Question 1

Chiara records the length of time in minutes that it takes for members of a population of cats to find their ways from the edge of a maze to its centre. The time taken, T, follows a normal distribution where $T \sim N(\mu, \sigma^2)$.

Chiara selects a random sample of 7 of the results, these results are displayed below.

4.3, 2.9, 3.5, 4.1, 3.6, 3.1, 3.7

(a) Determine

(i) the mean of the sample

(ii) the variance of the sample.

[3]

(b) Hence find

(i) an unbiased estimate of the mean of the population

(ii) an unbiased estimate of the variance, ${s_{n-1}}^2$ of the population.

[3]

It is subsequently discovered that the true standard deviation, σ, for the population is 0.23 minutes.

(c) Find a 95% confidence interval for the mean of the population.

[2]

a) $n = 7, \ \Sigma x = 25.2, \ \Sigma x^2 = 92.22$

(i) mean $= \frac{25.2}{7} = \boxed{3.6 \text{ minutes}}$

(ii) Remember, variance is (standard deviation)^2

The exact value is:

$$\text{variance} = \frac{92.22 - \frac{25.2^2}{7}}{7} = \frac{3}{14} = 0.214285...$$

s_n from GDC

$$s_n^2 = (0.46291004)^2 = 0.214285...$$

$$\boxed{0.214 \text{ minutes}^2 \ (3 \text{ s.f.})}$$

Chiara records the length of time in minutes that it takes for members of a population of cats to find their ways from the edge of a maze to its centre. The time taken, T, follows a normal distribution where $T \sim N(\mu, \sigma^2)$.

$n = 7$

Chiara selects a random sample of 7 of the results, these results are displayed below.

4.3, 2.9, 3.5, 4.1, 3.6, 3.1, 3.7

(a) Determine

(i) the mean of the sample $\boxed{3.6 \text{ minutes}}$

(ii) the variance of the sample.

$\boxed{(0.46291004)^2 = 0.214 \text{ minutes}^2 \ (3 \text{ s.f.})}$

[3]

(b) Hence find

(i) an unbiased estimate of the mean of the population

(ii) an unbiased estimate of the variance, ${s_{n-1}}^2$ of the population.

[3]

It is subsequently discovered that the true standard deviation, σ, for the population is 0.23 minutes.

(c) Find a 95% confidence interval for the mean of the population.

[2]

Unbiased estimate of population variance s_{n-1}^2	$s_{n-1}^2 = \frac{n}{n-1} s_n^2$

b) (i) The sample mean is an unbiased estimator of the population mean.

$$\boxed{3.6 \text{ minutes}}$$

(ii) $s_{n-1}^2 = \frac{7}{6}(0.46291004)^2 = 0.249999...$

$$\boxed{0.250 \text{ minutes}^2 \ (3 \text{ s.f.})}$$

Exact value is $\frac{7}{6} \times \frac{3}{14} = 0.25$

Chiara records the length of time in minutes that it takes for members of a population of cats to find their ways from the edge of a maze to its centre. The time taken, T, follows a normal distribution where $T \sim N(\mu, \sigma^2)$.

Chiara selects a random sample of 7 of the results, these results are displayed below.

4.3, 2.9, 3.5, 4.1, 3.6, 3.1, 3.7

(a) Determine

(i) the mean of the sample

(ii) the variance of the sample.

[3]

(b) Hence find

(i) an unbiased estimate of the mean of the population $\boxed{3.6 \text{ minutes}}$

(ii) an unbiased estimate of the variance, $s_{n-1}{}^2$ of the population.

[3]

It is subsequently discovered that the true standard deviation, σ, for the population is 0.23 minutes.

(c) Find a 95% confidence interval for the mean of the population.

[2]

c) The population variance is known, so use the Normal (Z) version of the confidence interval calculator on your GDC.

$n = 7 \quad \bar{x} = 3.6 \quad \sigma = 0.23$

Lower $= 3.42961675$

Upper $= 3.77038325$

In minutes to 3 s.f., the 95% confidence interval is

$(3.43,\ 3.77)$

Question 2

Let D be a normally distributed random variable that represents the distance travelled in metres by a slug in one day. The distance covered by a random sample of 21 slugs on a randomly selected day can be summarized as follows

$n = 21$

$$\Sigma d = 341, \quad \Sigma d^2 = 5881.$$

(a) Find an unbiased estimate of the mean, μ, of D.

[1]

(b) Use the formula $s_{n-1}{}^2 = \dfrac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$ to find an unbiased estimate of the variance of D.

[2]

(c) Find a 95% confidence interval for μ.

[2]

Justin believes that the average slug travels 15 m per day.

(d) State whether or not Justin's statement is valid. Give a reason for your answer.

[2]

a) The sample mean is an unbiased estimator of the population mean.

$$\mu = \frac{\Sigma d}{n} = \frac{341}{21} = 16.238095\ldots$$

16.2 metres (3 s.f.)

Let D be a normally distributed random variable that represents the distance travelled in metres by a slug in one day. The distance covered by a random sample of 21 slugs on a randomly selected day can be summarized as follows

$n = 21$

$$\Sigma d = 341, \quad \Sigma d^2 = 5881.$$

(a) Find an unbiased estimate of the mean, μ, of D.

[1]

(b) Use the formula $s_{n-1}{}^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$ to find an unbiased estimate of the variance of D.

[2]

(c) Find a 95% confidence interval for μ.

[2]

Justin believes that the average slug travels 15 m per day.

(d) State whether or not Justin's statement is valid. Give a reason for your answer.

[2]

b) $s_{n-1}{}^2 = \dfrac{5881 - \frac{341^2}{21}}{20} = \dfrac{361}{21} = 17.190476...$

$\boxed{17.2 \text{ metres}^2 \ (3 \text{ s.f.})}$

Let D be a normally distributed random variable that represents the distance travelled in metres by a slug in one day. The distance covered by a random sample of 21 slugs on a randomly selected day can be summarized as follows

$n = 21$

$$\Sigma d = 341, \quad \Sigma d^2 = 5881.$$

(a) Find an unbiased estimate of the mean, μ, of D.

$\boxed{\frac{341}{21} = 16.2 \text{ metres } (3 \text{ s.f.})}$

[1]

(b) Use the formula $s_{n-1}{}^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$ to find an unbiased estimate of the variance of D.

$\boxed{\frac{361}{21} = 17.2 \text{ metres}^2 \ (3 \text{ s.f.})}$

[2]

(c) Find a 95% confidence interval for μ.

[2]

Justin believes that the average slug travels 15 m per day.

(d) State whether or not Justin's statement is valid. Give a reason for your answer.

[2]

c) The population variance is unknown, so use the t-test version of the confidence interval calculator on your GDC.

$n = 21 \qquad \bar{x} = \frac{341}{21} \qquad s_{n-1} = \sqrt{\frac{361}{21}}$

Lower $= 14.350795$

Upper $= 18.1253955$

In metres to 3 s.f., the 95% confidence interval is

$(14.4, 18.1)$

Let D be a normally distributed random variable that represents the distance travelled in metres by a slug in one day. The distance covered by a random sample of 21 slugs on a randomly selected day can be summarized as follows

$$\Sigma d = 341, \quad \Sigma d^2 = 5881.$$

(a) Find an unbiased estimate of the mean, μ, of D. $\boxed{\frac{341}{21} = 16.2 \text{ metres } (3 \text{ s.f.})}$

[1]

(b) Use the formula $s_{n-1}{}^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$ to find an unbiased estimate of the variance of D.

[2]

(c) Find a 95% confidence interval for μ.

In metres to 3 s.f., the 95% confidence interval is

$(14.4, 18.1)$

[2]

Justin believes that the average slug travels 15 m per day.

(d) State whether or not Justin's statement is valid. Give a reason for your answer.

[2]

d) 15 metres is less than the oberved mean, but it is within the 95% confidence interval.

Therefore Justin's belief is valid.

Question 3

A farm grows pumpkins and transports them in crates of 24. The mass of the pumpkins follows a normal distribution with mean 7.9 kg and standard deviation 0.4 kg.

(a) Find the mean mass of a crate of pumpkins.

[2]

a) $24 \times 7.9 = \boxed{189.6 \text{ kg}}$

(b) Find the standard deviation of the mass of a crate of pumpkins.

[2]

(c) Find the probability that a crate selected at random has a mass of between 170 kg and 190 kg.

[2]

Here $X_1, X_2, \ldots, X_{24} \sim N(7.9, 0.4^2)$

Linear combinations of n independent random variables, $X_1, X_2, \ldots, X_n$	$E(a_1X_1 \pm a_2X_2 \pm \ldots \pm a_nX_n) = a_1E(X_1) \pm a_2E(X_2) \pm \ldots \pm a_nE(X_n)$ $\text{Var}(a_1X_1 \pm a_2X_2 \pm \ldots \pm a_nX_n)$ $= a_1^2\,\text{Var}(X_1) + a_2^2\,\text{Var}(X_2) + \ldots + a_n^2\,\text{Var}(X_n)$

A farm grows pumpkins and transports them in crates of 24. The mass of the pumpkins follows a normal distribution with mean 7.9 kg and standard deviation 0.4 kg.

(a) Find the mean mass of a crate of pumpkins.

[2]

(b) Find the standard deviation of the mass of a crate of pumpkins.

[2]

(c) Find the probability that a crate selected at random has a mass of between 170 kg and 190 kg.

[2]

b) $\sigma^2 = 24 \times (0.4)^2 = 3.84$

$\sigma = \sqrt{3.84} = \frac{4\sqrt{6}}{5} = 1.959591\ldots$

$\boxed{1.96 \text{ kg } (3 \text{ s.f.})}$

Here $X_1, X_2, \ldots, X_{24} \sim N(7.9, 0.4^2)$

Linear combinations of n independent random variables, $X_1, X_2, \ldots, X_n$	$E(a_1X_1 \pm a_2X_2 \pm \ldots \pm a_nX_n) = a_1E(X_1) \pm a_2E(X_2) \pm \ldots \pm a_nE(X_n)$ $\text{Var}(a_1X_1 \pm a_2X_2 \pm \ldots \pm a_nX_n)$ $= a_1^2\text{Var}(X_1) + a_2^2\text{Var}(X_2) + \ldots + a_n^2\text{Var}(X_n)$

A farm grows pumpkins and transports them in crates of 24. The mass of the pumpkins follows a normal distribution with mean 7.9 kg and standard deviation 0.4 kg.

(a) Find the mean mass of a crate of pumpkins. $\boxed{189.6 \text{ kg}}$

[2]

(b) Find the standard deviation of the mass of a crate of pumpkins.

$\boxed{\sqrt{3.84} = 1.96 \text{ kg } (3 \text{ s.f.})}$ [2]

(c) Find the probability that a crate selected at random has a mass of between 170 kg and 190 kg.

[2]

c) A linear combination of independent normal random variables also has a normal distribution.

Let X represent the mass of a crate of pumpkins.

Then $X \sim N(189.6, 3.84)$, and

$P(170 < X < 190) = 0.58087175$ from GDC

$\boxed{0.581 \text{ } (3 \text{ s.f.})}$

Question 4

The time taken for a customer services advisor to complete a phone call follows a normal distribution with mean 4.2 minutes and standard deviation 1.3 minutes.

A customer service advisor deals with 5 phone calls one after the other. It is assumed that the phone calls are independent events.

(a) Find the expected total time to complete the 5 phone calls.

[2]

(b) Find the variance of the total time to complete the 5 phone calls.

[2]

(c) Find the probability that the total time taken to complete the 5 phone calls will be more than 25 minutes.

[2]

a) $5 \times 4.2 = \boxed{21 \text{ minutes}}$

Here $X_1, X_2, ..., X_5 \sim N(4.2, 1.3^2)$

Linear combinations of n independent random variables, $X_1, X_2, ..., X_n$	$E(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n) = a_1E(X_1) \pm a_2E(X_2) \pm ... \pm a_nE(X_n)$ $Var(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n)$ $= a_1^2 Var(X_1) + a_2^2 Var(X_2) + ... + a_n^2 Var(X_n)$

The time taken for a customer services advisor to complete a phone call follows a normal distribution with mean 4.2 minutes and standard deviation 1.3 minutes.

A customer service advisor deals with 5 phone calls one after the other. It is assumed that the phone calls are independent events.

(a) Find the expected total time to complete the 5 phone calls.

[2]

(b) Find the variance of the total time to complete the 5 phone calls.

[2]

(c) Find the probability that the total time taken to complete the 5 phone calls will be more than 25 minutes.

[2]

b) $5 \times (1.3)^2 = \boxed{8.45 \text{ minutes}^2}$

Here $X_1, X_2, ..., X_5 \sim N(4.2, 1.3^2)$

Linear combinations of n independent random variables, $X_1, X_2, ..., X_n$	$E(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n) = a_1E(X_1) \pm a_2E(X_2) \pm ... \pm a_nE(X_n)$ $Var(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n)$ $= a_1^2 Var(X_1) + a_2^2 Var(X_2) + ... + a_n^2 Var(X_n)$

The time taken for a customer services advisor to complete a phone call follows a normal distribution with mean 4.2 minutes and standard deviation 1.3 minutes.

A customer service advisor deals with 5 phone calls one after the other. It is assumed that the phone calls are independent events.

(a) Find the expected total time to complete the 5 phone calls. 21 minutes

[2]

(b) Find the variance of the total time to complete the 5 phone calls. 8.45 minutes^2

[2]

(c) Find the probability that the total time taken to complete the 5 phone calls will be more than 25 minutes.

[2]

c) A linear combination of independent normal random variables also has a normal distribution.

Let X represent the total length of the five calls.

so $\sigma = \sqrt{8.45}$

Then $X \sim N(21, 8.45)$ and

$P(X > 25) = 0.08440434$ from GDC

$\boxed{0.0844 \text{ (3 s.f.)}}$

Question 5

A fisherman catches 16 fish from a local population of mackerel. He measures the fish that he catches and finds that they have mean length of 30.5 cm with standard deviation 5 cm.

$n = 16$

$s_n = 5$

(a) Find $s_{n-1}{}^2$.

[2]

(b) Find a 95 % confidence interval for the population mean.

[2]

The fisherman advertises the population from which he fishes as having an average length of 34 cm.

(c) Comment on the fisherman's claim using your answer from part (b).

[2]

a) $s_{n-1}^2 = \frac{16}{15} \times 5^2 = \frac{80}{3} = 26.666\ldots$

$\boxed{26.7 \text{ cm}^2 \text{ (3 s.f.)}}$

Unbiased estimate of population variance s_{n-1}^2	$s_{n-1}^2 = \frac{n}{n-1} s_n^2$

A fisherman catches 16 fish from a local population of mackerel. He measures the fish that he catches and finds that they have mean length of 30.5 cm with standard deviation 5 cm.

$\rightarrow n = 16$

$\rightarrow \mu = 30.5$

(a) Find $s_{n-1}{}^2$.

$\boxed{\frac{80}{3} = 26.7 \text{ cm}^2 \text{ (3 s.f.)}}$

[2]

(b) Find a 95 % confidence interval for the population mean.

[2]

The fisherman advertises the population from which he fishes as having an average length of 34 cm.

(c) Comment on the fisherman's claim using your answer from part (b).

[2]

b) The population variance is unknown, so use the t-test version of the confidence interval calculator on your GDC.

$n = 16 \quad \bar{x} = 30.5 \quad s_{n-1} = \sqrt{\frac{80}{3}}$

Lower = 27.7483105

Upper = 33.2516895

In cm to 3 s.f., the 95% confidence interval is (27.7, 33.3)

A fisherman catches 16 fish from a local population of mackerel. He measures the fish that he catches and finds that they have mean length of 30.5 cm with standard deviation 5 cm.

(a) Find $s_{n-1}{}^2$.

In cm to 3 s.f., the 95% confidence interval is (27.7, 33.3)

[2]

(b) Find a 95 % confidence interval for the population mean.

[2]

The fisherman advertises the population from which he fishes as having an average length of 34 cm.

(c) Comment on the fisherman's claim using your answer from part (b).

[2]

c) 34 cm is outside the confidence interval, so the fisherman's claim is not justified.

Question 6

A gardener is laying a pathway of pebbles from a large sack of pebbles. The mass of the pebbles is normally distributed with mean 564 g and standard deviation 57g.

(a) Find the probability that a pebble that the gardener picks at random from the sack has a mass of less than 500 g.

[2]

The gardener decides that any pebbles that have a mass greater than 620 g are "oversized" and should not be used to create the pathway.

(b) Find the probability that a pebble selected at random from the sack will be considered "oversized".

[2]

The gardener decides to select 8 of the pebbles at random.

(c) Find the probability that the mean mass of the 8 pebbles selected would fall in the "oversized" range.

[3]

a) Let X represent the mass of a randomly-selected pebble.

Then $X \sim N(564, 57^2)$, and

$P(X < 500) = 0.13075973$ from GDC

$\boxed{0.131 \text{ (3 s.f.)}}$

A gardener is laying a pathway of pebbles from a large sack of pebbles. The mass of the pebbles is normally distributed with mean 564 g and standard deviation 57g.

(a) Find the probability that a pebble that the gardener picks at random from the sack has a mass of less than 500 g.

[2]

The gardener decides that any pebbles that have a mass greater than 620 g are "oversized" and should not be used to create the pathway.

(b) Find the probability that a pebble selected at random from the sack will be considered "oversized".

[2]

The gardener decides to select 8 of the pebbles at random.

(c) Find the probability that the mean mass of the 8 pebbles selected would fall in the "oversized" range.

[3]

b) Let X represent the mass of a randomly-selected pebble.

Then $X \sim N(564, 57^2)$, and

$P(X > 620) = 0.16293759$ from GDC

$\boxed{0.163 \text{ (3 s.f.)}}$

A gardener is laying a pathway of pebbles from a large sack of pebbles. The mass of the pebbles is normally distributed with mean 564 g and standard deviation 57g.

(a) Find the probability that a pebble that the gardener picks at random from the sack has a mass of less than 500 g.

[2]

The gardener decides that any pebbles that have a mass greater than 620 g are "oversized" and should not be used to create the pathway.

(b) Find the probability that a pebble selected at random from the sack will be considered "oversized". $\boxed{0.163 \text{ (3 s.f.)}}$

[2]

$n = 8$

The gardener decides to select 8 of the pebbles at random.

(c) Find the probability that the mean mass of the 8 pebbles selected would fall in the "oversized" range.

c) If $X \sim N(\mu, \sigma^2)$, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\bar{X} \sim N\left(564, \frac{57^2}{8}\right) = N\left(564, \left(\frac{57\sqrt{2}}{4}\right)^2\right)$

$P(\bar{X} > 620) = 0.002728$ from GDC

$\boxed{0.00273 \text{ (3 s.f.)}}$

[3]

Question 7

In a busy office all workers are able to send jobs to the printer to be printed. It is assumed that each print job is an independent event, and that more than one print job does not arrive in the print queue at the same time. The number of jobs arriving in the print queue in 1 hour follows a Poisson distribution given by $X \sim Po(17)$.

(a) Find the probability that the number of print jobs sent to the printer in 1 hour is less than 15.

[2]

Helen wants to investigate the mean number of print jobs sent to the printer in an hour over the course of a working week of 35 hours.

(b) Using the central limit theorem, define a probability distribution that may be used to model the distribution of the random variable $\bar{X}$.

[2]

(c) Using the answer to part (b), find the probability that in a working week of 35 hours the mean number of print jobs in a single hour is less than 15.

[2]

Poisson distribution	$X \sim \mathrm{Po}(m)$
Mean	$\mathrm{E}(X) = m$
Variance	$\mathrm{Var}(X) = m$

a) $P(X < 15) = P(0 \le X \le 14) = 0.28083283$ from GDC

$\boxed{0.281 \ (3 \text{ s.f.})}$

In a busy office all workers are able to send jobs to the printer to be printed. It is assumed that each print job is an independent event, and that more than one print job does not arrive in the print queue at the same time. The number of jobs arriving in the print queue in 1 hour follows a Poisson distribution given by $X \sim Po(17)$.

(a) Find the probability that the number of print jobs sent to the printer in 1 hour is less than 15.

[2]

Helen wants to investigate the mean number of print jobs sent to the printer in an hour over the course of a working week of 35 hours. $n = 35$

(b) Using the central limit theorem, define a probability distribution that may be used to model the distribution of the random variable $\bar{X}$.

[2]

(c) Using the answer to part (b), find the probability that in a working week of 35 hours the mean number of print jobs in a single hour is less than 15.

[2]

Poisson distribution	$X \sim \mathrm{Po}(m)$
Mean	$\mathrm{E}(X) = m$
Variance	$\mathrm{Var}(X) = m$

b) $\frac{\sigma^2}{n} = \frac{17}{35}$ ← Variance is same as mean for Poisson

Because $n > 30$, the central limit theorem tells us that a good approximation for the distribution of the mean $\bar{X}$ will be

$\boxed{\bar{X} \sim N\left(17, \frac{17}{35}\right)}$

$N\left(\mu, \frac{\sigma^2}{n}\right)$

In a busy office all workers are able to send jobs to the printer to be printed. It is assumed that each print job is an independent event, and that more than one print job does not arrive in the print queue at the same time. The number of jobs arriving in the print queue in 1 hour follows a Poisson distribution given by $X \sim Po(17)$.

(a) Find the probability that the number of print jobs sent to the printer in 1 hour is less than 15.

$\boxed{0.281 \text{ (3 s.f.)}}$

[2]

Helen wants to investigate the mean number of print jobs sent to the printer in an hour over the course of a working week of 35 hours.

(b) Using the central limit theorem, define a probability distribution that may be used to model the distribution of the random variable $\bar{X}$.

$\boxed{\bar{X} \sim N\left(17, \frac{17}{35}\right)}$

[2]

(c) Using the answer to part (b), find the probability that in a working week of 35 hours the mean number of print jobs in a single hour is less than 15.

[2]

c) For $\bar{X} \sim N\left(17, \frac{17}{35}\right)$, so $\sigma = \sqrt{\frac{17}{35}}$

$P(\bar{X} < 15) = 0.0020542$ from GDC

$\boxed{0.00205 \text{ (3 s.f.)}}$

Remember that this is only an approximation.
The exact probability, which may be calculated using the Poisson distribution, is
$0.00162348... = 0.00162$ (3 s.f.)

Question 8

A population of leatherback turtles has a mean swimming speed of 31 km/h with standard deviation 2.3 km/h. $\rightarrow \sigma = 2.3$ $\quad \mu = 31$

(a) Using the central limit theorem, find an estimate for the probability that a sample of 32 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

[3]

(b) Also using the central limit theorem, find an estimate for the probability that a sample of 50 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

[2]

(c) Explain why there is a difference between your answers to part (a) and part (b).

[2]

a) If X has mean μ and variance σ^2, and if $n > 30$, then the central limit theorem says we can approximate the distribution of $\bar{X}$ by $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

By the central limit theorem, let $\bar{X}$ be approximated by $\bar{X} \sim N\left(31, \frac{2.3^2}{32}\right)$.

$\sigma = \frac{2.3}{\sqrt{32}}$

$P(\bar{X} > 31.8) = 0.02455695$ from GDC

$\boxed{0.0246 \text{ (3 s.f.)}}$

A population of leatherback turtles has a mean swimming speed of 31 km/h with standard deviation 2.3 km/h. $\rightarrow \sigma = 2.3$ $\quad \mu = 31$

(a) Using the central limit theorem, find an estimate for the probability that a sample of 32 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

$\boxed{0.0246 \text{ (3 s.f.)}}$

[3]

(b) Also using the central limit theorem, find an estimate for the probability that a sample of 50 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

[2]

(c) Explain why there is a difference between your answers to part (a) and part (b).

[2]

b) If X has mean μ and variance σ^2, and if $n > 30$, then the central limit theorem says we can approximate the distribution of $\bar{X}$ by $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

By the central limit theorem, let $\bar{X}$ be approximated by $\bar{X} \sim N\left(31, \frac{2.3^2}{50}\right)$.

$\sigma = \frac{2.3}{\sqrt{50}}$

$P(\bar{X} > 31.8) = 0.0069565$ from GDC

$\boxed{0.00696 \text{ (3 s.f.)}}$

A population of leatherback turtles has a mean swimming speed of 31 km/h with standard deviation 2.3 km/h.

(a) Using the central limit theorem, find an estimate for the probability that a sample of 32 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

$\boxed{0.0246 \text{ (3 s.f.)}}$

[3]

(b) Also using the central limit theorem, find an estimate for the probability that a sample of 50 leatherback turtles have a mean swimming speed greater than 31.8 km/h.

$\boxed{0.00696 \text{ (3 s.f.)}}$

[2]

(c) Explain why there is a difference between your answers to part (a) and part (b).

[2]

c) When n increases, the variance and standard deviation of $\bar{X} \sim \left(\mu, \frac{\sigma^2}{n}\right)$ decrease.

Therefore for larger n, the probability of the sample mean $\bar{X}$ being different from the population mean μ will be smaller.

Question 9

On a flower farm the height of a tulip, in centimetres, is normally distributed with mean μ and standard deviation σ. A random sample of 60 flowers is taken from the farm and can be summarised as follows

$\hookrightarrow n = 60$

$$\Sigma h = 1950, \quad \Sigma h^2 = 66075.$$

(a) Find an unbiased estimate for μ.

[1]

a) The sample mean is an unbiased estimator of the population mean.

$$\frac{\Sigma h}{n} = \frac{1950}{60} = \boxed{32.5 \text{ cm}}$$

(b) Given that $s_n{}^2 = 45$, find an unbiased estimate for the variance of the height of the tulips.

[2]

It is subsequently discovered that the actual standard deviation, σ, of the tulip population is 6.47 cm.

A second sample of 40 tulips is picked.

$\bar{H}$ denotes the mean height of the new sample.

(c) State a distribution that may reasonably be used to model $\bar{H}$.

[2]

(d) Using the answer to part (c), find an estimate for the probability that the mean height of the flowers in the new sample is between 20 and 30 cm.

[2]

On a flower farm the height of a tulip, in centimetres, is normally distributed with mean μ and standard deviation σ. A random sample of 60 flowers is taken from the farm and can be summarised as follows → $n = 60$

$$\Sigma h = 1950, \quad \Sigma h^2 = 66075.$$

(a) Find an unbiased estimate for μ. $\boxed{32.5 \text{ cm}}$

[1]

(b) Given that $s_n{}^2 = 45$, find an unbiased estimate for the variance of the height of the tulips.

[2]

It is subsequently discovered that the actual standard deviation, σ, of the tulip population is 6.47 cm.

A second sample of 40 tulips is picked.

$\bar{H}$ denotes the mean height of the new sample.

(c) State a distribution that may reasonably be used to model $\bar{H}$.

[2]

(d) Using the answer to part (c), find an estimate for the probability that the mean height of the flowers in the new sample is between 20 and 30 cm.

[2]

Unbiased estimate of population variance s_{n-1}^2	$s_{n-1}^2 = \dfrac{n}{n-1} s_n^2$

b) $s_{n-1}^2 = \dfrac{60}{59} \times 45 = \boxed{\dfrac{2700}{59} \text{ cm}^2}$

$= 45.762711...$

On a flower farm the height of a tulip, in centimetres, is normally distributed with mean μ and standard deviation σ. A random sample of 60 flowers is taken from the farm and can be summarised as follows

$$\Sigma h = 1950, \quad \Sigma h^2 = 66075.$$

This means the mean height will also be normally distributed.

(a) Find an unbiased estimate for μ. $\boxed{32.5 \text{ cm}}$

[1]

(b) Given that $s_n{}^2 = 45$, find an unbiased estimate for the variance of the height of the tulips. $\boxed{\dfrac{2700}{59} \text{ cm}^2}$

[2]

It is subsequently discovered that the actual standard deviation, σ, of the tulip population is 6.47 cm.

A second sample of 40 tulips is picked. $n = 40$

$\bar{H}$ denotes the mean height of the new sample.

(c) State a distribution that may reasonably be used to model $\bar{H}$.

[2]

(d) Using the answer to part (c), find an estimate for the probability that the mean height of the flowers in the new sample is between 20 and 30 cm.

[2]

c) If $X \sim N(\mu, \sigma^2)$, then $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$\bar{H}$ will be normally distributed. Use the true population standard deviation for σ (not an estimate based on the part (b) answer!), and use the sample mean from part (a) as an estimate for μ.

A good approximation for $\bar{H}$ will be

$$\bar{H} \sim N\left(32.5, \frac{6.47^2}{40}\right)$$

On a flower farm the height of a tulip, in centimetres, is normally distributed with mean μ and standard deviation σ. A random sample of 60 flowers is taken from the farm and can be summarised as follows

$$\Sigma h = 1950,\ \ \Sigma h^2 = 66075.$$

(a) Find an unbiased estimate for μ.

[1]

(b) Given that $s_n{}^2 = 45$, find an unbiased estimate for the variance of the height of the tulips.

[2]

It is subsequently discovered that the actual standard deviation, σ, of the tulip population is 6.47 cm.

A second sample of 40 tulips is picked.

$\bar{H}$ denotes the mean height of the new sample.

(c) State a distribution that may reasonably be used to model $\bar{H}$.

$\boxed{\bar{H} \sim N\left(32.5, \frac{6.47^2}{40}\right)}$

[2]

(d) Using the answer to part (c), find an estimate for the probability that the mean height of the flowers in the new sample is between 20 and 30 cm.

[2]

d) For $\bar{H} \sim N\left(32.5, \frac{6.47^2}{40}\right)$

$\hookrightarrow \sigma = \frac{6.47}{\sqrt{40}}$

$P(20 < \bar{H} < 30) = 0.0072667$ from GDC

$\boxed{0.00727 \ (3 \text{ s.f.})}$

Question 10

Jessamy is interested in the quality of the soil in her local area and decides to test 100 soil samples for levels of nitrogen. From her past research Jessamy knows that the level of nitrogen in an individual sample, N_i, has a mean of 41 ppm and a standard deviation of 7 ppm.

$\rightarrow \mathrm{Var}(N_i) = 7^2$ $\quad \hookrightarrow E(N_i) = 41$

Let $X = \sum_{i=1}^{100} N_i$ be the total of the levels of nitrogen in Jessamy's batch of samples.

$\hookrightarrow n = 100$

(a) Find

(i) $E(X)$

(ii) $\mathrm{Var}(X)$.

[3]

(b) Explain why a normal distribution can be used to give an approximate model for X.

[2]

(c) Use the model to find the an estimate for the value of a such that $P(X < a) = 0.1$.

[2]

a) (i) $E(X) = E(N_1 + N_2 + ... + N_{100}) = E(N_1) + E(N_2) + ... + E(N_{100})$

$\Rightarrow E(X) = 100\, E(N_i)$

$E(X) = 100 \times 41 = \boxed{4100}$

(ii) $\mathrm{Var}(X) = \mathrm{Var}(N_1 + N_2 + ... + N_{100}) = \mathrm{Var}(N_1) + \mathrm{Var}(N_2) + ... + \mathrm{Var}(N_{100})$

$\Rightarrow \mathrm{Var}(X) = 100\, \mathrm{Var}(N_i)$

$\mathrm{Var}(X) = 100 \times 7^2 = \boxed{4900}$

Linear combinations of n independent random variables, $X_1, X_2, ..., X_n$	$E(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n) = a_1E(X_1) \pm a_2E(X_2) \pm ... \pm a_nE(X_n)$ $\mathrm{Var}(a_1X_1 \pm a_2X_2 \pm ... \pm a_nX_n)$ $= a_1^2\,\mathrm{Var}(X_1) + a_2^2\,\mathrm{Var}(X_2) + ... + a_n^2\,\mathrm{Var}(X_n)$

Jessamy is interested in the quality of the soil in her local area and decides to test 100 soil samples for levels of nitrogen. From her past research Jessamy knows that the level of nitrogen in an individual sample, N_i, has a mean of 41 ppm and a standard deviation of 7 ppm.

Let $X = \sum_{i=1}^{100} N_i$ be the total of the levels of nitrogen in Jessamy's batch of samples.

$\hookrightarrow n = 100$

(a) Find

(i) E(X)

(ii) Var(X).

[3]

(b) Explain why a normal distribution can be used to give an approximate model for X.

[2]

(c) Use the model to find the an estimate for the value of a such that $\mathrm{P}(X < a) = 0.1$.

[2]

b) If X has mean μ and variance σ^2, and if $n > 30$, then the central limit theorem says we can approximate the distribution of $X_1 + X_2 + \ldots + X_n$ by $N(n\mu, n\sigma^2)$.

Because $n = 100 > 30$, the central limit theorem says we can approximate the mean of $N_1 + N_2 + \ldots + N_{100}$ by the normal distribution

$$N(100\,E(N_i),\ 100\,Var(N_i))$$
$$= N(E(X),\ Var(X))$$

Jessamy is interested in the quality of the soil in her local area and decides to test 100 soil samples for levels of nitrogen. From her past research Jessamy knows that the level of nitrogen in an individual sample, N_i, has a mean of 41 ppm and a standard deviation of 7 ppm.

Let $X = \sum_{i=1}^{100} N_i$ be the total of the levels of nitrogen in Jessamy's batch of samples.

(a) Find

(i) E(X) $\boxed{4100}$

(ii) Var(X). $\boxed{4900}$

[3]

(b) Explain why a normal distribution can be used to give an approximate model for X.

[2]

(c) Use the model to find the an estimate for the value of a such that $\mathrm{P}(X < a) = 0.1$.

[2]

c) Let the distribution of X be approximated by $X \sim N(4100, 4900)$.

$\hookrightarrow \sigma = \sqrt{4900} = 70$

Then

$$P(X < 4010.29139) = 0.1$$

Use inverse normal function on GDC to find this.

$$\boxed{a = 4010 \text{ (3 s.f.)}}$$