$

$= {}^6C_6 \left(\frac{1}{8}\right)^6 \left(\frac{7}{8}\right)^0$

$= 0.000\ 003\ 81\ \ldots$

$\doteqdot 0.0000$

These results are summarised in the following distribution table.

x	0	1	2	3	4	5	6
$P(X = x)$	0.4488	0.3847	0.1374	0.0262	0.0028	0.0002	0.0000

5c $n = 1000;\ p \doteqdot 0.0262$

$\mathrm{E}(X = 3)$

$= np$

$\doteqdot 1000 \times 0.0262$

$= 26.2$

In 1000 throws of the six dice, you would expect to get exactly three eights about 26 times.

5d $n = 1000$

$p = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$

$= 0.0262 + 0.0028 + 0.0002 + 0$

$= 0.0292$

$\mathrm{E}(X \geq 3)$

$= np$

$\doteqdot 1000 \times 0.0292$

$= 29.2$

In 1000 throws of the six dice, you would expect to get three or more eights about 29 times.

6a Yes. There are only two possible outcomes: heads and tails. Each coin toss is independent of the other and has a probability of 0.5 for each outcome for every trial.

$p = q = 0.5$

6b Yes. There are only two possible outcomes: winning if the sum is more than 10 and losing. Every time the dice are thrown, the sum of the numbers appearing is independent of the previous throw's outcome.

Successful outcomes are (5, 6), (6, 5), (6, 6). So there are 3 successful outcomes out of 36 possible outcomes.

$p = \frac{3}{36} = \frac{1}{12}; \quad q = \frac{11}{12}$

6c Yes. There are only two possible outcomes: passing quality control and not passing. The selection of the item is random and the chance of an item passing is independent of the result of the previous item.

There are 4 successful outcomes out of 1000 possible outcomes.

$p = \frac{4}{1000} = 0.004; \quad q = 0.996$

6d No. It is not mentioned if the card is put back or not. Also, there are four choices for the suit of the card but there should be only two possible outcomes for a Bernoulli trial.

7a $B(n, p) = B(20, 0.2)$

$\mu = np$

$= 20 \times 0.2$

$= 2$

$\sigma^2 = npq$

$= 20 \times 0.2 \times 0.8$

$= 3.2$

$\sigma = \sqrt{npq}$

$= \sqrt{3.2}$

$= 1.788\,85\ldots$

$\doteqdot 1.79$

7b $B(n, p) = B(70, 0.5)$

$\mu = np$

$= 70 \times 0.5$

$= 35$

$\sigma^2 = npq$

$= 70 \times 0.5 \times 0.5$

$= 17.5$

$\sigma = \sqrt{npq}$

$= \sqrt{17.5}$

$= 4.183\,30\ldots$

$\doteqdot 4.18$

7c $B(n,p) = B(6, 0.8)$

$\mu = np$

$= 6 \times 0.8$

$= 4.8$

$\sigma^2 = npq$

$= 6 \times 0.8 \times 0.2$

$= 0.96$

$\sigma = \sqrt{npq}$

$= \sqrt{0.96}$

$= 0.979\,79\ldots$

$\doteqdot 0.98$

7d $B(n,p) = B(120, 0.4)$

$\mu = np$

$= 120 \times 0.4$

$= 48$

$\sigma^2 = npq$

$= 120 \times 0.4 \times 0.6$

$= 28.8$

$\sigma = \sqrt{npq}$

$= \sqrt{28.8}$

$= 5.366\,56\ldots$

$\doteqdot 5.37$

7e $B(n,p) = B(300, 0.1)$

$\mu = np$

$= 300 \times 0.1$

$= 30$

$\sigma^2 = npq$

$= 300 \times 0.1 \times 0.9$

$= 27$

$\sigma = \sqrt{npq}$

$= \sqrt{27}$

$= 5.196\,15\ldots$

$\doteqdot 5.20$

7f $B(n, p) = B(5, 0.25)$

$\mu = np$

$= 5 \times 0.25$

$= 1.25$

$\sigma^2 = npq$

$= 5 \times 0.25 \times 0.75$

$= 0.9375$

$\sigma = \sqrt{npq}$

$= \sqrt{0.9375}$

$= 0.968\,24\ldots$

$\doteqdot 0.97$

8 $n = 60;\ X = \text{number of cases that do not pass inspection};\ p = 0.05;\ q = 0.95$

8a $B(60, 0.05)$

$\mu = np$

$= 60 \times 0.05$

$= 3$

$\sigma^2 = npq$

$= 60 \times 0.05 \times 0.95$

$= 2.85$

$\sigma = \sqrt{npq}$

$= \sqrt{2.85}$

$= 1.688\ 19\ ...$

$\doteqdot 1.7$

8b $\mu = 3;\ \sigma \doteqdot 1.7$

$\mu - \sigma \doteqdot 3 - 1.7 = 1.3$

$\mu + \sigma \doteqdot 3 + 1.7 = 4.7$

$P(1.3 \le X \le 4.7)$

$= P(X = 2, 3 \text{ or } 4)$

$= P(X = 2) + P(X = 3) + P(X = 4)$

$= {}^{60}C_2\ (0.05)^2(0.95)^{58} + {}^{60}C_3\ (0.05)^3(0.95)^{57} + {}^{60}C_4\ (0.05)^4(0.95)^{56}$

$= 0.628\ 111\ ...$

$\doteqdot 62.8\%$

8c $n = 60;\ p = 0.05;\ q = 0.95;\ \mu + \sigma \doteqdot 3 + 1.7 = 4.7$

$P(\text{batch rejected})$

$= P(X > 4)$

$= 1 - P(X \le 4)$

$= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)]$

$= 1 - [{}^{60}C_0\ (0.05)^0(0.95)^{60} + {}^{60}C_1\ (0.05)^1(0.95)^{59} + {}^{60}C_2\ (0.05)^2(0.95)^{58}$

$+ {}^{60}C_3\ (0.05)^3(0.95)^{57} + {}^{60}C_4\ (0.05)^4(0.95)^{56}]$

$= 1 - 0.819\ 66\ ...$

$= 0.180\ 33\ ...$

$\doteqdot 18\%$

The probability of rejecting the batch is 18%.

8d $n = 60;\ p = 0.02;\ q = 0.98$

$\mu = np$

$= 60 \times 0.02$

$= 1.2$

$\sigma^2 = npq$

$= 60 \times 0.02 \times 0.98$

$= 1.176$

$\sigma = \sqrt{npq}$

$= \sqrt{1.176}$

$= 1.084\ 43\ \ldots$

$\doteqdot 1.08$

$\mu + \sigma \doteqdot 1.2 + 1.08 = 2.28$

$P(\text{batch rejected})$

$= P(X > 2)$

$= 1 - P(X \le 2)$

$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$

$= 1 - [{}^{60}\mathrm{C}_0\ (0.02)^0(0.98)^{60} + {}^{60}\mathrm{C}_1\ (0.02)^1(0.98)^{59} + {}^{60}\mathrm{C}_2\ (0.02)^2(0.98)^{58}]$

$= 1 - 0.881\ 25\ \ldots$

$= 0.118\ 74\ \ldots$

$\doteqdot 12\%$

The probability of rejecting the batch is 12%.

9a $n = 80;\ X = \text{number of heads in 80 tosses of coin};\ p = 0.5;\ q = 0.5$

$P(X = 38, 39 \text{ or } 40)$

$= P(X = 38) + P(X = 39) + P(X = 40)$

$= {}^{80}\mathrm{C}_{38}\ (0.5)^{38}(0.5)^{42} + {}^{80}\mathrm{C}_{39}\ (0.5)^{39}(0.5)^{41} + {}^{80}\mathrm{C}_{40}\ (0.5)^{40}(0.5)^{40}$

$= 0.256\ 24\ \ldots$

$\doteqdot 25.62\%$

9b $\mu = np$

$= 80 \times 0.5$

$= 40$

$\sigma = \sqrt{npq}$

$= \sqrt{40 \times 0.5}$

$= \sqrt{20}$

$= 4.472\ 13 \ldots$

$\doteqdot 4.47$

9c $np = 80 \times 0.5 = 40$

$nq = 80 \times 0.5 = 40$

$np > 5$ and $nq > 5$

Hence, a normal approximation to the binomial may be used.

9d The probability of a binomial distribution is the area under the histogram of the binomial distribution with unit width of each bar. Normal distribution is a continuous curve which approximates the binomial distribution (a discrete distribution). To account for the rectangles of unit width and the small triangular portions of area that are out of the normal distribution curve, we calculate the area for $P(37.5 \leq X \leq 40.5)$ rather than $P(38 \leq X \leq 40)$.

This is called a continuity correction and occurs because we are approximating a discrete distribution by a continuous curve.

9e $P(37.5 \leq X \leq 40.5)$

$\doteqdot P\left(\frac{37.5 - 40}{\sqrt{20}} \leq Z \leq \frac{40.5 - 40}{\sqrt{20}}\right)$

$\doteqdot P(-0.56 \leq Z \leq 0.11)$

$= P(Z \leq 0.11) - [1 - P(Z \leq 0.56)]$

$= 0.5438 - 0.2887$

$= 0.2561$

$\doteqdot 25.6\%$

9f Percentage error

$= \dfrac{0.2562 - 0.2561}{0.2562} \times 100\%$

$= 0.039\,03\,...\%$

$\doteqdot 0.04\%$

The percentage error is less than 0.1%.

9g $P(X \geq 50)$

$\doteqdot P\left(Z \geq \dfrac{49.5 - 40}{\sqrt{20}}\right)$

$\doteqdot P(Z \geq 2.12)$

$= 1 - P(Z \leq 2.12)$

$= 1 - 0.9830$

$= 0.0170$

$\doteqdot 1.7\%$

10a Outcomes when sum is at least 10: $\{(4,6),(5,5),(6,4),(5,6),(6,5),(6,6)\}$

There are 6 outcomes out of a possible 36 outcomes.

$P(X \geq 10) = \dfrac{6}{36} = \dfrac{1}{6}$

10b $n = 80; X =$ number of times the sum of two dice is at least 10; $p = \frac{1}{6}; q = \frac{5}{6}$

$\mu = np$

$= 80 \times \dfrac{1}{6}$

$= \dfrac{40}{3}$

$\sigma = \sqrt{npq}$

$= \sqrt{\dfrac{40}{3} \times \dfrac{5}{6}}$

$= \sqrt{\dfrac{100}{9}}$

$= \dfrac{10}{3}$

$P(X \geq 15)$

$\doteqdot P\left(Z > \dfrac{14.5 - \frac{40}{3}}{\frac{10}{3}}\right)$

$= P(Z \geq 0.35)$

$= 1 - P(Z \leq 0.35)$

$= 1 - 0.6368$

$= 0.3632$

$\doteqdot 36\%$

11a $n = 100$; $n(\text{intending to vote yes}) = 35$

$p = \dfrac{35}{100} = 0.35$

11b Not necessarily. A voter's choice of a party is dependent on their personal interest and the ideology of the party, both of which are subject to change.

11c It will be a binomial distribution with $n = 100$ and the probability of success equal to the (unknown) proportion of the population intending to vote for the WTP.

12a $n = 5$; $X =$ number of red balls in selection of 5 balls; $p = \frac{3}{5}$; $q = \frac{2}{5}$

When $X = 0$, $\hat{p} = \dfrac{0}{5} = 0$.

$P(\hat{p} = 0)$

$= {}^5C_0 \left(\dfrac{3}{5}\right)^0 \left(\dfrac{2}{5}\right)^5$

$= 1 \times 1 \times \frac{32}{3125}$

$= \frac{32}{3125}$

When $X = 1,\ \hat{p} = \frac{1}{5} = 0.2.$

$P(\hat{p} = 0.2)$

$= {}^5C_1 \left(\frac{3}{5}\right)^1 \left(\frac{2}{5}\right)^4$

$= 5 \times \frac{3}{5} \times \frac{16}{625}$

$= \frac{48}{625}$

When $X = 2,\ \hat{p} = \frac{2}{5} = 0.4$.

$P(\hat{p} = 0.4)$

$= {}^5C_2 \left(\frac{3}{5}\right)^2 \left(\frac{2}{5}\right)^3$

$= 10 \times \frac{9}{25} \times \frac{8}{125}$

$= \frac{144}{625}$

When $X = 3,\ \hat{p} = \frac{3}{5} = 0.6$.

$P(\hat{p} = 0.6)$

$= {}^5C_3 \left(\frac{3}{5}\right)^3 \left(\frac{2}{5}\right)^2$

$= 10 \times \frac{27}{125} \times \frac{4}{25}$

$= \frac{216}{625}$

When $X = 4$, $\hat{p} = \dfrac{4}{5} = 0.8$.

$P(\hat{p} = 0.8)$

$= {}^5C_4 \left(\frac{3}{5}\right)^4 \left(\frac{2}{5}\right)^1$

$= 5 \times \dfrac{81}{625} \times \dfrac{2}{5}$

$= \dfrac{162}{625}$

When $X = 5$, $\hat{p} = \dfrac{5}{5} = 1$.

$P(\hat{p} = 1)$

$= {}^5C_5 \left(\frac{3}{5}\right)^5 \left(\frac{2}{5}\right)^0$

$= 1 \times \dfrac{243}{3125} \times 1$

$= \dfrac{243}{3125}$

The results are summarised in the following table.

$\hat{p}$	0	0.2	0.4	0.6	0.8	1
$P(\hat{p})$	$\dfrac{32}{3125}$	$\dfrac{48}{625}$	$\dfrac{144}{625}$	$\dfrac{216}{625}$	$\dfrac{162}{625}$	$\dfrac{243}{3125}$

12b i $\hat{p} = 40\% = 0.4$ means $X = 0.4 \times 5 = 2$

$P(\hat{p} < 0.4)$

$= P(X < 2)$

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

$= P(\hat{p} = 0) + P(\hat{p} = 0.2)$

$= \dfrac{32}{3125} + \dfrac{48}{625}$

$= \frac{272}{3125}$

12b ii $\hat{p} = 50\% = 0.5$ means $X = 0.5 \times 5 = 2.5$

$P(\hat{p} < 0.5)$

$= P(X < 2.5)$

$= P(X = 0) + P(X = 1) + P(X = 2)$

$= P(\hat{p} = 0) + P(\hat{p} = 0.2) + P(\hat{p} = 0.4)$

$= \frac{32}{3125} + \frac{48}{625} + \frac{144}{625}$

$= \frac{992}{3125}$

12c $n = 5$; $p = \frac{3}{5}$; $q = \frac{2}{5}$

For the sample proportion:

$\mu_{\hat{p}} = p = \frac{3}{5}$

$\sigma_{\hat{p}}{}^2 = \frac{pq}{n}$

$= \frac{\frac{3}{5} \times \frac{2}{5}}{5}$

$= \frac{6}{125}$

$\sigma_{\hat{p}} = \sqrt{\frac{6}{125}}$

$= \frac{1}{25}\sqrt{30}$

13 a $n = 500$; $X =$ number of sixes thrown; $p = \frac{1}{6}$; $q = \frac{5}{6}$

For the sample proportion:

$$\mu_{\hat{p}} = p = \frac{1}{6}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{\frac{1}{6} \times \frac{5}{6}}{500}}$$

$$= \sqrt{\frac{1}{3600}}$$

$$= \frac{1}{60}$$

13b $n = 500$; $X = 70$; $p = \frac{1}{6}$

$$\hat{p} = \frac{70}{500} = \frac{7}{50};\ \mu_{\hat{p}} = \frac{1}{6};\ \sigma = \frac{1}{60}$$

Number of standard deviations away

$$= \frac{\frac{7}{50} - \frac{1}{6}}{\frac{1}{60}}$$

$$= -\frac{8}{5}$$

$$= -1.6$$

Hence, the result is 1.6 standard deviations below the mean.

14 $n = 653$; $X =$ number of male births; $p = 0.53$

$$\hat{p} = 54\% = 0.54$$

$$\text{or } \hat{p} = \frac{0.54 \times 653}{653} = 0.54$$

$$\mu_{\hat{p}} = p = 0.53$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{0.53 \times 0.47}{653}}$$

$$= 0.019\,53\,\ldots$$

$$\doteqdot 0.0195$$

$$P(\hat{p} > 0.54)$$

$$\doteqdot P\left(Z > \frac{0.54 - 0.53}{0.0195}\right)$$

$$\doteqdot P(Z > 0.51)$$

$$= 1 - P(Z < 0.51)$$

$$= 1 - 0.6950$$

$$= 0.3050$$

$$\doteqdot 31\%$$