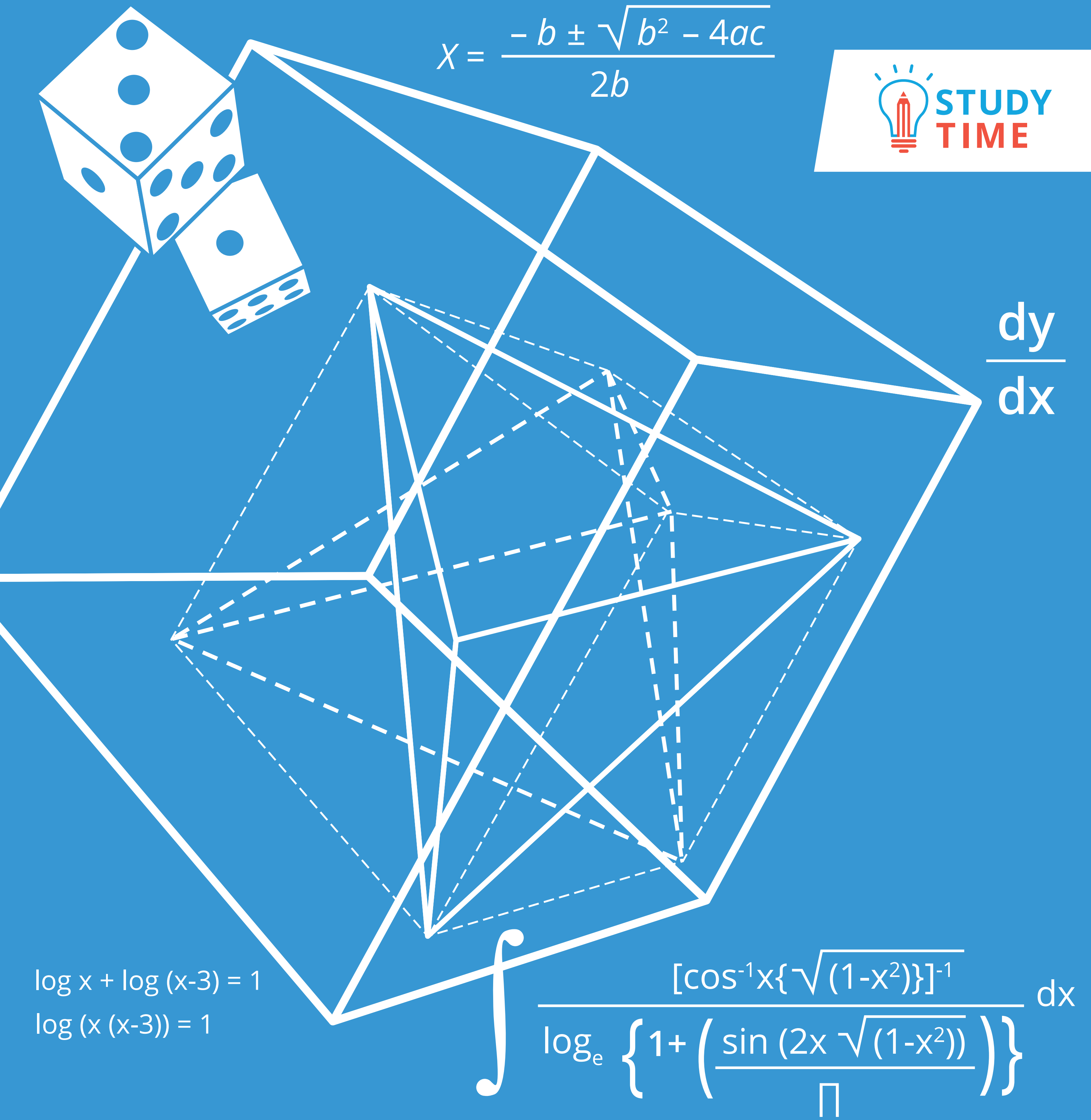
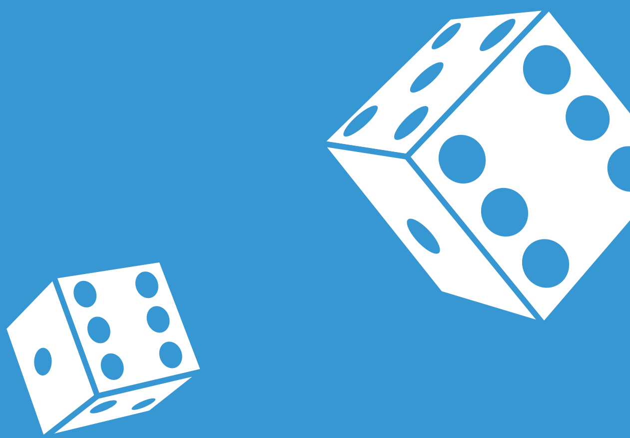
STUDY TIME



LEVEL 2 MATHS

PROBABILITY

NCEA Workbook Answers



Section One
The Foundations

1. Key Terms

a. Claim: a statement based on probability.

b. Population: all of the members of a defined group.

c. Sample: a smaller section of an overall population.

d. Favourable outcome: the outcome we are investigating (not necessarily a 'good outcome).

e. Probability: the chance of something happening (measured between 0 and 1).

f. Proportion: how much of a population had a specific outcome, measured as a decimal.

g. Conditional probability: probability based on two independent situations occurring (It always has the words 'if' or 'given that' in the question).

h. Risk: a probability of an event occurring.

i. Relative risk: a risk calculated by comparing two or more absolute risks.

j. Mean (μ): the average value of a measurement.

k. Standard Deviation (σ): a measure of how far the values of a measurement are from the mean.

l. Validity: a measure of how reliable a statistic is.

2. Key Equations to Remember

a. $\frac{\textit{number of favourable outcomes}}{\textit{total number of outcomes}}$

b. $\frac{\textit{number of favourable outcomes}}{\textit{total number of outcomes}} \times 100$

c. $\frac{\textit{number of favourable outcomes}}{\textit{total number of outcomes}}$, then convert to a decimal

d. $\frac{\textit{number of favourable outcomes}}{\textit{total number of trials}}$

e. theoretical probability × number of Times you attempt to get the favourable outcome

f. $\frac{\textit{Probability of A and B occurring}}{\textit{Probability of B}} = P(A|B) = \frac{P(A \cap B)}{P(B)}$

g. $\frac{\textit{Risk of A}}{\textit{Risk of B}}$

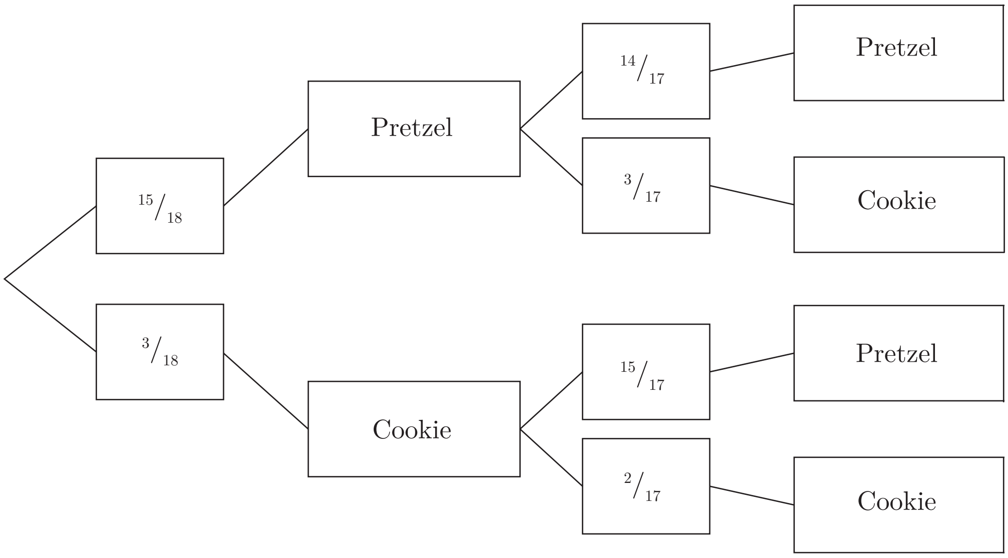
Answers

3. Probability Trees

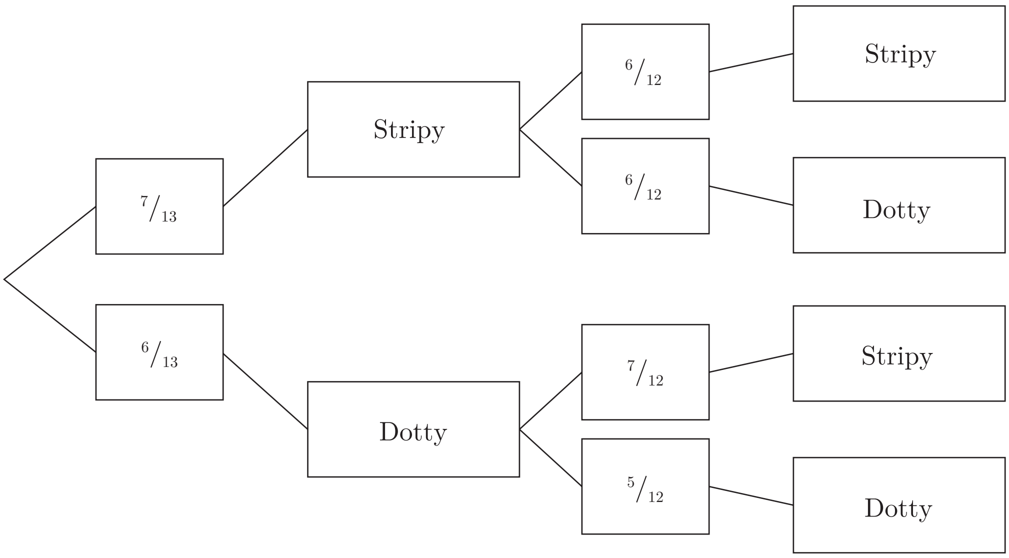
a.

Fraction	Percentage (1dp)	Proportion
$^{8}/_{31}$	25.8%	0.258
$^{24}/_{100} = ^{6}/_{25}$	24.0%	0.240
$^{75}/_{100} = ^{3}/_{4}$	75.0%	0.750

b.



c. After doing the laundry, Jess realised that she had lost one of her socks, leaving her with 7 stripy socks and 6 dotty socks. On Saturday she picks out two socks at random to wear. Using the information provided and the probability tree answer the following questions.



i. $^{7}/_{13} \times ^{6}/_{12} = ^{42}/_{156} = ^{7}/_{26} = 0.2692$ (4 d.p.)

ii. ${}^{6}/_{13} \times {}^{5}/_{12} = {}^{30}/_{156} = {}^{5}/_{26} = 0.1923\ (4\text{ d.p.})$

iii. She could pick out two stripy socks OR two dotty socks. This or means that we add together the two probabilities. ${}^{7}/_{26} + {}^{5}/_{26} = {}^{12}/_{26} = {}^{6}/_{13} = 0.4615\ (4\text{ d.p.})$

iv. Since she either has matching socks or odd socks, the easiest way to calculate this is with:
1 - P(matching socks)

As the probabilities of all outcomes add to 1. So, our answer is:
1 - ${}^{12}/_{26} = {}^{14}/_{26} = {}^{7}/_{13} = 0.5385\ (4\text{ d.p.})$

v. ${}^{6}/_{13} \times {}^{7}/_{12} = {}^{42}/_{156} = {}^{7}/_{26} = 0.2692\ (4\text{ d.p.})$

4. Two-way Tables

a. *i.* ${}^{7}/_{32} = 0.2186\ (4\text{ d.p.})$

ii. ${}^{6}/_{18} = 0.3333\ (4\text{ d.p.})$

iii. ${}^{8}/_{15} = 0.5333\ (4\text{ d.p.})$

i. P(female AND licence AND accident)
= P(female) × P(female with licence) × P(female with accident)
= ${}^{55}/_{100} \times {}^{35}/_{55} \times {}^{6}/_{55}$
= $0.0382\ (4\text{ d.p.})$

ii. P(male AND no licence AND accident)
= P(male) × P(male with no licence) × P(male with accident)
= ${}^{45}/_{100} \times {}^{15}/_{45} \times {}^{9}/_{45}$
= 0.03

iii. P(female AND no licence AND no accident)
= P(female) × P(female with no licence) × P(female with no accident)
= ${}^{55}/_{100} \times {}^{20}/_{55} \times {}^{49}/_{55}$
= $0.1782\ (4\text{ d.p.})$

iv. P(male has licence AND no accident)
= P(male with licence) × P(male with no accident)
= ${}^{30}/_{45} \times {}^{36}/_{45}$
= $0.5333\ (4\text{ d.p.})$

v. P(student who has accident is female AND has licence)
= P(student who has accident is female) x P(female with licence)
= ${}^{6}/_{15} \times {}^{35}/_{55}$
= 0.2546

5. Expected Value and Conditional Probability

a. *i.* Expected value is ${}^{2}/_{56} = 2.4$
Therefore, she should expect to get 2 prizes.

ii. ${}^{3}/_{6} = {}^{1}/_{2}$

iii. To make the experimental probability agree with the theoretical probability you need to conduct more trials.

b. *i.* ${}^{4}/_{7} = 0.5714$ (4 d.p.)

ii. ${}^{11}/_{13} = 0.8462$ (4 d.p.)

iii. ${}^{4}/_{5} = 0.8$

iv. ${}^{5}/_{20} = {}^{1}/_{4} = 0.25$

c. *i.* 0.45

ii. 0.35

iii. P(no raincoat | rain) = $\frac{\text{P(no raincoat AND rain)}}{\text{P(rain)}}$

P(no raincoat AND rain) $= 0 \times 20 \times 65 = 0.13$
P(rain) = P(no raincoat AND rain) + P(raincoat AND rain)
$= 0 \times 20 \times 65 + 0 \times 80 \times 45$
$= 0.49$

P(no raincoat | rain) = ${}^{0.13}/_{0.49}$
$= 0.2653$ (4 d.p.)

iv. P(raincoat | fine) = $\frac{\text{P(raincoat AND fine)}}{\text{P(fine)}}$

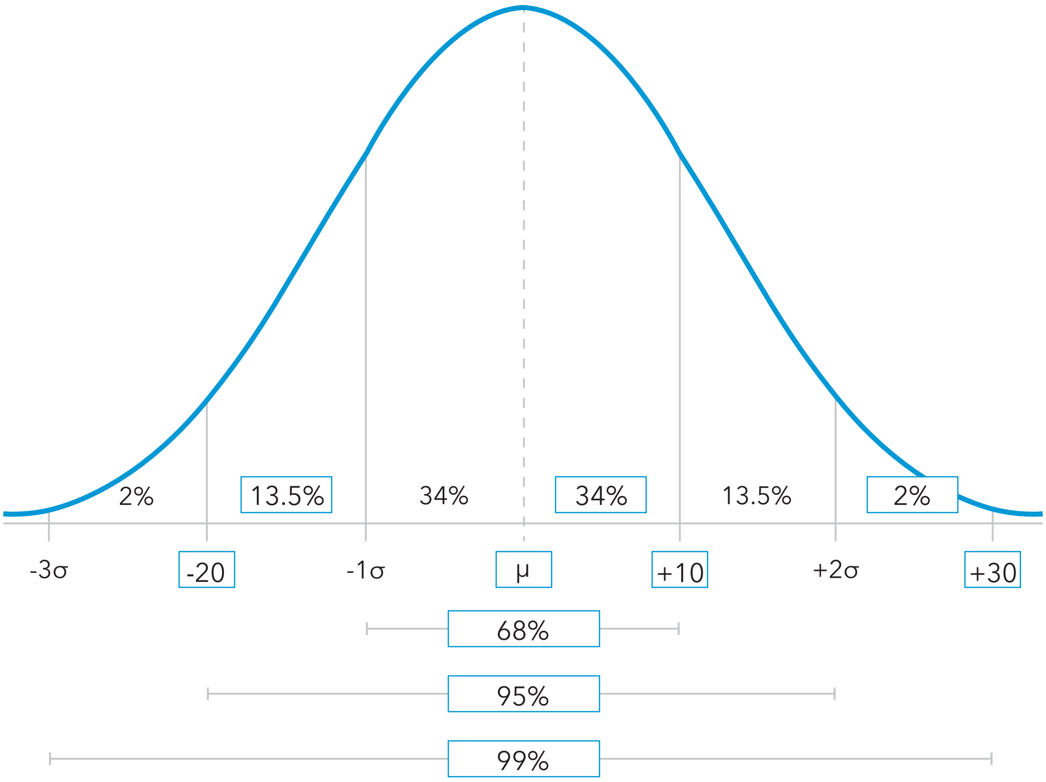
P(raincoat AND fine) $= 0 \times 80 \times 55 = 0.44$
P(fine) = P(no raincoat AND fine) + P(raincoat AND fine)
$= 0 \times 20 \times 35 + 0 \times 80 \times 55$
$= 0.51$

P(no raincoat | rain) $= {}^{0.44}/_{0.51}$
$= 0.8627$ (4 d.p.)

6. Normal Distribution

6.1 Normal Distribution

a.



b. *i.* Between 64 and 80 seconds.

ii. Billy's data is not symmetric. The data is skewed to the right which means the variation in the longer wait time values is larger than the shorter wait time values. If the data was symmetric the variation in the shorter wait times would be similar to the variation in the longer wait times.

iii. Billy's data is not normally distributed because it is not symmetrically distributed around the mean of the data.

c. *i.* $82 - 90 \text{ grams}$

ii. Figure one range: $55 - 96 \text{ grams}$
Figure two range: $58 - 94 \text{ grams}$

iii. The distribution of the weights of Lauren's donuts is not symmetrical. The data has a left skew, which means there is a tail in the data towards the left. This means that there is more variation in the lighter weight donuts than there is in the heavier weight donuts. In order for the weight of Lauren's donuts to be symmetrical, then the variation in the lighter weights needs to be the same as the heavier weights.

iv. The mean weight if Figure two will be lower than the median. This will be lower because there is more variation in the weights to the left of the mean which pulls the mean to the left of the median. The median is not effected by the skew in the data because it takes the middle value, whereas to calculate the mean you need to sum all the values and divide by the number of points so the presence of small and high values will influence the final mean.

v. The mean will be the same as the median because the data is symmetrical. There are no extreme weights in the dataset that pull the mean to the left or the right of the mean.

6.2 Standard Normal Distribution

a. Z: z-score.

X: original random variable (the value we are trying to calculate).

μ: mean.

σ: standard deviation.

b. *i.* $\mathrm{P}(11.0 < \mathrm{X} < 11.5)$

ii. $\mathrm{Z}_1 = \frac{11.0 \quad - 10.7}{0.3} = 1$

$\mathrm{Z}_2 = \frac{11.0 \quad - 10.7}{0.3} = 2.667 \text{ (3 d.p.)}$

Note: we use 3 decimal places as that is the accuracy that the z-table provided has.
Probability statement with standardised values:

$\mathrm{P}(1 < \mathrm{Z} < 2.667)$

iii. $\mathrm{P}(11.0 < \mathrm{X} < 11.5) = \mathrm{P}(1 < \mathrm{Z} < 2.667)$
$= \mathrm{P}(\mathrm{Z} < 2.667) - \mathrm{P}(\mathrm{Z} < 1)$

From the z-table:
$\mathrm{P}(\mathrm{Z} < 2.667) = 0.4962$
$\mathrm{P}(\mathrm{Z} < 1) = 0.3413$
$\mathrm{P}(\mathrm{Z} < 2.67) - \mathrm{P}(\mathrm{Z} < 1) = 0.4962 - 0.3413 \quad = 0.1549$

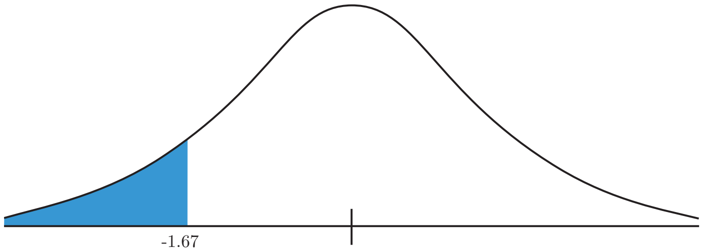
c. i. $\mathrm{P}(\mathrm{X} < 825)$

ii. $\mathrm{Z} = \frac{825 \quad - 950}{75} = -\ 1.667 \quad \text{(2 d.p.)}$

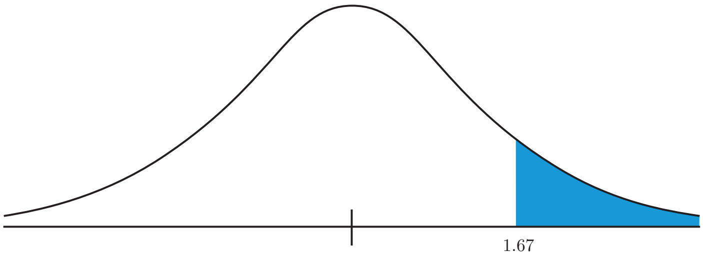
$\mathrm{P}(\mathrm{Z} < \text{-}1.667)$

iii. $P(Z < -1.667) = 0.5 - P(Z < 1.667)$ on the z-table.
$P(Z < -1.667) = 0.5 - 0.4522$
$P(Z < -1.667) = 0.0478$

If you struggle with this, try to visualise the normal distribution in your head. We are trying to find the shaded area:



But the z-tables only show probabilities between positive z values and 0. Luckily, our shaded area is equivalent to:



And this area is half of the area under the curve minus the area given in the z-table, hence we have:
$P(Z < -1.667) = 0.5 - P(Z < 1.667)$

d. The minimum weight of an extra-large loaf is $801.26g$.

We are interested in finding x, where:
$P(X > x) = 0.1$

If we standardise the distribution, we can use the z-table to find a value z, where:
$P(Z > z) = 0.1$
$P(Z < z) = 0.9 = 0.5 + 0.4$

So we are looking for the z value that gives 0.4 on the z-table:
$z = 1.2815$

Now we can find x from z:
$z = \frac{X - \mu}{\sigma}$

$x = z\sigma + \mu = 1.2815 \times 40 + 750 = 801.26g$

Answers

Answers

e. The mean weight of the cookies is 59.82g (2 d.p.)

$P(X > 65) = 0.15$
$P(X < 65) = 0.85 = 0.5 + 0.35$

We can use the z-table to find the z value that 65 corresponds to. The z value that we are looking for has a probability of 0.35 on the z-table:

$z \approx 1.0365$

Now we can find the mean:

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

$\mu = x - z\sigma$
$= 65 - 1.0365 \times 5 = 59.82\text{g}$ (2 d.p.)

f. The standard deviation of the brownies weight is 35.67g (2 d.p.).

$P(X < 100) = 0.2$

Again, we need to find the z value that corresponds to the given x value.
Since x in this case is less than the mean, the z-tables as they're given on the formula sheet don't apply. We have to find the equivalent z value that is greater than the mean. What the z-tables look at is the area between the z value and the mean. Let's look at what this probability is for our value of x:

$P(100 < X < 130) = 0.3$
So, let's find the z value that corresponds to 0.3:
$z \approx 0.841$

But, since the x value is less than the mean, we want a negative z value:
$z \approx -0.841$

Now we can use this to find the standard deviation:
$z = \frac{X - \mu}{\sigma}$

$\sigma = \frac{X - \mu}{z} = \frac{100 - 130}{-0.841} = 35.67\text{g}$ (2 d.p.)

7. Risk/Relative Risk and Analysing Claims

a. *i.* $^{591}/_{1500} = 0.394$
There is a 39.4% risk of a person needing glasses.

ii. Risk of someone needing glasses if they eat carrots:
$^{237}/_{1020}$ or 0.2324

Risk of someone needing glasses if they don't eat carrots:
$^{354}/_{480}$ or 0.7375

Relative risk:
$\frac{0.2324}{0.7375} = 0.3151 \text{ (4 d.p.)}$
OR $\frac{0.7375}{0.2324} = 3.1741 \text{ (4 d.p.)}$

EITHER: It seems that people who eat carrots daily are 0.3151 times more likely to need glasses than people who don't.
OR: It seems that people who do not eat carrots daily are 3.1741 times more likely to need glasses than people who do.

iii. Risk of someone not needing glasses if they eat carrots: $783/1020$
Risk of someone not needing glasses if they don't eat carrots: $126/480$

Relative risk:
$\frac{783\ /\ 1020}{126\ /\ 480} = 2.9244 \text{ (4 d.p.)}$
OR $\frac{126\ /\ 480}{783\ /\ 1020} = 0.3420 \text{ (4 d.p.)}$

EITHER: It seems that people who eat carrots daily are 0.3151 times more likely to not need glasses than people who don't.
OR: It seems that people who do not eat carrots daily are 3.1741 times more likely to not need glasses than people who do.

b. The risk of someone who does not eat carrots every day needing glasses is 3.174 times that of someone who eats carrots every day. The claim made by the farmer is greatly exaggerated as the claim is almost double what the relative risk was calculated to be.

c. By standardising the distribution, we can calculate what proportion of chocolate bars weigh at least 177g:

$\text{z} = \frac{\text{X} - \mu}{\sigma} = \frac{177 - 180}{5} = -0.6$
$\text{P}(\text{Z} > -0.6) = \text{P}(\text{Z} < 0.6) + 0.5$

From the z-tables:
$\text{P}(\text{Z} > -0.6) = 0.2258 + 0.5 = 0.7258$
So 72.58% of the chocolate bars weigh at least 177g. Since this is less than 85%, the claim is false.

Section Two
Putting it into Context

1. Calculating probabilities from two-way tables:

1. A sheep farmer wants to see if there is a connection between the weather and the number of lambs a ewe has. He records the data in the table below.

	Fine	Rain	Snow	**Total**
No One lamb	250	300	150	700
Two lambs	75	180	20	275
Three or more lambs	5	13	9	27
Total	330	493	179	1002

a. *i.* $^{250}/_{330}$ or 0.7576

ii. $^{20}/_{179}$

iii. Multiply the two probabilities.

iv. $(^{250}/_{330}) \times (^{20}/_{179}) = 0.0843$ (4 d.p.)

b. *i.* $^{300}/_{493}$ or 0.6085

ii. $^{180}/_{493}$ or 0.3651

iii. $\frac{0.6085}{0.3651} = 1.6667$ (4 d.p.)

iv. I disagree with the statement made by the farmer as it is 1.667 times more likely to have one lamb born in the rain than having 2 lambs born in the rain, which is not 3 times more likely.

c. *i.* $^{5}/_{27} = 0.8$

ii. $^{250}/_{700} = 0.3571$ (4 d.p.)

iii. It is more like to be raining when three or more lambs are born, as 0.1851 < 0.3571

Answers

2. A survey of Year 9 and Year 13 students was undertaken to see whether the students make their own lunch or whether their parents make their lunch for school. The results are displayed in the table below.

	Makes their own lunch	Parents make their lunch	Total
Year 9	66	143	209
Year 13	64	78	142
Total	130	221	351

a. *i.* ${}^{64}/_{130}$ or 0.4923

ii. ${}^{66}/_{209}$ or 0.3158

iii. Multiply the two probabilities.

iv. $({}^{64}/_{130}) \times ({}^{66}/_{209}) = 0.8081$ (4 d.p.)

b. *i.* ${}^{143}/_{209}$ or 0.6842

ii. ${}^{78}/_{142}$ or 0.5493

iii. $\frac{0.6842}{0.5493} = 1.2456$ (4 d.p.)

iv. I agree with the claim. The parent of a Year 9 student making their lunch is 1.2456 times more likely than the parent of a Year 13 student, or roughly 25% more likely.

c. *i.* The probability of an event occurring multiplied by the number of times the test is run in this case, the number of students.

ii. ${}^{78}/_{142}$ or 0.5493

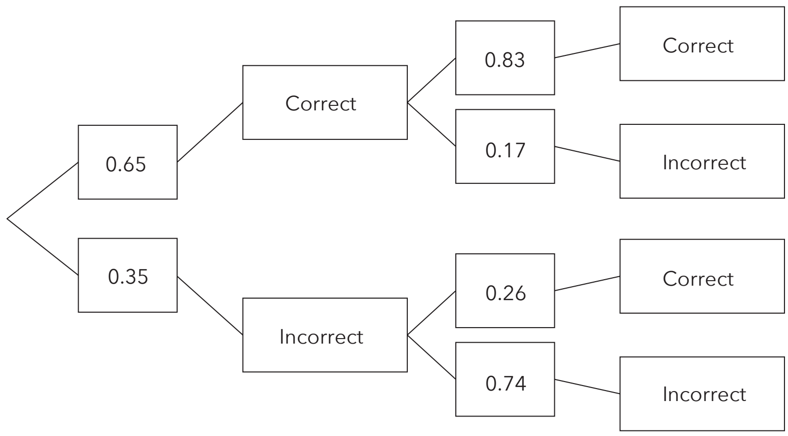
iii. $1450 \times ({}^{78}/_{142}) = 796.4789 \approx 796$

iv. In the city we would expect that 796 Year 13 students have their lunch made by their parents.

2. Calculating probabilities from probability trees:

1. A meteorologist will get the weather forecast correct 65% of the time. If he correctly predicted the weather, then the probability he correctly predicts the next day increases to 83%. If he did not correctly predict the weather, then the probability that he gets the next day wrong as well as 74%.

a. *i.*



ii. He could be correct the first day and incorrect the second day, or incorrect the first day and correct the second day.

iii. $0.65 \times 0.17 = 0.1105$

iv. $0.35 \times 0.26 = 0.091$

v. Add the two probabilities.

vi. $0.1105 + 0.091 = 0.2015$

b. *i.* $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

A is predicting correctly on the first day.
B is predicting correctly on the second day.

ii. $0.65 \times 0.83 = 0.5395$

iii. $0.35 \times 0.26 = 0.0910$

iv. $(0 \times 650 \times 83) + (0 \times 350 \times 26) = 0.6305$

v. $\frac{(0.65 \times 0.83)}{0.6305} = 0.8557 \text{ (4 d.p.)}$

Answers

c. *i.* $\text{P(Incorrect 3rd day)} = \frac{0.15}{0.35 \times 0.74} = 0.5792 \text{ (4 d.p.)}$

ii. $\text{Correct} = 1 - 0.5792 = 0.4208 \text{ (4 d.p.)}$

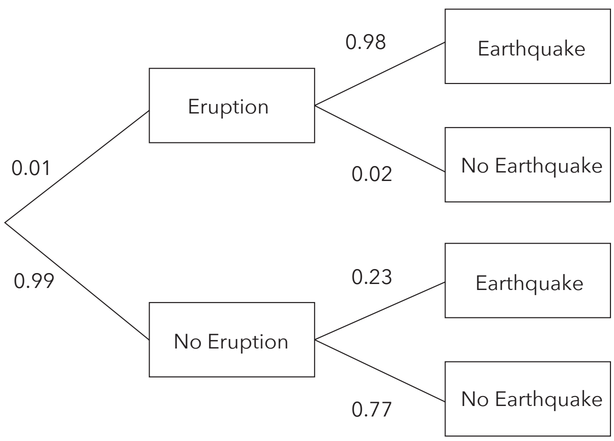
iii. $0.35 \times 0.74 \times \text{P(Incorrect 3rd day)} = 0.25$
$\text{P(Incorrect 3rd day)} = \frac{0.15}{0.35 \times 0.74} = 0.9653 \text{ (4 d.p.)}$

iv. $\text{Correct} = 1 - 0.9653 = 0.0347 \text{ (4 d.p.)}$

v. The largest probability that he is correct on the third day is 0.4208 (4 d.p)

2. The probability that Mt Ruapehu erupts on any given day is 0.01. If Mt Ruapehu erupts the probability of a minor earthquake occurring is 0.98, but if it does not erupt then the probability of an earthquake swarm occurring is 0.23.

a. *i.*



ii. $0.01 \times 0.98 = 9.8 \times 10^{-3}$

iii. $0.99 \times 0.23 = 0.2277$

iv. Add the two probabilities.

v. $9.8 \times 10^{-3} + 0.2277 = 0.2375$

b. *i.* $0.99 \times 0.23 \times 0.2 = 0.0455 \text{ (4 d.p.)}$

ii. $0.99 \times 0.23 \times 0.22 = 0.0501 \text{ (4 d.p.)}$

ii. The probability of there being no eruption, an earthquake swarm and a large magnitude earthquake is between 0.0455 and 0.0501.

Answers

c. *i.* The probability of the event occurring multiplied by the number of times the test is run. In this case, the number of days.

ii. $0.01 \times 0.98 = 9.8 \times 10^{-3}$ (1 d.p.)

iii. $9.8 \times 10^{-3} \times (365 \times 5) = 17.885 \approx 18$

iv. Over the 5-year period you would expect to observe an eruption and a minor earthquake on 18 days.

3. Using the Normal Distribution to Calculate Probability:

1. A candy factory has a machine that automatically packages lollies. The weight of the packets is normally distributed with a mean of 180g and a standard deviation of 14g.

a. *i.* $P(160 < x < 185)$

ii. $Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{160 - 180}{14} = -1.429$ (3 d.p.)

$Z_2 = \frac{X_1 - \mu}{\sigma} = \frac{185 - 180}{14} = 0.357$ (3 d.p.)

iii. $P(160 < x < 185) = P(-1.429 < z < 0.357)$

$P(160 < x < 185) = P(z < 0.357) - P(z < -1.429)$

$P(160 < x < 185) = P(z < 0.357) + P(z < 1.429)$

From the z-tables:

$P(z < 0.3571) = 0.13706$

$P(z < 1.429) = 0.42233$

$P(160 < x < 185) = 0.13706 + 0.42233 = 0.5594$ (4 d.p.)

b. *i.* Probability is $0.5 - 0.12 = 0.38$

ii. From the z-tables:

$z = 1.175$

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

$x = z\sigma + \mu$

$= 1.175 \times 14 + 180 = 196.45$

iv. The minimum weight of the heaviest 12% of bags is 196.45g

c. *i.* Probability is $0.5 - 0.3 = 0.2$

Answers

ii. From the z-tables:

$z \approx 0.524$

iii. $\sigma = \frac{x - \mu}{z} = \frac{185 - 180}{0.524} = 9.5420$ (4 d.p.)

iv. The range of the standard deviations for the machine is between $9.5420g$ to $14g$ because the standard deviation is smaller for this machine than the old machine.

2. A farmer is interested in how much pollen each bee is bringing back to the hive. He sets up a set of very sensitive scales at the entrance of the hive and automatically weighs each bee as they re-enter the hive. He observes that the amount of pollen each bee brings back is normally distributed with a mean of $2.2g$ and a standard deviation of $0.1g$.

a. *i.* $P(x < 1.9)$

ii. $z = \frac{x - \mu}{\sigma} = \frac{1.9 - 2.2}{0.1} = -3$

iii. $P(x < 1.9) = P(z < -3)$

Converting this into a format that can be used with the z-table:

$P(x < 1.9) = 0.5 - P(z < 3)$

$P(x < 1.9) = 0.5 - 0.4987$

$P(x < 1.9) = 1.3 \times 10^{-3}$

b. *i.* The probability we need is also 0.25. However since we are interested in a value less than the mean, once we find a z value, we will need to make it negative.

ii. From the z-table:

$z = -0.662$

iii. $x = z\sigma + \mu = -0.662 \times 0.1 + 2.2 = 2.1338g$

iv. The maximum weight of the lightest 25% of pollen brought back to the hive is 2.1338g.

c. *i.* $z = \frac{x - \mu}{\sigma} = \frac{2.35 - 2.2}{0.1} = 1.5$

ii. $P(x > 2.45) = P(z > 1.5)$

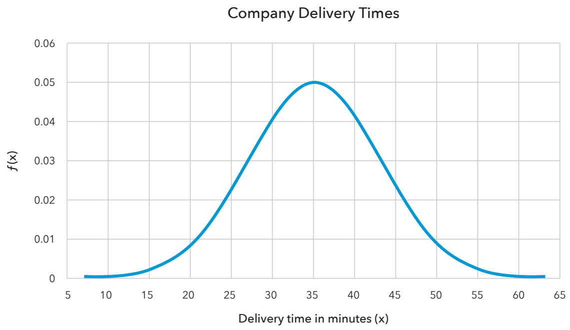
$P(x > 2.45) = 0.5 - P(z < 1.5)$

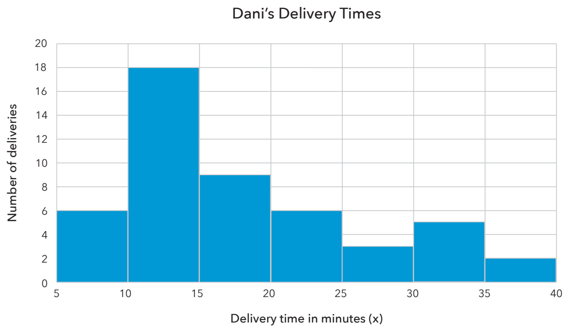
$P(x > 2.45) = 0.5 - 0.4332 = 0.0668$

iii. Multiply the two probabilities.

iv. $0.0668 \times 0.0668 = 4.4622 \times 10^{-3}$ (4 d.p.)

3. Dani delivers pizzas to save some money for university. The company he works for says that the mean delivery time for the pizza is 35 minutes with a standard deviation of 8 minutes. This is displayed in Figure one. Dani wants to see how his delivery times compare to what the company states, so recorded the time it took to make his deliveries over the week. Figure two is a histogram of his delivery times.





a. *i.* In Figure one, the shape of the probability distribution is symmetrical. This means that the probability of having a very short or very long delivery time is very small and the probability of the pizza delivery time being close to the mean delivery time of 35 minutes is high.

ii. The shape of the histogram of Dan's delivery times is not symmetrical. It is skewed to the right. This indicates that the probability that the delivery time is short is high and the probability of a long delivery time is low. There is more variation in the higher delivery times than in the shorter delivery times.

iii. Due to the symmetry in the probability distribution, the mean and the median are going to be the same. The mean/median of the distribution in 35 minutes.

iv. Due to the skew in the data the mean and median are going to be different. Dani's median delivery time is between 10 and 15 minutes, whereas the mean delivery time will be slightly longer (between 15 and 20 minutes).

v. The spread of the probability distribution is between in Figure one is between 7 and 63 minutes.

iv. The spread of Dan's delivery times in Figure two is between 5 and 40 minutes.

b. The company wants Dani's to alter his delivery times so that the times closely match the probability distribution. What does Dani have to do to make this happen and include a sketch of a potential new histogram of Dani's new delivery times?

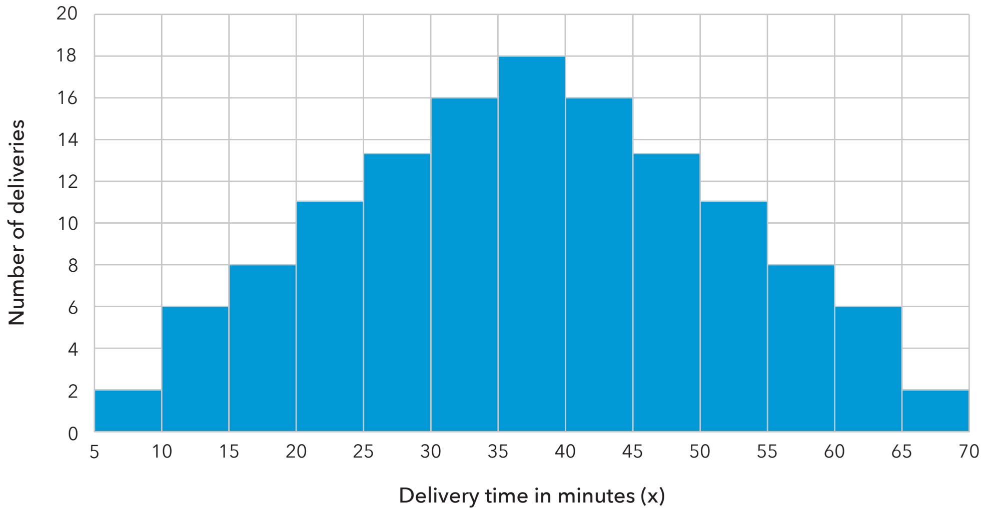
i. Dani's mean delivery times are shorter than the companies mean delivery times. Dani's delivery times are not symmetrical, so are skewed to the right resulting in higher variation in the longer delivery times than the shorter delivery times.

ii. Dani's delivery times should be centred around 35 minutes.

iii. Dani needs to increase the length of time it takes to deliver the pizza. By increasing the length of time it takes to deliver the pizza the centre of the distribution will shift to the right to better reflect the probability distribution.

iv. Dani needs to make sure that most of his deliveries are around 35minutes. However, he needs to make sure that a small number of deliveries occur in a short time (less than 20 minutes) as well as ensure that a small number of deliveries occur in a much longer time frame (greater than 50 minutes).

v. Finally, include a sketch of one possible histogram of Dani's new delivery times?



Section Three
Practice Exam

Question One:

a. i. P(no eggs from two egg nests = $0.43 \times 0.62 = 0.2666$

ii. P(at least one chick) = P(one egg and hatch) + P(two eggs and one hatch) + P(two eggs and two hatch)
$= 0.57 \times 0.54 + 0.43 \times 0.25 + 0.43 \times 0.13 = 0.4712$

iii. P(two eggs | one hatch) = P(two eggs and one hatch) / P(one hatch)
P(two eggs | one hatch) = $\frac{0.43 \times 0.25}{0.57 \times 0.54 + 0.43 \times 0.25}$
P(two eggs | one hatch) = 0.2588 (4 d.p.)

b. Let's consider the two ends of the range we are given:

P(one egg and hatch and maturity) = 0.2
P(one egg) × P(hatch) × P(maturity) = 0.2
P(maturity) = $\frac{0.2}{0.57 \times 0.54} = 0.6498$ (4 d.p.)
P(not making it to maturity) = 1 - 0.6498 = 0.3502 (4 d.p.)

P(one egg and hatch and maturity) = 0.25
P(one egg)P(hatch)P(maturity) = 0.25
P(maturity) = $\frac{0.2}{0.57 \times 0.54} \times 0.54 = 0.8122$ (4 d.p.)
P(not making it to maturity) = 1 - 0.8122 = 0.1878 (4 d.p.)

The smallest probability that a kororā chick will not make it to maturity is 0.1878.

c. C = chick and N= no chick
P(successful) = P(CC) + P(CNC) + P(CNNC) + P(NCC) + P(NCNC) + P(NNCC)
$= (0.63^2) + (0.63 \times 0.370.63) + (0.63 \text{ x } 0.37^2 \times 0.63) + (0.37 \times 0.63^2) + (0.37 \times 0.63 \times 0.37 \times 0.63) + (0.37^2 + 0.632)$
= 0.8536 (4 d.p.)

Question Two:

a. *i.* P(underweight) = $^{217}/_{800} = 0.2713$ (4 d.p.)

ii. P(female) = $^{379}/_{800} = 0.47375$ (5 d.p.)
Expected number of females = $0.47375 \times 1400 = 663.25 = 663$

iii. P(female is healthy weight) = 266/379 = 0.7018 (4 d.p.)
P(male is healthy weight) = 317/421 = 0.7530 (4 d.p.)

It is more likely that male kororā chicks are going to be of a healthy weight than female kororā chicks.

iv. $P(\text{female is underweight}) = {}^{113}/_{379} = 0.2982$ (4 d.p.)
$P(\text{male is underweight}) = {}^{104}/_{421} = 0.2470$ (4 d.p.)
Relative risk $= \frac{0.2982}{0.2470} = 1.2069$ (4 d.p.)

The claim made by DOC is incorrect. While female kororā chicks are more likely to be under-weight they are not 30% more likely to be under-weight.

b. i. If 32 died, then 583 - 32 = 551 survived.

$P(\text{healthy weight kororā chick survived}) = {}^{551}/_{583} = 0.9451$ (4 d.p.).

ii. $P(\text{under-weight and male}) = {}^{104}/_{800}$
$P(\text{under-weight and survives}) = {}^{60}/_{217}$
$P(\text{under-weight and male and survives}) = ({}^{104}/_{800}) \times ({}^{60}/_{217}) = 0.0359$ (4 d.p.)

iii. $P(\text{under-weight and survives}) = {}^{60}/_{217}$
$P(\text{healthy weight and survives}) = {}^{551}/_{583}$
Relative risk $= \frac{0.9451}{0.2765} = 3.4182$ (4 d.p.)

The claim made by the bird expert is incorrect as healthy weight kororā chicks are only 3.4182 times more likely to survive than under-weight kororā chicks.

Question Three:

a. i. $Z_1 = \frac{x_1 - \mu}{\sigma} = \frac{980 - 975}{34} = 0.147$ (3 d.p.)
$Z_2 = \frac{x_2 - \mu}{\sigma} = \frac{1005 - 975}{34} = 0.882$ (3 d.p.)
$P(980 < x < 1005) = P(0.147 < z < 0.882)$
$P(980 < x < 1005) = P(z < 0.882) - P(z < 0.147)$
$P(980 < x < 1005) = 0.3112 - 0.05598$
$P(980 < x < 1005) = 0.2552$ (4 d.p.)

ii. $Z = X\text{-}\mu = 900\text{-}97534 = -2.206$ (3 d.p.)
$P(x < 900) = P(z < -2.206)$

Converting this to a form that we can use with the provided z-table:
$P(x < 900) = 0.5 - P(z < 2.206)$
$P(x < 900) = 0.5 - 0.4863 = 0.0137$

b. To find the z value for heaviest 5%, we need to find the z value that corresponds to the probability of 0.45 in the z-table:
$z_{max} = 1.645$
To find the z value for lightest 20%, we need to find the z value that corresponds to the probability of 0.3 in the z-table, then make it negative:
$z_{min} \approx -0.841$

Minimum weight of overweight chicks:

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

$x = z\sigma + \mu = 1.645 \times 34 + 975 = 1030.93$

Maximum weight of underweight chicks:

$x = z\sigma + \mu = -0.841 \times 34 + 975 = 946.406$

The zoo considers a healthy weight range for the chicks to be between 946.406g and 1030.93g.

c. *i.* The chicks' weights have the same shaped distribution that has the same variation around the mean. However, the centre of the distribution (or the mean of the distribution) is shifted to the left of the zoo's distribution of weights.

ii. 40% of the New Zealand chicks weigh more than 975g. What is the mean weight of the New Zealand chicks?

To find the z value for heaviest 40%, we need to find the z value that corresponds to the probability of 0.1 in the z-table:

$z \approx 0.5398$

Now finding the mean weight of the New Zealand chicks:

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

$- \mu = z\sigma - x$

$\mu = x - z\sigma = 975 - (0.251 \times 34) = 966.466$

Possible comments:

Shape:

- Symmetrical vs not symmetrical.
- Higher probabilities of smaller and larger weights in Figure Two compared to Figure One.
- Both are unimodal.

Centre:

- Medians [975g vs less than 967] (do not accept modes)
- Means are very similar [975 vs 976.13]

Spread:

- Ranges of the two data sets [856-1096 vs 907-1067]