EXERCISE 27 <SELECTED QUESTIONS>

① a) $p = \frac{58}{100}$

$= 0.58$

b) $\hat{p} = \frac{65}{100}$

$= 0.65$

⑤ a)

x	$\hat{p}$	$Pr(X=x) = Pr(\hat{P}=\hat{p})$
0	0	$\frac{^{47}C_0 \times {}^{73}C_4}{^{120}C_4} = 0.1325$ <4 dp>
1	$\frac{1}{4}$	$\frac{^{47}C_1 \times {}^{73}C_3}{^{120}C_4} = 0.3559$ <4dp>
2	$\frac{1}{2}$	$\frac{^{47}C_2 \times {}^{73}C_2}{^{120}C_4} = 0.3458$ <4dp>
3	$\frac{3}{4}$	$\frac{^{47}C_3 \times {}^{73}C_1}{^{120}C_4} = 0.1441$ <4dp>
4	1	$\frac{^{47}C_4 \times {}^{73}C_0}{^{120}C_4} = 0.0217$ <4dp>

b) $n = 4$, population size $= 120$

$n < \frac{\text{population size}}{10}$

∴ Population is 'large' in comparison to sample size and we can approximate $\hat{p}$ well by assuming it has binomial distribution.

Let X^* be a binomial random variable with $n = 4$, $p = \frac{47}{120}$ and $\hat{p}^* = \frac{X^*}{4}$

x	$\hat{p}$	$Pr(X^*=x) = Pr(\hat{P}^*=\hat{p})$
0	0	$^4C_0\left(\frac{47}{120}\right)^0\left(\frac{73}{120}\right)^4 = 0.1370$ <4dp>
1	$\frac{1}{4}$	$^4C_1\left(\frac{47}{120}\right)^1\left(\frac{73}{120}\right)^3 = 0.3527$ <4dp>
2	$\frac{1}{2}$	$^4C_2\left(\frac{47}{120}\right)^2\left(\frac{73}{120}\right)^2 = 0.3406$ <4dp>
3	$\frac{3}{4}$	$^4C_3\left(\frac{47}{120}\right)^3\left(\frac{73}{120}\right)^1 = 0.1462$ <4dp>
4	1	$^4C_4\left(\frac{47}{120}\right)^4\left(\frac{73}{120}\right)^0 = 0.0235$ <4dp>

⑥ a) $\frac{^{3000}C_{12} \times {}^{7000}C_{13}}{^{10000}C_{25}} = 0.0267$ <4dp>

b) $^{25}C_{12} \times \left(\frac{3000}{10000}\right)^{12} \times \left(\frac{7000}{10000}\right)^{13} = 0.0268$ <4dp>

⑦ a) $E(\hat{p}) = p = \frac{75}{100} = 0.75$

b)
$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$
$$= \sqrt{\frac{0.75(1-0.75)}{25}}$$
$$= 0.0866$$ <4dp>

⑫ a) $n = 200$

$\hat{p} = \frac{37}{200} = 0.185$

$n\hat{p} = 200 \times 0.185 = 37$

$n(1-\hat{p}) = 200(1-0.185) = 163$

Both $n\hat{p}$ and $n(1-\hat{p})$ are greater than 5, we can use z_k-value obtained from normal tables.

$k = 90 \Rightarrow Z_k = 1.65$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$
$$= \left(\frac{37}{200} - 1.65\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}, \ \frac{37}{200} + 1.65\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}\right)$$
$$= (0.1396963306, \ 0.2303036694)$$
$$\approx (0.140, \ 0.230)$$ <3dp>

b) $k = 95 \Rightarrow Z_k = 1.96$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$
$$= \left(\frac{37}{200} - 1.96\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}, \ \frac{37}{200} + 1.96\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}\right)$$
$$= (0.1311847029, \ 0.2388152971)$$
$$\approx (0.131, \ 0.239)$$ <3dp>

c) $k = 99 \Rightarrow Z_k = 2.58$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$
$$= \left(\frac{37}{200} - 2.58\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}, \ \frac{37}{200} + 2.58\sqrt{\frac{\frac{37}{200}(1-\frac{37}{200})}{200}}\right)$$

$= (0.1141614967, 0.2558385033)$

$\approx (0.114, 0.256)$ (3dp)

EXERCISE SHEET 27

① <Selected Question>

② a) population proportion, p //

b) sample proportion, $\hat{p}$ //

③ No, we cannot have $\hat{p}=1$ because there are only 7 peppermint flavoured chocolate bars in total. //

④ a) Hypergeometric ~ if the selection of people to do the survey was done <u>without replacement</u> //

b) Binomial ~ if the selection of people to do the survey was done <u>with replacement</u>.
Approximated with Binomial ~ if the selection of people to do the survey was done <u>without replacement</u> AND sample size is small $\left(n < \frac{\text{population size}}{10}\right)$ //

c) Let $\hat{p}$ be the sample proportion of people who travel to Melbourne to work.
Let n be the sample size
$\hat{p}$ is approximately normal ~ if $n\hat{p} > 15$ and $n(1-\hat{p}) > 15$ //

⑤ <Selected question>

⑥ <Selected question>

⑦ <Selected question>

⑧ a) 5 samples of 10 : 0.3, 0.2, 0.5, 0.3, 0.1

$\Rightarrow$ mean, $\mu = \frac{0.3+0.2+0.5+0.3+0.1}{5}$

$= 0.28$ //

$\Rightarrow$ standard deviation, $\sigma = 0.1483239697$ <Used calculator>

≈ 0.15 //

5 samples of 100 : 0.25, 0.32, 0.27, 0.35, 0.31

$\Rightarrow$ mean, $\mu = \frac{0.25+0.32+0.27+0.35+0.31}{5}$

$= 0.3$

$\Rightarrow$ standard deviation $= 0.04$ <Used calculator>

b) i) $E(\hat{p}) = p = 0.31$

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

$$= \sqrt{\frac{0.31(1-0.31)}{10}}$$

$= 0.15$ <2dp>

ii) $E(\hat{p}) = 0.31$

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

$$= \sqrt{\frac{0.31(1-0.31)}{100}}$$

$= 0.05$ <2dp>

iii) The mean found in (a) with the larger sample size was closer to the long-term average found in (b) than the mean of the data with the smaller sample-size.
The standard deviation found in (a) with the larger sample size was smaller than the standard deviation of the data with the smaller sample size.
$\therefore$ The larger sample-size tends to lead to sample proportions which are clustered closer to the mean

9) $n = 200$

$$\hat{p} = \frac{200-40}{200} = 0.8$$

$n\hat{p} = 200 \times 0.8 = 160$ $\qquad$ $n(1-\hat{p}) = 200(1-0.8) = 40$

$n\hat{p}$ and $n(1-\hat{p})$ are greater than 15, we can use z_k values obtained from normal tables.

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.8 - 1.96\sqrt{\frac{0.8(1-0.8)}{200}},\ 0.8 + 1.96\sqrt{\frac{0.8(1-0.8)}{200}}\right)$$

$= (0.7446, 0.8554)$ <4dp>

10 $n = 1000$

$\hat{p} = 0.38$

$n\hat{p} = (1000)(0.38) = 380$ $n(1-\hat{p}) = (1000)(1-0.38) = 620$

$n\hat{p}$ and $n(1-\hat{p})$ are larger than 15, we can use Z_k values obtained from the normal tables.

$$\left(\hat{p} - Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.38 - 1.96\sqrt{\frac{0.38(1-0.38)}{1000}},\ 0.38 - 1.96\sqrt{\frac{0.38(1-0.38)}{1000}}\right)$$

$= (0.35, 0.41)$ <2dp> //

11 a) $\hat{p} = \frac{84}{100} = 0.84$

b) $n = 100$

$\hat{p} = 0.84$

$n\hat{p} = (100)(0.84) = 84$ $n(1-\hat{p}) = 100(1-0.84) = 16$

$n\hat{p}$ and $n(1-\hat{p})$ are greater than 15, we can use Z_k values from the normal table.

$$\left(\hat{p} - Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.84 - 1.96\sqrt{\frac{0.84(1-0.84)}{100}},\ 0.84 + 1.96\sqrt{\frac{0.84(1-0.84)}{100}}\right)$$

$= (0.76815, 0.91185)$ <5dp>

c) If many samples were considered and used to calculate many 95% confidence intervals in this same way, then approximately 95% of those confidence intervals will include the population proportion.

d) <Not examinable>

12 $n = 200$

$\hat{p} = \frac{37}{200} = 0.185$

$n\hat{p} = (200)(0.185)$ $\qquad$ $n(1-\hat{p}) = (200)(1-0.185)$

$= 37$ $\qquad\qquad$ $= 163$

$n\hat{p}$ and $n(1-\hat{p})$ are greater than 15, we can use z_k values from normal table

a) 90% confidence interval, $z_k = 1.65$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.185 - 1.65\sqrt{\frac{0.185(1-0.185)}{200}},\ 0.185 + 1.65\sqrt{\frac{0.185(1-0.185)}{200}}\right)$$

$$= (0.140, 0.230) \ \langle 3dp \rangle //$$

b) 95% confidence interval, $z_k = 1.96$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.185 - 1.96\sqrt{\frac{0.185(1-0.185)}{200}},\ 0.185 + 1.96\sqrt{\frac{0.185(1-0.185)}{200}}\right)$$

$$= (0.131, 0.239) \ \langle 3dp \rangle //$$

c) 99% confidence interval, $z_k = 2.58$

$$\left(\hat{p} - z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.185 - 2.58\sqrt{\frac{0.185(1-0.185)}{200}},\ 0.185 + 2.58\sqrt{\frac{0.185(1-0.185)}{200}}\right)$$

$$= (0.114, 0.256) \ \langle 3dp \rangle //$$

d) $p = 0.24$ is only in the 99% confidence interval. The higher the percentage is, the wider the confidence interval.

13 a) The sample proportion is the midpoint of the confidence interval:

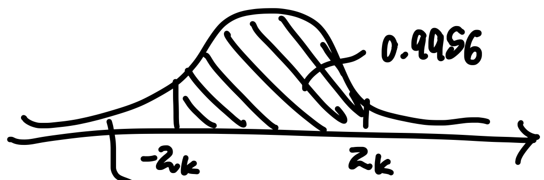
$$\hat{p} = \frac{0.2835 + 0.3165}{2} = 0.3 //$$

b) $$0.3 - 1.65\sqrt{\frac{0.3(1-0.3)}{n}} = 0.2835$$

$$n = 2100 //$$

(14) a) $\hat{p} = \dfrac{0.743 + 0.857}{2} = 0.8$ //

b) $Pr(-z_k < Z < z_k) = 0.9956$



$\dfrac{1 - 0.9956}{2} = 0.0022$

<Check normal table>

$-z_k = -2.85$

$z_k = 2.85$ //

c) $0.8 - 2.85\sqrt{\dfrac{0.8(1-0.8)}{n}} = 0.743$

$n = 400$ //

(15) a) $n = 1000$

$\hat{p} = \frac{368}{1000} = 0.368$

95% CI 1.96

$\left(\hat{p} - z_k\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}, \ \hat{p} + z_k\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right)$

$= \left(0.368 - 1.96\sqrt{\dfrac{0.368(1-0.368)}{1000}}, \ 0.368 + 1.96\sqrt{\dfrac{0.368(1-0.368)}{1000}}\right)$

$= (0.338, 0.398)$ <3dp> //

b) $n = 3786$

$k\%$ CI $(0.33, 0.37)$

$\hat{p} = \dfrac{0.33 + 0.37}{2} = 0.35$

$0.35 - z_k\sqrt{\dfrac{0.35(1-0.35)}{3786}} = 0.33$

$z_k = 2.5801$ <4dp>

<Check normal table> $k = 0.9951$

$k\% = 99\%$

(16) a) $n = 500$

$\hat{p} = \frac{120}{500} = 0.24$

90% CI, $Z_k = 1.65$

$$\left(\hat{p} - Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$$= \left(0.24 - 1.65\sqrt{\frac{0.24(1-0.24)}{500}},\ 0.24 + 1.65\sqrt{\frac{0.24(1-0.24)}{500}}\right)$$

$= (0.208, 0.272)$ <3dp> //

b) $1.65\sqrt{\frac{0.24(1-0.24)}{n}} = 0.01$

$n = 4966$ //

(17) $\hat{p} = 20\% = 0.2$

$$Z_k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq 0.01$$

$$1.65\sqrt{\frac{0.2(1-0.2)}{n}} \leq 0.01$$

$n \geq 4356$

The farmer inspected at least 4356 apples.

(18) 90% CI, $Z_k = 1.65$

$$\hat{p} = \frac{0.35 + 0.45}{2}$$

$= 0.4$

$$0.4 - 1.65\sqrt{\frac{0.4(1-0.4)}{n}} = 0.35$$

$n = 261.36$

≈ 262 <Rounded to next integer> //

(19) a) $\hat{p} = \frac{0.29277 + 0.56723}{2}$

$= 0.43$ //

b) 95% CI, $z_K = 1.96$

$$0.43 - 1.96\sqrt{\frac{0.43(1-0.43)}{n}} = 0.29277$$

$$n = 49.99854092$$

$$\approx 50$$