Chapter 13 — Sampling and confidence intervals

Exercise 13.2 – Sample statistics

1 Mr Parker teaches on average 120 students per day. This is the population size. $N = 120$
He asks one class of 30 about the amount of homework they have that night. This is the sample size. $n = 30$

2 Bruce is able to hem 100 shirts per day. This is the population size. $N = 100$
Each day he checks 5 to make sure that they are suitable. This is the sample size. $n = 5$

3 Ms Lane tests her joke on this year's class (15 students). This is the sample size. $n = 15$
We don't know how many students Ms Lane will teach. The population size is unknown.

4 Lee-Yin asks 9 friends what they think. This is the sample size. $n = 9$
We don't know how many people will eventually eat Lee-Yin's cake pops. The population size is unknown.

5 **a** Population parameter
b Sample statistic

6 **a** Population parameter
b Sample statistic

7 Number of boys: $\dfrac{523}{523 + 621} \times 75 = 34.29$. Therefore 34 boys.
Number of girls: $\dfrac{621}{523 + 621} \times 75 = 40.71$. Therefore 41 girls.

8 Number of boarders: 23% of 90 = 20.7. Therefore 21 boarders.
The rest of the sample will be day students. $90 - 21 = 69$ day students.

9 **a** We don't know how many people will eventually eat the pudding. The population size is unknown.
b 40 volunteers to taste test your recipe. This is the sample size. $n = 40$

10 **a** We don't know how many people will eventually receive the vaccine. The population size is unknown.
b 247 suitable people test the vaccine. This is the sample size. $n = 247$

11 Sample statistic

12 Population parameter

13 Sample statistics

14 Population parameter

15 **a** A systematic sample with $k = 10$
b Yes – assuming that the order of patients is random

16 The sample is not random, therefore the results are not likely to be random

17 It is probably not random. Tony is likely to ask people that he knows, or people that approach him.

18 Number of male staff: 60% of 1500 = 900
Number of full time male staff: 95% of 900 = 855
Number to sample: $\dfrac{855}{1500} \times 80 = 45.6$
Number of part-time male staff: $900 - 855 = 45$
Number to sample: $\dfrac{45}{1500} \times 80 = 2.4$
Number of female staff: $1500 - 900 = 600$
Number of full time female staff: 78% of 600 = 460
Number to sample: $\dfrac{460}{1500} \times 80 = 24.96$
Number of part time female staff: $600 - 460 = 140$
Number to sample: $\dfrac{140}{1500} \times 80 = 7.47$
The sample consists of:
Full time male staff: 46
Part time male staff: 2
Full time female staff: 25
Part time female staff: 7

19 Use the random number generator on your calculator to produce numbers from 1 to 100. Keep generating numbers until you have 10 different numbers. Answers will vary.

20 First assign every person in your class a number e.g. 1–25. Decide how many students will be in your sample, e.g. 5. Then use the random number generator on your calculator to produce numbers from 1 to 25. Keep generating numbers until you have 5 different numbers. The students that were assigned these numbers are the 5 students in your random sample. Answers will vary.

Exercise 13.3 – The distribution of $\hat{p}$

1
$$\begin{aligned}\hat{p} &= \frac{6}{15}\\ &= \frac{2}{5}\end{aligned}$$

2
$$\begin{aligned}\hat{p} &= \frac{6}{20}\\ &= \frac{3}{10}\end{aligned}$$

3 $N = 1000$
$n = 50$
$p = 0.85$
Is $10n \le N$? $10n = 500$, therefore $10n \le N$.
Is $np \ge 10$? $\begin{aligned}np &= 50 \times 0.85\\ &= 42.5\end{aligned}$ therefore $np \ge 10$.
Is $nq \ge 10$? $\begin{aligned}nq &= 50 \times 0.15\\ &= 7.5\end{aligned}$ therefore $nq \not\geq 10$
The sample is not large.
For the sample to be large,
$$\begin{aligned}nq &= 10\\ 0.15n &= 10\\ n &= 66.7\end{aligned}$$
$n = 67$ is the smallest sample size that can be considered large.

4
$$\begin{aligned}np &= 10\\ 0.05n &= 10\\ n &= 200\end{aligned}$$
As $p < q$, if $np \le 10$, then $np \le 10$.
Is $10n \le N$? $10n = 2000$, therefore $n = 200$ is a large sample.

5 **a**
$$\begin{aligned}\mu_{\hat{p}} &= p\\ &= 0.5\end{aligned}$$

b $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$
$= \sqrt{\dfrac{0.5 \times 0.5}{50}}$
$= 0.07$

6 If $N = 1000, n = 100$ and $p = 0.8$.

a $\mu_{\hat{p}} = p$
$= 0.8$

b $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$
$= \sqrt{\dfrac{0.8 \times 0.2}{100}}$
$= 0.04$

7 $\hat{p} = \dfrac{27}{30}$
$= \dfrac{9}{10}$

8 $\hat{p} = \dfrac{147}{537}$

9 a $p = \dfrac{12}{21} = \dfrac{4}{7}$

b 0 females chosen out of 4, 1 chosen out of 4, 2 chosen out of 4, 3 chosen out of 4 or 4 chosen out of 4.

Therefore the possible values for $\hat{p}$ are $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$.

c

X	$\hat{p}$	Number of samples	Relative frequency
0	0	${}^{12}C_0\,{}^9C_4 = 126$	0.021
1	$\frac{1}{4}$	${}^{12}C_1\,{}^9C_3 = 1008$	0.168
2	$\frac{1}{2}$	${}^{12}C_2\,{}^9C_2 = 2376$	0.397
3	$\frac{3}{4}$	${}^{12}C_3\,{}^9C_1 = 1980$	0.331
4	1	${}^{12}C_4\,{}^9C_0 = 495$	0.083
	TOTAL samples	5985	

Probability distribution table:

$\hat{p}$	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1
$P(\hat{P} = \hat{p})$	0.021	0.168	0.397	0.331	0.083

d $P(\hat{P} > 0.6) = P\left(\hat{P} = \dfrac{3}{4}\right) + P(\hat{P} = 1)$
$= 0.331 + 0.083$
$= 0.414$

e $P(\hat{P} > 0.5 | \hat{P} > 0.3) = \dfrac{P(\hat{P} > 0.5)}{P(\hat{P} > 0.3)}$
$= \dfrac{0.414}{0.414 + 0.397}$
$= 0.510$

10 a $0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$

b

X	$\hat{p}$	$P(\hat{P} = \hat{p})$
0	0	${}^5C_0(0.62)^0(0.38)^5 = 0.008$
1	$\frac{1}{5}$	${}^5C_1(0.62)^1(0.38)^4 = 0.064$
2	$\frac{2}{5}$	${}^5C_2(0.62)^2(0.38)^3 = 0.211$
3	$\frac{3}{5}$	${}^5C_3(0.62)^3(0.38)^2 = 0.344$
4	$\frac{4}{5}$	${}^5C_4(0.62)^4(0.38)^1 = 0.281$
5	1	${}^5C_5(0.62)^5(0.38)^0 = 0.092$
	TOTAL samples	5985

Probability distribution table:

$\hat{p}$	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	1
$P(\hat{P} = \hat{p})$	0.008	0.064	0.211	0.344	0.281	0.092

c $P(\hat{P} > 0.5) = P\left(\hat{P} = \dfrac{3}{5}\right) + P\left(\hat{P} = \dfrac{4}{5}\right) + P(\hat{P} = 1)$
$= 0.344 + 0.281 + 0.092$
$= 0.717$

11 $np = 10$ As $p < q$, if $np \le 10$, then $nq \le 10$.
$0.01n = 10$
$n = 1000$
Is $10n \le N$? $10n = 10\,000$, therefore $n = 1000$ is a large sample.

12 $\mu_{\hat{p}} = p$
$= 0.15$
$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$
$= \sqrt{\dfrac{0.15 \times 0.85}{150}}$
$= 0.029$

13 $\mu_{\hat{p}} = p$
$= 0.75$
$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$
$= \sqrt{\dfrac{0.75 \times 0.25}{100}}$
$= 0.043$

14 $\mu_{\hat{p}} = p$
$p = 0.12$
$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$
$0.0285 = \sqrt{\dfrac{0.12 \times 0.88}{n}}$
$8.1225 \times 10^{-4} = \dfrac{0.1056}{n}$
$n = \dfrac{0.1056}{8.1225 \times 10^{-4}}$
$= 130$

15 $\mu_{\hat{p}} = p$

$p = 0.81$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$0.0253 = \sqrt{\frac{0.81 \times 0.19}{n}}$$

$$6.4009 \times 10^{-4} = \frac{0.1539}{n}$$

$$n = \frac{0.1539}{6.4009 \times 10^{-4}}$$

$$= 240.4$$

Therefore $n = 240$

16 See figure at the foot of the page*

17

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$0.015 = \sqrt{\frac{p(1-p)}{510}}$$

$$2.25 \times 10^{-4} = \frac{p(1-p)}{510}$$

$$0.11475 = p(1-p)$$

$$= p - p^2$$

The quadratic $p^2 - p + 0.11\,475 = 0$ can be solved using the quadratic formula.

$$p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$= \frac{1 \pm \sqrt{1 - 4 \times 1 \times 0.11\,475}}{2}$$

$$= \frac{1 \pm \sqrt{0.541}}{2}$$

$p = 0.87$ or $p = 0.13$

As $\hat{p} > 0.5$, the population proportion is $p = 0.87$.

18

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$0.0255 = \sqrt{\frac{p(1-p)}{350}}$$

$$6.5025 \times 10^{-4} = \frac{p(1-p)}{350}$$

$$0.227\,587\,5 = p(1-p)$$

$$= p - p^2$$

The quadratic $p^2 - p + 0.227\,587\,5 = 0$ can be solved using the quadratic formula.

$$p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$= \frac{1 \pm \sqrt{1 - 4 \times 1 \times 0.227\,587\,5}}{2}$$

$$= \frac{1 \pm \sqrt{0.08\,965}}{2}$$

$p = 0.65$ or $p = 0.35$

Exercise 13.4 – Confidence intervals

1 $n = 30$

$\hat{p} = 0.78$

$z = 1.96$

The 95% confidence interval is

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.78 \pm 1.96\sqrt{\frac{0.78 \times 0.22}{30}}$$

So, $C.I = (0.63, 0.93)$

2 $n = 53$

$\hat{p} = 0.82$

$z = 1.96$

The 95% confidence interval is

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.82 \pm 1.96\sqrt{\frac{0.82 \times 0.18}{53}}$$

So, $C.I. = (0.72, 0.92)$

3 $n = 116$

$\hat{p} = 0.86$

$z = 2.58\ (P(Z < z) = 0.005)$

The 99% confidence interval is

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.86 \pm 2.58\sqrt{\frac{0.86 \times 0.14}{116}}$$

So, $C.I. = (0.78, 0.94)$

4 $n = 95$

$\hat{p} = 0.3$

$z = 1.64\ (P(Z < z) = 0.05)$

The 90% confidence interval is

Which gives $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3 \pm 1.64\sqrt{\frac{0.3 \times 0.7}{95}}$

So $C.I. = (0.22, 0.38)$

5 95% confidence interval $z = 1.96$

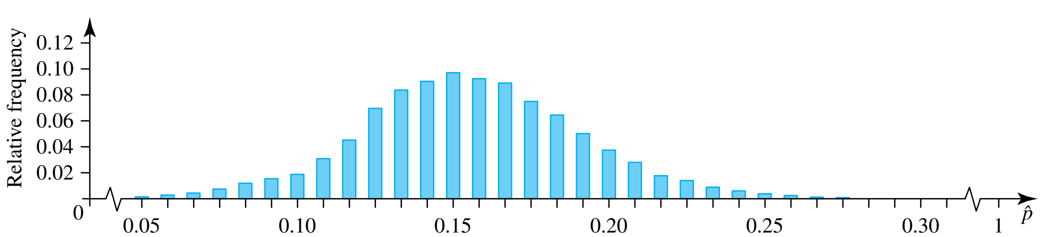
$\hat{p}$ will be at the center of the interval, $\hat{p} = 0.4$

The confidence interval is $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This means that

$$z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05.$$

$$z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$$

16*



$$1.96\sqrt{\frac{0.4\times 0.6}{n}} = 0.05$$
$$\sqrt{\frac{0.24}{n}} = 0.0255$$
$$\frac{0.24}{n} = 6.5077\times 10^{-4}$$
$$n = \frac{0.24}{6.5077\times 10^{-4}}$$
$$= 368.8$$

A sample of size 369 is needed.

6 90% confidence interval $z = 1.64$

$\hat{p}$ will be at the center of the interval, $\hat{p} = 0.8$

The confidence interval is $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This means that $z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$.

$$z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$$
$$1.64\sqrt{\frac{0.8\times 0.2}{n}} = 0.05$$
$$\sqrt{\frac{0.16}{n}} = 0.0305$$
$$\frac{0.16}{n} = 9.285\times 10^{-4}$$
$$n = \frac{0.16}{9.285\times 10^{-4}}$$
$$= 172.1$$

A sample of size 173 is needed.

7 $n = 250$

$$\hat{p} = \frac{20}{250}$$
$$= 0.08$$
$$z = 1.96$$

The 95% confidence interval is $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.08 \pm 1.96\sqrt{\frac{0.08\times 0.92}{250}}$$

So $C.I. = (0.46, 0.114)$

4.6% − 11.4% of complaints take more than 1 day to resolve.

8 $n = 250$

$$\hat{p} = \frac{230}{250}$$
$$= 0.92$$
$$z = 2.58$$

$$0.92 \pm 2.58\sqrt{\frac{0.08\times 0.92}{250}}$$

So, $C.I = (0.876, 0.964)$

87.6%–96.4% of complaints are resolved within 1 day.

9 95% confidence interval means that $z = 1.96$

The interval is $(0, 0.5)$

$\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.025 = 2.5\%$

The confidence interval is $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This means that $z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.025$.

$$z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.025$$
$$1.96\sqrt{\frac{0.025\times 0.975}{n}} = 0.025$$
$$\sqrt{\frac{0.024\,375}{n}} = 0.012\,755$$
$$\frac{0.024\,375}{n} = 1.6269\times 10^{-4}$$
$$n = \frac{0.024\,375}{1.6269\times 10^{-4}}$$
$$= 149.8$$

A sample of size 150 is needed.

10 $n = 250$

$$\hat{p} = \frac{92}{250}$$
$$= 0.368$$
$$z = 1.64\,485$$

The 90% confidence interval $0.368 \pm 1.64\sqrt{\frac{0.368\times 0.632}{250}}$

So $C.I. = (0.318, 0.418)$

31.8%–41.8% of Australians have Type A blood.

11 $n = 250,\ p = 0.65$

Since n is large, we can approximate the distribution of $\hat{P}$ to that of a normal curve. Therefore $\mu = p = 0.65$ and

$$\sigma = \sqrt{\frac{0.65\times 0.35}{250}} = 0.030$$

$\mathrm{P}\left(\hat{P} < 0.6\right) = 0.0487$

12 $n = 200,\ p = 0.8$

Since n is large, we can approximate the distribution of $\hat{P}$ to that of a normal curve. Therefore $\mu = p = 0.8$ and

$$\sigma = \sqrt{\frac{0.8\times 0.2}{200}} = 0.0283$$

$$P\left(0.8 < \hat{P} < 0.9 | \hat{P} > 0.65\right) = \frac{P\left(0.8 < \hat{P} < 0.9\right)}{P\left(\hat{P} > 0.65\right)}$$
$$= \frac{0.4998}{0.9999}$$
$$= 0.4998$$

13 $z = 1.96$

$\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.3$

The confidence interval is $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This means that $z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$.

$$z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$$
$$1.96\sqrt{\frac{0.3\times 0.7}{n}} = 0.05$$
$$\sqrt{\frac{0.21}{n}} = 0.0255$$
$$\frac{0.21}{n} = 6.5077\times 10^{-4}$$
$$n = \frac{0.21}{6.5077\times 10^{-4}}$$
$$= 322.7$$

A sample of size 323 is needed.

14 $z = 2.58$

$\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.25$

The confidence interval is $\hat{p} \pm z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$. This means that

$z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} = 0.05.$

$$
\begin{aligned}
z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.05 \\
2.58\sqrt{\frac{0.25 \times 0.75}{n}} &= 0.05 \\
\sqrt{\frac{0.1875}{n}} &= 0.0194 \\
\frac{0.1875}{n} &= 3.756 \times 10^{-4} \\
n &= \frac{0.1875}{3.756 \times 10^{-4}} \\
&= 497.62
\end{aligned}
$$

A sample of size 498 is needed.

15 99% confidence interval means that $z = 2.58$

$\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.94\,(94\%)$

The confidence interval is $\hat{p} \pm z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$. This means that

$z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} = 0.04.$

$$
\begin{aligned}
z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.04 \\
2.58\sqrt{\frac{0.94 \times 0.06}{n}} &= 0.04 \\
\sqrt{\frac{0.0564}{n}} &= 0.0155 \\
\frac{0.0564}{n} &= 2.404 \times 10^{-4} \\
n &= \frac{0.0564}{2.404 \times 10^{-4}} \\
&= 234.6
\end{aligned}
$$

A sample of size 235 is needed.

16
$$
\begin{aligned}
M &= 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\
0.03 &= 1.96\sqrt{\frac{0.15 \times 0.85}{n}} \\
n &= 544
\end{aligned}
$$

The sample size needed is 544 people.

17 $\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.90$

The confidence interval is $\hat{p} \pm z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$. This means that

$z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} = 0.05.$

$$
\begin{aligned}
z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.05 \\
z\sqrt{\frac{0.9 \times 0.1}{100}} &= 0.05 \\
0.03z &= 0.05 \\
z &= 1.67
\end{aligned}
$$

$P(-1.67 < z < 1.67) = 0.9$

Bentons are 90% sure of their claim.

18 $\hat{p}$ will be at the centre of the interval, $\hat{p} = 0.775$

The confidence interval is $\hat{p} \pm z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$. This means that

$z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} = 0.025.$

$$
\begin{aligned}
z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0.025 \\
z\sqrt{\frac{0.775 \times 0.225}{250}} &= 0.025 \\
0.026z &= 0.025 \\
z &= 0.947
\end{aligned}
$$

$P(-0.947 < z < 0.947) = 0.66$

The Brisbane Lions are 66% sure of their claim.

13.5 Review: exam practice

1 b is a population parameter

2 a Population parameter

b Population parameter

c Sample statistic

3 $X = 132;\ n = 150$

$$
\begin{aligned}
\hat{p} &= \frac{X}{n} \\
&= \frac{132}{150} \\
\hat{p} &= 0.88
\end{aligned}
$$

4 a Systematic sample

b Yes, as the clients are likely to be in a random order.

5 $p = 0.25;\ n = 30$

$$
\begin{aligned}
\sigma_{\hat{p}} &= \sqrt{\frac{p(1-p)}{n}} \\
&= \sqrt{\frac{0.25(1-0.25)}{30}} \\
&= \sqrt{\frac{0.1875}{30}} \\
\sigma_{\hat{p}} &= 0.079
\end{aligned}
$$

The standard deviation of $\hat{p}$ is 0.079

6 a As the confidence interval is symmetric about $\hat{p}$,

$$
\begin{aligned}
\hat{p} &= \frac{\textit{upper limit} + \textit{lower limit}}{2} \\
&= \frac{0.58 + 0.66}{2} \\
\hat{p} &= 0.62
\end{aligned}
$$

b
$$
\begin{aligned}
E &= \left(\hat{p} + z\sigma_{\hat{p}}\right) - \hat{p} \\
&= 0.66 - 0.62 \\
&= 0.04
\end{aligned}
$$

The margin of error is 0.04

7 a Find the total number of students.

Gender	Middle School	Senior School
Male	253	342
Female	287	323
Total	540	665

$$
\begin{aligned}
N &= 540 + 665 \\
&= 1205
\end{aligned}
$$

There are 1205 students in the population

b $n = 50$

c

Gender	Middle School	Senior School
Male	$\frac{253}{1205} \times 50 = 10.5$ 11 students	$\frac{342}{1205} \times 50 = 14.1$ 14 students
Female	$\frac{287}{1205} \times 50 = 11.9$ 12 students	$\frac{323}{1205} \times 50 = 13.4$ 13 students

$\therefore$ The sample should contain 11 male and 12 female Middle school students and 14 male and 13 female Senior school students.

8 $N = 423$; $n = 52$; $X = 52 - 23 = 29$

$$\hat{p} = \frac{X}{n} = \frac{29}{52}$$
$$= 0.56$$

9 a $\mu_{\hat{p}} = p = 0.37$

b $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

$$= \sqrt{\frac{0.37(1-0.37)}{120}}$$
$$= 0.044$$

10 $\hat{p} = \frac{65+75}{2}$

$$= 70\%$$
$$= 0.7$$

Confidence level of 95% means $P(Z \le z) = 0.025$
using graphics calculator: InvNorm
(0.025, 0, 1, Left) $z = 1.96$
Using the lower confidence level:

$$0.65 = \hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$0.65 = 0.7 - 1.96\sqrt{\frac{0.7(1-0.7)}{n}}$$
$$\sqrt{\frac{0.21}{n}} = \frac{0.05}{1.96}$$
$$\frac{0.21}{n} = (0.0255)^2$$
$$n = 322.69$$

The sample size needed was 323

11 a $N = 52\,000$

b $n = \frac{52\,000}{25} = 2080$

c $X = 1600$

$$\hat{p} = \frac{X}{n}$$
$$= \frac{1600}{2080}$$
$$= \frac{10}{13} = 0.77$$

12 For large samples, $\hat{p}$ has an approximately normal distribution. This is best presented by B.

Complex familiar

13 a $\hat{p} = \frac{10}{100} = 0.1$

b For the 85% confidence level, $z = 1.44$ (InvNorm (0.075, 0, 1, Left))

$$\hat{p} \pm z\sigma_{\hat{p}} = \hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.1 \pm 1.44\sqrt{\frac{0.1(1-0.1)}{100}}$$
$$= 0.1 \pm 1.44 \times 0.03$$
$$= 0.1 \pm 0.0432$$

c $E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$$= 1.44\sqrt{\frac{0.1(1-0.1)}{100}}$$
$$= 0.0432$$

d As $E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ then $E \propto \frac{1}{\sqrt{n}}$ or, for a given distribution, $E\sqrt{n} = k$ where k is some constant.
As $E_{100}\sqrt{100} = k$ and $E_{50}\sqrt{50} = k$,

$$E_{100}\sqrt{100} = E_{50}\sqrt{50}$$
$$\Rightarrow E_{50} = \frac{E_{100}\sqrt{100}}{\sqrt{50}}$$
$$= E_{100}\sqrt{2}$$

Therefore, decreasing the sample to 50 increases the margin of error by a factor of $\sqrt{2}$.

14 $n = 50$; $\sigma_{\hat{p}} = 0.05$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$
$$0.05 = \sqrt{\frac{p(1-p)}{50}}$$
$$0.0025 = \frac{p(1-p)}{50}$$
$$0.125 = p(1-p)$$
$$0 = p^2 - p + 0.125$$

Solving for p using the quadratic solution:

$$p = \frac{1 \pm \sqrt{(-1)^2 - 4(1)(0.125)}}{2}$$
$$= \frac{1 \pm \sqrt{(0.5)}}{2}$$

$p = 0.15$ or 0.85

15 $n = 100$

$$\hat{p} = \frac{0.85 + 0.9}{2}$$
$$= 0.875$$
$$\hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.9$$
$$0.875 + z\sqrt{\frac{0.875(0.125)}{100}} = 0.9$$
$$z\sqrt{1.093\,75 \times 10^{-3}} = 0.025$$
$$z = 0.756$$

$P(-0.756 \le Z \le 0.756) = 0.55$
There is a 55% likelihood that the proportion will lie in this interval.

16 $z = 2.58$

$$\hat{p} = \frac{0.67 + 0.83}{2} = 0.75$$

$$\hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.83$$

$$0.75 + 2.58\sqrt{\frac{0.75\,(0.25)}{n}} = 0.83$$

$$2.58\sqrt{\frac{0.1875}{n}} = 0.08$$

$$\sqrt{\frac{0.1875}{n}} = 0.031$$

$$\frac{0.1875}{n} = 9.61 \times 10^{-4}$$

$$n = 195$$

Krypton Industries surveyed 195 people.

Complex unfamilar

17 $n = 50$

$\hat{p} = 87\% = 0.87$

For the 90% confidence level, $z = 1.64$ (using graphics calculator InvNorm (0.05, 0, 1, Left))

$$z\sigma_{\hat{p}} = z\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = 1.64\sqrt{\frac{0.87\,(1-0.87)}{50}} = 0.078$$

$\hat{p} - z\sigma_{\hat{p}} = 0.87 - 0.078 = 0.792$

$\hat{p} + z\sigma_{\hat{p}} = 0.87 + 0.078 = 0.948$

$C.I = (0.79, 0.95)$

We can be 90% confident that between 79% and 95% of drivers would rate the Twelve Apostles as the highlight of their drive.

18 $p = 0.4$

$n = 400$

$N = \textit{population of Australia} \approx 24\,\textit{million}$

As $np = 400 \times 0.4 = 160 > 10$

and $10n = 10 \times 400 = 4\,000 < N$

$\Rightarrow$ this can be treated as a large sample.

$\mu_{\hat{p}} = p = 0.4$

$$\sigma_{\hat{p}} = \sqrt{\frac{p\,(1-p)}{n}} = \sqrt{\frac{0.4\,(1-0.4)}{400}} = 0.02449$$

$$z = \frac{0.45 - 0.4}{0.024\,49} = 2.041$$

$$\begin{aligned} P(\hat{p} > 0.45) &= P(z > 2.041) \\ &= 1 - P(z \le 2.041) \\ &= 1 - 0.9794 \text{ (using the normal cdf)} \\ &= 0.0206 \end{aligned}$$

19 $n = 100$

For 95% confidence level, $z = 1.96$

$$0.13 = \hat{p} - z\sigma_{\hat{p}}$$

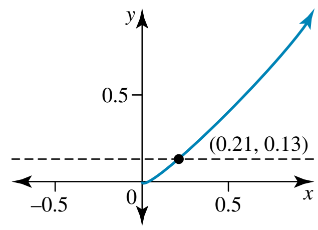
$$0.13 = \hat{p} - z\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$$

$$0.13 = \hat{p} - 1.96\sqrt{\frac{\hat{p}\,(1-\hat{p})}{100}}$$

$$0.13 = \hat{p} - 1.96\frac{\sqrt{\hat{p}\,(1-\hat{p})}}{10}$$

$$0.13 = \hat{p} - 0.196\sqrt{\hat{p}\,(1-\hat{p})}$$

Solve graphically, finding the intersection of $y = 0.13$ and $y = x - 0.196\sqrt{x - x^2}$:



Intersection occurs at $x = 0.21 \Rightarrow \hat{p} = 0.21$

Therefore, the sample proportion is 0.21

20 For a confidence level of 95%, $z = 1.96$

a Using Breanna's method of averaged individual intervals:

Kayley's data:

$n = 100; X = 20$

$$\hat{p} = \frac{20}{100} = 0.2$$

$$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.2\,(1-0.2)}{100}} = 0.04$$

$\hat{p} - z\sigma_{\hat{p}} = 0.2 - 1.96 \times 0.04 = 0.1216$

$\hat{p} + z\sigma_{\hat{p}} = 0.2 + 1.96 \times 0.04 = 0.2784$

Kayley's confidence interval is $(0.12, 0.28)$

Breanna's data:

$n = 100; X = 23$

$$\hat{p} = \frac{23}{100} = 0.23$$

$$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.23\,(1-0.23)}{100}} = 0.042$$

$\hat{p} - z\sigma_{\hat{p}} = 0.23 - 1.96 \times 0.042 = 0.1475$

$\hat{p} + z\sigma_{\hat{p}} = 0.23 + 1.96 \times 0.042 = 0.3125$

Breanna's confidence interval is $(0.15, 0.31)$

Teagan's data:

$n = 100; X = 19$

$$\hat{p} = \frac{19}{100} = 0.19$$

$$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.19\,(1-0.19)}{100}} = 0.039$$

$\hat{p} - z\sigma_{\hat{p}} = 0.19 - 1.96 \times 0.039 = 0.1131$

$\hat{p} + z\sigma_{\hat{p}} = 0.19 + 1.96 \times 0.039 = 0.2669$

Teagan's confidence interval is $(0.11, 0.27)$

Average lower confidence limit =

$$\frac{12.16 + 14.74 + 11.31}{3} = 12.74\%$$

Average upper confidence limit =

$$\frac{27.84 + 31.25 + 26.69}{3} = 28.59\%$$

Using Kayley's method of combined data:
$n = 300;$
$X = (0.23 + 0.2 + 0.19)\,100 = 62$

$$\hat{p} = \frac{62}{300} = 0.21$$

$$\begin{aligned}\sigma_{\hat{p}} &= \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\\ &= \sqrt{\frac{0.21\,(1-0.21)}{300}}\\ &= 0.0235\end{aligned}$$

$\hat{p} - z\sigma_{\hat{p}} = 0.21 - 1.96 \times 0.0235 = 0.1639$
$\hat{p} + z\sigma_{\hat{p}} = 0.21 + 1.96 \times 0.0235 = 0.2561$

$Combined\ data\ lower\ interval\ limit = 16.39\%$
$Combined\ data\ upper\ interval\ limit = 25.61\%$
$C.I. = (0.16, 0.26)$

As the interval limits for Breanna's and Kayley's methods are not the same, Teagan is incorrect.

b Kayley's method is more reliable as, despite having three survey takers, 300 people were actually sampled in total. As a larger sample size is more likely to have similar proportions to the population, the confidence interval can be smaller.

c The best estimate of the population parameter is 16%–26%.