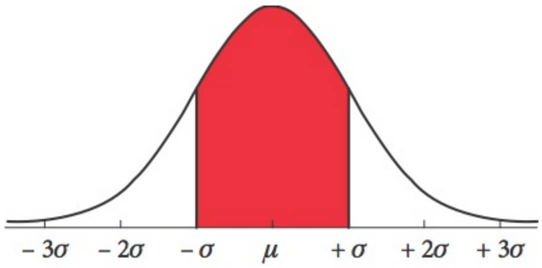
Statistics IV – Inferential Statistics

Practice Questions 25.1

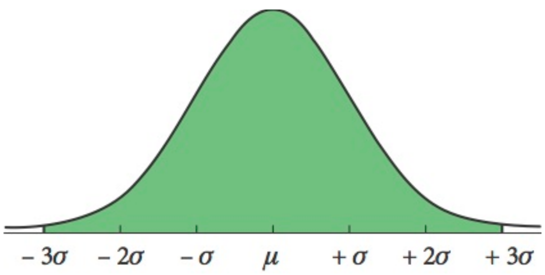
1. What percentage of the following normal distribution curves is shaded?

(i)



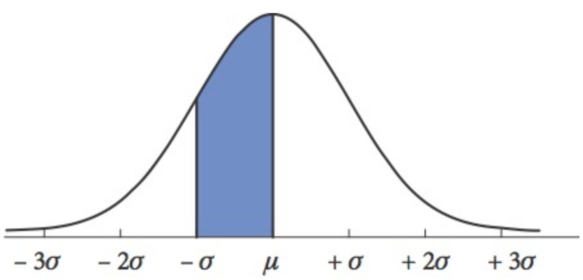
$\mu \pm \sigma = 68\%$

(ii)



$\mu \pm 3\sigma = 99{\cdot}7\%$

(iii)

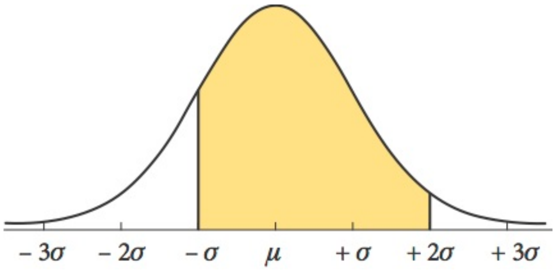


$\mu \pm \sigma = 68\%$.

Since the curve is symmetrical,

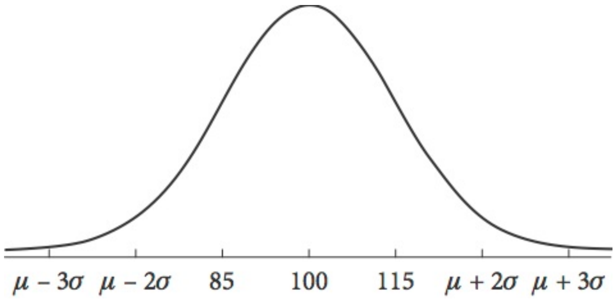
$\mu - \sigma = \dfrac{68}{2} = 34\%$

(iv)



81·5% (68% + 13·5%)

2. The normal curve shown below represents the distribution of IQ scores for a population.



(i) What is the mean?

Mean = 100 (Centre of normal curve)

(ii) What is the standard deviation?

Standard deviation = 115 − 100 = 15

(iii) Calculate the value of $\mu - 2\sigma$.

$$\begin{aligned} \mu - 2\sigma &= 100 - 2(15) \\ &= 100 - 30 \\ &= 70 \end{aligned}$$

(iv) What percentage of the population have an IQ score between 85 and 115?

$85 \rightarrow 115 = \mu - \sigma \rightarrow \mu + \sigma$

$= 68\%$

3. Given an approximately normal distribution, what percentage of all values are:

(i) above the mean?

50% (the normal curve is symmetrical)

(ii) below the mean?

50% (the normal curve is symmetrical)

(iii) within one standard deviation from the mean?

68%

(iv) within two standard deviations from the mean?

95%

(v) within three standard deviations from the mean?

99·7%

and

(vi) what interval contains 99·7% of all values?

$\mu - 3\sigma \longrightarrow \mu + 3\sigma$

4. For each of the following, find the interval in which **(a)** 68%, **(b)** 95% and **(c)** 99·7% of data lies.

(i) Mean = 16 Standard deviation = 3

(a) $68\% = \mu \pm \sigma$

$= 16 - 3 \rightarrow 16 + 3$

$= 13 \rightarrow 19$

(b) $95\% = \mu \pm 2\sigma$

$= 16 - 6 \rightarrow 16 + 6$

$= 10 \rightarrow 22$

(c) $99{\cdot}7\% = \mu \pm 3\sigma$

$= 16 - 9 \rightarrow 16 + 9$

$= 7 \rightarrow 25$

(ii) Mean = 5·2 Standard deviation = 0·8

(a) $68\% = \mu \pm \sigma$

$= 5\cdot2 - 0\cdot8 \rightarrow 5\cdot2 + 0\cdot8$

$= 4\cdot4 \rightarrow 6$

(b) $95\% = \mu \pm 2\sigma$

$= 5\cdot2 - 1\cdot6 \rightarrow 5\cdot2 + 1\cdot6$

$= 3\cdot6 \rightarrow 6\cdot8$

(c) $99\cdot7\% = \mu \pm 3\sigma$

$= 5\cdot2 - 2\cdot4 \rightarrow 5\cdot2 + 2\cdot4$

$= 2\cdot8 \rightarrow 7\cdot6$

(iii) Mean = 148 Standard deviation = 6

(a) $68\% = \mu \pm \sigma$

$= 148 - 6 \rightarrow 148 + 6$

$= 142 \rightarrow 154$

(b) $95\% = \mu \pm 2\sigma$

$= 148 - 12 \rightarrow 148 + 12$

$= 136 \rightarrow 160$

(c) $99\cdot7\% = \mu \pm 3\sigma$

$= 148 - 18 \rightarrow 148 + 18$

$= 130 \rightarrow 166$

(iv) Mean = 68 Standard deviation = 4·6

(a) $68\% = \mu \pm \sigma$

$= 68 - 4\cdot6 \rightarrow 68 + 4\cdot6$

$= 63\cdot4 \rightarrow 72\cdot6$

(b) $95\% = \mu \pm 2\sigma$

$= 68 - 9\cdot2 \rightarrow 68 + 9\cdot2$

$= 58\cdot8 \rightarrow 77\cdot2$

(c) $99{\cdot}7\% = \mu + 3\sigma$

$= 68 - 13{\cdot}8 \rightarrow 68 + 13{\cdot}8$

$= 54{\cdot}2 \rightarrow 81{\cdot}8$

(v) Mean = 250 Standard deviation = $31{\cdot}5$

(a) $68\% = \mu \pm \sigma$

$= 250 - 31{\cdot}5 \rightarrow 250 + 31{\cdot}5$

$= 218{\cdot}5 \rightarrow 281{\cdot}5$

(b) $95\% = \mu \pm 2\sigma$

$= 250 - 63 \rightarrow 250 + 63$

$= 187 \rightarrow 313$

(c) $99{\cdot}7\% = \mu \pm 3\sigma$

$= 250 - 94{\cdot}5 \rightarrow 250 + 94{\cdot}5$

$= 115{\cdot}5 \rightarrow 344{\cdot}5$

5. The heights of male students is normally distributed with a mean of 170 cm and a standard deviation of 8 cm. Find the percentage of male students whose height is:

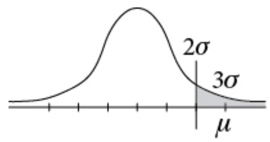
Mean = 170 cm (μ) Standard deviation = 8 cm (σ)

(i) greater than 186 cm

$186 = 170 + 16$

$170 + 2(8)$

This means that greater than 186 cm is the area above $\mu + 2\sigma$



Percentage above 2σ $= 2{\cdot}35\% + 0{\cdot}15\%$ (from graph)

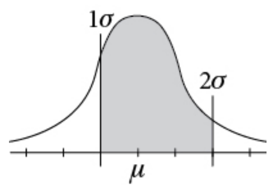
$= 2{\cdot}5\%$

(ii) between 162 cm and 186 cm

$170 - 162 = 8$

$162 = \mu - \sigma$

The area between 162 and 186 is the area from $\mu - \sigma \rightarrow \mu + 2\sigma$



Percentage from $\mu - \sigma \rightarrow \mu + 2\sigma$ $= 68\% + 13{\cdot}5\%$ (From graph)

$= 81{\cdot}5\%$

(iii) between 154 cm and 178 cm.

$170 - 154 = 16$

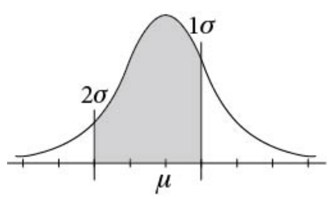
$= 2(8)$

$154 = \mu - 2\sigma$

$178 - 170 = 8$

$178 = \mu + \sigma$

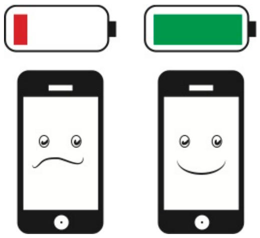
The area between 154 and 178 is the area from $\mu - 2\sigma \rightarrow \mu + \sigma$



Percentage from $\mu - 2\sigma \rightarrow \mu + \sigma$ $= 13{\cdot}5\% + 68\%$

$= 81{\cdot}5\%$

6. The mean battery life of a mobile phone battery is 14 hours. Assuming that the battery life is normally distributed and the standard deviation is two hours,



Mean = 14 hours (μ) Standard deviation = 2 hours (σ)

(i) what percentage of the batteries have a life of between 12 and 18 hours?

12 − 18 hours:

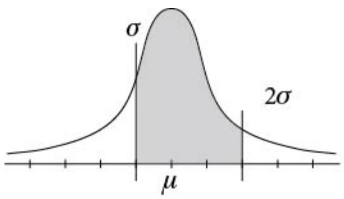
12: 14 − 12 = 2 hours = σ

$12 = \mu - \sigma$

18: 18 − 14 = 4 hours = 2σ

$18 = \mu + 2\sigma$

The area between 12 and 18 hours is the area from $\mu - \sigma \rightarrow \mu + 2\sigma$



$12 \rightarrow 18$ hours $= \mu - \sigma \rightarrow \mu + 2\sigma$

$= 68\% + 13{\cdot}5\%$

$= 81{\cdot}5\%$

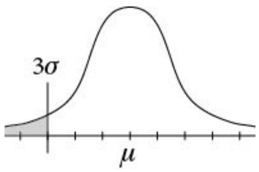
Batteries that have a life of less than eight hours are rejected. On a particular day a mobile battery production unit makes 1,200 batteries.

(ii) How many batteries would you expect to be rejected on this day?

Less than 8 hours:

$14 - 8 = 6 = 3(2)$

$8 = \mu - 3\sigma$



Percentage < 8 hours = 0·15%

1,200 batteries a day

0·15% of 1,200 = 1·8 batteries rejected per day.

7. The results of 20 students in their end-of-term maths exam were:

45, 56, 67, 85, 28, 64, 72, 83, 48, 96, 47, 71, 59, 63, 54, 66, 79, 80, 42, 50

(i) Find the mean of this data.

$$\text{Mean} = \frac{\text{sum of numbers}}{\text{number of numbers}}$$

$$= \frac{45+56+67+85+28+64+72+83+48+96+47+71+59+63+54+66+79+80+42+50}{20}$$

$$= \frac{1255}{20}$$

$$= 62{\cdot}75$$

(ii) Find the standard deviation to 2 decimal places.

Standard deviation: 16·54 (by calculator)

(iii) Assuming that the results are normally distributed, find the interval in which:

(a) 68% of the data lies

68% = $\mu - \sigma \rightarrow \mu + \sigma$

$\mu = 62{\cdot}75$ $\sigma = 16{\cdot}54$

$\mu - \sigma = 62{\cdot}75 - 16{\cdot}54$

$= 46{\cdot}21$

$\mu + \sigma = 62{\cdot}75 + 16{\cdot}54$

$= 79{\cdot}29$

68% of the data lies between 46·21 → 79·29

(b) 95% of the data lies

95% = $\mu - 2\sigma \rightarrow \mu + 2\sigma$

$\mu - 2\sigma = 62{\cdot}75 - 2(16{\cdot}54)$

$= 29{\cdot}67$

$\mu + 2\sigma = 62{\cdot}75 + 2(16{\cdot}54)$

$= 95{\cdot}83$

95% of the data lies between 29·67 → 95·83

(c) 99·7% of the data lies.

99·7% = $\mu - 3\sigma \rightarrow \mu + 3\sigma$

$\mu - 3\sigma = 62{\cdot}75 - 3(16{\cdot}54)$

$= 13{\cdot}13$

$\mu + 3\sigma = 62{\cdot}75 + 3(16{\cdot}54)$

$= 112{\cdot}37$

99·7% of the data lies between 13·13 → 112·37

8. At a certain IT company, the ages of all new employees hired during the last five years are normally distributed. Within this curve, 95% of the ages are between 24·6 and 37·4 years.
Find the mean age and the standard deviation of the data.

95% $= 24{\cdot}6 - 37{\cdot}4$

$= \mu - 2\sigma \rightarrow \mu + 2\sigma$

$24{\cdot}6 = \mu - 2\sigma$ ① $37{\cdot}4 = \mu + 2\sigma$ ②

Solve these equations simultaneously:

$$24\cdot 6 = \mu - \cancel{2\sigma} \quad ①$$

$$\underline{37\cdot 4 = \mu + \cancel{2\sigma}} \quad ②$$

$62 = 2\mu$ (divide both sides by 2)

$31 = \mu$

① $24{\cdot}6 = (31) - 2\sigma$ (substitute μ into equation ①)

$2\sigma = 31 - 24{\cdot}6$

$2\sigma = 6{\cdot}4$

$\sigma = 3{\cdot}2$

9. Given $\mu + 1\sigma = 247$ and $\mu + 2\sigma = 428$ find:

(i) the standard deviation

① $\mu + 1\sigma = 247$ ② $\mu + 2\sigma = 428$

Solve these equations simultaneously:

① $\cancel{\mu} + \sigma = 247$

$-$② $\underline{-\cancel{\mu} - 2\sigma = -428}$

$-\sigma = -181$

$\sigma = 181$

(ii) the mean

① $\mu + (181) = 247$ (substitute σ into equation ①)

$\mu = 247 - 181$

$\mu = 66$

(iii) the approximate percentage of the distribution that lies between 66 and 428.

$66 - 428 = \%?$

66: $= \mu$

428: $428 - 66$

$= 362$

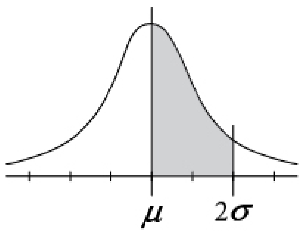
$= 2(181)$

$= 2\sigma$

$\mu \rightarrow \mu + 2\sigma$

$34\% + 13 \cdot 5\%$

$= 47 \cdot 5\%$



Practice Questions 25.2

1. Calculate the margin of error (correct to two decimal places) for the following samples:

(i) 150

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

$$= \frac{1}{\sqrt{150}}$$

$$= 0 \cdot 081$$

$$\simeq 0 \cdot 08$$

(ii) 800

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

$$= \frac{1}{\sqrt{800}}$$

$$= 0 \cdot 035$$

$$\simeq 0 \cdot 04$$

(iii) 1,600

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

$$= \frac{1}{\sqrt{1,600}}$$

$$= 0 \cdot 025$$

$$\simeq 0 \cdot 03$$

(iv) 3,600

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

$$= \frac{1}{\sqrt{3,600}}$$

$$= 0 \cdot 01\dot{6}$$

$$\simeq 0 \cdot 02$$

(v) 5,000

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

$$= \frac{1}{\sqrt{5,000}}$$

$$= 0 \cdot 014$$

$$\simeq 0 \cdot 01$$

2. In a random sample of machines, 20 out of 220 were found to have been damaged in shipment.

(i) What is the sample size?

Sample size = 220

(ii) What is the sample proportion, $\hat{p}$, to two decimal places?

Sample proportion $(\hat{p})$

$$= \frac{20}{220}$$
$$= 0 \cdot 09$$

(iii) What is the margin of error, to two decimal places?

Margin of error

$$= \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{220}}$$
$$= 0 \cdot 07$$

(iv) Find the 95% confidence interval.

95% confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$= 0{\cdot}09 - 0{\cdot}07 < p < 0{\cdot}09 + 0{\cdot}07$

$= 0{\cdot}02 < p < 0{\cdot}16$

(v) If you increase the sample size to 300, what effect would this have on the margin of error?

Sample size $= 300$

Margin of error

$$= \frac{1}{\sqrt{300}}$$
$$= 0 \cdot 057$$
$$\simeq 0{\cdot}06$$

The margin would decrease from 0·07 to 0·06 or 7% to 6%

3. An assembly line does a quality check by sampling 50 of its products. It finds that 16% of the parts are defective.

(i) What is the sample size?

Sample size = 50

(ii) What is the sample proportion, $\hat{p}$?

Sample proportion ($\hat{p}$) = 16% of 50 = 8

$$\frac{8}{50} = 0 \cdot 16$$

(iii) What is the margin of error, to two decimal places?

Margin of error:

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

$$= \frac{1}{\sqrt{50}}$$

$$= 0 \cdot 141$$

$$\simeq 0 \cdot 14$$

(iv) Find the 95% confidence interval.

95% confidence interval

$$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

= 0·16 − 0·14 < p < 0·16 + 0·14

= 0·02 < p < 0·3

(v) How could the company decrease the margin of error?

Increase the sample size.

4. 60% of a sample of 500 people leaving a shopping centre claim to have spent over €25.

(i) What is the sample size?

Sample size: 500

(ii) What is the sample proportion, $\hat{p}$?

Sample proportion ($\hat{p}$): 60% of 500 = 300

$$\frac{300}{500} = 0 \cdot 6$$

(iii) What is the margin of error, to two decimal places?

Margin of error:

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

$$= \frac{1}{\sqrt{500}}$$

$$= 0 \cdot 044$$

$$\simeq 0 \cdot 04$$

(iv) Find the 95% confidence interval.

95% confidence interval

$$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$= 0 \cdot 6 - 0 \cdot 04 < p < 0 \cdot 6 + 0 \cdot 04$$

$$= \quad 0 \cdot 56 < p < 0 \cdot 64$$

5. A nationwide poll was taken of 1,400 teenagers (ages 13–18). 600 of them said they have a TV in their bedroom.

(i) Construct the 95% confidence interval by calculating $\hat{p}$ and margin of error to two decimal places.

$n = 1400$

$$\hat{p} = \frac{600}{1400}$$
$$= 0 \cdot 428$$
$$\simeq 0 \cdot 43$$

$$\frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{1400}}$$
$$= 0 \cdot 026$$
$$\simeq 0 \cdot 03$$

$$95\% \text{ confidence interval}: \hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$= 0·43 - 0·03 < p < 0·43 + 0·03$$

$$= 0·4 < p < 0·46$$

(ii) If the sample was increased to 2,000, but none of the additional people had a TV, find the new confidence interval.

Sample increased to 2000

$$\hat{p} = \frac{600}{2000} = 0 \cdot 3$$

$$\text{Margin of error} = \frac{1}{\sqrt{2000}} = 0 \cdot 02$$

Confidence interval:

$$0·3 - 0·02 < p < 0·3 + 0·02$$

$$0·28 < p < 0·32$$

(iii) What happens to the confidence interval as the sample size increases?

As the sample size increases, the margin of error decreases and so the confidence interval shrinks.

6. What sample size, to the nearest whole number, is needed to have the following margins of error:

(i) 1%

$$\frac{1}{\sqrt{n}} = 1\%$$

$$\frac{1}{\sqrt{n}} = 0 \cdot 01$$

$$\left(\frac{1}{\sqrt{n}}\right)^2 = (0 \cdot 01)^2 \qquad \text{(Square both sides)}$$

$$\frac{1}{n} = 0 \cdot 0001$$

$1 = 0{\cdot}0001n$ (Multiply both sides by n)

$n = 10{,}000$ (Divide both sides by 0·0001)

(ii) 6%

$$\frac{1}{\sqrt{n}} = 6\%$$

$$\frac{1}{\sqrt{n}} = 0 \cdot 06 \qquad \text{(square both sides)}$$

$$\frac{1}{n} = 0 \cdot 0036 \qquad \text{(multiply both sides by } n\text{)}$$

$1 = 0{\cdot}0036n$ (divide both sides by 0·0036)

$n = 277{\cdot}\dot{7}$

$n = 278$

(iii) 10%

$$\frac{1}{\sqrt{n}} = 10\%$$

$$\frac{1}{\sqrt{n}} = 0 \cdot 1 \qquad \text{(square both sides)}$$

$$\frac{1}{n} = 0 \cdot 01 \qquad \text{(multiply both sides by } n\text{)}$$

$$1 = 0 \cdot 01n \qquad \text{(divide both sides by } 0 \cdot 01\text{)}$$

$$n = 100$$

(iv) 25%

$$\frac{1}{\sqrt{n}} = 25\%$$

$$\frac{1}{\sqrt{n}} = 0 \cdot 25$$

$$\frac{1}{n} = 0 \cdot 0625$$

$$1 = 0 \cdot 0625n$$

$$n = 16$$

(v) 3·5%

$$\frac{1}{\sqrt{n}} = 3 \cdot 5\%$$

$$\frac{1}{\sqrt{n}} = 0 \cdot 035$$

$$\frac{1}{n} = 0 \cdot 001225$$

$$1 = 0 \cdot 001225n$$

$$n = 816 \cdot 32653$$

$$n = 816$$

7. Several factors are involved in the creation of a confidence interval. Among them are the sample size and the margin of error. Which statements are correct?

(i) Larger samples provide smaller margins of error.

Yes. This statement is correct.

Example $\frac{1}{\sqrt{100}} = 0 \cdot 1 \qquad \frac{1}{\sqrt{10,000}} = 0 \cdot 01$

(ii) Halving the margin of error requires a sample twice as large.

No. This statement is incorrect.

Example $\frac{1}{\sqrt{100}} = 0 \cdot 1 \qquad \frac{1}{\sqrt{50}} = 0 \cdot 14$

$2(0 \cdot 1) \neq 0 \cdot 14$

(iii) You can get a smaller margin of error by selecting a bigger sample.

Yes. This statement is correct.

Example see (i)

(iv) A sample nine times as large will make a margin of error one third as big.

Yes. This statement is correct.

Example $\frac{1}{\sqrt{100}} = 0 \cdot 1 \qquad \frac{1}{\sqrt{900}} = 0 \cdot 0\dot{3}$

$0 \cdot 1 = 3(0 \cdot 0\dot{3})$

8. A random sample of 100 cinemas showed that the mean price of a student ticket was €7·00, with a standard deviation of €0·80. Assuming the data is normally distributed construct a 95% confidence interval for this data.

$n = 100$

$\sigma = €0{\cdot}80$

$\mu = €7$

Assuming this data is normally distributed, we can use the empirical rule to find the range of prices within 95% of the tickets are sold for.

95% confidence interval = $\mu - 2\sigma \rightarrow \mu + 2\sigma$

$7 - 2(0{\cdot}80) \rightarrow 7 + 2(0{\cdot}80)$

$€5{\cdot}40 \rightarrow €8{\cdot}60$

9. In an RTÉ news telephone poll of 1,012 adults, 11% of the respondents said that they were happy with the work of the current president. Construct a 95% confidence interval for the study.

Is this a good reflection of the true population in your opinion? Justify your answer.

Sample: 1012

11% happy $\Rightarrow$ 11% of 1012 = 111·32

Sample proportion $(\hat{p})$: $\dfrac{111\cdot 32}{1012} = 0\cdot 11$

Margin of error :

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

$$= \frac{1}{\sqrt{1012}}$$

$$= 0\cdot 0314$$

$$\simeq 0\cdot 03$$

95% confidence interval:

$$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0\cdot 11 - 0\cdot 03 < p < 0\cdot 11 + 0\cdot 03$$

$$0\cdot 08 < p < 0\cdot 14$$

Yes, this is a good reflection of the true population because the sample size is quite large.

10. A random sample of 100 students is selected from a certain school. They are given an IQ test which has a known standard deviation of 11. The sample mean is found to be 112. Assuming the data is normally distributed, determine a 95% confidence interval for estimating the mean school intelligence.

Sample: 100

Standard deviation $\sigma = 11$

Mean $\mu = 112$

Assuming this data is normally distributed, we can use the empirical rule to find the IQ scores within 95% of the population fall.

95% confidence interval: $\mu - 2\sigma < p < \mu + 2\sigma$

$$112 - 2(11) < p < 112 + 2(11)$$

$$112 - 22 < p < 112 + 22$$

$$90 < p < 134$$

Practice Questions 25.3

1. State the null and alternative hypotheses used to test the following claims:

(i) A school publicises that the percentage of its students who are involved in at least one extracurricular activity is 61%.

H_0: The percentage of students involved in at least one extracurricular activity is 61%.

H_1: The percentage of students involved in at least one extracurricular activity is not 61%.

(ii) A car garage announces that the mean time for a wheel change is less than 15 minutes.

H_0: The mean time for a wheel change is less than 15 minutes.

H_1: The mean time for a wheel change is not less than 15 minutes.

(iii) A company advertises that the mean life of its products is more than eight years.

H_0: The mean life of the products is more than 8 years.

H_1: The mean life of the products are not more than 8 years.

(iv) A drug company announces that a new vaccine is 10% more effective than its predecessor.

H_0: The new product is 10% more effective than the previous.

H_1: The new product is not 10% more effective than the previous.

(v) A skincare company advertise that 35% of people surveyed thought that their product reduced skin oiliness.

H_0: The percentage of people who thought the product reduced skin oiliness is 35%.

H_1: The percentage of people who thought the product reduced skin oiliness is not 35%.

2. An advertisement states that 8 out of 10 cats prefer Catty Cat cat food.
To test this claim, a researcher carried out an experiment using 150 cats and found that 75 cats preferred Catty Cat.

(i) State the null and alternative hypotheses.

H_0: The proportion of cats who prefer Catty Cat is 0·8

H_1: The proportion of cats who prefer Catty Cat is not 0·8.

(ii) Calculate $\hat{p}$.

$$\hat{p} : \frac{75}{150} = 0 \cdot 5$$

(iii) Calculate the margin of error to two decimal places.

Margin of error:

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

$$= \frac{1}{\sqrt{150}}$$

$$= 0 \cdot 081$$

$$\simeq 0 \cdot 08$$

(iv) Set up the 95% confidence interval.

95% confidence interval:

$$\hat{p}-\frac{1}{\sqrt{n}}<p<\hat{p}+\frac{1}{\sqrt{n}}$$

$= 0\cdot5 - 0\cdot08 < p < 0\cdot5 + 0\cdot08$

$= \quad 0\cdot42 < p < 0\cdot58$

(v) Decide whether to reject or fail to reject the null hypothesis.

H_0 states the proportion is 0·8 which does not fall within the confidence interval so we reject H_0.

(vi) What does this mean in relation to the claim?

It appears the claim that 8 out of 10 cats prefer Catty Cat food is untrue.

3. A hotel chain claims that 35% of its online reviews are positive. A sample of 200 reviews shows that 120 are positive.

(i) State the null and alternative hypotheses.

H_0: The percentage of positive reviews is 35%.

H_1: The percentage of positive reviews is not 35%.

(ii) Calculate $\hat{p}$.

$$\hat{p}:\frac{120}{200}=0\cdot6$$

(iii) Calculate the margin of error to two decimal places.

Margin of error:

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

$$=\frac{1}{\sqrt{200}}$$

$$=0\cdot07$$

(iv) Set up the 95% confidence interval.

Confidence interval: $\hat{p}-\frac{1}{\sqrt{n}}<p<\hat{p}+\frac{1}{\sqrt{n}}$

$$0{\cdot}6 - 0{\cdot}07 < p < 0{\cdot}6 + 0{\cdot}07$$

$$0{\cdot}53 < p < 0{\cdot}67$$

(v) Decide whether to reject or fail to reject the null hypothesis.

We reject the null hypothesis because 35% = 0·35 which does not fall within the confidence interval.

(vi) What does this mean in relation to the claim?

This means the claim that 35% of online reviews are positive is untrue.

In fact it appears that the number of positive reviews are actually higher than the company claims.

4. An opinion poll of 1,000 voters is carried out prior to an election. 25% of those polled said they would vote for the Left Party. Left Party management believe they have 35% support.

(i) Set up the 95% confidence interval.

H_0: Left party receive 35% of the votes in the election.

H_1: Left party do not receive 35% of the votes in the election.

$\hat{p}$ = 0·25 (given in question)

Margin of error :

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

$$=\frac{1}{\sqrt{1000}}$$

$$=0\cdot0316$$

$$\simeq 0\cdot03$$

Confidence interval: $\hat{p}-\frac{1}{\sqrt{n}}<p<\hat{p}+\frac{1}{\sqrt{n}}$

$$0{\cdot}25 - 0{\cdot}03 < p < 0{\cdot}25 + 0{\cdot}03$$

$$0{\cdot}22 < p < 0{\cdot}28$$

(ii) Decide whether to reject or fail to reject the null hypothesis, using a 5% level of significance.

We reject the null hypothesis as $0{\cdot}35$ does not fall within the confidence interval.

(iii) What does this mean in relation to the claim?

This means that the claim is incorrect.

5. Evie rolled a die 360 times and got 72 fives. She suspects that the die is biased and carries out a hypothesis test to check.

(i) What proportion of fives would Evie expect if the die was unbiased?

$\frac{1}{6}$ should be 5's $= 0{\cdot}1\dot{6}$ (Probability of rolling a 5)

(ii) What is the null hypothesis?

H_0: The die is unbiased [Die is always assumed to be fair]

(iii) What is the alternative hypothesis?

H_1: The die is biased

(iv) Use a hypothesis test at the 95% confidence level to determine if the die is biased.

$$\hat{p} = \frac{72}{360} = 0 \cdot 2$$

Margin of error:

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

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

$$= 0 \cdot 0527$$

$$\simeq 0 \cdot 05$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0{\cdot}2 - 0{\cdot}05 < p < 0{\cdot}2 + 0{\cdot}05$$

$$0{\cdot}15 < p < 0{\cdot}25$$

The null hypothesis is not rejected as $0{\cdot}1\dot{6}$ falls within the confidence interval. Therefore, the die is unbiased (fair).

6. A community sports organisation claims that it has equal numbers of males and females. A local councillor wanted to investigate this claim. He took a sample of 80 members and found that 65 of them were male. Determine whether the sports organisation's claim is true, using 5% level of significance.

H_0: The percentage of members that are male is 50% (0·5).

H_1: The percentage of members that are male is not 50%.

$\hat{p}: \ \dfrac{65}{80} = 0\cdot 8125 \simeq 0\cdot 81$

Margin of error: $\dfrac{1}{\sqrt{n}} = \dfrac{1}{\sqrt{80}} = 0\cdot 1118\ldots \ \simeq 0\cdot 11$

Confidence interval: $\hat{p} - \dfrac{1}{\sqrt{n}} < p < \hat{p} + \dfrac{1}{\sqrt{n}}$

$$0\cdot 81 - 0\cdot 11 < p < 0\cdot 81 + 0\cdot 11$$

$$0.7 < p < 0.92$$

We reject the null hypothesis as 0·5 does not fall within the confidence interval. Therefore, the claim that the organisation has equal numbers of males and females is untrue.

7. The Irish Tourist Board claimed that 65% of the tourists who visited Ireland last year were return visitors. The Department of Transport, Tourism and Sport wanted to test this claim. They asked a random sample of 800 tourists if they had been to Ireland before. 300 responded that they had.

(i) Using a 95% confidence interval, decide whether to reject or fail to reject the null hypothesis.

H_0: The percentage of first time visitors to Ireland last year is 65% (0·65)

H_1: The percentage of first time visitors to Ireland last year is not 65%.

$$\begin{aligned}\hat{p} &= \frac{300}{800}\\ &= 0\cdot 375\\ &\simeq 0\cdot 38\end{aligned}$$

$$\begin{aligned}\text{Margin of error} &= \frac{1}{\sqrt{n}}\\ &= \frac{1}{\sqrt{800}}\\ &= 0\cdot 035\ldots\\ &\simeq 0\cdot 04\end{aligned}$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0\cdot 38 - 0\cdot 04 < p < 0\cdot 38 + 0\cdot 04$$

$$0\cdot 34 < p < 0\cdot 42$$

We reject the null hypothesis as 0·65 does not fall in the confidence interval.

(ii) What does this mean in relation to the Irish Tourist Board's claim?

This means the claim is untrue. The percentage of first time visitors is lower than 65%.

8. A popular mobile phone provider claims that 70% of their customers are satisfied with the service provided by them. A consumer agency decides to investigate the claim. They survey 900 of the mobile company's customers and find that 650 people said they were satisfied.

Test the company's claim using a 5% level of significance. As a result of your findings, what advice would you give to the mobile company?

H_0: The proportion of satisfied customers is 70% (0·7)

H_1: The proportion of satisfied customers is not 70%.

$$\begin{aligned}\hat{p} &= \frac{650}{900}\\ &= 0\cdot722\\ &\simeq 0\cdot72\end{aligned}$$

$$\begin{aligned}\text{Margin of error} &= \frac{1}{\sqrt{n}}\\ &= \frac{1}{\sqrt{900}}\\ &= 0\cdot033\\ &\simeq 0\cdot03\end{aligned}$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0\cdot72 - 0\cdot03 < p < 0\cdot72 + 0\cdot03$$

$$0\cdot69 < p < 0\cdot75$$

We fail to reject the null hypothesis as 0·7 falls within the confidence interval.

It appears the company's satisfaction rate is higher than claimed.

Advice to the company: They should continue to provide the services they do.

9. A company manufacturing a new type of chocolate bar claims that 85% of the people who tried their product would recommend it to a friend. An independent study of 625 people who tried the bar, found that only 125 would recommend it. Use a 95% confidence interval to investigate whether the company's claim is true.

H_0: The percentage of people that would recommend the new chocolate bar to a friend is 85%.

H_1: The percentage of people that would recommend the new chocolate bar to a friend is not 85%.

$$\hat{p} = \frac{125}{625}$$
$$= 0 \cdot 2$$

$$\text{Margin of error } = \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{625}}$$
$$= 0 \cdot 04$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0{\cdot}2 - 0{\cdot}04 < p < 0{\cdot}2 + 0{\cdot}04$$

$$0{\cdot}16 < p < 0{\cdot}24$$

We reject the null hypothesis as 0·85 falls outside the confidence interval. Therefore, the claim is untrue.

10. The Union of Secondary Students of Ireland claims that 65% of Leaving Cert students spend more than three hours a night on study or homework. The principal of a school wants to investigate this claim.

She surveys the 200 Leaving Cert students in her school and finds that 85 of them report spending more than three hours on study or homework. Using a 95% confidence interval, determine whether that principal is likely to accept the claim of the Students' Union.

H_0: The percentage of Leaving Certificate students studying more than 3 hours a night is 65% (0·65).

H_1: The percentage of Leaving Certificate students studying more than 3 hours a night is not 65% (0·65).

$$\hat{p} = \frac{85}{200} = 0\cdot 425 \simeq 0\cdot 43$$

$$\text{Margin of error} = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{200}} = 0\cdot 0707.. \simeq 0\cdot 07$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0\cdot 43 - 0\cdot 07 < p < 0\cdot 43 + 0\cdot 07$$

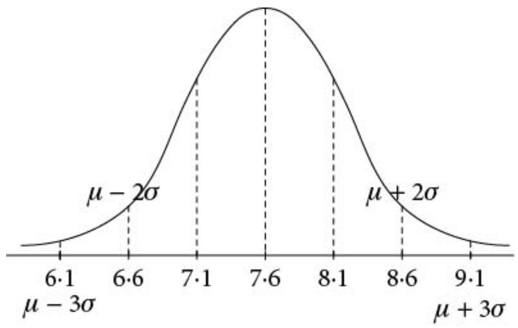
$$0\cdot 36 < p < 0\cdot 5$$

The null hypothesis is rejected as 0·65 does not fall within the confidence interval. Therefore, the principal is unlikely to accept the claim.

Revision and Exam Style Questions − Section A

1. The time, in hours, that each student spent sleeping on a school night was recorded for 550 secondary school students. The distribution of these times was found to be approximately normal with a mean of 7·6 hours and a standard deviation of 0·5 hours.

(i) Draw the normal distribution curve to show this data.



(ii) What amount of time did 95% of these students spend sleeping?

95% of the students are between $\mu - 2\sigma \rightarrow \mu + 2\sigma$

$= 6{\cdot}6 \rightarrow 8{\cdot}6$ hrs

(iii) What percentage of students spent between 7·1 and 8·1 hours sleeping?

7·1 − 8·1 hrs

$= \mu - \sigma \rightarrow \mu + \sigma$

= 68%

2. A large group of students have a mean weight of 68 kg, with a standard deviation of 3 kg. Use the empirical rule to find the weight interval that contains 95% of the students.

$\mu = 68$ kg $\sigma = 3$ kg

95% are within: $\mu - 2\sigma \rightarrow \mu + 2\sigma$

$68 - 2(3) \rightarrow 68 + 2(3)$

$68 - 6 \rightarrow 68 + 6$

$62\text{kg} \rightarrow 74\text{kg}$

3. Deirdre claims that the weather forecasts on the local radio are no better than rolling a fair die. She predicts rain if the result is odd and no rain if the result is even. She records the weather for 30 days and finds that the forecast is correct on 20 of the days.

$$\left.\begin{aligned}\text{Deirdre predicts rain} = \text{odd} \quad & \frac{3}{6} = 50\% \\ \text{no rain} = \text{even} \quad & \frac{3}{6} = 50\%\end{aligned}\right\}\text{Probability}$$

(i) State the null and alternative hypotheses.

H_0: The percentage of times her forecast is correct is 50%

H_1: The percentage of times her forecast is correct is not 50%

(ii) Construct the confidence interval.

$$\begin{aligned}\hat{p} &= \frac{20}{30} \\ &= 0\cdot\dot{6} \\ &= 0\cdot 67\end{aligned}$$

$$\begin{aligned}\text{Margin of error} &= \frac{1}{\sqrt{n}} \\ &= \frac{1}{\sqrt{30}} \\ &= 0\cdot 18257 \\ &\simeq 0\cdot 18\end{aligned}$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0\cdot 67 - 0\cdot 18 < p < 0\cdot 67 + 0\cdot 18$$

$$0\cdot 49 < p < 0\cdot 85$$

(iii) Is Deirdre’s claim correct?

We fail to reject the null hypothesis as $0\cdot 5$ falls within the interval.

Therefore Deirdre’s claim is true.

4. A candidate for election claims that 45% of voters will vote for him. His opponent wants to test this claim. She takes a random sample of 120 voters and finds that 48 people said they would vote for him.

(i) State the null and alternative hypotheses.

H_0: The percentage of voters that will vote for the candidate is 45%.

H_1: The percentage of voters that will vote for the candidate is not 45%.

(ii) Calculate the confidence interval.

$$\hat{p} = \frac{48}{120}$$
$$= 0\cdot 4$$

$$\text{Margin of error} = \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{120}}$$
$$= 0\cdot 091287$$
$$\simeq 0\cdot 09$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0\cdot 4 - 0\cdot 09 < p < 0\cdot 4 + 0\cdot 09$$

$$0\cdot 31 < p < 0\cdot 49$$

(iii) Is the candidate's claim upheld?

We fail to reject the null hypothesis as 0·45 falls within the confidence interval.

Therefore, the candidates claim is upheld.

5. An estate agent claims to have an 80% success rate for selling a house within two months. A client wants to test this hypothesis. He looks at the last 60 houses they have sold and found that 50 of them had been sold within two months.

(i) State the null and alternative hypotheses for this test.

H_0: The percentage of houses sold within 2 months is 80%.

H_1: The percentage of houses sold within 2 months is not 80%.

(ii) Set up the confidence interval for this test.

$$\hat{p} = \frac{50}{60} = 0 \cdot 83333 \simeq 0 \cdot 83$$

$$\text{Margin of error } = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{60}} = 0 \cdot 129 \simeq 0 \cdot 13$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0 \cdot 83 - 0 \cdot 13 < p < 0 \cdot 83 + 0 \cdot 13$$

$$0 \cdot 7 < p < 0 \cdot 96$$

(iii) Is the estate agent's claim correct?

We fail to reject the null hypothesis as 0·8 falls within the confidence interval.

Therefore the estate agents claim is correct.

6. Adult IQ scores have a normal distribution with mean of 100 and a standard deviation of 15.

(i) Use the empirical rule to find the percentage of adults with scores between 70 and 130.

$\mu = 100 \qquad \sigma = 15$

70: $100 - 70 = 30$

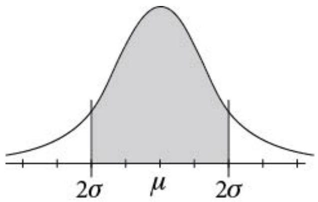
$= 2(15)$

$70 = \mu - 2\sigma$

130 : $130 - 100 = 30$

$= 2(15)$

$130 = \mu + 2\sigma$



$70 - 130$

$= \mu - 2\sigma \rightarrow \mu + 2\sigma$

$= 95\%$

(ii) If 250 adults are randomly selected, about how many of them have an IQ between 85 and 130?

85: $100 - 85 = 15$

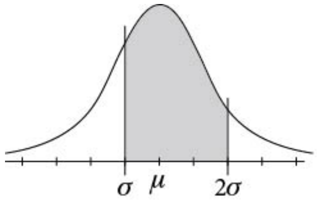
$85 = \mu - \sigma$

$130 = \mu + 2\sigma$

$85 – 130 \qquad = \mu - \sigma \rightarrow \mu + 2\sigma$

$= 68\% + 13{\cdot}5\%$

$= 81{\cdot}5\%$



$81{\cdot}5\%$ of $250 = 203{\cdot}75$

Since it is not possible to have $0{\cdot}75$ of a person, the number of adults that have an IQ between 85 and 130 is 203.

7. A machine fills packets with roast peanuts.

The line manager takes a sample of 10 packets every hour.

The mean weights of the samples are normally distributed with a mean of 106 g and a standard deviation of 2 g.

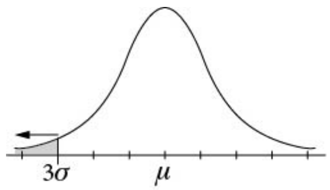
(i) Write down the percentage of the samples that are likely to have a sample mean of less than 100 g.

$\mu = 106$ g $\quad\quad \sigma = 2$ g

100: $106 - 100 = 6$

$= 3(2)$

$100 = \mu - 3\sigma$



$0{\cdot}15\% < \mu - 3\sigma$

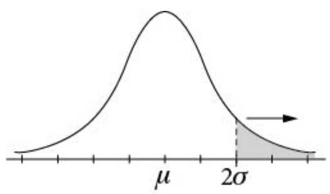
0·15% of the samples have a weight < 100 g

(ii) Write down the percentage of the samples that are likely to have a sample mean of more than 110 g.

$110\ g = 110 - 106 = 4$

$= 106 + 2(2)$

$110 = \mu + 2\sigma$



$2{\cdot}35\% + 0{\cdot}15\% > \mu + 2\sigma$

2·5% of the samples have a weight > 110 g

Section B − More challenging problems

1. The time spent by shoppers at a furniture shop on a Saturday is approximately normally distributed with a mean of 30 minutes and a standard deviation of 5 minutes. If 2,850 shoppers are expected to visit the shop on a Saturday, how many shoppers are expected to spend:

(i) more than 40 minutes in the shop?

$\mu = 30$ min $\quad \sigma = 5$ min

40: $40 - 30$

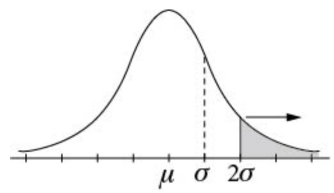
$= 10$

$= 2(5)$

40: $\mu + 2\sigma$

$2{\cdot}35\% + 0{\cdot}15\% > 40$ mins

$2{\cdot}5\% > 40$ mins



$2{\cdot}5\%$ of $2850 = 71{\cdot}25$

Since it is not possible to have $0 \cdot 25$ of a person, the number of shoppers expected to spend more than 40 minutes in the shop is 71 shoppers.

(ii) between 20 and 35 minutes in the shop?

20: $30 - 20$

$= 10$

$= 2(5)$

20: $\mu - 2\sigma$

35: $35 - 30$

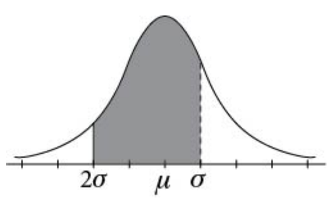
$= 5$

$= \sigma$

35: $\mu + \sigma$

Between 20 and 35 minutes is:

$\mu - 2\sigma \rightarrow \mu + \sigma$



$68\% + 13{\cdot}5\% = 81{\cdot}5\%$

20 min $< 81{\cdot}5\% <$ 35 min

$81 \cdot 5\%$ of $2850 = 2322 \cdot 75$

Since it is not possible to have $0 \cdot 75$ of a person, the number of shoppers are expected to spend between 20 and 35 minutes in the shop is 2322 shoppers.

2. A machine fills packets with corn flakes. The label on each packet says the contents weigh 1·5 kg.

At regular intervals a sample of 10 packets is taken from the machine.

The sample mean is calculated. The sample means are normally distributed with a mean of 1,520 g and a standard deviation of 4 g.

(i) Give one reason why it may not be practical to check the weight of each packet.

It would be time-consuming and very expensive.

(ii) Between what limits would you expect 95% of the sample means to lie?

95% : $\mu - 2\sigma \rightarrow \mu + 2\sigma$ (normally distributed)

$\mu = 1{,}520$ g $\sigma = 4$ g

$1520 - 2(4) \rightarrow 1520 + 2(4)$

$1520 - 8 \rightarrow 1520 + 8$

1512 g $\rightarrow 1528$ g

(iii) Martha takes a sample of 10 packets. She finds the weight, in grams, of each packet.

Here are her results:

1,517	1,525	1,534	1,527	1,531
1,521	1,532	1,535	1,526	1,532

The allowable limits for the weights are 1,520 g ± 3 standard deviations. How many of the sample would be rejected on these grounds?

$1520 - 3\sigma \rightarrow 1520 + 3\sigma$

$1520 - 3(4) \rightarrow 1520 + 3(4)$

$1520 - 12 \rightarrow 1520 + 12$

1508 g $\rightarrow 1532$ g

Two of the samples (1534 g, 1535 g) are outside the limits set and would be excluded.

(iv) What could be done to ensure weights are within the acceptable range?

The machine could be set to a lower weight.

3. A drug company claims that the success rate of a treatment for a disease is 47%. Dr. Green doubts this claim and wants to test it. He uses the treatment on 50 patients with the disease and cures 24 of them. Is the company's claim correct? Justify your answer.

H_0: The success rate for the treatment is 47%.

H_1: The success rate for the treatment is not 47%.

$$\hat{p} = \frac{24}{50}$$
$$= 0\cdot 48$$

$$\text{Margin of error} = \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{50}}$$
$$= 0\cdot 14$$

95% confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0\cdot 48 - 0\cdot 14 < p < 0\cdot 48 + 0\cdot 14$$

$$0\cdot 34 < p < 0\cdot 62$$

We fail to reject H_0 as $0\cdot 47$ falls within the confidence interval. The company's claim appears to be true.

4. A survey is being conducted of voters' opinions on several different issues.

(i) What is the overall margin of error of the survey, at 95% confidence, if it is based on a simple random sample of 1,000 voters?

Margin of error:

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

$$= \frac{1}{\sqrt{1000}}$$

$$= 0 \cdot 0316$$

$$\simeq 0 \cdot 03$$

A political party had claimed that it has the support of 26% of the electorate. Of the voters in the sample above, 238 stated that they support the party.

(ii) Is there sufficient evidence to reject the party's claim? Justify your answer.

H_0: Support for the party is 26%.

H_1: Support for the party is not 26%.

$$\hat{p} = \frac{238}{1000}$$

$$= 0 \cdot 238$$

$$\simeq 0 \cdot 24$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0{\cdot}24 - 0{\cdot}03 < p < 0{\cdot}24 + 0{\cdot}03$$

$$0{\cdot}21 < p < 0{\cdot}27$$

Therefore, we fail to reject the H_0 as 26% falls within the confidence interval.

So, the party's claim appears to be true.

5. A manufacturing company claims that 75% of suppliers recommend their brand of product. In a survey of 42 traders, 28 said that they would recommend the company's product. Use a hypothesis test at the 95% level of confidence to decide whether there is sufficient evidence to reject the company's claim. State clearly the null hypothesis and your conclusion.

H_0: The percentage of suppliers that would recommend the brand is 75%.

H_1: The percentage of suppliers that would recommend the brand is not 75%.

$$\hat{p} = \frac{28}{42}$$
$$= 0\cdot\dot{6}$$
$$= 0\cdot 67$$

$$\text{Margin of error} = \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{42}}$$
$$= 0\cdot 1543$$
$$\simeq 0\cdot 15$$

Confidence interval: $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$0\cdot 67 - 0\cdot 15 < p < 0\cdot 67 + 0\cdot 15$$

$$0\cdot 52 < p < 0\cdot 82$$

We fail to reject the null hypothesis as 0·75 falls within the confidence interval. Therefore, we can see that the company's claim appears to be true.

6. A boy decides to test whether a coin is fair by tossing it 20 times and recording the number of heads. He suspects that it is more likely to show a head than a tail and when he tosses it he gets 11 heads.

(i) State suitable hypotheses to use for this test.

H_0: The coin is fair and the chance of a tail is 0·5.

H_1: The coin is unfair and the chance of a head is not 0·5.

(ii) Use a 5% level of significance to test whether the coin is fair.

$$\hat{p} = \frac{11}{20}$$
$$= 0 \cdot 55$$

Margin of error:

$$= \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{20}}$$
$$= 0 \cdot 2236$$
$$\simeq 0 \cdot 22$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0{\cdot}55 - 0{\cdot}22 < p < 0{\cdot}55 + 0{\cdot}22$$

$$0{\cdot}33 < p < 0{\cdot}77$$

We fail to reject the null hypothesis as 0·5 falls within the confidence interval.

Therefore, the coin is fair.

(iii) Do you think his results are reliable? Give a reason for the answer.

No, the results are not reliable. The number of tosses is very small. 20 is too small a sample size.

7. In general, 55% of driving tests result in a pass. It is suspected that a particular examiner is tougher than usual. To investigate this, records are kept on the next 10 people tested by this examiner. Four of them pass.

(i) Stating your hypotheses carefully and showing your working in full, test at the 5% level whether or not the examiner is tougher than normal.

H_0: The examiners pass rate is 55%.

H_0: The examiners pass rate is not 55%.

$$\hat{p} = \frac{4}{10}$$
$$= 0 \cdot 4$$

$$\text{Margin of error} = \frac{1}{\sqrt{n}}$$
$$= \frac{1}{\sqrt{10}}$$
$$= 0 \cdot 3162$$
$$\simeq 0 \cdot 32$$

Confidence interval: $$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

$$0{\cdot}4 - 0{\cdot}32 < p < 0{\cdot}4 + 0{\cdot}32$$

$$0{\cdot}08 < p < 0{\cdot}72$$

We fail to reject the null hypothesis as 0·55 falls within the confidence interval.

Therefore, the examiner is not tougher than normal.

(ii) Comment on the validity of this method of assessing the examiner.

This is a very small sample to assess the examiner by and so the results presented are not necessarily valid.

The small sample size leads to a high margin of error.

8. On a production line in a factory, pineapple rings in syrup are put in tins. The label on each tin says that the contents weigh 415 g.

(i) Give two reasons why it is not practical to check the weight of the contents of each tin.

It would be very time consuming and very expensive.

Samples of tins are taken at intervals and the weights of the contents are found. It has been found that the mean weight of the samples is 417 g and the standard deviation is 0·6 g. The mean weights of the samples are normally distributed.

(ii) Between what limits would you expect 99·7% of the sample means to lie?

$\mu = 417$ g $\quad \sigma = 0{\cdot}6$ g

$99{\cdot}7\% = \mu - 3\sigma \rightarrow \mu + 3\sigma$

$$\mu - 3\sigma \rightarrow \mu + 3\sigma$$

$$417 - 3(0{\cdot}6) \rightarrow 417 + 3(0{\cdot}6)$$

$$417 - 1{\cdot}8 \rightarrow 417 + 1{\cdot}8$$

$$415{\cdot}2 \text{ g} \rightarrow 418{\cdot}8 \text{ g}$$

(iii) A consumer authority inspects the factory and decides that the labels on the tins should be altered from 415 g. Use your answer from **(ii)** to explain why this is the case.

Only a very small number will fall below 415·2 (0·15%) and therefore, the majority of tins weigh > 415 g.