Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 1

Question:

The random variable X is binomially distributed. A sample of 10 is taken, and it is desired to test $H_0 : p = 0.25$ against $H_1 : p > 0.25$, using a 5% level of significance.

a Calculate the critical region for this test.

b State the probability of a type I error for this test and, given that the true value of p was later found to be 0.30, calculate the probability of a type II error.

Solution:

$H_0 : p = 0.25 \qquad H_1 : p > 0.25$

a Seek c such that $P(X \geq c) < 0.05$ where $X \sim B(10, 0.25)$

Tables give:

$P(X \leq 5) = 0.9803$

$\therefore P(X \geq 6) = 0.0197$

$\therefore$ critical region is $X \geq 6$

b $P(\text{Type I error}) = P(X \geq 6) = 0.0197$

$P(\text{Type II error} \mid p = 0.3) = P(X \leq 5 \mid p = 0.3)$

$= 0.9527$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 2

Question:

The random variable X is binomially distributed. A sample of 20 is taken, and it is desired to test $H_0: p = 0.30$ against $H_1: p < 0.30$, using a 1% level of significance.

a Calculate the critical region for this test.

b State the probability of a type I error for this test and, given that the true probability was later found to be 0.25, calculate the probability of a type II error.

Solution:

a $H_0: p = 0.30 \quad H_1: p < 0.30$

Seek c such that $P(X \leq c) < 0.01$ where $X \sim B(20, 0.30)$

From tables

$P(X \leq 1) = 0.0076$

and $P(X \leq 2) = 0.0355$

$\therefore$ critical region is $X \leq 1$

b $P(\text{Type I error}) = 0.0076$

$$\begin{aligned} P(\text{Type II error}) &= P(X \geq 2 \mid p = 0.25) \\ &= 1 - P(X \leq 1 \mid p = 0.25) \\ &= 1 - 0.0243 \\ &= 0.9757 \end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 3

Question:

The random variable X is binomially distributed. A sample of 10 is taken, and it is desired to test $H_0 : p = 0.45$ against $H_1 : p \neq 0.45$, using a 5% level of significance.

a Calculate the critical region for this test.

b State the probability of a type I error for this test and, given that the true probability was later found to be 0.40, calculate the probability of a type II error.

Solution:

a $H_0 : p = 0.45 \quad H_1 : p \neq 0.45$

Seek c_1 and c_2 such that $P(X \le c_1) < 0.025$ and $P(X \ge c_2) < 0.025$ where

$X \sim B(10, 0.45)$

From tables

$P(X \le 1) = 0.0233$

$P(X \le 7) = 0.9726 \Rightarrow P(X \ge 8) = 0.0274$

$P(X \le 8) = 0.9955 \Rightarrow P(X \ge 9) = 0.0045$

$\therefore$ critical region is $\{X \le 1\} \cup \{X \ge 9\}$

b
$$\begin{aligned} P(\text{Type I error}) &= P(X \le 1) + P(X \ge 9) \\ &= 0.0233 + 0.0045 \\ &= 0.0278 \end{aligned}$$

$$\begin{aligned} P(\text{Type II error}) &= P(2 \le X \le 8 \mid X \sim B(10, 0.40)) \\ &= P(X \le 8) - P(X \le 1) \\ &= 0.9983 - 0.0464 \\ &= 0.9519 \end{aligned}$$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 4

Question:

The random variable X has a Poisson distribution. A sample is taken, and it is desired to test $H_0 : \lambda = 6$ against $H_1 : \lambda > 6$, using a 5% level of significance.

a Find the critical region for this test.

b Calculate the probability of a type I error and, given that the true value of λ was later found to be 7, calculate the probability of a type II error.

Solution:

$H_0 : \lambda = 6 \qquad H_1 : \lambda > 6$

a Seek c such that $P(X \geq c) < 0.05$ where $X \sim Po(6)$

From tables:

$P(X \leq 10) = 0.9574$

$\therefore P(X \geq 11) = 0.0426$

$\therefore$ critical region is $X \geq 11$

b $P(\text{Type I error}) = P(X \geq 11 \mid X \sim Po(6))$

$= 0.0426$

$P(\text{Type II error}) = P(X \leq 10 \mid \lambda = 7)$

$= 0.9015$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 5

Question:

The random variable X has a Poisson distribution. A sample is taken, and it is desired to test $H_0 : \lambda = 4.5$ against $H_1 : \lambda < 4.5$, using a 5% level of significance.

a Find the critical region for this test.

b Calculate the probability of a type I error and, given that the true value of λ was later found to be 3.5, calculate the probability of a type II error.

Solution:

a $H_0 : \lambda = 4.5 \quad H_1 : \lambda < 4.5$

Seek c such that $P(X \le c) < 0.05$ where $X \sim \text{Po}(4.5)$

Tables give:

$P(X \le 1) = 0.0611$

$P(X = 0) = 0.0111$

$\therefore$ critical region is $X = 0$

b $P(\text{Type I error}) = 0.0111$

$$\begin{aligned} P(\text{Type II error}) &= P(X \ge 1 \mid \lambda = 3.5) \\ &= 1 - P(X = 0 \mid \lambda = 3.5) \\ &= 1 - 0.0302 \\ &= 0.9698 \end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 6

Question:

The random variable X has a Poisson distribution. A sample is taken, and it is desired to test $H_0: \lambda = 9$ against $H_1: \lambda \neq 9$, using a 5% level of significance.

a Find the critical region for this test.

b Calculate the probability of a type I error and, given that the true value of λ was later found to be 8, calculate the probability of a type II error.

Solution:

$H_0: \lambda = 9 \quad H_1: \lambda \neq 9$.

a Seek c_1 and c_2 such that $P(X \leq c_1) < 0.025$ and $P(X \geq c_2) < 0.025$

where $X \sim Po(9)$

From tables:

$P(X \leq 3) = 0.0212$

$P(X \leq 4) = 0.0550$

$P(X \leq 15) = 0.9780 \Rightarrow P(X \geq 16) = 0.0220$

$\therefore$ critical region is $\{X \leq 3\} \cup \{X \geq 16\}$

b $P(\text{Type I error}) = 0.0212 + 0.0220$

$= 0.0432$

$P(\text{Type II error}) = P(4 \leq X \leq 15 \mid \lambda = 8)$

$= P(X \leq 15) - P(X \leq 3)$

$= 0.9918 - 0.0424$

$= 0.9494$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 1

Question:

The random variable $X \sim N(\mu, 3^2)$. A random sample of 20 observations of X is taken, and the sample mean $\bar{x}$ is taken to be the test statistic. It is desired to test $H_0 : \mu = 50$ against $H_1 : \mu > 50$, using a 1% level of significance.

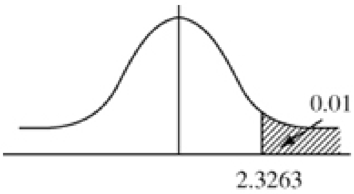
a Find the critical region for this test.
b State the probability of a type I error for this test.
Given that the true mean was later found to be 53,
c find the probability of a type II error.

Solution:

$H_0 : \mu = 50 \qquad H_1 : \mu > 50$

a Critical region when $Z = \dfrac{\bar{x} - 50}{\frac{3}{\sqrt{20}}} > 2.3263$

i.e. $\bar{X} > 51.5605\ldots$



b $P(\text{Type I error}) = \text{significance level} = 0.01$

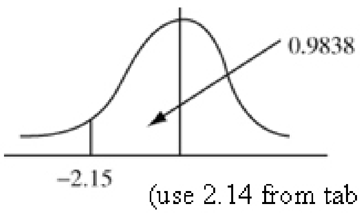
c $P(\text{Type II error}) = P(\bar{X} \le 51.5605\ldots \mid \mu = 53)$

$$= P\left(Z < \frac{51.5605\ldots - 53}{\frac{3}{\sqrt{20}}}\right)$$

$$= P(Z < -2.1458\ldots)$$

$$= 1 - 0.9838$$

$$= 0.0162$$



(use 2.14 from tables)

(Calculator gives $0.01594\ldots$ so accept awrt 0.016)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 2

Question:

The random variable $X \sim N(\mu, 2^2)$. A random sample of 16 observations of X is taken, and the sample mean $\bar{x}$ is taken to be the test statistic. It is desired to test $H_0 : \mu = 30$ against $H_1 : \mu < 30$, using a 5% level of significance.

a Find the critical region for this test.
b State the probability of a type I error for this test.
Given that the true mean was later found to be 28.5,
c find the probability of a type II error.

Solution:

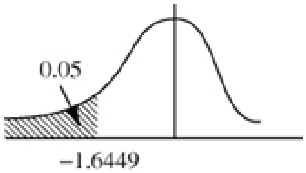
$H_0 : \mu = 30 \qquad H_1 : \mu < 30$

a critical region when

$$Z = \frac{\bar{x} - 30}{\frac{2}{\sqrt{16}}} < -1.6449$$

i.e. $\bar{X} < 30 - 1.6449 \times \frac{1}{2} = 29.17755\ldots$

$\bar{X} < 29.178$



b $P(\text{Type I error}) = 0.05$

c $P(\text{Type II error}) = P(\bar{X} \ge 29.1775\ldots \mid \mu = 28.5)$

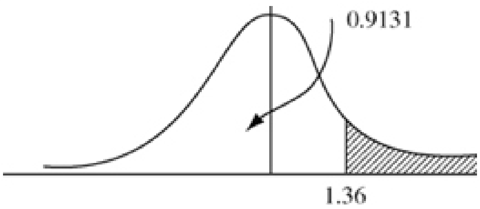
$$= P\left(Z > \frac{29.17755\ldots - 28.5}{\frac{2}{\sqrt{16}}}\right)$$

$= P(Z > 1.3551)$

$= 1 - 0.9131$

$= 0.0869$

(Calculator gives $0.08769\ldots$ so accept answer in range awrt $0.087 \sim 0.088$).



Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 3

Question:

The random variable $X \sim N(\mu, 4^2)$. A random sample of 25 observations of X is taken, and the sample mean $\bar{x}$ is taken to be the test statistic. It is desired to test $H_0: \mu = 40$ against $H_1: \mu \neq 40$, using a 1% level of significance.

a Find the critical region for this test.

b State the probability of a type I error.

Given that the true mean was later found to be 42,

c find the probability of a type II error.

Solution:

$H_0: \mu = 40 \qquad H_1: \mu \neq 40$

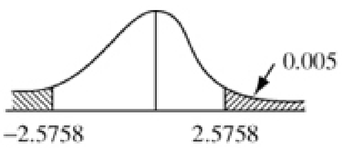
a Critical region $Z < -2.5758$ or $Z > 2.5758$

where $Z = \dfrac{\bar{x} - 40}{\frac{4}{\sqrt{25}}}$

$\therefore \bar{X} > 40 + 0.8 \times 2.5758 = 42.0606\ldots$

or $\bar{X} < 40 - 0.8 \times 2.5758 = 37.9393\ldots$

i.e. $\{\bar{X} < 37.939\} \cup \{\bar{X} > 42.061\}$



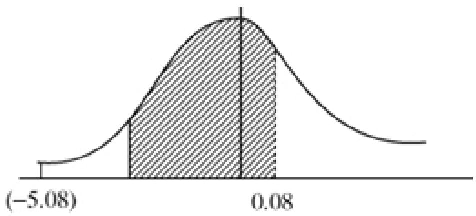
b $P(\text{Type I error}) = 0.01$

c $P(\text{Type II error}) = P(37.939 < \bar{X} < 42.061 \mid \mu = 42)$

$= P(-5.076\ldots < Z < 0.07625)$

$= 0.5319$

(Calculator gives $0.530389\ldots$ so accept awrt 0.53)



Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 4

Question:

A manufacturer claims that the average outside diameter of a particular washer produced by his factory is 15 mm. The diameter is assumed to be normally distributed with a standard deviation of 1. The manufacturer decides to take a random sample of 25 washers from each day's production in order monitor any changes in the mean diameter.

a Using a significance level of 5% find the critical region to be used for this test. Given that the average diameter had in fact increased to 15.6 mm

b find the probability that the day's production would be wrongly accepted.

Solution:

a $D \sim \mathrm{N}(\mu, 1^2)$

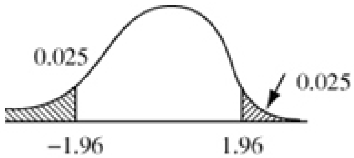
$\mathrm{H}_0 : \mu = 15$ (no change) $\mathrm{H}_1 : \mu \neq 15$ (change in D's mean)

$n = 25$

Critical region $|Z| > 1.96$

i.e. $\dfrac{\bar{X} - 15}{\frac{1}{5}} < -1.96$ or $\dfrac{\bar{X} - 15}{\frac{1}{5}} > 1.96$

i.e. $\bar{X} < 14.608$ or $\bar{X} > 15.392$



b $\mathrm{P}(\text{Accepting production} \mid \mu = 15.6)$

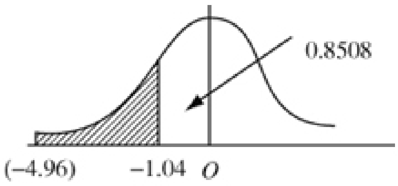
$= \mathrm{P}(14.608 < \bar{X} < 15.392 \mid \mu = 15.6)$

$= \mathrm{P}(-4.96 < Z < -1.04)$

$= 1 - 0.8508$

$= 0.1492$

(Calculator gives 0.1491699… so accept awrt 0.1492)



Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 5

Question:

The number of petrie dishes that a laboratory technician can deal with in one hour can be modelled by a normal distribution with mean 40 and standard deviation 8. A producer of glass pipettes claims that a new type of pipette will speed up the rate at which the technician works.
A random sample of 30 technicians tried out the new pipettes and the average number of petrie dishes they dealt with per hour $\bar{X}$ was recorded.

a Using a 5% significance level find the critical value of $\bar{X}$.

The average number of petrie dishes dealt with per hour using the new pipettes was in fact 42.

b Find the probability of making a type II error.

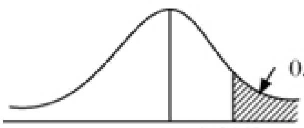
Solution:

a X = number of dishes per hour $\sim N(\mu, 8^2)$

$H_0: \mu = 40 \quad H_1: \mu > 40$

critical region $Z = \dfrac{\bar{X} - 40}{\frac{8}{\sqrt{30}}} > 1.6449$

i.e. $\bar{X} > 42.4025\ldots$

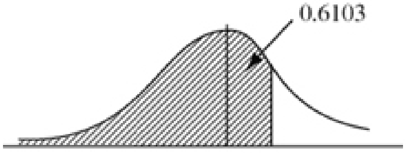


b $P(\text{Type II error}) = P(\bar{X} < 42.4025\ldots \mid \mu = 42)$

$= P(Z < 0.2755\ldots)$

$= 0.6103$

(Calculator gives 0.60856…)
So accept awrt 0.61



c Increasing P(Type II error) will decrease P(Type I error)
Decreasing P(Type II error) will increase P(Type I error)
So only way of reducing P(Type II error) and changing significance level is to increase sample size.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 1

Question:

The random variable $X \sim N(\mu, 3^2)$. A random sample of 25 observations of X is taken and the sample mean $\bar{x}$ is taken as the test statistic. It is desired to test $H_0 : \mu = 20$ against $H_1 : \mu > 20$ using a 5% level of significance.

a Find the critical region for this test.

b Given that $\mu = 20.8$ find the power of this test.

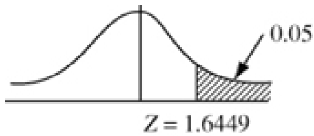
Solution:

a $H_0 : \mu = 20 \qquad H_1 : \mu > 20$

critical region $Z = \dfrac{\bar{x} - 20}{\frac{3}{\sqrt{25}}} > 1.6449$

$\therefore \bar{X} > 20 + 0.6 \times 1.6449$

$\bar{X} > 20.9869\ldots$



b Power $= 1 - P(\text{Type II error})$

$= 1 - P(\bar{X} \le 20.9869\ldots \mid \mu = 20.8)$

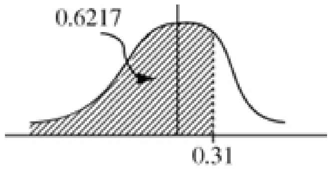
$= 1 - P\left(Z \le \dfrac{20.9869\ldots - 20.8}{0.6}\right)$

$= 1 - P(Z \le 0.3115\ldots)$

$= 1 - 0.6217$

$= 0.3783$

(Calculator gives $0.37768\ldots$ so accept awrt 0.378)



Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 2

Question:

The random variable X is a binomial distribution. A sample of 20 is taken from it. It is desired to test $\mathrm{H}_0 : p = 0.35$ against $\mathrm{H}_1 : p > 0.35$ using a 5% level of significance.

a Calculate the size of this test.

b Given that $p = 0.36$ calculate the power of this test.

Solution:

a $\mathrm{H}_0 : p = 0.35 \qquad \mathrm{H}_1 : p > 0.35$

Seek c such that $\mathrm{P}(X \geq c) < 0.05$ where $X \sim B(20, 0.35)$

Tables give

$\mathrm{P}(X \leq 10) = 0.9468$

$\mathrm{P}(X \leq 11) = 0.9804$

$\therefore \mathrm{P}(X \geq 12) = 1 - 0.9804 = 0.0196$

$\therefore$ size of test is 0.0196

b $\text{Power} = 1 - \mathrm{P}(\text{Type II error})$

critical region is $X \geq 12$

$\therefore \text{Power} = \mathrm{P}(X \geq 12 \mid p = 0.36)$

$\therefore \text{Power} = 1 - \mathrm{P}(X \leq 11 \mid p = 0.36)$

$= 1 - 0.9753$

$= 0.0247$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 3

Question:

The random variable X has a Poisson distribution. A sample is taken and it is desired to test $H_0 : \lambda = 4.5$ against $H_1 : \lambda < 4.5$. If a 5% significance level is to be used,

a find the size of this test.

b Given that $\lambda = 4.1$ find the power of the test.

Solution:

a $H_0 : \lambda = 4.5 \quad H_1 : \lambda < 4.5$

critical region seek c such that

$P(X \le c) < 0.05$ where $X \sim \text{Po}(4.5)$

Tables give:

$P(X \le 1) = 0.0611$

$P(X = 0) = 0.0111$

$\therefore$ critical region is $X = 0$

Size is 0.0111

b Power $= P(X = 0 \mid \lambda = 4.1)$

$= e^{-4.1}$

$= 0.016572\ldots$

$= 0.0166 \quad (3\text{ s.f.})$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 4

Question:

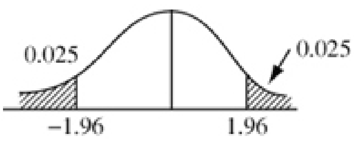
A manufacturer claims that a particular rivet produced in his factory has a diameter of 2 mm, and that the diameter is normally distributed with a variance of 0.004 mm^2. A random sample of 25 rivets is taken from a day's production to test whether the mean diameter had altered, up or down, from the stated figure. A 5% significance level is to be used for this test.
If the mean diameter had in fact altered to 2.02 mm, calculate the power of this test.

Solution:

$D = \text{diameter} \sim \text{N}(\mu, 0.004)$

$\text{H}_0 : \mu = 2 \qquad \text{H}_1 : \mu \neq 2$

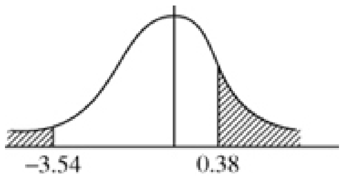
Critical region is $|Z| > 1.96$



$\therefore \dfrac{\overline{X} - 2}{\sqrt{\dfrac{0.004}{25}}} > 1.96$ or $\dfrac{\overline{X} - 2}{\sqrt{\dfrac{0.004}{25}}} < -1.96$

i.e. $\overline{X} < 1.9752\ldots$ or $\overline{X} > 2.0247\ldots$

$$\begin{aligned}\text{Power} &= \text{P}(\overline{X} < 1.9752\ldots \mid \mu = 2.02) + \text{P}(\overline{X} > 2.0247\ldots \mid \mu = 2.02)\\ &= \text{P}(Z < -3.54\ldots) + \text{P}(Z > 0.3788\ldots)\\ &= 0.0002 + (1 - 0.6480)\\ &= 0.0002 + 0.352\\ &= 0.3522\end{aligned}$$



Calculator gives $0.352596\ldots$
so accept awrt $0.352 \sim 0.353$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 5

Question:

In a binomial experiment consisting of 10 trials the random variable X represents the number of successes, and p is the probability of a success.
In a test of $\mathrm{H}_0: p = 0.3$ against $\mathrm{H}_1: p > 0.3$, a critical region of $X \geq 7$ is used.
Find the power of this test when
a $p = 0.4$,
b $p = 0.8$.
c Comment on your results. **[E]**

Solution:

$\mathrm{H}_0: p = 0.3 \qquad \mathrm{H}_1: p > 0.3$
Critical region is $X \geq 7 \qquad n = 10$

a $\text{Power} = \mathrm{P}(X \geq 7 \mid p = 0.4)$
$= 1 - \mathrm{P}(X \leq 6)$
$= 1 - 0.9452$
$= 0.0548$

b $\text{Power} = \mathrm{P}(X \geq 7 \mid p = 0.8)$
Let $Y \sim \mathrm{B}(10, 0.2)$
$= \mathrm{P}(Y \leq 3)$
$= 0.8791$

c The test is more powerful for values of p further away from $p = 0.3$.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 6

Question:

Explain briefly what you understand by

a a type I error,

b the size of a significance test.

A single observation is made on a random variable X, where $X \sim N(\mu, 10)$.

The observation, x, is to be used to test $H_0: \mu = 20$ against $H_1: \mu > 20$. The critical region is chosen to be $X \geq 25$.

c Find the size of the test.

Solution:

a Type I error is when H_0 is rejected when H_0 is in fact true.

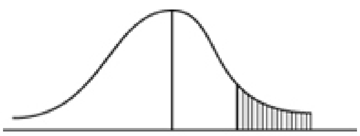
b Size = P(Type I error)

c $H_0: \mu = 20 \qquad H_1: \mu > 20$

Critical region is $X \geq 25 \qquad X \sim N(\mu, 10)$

$$\begin{aligned}\text{Size} &= P(\text{Reject } H_0 \text{ when } H_0 \text{ is true}) \\ &= P(X \geq 25 \mid \mu = 20) \\ &= P\left(Z > \frac{25-20}{\frac{\sqrt{10}}{\sqrt{1}}}\right) \\ &= P(Z > 1.58\ldots) \\ &= 1 - 0.9429 \\ &= 0.0571\end{aligned}$$

$n = 1$ (single observation)



Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 1

Question:

A single observation x is taken from a Poisson distribution with parameter λ. This observation is to be used to test $H_0 : \lambda = 6.5$ against $H_1 : \lambda < 6.5$. The critical region chosen was $X \leq 2$.

a Find the size of the test.

b Show that the power function of this test is given by $e^{-\lambda}\left(1+\lambda+\frac{1}{2}\lambda^2\right)$.

The table below gives the value of the power function to two decimal places.

λ	1	2	3	4	5	6
Power	0.92	s	0.42	0.24	t	0.06

c Calculate values for s and t.

d Draw a graph of the power function.

e Find the values of λ for which the test is more likely than not to come to the correct conclusion.

Solution:

$H_0: \lambda = 6.5 \qquad H_1: \lambda < 6.5$

Critical region $X \leq 2$

a Size $= P(X \leq 2 \mid \lambda = 6.5) = 0.0430$

b Power $= P(X \leq 2 \mid \lambda)$

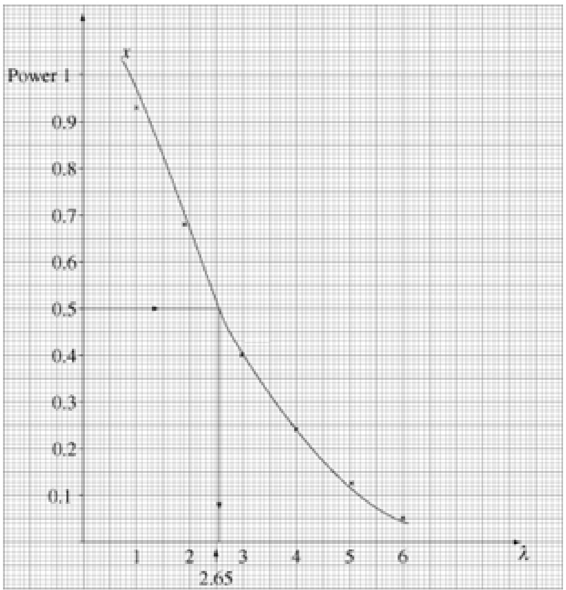
$$= e^{-\lambda} + \frac{e^{-\lambda} \cdot \lambda^1}{1!} + \frac{e^{-\lambda}\lambda^2}{2!}$$

$$= e^{-\lambda}\left(1 + \lambda + \frac{1}{2}\lambda^2\right)$$

c $\lambda = 2 \Rightarrow s = 0.6767$ (tables) $= 0.68$ (2 d.p.)

$\lambda = 5 \Rightarrow t = 0.1247$ (tables) $= 0.12$ (2 d.p.)

d See graph.



e Correct conclusion is arrived at when: $\lambda = 6.5$,
H_0 is accepted. So since size is 0.0430
probability of accepting $\lambda = 6.5$ is 0.957
$\therefore \lambda = 6.5$
or for $\lambda < 6.5$, correct conclusion is to reject H_0.
So require where power > 0.5 i.e. $\lambda < 2.65$
(from graph)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 2

Question:

In a binomial experiment consisting of 12 trials X represents the number of successes and p the probability of a success.
In a test of $\mathrm{H}_0 : p = 0.45$ against $\mathrm{H}_1 : p < 0.45$ the null hypothesis is rejected if the number of successes is 2 or less.

a Find the size of this test.

b Show that the power function for this test is given by
$(1-p)^{12} + 12p(1-p)^{11} + 66p^2(1-p)^{10}$.

c Find the power of this test when p is 0.3.

Solution:

$\mathrm{H}_0 : p = 0.45 \quad \mathrm{H}_1 : p < 0.45$

Critical region $X \leq 2$, where $X \sim \mathrm{B}(12, 0.45)$

a $\text{Size} = \mathrm{P}(X \leq 2)$
$= 0.0421$

b $\text{Power} = \mathrm{P}(X \leq 2 \mid X \sim \mathrm{B}(12, p))$

$$= (1-p)^{12} + 12p(1-p)^{11} + \binom{12}{2} p^2 (1-p)^{10}$$

$$= (1-p)^{12} + 12p(1-p)^{11} + 66p^2(1-p)^{10}$$

c $p = 0.3$

$\text{Power} = 0.2528$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 3

Question:

In a binomial experiment consisting of 10 trials the random variable X represents the number of successes and p the probability of a success.

In a test of $H_0: p = 0.4$ against $H_1: p > 0.4$, a critical region of $X \geq 8$ was used.

Find the power of this test when

a $p = 0.5$

b $p = 0.8$

c Comment on your results.

Solution:

$H_0: p = 0.4 \quad H_1: p > 0.4$

Critical region $X \geq 8$

a $\text{Power} = P(X \geq 8 \mid X \sim B(10, 0.5))$

$= 1 - P(X \leq 7)$

$= 1 - 0.9453 = 0.0547$

b $\text{Power} = P(X \geq 8 \mid X \sim B(10, 0.8))$

Let $Y \sim B(10, 0.2)$ then

$\text{Power} = P(Y \leq 2 \mid Y \sim B(10, 0.2))$

$= 0.6778$

c The test is more powerful for values of p further away from 0.4.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 4

Question:

A certain gambler always calls heads when a coin is tossed. Before he uses a coin he tests it to see whether or not it is fair and uses the following hypotheses:

$H_0 : p = \frac{1}{2} \quad H_1 : p < \frac{1}{2}$

where p is the probability that the coin lands heads on a particular toss. Two tests are proposed.

In test A the coin is tossed 10 times and H_0 is rejected if the number of heads is 2 or fewer.

a Find the size of test A.

b Explain why the power of test A is given by
$(1-p)^{10} + 10p(1-p)^9 + 45p^2(1-p)^8$.

In test B the coin is first tossed 5 times. If no heads result H_0 is immediately rejected. Otherwise the coin is tossed a further 5 times and H_0 is rejected if no heads appear on this second occasion.

c Find the size of test B.

d Find an expression for the power of test B in terms of p.

The power for test A and the power for test B are given in the table for various values of p.

p	0.1	0.2	0.25	0.3	0.35	0.4
Power for test A	0.9298	0.6778		0.3828		0.1673
Power for test B	0.8323	0.5480	0.4183	0.3079	0.2186	0.1495

e Find the power for test A when p is 0.25 and 0.35.

f Giving a reason, advise the gambler about which test he should use. **[E]**

Solution:

$H_0: p = \frac{1}{2} \quad H_1: p < \frac{1}{2} \qquad (n = 10)$

Test A Critical region $X \le 2$ where $X \sim B(10, p)$

a Size $= P(X \le 2 \mid X \sim B(10, 0.5))$

$= 0.0547$

b Power $= P(X \le 2 \mid X \sim B(10, p))$

$$= (1-p)^{10} + 10p(1-p)^9 + \binom{10}{2} p^2 (1-p)^8$$

$$= (1-p)^{10} + 10p(1-p)^9 + 45p^2(1-p)^8$$

Test B Let $Y \sim B(5, p)$.

c Size $= P(Y = 0) + [1 - P(Y = 0)]\, P(Y = 0)$ where $p = 0.5$

$= 0.0312 + [1 - 0.0312] \times 0.0312$

$= 0.06142$

NB calculator gives 0.06152

d Power $= (1-p)^5 + \left[1 - (1-p)^5\right](1-p)^5$

$$= (1-p)^5 \left[2 - (1-p)^5\right]$$

e $p = 0.25 \Rightarrow \text{power}_A = 0.5256$
$p = 0.35 \Rightarrow \text{power}_A = 0.2616$ } from calculator

f Use test A as this is always more powerful

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 1

Question:

A random sample of size 3 is taken without replacement, from a population with mean μ and variance σ^2. Two unbiased estimators of the mean of the population are

$\hat{\mu}_1 = \frac{1}{3}(X_1 + X_2 + X_3)$ and $\hat{\mu}_2 = \frac{1}{4}(X_1 + 2X_2 + X_3)$.

a Calculate $\text{Var}(\hat{\mu}_1)$ and $\text{Var}(\hat{\mu}_2)$.

b Hence state, giving a reason, which estimator you would recommend. **[E]**

Solution:

a
$$\text{Var}(\hat{\mu}_1) = \frac{1}{9}\left[\text{Var}(X_1) + \text{Var}(X_2) + \text{Var}(X_3)\right]$$
$$= \frac{\sigma^2 + \sigma^2 + \sigma^2}{9} \quad = \frac{\sigma^2}{3}$$
$$\text{Var}(\hat{\mu}_2) = \frac{1}{16}\left[\text{Var}(X_1) + 2^2\text{Var}(X_2) + \text{Var}(X_3)\right]$$
$$= \frac{\sigma^2 + 4\sigma^2 + \sigma^2}{16} = \frac{3\sigma^2}{8}$$

b Recommend $\hat{\mu}_1 \because \text{Var}(\hat{\mu}_1) < \text{Var}(\hat{\mu}_2)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 2

Question:

If X_1, X_2, X_3, is a random sample from a population with mean μ and variance σ^2, find which of the following estimators of μ are unbiased. If any are biased find an expression for the bias.

a $\frac{1}{8}X_1+\frac{3}{8}X_2+\frac{1}{2}X_3$

b $\frac{1}{4}X_1+\frac{1}{2}X_2$

c $\frac{1}{3}X_1+\frac{2}{3}X_2$

d $\frac{1}{3}(X_1+X_2+X_3)$

e $\frac{1}{5}X_1+\frac{2}{5}X_2+\frac{3}{5}X_3$

Solution:

a $\mathrm{E}\left(\frac{1}{8}X_1+\frac{3}{8}X_2+\frac{1}{2}X_3\right)=\frac{1}{8}\mu+\frac{3}{8}\mu+\frac{1}{2}\mu$

$=\frac{8}{8}\mu=\mu \therefore$ unbiased

b $\mathrm{E}\left(\frac{1}{4}X_1+\frac{1}{2}X_2\right)=\frac{1}{4}\mu+\frac{1}{2}\mu=\frac{3}{4}\mu$

$\therefore$ bias $=\frac{3}{4}\mu-\mu=-\frac{1}{4}\mu$

c $\mathrm{E}\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=\frac{1}{3}\mu+\frac{2}{3}\mu=\mu$

$\therefore$ unbiased

d $\mathrm{E}\left\{\frac{1}{3}(X_1+X_2+X_3)\right\}=\frac{1}{3}(\mu+\mu+\mu)=\mu$

$\therefore$ unbiased

e $\mathrm{E}\left(\frac{1}{5}X_1+\frac{2}{5}X_2+\frac{3}{5}X_3\right)=\frac{1}{5}\mu+\frac{2}{5}\mu+\frac{3}{5}\mu$

$=\frac{6}{5}\mu$

$\therefore$ bias $=\frac{1}{5}\mu$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 3

Question:

Find which one of the estimators in question 2 is the best.

Solution:

a $\text{Var}\left(\frac{1}{8}X_1+\frac{3}{8}X_2+\frac{1}{2}X_3\right)$

$=\frac{1}{64}\text{Var}(X_1)+\frac{9}{64}\text{Var}(X_2)+\frac{1}{4}\text{Var}(X_3)$

$=\frac{26}{64}\sigma^2$ or $\frac{13}{32}\sigma^2\ (=0.40625\sigma^2)$

b estimator is biased so would not prefer

c $\text{Var}\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=\frac{1}{9}\sigma^2+\frac{4}{9}\sigma^2=\frac{5}{9}\sigma^2$ or $0.555\sigma^2$

d $\text{Var}\left(\frac{1}{3}[X_1+X_2+X_3]\right)=\frac{1}{9}(\sigma^2+\sigma^2+\sigma^2)=\frac{3}{9}\sigma^2$ or $\frac{1}{3}\sigma^2$ or $0.333\sigma^2$

e estimator is biased
Best estimator is unbiased with smallest variance.
Since $\frac{1}{3}\sigma^2<\frac{13}{32}\sigma^2<\frac{5}{9}\sigma^2$

$\therefore$ Choose $\frac{1}{3}(X_1+X_2+X_3)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 4

Question:

A uniform distribution on the interval $[0, a]$ has a mean of $\frac{a}{2}$, and a variance of $\frac{a^2}{12}$.
Three single samples X_1, X_2 and X_3 are taken from this distribution, and are to be used to estimate a. The following estimators are proposed.

i $X_1 + X_2 + X_3$

ii $\frac{2}{3}(X_1 + X_2 + X_3)$

iii $2(X_1 + 2X_2 + X_3)$

a Determine the bias, if any of each of these estimators.
b Find the variance of each of these estimators.
c State, giving reasons, which of these estimators you would use.
d If $x_1 = 2, x_2 = 2.5$ and $x_3 = 3.2$, calculate the best estimate of a.

Solution:

a i $\mathrm{E}(X_1 + X_2 + X_3) = \frac{a}{2} + \frac{a}{2} + \frac{a}{2} = \frac{3a}{2} \therefore$ bias $= \frac{a}{2}$

ii $\mathrm{E}\left(\frac{2}{3}[X_1 + X_2 + X_3]\right) = \frac{2}{3}\left[\frac{3a}{2}\right] = a \therefore$ unbiased

iii $\mathrm{E}(2[X_1 + 2X_2 + X_3]) = 2\left[\frac{a}{2} + 2\frac{a}{2} + \frac{a}{2}\right] = 4a \therefore$ bias $= 3a$

b i $\mathrm{Var}(X_1 + X_2 + X_3) = \frac{a^2}{12} + \frac{a^2}{12} + \frac{a^2}{12} = \frac{a^2}{4}$

ii $\mathrm{Var}\left(\frac{2}{3}[X_1 + X_2 + X_3]\right) = \frac{4}{9}\left[\frac{a^2}{12} + \frac{a^2}{12} + \frac{a^2}{12}\right] = \frac{a^2}{9}$

iii $\mathrm{Var}[2(X_1 + 2X_2 + X_3)] = 4\left[\frac{a^2}{12} + 4\frac{a^2}{12} + \frac{a^2}{12}\right] = 2a^2$

c Use $\frac{2}{3}(X_1 + X_2 + X_3)$ since it is unbiased (and has the smallest variance)

d $x_1 = 2, x_2 = 2.5, x_3 = 3.2$

$$\Rightarrow \frac{2}{3}(x_1 + x_2 + x_3) = \frac{2}{3}(2 + 2.5 + 3.2) = \frac{2}{3}(7.7)$$

$$\hat{a} = 5.13$$

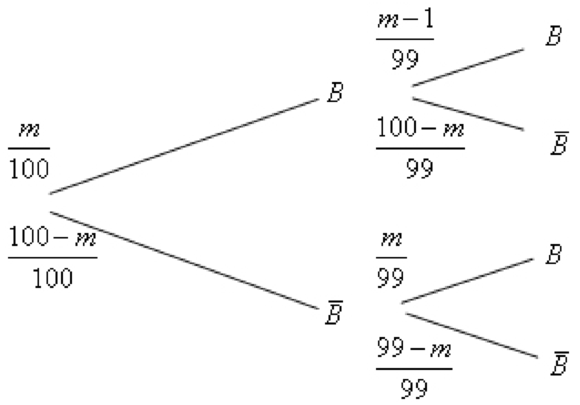
Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 5

Question:

A bag contains 100 counters of which an unknown number m are blue. It is known that $2 \leq m \leq 98$. Two discs are drawn simultaneously from the bag and the number n of blue ones counted. It is desired to estimate m by $\hat{m} = cn$ where c is an unknown constant. Find the value of c given that the estimate is unbiased.

Solution:



$X =$ number of blue ones chosen

n	0	1	2
$P(X = n)$	$\frac{(100-m)(99-m)}{100\times 99}$	$\frac{200m - 2m^2}{100\times 99}$	$\frac{m(m-1)}{100\times 99}$

$$\therefore E(X) = \frac{200m - 2m^2 + 2m^2 - 2m}{100 \times 99}$$

$$= \frac{198m}{100\times 99}$$

$$= \frac{m}{50}$$

Using $\hat{m} = cX$

$$E(\hat{m}) = cE(X) = c \times \frac{m}{50}$$

$\therefore$ for $\hat{m}$ to be unbiased you need $c = 50$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 6

Question:

A sample of size n is taken from a population with a mean of μ and variance of σ^2.

a Show that the sample mean $\bar{X}$ is an unbiased estimator of μ.

b Show that as n increases $\text{Var}(\bar{X})$ decreases.

c Show that $S^2 = \dfrac{\sum X_i^2 - n\bar{X}^2}{n-1}$ is an unbiased estimator of σ^2, but that $T = \dfrac{\sum X_i^2 - n\bar{X}^2}{n}$ is a biased estimator of σ^2.

Solution:

a $\text{E}(\bar{X}) = \frac{1}{n}\text{E}\left(X_1 + X_2 + \cdots + X_n\right) = \frac{1}{n}(\mu + \mu + \cdots + \mu) = \frac{n\mu}{n} = \mu$

$\therefore \bar{X}$ is unbiased estimator of μ.

b $\text{Var}\left(\bar{X}\right) = \frac{1}{n^2}\text{Var}\left(X_1 + \cdots + X_n\right) = \frac{1}{n^2}\left(\sigma^2 + \cdots + \sigma^2\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$

$\therefore$ as $n \to \infty \ \text{Var}\left(\bar{X}\right) \to 0$.

c $\text{E}\left(S^2\right) = \frac{1}{n-1}\left\{\text{E}\left(X_1^2 + X_2^2 + \cdots + X_n^2\right) - n\text{E}\left(\bar{X}^2\right)\right\}$

NB $\sigma^2 = \text{E}\left(X^2\right) - \mu^2 \quad \therefore \text{E}\left(X^2\right) = \mu^2 + \sigma^2$

$\frac{\sigma^2}{n} = \text{E}\left(\bar{X}^2\right) - \mu^2 \quad \therefore \text{E}\left(\bar{X}^2\right) = \mu^2 + \frac{\sigma^2}{n}$

So $\text{E}\left(S^2\right) = \frac{1}{n-1}\left\{n\left[\mu^2 + \sigma^2\right] - n\left[\mu^2 + \frac{\sigma^2}{n}\right]\right\}$

$= \frac{1}{n-1}\left\{n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right\}$

$= \frac{(n-1)}{n-1}\sigma^2 = \sigma^2$

$\text{E}(T) = \frac{1}{n} \times (n-1)\sigma^2 \neq \sigma^2$

$\therefore T$ is not unbiased for σ^2.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 7

Question:

A six-sided die has some of its faces showing the number 0 and the rest showing the number 1 so that p is the probability of getting a 1 when the die is thrown and q is the probability of getting a 0. If the random variable X is the value showing when the die is rolled,

a find E(X) and Var(X).

A random sample is now taken by rolling the die three times in order to get an estimate for p.

b Show that if $a_1X_1 + a_2X_2 + a_3X_3$ is to be an unbiased estimator of p then $a_1 + a_2 + a_3 = 1$.

c Find the variance of this estimator.

The following estimators of p are proposed.

i $\frac{1}{5}X_1 + \frac{2}{5}X_2 + \frac{2}{5}X_3$

ii $\frac{1}{4}X_1 + \frac{3}{8}X_2 + \frac{1}{4}X_3$

iii $\frac{4}{9}X_1 + \frac{5}{9}X_3$

d Find which of these is the best unbiased estimator.

Solution:

x	0	1
$P(X=x)$	q	p

$p+q=1$

a $E(X)=p$

$\text{Var}(X)=0+p-p^2=p(1-p)$ or pq

b $E(a_1X_1+a_2X_2+a_3X_3)=a_1p+a_2p+a_3p$

$\therefore$ if unbiased $a_1+a_2+a_3=1$

c $\text{Var}(a_1X_1+a_2X_2+a_3X_3)=a_1^2pq+a_2^2pq+a_3^2pq$

$=pq(a_1^2+a_2^2+a_3^2)$

d i $E\left(\frac{1}{5}X_1+\frac{2}{5}X_2+\frac{2}{5}X_3\right)=\frac{p+2p+2p}{5}=p$ $\therefore$ unbiased

$\text{Var}\left(\frac{1}{5}X_1+\frac{2}{5}X_2+\frac{2}{5}X_3\right)=\frac{pq}{25}(1+4+4)=\frac{9pq}{25}$

ii $\frac{1}{4}+\frac{3}{8}+\frac{1}{4}=\frac{7}{8}\neq 1$ $\therefore$ not unbiased

iii $\frac{4}{9}+\frac{5}{9}=\frac{9}{9}=1$ $\therefore$ unbiased

$\text{Var}\left(\frac{4}{9}X_1+\frac{5}{9}X_3\right)=\frac{pq}{81}(16+25)=\frac{41pq}{81}$

$\therefore$ Best estimator is $\frac{1}{5}X_1+\frac{2}{5}X_2+\frac{2}{5}X_3$ since it is unbiased and has smaller variance, $\left(\frac{9}{25}<\frac{41}{81}\right)$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 1

Question:

A biased die has probability of a six equal to p. The die is rolled n times and the number of sixes recorded. The die is then rolled a further n times and the number of sixes recorded. The proportion of the $2n$ rolls that were sixes is called R.

a Show that R is a consistent estimator of p.

The die is rolled a total of 50 times and 18 sixes are recorded.

b Find an estimate of p.

Solution:

$X =$ number of sixes in n rolls $\quad X \sim \mathrm{B}(n, p)$

$R = \dfrac{X_1 + X_2}{2n} \qquad \mathrm{E}(X) = np, \mathrm{var}(X) = np(1-p)$

a $\mathrm{E}(R) = \dfrac{1}{2n}\mathrm{E}(X_1 + X_2)$

$= \dfrac{1}{2n}[np + np] = \dfrac{2np}{2n} = p$

$\therefore$ R is an unbiased estimator of p.

$\mathrm{Var}(R) = \dfrac{1}{4n^2}\mathrm{Var}(X_1 + X_2)$

$= \dfrac{1}{4n^2}[np(1-p) + np(1-p)]$

$= \dfrac{2np(1-p)}{4n^2}$

$= \dfrac{p(1-p)}{2n}$

$\therefore$ as $n \to \infty \quad \mathrm{Var}(R) \to 0$

$\therefore$ R is a consistent estimator for p.

b $\hat{p} = \dfrac{18}{50}$ or $\dfrac{9}{25}$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 2

Question:

The continuous random variable $X \sim \text{U}[0, a]$.

a Show that $2\bar{X}$ is an unbiased estimator of a.

b Determine whether or not $2\bar{X}$ is a consistent estimator of a.

Solution:

$X \sim \text{U}[0, a]$

$\text{E}(X) = \mu = \frac{a}{2} \quad \text{Var}(X) = \sigma^2 = \frac{a^2}{12}$

a $\text{E}(2\bar{X}) = 2\text{E}(\bar{X}) = 2\mu = 2 \times \frac{a}{2} = a$

$\therefore 2\bar{X}$ is an unbiased estimator of a.

b $\text{Var}(2\bar{X}) = 4\text{Var}(\bar{X}) = 4\frac{\sigma^2}{n}$

$= \frac{4a^2}{12n} = \frac{a^2}{3n}$

$\therefore$ as $n \to \infty \quad \text{Var}(2\bar{X}) \to 0$

$\therefore 2\bar{X}$ is a consistent estimator of a

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 3

Question:

Using the information and results from Example 16 show that M is a consistent estimator of a.

Solution:

From Example 16 $M = \max\{X_1 \dots X_n\}$

$$\mathrm{E}(M) = \frac{n}{n+1}a = \left(\frac{n+1}{n+1}\right)a - \left(\frac{1}{n+1}\right)a$$

$$= a - \left(\frac{1}{n+1}\right)a$$

as $n \to \infty$ $\mathrm{E}(M) \to a$

$\therefore M$ is an asymptotically unbiased estimator of a.

$$\mathrm{Var}(M) = \frac{na^2}{(n+2)(n+1)^2}$$

as $n \to \infty$ $\mathrm{Var}(M) \to 0$

$\therefore M$ is a consistent estimator of a.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 4

Question:

If a random sample $X_1, X_2, X_3, \ldots, X_n$, is taken from a population with mean μ and standard deviation σ, show that both,

$\frac{1}{n}(X_1 + X_2 + \ldots + X_{n-1} + X_n)$, and

$2\frac{(nX_1 + (n-1)X_2 + \ldots + 2X_{n-1} + 1X_n)}{n(n+1)}$

are unbiased and consistent estimators for μ.

You may use $\sum_{r=1}^{n} r = \frac{1}{2}n(n+1)$ and $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$

Solution:

a $\frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \bar{X}$

$\mathrm{E}(\bar{X}) = \mu$ and $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$

$\therefore \bar{X}$ is an unbiased estimator of μ and $\because \mathrm{Var}(\bar{X}) \to 0$ an $n \to \infty, \bar{X}$ is a consistent estimator of μ

b Let $Y = \frac{2(nX_1 + (n-1)X_2 + \cdots + 1X_n)}{n(n+1)}$

$$\mathrm{E}(Y) = \frac{2}{n(n+1)}[n\mathrm{E}(X_1) + (n-1)\mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n)]$$

$$= \frac{2\mu}{n(n+1)}[n + (n-1) + \cdots + 1]$$

But $\sum_{r=1}^{n} r = \frac{n(n+1)}{2}$

$$\therefore \mathrm{E}(Y) = \frac{2\mu}{n(n+1)} \times \frac{n(n+1)}{2} = \mu$$

$\therefore Y$ is an unbiased estimator of μ

$$\mathrm{Var}(Y) = \frac{4}{n^2(n+1)^2}\left[n^2\sigma^2 + (n-1)^2\sigma^2 + \cdots + 1^2\sigma^2\right]$$

$$= \frac{4\sigma^2}{n^2(n+1)^2} \times \frac{n}{6}(n+1)(2n+1) \qquad \because \sum_{1}^{n} r^2 = \frac{n}{6}(n+1)(2n+1)$$

$$= \frac{4\sigma^2(2n+1)}{6n(n+1)}$$

As $n \to \infty$ $\mathrm{Var}(Y) \to 0$ Y is in consistent estimator of μ.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 5

Question:

The random variable $X \sim \text{U}[0, a]$.

a Show that $\text{E}(X^n) = \frac{a^n}{n+1}$

A random sample of 3 readings is taken from X and the statistic $S = X_1^2 + X_2^2 + X_3^2$ is calculated.

b Show that S is an unbiased estimator of a^2.

c Show that $\text{Var}(X^2) = \frac{4}{45}a^4$

A random sample of size n is taken of X.

d Show that $T = \frac{3}{n}(X_1^2 + X_2^2 + \ldots + X_n^2)$ is a consistent estimator of a^2.

Solution:

$X \sim \text{U}[0, a]$

a $\text{E}(X^n) = \int_0^a x^n \times \frac{1}{a}\,\text{d}x = \left[\frac{x^{n+1}}{a(n+1)}\right]_0^a = \left(\frac{a^{n+1}}{a(n+1)}\right) - (0)$

$= \frac{a^n}{n+1}$

b $S = X_1^2 + X_2^2 + X_3^2$

$\text{E}(X^2) = \frac{a^2}{3}$ (by **a**)

$\therefore \text{E}(S) = \frac{a^2}{3} + \frac{a^2}{3} + \frac{a^2}{3} = a^2$

$\therefore S$ is an unbiased estimator of a^2

c $\text{Var}(X^2) = \text{E}(X^4) - \left[\text{E}(X^2)\right]^2 = \frac{a^4}{5} - \left[\frac{a^2}{3}\right]^2$

$= \frac{9a^4 - 5a^4}{45} = \frac{4a^4}{45}$

d $\text{E}(T) = \frac{3}{n}\text{E}\left[X_1^2 + \cdots + X_n^2\right] = \frac{3}{n}\left[\frac{a^2}{3} + \frac{a^2}{3} + \cdots + \frac{a^2}{3}\right]$

$= \frac{3}{n} \times \frac{na^2}{3} = a^2$

$\therefore T$ is an unbiased estimator of a^2

$\text{Var}(T) = \frac{9}{n^2}\left[\text{Var}(X_1^2) + \text{Var}(X_2^2) + \cdots + \text{Var}(X_n^2)\right]$

$= \frac{9}{n^2}\left[\frac{4a^4}{45} \times n\right] = \frac{4a^4}{5n}$

$\therefore$ as $n \to \infty$ $\text{Var}(T) \to 0$

$\therefore T$ is a consistent estimator for a^2.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 6

Question:

When a die is rolled the probability of obtaining a six is an unknown constant p. In order to estimate p the die is rolled n times and the number, X, of sixes is recorded. A second trial is then done with the die being rolled the same number of times, and the number of sixes, Y, is again recorded. Show that

a $\hat{p}_1 = \frac{3\bar{X} + 4\bar{Y}}{7n}$, and $\hat{p}_2 = \frac{\bar{X} + \bar{Y}}{2n}$, are unbiased and consistent estimators of p.

b State, giving reasons, which of the two estimators is the better one.

Solution:

$X \sim \mathrm{B}(n, p) \quad Y \sim \mathrm{B}(n, p)$

$\mathrm{E}(X) = \mu = np \quad \mathrm{Var}(X) = \sigma^2 = np(1-p)$

a $\mathrm{E}(\hat{p}_1) = \frac{3np + 4np}{7n} = \frac{7np}{7n} = p$

$$\mathrm{Var}(\hat{p}_1) = \frac{9\frac{np(1-p)}{n} + 16\frac{np(1-p)}{n}}{49n^2} = \frac{25\frac{np(1-p)}{n}}{49n^2} = \frac{25p(1-p)}{49n^2}$$

$\therefore \hat{p}_1$ is unbiased and $\mathrm{Var}(\hat{p}_1) \to 0$ an $n \to \infty \therefore \hat{p}_1$ is consistent for p

$\mathrm{E}(\hat{p}_2) = \frac{np + np}{2n} = \frac{2np}{2n} = p$

$$\mathrm{Var}(\hat{p}_2) = \frac{\frac{np(1-p)}{n} + \frac{np(1-p)}{n}}{4n^2} = \frac{p(1-p)}{2n^2}$$

$\therefore \hat{p}_2$ is unbiased and $\mathrm{Var}(\hat{p}_2) \to 0$ as $n \to \infty \therefore \hat{p}_2$ is consistent for p.

b $\because \frac{25}{49} > \frac{1}{2}$

$\therefore$ Choose $\frac{\bar{X} + \bar{Y}}{2n}$ since it has smaller variance.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 1

Question:

The random variable X is binomially distributed. A sample of 15 observations is taken and it is desired to test $H_0: p = 0.35$ against $H_1: p > 0.35$ using a 5% significance level.

a Find the critical region for this test.

b State the probability of making a type I error for this test.

The true value of p was found later to be 0.5.

c Calculate the power of this test.

Solution:

$H_0: p = 0.35 \quad H_1: p > 0.35 \quad X \sim B(15, p)$

a Seek c such that $P(X \geq c) < 0.05$

Tables give: $P(X \leq 8) = 0.9578$

$\therefore P(X \geq 9) = 0.0422$

So critical region is $X \geq 9$

b $P(\text{Type I error}) = 0.0422$

c $\text{Power} = P(X \geq 9 \mid p = 0.5)$

$= 1 - P(X \leq 8 \mid p = 0.5)$

$= 1 - 0.6964$

$= 0.3036$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 2

Question:

The random variable X has a Poisson distribution. A sample is taken and it is desired to test $H_0 : \lambda = 3.5$ against $H_1 : \lambda < 3.5$ using a 5% significance level.

a Find the critical region for this test.

b State the probability of committing a type I error for this test.

Given that the true value of λ is 3.0,

c find the power of this test.

Solution:

$H_0 : \lambda = 3.5 \quad H_1 : \lambda < 3.5$

a Seek c: $P(X \leq c) < 0.05$ where $X \sim Po(\lambda)\ (\lambda = 3.5)$

Tables: $P(X \leq 1) = 0.1359 > 0.05$

$P(X \leq 0) = 0.0302 < 0.05$

$\therefore$ critical region $X = 0$

b $P(\text{Type I error}) = 0.0302$

c $\text{Power} = P(X = 0 \mid \lambda = 3.0)$

$= 0.0498$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 3

Question:

The random variable $X \sim \mathrm{N}(\mu, 9)$. A random sample of 18 observations is taken, and it is desired to test $\mathrm{H}_0: \mu = 8$ against $\mathrm{H}_1: \mu \neq 8$, at the 5% significance level. The test statistic to be used is $Z = \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$.

a Find the critical region for this test.
b State the probability of a type I error for this test.
Given that μ was later found to be 7,
c find the probability of making a type II error.
d State the power of this test.

Solution:

$X \sim \mathrm{N}(\mu, 3^2)$ $\quad n = 18$
$\mathrm{H}_0: \mu = 8 \quad \mathrm{H}_1: \mu \neq 8$

a critical region when $|Z| > 1.96$

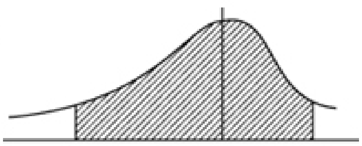
i.e. $\dfrac{\bar{X} - 8}{\frac{3}{\sqrt{18}}} < -1.96$ or $\dfrac{\bar{X} - 8}{\frac{3}{\sqrt{18}}} > 1.96$

$\Rightarrow \bar{X} < 6.614\ldots$ or $\bar{X} > 9.3859\ldots$

b $\mathrm{P}(\text{Type I error}) = \text{significance level} = 0.05$

c $\mathrm{P}(\text{Type II error}) = \mathrm{P}(6.614\ldots < \bar{X} < 9.3859\ldots \mid \mu = 7)$

$$= \mathrm{P}\left(\frac{6.614\ldots - 7}{\frac{3}{\sqrt{18}}} < Z < \frac{9.3859\ldots - 7}{\frac{3}{\sqrt{18}}}\right)$$
$$= \mathrm{P}(-0.545\ldots < Z < 3.374\ldots)$$
$$= 0.9996 - (1 - 0.7088)$$
$$= 0.7084$$



Calculator gives 0.707023 so accept awrt $0.707 \sim 0.708$

d $\text{Power} = 1 - \mathrm{P}(\text{Type II error})$
$= 0.293 \sim 0.292$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 4

Question:

A single observation, x, is taken from a Poisson distribution with parameter λ. The observation is used to test $H_0 : \lambda = 4.5$ against $H_1 : \lambda > 4.5$. The critical region chosen for this test was $x \geq 8$.

a Find the size of this test.

b The table below gives the power of the test for different values of λ.

λ	1	2	3	4	5	6	7	8	9	10
Power	0	0.0011	0.0019	r	0.1334	s	0.4013	0.5470	t	0.7798

i Find values for r, s and t.

ii Using graph paper, plot the power function against λ.

Solution:

$H_0 : \lambda = 4.5 \quad H_1 : \lambda > 4.5$

Critical region $X \geq 8$

a Size $= P(X \geq 8 \mid \lambda = 4.5)$

$= 1 - P(X \leq 7 \mid \lambda = 4.5) = 1 - 0.9134$

$= 0.0866$

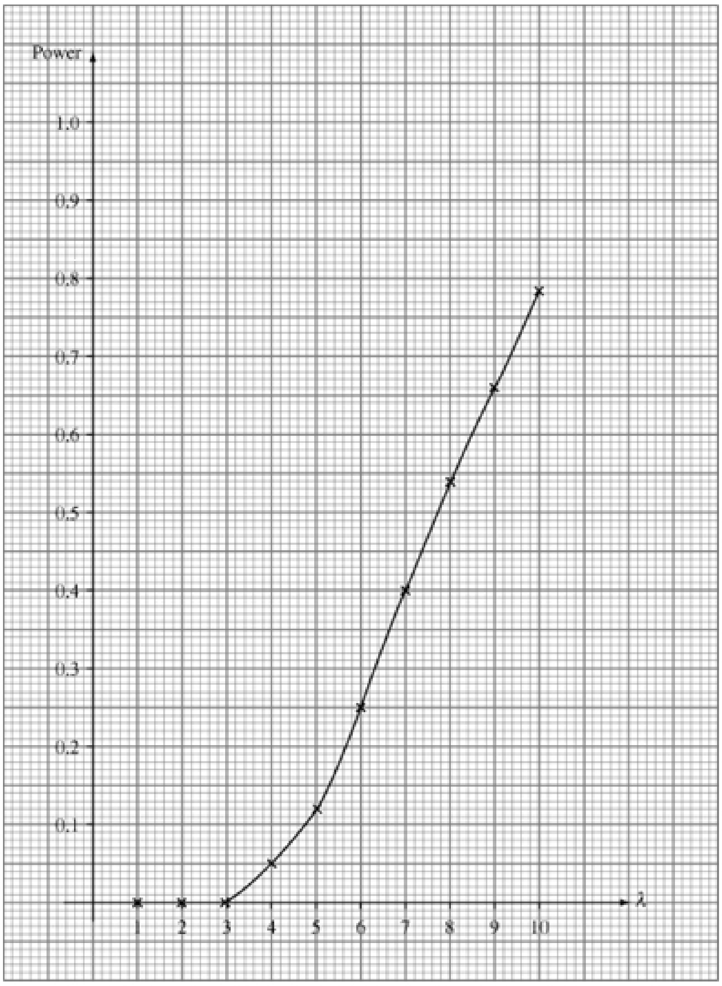
b i Power $= P(X \geq 8 \mid \lambda)$

$\therefore r = 1 - 0.9489 = 0.0511$

$s = 1 - 0.7440 = 0.2560$

$t = 1 - 0.3239 = 0.6761$

ii See graph.



Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 5

Question:

In a binomial experiment consisting of 15 trials X represents the number of successes and p the probability of success.

In a test of $H_0 : p = 0.45$ against $H_1 : p < 0.45$ the critical region for the test was $X \leq 3$.

a Find the size of the test.

b Use the table of the binomial cumulative distribution function to complete the table given below.

p	0.1	0.2	0.3	0.4	0.5
Power	0.944	s	0.2969	t	0.0176

c Draw the graph of the power function for this test

Solution:

$H_0 : p = 0.45 \quad H_1 : p < 0.45$

Critical region $X \le 3$

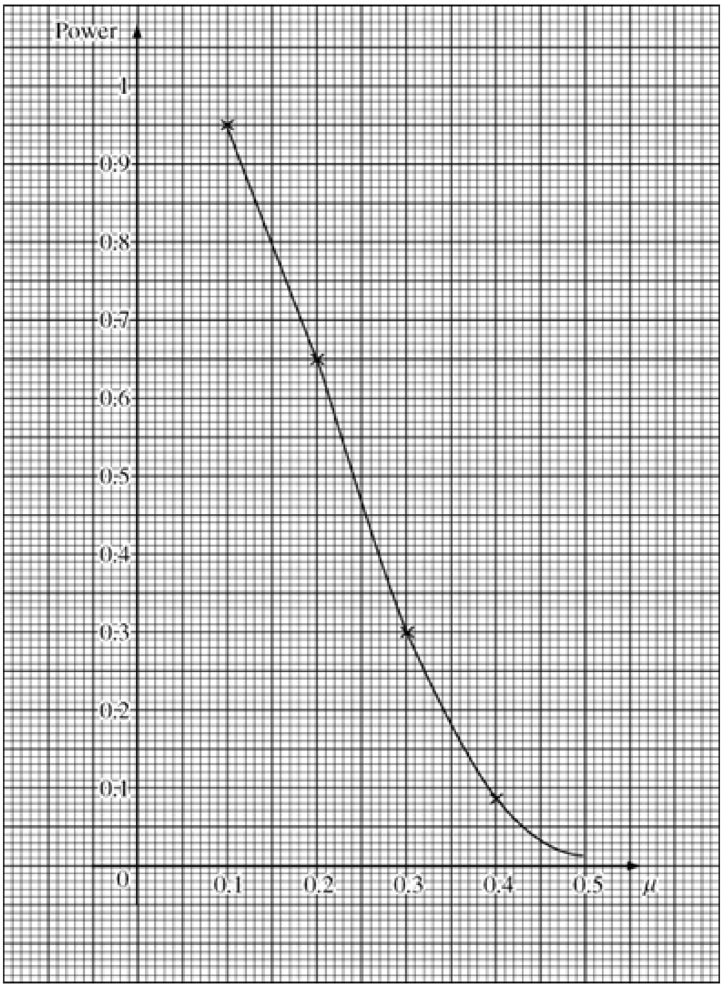
a $\text{Size} = P\left(X \le 3 \mid X \sim B\left(15, 0.45\right)\right) = 0.0424$

b $\text{Power} = P\left(X \le 3 \mid X \sim B(15, p)\right)$

$p = 0.2 \Rightarrow s = 0.6482$

$p = 0.4 \Rightarrow t = 0.0905$

c See Graph



Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 6

Question:

A bag contains 25 balls of which an unknown number, m, are coloured red $(3 < m \le 22)$. Two of the balls are drawn from the bag and the number of red balls, X, is noted. It is desired to estimate m by $\hat{m} = cX$.

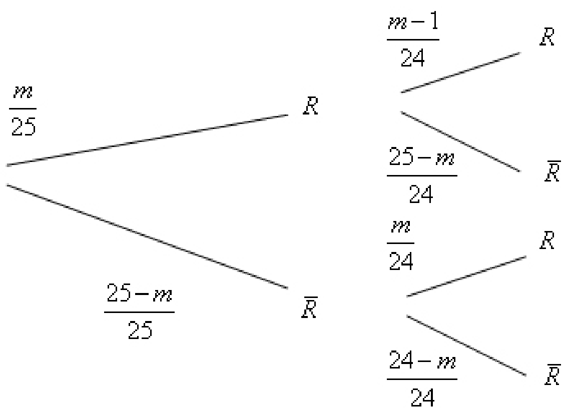
a Calculate a value for c if the estimate is to be unbiased.

The balls are replaced and a second draw is made and the number of red balls, Y, is noted.

b Write down $\mathrm{E}(Y)$.

c Show that $Z = (5X + 7.5Y)$ is an unbiased estimator of m.

Solution:



$X =$ number of reds

x	0	1	2
$\mathrm{P}(X = x)$	$\frac{(25-m)(24-m)}{25\times 24}$	$\frac{m(25-m)+(25-m)m}{25\times 24}$	$\frac{m(m-1)}{25\times 24}$

$$\mathrm{E}(X) = \frac{50m - 2m^2}{25\times 24} + \frac{2m^2 - 2m}{25\times 24} = \frac{48m}{25\times 24} = \frac{2m}{25}$$

a $\mathrm{E}(\hat{m}) = c\mathrm{E}(X) = c \times \frac{2m}{25}$

$\therefore$ for $\hat{m}$ to be unbiased for m, we need $c = \frac{25}{2}$

b $\mathrm{E}(Y) = \mathrm{E}(X) = \frac{2m}{25}$

c $\mathrm{E}(Z) = 5\mathrm{E}(X) + 7.5\,\mathrm{E}(Y)$

$$= 5 \times \frac{2m}{25} + \frac{7.5\times 2m}{25}$$

$$= \frac{25m}{25} = m$$

$\therefore Z$ is an unbiased estimater of m.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 7

Question:

A bag contains 25 balls of which an unknown number, m, are green, $(4 < m \leq 21)$.
Three balls are drawn from the bag and the number, X, of green balls is recorded. The balls are replaced and four balls are drawn with the number, Y, of green balls noted.
Three estimators of p, the probability of getting a green ball, are proposed

i $\frac{X+Y}{7}$

ii $\frac{3X+4Y}{25}$

iii $\frac{4X+3Y}{24}$.

a Show that all three are unbiased estimators of p.

b Find which is the best estimator.

Solution:

If 3 balls are selected $\mathrm{E}(X) = \frac{3m}{25}$ (compare with np for a binomial)

If 4 balls are selected $\mathrm{E}(Y) = \frac{4m}{25}$

a i $\mathrm{E}\left(\frac{X+Y}{7}\right) = \frac{\mathrm{E}(X)+\mathrm{E}(Y)}{7} = \frac{\frac{3m}{25}+\frac{4m}{25}}{7} = \frac{m}{25} = p$

ii $\mathrm{E}\left(\frac{3X+4Y}{25}\right) = \frac{\frac{9m}{25}+\frac{16m}{25}}{25} = \frac{m}{25} = p$

iii $\mathrm{E}\left(\frac{4X+3Y}{24}\right) = \frac{\frac{12m}{25}+\frac{12m}{25}}{24} = \frac{\frac{24m}{25}}{24} = \frac{m}{25} = p$

$\therefore$ all 3 are unbiased estimators of p

Similarly $\mathrm{Var}(X) = \frac{3m}{25}\frac{(25-m)}{25} = 3p(1-p)$

$\mathrm{Var}(Y) = \frac{4m}{25}\left(\frac{25-m}{25}\right) = 4p(1-p)$

b $\mathrm{Var}\left(\frac{X+Y}{7}\right) = \frac{\mathrm{Var}(X)+\mathrm{Var}(Y)}{49} = \frac{p(1-p)}{7} = 0.142p(1-p)$

$\mathrm{Var}\left(\frac{3X+4Y}{25}\right) = \frac{9\mathrm{Var}(X)+16\mathrm{Var}(Y)}{25^2} = \frac{(27+64)}{625}p(1-p) = \frac{91}{625}p(1-p)$

$= 0.1456p(1-p)$

$\mathrm{Var}\left(\frac{4X+3Y}{24}\right) = \frac{16\mathrm{Var}(X)+9\mathrm{Var}(Y)}{24^2} = \frac{48+36}{576}p(1-p) = \frac{84}{576}p(1-p)$

$= 0.1458p(1-p)$

$\because \frac{1}{7} < \frac{91}{625} < \frac{84}{576}$

$\therefore$ Choose $\frac{X+Y}{7}$ $\because$ variance is smallest.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 8

Question:

A company buys rope from Bindings Ltd and it is known that the number of faults per 100 m of their rope follows a Poisson distribution with mean 2. The company is offered 100 m of rope by Tieup, a newly established rope manufacturer. The company is concerned that the rope from Tieup might be of poor quality.

a Write down the null and alternative hypotheses appropriate for testing that rope from Tieup is in fact as reliable as that from Bindings Ltd.

b Derive a critical region to test your null hypothesis with a size of approximately 0.05.

c Calculate the power of this test if rope from Tieup contains an average of 4 faults per 100 m. **[E]**

Solution:

a $H_0: \lambda = 2$ $\quad$ $H_1: \lambda > 2$

(Quality the same) $\quad$ (Quality is poorer)

b Seek c such that $P(X \geq c) \approx 0.05$ where $X \sim Po(2)$

Tables $P(X \leq 4) = 0.9473 \Rightarrow P(X \geq 5) = 0.0527$

$P(X \leq 5) = 0.9834 \Rightarrow P(X \geq 6) = 0.0166$

Nearest to 0.05 is $X \geq 5$

critical region is $X \geq 5$

c Power $= P(X \geq 5 \mid \lambda = 4)$

$= 1 - P(X \leq 4 \mid \lambda = 4)$

$= 1 - 0.6288$

$= 0.3712$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 9

Question:

The number of faulty garments produced per day by machinists in a clothing factory has a Poisson distribution with mean 2. A new machinist is trained and the number of faulty garments made in one day by the new machinist is counted.

a Write down the appropriate null and alternative hypotheses involved in testing the theory that the new machinist is at least as reliable as the other machinists.

b Derive a critical region, of size approximately 0.05, to test the null hypothesis.

c Calculate the power of this test if the new machinist produces an average of 3 faulty garments per day.

The number of faulty garments produced by the new machinist over three randomly selected days is counted.

d Derive a critical region, of approximately the same size as in part **b**, to test the null hypothesis.

e Calculate the power of this test if the machinist produces an average of 3 faulty garments per day.

f Comment briefly on the difference between the two tests. **[E]**

Solution:

a $H_0: \lambda = 2$ (as good) $H_1: \lambda > 2$ (worse)

b Seek c such that $P(X \geq c) \approx 0.05$ where $X \sim Po(2)$

Tables $P(X \leq 4) = 0.9473 \Rightarrow P(X \geq 5) = 0.0527$ closest to 0.05

$P(X \leq 5) = 0.9834 \Rightarrow P(X \geq 6) = 0.0166$

$\therefore$ critical region is $X \geq 5$

c Power $= P(X \geq 5 \mid \lambda = 3)$

$= 1 - P(X \leq 4 \mid \lambda = 3) = 1 - 0.8153 = 0.1847$

d Seek d such that $P(X \geq d) \approx 0.05$ where $X \sim Po(6)$

Tables $P(X \leq 10) = 0.9574 \Rightarrow P(X \geq 11) = 0.0426$

$\therefore$ critical region is $X \geq 11$

e Power $= P(X \geq 11 \mid \lambda = 9)$ [3 days has mean $= 3 \times 3 = 9$]

$= 1 - P(X \leq 10 \mid \lambda = 9)$

$= 1 - 0.7060$

$= 0.294$

f Second test is more powerful as it uses more days.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 10

Question:

A single observation, x, is to be taken from a Poisson distribution with parameter μ. This observation is to be used to test $H_0 : \mu = 6$ against $H_1 : \mu < 6$. The critical region is chosen to be $x \leq 2$.

a Find the size of the critical region.

b Show that the power function for this test is given by $\frac{1}{2}e^{-\mu}(2+2\mu+\mu^2)$

The table below gives the values of the power function to 2 decimal places.

μ	1.0	1.5	2.0	4.0	5.0	6.0	7.0
Power	0.92	0.81	s	0.24	t	0.06	0.03

c Calculate the values of s and t.

d Draw a graph of the power function.

e Find the range of values of μ for which the power of this test is greater than 0.8.

[E]

Solution:

$H_0 : \mu = 6 \quad H_1 : \mu < 6$

Critical region is $X \leq 2$

a Size $= P(X \leq 2 \mid X \sim Po(6))$

$= 0.0620$

b Power $= P(X \leq 2 \mid X \sim Po(6))$

$$= e^{-\mu} + \mu e^{-\mu} + \frac{\mu^2}{2} e^{-\mu}$$

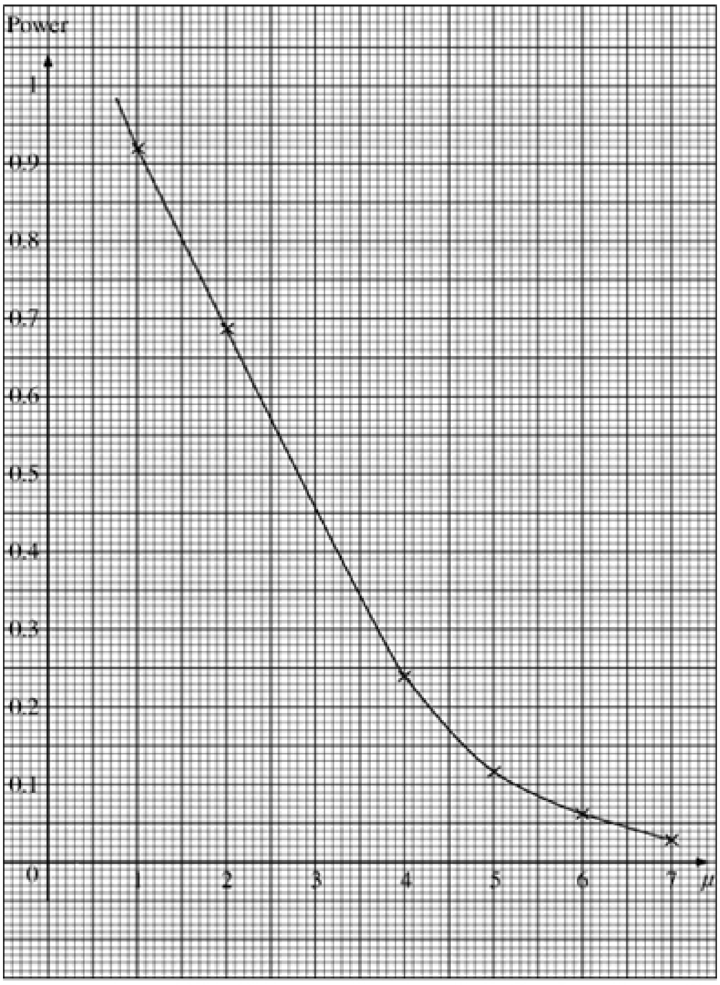
$$= e^{-\mu}\left(1 + \mu + \frac{\mu^2}{2}\right)$$

$$= \frac{e^{-\mu}}{2}\left(2 + 2\mu + \mu^2\right)$$

c $s = 0.6767$

$t = 0.1247$

d See Graph.



e From Graph

$\mu < 1.55$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise G, Question 11

Question:

The random variable X has the following distribution:

x	0	1
$P(X = x)$	q	p

a Find E(X) and Var(X).

A random sample X_1, X_2, X_3, is taken from the distribution in order to estimate p.

b Find the condition which must be satisfied by the constants a_1, a_2, a_3, if $a_1X_1 + a_2X_2 + a_3X_3$ is to be an unbiased estimator of p.

c Find the variance of this estimator.

The following estimators are proposed:

i $\frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3$

ii $\frac{1}{3}X_1 + \frac{1}{6}X_2 + \frac{5}{12}X_3$

iii $\frac{7}{12}X_1 + \frac{5}{12}X_2$

d Of these three estimators, find the best unbiased estimator. **[E]**

Solution:

a $\mathrm{E}(X) = 0 + p = p$

$\mathrm{E}(X^2) = 0 + 1^2 p = p \quad \therefore \mathrm{Var}(X) = p - p^2 = p(1-p)$

b $Y = a_1X_1 + a_2X_2 + a_3X_3$

$\mathrm{E}(Y) = (a_1 + a_2 + a_3)p \quad \therefore$ for Y to be unbiased estimator of p

You need $a_1 + a_2 + a_3 = 1$

c $\mathrm{Var}(Y) = a_1^2 p(1-p) + a_2^2 p(1-p) + a_3^2 p(1-p)$

$= (a_1^2 + a_2^2 + a_2^2)pq$

d i $\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1 \therefore$ unbiased

Variance $= \left(\frac{1}{36} + \frac{1}{9} + \frac{1}{4}\right)pq = \left(\frac{1+4+9}{36}\right)pq = \frac{14}{36}pq = \frac{28}{72}pq$

ii $\frac{1}{3} + \frac{1}{6} + \frac{5}{12} \neq 1 \therefore$ biased

iii $\frac{7}{12} + \frac{5}{12} = 1 \therefore$ unbiased

Variance $= \left(\frac{49}{144} + \frac{25}{144}\right)pq = \frac{74}{144}pq = \frac{37}{72}pq$

$\therefore$ best estimator is $\frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3$ as it has the smallest variance

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 12

Question:

Two sets of binomial trials were carried out and in both sets the probability of success is p. In the first set there were X successes out of n trials and in the second set there were Y successes out of m trials.

Possible estimators for p are $\hat{p}_1 = \frac{1}{2}\left(\frac{X}{n} + \frac{Y}{m}\right)$ and $\hat{p}_2 = \frac{X+Y}{n+m}$

a Show that both $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators of p.

b Find the variances of $\hat{p}_1$ and $\hat{p}_2$

c If $n = 10$ and $m = 20$ state, giving a reason, which estimator you would use. **[E]**

Solution:

$X \sim \mathrm{B}(n, p) \Rightarrow \mu_x = np \quad \sigma_x^2 = np(1-p)$

$Y \sim \mathrm{B}(m, p) \Rightarrow \mu_y = mp \quad \sigma_y^2 = mp(1-p)$

a $\mathrm{E}(\hat{p}_1) = \frac{1}{2}\left[\frac{\mathrm{E}(X)}{n} + \frac{\mathrm{E}(Y)}{m}\right] = \frac{1}{2}\left[\frac{np}{n} + \frac{mp}{m}\right] = p$

$\mathrm{E}(\hat{p}_2) = \frac{\mathrm{E}(X) + \mathrm{E}(Y)}{n+m} = \frac{np + mp}{n+m} = \frac{(n+m)p}{n+m} = p$

$\therefore$ both $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators of p

b $\mathrm{Var}(\hat{p}_1) = \frac{1}{4}\left[\frac{\mathrm{Var}(X)}{n^2} + \frac{\mathrm{Var}(Y)}{m^2}\right] = \frac{1}{4}\left[\frac{np(1-p)}{n^2} + \frac{mp(1-p)}{m^2}\right] = \frac{(m+n)p(1-p)}{4mn}$

$\mathrm{Var}(\hat{p}_2) = \frac{\mathrm{Var}(X) + \mathrm{Var}(Y)}{(n+m)^2} = \frac{np(1-p) + mp(1-p)}{(n+m)^2} = \frac{p(1-p)}{n+m}$

c $n = 10, m = 20 \Rightarrow \mathrm{Var}(\hat{p}_1) = \frac{(20+10)p(1-p)}{4(20)(10)}$

$= \frac{3p(1-p)}{80}$

and $\mathrm{Var}(\hat{p}_2) = \frac{p(1-p)}{30}$

$\because \frac{1}{30} < \frac{3}{80} \quad (\because 80 < 90)$

$\therefore$ use $\hat{p}_2$ $\because$ unbiased and has smaller variance.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise G, Question 13

Question:

(In this question $\max(a, b)$ = the greater of the two values a and b.)
A palaeontologist was attempting to estimate the length of time, T, in years, during which a small herbivorous dinosaur existed on Earth. He believed from other evidence that the earliest existence of the animal had been at the start of the Jurassic period.

Two examples of the animal had been discovered in the fossil record, at times t_1 and t_2 after the start of the Jurassic period. His model assumed that these times were values of two independent random variables T_1 and T_2 each having a continuous uniform distribution on the interval $[0, \tau]$. He considered three estimators for τ:
$\tau_1 = T_1 + T_2$, $\tau_2 = \sqrt{3}|T_2 - T_1|$, $\tau_3 = 1.5\max(T_1, T_2)$
He used appropriate probability theory and calculated the results shown in the table.

Variable	Expectation	Variance
T_1	$\frac{\tau}{2}$	$\frac{\tau^2}{12}$
$\|T_2 - T_1\|$	$\frac{\tau}{3}$	$\frac{\tau^2}{18}$
$\max(T_1, T_2)$	$\frac{2\tau}{3}$	$\frac{\tau^2}{18}$

Using these results,
a determine the bias of each of his estimators,
b find the variance of each of his estimators.
Using your results from **a** and **b**, state, giving a reason,
c which estimator is the best of the three,
d which estimator is the worst.

Solution:

a $\mathrm{E}(\tau_1) = \mathrm{E}(T_1 + T_2) = \frac{\tau}{2} + \frac{\tau}{2} = \tau \therefore$ unbiased

$\mathrm{E}(\tau_2) = \sqrt{3}\mathrm{E}\,|T_2 - T_1| = \sqrt{3} \cdot \frac{\tau}{3} \therefore \text{bias} = \frac{\sqrt{3}}{3}\tau - \tau = \tau\left(\frac{\sqrt{3}}{3} - 1\right)$

$\mathrm{E}(\tau_3) = 1.5\mathrm{E}(\max(T_1, T_2)) = 1.5 \cdot \frac{2\tau}{3} = \tau \therefore$ unbiased.

$\mathrm{Var}(\tau_1) = \mathrm{Var}(T_1) + \mathrm{Var}(T_2) = \frac{\tau^2}{12} + \frac{\tau^2}{12} = \frac{\tau^2}{6}$

$\mathrm{Var}(\tau_2) = \sqrt{3}^2 \times \frac{\tau^2}{18} = \frac{\tau^2}{6}$

$\mathrm{Var}(\tau_3) = \frac{9}{4} \times \frac{\tau^2}{18} = \frac{\tau^2}{8}$

c τ_3 is best since it is unbiased and it has the smallest variance.

d τ_2 is worst since it is biased (and variance is just the same as τ_1)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 1

Question:

Given that the random variable X has a t_{12}-distribution, find values of t such that,

a $P(X < t) = 0.025$,

b $P(X > t) = 0.05$,

c $P(|X| > t) = 0.95$.

Solution:

a $P(X > t) = 0.025$ when $t = 2.179$ so $P(X < t) = 0.025$ when $t = -2.179$

b $P(X > t) = 0.05$ when $t = 1.782$

c $P(X > t) = 0.025$ when $t = 2.179$

$P(|X| > t) = 0.95$ when $|t| = 2.179$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 2

Question:

Given that the random variable X has a t_{26}-distribution. Find

a $t_\nu(0.01)$,

b $t_\nu(0.05)$.

Solution:

a 2.479 (from tables)
b 1.706 (from tables)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 3

Question:

The random variable Y has a t_ν-distribution. Find a value (or values) of t in each of the following.

a $\nu = 10, \text{P}(Y < t) = 0.95$

b $\nu = 32, \text{P}(Y < t) = 0.005$

c $\nu = 5, \text{P}(Y < t) = 0.025$

d $\nu = 16, \text{P}(|Y| < t) = 0.98$

e $\nu = 18, \text{P}(|Y| > t) = 0.10$

Solution:

a $\text{P}(Y > t) = 0.05$ when $t = 1.812$ so $\text{P}(Y < t) = 0.95$ when $t = 1.812$

b $\text{P}(Y > t) = 0.005$ when $t = 2.738$ so $\text{P}(Y < t) = 0.005$ when $t = -2.738$

c $\text{P}(Y > t) = 0.025$ when $t = 2.571$ so $\text{P}(Y < t) = 0.025$ when $t = -2.571$

d $\text{P}(Y > t) = 0.01$, when $t = 2.583$, and $\text{P}(Y < t) = 0.01$ when $t = -2.583$ so $\text{P}(|Y| < t) = 0.98$ when $|t| = 2.583$

e $\text{P}(Y > t) = 0.05$ when $t = 1.734$ and $\text{P}(Y < t) = 0.05$ when $t = -1.734$ so $\text{P}(|Y| > t) = 0.10$ when $|t| = 1.734$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 1

Question:

A test on the life (in hours) of a certain make of torch batteries gave the following results:

20.3 17.3 25.0 18.4 16.3 24.8 24.3 21.2

Assuming that the lifetime of batteries is normally distributed, find a 90% confidence interval for the mean.

Solution:

$\bar{x} = 20.95 \quad s = 3.4719\ldots n = 8 \quad \nu = 7$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 20.95 \pm 1.895 \times \frac{3.4719\ldots}{\sqrt{8}} = 18.624$ and 23.276

Confidence interval $= (18.624, 23.276)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 2

Question:

A sample of size 16 taken from a normal population with unknown variance gave the following sample values $\bar{x} = 12.4, s^2 = 21.0$. Find a 95% confidence interval on the population mean.

Solution:

$\bar{x} = 12.4 \quad s = \sqrt{21} \quad n = 16 \quad v = 15$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 12.4 \pm 2.131 \times \frac{\sqrt{21}}{\sqrt{16}} = 9.9586\ldots$ and

$14.8413\ \ldots$

Confidence interval $= (9.959, 14.841)$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 3

Question:

The mean heights (measured in centimetres) of six male students at a college were as follows:

182 178 183 180 169 184

Calculate,

a a 90% confidence interval and

b a 95% confidence interval for the mean height of male students at the college.

You may assume that the heights are normally distributed.

Solution:

a $\bar{x} = 179.333333 \quad s = 5.5015\ldots n = 6 \quad \nu = 5$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 179.333.. \pm 2.015 \times \frac{5.5015\ldots}{\sqrt{6}} = 174.808$

and 183.859

Confidence interval $= (174.808, 183.859)$

b confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 179.333.. \pm 2.571 \times \frac{5.5015\ldots}{\sqrt{6}} = 173.559$

and 185.108

Confidence interval $= (173.559, 185.108)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 4

Question:

The masses (in grams) of 10 nails selected at random from a bin of 90 cm long nails were:

9.7 10.2 11.2 9.4 11.0 11.2 9.8 9.8 10.0 11.3

Calculate a 98% confidence interval for the mean mass of the nails, assuming that their mass is normally distributed.

Solution:

$\bar{x} = 10.36 \quad s = 0.73363\ldots \; n = 10 \quad v = 9$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 10.36 \pm 2.821 \times \frac{0.73363\ldots}{\sqrt{10}} = 9.706$ and 11.014

Confidence interval $= (9.706, 11.014)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 5

Question:

It is known that the length of men's feet is normally distributed. A random sample of the feet of 8 adult males gave the following summary statistics of length x (in cm):

$\sum x = 224.1 \quad \sum x^2 = 6337.39$

Calculate a 99% confidence interval for the mean lengths of men's feet based upon these results.

Solution:

$\bar{x} = \frac{224.1}{8} = 28.0125 \quad s^2 = \frac{1}{7}\left\{6337.39 - \frac{224.1^2}{8}\right\} = 8.54125$

$s = 2.92254\ldots n = 8 \quad \nu = 7$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 28.0125 \pm 3.499 \times \frac{2.92254\ldots}{\sqrt{8}} = 24.397$

and 31.628

Confidence interval = (24.397, 31.628)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 6

Question:

A random sample of 26 students from the sixth form of a school sat an intelligence test that measured their IQs. The result are summarised below

$\bar{x} = 122 \quad s^2 = 225$

Assuming that the IQ is normally distributed, calculate a 95% confidence interval for the mean IQ of the students.

Solution:

$\bar{x} = 122 \quad s = \sqrt{225} = 15 \quad \nu = 25$

confidence limits $= \bar{x} \pm t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} = 122 \pm 2.060 \times \frac{\sqrt{225}}{\sqrt{26}} = 115.940$ and 128.060

Confidence interval $= (115.94, 128.06)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 1

Question:

Given that the observations 9, 11, 11, 12, 14, have been drawn from a normal distribution, test $H_0 : \mu = 11$ against $H_1 : \mu > 11$. Use a 5% significance level.

Solution:

$\bar{x} = 11.4 \quad s = 1.816\ldots$

$H_0 : \mu = 11 \quad H_1 : \mu > 11$

Critical region $t > 2.132$

Test statistic $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}} = \dfrac{11.4 - 11.0}{\dfrac{1.816\ldots}{\sqrt{5}}} = 0.492$

The result is not in critical region.

No evidence that μ is not 11

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 2

Question:

A random sample of size 28 taken from a normally distributed variable gave the following sample values $\bar{x} = 17.1$ and $s^2 = 4$. Test $H_0 : \mu = 19$ against $H_1 : \mu < 19$. Use a 1% level of significance.

Solution:

$\bar{x} = 17.1 \quad s = 2$

$H_0 : \mu = 19 \quad H_1 : \mu < 19$

Critical region $t < -2.473$

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{17.1 - 19}{\frac{2}{\sqrt{28}}} = -5.027$

The result is in the critical region.

There is evidence that μ is less than 19

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 3

Question:

A random sample of size 13 taken from a normally distributed variable gave the following sample values $\bar{x} = 3.26, s^2 = 0.64$. Test $H_0 : \mu = 3$ against $H_1 : \mu \neq 3$. Use a 5% significance level.

Solution:

$\bar{x} = 3.26 \quad s = 0.8$

$H_0 : \mu = 3 \quad H_1 : \mu \neq 3$

Critical values ± 2.179

Critical region $t < -2.179$ or $t > 2.179$

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{3.26 - 3}{\frac{0.8}{\sqrt{13}}} = 1.172$

The result is not in the critical region

There is no evidence that μ is not 3

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 4

Question:

A certain brand of blanched hazelnuts for use in cooking is sold in packets. The weights of the packets of hazelnuts follow a normal distribution with mean μ. The manufacturer claims that $\mu = 100$ g. A sample of 15 packets was taken and the weight x of each was measured. The results are summarised by the following statistics $\sum x = 1473$, $\sum x^2 = 148119$.

Test at the 5% significance level whether or not there is evidence to justify the manufacturer's claim.

Solution:

$\bar{x} = 98.2 \quad s = 15.744\ldots$

$H_0 : \mu = 100 \quad H_1 : \mu \neq 100$

Critical region < -2.145 or > 2.145

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{98.2 - 100}{\frac{15.74\ldots}{\sqrt{15}}} = -0.443$

The result is not in the critical region

There is no evidence that μ is not 100

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 5

Question:

A manufacturer claims that the lifetimes of its 100 watt bulbs are normally distributed with a mean of 1000 hours. A laboratory tests 8 bulbs and finds their lifetimes to be 985, 920, 1110, 1040, 945, 1165, 1170, and 1055 hours.
Stating your hypotheses clearly, examine whether or not the bulbs have a longer mean life than that claimed. Use a 5% level of significance.

Solution:

$\bar{x} = 1048.75 \quad s = 95.2346\ldots$

$H_0 : \mu = 1000 \quad H_1 : \mu > 1000$

Critical region $t > 1.895$

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{1048.75 - 1000}{\frac{95.234\ldots}{\sqrt{8}}} = 1.448$

The result is not in the critical region

There is no evidence that μ is not 1000

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 6

Question:

A fertiliser manufacturer claims that by using brand F fertiliser the yield of fruit bushes will be increased. A random sample of 14 fruit bushes was fertilised with brand F and the resulting yields, x, were summarised by $\sum x = 90.8, \sum x^2 = 600$. The yield of bushes fertilised by the usual fertiliser was normally distributed with a mean of 6 kg per bush.

Test, at the 2.5% significance level, the manufacturer's claim.

Solution:

$\bar{x} = 6.4857\ldots \quad s^2 = 0.853626\ldots \quad s = 0.923919\ldots$

$H_0 : \mu = 6 \quad H_1 : \mu > 6$

Critical region $t > 2.160$

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{6.4857142 - 6}{\frac{0.923919\ldots}{\sqrt{14}}} = 1.967$

The result is not in the critical region

There is no evidence supporting manufacturer's claim.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 7

Question:

A nuclear reprocessing company claims that the amount of radiation within a reprocessing building in which there had been an accident had been reduced to an acceptable level by their clean up team. The amount of radiation, x, at 20 sites within the building in suitable units are summarised by $\sum x = 21.7$, $\sum x^2 = 28.4$. Before the accident the level of radiation in the building was normally distributed with a mean of 1.00. Test, at the 0.10 level whether or not the claim is justified.

Solution:

$\bar{x} = 1.085 \quad s^2 = \frac{28.4 - 20 \times 1.085^2}{19} = 0.2555\ldots \quad s = 0.5055\ldots$

$H_0: \mu = 1.00 \quad H_1: \mu > 1.00$

Critical values $t < 1.328$

Test statistic $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{1.085 - 1}{\frac{0.5055..}{\sqrt{20}}} = 0.752$

The result is not in the critical region

There is no evidence that μ is not 1.00 so the claim is not justified

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 1

Question:

A random sample of 15 observations of a normal population gave an unbiased estimate for the variance of the population of $s^2 = 4.8$. Calculate a 95% confidence interval for the population variance.

Solution:

$$\text{Confidence interval} = \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)} \right) = \left(\frac{14 \times 4.8}{26.119}, \frac{14 \times 4.8}{5.629} \right)$$
$$= (2.573, 11.938)$$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 2

Question:

A random sample of 20 observations of a normally distributed variable X is summarised by $\sum x = 132.4$ and $\sum x^2 = 884.3$. Calculate a 90% confidence interval for the variance of X.

Solution:

$$\bar{x} = 6.62 \quad s^2 = \frac{1}{19}\left(884.3 - \frac{132^2}{20}\right) = 0.4111\ldots$$

$$\text{Confidence interval} = \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right) = \left(\frac{19\times 0.411\ldots}{30.144}, \frac{19\times 0.411..}{10.117}\right)$$

$$= (0.259, 0.772)$$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 3

Question:

A random sample of 14 observations is taken from a population that is assumed to be normally distributed. The resulting values were:

2.3, 3.9, 3.5, 2.2, 2.6, 2.5, 2.3, 3.9, 2.1, 3.6, 2.1, 2.7, 3.2, 3.4

Calculate a 95% confidence interval for the population variance.

Solution:

$\bar{x} = 2.8785 \quad \ldots. \quad s^2 = 0.45873\ldots$

$$\text{Confidence interval} = \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)} \right) = \left(\frac{13 \times 0.458\ldots}{24.736}, \frac{13 \times 0.458\ldots}{5.009} \right)$$

$$= (0.241, 1.191)$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 4

Question:

A random sample of female voles was trapped in a wood. Their lengths, in centimetres (excluding tails) were 7.5, 8.4, 10.1, 6.2, and 8.4 cm.
Assuming that this is a sample from a normal distribution, calculate 95% confidence intervals for:

a the mean length,

b the variance of the lengths of female voles.

Solution:

$\bar{x} = 8.12 \quad s^2 = 2.037\ldots\ldots s = 1.427\ldots$

a Confidence interval $= \bar{x} \pm t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} =$

$$\left(8.12 - 2.776 \times \frac{1.427..}{\sqrt{5}},\ 8.12 + 2.776 \times \frac{1.427..}{\sqrt{5}}\right) = (6.348, 9.892)$$

b Confidence interval $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right) = \left(\frac{4 \times 2.037}{11.143}, \frac{4 \times 2.037}{0.484}\right)$

$= (0.731, 16.835)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 5

Question:

a A random sample of 10 is taken from the annual rainfall figures, x cm, in a certain district. The result is summarised by $\sum x = 621$ and $\sum x^2 = 38\,938$.

Calculate 90% confidence limits for,

i the mean annual rainfall,

ii the variance of the annual rainfall.

b What assumption have you made about the distribution of the annual rainfall in part **a**?

Solution:

a $\bar{x} = 62.1 \qquad s^2 = \frac{1}{9}\left(38\,938 - \frac{621^2}{10}\right) = 41.544\ldots \qquad s = 6.445\ldots$

i $$\text{Confidence interval} = \bar{x} \pm t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}$$
$$= \left(62.1 - 1.833 \times \frac{6.445\ldots}{\sqrt{10}}, 62.1 + 1.833 \times \frac{6.445\ldots}{\sqrt{10}}\right)$$
$$= (58.364, 65.836)$$

ii $$\text{Confidence interval} = \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right) = \left(\frac{9 \times 41.544\ldots}{16.919}, \frac{9 \times 41.544}{3.325}\right)$$
$$= (22.099, 112.450)$$

b Normal distribution

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 6

Question:

A new variety of small daffodil is grown in the trial ground of a nursery. During the flowering period a random sample of 10 flowers was taken and the lengths, in millimetres, of their stalks were measured. The results were as follows:

266, 254, 215, 220, 253, 230, 216, 248, 234, 244 mm.

Assuming that the lengths are normally distributed, calculate 95% confidence intervals for the mean and variance of the lengths.

Solution:

$\bar{x} = 238 \; s = 17.694\ldots \quad s^2 = 313.111\ldots$

$$\text{Confidence interval mean} = \left(\bar{x} - t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}},\ \bar{x} + t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}\right)$$

$$= \left(238 - 2.262 \times \frac{17.694\ldots}{\sqrt{10}}, 238 + 2.262 \times \frac{17.694\ldots}{\sqrt{10}}\right)$$

$$= (225.343, 250.657)$$

$$\text{Confidence interval variance} = \left(\frac{(n-1)s^2}{\chi_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right)}\right)$$

$$= \left(\frac{9 \times 313.111\ldots}{19.023}, \frac{9 \times 313.11\ldots}{2.700}\right)$$

$$= (148.136, 1043.704)$$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 1

Question:

Twenty random observations (x) are taken from a normal distribution with variance σ^2. The results are summarised as follows:

$\sum x = 332.1, \sum x^2 = 5583.63$

a Calculate an unbiased estimate for the population variance.

b Test, at the 5% significance level, the hypothesis $H_0 : \sigma^2 = 1.5$ against the hypothesis $H_0 : \sigma^2 > 1.5$.

Solution:

a $\bar{x} = 16.605 \quad s^2 = \frac{5583.63 - 20(16.605)^2}{19} = 3.637\ldots$

b $H_0 : \sigma^2 = 1.5 \quad H_1 : \sigma^2 > 1.5$

Critical region > 30.144

Test statistic $= \frac{(n-1)s^2}{\sigma^2} = \frac{19 \times 3.637..}{1.5} = 46.072$

The test statistic is in the critical region

There is evidence to suggest $\sigma^2 > 1.5$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 2

Question:

A random sample of 10 observations is taken from a normal distribution with variance σ^2 which is thought to be equal to 0.09. The results were as follows:

0.35, 0.42, 0.30, 0.26, 0.31, 0.30, 0.40, 0.33, 0.30, 0.40

Test, at the 0.025% level of significance, the hypothesis $H_0 : \sigma^2 = 0.09$ against the hypothesis $H_0 : \sigma^2 < 0.09$.

Solution:

$\bar{x} = 0.337 \qquad s^2 = 0.0028677\ldots$

$H_0 : \sigma^2 = 0.09 \qquad H_1 : \sigma^2 < 0.09$

Critical region < 2.700

$$\text{Test statistic} = \frac{(n-1)s^2}{\sigma^2} = \frac{9 \times 0.0028677\ldots}{0.09} = 0.287$$

The test statistic is in the critical region.

There is evidence to suggest that variance is less than 0.09.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 3

Question:

The following random observations are taken from a normal distribution which is thought to have a variance of 4.1:

2.1, 2.3, 3.5, 4.6, 5.0, 6.4, 7.1, 8.6, 8.7, 9.1

Test, at the 5% significance level, the hypothesis $H_0 : \sigma^2 = 4.1$ against the hypothesis $H_1 : \sigma^2 \neq 4.1$

Solution:

$H_0 : \sigma^2 = 4.1 \quad H_1 : \sigma^2 \neq 4.1$

$\bar{x} = 5.74 \quad s^2 = 6.940\ldots$

Critical region < 2.7 and > 19.023

Test statistic $= \frac{(n-1)s^2}{\sigma^2} = \frac{9 \times 6.940\ldots}{4.1} = 15.235$

The test statistic is not in the critical region.

There is no evidence the variance does not equal 4.1.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 4

Question:

It is claimed that the masses of a particular component produced in a small factory are normally distributed and have a mean mass of 10 g and a standard deviation of 1.12 g. A random sample of 20 such components was found to have a variance of 1.15 g.

Test, at the 5 % significance level, the hypothesis $H_0 : \sigma^2 = 1.12^2$ against the hypothesis $H_1 : \sigma^2 \neq 1.12^2$

Solution:

$H_0 : \sigma^2 = 1.12^2 \quad H_1 : \sigma^2 \neq 1.12^2$

Critical region < 8.907 and > 32.852

Test statistic $= \frac{(n-1)s^2}{\sigma^2} = \frac{19 \times 1.15}{1.12^2} = 17.419$

The test statistic is not in the critical region

There is no evidence the variance does not equal 1.12

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 5

Question:

Rollers for use in roller bearings are produced on a certain machine. The rollers are supposed be normally distributed and to have a mean diameter (μ) of 10 mm with a variance (σ^2) of 0.04 mm^2.

A random sample of 15 rollers from the machine have their diameters, x in millimetres, measured. The results are summarised below:

$\sum x = 149.941 \quad \sum x^2 = 1498.83$

a Calculate unbiased estimates for μ and σ^2.

b Test at the 5 % significance level,

i the hypothesis $\mu = 10$ against the hypothesis $\mu \neq 10$, using your estimate for σ^2 as the true variance of the population

ii the hypothesis $\sigma^2 = 0.04$ against the hypothesis $\sigma^2 \neq 0.04$

Solution:

a $\bar{x} = \frac{149.941}{15} = 9.996\ldots, s^2 = \frac{1498.83 - 15 \times 9.996\ldots^2}{14} = 0.0006977\ldots$

$s = 0.0264$

b **i** $H_0 : \mu = 10 \quad H_1 : \mu \neq 10$

Critical region < -2.145 and > 2.145.

$$\text{Test statistic} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{9.996.. - 10}{\frac{0.0264\ldots}{\sqrt{15}}} = -0.587 \text{ (or } -0.577 \text{ of full calculator accuracy used)}$$

The test statistic is not in the critical region

There is no evidence that the mean does not equal 10

ii $H_0 : \sigma^2 = 0.04 \quad H_1 : \sigma^2 \neq 0.04$

Critical region < 5.629 and > 26.119

$$\text{Test statistic} = \frac{(n-1)s^2}{\sigma^2} = \frac{14 \times 0.0006977}{0.04} = 0.244$$

This is in the critical region

There is evidence that the variance is not 0.04

It is less than 0.04 which is good in this context because it means that there is very little variability.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 6

Question:

The diameters of the eggs of the little gull are approximately normally distributed with mean 4.11 cm with a variance of 0.19 cm^2.

A sample of 8 little gulls eggs from a particular island which were measured had diameters in centimetres as follows:

4.4, 4.5, 4.1, 3.9, 4.4, 4.6, 4.5, 4.1

a Calculate an unbiased estimate for the variance of the population of little gull eggs on the island.

b Calculate an unbiased estimate of the mean diameter of the eggs and, test, at the 5% level, the hypothesis $\mu = 4.11$ against the hypothesis $\mu > 4.11$

c Test, at the 10% significance level, the hypothesis $\sigma^2 = 0.19$ against the hypothesis $\sigma^2 \neq 0.19$

Solution:

a $s^2 = 0.06125$

b $\bar{x} = 4.3125$

$H_0 : \mu = 4.11 \quad H_1 : \mu > 4.11$

Critical region > 1.895

$$\text{Test statistic} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{4.3125 - 4.11}{\frac{\sqrt{0.06125}}{\sqrt{8}}} = 2.3143$$

The test statistic is in the critical region

The mean weight is greater than 4.11

c $H_0 : \sigma^2 = 0.19 \quad H_1 : \sigma^2 \neq 0.19$

Critical region < 2.167 and > 14.067

$$\text{Test statistic} = \frac{(n-1)s^2}{\sigma^2} = \frac{7 \times 0.06125}{0.19} = 2.256$$

The test statistic is not in the critical region.

There is no evidence that σ^2 does not equal 0.19

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 7

Question:

Climbing rope produced by a certain manufacturer is known to have a mean tensile breaking strength (μ) of 170.2 kg and standard deviation 10.5 kg. The breaking strength of the rope is normally distributed.
A new component is added to the material which will, it is claimed, decrease the standard deviation without altering the tensile strength. A random sample of 20 pieces of the new rope is selected and each is tested to destruction. The tensile strength of each piece is noted. The results are used to calculate unbiased estimates of the mean strength and standard deviation of the population of new rope. These were found to be 172.4 kg and 8.5 kg.

a Test at the 5% level whether or not the variance has been reduced.

b What recommendation would you make to the manufacturer?

Solution:

a $H_0 : \sigma^2 = 110.25 \quad H_1 : \sigma^2 < 110.25$ [$10.5^2 = 110.25$]

Critical region < 10.117

$$\text{Test statistic} = \frac{(n-1)s^2}{\sigma^2} = \frac{19 \times 8.5^2}{110.25} = 12.451$$

The test statistic is not in the not critical region
There is no evidence that the variance has reduced.

b Take a larger sample before committing to new component.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 1

Question:

A random sample of 14 observations is taken from a normal distribution. The sample has a mean $\bar{x} = 30.4$ and a sample variance $s^2 = 36$.
It is suggested that the population mean is 28. Test this hypothesis at the 5% level of significance.

Solution:

$H_0 : \mu = 28 \quad H_1 : \mu \neq 28$

Critical region < -2.160 or > 2.160

Test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{30.4 - 28}{\frac{6}{\sqrt{14}}} = 1.4967$

The test statistic is not in the critical region

There is no evidence to suggest that μ does not equal 28

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 2

Question:

A random sample of 8 observations is taken from a random variable X that is normally distributed. The sample gave the following summary statistics

$\sum x^2 = 970.25 \quad \sum x = 85$

The population mean is thought to be 10. Test this hypothesis against the alternative hypothesis that the mean is greater than 10. Use the 5% level of significance.

Solution:

$H_0 : \mu = 10 \quad H_1 : \mu > 10$

Critical Region > 1.895

$$\bar{x} = \frac{85}{8} = 10.625$$

$$s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{970.25 - 8 \times 10.625^2}{7} = 9.589\ldots$$

$$\text{Test statistic} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{10.625 - 10}{\sqrt{\frac{9.589..}{8}}} = 0.571$$

not critical – no evidence to suggest that $\mu > 10$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 3

Question:

Six eggs selected at random from the daily output of a battery of hens had the following weights in grams.

55, 50, 53, 53, 52, 54

Calculate 95% confidence intervals for

a the mean,

b the variance of the population from which these eggs were taken.

c What assumption have you made about the distribution of the weights of eggs?

Solution:

$\bar{x} = 52.833\ldots \quad s = 1.722\ldots$

a confidence interval $= \left(\bar{x} - t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}, \bar{x} + t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}\right)$

$= \left(52.833\ldots - 2.571 \times \frac{1.722\ldots}{\sqrt{6}}, 52.833\ldots + 2.571 \times \frac{1.722\ldots}{\sqrt{6}}\right)$

$= (51.025, 54.641)$

b confidence interval $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right)}\right) = \left(\frac{5 \times 1.722\ldots^2}{12.832}, \frac{5 \times 1.722\ldots^2}{0.831}\right)$

$= (1.156, 17.850)$

c They are normally distributed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 4

Question:

A sample of size 18 was taken from a random variable X which was normally distributed, producing the following summary statistics.

$\bar{x} = 9.8 \quad s^2 = 0.49$

Calculate 95% confidence intervals for

a the mean,

b the variance of the population.

Solution:

a confidence interval $= \left(\bar{x} - t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}, \bar{x} + t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} \right)$

$= \left(9.8 - 2.110 \times \frac{0.7}{\sqrt{18}}, 9.8 + 2.110 \times \frac{0.7}{\sqrt{18}} \right)$

$= (9.451, 10.148)$

b confidence interval $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right)} \right) = \left(\frac{17 \times 0.49}{30.191}, \frac{17 \times 0.49}{7.564} \right)$

$= (0.276, 1.101)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 5

Question:

A random sample of 14 observations was taken of a random variable X which was normally distributed. The sample had a mean $\bar{x} = 23.8$, and a variance $s^2 = 1.8$.
Calculate,

a a 95% confidence interval for the variance of the population,

b a 90% confidence interval for the variance of the population.

Solution:

a confidence interval $= \left(\dfrac{(n-1)s^2}{\chi^2_{n-1}\left(\dfrac{\alpha}{2}\right)}, \dfrac{(n-1)s^2}{\chi^2_{n-1}\left(1-\dfrac{\alpha}{2}\right)} \right) = \left(\dfrac{13 \times 1.8}{24.736}, \dfrac{13 \times 1.8}{5.009} \right)$

$= (0.946, 4.672)$

b confidence interval $= \left(\dfrac{(n-1)s^2}{\chi^2_{n-1}\left(\dfrac{\alpha}{2}\right)}, \dfrac{(n-1)s^2}{\chi^2_{n-1}\left(1-\dfrac{\alpha}{2}\right)} \right) = \left(\dfrac{13 \times 1.8}{22.362}, \dfrac{13 \times 1.8}{5.892} \right)$

$= (1.046, 3.971)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 6

Question:

A manufacturer claims that the lifetime of its batteries is normally distributed with mean 21.5 hours. A laboratory tests 8 batteries and finds the lifetimes of these batteries to be as follows:

19.7 18.4 22.2 20.8 16.9 25.3 23.2 21.1

Stating clearly your hypotheses, examine whether or not these lifetimes indicate that the batteries have a shorter mean lifetime than that claimed by the company. Use a 5% level of significance. **[E]**

Solution:

$\bar{x} = 20.95 \quad s = 2.674\ldots$

$H_0 : \mu = 21.5 \quad H_1 : \mu < 21.5$

critical region < -1.895

test statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{20.95 - 21.5}{\frac{2.674\ldots}{\sqrt{8}}} = -0.5817$

The test statistic is not in the critical region

There is no evidence to reject claim.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 7

Question:

A diabetic patient monitors his blood glucose in mmol/l at random times of the day over several days. The following is a random sample of the results for this patient.

5.1 5.8 6.1 6.8 6.2 5.1 6.3 6.6 6.1 7.9 5.8 6.5

Assuming the data to be normally distributed, calculate a 95% confidence interval for

a the mean of the population of blood glucose readings,

b the standard deviation of the population of blood glucose readings.

The level of blood glucose varies throughout the day according to the consumption of food and the amount of exercise taken during the day.

c Comment on the suitability of the patient's method of data collection. **[E]**

Solution:

$\bar{x} = 6.1916\ldots\ s = 0.7549\ldots s^2 = 0.5699\ldots$

a confidence interval $= \left(\bar{x} - t_{(n-1)}\left(\frac{\alpha}{2}\right)\times\frac{s}{\sqrt{n}}, \bar{x} + t_{(n-1)}\left(\frac{\alpha}{2}\right)\times\frac{s}{\sqrt{n}}\right)$

$= \left(6.1916\ldots - 2.201\times\frac{0.7549\ldots}{\sqrt{12}}, 6.1916\ldots + 2.201\times\frac{0.7549\ldots}{\sqrt{12}}\right)$

$= (5.712, 6.671)$

b confidence interval Var. $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right)$

$= \left(\frac{11\times 0.5699\ldots}{21.920}, \frac{11\times 0.5699\ldots}{3.816}\right)$

$= (0.286, 1.643)$

confidence interval s.d. $= (0.535, 1.28)$

c He should measure his blood glucose at the same time each day.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 8

Question:

A woollen mill produces scarves. The mill has several machines each operated by a different person. Jane has recently started working at the mill and the supervisor wishes to check the lengths of the scarves Jane is producing. A random sample of 20 scarves is taken and the length, x cm, of each scarf is recorded. The results are summarised as.

$\sum x = 1428, \quad \sum x^2 = 102\,286$

Assuming that the lengths of scarves produced by any individual follow a normal distribution,

a calculate a 95% confidence interval for the variance σ^2 of the lengths of scarves produced by Jane.

The mill's owners require that 90% of scarves should be within 10 cm of the mean length.

b Find the value of σ that would satisfy this condition.

c Explain whether or not the supervisor should be concerned about the scarves Jane is producing. **[E]**

Solution:

$\bar{x} = \frac{1428}{20} = 71.4 \quad s^2 = \frac{102\ 286 - 20 \times 71.4^2}{19} = 17.2$

a confidence interval $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)} \right) = \left(\frac{19 \times 17.2}{32.852}, \frac{19 \times 17.2}{8.907} \right)$

$= (9.948, 36.69)$

b $10 = 1.6449 \times \sigma$ so $\sigma = \frac{10}{1.6449} = 6.079$

c $\sqrt{36.69} < 6.079$ so the supervisor should not be concerned.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 9

Question:

In order to discover the possible error in using a stop-watch, a student started the watch and stopped it again as quickly as she could. The times taken in centiseconds for 6 such attempts are recorded below:

10, 13, 14, 10, 13, 9

Assuming that the times are normally distributed, find 95% confidence limits for

a the mean,

b the variance. **[E]**

Solution:

$\bar{x} = 11.5 \qquad s = 2.073\ldots$

a confidence interval $= \left(\bar{x} - t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}, \bar{x} + t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}} \right)$

$= \left(11.5 - 2.571 \times \frac{2.073\ldots}{\sqrt{6}}, 11.5 + 2.571 \times \frac{2.073\ldots}{\sqrt{6}} \right)$

$= (9.324, 13.675)$

b confidence interval $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right)} \right) = \left(\frac{5 \times 2.073\ldots^2}{12.832}, \frac{5 \times 2.073\ldots^2}{0.831} \right)$

$= (1.675, 25.872)$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 10

Question:

A manufacturer claims that the car batteries which it produces have a mean lifetime of 24 months, with a standard deviation of 4 months. A garage selling the batteries doubts this claim and suggests that both values are in fact higher.

The garage monitors the lifetimes of 10 randomly selected batteries and finds that they have a mean lifetime of 27.2 months and a standard deviation of 5.2 months.

Stating clearly your hypotheses and using a 5% level of significance, test the claim made by the manufacturer for

a the standard deviation,

b the mean,

c State an assumption which has to be made when carrying out these tests. **[E]**

Solution:

a $H_0 : \sigma = 4 \quad H_1 : \sigma > 4$

Critical region $\chi^2 > 16.919$

Test statistic $= \frac{(n-1)s^2}{\sigma^2} = \frac{9 \times 5.2^2}{4^2} = 15.21$

The test statistic is not in the critical region.

standard deviation could be 4 months

b $H_0 : \mu = 24 \quad H_1 : \mu > 24$

Critical region $t > 1.833$

Test statistic $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{27.2 - 24}{\frac{5.2}{\sqrt{10}}} = 1.946$

There is evidence to reject H_0

The mean battery life is greater than 24 months

c Lifetime is normally distributed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 11

Question:

The distance to 'take-off' from a standing start of an aircraft was measured on twenty occasions. The results are summarised in the following table.

Distance (m)	Frequency
700–	3
710–	5
720–	9
730–	2
740–750	1

Assuming that distance to 'take-off' is normally distributed, find 95% confidence intervals for

a the mean,

b the standard deviation.

It has been hypothesised that the mean distance to 'take-off' is 725 m.

c Comment on this hypothesis in the light of your interval from part **a**. **[E]**

Solution:

$\bar{x} = 721.5 \quad s = 10.399\ldots$

a confidence interval $= \left(\bar{x} - t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}, \bar{x} + t_{(n-1)}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}\right)$

$= \left(721.5 - 2.093 \times \frac{10.399\ldots.}{\sqrt{20}}, 721.5 + 2.093 \times \frac{10.399\ldots}{\sqrt{20}}\right)$

$= (717, 726)$

b confidence interval variance $= \left(\frac{(n-1)s^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)}, \frac{(n-1)s^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right)$

$= \left(\frac{19 \times 10.399\ldots^2}{32.852}, \frac{19 \times 10.399\ldots^2}{8.907}\right)$

$= (62.553, 230.717)$

confidence interval standard deviation $= (7.909, 15.189)$

c 725 within confidence interval,
There is no evidence to reject this hypothesis.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 12

Question:

The maximum weight that 50 cm lengths of a certain make of string can hold before breaking (the breaking strain) has a normal distribution with mean 40 kg and standard deviation 5 kg. The manufacturer of the string has developed a new process which should increase the mean breaking strain of the string but should not alter the standard deviation. Ten randomly selected pieces of string are tested and their breaking strains, in kg, are:

51, 48, 37, 46, 36, 53, 34, 49, 47, 50

a Stating your hypotheses clearly test, at the 5% level of significance, whether or not the new process has altered the variance.

In the light of your conclusion to the test in part **a**,

b test whether or not there is evidence that the new process has increased the mean breaking strain. State your hypotheses clearly and use a 5% level of significance.

c Explain briefly your choice of test in part **b**. **[E]**

Solution:

$\bar{x} = 45.1 \quad s = 6.838\ldots$

$H_0 : \sigma = 5 \quad H_1 : \sigma \neq 5$

a Critical region > 19.023 and < 2.700

Test statistic $= \frac{(n-1)s^2}{\sigma^2} = \frac{9 \times 6.838\ldots^2}{5^2} = 16.836$

There is insufficient evidence to reject H_0

$\sigma = 5$

b Critical region $z > 1.6449$

Test statistic $z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{45.1 - 40}{\frac{5}{\sqrt{10}}} = 3.225$

The test statistic is in the critical region.

There is evidence to suggest there is an increase in breaking strain.

c In **a** there was no change in σ so assume $\sigma = 5 \therefore$ use z not t

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 13

Question:

A company knows from previous experience that the time taken by maintenance engineers to repair a particular electrical fault on a complex piece of electrical equipment is 3.5 hours on average with a standard deviation of 0.5 hours.

A new method of repair has been devised, but before converting to this new method the company took a random sample of 10 of its engineers and each engineer carried out a repair using the new method. The time, x hours, it took each of them to carry out the repair was recorded and the data are summarised below:

$\sum x = 34.2 \quad \sum x^2 = 121.6$

Assume that the data can be regarded as a random sample from a normal population.

a For the new repair method, calculate an unbiased estimate of the variance.

b Use your estimate from **a** to calculate for the new repair method a 95% confidence interval for

  i the mean,

  ii the standard deviation.

c Use your calculations and the given data to compare the two repair methods in order to advise the company as to which method to use.

d Suggest an alternative way of comparing the two methods of repair using the 10 randomly chosen engineers. **[E]**

Solution:

a $\bar{x} = \frac{34.2}{10} = 3.42 \quad s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{121.6 - 10 \times 3.42^2}{9} = 0.5151\ldots$

b i confidence interval mean $= \left(\bar{x} - t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{s}{\sqrt{n}}\right)$

$= (3.42 - 2.262 \times \frac{0.7177\ldots}{\sqrt{10}}, 3.42 + 2.262 \times \frac{0.7177\ldots}{\sqrt{10}})$

$= (2.906, 3.933)$

ii confidence interval variance $= \left(\frac{(n-1)s^2}{\chi_9^2(0.025)}, \frac{(n-1)s^2}{\chi_9^2(0.975)}\right)$

$= \left(\frac{9 \times 0.515\ldots}{19.023}, \frac{9 \times 0.515\ldots}{2.700}\right) = (0.244, 1.717)$

confidence interval standard deviation $= (0.4937, 1.3103)$

c 3.5 hours is inside the confidence interval on the mean, so there is no evidence of a change in the mean time.
0.5 hours is inside the confidence interval on the standard deviation so there is no evidence of a change in the variability of the time.
There is no reason to change the repair method.

d Use a 'matched pairs' experiment, getting each engineer to carry out a similar repair using the old method and the new method and use a paired t-test.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 1

Question:

Find the upper 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 12, b = 18$,

b $a = 4, b = 11$,

c $a = 6, b = 9$.

Solution:

a 2.34
b 3.36
c 3.37

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 2

Question:

Find the lower 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 6, b = 8$,

b $a = 25, b = 12$,

c $a = 5, b = 5$.

Solution:

a $\frac{1}{F_{8,6}} = 0.241$

b $\frac{1}{F_{12,25}} = 0.463$

c $\frac{1}{F_{5,5}} = 0.198$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 3

Question:

Find the upper 1% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 12, b = 18$,

b $a = 6, b = 16$,

c $a = 5, b = 9$.

Solution:

a 3.37

b 4.20

c 6.06

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 4

Question:

Find the lower 1% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 3, b = 12$,

b $a = 8, b = 12$,

c $a = 5, b = 12$.

Solution:

a $\frac{1}{F_{12,3}} = 0.0370$

b $\frac{1}{F_{12,8}} = 0.176$

c $\frac{1}{F_{12,5}} = 0.101$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 5

Question:

Find the lower and upper 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 8, b = 10$,

b $a = 12, b = 10$,

c $a = 3, b = 5$.

Solution:

a $3.07, \frac{1}{F_{10,8}} = 0.299$

b $2.91, \frac{1}{F_{10,12}} = 0.364$

c $5.41, \frac{1}{F_{5,3}} 0.111$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 6

Question:

The random variable X follows a $F_{40,12}$-distribution.
Find $P(X < 0.5)$.

Solution:

$$P(X < 0.5) = P(F_{40,12} < 0.5)$$
$$= P(F_{12,40} > \frac{1}{0.5})$$
$$= P(F_{12,40} > 2)$$

From the tables $F_{12,40}(0.05) = 2$

$\therefore P(F_{12,40} > 2) = P(F_{40,12} < 0.5) = 0.05$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 7

Question:

The random variable X follows a $F_{12,8}$-distribution.

Find $\mathrm{P}\left(\frac{1}{2.85} < X < 3.28\right)$

Solution:

$$\begin{aligned}\mathrm{P}(X < 3.28) &= 1 - \mathrm{P}(F_{12,8} > 3.28)\\ &= 1 - 0.05 = 0.95\end{aligned}$$

$$\begin{aligned}\mathrm{P}\left(X < \frac{1}{2.85}\right) &= \mathrm{P}\left(F_{12,8} < \frac{1}{2.85}\right)\\ &= \mathrm{P}(F_{8,12} > 2.85)\end{aligned}$$

$$\therefore \mathrm{P}\left(X < \frac{1}{2.85}\right) = 0.05$$

$$\begin{aligned}\mathrm{P}\left(\frac{1}{2.85} < X < 3.28\right) &= \mathrm{P}(X < 3.28) - \mathrm{P}\left(X < \frac{1}{2.85}\right)\\ &= 0.95 - 0.05\\ &= 0.90\end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 8

Question:

The random variable X has an F-distribution with 2 and 7 degrees of freedom.
Find $P(X < 9.55)$. **[E]**

Solution:

$$\begin{aligned} P(X < 9.55) &= 1 - P(F_{2,7} > 9.55) \\ &= 1 - 0.01 \\ &= 0.99 \end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 9

Question:

The random variable X follows an F-distribution with 6 and 12 degrees of freedom.

a Show that $P(0.25 < X < 3.00) = 0.9$.

A large number of values are randomly selected from an F-distribution with 6 and 12 degrees of freedom.

b Find the probability that the seventh value to be selected will be the third value to lie between 0.25 and 3.00. **[E]**

Solution:

a
$$\begin{aligned} P(X < 3.00) &= 1 - P(F_{6,12} > 3.00) \\ &= 1 - 0.05 = 0.95 \\ P(X > 0.25) &= P(F_{12,6} > \frac{1}{0.25}) \\ &= P(F_{12,6} > 4) \\ &= 0.95 \\ P(0.25 < X < 3) &= 0.95 - 0.05 \\ &= 0.90 \end{aligned}$$

b $^6C_2 (0.9)^2 (0.1)^4 \times 0.9 = 0.00109$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 1

Question:

Random samples are taken from two normally distributed populations. There are 11 observations from the first population and the best estimate for the population variance is $s^2 = 7.6$. There are 7 observations from the second population and the best estimate for the population variance is $s^2 = 6.4$.
Test, at the 5% level of significance, the hypothesis $H_0 : \sigma_1{}^2 = \sigma_2{}^2$ against the alternative hypothesis $H_1 : \sigma_1{}^2 > \sigma_2{}^2$.

Solution:

Critical value is $F_{10,6} = 4.06$

$F_{\text{test}} = \frac{7.6}{6.4} = 1.1875$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 2

Question:

Random samples are taken from two normally distributed populations. There are 25 observations from the first population and the best estimate for the population variance is $s^2 = 0.42$. There are 41 observations from the second population and the best estimate for the population variance is $s^2 = 0.17$.

Test, at the 1% significance level, the hypothesis $H_0 : \sigma_1{}^2 = \sigma_2{}^2$ against the alternative hypothesis $H_1 : \sigma_1{}^2 > \sigma_2{}^2$.

Solution:

Critical value is $F_{24,40} = 2.29$

$F_{\text{test}} = \dfrac{0.42}{0.17} = 2.4706$

In critical region

reject H_0 – there is evidence to suggest that $\sigma_1^2 > \sigma_2^2$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 3

Question:

The variance of the lengths of a sample of 9 tent-poles produced by a machine was 63 mm^2. A second machine produced a sample of 13 tent-poles with a variance of 225 mm^2. Both these values are unbiased estimates of the population variances.

a Test, at the 10 % level, whether there is evidence that the machines differ in variability, stating the null and alternative hypotheses.

b State the assumption you have made about the distribution of the populations in order to carry out the test in part **a**. **[E]**

Solution:

a $H_0: \sigma_1^2 = \sigma_2^2 \quad H_1: \sigma_1^2 \neq \sigma_2^2$

Critical value is $F_{12,8} = 3.28$

$$F_{\text{test}} = \frac{225}{63} = 3.57$$

In critical region

reject H_0 – There is evidence to suggest that the machines differ in variability

b Population distributions are assumed to be normal

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 4

Question:

Random samples are taken from two normally distributed populations. The size of the sample from the first population is $n_1 = 13$ and this gives an unbiased estimate for the population variance $s_1^2 = 36.4$. The figures for the second population are $n_2 = 9$ and $s_2^2 = 52.6$.

Test, at the 5% significance level, whether $\sigma_1^2 = \sigma_2^2$ or if $\sigma_1^2 > \sigma_2^2$.

Solution:

Critical value is $F_{8,12} = 2.85$

$F_{\text{test}} = \frac{52.6}{36.4} = 1.445$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 5

Question:

Dining chairs Ltd are in the process of selecting a make of glue for using on the joints of their furniture. There are two possible contenders – Goodstick which is the more expensive, and Holdtight, the cheaper of the two.

The company are concerned that, while both glues are said to have the same adhesive power, one might be more variable than the other.

A series of trials are carried out with each glue and the joints tested to destruction. The force in newtons at which each joint failed is recorded. The results are as follows:

Goodstick	10.3	8.2	9.5	9.9	11.4	
Holdtight	9.6	10.8	9.9	10.8	10.0	10.2

a Test, at the 10% significance level, whether or not the variances are equal.
b Which glue would you recommend and why?

Solution:

a $\sigma^2_{\text{goodstick}} = 1.363$

$\sigma^2_{\text{Holdtight}} = 0.24167$

Critical value is $F_{4,5} = 5.19$

$F_{\text{test}} = \dfrac{1.363}{0.24167} = 5.64$

In critical region

reject H_0 – there is evidence to suggest that the variances are not equal.

b Holdtight as it is less variable and cheaper.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 6

Question:

The closing balances, £x, of a number of randomly chosen bank current accounts of two different types, Chegrit and Dicabalk, are analysed by a statistician. The summary statistics are given in the table below.

	Sample size	$\sum x$	$\sum x^2$
Chegrit	7	276	143 742
Dicaballk	15	394	102 341

Stating clearly your hypotheses test, at the 10% significance level, whether or not the two distributions have the same variance. (You may assume that the closing balances of each type of account are normally distributed.) **[E]**

Solution:

$\sigma^2_{\text{Chegrit}} = 22\,143.286$

$\sigma^2_{\text{Dicabalk}} = 6570.85238$

Critical value is $F_{6,14} = 2.85$

$$F_{\text{test}} = \frac{22143.286}{6570.85238} = 3.3699$$

In critical region − there is evidence to suggest that their variances differ

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 7

Question:

Bigborough council wishes to change the bulbs in their traffic lights at regular intervals so that there is a very small probability that any light bulb will fail in service. The council are anxious that the length of time between changes should be as long as possible, and to this end they have obtained a sample of bulbs from another manufacturer, who claims the same bulb life as their present manufacturer. The council wishes therefore to select the manufacturer whose bulbs have the smallest variance.

When they last tested a random sample of 9 bulbs from their present supplier the summary results were $\sum x = 9415$ hours, $\sum x^2 = 9\,863\,681$, where x represents the lifetime of a bulb.

A random sample of 8 bulbs from the prospective new supplier gave the following bulb lifetimes in hours: 1002, 1018, 943, 1030, 984, 963, 1048, 994.

a Calculate unbiased estimates for the means and variances of the two populations.

Assuming that the lifetimes of bulbs are normally distributed,

b test, at the 10% significance level, whether or not the two variances are equal.

c State your recommendation to the council, giving reasons for your choice.

Solution:

a $\mu_1 = 1046$ $\quad s_1^2 = 1818.11$ and $\mu_2 = 997.75$ $\quad s_2^2 = 1200.21$

b Critical value is $F_{8,7} = 3.73$

$$F_{\text{test}} = \frac{1818.111}{1200.21} = 1.5148$$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

c Use present supplier who appears to have a higher mean.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 1

Question:

A random sample of 10 toothed winkles was taken from a sheltered shore, and a sample of 15 was taken from a non-sheltered shore. The maximum basal width, (x mm), of the shells was measured and the results are summarised below.

Sheltered shore: $\bar{x} = 25$, $s^2 = 4$. Non-sheltered shore: $\bar{x} = 22$, $s^2 = 5.3$.

a Find a 95% confidence interval for the difference between the means.

b State an assumption that you have made when calculating this interval.

Solution:

a

$$s_p^2 = \frac{(9\times4)+(14\times5.3)}{10+15-2} = 4.7913$$

$$S_p = 2.189$$

$$t_{23}(2.5\%) = 2.069$$

$$(25-22) \pm 2.069\times2.189\sqrt{\frac{1}{15}+\frac{1}{10}} = 3\pm1.849$$

$$= (1.151, 4.849)$$

b Independent random samples, normal distributions, common variance

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 2

Question:

A packet of plant seeds was sown and, when the seeds had germinated and begun to grow, 8 were transferred into pots containing a soil-less compost and 10 were grown on in a soil-based compost. After 6 weeks of growth the heights, x, in cm of the plants were measured with the following results:
Soil-less compost: 9.3, 8.7, 7.8, 10.0, 9.2, 9.5, 7.9, 8.9.
Soil-based compost: 12.8, 13.1, 11.2, 10.1, 13.1, 12.0, 12.5, 11.7, 11.9, 12.0.
Assuming that the populations are normally distributed, and that there is a difference between the two means calculate a 90% confidence interval for this difference.

Solution:

$$\bar{x}_s = 8.9125, \quad s_s^2 = 0.58125$$
$$\bar{x}_{ns} = 12.04 \quad s_{ns}^2 = 0.84933$$
$$s_p^2 = \frac{(7\times 0.58125)+(9\times 0.84933)}{10+8-2} = 0.7319$$
$$s_p = 0.855$$
$$t_{16}(5\%) = 1.746$$
$$(12.04-8.9125)\pm 1.746\times 0.855\sqrt{\frac{1}{10}+\frac{1}{8}} = 3.1275\pm 0.7081$$
$$= (2.419, 3.836)$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 3

Question:

Forty children were randomly selected from all 12-year-old children in a large city to compare two methods of teaching the spelling of 50 words which were likely to be unfamiliar to the children. Twenty children were randomly allocated to each method. Six weeks later the children were tested to see how many of the words they could spell correctly. The summary statistics for the two methods are given in the table below, where $\bar{x}$ is the mean number of words spelt correctly, s^2 is an unbiased estimate of the variance of the number of words spelt correctly and n is the number of children taught using each method.

	$\bar{x}$	s^2	n
Method A	32.7	6.1^2	20
Method B	38.2	5.2^2	20

a Calculate a 99% confidence interval for the difference between the mean numbers of words spelt correctly by children who used Method B and Method A.

b State two assumptions you have made in carrying out part **a**.

c Interpret your result.

Solution:

a

$$s_p^2 = \frac{(19\times 6.1^2)+(19\times 5.2^2)}{20+20-2} = 32.125$$

$$s_p = 5.66789$$

$$t_{38}(0.5\%) = 2.712$$

$$(38.2-32.7)\pm 2.712\times 5.66789\sqrt{\frac{1}{20}+\frac{1}{20}} = 5.5\pm 4.86083$$

$$= (0.6392, 10.3608)$$

b normality and equal variances

c zero not in interval $\Rightarrow$ method B seems better than method A

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 4

Question:

The table below shows summary statistics for the mean daily consumption of cigarettes by a random sample of 10 smokers before and after their attendence at an anti-smoking workshop with $\bar{x}$ representing the means and s^2 representing the unbiased estimates of population variance in each case.

	$\bar{x}$	s^2	n
Mean daily consumption before the workshop	18.6	32.488	10
Mean daily consumption after the workshop	14.3	33.344	10

Stating clearly any assumption you make, calculate a 90% confidence interval for the difference in the mean daily consumption of cigarettes before and after the workshop.

Solution:

Assume same variances and that the population of differences is normally distributed

$$s_p^2 = \frac{(9 \times 32.488) + (9 \times 33.344)}{10 + 10 - 2} = 32.916$$

$$s_p = 5.73725$$

$$t_{18}(5\%) = 1.734$$

$$(18.6 - 14.3) \pm 1.734 \times 5.73725 \sqrt{\frac{1}{10} + \frac{1}{10}} = 4.3 \pm 4.44905$$

$$= (-0.149, 8.749)$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 1

Question:

A random sample of size 20 from a normal population gave $\bar{x} = 16, s^2 = 12$.

A second random sample of size 11 from a normal population gave $\bar{x} = 14, s^2 = 12$.

a Assuming that the both populations have the same variance, find an unbiased estimate for that variance.

b Test, at the 5% level of significance, the suggestion that the two populations have the same mean.

Solution:

a $s_p^2 = \dfrac{(19 \times 12) + (10 \times 12)}{20 + 11 - 2} = 12$ so $s_p = \sqrt{12} = 3.464$

b $H_0 : \mu_{1st} = \mu_{2nd}$ $H_1 : \mu_{1st} \neq \mu_{2nd}$

critical value $t_{29}(0.025) = 2.045$

critical region is $t \leq -2.045$ and $t \geq 2.045$

$$t = \frac{(16 - 14) - 0}{3.464\sqrt{\frac{1}{20} + \frac{1}{11}}} = 1.538$$

Not in critical region – do not reject H_0

There is evidence to suggest that the populations have the same mean

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 2

Question:

Salmon reared in Scottish fish farms are generally larger than wild salmon.
A fisherman measured the length of the first 6 salmon caught on his boat at a fish farm. Their lengths in centimetres were

42.8, 40.0, 38.2, 37.5, 37.0, 36.5.

Chefs prefer wild salmon to fish-farmed salmon because of their better flavour. A chef was offered 4 salmon that were claimed to be wild. Their lengths in centimetres were

42.0, 43.0, 41.5, 40.0.

Use the information given above and a suitable t-test at the 5% level of significance to help the chef to decide if the claim is likely to be correct. You may assume that the populations are normally distributed.

Solution:

$\mathrm{H_0}: \mu_c = \mu_F \quad \mathrm{H_1}: \mu_c > \mu_F$

$\bar{x}_F = 38.67, \quad s_F^2 = 5.5827$

$\bar{x}_c = 41.625 \quad s_c^2 = 1.5625$

$$s_p^2 = \frac{(5\times 5.5827)+(3\times 1.5625)}{6+4-2} = 4.075 \quad \text{so } s_p = \sqrt{4.075} = 2.0187$$

critical value $t_8(0.05) = 1.86$

critical region $t \geq 1.860$

$$t = \frac{(41.625-38.67)-0}{2.0187\sqrt{\frac{1}{6}+\frac{1}{4}}} = 2.270$$

In the critical region − reject $\mathrm{H_0}$ there is evidence to suggest that the salmon are wild

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 3

Question:

In order to check the effectiveness of three drugs against the E. *coli* bacillus, 15 cultures of the bacillus (5 for each of 3 different antibiotics) had discs soaked in the antibiotics placed in their centre. The 15 cultures were left for a time and the area in cm^2 per microgram of drug where the E. *coli* was killed was measured. The results for three different drugs are given below:

Streptomycin	0.210, 0.252, 0.251, 0.210, 0.256, 0.253
Tetracycline	0.123, 0.090, 0.123, 0.141, 0.142, 0.092
Erythromycin	0.134, 0.120, 0.123, 0.210, 0.134, 0.134

a It was thought that Tetracycline and Erythromycin seemed equally as effective. Assuming that the populations are normally distributed, test this at the 5% significance level.

b Streptomycin was thought to be more effective than either of the others. Treating the other 2 as being a single sample of 12, test this assertion at the same level of significance.

Solution:

a $\bar{x}_t = 0.1185, \quad s_t^2 = 0.0005227.$

$\bar{x}_e = 0.1425 \quad s_e^2 = 0.0011319.$

$$s_p^2 = \frac{(5\times 0.0005227)+(5\times 0.0011319)}{6+6-2} = 0.000827 \text{ so } s_p = \sqrt{0.000827} = 0.02876$$

$\text{H}_0 : \mu_t = \mu_e \quad \text{H}_1 : \mu_t \neq \mu_e$

critical value $t_{10}(0.025) = 2.228$

critical region $t \leq -2.228$ and $t \geq 2.228$

$$t = \frac{(0.1425 - 0.1185) - 0}{0.02876\sqrt{\frac{1}{6}+\frac{1}{6}}} = 1.445$$

not in the critical region – accept H_0; there is evidence to suggest that Tetracycline and Erythromycin are equally as effective

b $\bar{x}_s = 0.2387, \quad s_s^2 = 0.0004959$

$\bar{x}_2 = 0.1305, \quad s_2^2 = 0.000909$

$$s_p^2 = \frac{(11\times 0.000909)+(5\times 0.0004959)}{12+6-2} = 0.000780, \text{ so } s_p = \sqrt{0.000780} = 0.0279$$

$\text{H}_0 : \mu_s = \mu_2 \quad \text{H}_1 : \mu_s > \mu_2$

critical value $t_{16}(0.05) = 1.746$

critical region $t \geq 1.746$

$$t = \frac{(0.2387 - 0.1305) - 0}{0.0279\sqrt{\frac{1}{12}+\frac{1}{6}}} = 7.75$$

in the critical region – reject H_0. There is evidence to suggest that Streptomycin is more effective than the others.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 4

Question:

To test whether a new version of a computer programming language enabled faster task completion, the same task was performed by 16 programmers, divided at random into two groups. The first group used the new version of the language, and the time for task completion, in hours, for each programmer was as follows:

4.9 6.3 9.6 5.2 4.1 7.2 4.0.

The second group used the old version, and their times were summarised as follows:

$n = 9, \sum x = 71.2, \sum x^2 = 604.92$.

a State the null and alternative hypotheses.

b Perform an appropriate test at the 5% level of significance.

In order to compare like with like, experiments such as this are often performed using the same individuals in the first and the second groups.

c Give a reason why this strategy would not be appropriate in this case. **[E]**

Solution:

a $H_0 : \mu_{old} = \mu_{new}$, $H_1 : \mu_{old} > \mu_{new}$

b $\bar{x}_{old} = 7.911 \quad s^2_{old} = 5.206$

$\bar{x}_{new} = 5.9, \quad s^2_{new} = 3.98$

$$s_p{}^2 = \frac{(6 \times 3.98) + (8 \times 5.206)}{7 + 9 - 2} = 4.6806 \quad \text{so } s_p = \sqrt{4.6806} = 2.1635$$

Critical value $t_{14} = 1.761$

critical region $t \geq 1.761$

$$\text{Test statistic } t = \frac{(7.911 - 5.9) - 0}{2.1635\sqrt{\frac{1}{9} + \frac{1}{7}}} = 1.84446$$

Significant – there is evidence to suggest that new language does improve time.

c Once task is solved the programmer should be quicker next time with either language.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 5

Question:

A company undertakes investigations to compare the fuel consumption, x, in miles per gallon, of two different cars, the Volcera and the Spintono, with a view to purchasing a number as company cars.

For a random sample of 12 Volceras the fuel consumption is summarised by

$\sum V = 384$ and $\sum V^2 = 12\,480$.

A statistician incorrectly combines the figures for the sample of 12 Volceras with those of a random sample of 15 Spintonos, then carries out calculations as if they are all one larger sample and obtains the results $\bar{y} = 34$ and $s^2 = 23$.

a Show that, for the sample of 15 Spintonos, $\sum x = 534$ and $\sum x^2 = 19\,330$.

Given that the variance of the fuel consumption for each make of car is σ^2

b obtain an unbiased estimate for σ^2.

c Test, at the 5% level of signifance, whether there is a difference between the mean fuel consumption of the two models of car. State your hypothese and conclusion clearly.

d State any further assumption you made in order to be able to carry out your test in c.

e Give two precautions which could be taken when undertaking an investigation into the fuel consumption of two models of car to ensure that a fair comparison is made.

[E]

Solution:

a $27 \times 34 - 384 = 534 = \sum x$

$$23 = \frac{\sum y^2}{27-1} - \frac{918^2}{27(27-1)}$$

$384 + 534 = 918$

$\sum y^2 = 31\,810$

$31\,810 - 12\,480 = 19\,330 = \sum x^2$

b $\bar{x}_v = 32 \quad s_v^2 = 17.45$

$\bar{x}_s = 35.6, \quad s_s^2 = 22.829$

$$s_p^2 = \frac{(14 \times 22.829) + (11 \times 17.45)}{12 + 15 - 2} = 20.464$$

c $H_0 : \mu_v = \mu_s, \quad H_1 : \mu_v \neq \mu_s$

Critical value $t_{25}(0.025) = 2.060$

Critical region $t \leq -2.060$ and $t \leq 2.060$

$$t = \frac{(35.6 - 32) - 0}{4.524\sqrt{\frac{1}{15} + \frac{1}{12}}} = 2.0547$$ – accept H_0 – no evidence to suggest difference in means

d normality

e same types of driving, roads and weather

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 1

Question:

It is claimed that completion of a shorthand course has increased the shorthand speeds of the students.

a If the suggestion that the mean speed of the students has not altered is to be tested, write down suitable hypotheses for which **i** a two-tailed test is appropriate, and **ii** a one-tailed test is appropriate.

The table below gives the shorthand speeds of students before and after the course.

Student	A	B	C	D	E	F
Speed before in words/minute	35	40	28	45	30	32
Speed after	42	45	28	45	40	40

b Carry out a paired t-test, at the 5% significance level, to determine whether or not there has been an increase in shorthand speeds.

Solution:

a **i** $H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

ii $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

b $\sum d = 30 \quad \sum d^2 = 238$

$\bar{d} = 5$

$s^2 = \frac{238 - 6(5)^2}{5} = 17.6$

$s = 4.195$

Critical value $t_5(5\%) = 2.015$

The critical region is $t > 2.015$

$t = \frac{5-0}{\frac{4.195}{\sqrt{6}}}$

$= 2.919$

In the critical region – reject H_0.

There is evidence to suggest that there has been an increase in shorthand speed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 2

Question:

A large number of students took two General Studies papers that were supposed to be of equal difficulty. The results for 10 students chosen at random are shown below:

Candidate	A	B	C	D	E	F	G	H	I	J
Paper 1	18	25	40	10	38	20	25	35	18	43
Paper 2	20	27	39	12	40	23	20	35	20	41

The teacher looked at the marks of a random sample of 10 students, and decided that paper 2 was easier than paper 1.
Given that the marks on each paper are normally distributed, carry out an appropriate test, at the 1% level of significance.

Solution:

$H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$\sum d = 5 \quad \sum d^2 = 59$

$\bar{d} = 0.5$

$$s^2 = \frac{59 - 10(0.5)^2}{9} = 6.278$$

$s = 2.50555$

Critical value $t_9(1\%) = 2.821$

The critical region is $t > 2.821$.

$$t = \frac{0.5 - 0}{\frac{2.50555}{\sqrt{10}}}$$

$= 0.631$

Not in the critical region. Do not reject H_0.

There is insufficient evidence to suggest that paper 2 is easier than paper 1 so the teacher is not correct.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 3

Question:

It is claimed by the manufacturer that by chewing a special flavoured chewing gum smokers are able to reduce their craving for cigarettes, and thus cut down on the number of cigarettes smoked per day. In a trial of the gum on a random selection of 10 people the no-gum smoking rate and the smoking rate when chewing the gum were investigated, with the following results:

Person	A	B	C	D	E	F	G	H	I	J
Without gum smoking rate cigs./day	20	35	40	32	45	15	22	30	34	40
With gum smoking rate cigs./day	15	25	35	30	45	15	14	25	28	34

a Use a paired t-test at the 5% significance level to test the manufacturer's claim.

b State any assumptions you have had to make.

Solution:

a $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$$\sum d = 47 \quad \sum d^2 = 315$$

$$\bar{d} = 4.7$$

$$s^2 = \frac{315 - 10(4.7)^2}{9} = 10.456$$

$$s = 3.234$$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$$t = \frac{4.7 - 0}{\frac{3.234}{\sqrt{10}}}$$

$$= 4.596$$

In the critical region. Reject H_0.

There is evidence to suggest that chewing the gum does reduce the craving for cigarettes.

b The differences are normally distributed

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 4

Question:

The council of Somewhere town are going to put a new traffic management scheme into operation in the hope that it will make travel to work in the mornings quicker for most people. Before the scheme is put into operation, 10 randomly selected workers are asked to record the time it takes them to come into work on a Wednesday morning. After the scheme is put into place, the same 10 workers are again asked to record the time it takes them to come into work on a particular Wednesday morning.
The times in minutes are shown in the table below:

Worker	A	B	C	D	E	F	G	H	I	J
Before	23	37	53	42	39	60	54	85	46	38
After	18	35	49	42	34	48	52	79	37	37

Test, at the 5% significance level, whether or not the journey time to work has decreased.

Solution:

$H_0 : \mu_D = 0, \quad H_1 : \mu_D > 0$

$\sum d = 46 \quad \sum d^2 = 336$

$\bar{d} = 4.6$

$s^2 = \frac{336 - 10(4.6)^2}{9} = 13.8222$

$s = 3.7178$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$t = \frac{4.6 - 0}{\frac{3.7178}{\sqrt{10}}}$

$= 3.913$

In the critical region. Reject H_0.

There is evidence to suggest that the journey times have decreased.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 5

Question:

A teacher is anxious to test the idea that students' results in mock examinations are good predictors for their results in actual examinations. He selects 8 students at random from those doing a mock Statistics examination and records their marks out of 100; later he collects the same students' marks in the actual examination. The resulting marks are as follows:

Student	A	B	C	D	E	F	G	H
Mock examination mark	35	86	70	91	45	64	78	38
Actual examination	45	77	81	86	53	71	68	46

a Use a paired t-test to investigate whether or not the mock examination is a good predictor. (Use a 10% significance level.)

b State any assumptions you have made.

Solution:

a $H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

$\sum d = 20 \quad \sum d^2 = 604$

$\bar{d} = 2.5$

$$s^2 = \frac{604 - 8(2.5)^2}{7} = 79.1429$$

$s = 8.896$

Critical value $t_7(5\%) = 1.895$

The critical regions are $t < -1.895$ and $t > 1.895$.

$$t = \frac{2.5 - 0}{\frac{8.896}{\sqrt{8}}}$$

$= 0.795$

Not in the critical region. Do not reject H_0.

The mock examination is a good predictor.

b The differences are normally distributed

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 6

Question:

The manager of a dress-making company took a random sample of 10 of his employees and recorded the number of dresses made by each. He discovered that the number of dresses made between 3.00 and 5.00 p.m. was fewer than the same employees achieved between 9.00 and 11.00 a.m. He wondered if a tea break from 2.45–3.00 p.m. would increase productivity during these last two hours of the day. The number of dresses made by these workers in the last two hours of the day before and after the introduction of the tea break were as shown below.

Worker	A	B	C	D	E	F	G	H	I	J
Before	75	73	75	81	74	73	77	75	75	72
After	80	84	79	84	85	84	78	78	80	83

a Why was the comparison made for the same ten workers?

b Conduct, at the 5% level of significance, a paired t-test to see if the introduction of a tea break has increased production between 3.00 and 5.00 p.m.

Solution:

a Different people will have different productivity rates. Need a common link if want to compare before and after. This reduces experimental error due to differences between individuals so that, if a difference does exist, it is more likely to be detected.

b $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$\sum d = 65 \quad \sum d^2 = 569$

$\bar{d} = 6.5$

$$s^2 = \frac{569 - 10(6.5)^2}{9} = 16.278$$

$s = 4.0346$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$$t = \frac{6.5 - 0}{\frac{4.0346}{\sqrt{10}}}$$

$= 5.095$

In the critical region. Reject H_0.

There is evidence to suggest a tea break increases the number of garments made.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 7

Question:

A drug administered in tablet form to help people sleep and a placebo was given for two weeks to a random sample of eight patients in a clinic. The drug and the placebo were given in random order for one week each. The average numbers of hours sleep that each patient had per night with the drug and with the placebo are given in the table below.

Patient	1	2	3	4	5	6	7	8
Hours of sleep with drug	10.5	6.7	8.9	6.7	9.2	10.9	11.9	7.6
Hours of sleep with placebo	10.3	6.5	9.0	5.3	8.7	7.5	9.3	7.2

Test, at the 1% level of significance, whether or not the drug increases the mean number of hours sleep per night. State your hypotheses clearly. **[E]**

Solution:

$H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$\sum d = 8.6 \sum d^2 = 20.78$

$\bar{d} = 1.075$

$s^2 = \frac{20.78 - 8(1.075)^2}{7} = 1.64786$

$s = 1.2837$

Critical value $t_7(1\%) = 2.998$

The critical region is $t > 2.998$

$t = \frac{1.075 - 0}{\frac{1.2837}{\sqrt{8}}}$

$= 2.3686$

Not in the critical region. Do not reject H_0.

There is no evidence to suggest that the drug increases the mean number of hours sleep per night.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 1

Question:

The random variable X has an F-distribution with 5 and 10 degrees of freedom. Find values of a and b such that $\mathrm{P}(a \le X \le b) = 0.90$ **[E]**

Solution:

$\mathrm{P}(F_{5,10} \ge 3.33) = 0.05 \Rightarrow b = 3.33$

$\mathrm{P}\left(F_{10,5} \ge 4.74\right) = 0.05 \Rightarrow \mathrm{P}\left(F_{5,10} \le \frac{1}{4.74}\right) = 0.05$

$\therefore a = 0.2110 \text{ (4 s.f.)}$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 2

Question:

A chemist has developed a fuel additive and claims that it reduces the fuel consumption of cars. To test this claim, 8 randomly selected cars were each filled with 20 litres of fuel and driven around a race circuit. Each car was tested twice, once with the additive and once without. The distance, in miles, that each car travelled before running out of fuel are given in the table below.

Car	1	2	3	4	5	6	7	8
Distance without additive	163	172	195	170	183	185	161	176
Distance with additive	168	185	187	172	180	189	172	175

Assuming that the distances travelled follow a normal distribution and stating your hypotheses clearly test, at the 10% level of significance, whether or not there is evidence to support the chemist's claim. **[E]**

Solution:

$d: 5, 13, -8, 2, -3, 4, 11, -1$

$(\Sigma d = 23, \Sigma d^2 = 409) \quad \bar{d} = 2.875, \ sd = 6.9987(\approx 7.00)$

$H_0: \mu_d = 0, H_1: \mu_d > 0$

$$t = \frac{(2.875 - 0)}{\frac{6.9987}{\sqrt{8}}} = 1.1618\ldots(\approx 1.16)$$

Critical value $t_7(10\%) = 1.415$ (1 tail)

Critical region is $t > 1.415$

Not significant

Insufficient evidence to support the chemist's claim

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 3

Question:

The standard deviation of the length of a random sample of 8 fence posts produced by a timber yard was 8 mm. A second timber yard produced a random sample of 13 fence posts with a standard deviation of 14 mm.

a Test, at the 10% significance level, whether or not there is evidence that the lengths of fence posts produced by these timber yards differ in variability. State your hypotheses clearly.

b State an assumption you have made in order to carry out the test in part a. **[E]**

Solution:

a $H_0 : \sigma_1^2 = \sigma_2^2$; $H_1 : \sigma_1^2 \neq \sigma_2^2$

$$\frac{s_1^2}{s_2^2} = \frac{14^2}{8^2} = 3.0625 \left(\text{or } \frac{8^2}{14^2} = 0.32653\ldots \right)$$

Critical value $F_{12,7} = 3.57$ $\left(\text{Critical value} : F_{7,12} = \frac{1}{3.57} = 0.28011 \right)$

Since 3.0625 is not in the critical region there is insufficient evidence to reject H_0.
There is insufficient evidence of a difference in the variances of the lengths of the fence posts.

b The distribution of the population of lengths of fence posts is normally distributed.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 4

Question:

A farmer set up a trial to assess the effect of two different diets on the increase in the weight of his lambs. He randomly selected 20 lambs. Ten of the lambs were given diet A and the other 10 lambs were given diet B. The gain in weight, in kg, of each lamb over the period of the trial was recorded.

a State why a paired t-test is not suitable for use with these data.

b Suggest an alternative method for selecting the sample which would make the use of a paired t-test valid.

c Suggest two other factors that the farmer might consider when selecting the sample.

The following paired data were collected.

Diet A	5	6	7	4.6	6.1	5.7	6.2	7.4	5	3
Diet B	7	7.2	8	6.4	5.1	7.9	8.2	6.2	6.1	5.8

d Using a paired t-test at the 5% significance level, test whether or not there is evidence of a difference in the weight gained by the lambs using diet A compared with those using diet B.

e State, giving a reason, which diet you would recommend the farmer to use for his lambs. **[E]**

Solution:

a The data were not collected in pairs.

b Use data from twin lambs.

c Age, weight, gender

d $d = B - A$

d : 2, 1.2, 1, 1.8, −1, 2.2, 2, −1.2, 1.1, 2.8

$\Sigma d = 11.9; \quad \Sigma d^2 = 30.01$

$\therefore \bar{d} = 1.19; s^2 = 1.761 \quad (s = 1.327)$

$H_0 : \mu_D = 0; H_1 : \mu_D \neq 0$

$$t = \frac{1.19 - 0}{\sqrt{\frac{1.761}{10}}} = 2.83574\ldots$$

$u = 9$; critical value : $t = 2.262$

Since $2.8357\ldots$ is in the critical region $(t > 2.262)$ there is evidence to reject H_0.

The (mean) weight gained by the lambs is different for each diet.

e Diet B - it has the higher mean

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 5

Question:

A medical student is investigating two methods of taking a person's blood pressure. He takes a random sample of 10 people and measures their blood pressure using an arm cuff and a finger monitor. The table below shows the blood pressure for each person, measured by each method.

Person	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Arm cuff	140	110	138	127	142	112	122	128	132	160
Finger monitor	154	112	156	152	142	104	126	132	144	180

a Use a paired t-test to determine, at the 10% level of significance, whether or not there is a difference in the mean blood pressure measured using the two methods. State your hypotheses clearly.

b State an assumption about the underlying distribution of measured blood pressure required for this test. **[E]**

Solution:

a $d: 14 \;\; 2 \;\; 18 \;\; 25 \;\; 0 \;\; -8 \;\; 4 \;\; 4 \;\; 12 \;\; 20$

$\left(\sum d = 91, \quad \sum x^2 = 1789\right)$

$\bar{d} = 9.1 \quad s = \sqrt{106.7} = 10.332..$

$H_0 : \mu_d = 0 \quad H_1 : \mu_d \neq 0$

$$t = \frac{(9.1 - 0)}{\frac{10.332}{\sqrt{10}}} = 2.785$$

Critical value $t_9 = \pm 1.833$

critical regions: $t \leq -1.833$ or $t \geq 1.833$

Significant. There is a difference between *blood pressure* measured by arm cuff and finger monitor.

b The *difference in measurements* of blood pressure is *normally distributed*

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 6

Question:

The lengths, x mm, of the forewings of a random sample of male and female adult butterflies are measured. The following statistics are obtained from the data.

	Number of butterflies	Sample mean $\bar{x}$	$\sum x^2$
Females	7	50.6	17 956.5
Males	10	53.2	28 335.1

a Assuming the lengths of the forewings are normally distributed, test, at the 10% level of significance, whether or not the variances of the two distributions are the same. State your hypotheses clearly.

b Stating your hypotheses clearly test, at the 5% level of significance, whether the mean length of the forewings of the female butterflies is less than the mean length of the forewings of the male butterflies. **[E]**

Solution:

a $H_0: \sigma^2_F = \sigma^2_M \quad H_1: \sigma^2_F \neq \sigma^2_M$

$$s^2_F = \frac{1}{6}(17\,956.5 - 7 \times 50.6^2) = \frac{33.98}{6} = 5.66333\ldots$$

$$s^2_M = \frac{1}{9}(28\,335.1 - 10 \times 53.2^2) = \frac{32.7}{9} = 3.63333\ldots$$

$$\frac{s^2_F}{s^2_M} = 1.5587\ldots \text{(Reciprocal 0.6415)}$$

$$F_{6,9} = 3.37 \text{(or 0.297)}$$

Not in critical region. There is no reason to doubt the *variances* of the two distributions *are the same*

b $H_0: \mu_F = \mu_M \quad H_1: \mu_F < \mu_M$

$$\text{Pooled estimate } s^2 = \frac{6 \times 5.66333\ldots + 9 \times 3.63333}{15}$$

$$= 4.44533$$

$$s = 2.11$$

$$t = \frac{50.6 - 53.2}{2.11\sqrt{\frac{1}{7} + \frac{1}{10}}} = -2.50$$

Critical value $t_{15}(5\%) = -1.753$

so critical region $t \le -1.753$

Significant. The mean length of the *female's forewing is* less than the length of the male's forewing

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 7

Question:

The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.
The ten mice are then fed on a special diet. They are weighted again after two weeks.
Their weights in grams are as follows:

Mouse	A	B	C	D	E	F	G	H	I	J
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice. **[E]**

Solution:

Differences 2.1 −0.7 2.6 −1.7 3.3 1.6 1.7 1.2 1.6 2.4

$\sum d = 14.1 \quad \sum d^2 = 40.65 \quad \bar{d} = 1.41$

$H_0 : \mu_d = 0 \quad H_1 : \mu_d > 0$

$$s = \sqrt{\frac{40.65 - 10 \times 1.41^2}{9}} = 1.5191\ldots$$

$$t = \frac{1.41}{\left(\frac{1.519\ldots}{\sqrt{10}}\right)} = 2.935$$

$t_9(1\%) = 2.821$

so critical region $t > 2.821$

$2.935.. > 2.821$ Evidence to reject H_0.

There has been an increase in the mean weight of the mice.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 8

Question:

A hospital department installed a new, more sophisticated piece of equipment to replace an ageing one in the hope that it would speed up the treatment of patients. The treatment times of random samples of patients during the last week of operation of the old equipment and during the first week of operation of the new equipment were recorded. The summary results, in minutes, were:

	n	$\sum x$	$\sum x^2$
Old equipment	10	225	5136.3
New equipment	9	234	6200.0

a Show that the values of s^2 for the old and new equipment are 8.2 and 14.5 respectively.

Stating clearly your hypotheses, test

b whether the variance of the times using the new equipment is greater than the variance of the times using the old equipment, using a 5% significance level,

c whether there is a difference between the mean times for treatment using the new equipment and old equipment, using a 2% significance level.

d Find 95% confidence limits for the mean difference in treatment times between the new and old equipment.

Even if the new equipment would eventually lead to a reduction in treatment times, it might be that to begin with treatment times using the new equipment would be higher than those using the old equipment.

e Give one reason why this might be so.

f Suggest how the comparison between the old and new equipment could be improved. **[E]**

Solution:

a $s_o^2 = \frac{5136.3}{9} - \frac{(225)^2}{10(10-1)} = 8.2$

$s_n^2 = \frac{6200}{8} - \frac{(234)^2}{9(9-1)} = 14.5$

b $H_0: \sigma_o^2 = \sigma_n^2 \quad H_1: \sigma_o^2 < \sigma_n^2$

Critical value is $F_{8,9} = 3.23$

so critical region, $F \geq 3.23$

$F_{test} = \frac{14.5}{8.2} = 1.768$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_o^2 = \sigma_n^2$

c $s_p^2 = \frac{(9 \times 8.2) + (8 \times 14.5)}{10 + 9 - 2} = 11.1647$

$H_0: \mu_0 = \mu_n \quad H_1: \mu_o \neq \mu_n$

critical value $t_{17}(0.01) = 2.567$

critical region $t \leq -2.567$ and $t \geq 2.567$

$t = \frac{(26 - 22.5) - 0}{\sqrt{11.1647}\sqrt{\frac{1}{10} + \frac{1}{9}}} = 2.2798$

Not in the critical region – do not reject H_0.

There is evidence to suggest that there is no difference in mean times between the old and new equipment.

d $t_{17}(2.5\%) = 2.110$

$(26 - 22.5) \pm 2.110 \times \sqrt{11.1647} \times \sqrt{\frac{1}{10} + \frac{1}{9}} = 3.5 \pm 3.2394$

$= (0.261, 6.739)$

e Need to learn how to use new equipment efficiently

f Gather data on new equipment after it has been mastered

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 1

Question:

Historical records from a large colony of squirrels show that the weight of squirrels is normally distributed with a mean of 1012 g. Following a change in the diet of squirrels, a biologist is interested in whether or not the mean weight has changed.
A random sample of 14 squirrels is weighed and their weights x, in grams, recorded.
The results are summarised as follows:

$$\sum x = 13\,700, \quad \sum x^2 = 13\,448\,750\,.$$

Stating your hypotheses clearly test, at the 5% level of significance, whether or not there has been a change in the mean weight of the squirrels. **[E]**

Solution:

$H_0 : \mu = 1012 \quad H_1 : \mu \neq 1012$

$$\bar{x} = \frac{13\,700}{14} (= 978.57\ldots)$$

$$S_x^2 = \frac{13\,448\,750 - 14\bar{x}^2}{13} (= 3255.49)$$

$$t_{13} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{978.6 - 1012}{\frac{57.06}{\sqrt{14}}} = -2.19\ldots$$

t_{13} (5%) two-tail critical value $= -2.160$

Significant result – there is evidence of a change in mean weight of squirrels

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 2

Question:

A random sample $X_1, X_2, \ldots, X_{10}$ is taken from a population with mean μ and variance σ^2.

a Determine the bias, if any, of each of the following estimators of μ.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}$$

b Find the variance of each of these estimators.

c State, giving reasons, which of these three estimators for μ is

i the best estimator,

ii the worst estimator. **[E]**

Solution:

a $\mathrm{E}(\theta_1) = \dfrac{\mathrm{E}\left((X_3) + \mathrm{E}(X_4) + \mathrm{E}(X_5)\right)}{3} = \dfrac{3\mu}{3} = \mu \quad \text{Bias} = 0$

$\mathrm{E}(\theta_2) = \dfrac{\mathrm{E}(X_{10}) - \mathrm{E}(X_1)}{3} = \dfrac{1}{3}(\mu - \mu) = 0 \quad \text{Bias} = -\mu$

$\mathrm{E}(\theta_3) = \dfrac{3\mathrm{E}(X_1) + 2\mathrm{E}(X_2) + \mathrm{E}(X_{10})}{6} = \dfrac{3\mu + 2\mu + \mu}{6} = \mu \quad \text{Bias} = 0$

b $\mathrm{Var}(\theta_1) = \dfrac{1}{9}\left(\sigma^2 + \sigma^2 + \sigma^2\right) = \dfrac{\sigma^2}{3}$

$\mathrm{Var}(\theta_2) = \dfrac{1}{9}\left(\sigma^2 + \sigma^2\right) = \dfrac{2\sigma^2}{9}$

$\mathrm{Var}(\theta_3) = \dfrac{1}{36}\left[9\sigma^2 + 4\sigma^2 + \sigma^2\right] = \dfrac{14\sigma^2}{36} = \dfrac{7\sigma^2}{18}$

c Don't use θ_2 as it is biased.

$\mathrm{Var}(\theta_1) = \dfrac{\sigma^2}{3} = \dfrac{6\sigma^2}{18}$

$\mathrm{Var}(\theta_3) = \dfrac{7\sigma^2}{18}$

i So choose θ_1 as it is unbiased and has the smallest variance, to be the best estimator

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 3

Question:

A random sample of 10 mustard plants had the following heights, in millimetres, after 4 days growth.

5.0, 4.5, 4.8, 5.2, 4.3, 5.1, 5.2, 4.9, 5.1, 5.0

Those grown previously had a mean height of 5.1 mm after 4 days. Using a 2.5% significance level, test whether or not the mean height of these plants is less than that of those grown previously.
(You may assume that the height of mustard plants after 4days follows a normal distribution.) **[E]**

Solution:

$H_0 : \mu = 5.1, H_1 : \mu < 5.1$

$\nu = 9$

Critical Region $t < -2.262$

$\bar{x} = 4.91$

$$s^2 = \frac{241.89 - 10 \times (4.91)^2}{9} = 0.0899$$

$s = 0.300$

$$t = \frac{4.91 - 5.1}{\frac{0.3}{\sqrt{10}}} = -2.00$$

There is no evidence to suggest that the mean height is less than those grown previously

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 4

Question:

A mechanic is required to change car tyres. An inspector timed a random sample of 20 tyre changes and calculated the unbiased estimate of the population variance to be 6.25 minutes^2. Test, at the 5% significance level, whether or not the standard deviation of the population of times taken by the mechanic is grater than 2 minutes. State your hypotheses clearly.

[E]

Solution:

$H_0 : \sigma^2 = 4; H_1 : \sigma^2 > 4$

$\nu = 19, X^2_{19}(0.05) = 30.144$

$$\frac{(n-1)S^2}{\sigma^2} = \frac{19 \times 6.25}{4} = 29.6875$$

Since $29.6875 < 30.144$ there is insufficient evidence to reject H_0.

There is insufficient evidence to suggest that the standard deviation is greater than 2.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 5

Question:

The value of orders, in £, made to a firm over the internet has distribution $N(\mu, \sigma^2)$.
A random sample of n orders is taken and $\bar{X}$ denotes the sample mean.

a Write down the mean and variance of $\bar{X}$ in terms of μ and σ^2.

A second sample of m orders is taken and $\bar{Y}$ denotes the mean of this sample.

An estimator of the population mean is given by $U = \dfrac{n\bar{X} + m\bar{Y}}{n+m}$.

b Show that U is an unbiased estimator for μ.

c Show that the variance of U is $\dfrac{\sigma^2}{n+m}$.

d State which of $\bar{X}$ or U is a better estimator for μ. Give a reason for your answer.

[E]

Solution:

a $E(\bar{X}) = \mu$

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

b
$$E(U) = \frac{1}{n+m}\left(nE\left(\bar{X}\right) + mE\left(\bar{Y}\right)\right)$$
$$= \frac{1}{n+m}(n\mu + m\mu)$$
$$= \mu \Rightarrow U \text{ is unbiased}$$

c
$$\mathrm{Var}\left(\bar{Y}\right) = \frac{\sigma^2}{m}$$
$$\mathrm{Var}(U) = \frac{n^2\mathrm{Var}\left(\bar{X}\right) + m^2\mathrm{Var}\left(\bar{Y}\right)}{(n+m)^2}$$
$$= \frac{n^2\frac{\sigma^2}{n} + m^2\frac{\sigma^2}{m}}{(n+m)^2}$$
$$= \frac{n\sigma^2 + m\sigma^2}{(n+m)^2}$$
$$= \frac{\sigma^2}{n+m}$$

d $\dfrac{n\bar{X} + m\bar{Y}}{n+m}$ is a better estimate since variance is smaller.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 6

Question:

A machine is set to fill bags with flour such that the mean weight is 1010 grams. To check that the machine is working properly, a random sample of 8 bags is selected. The weight of flour, in grams, in each bag is as follows.

1010 1015 1005 1000 998 1008 1012 1007

Carry out a suitable test, at the 5% significance level, to test whether or not the mean weight of flour in the bags is less than 1010 grams. (You may assume that the weight of flour delivered by the machine is normally distributed.)

Solution:

Let x represent weight of flour

$\sum x = 8055 \quad \therefore \bar{x} = 1006.875$

$\sum x^2 = 8\ 110\ 611 \therefore s^2 = \frac{1}{7}\left\{8\ 110\ 611 - \frac{8055^2}{8}\right\} = 33.26785\ldots$

$\therefore s = 5.767825$

$H_0 : \mu = 1010; H_1 : \mu < 1010$

critical value: $t = -1.895$ so critical region $t < -1.895$

$$t = \frac{(1006.875 - 1010)}{\left(\frac{5.7678}{\sqrt{8}}\right)} = -1.5324$$

Since -1.53 is not in the critical region there is insufficient evidence to reject H_0.
The mean weight of flour delivered by the machine is 1010g.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 7

Question:

A train company claims that the probability p of one of its trains arriving late is 10%. A regular traveler on the company's trains believes that the probability is greater than 10% and decides to test this by randomly selecting 12 trains and recording the number, X, of trains that were late. The traveller sets up the hypotheses $H_0: p = 0.1$ and $H_1: p > 0.1$ and accepts the null hypothesis if $x \leq 2$.

a Find the size of the test.

b Show that the power function of the test is $1-(1-p)^{10}(1+10p+55p^2)$.

c Calculate the power of the test when

 i $p = 0.2$

 ii $p = 0.6$

d Comment on your results from part **c**. **[E]**

Solution:

a $1-0.8891 = 0.1109$

b $1-(P(0)+P(1)+P(2))$

$$= 1-((1-p)^{12} + 12p(1-p)^{11} + 66p^2(1-p)^{10})$$
$$= 1-(1-p)^{10}((1-p)^2 + 12p(1-p) + 66p^2)$$
$$= 1-(1-p)^{10}\left[1-2p+p^2+12p-12p^2+66p^2\right]$$
$$= 1-(1-p)^{10}(1+10p+55p^2)$$

c $1-0.5583 = 0.442$

$1-0.00281 = 0.997$

d The test is more discriminating (powerful) for the larger value of p.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 8

Question:

It is suggested that a Poisson distribution with parameter λ can model the number of currants in a currant bun. A random bun is selected in order to test the hypotheses $H_0 : \lambda = 8$ against $H_1 : \lambda \neq 8$, using a 10% level of significance.

a Find the critical region for this test, such that the probability in each tail is as close as possible to 5%.

b Given that $\lambda = 10$, find

i the probability of a type II error,

ii the power of the test. **[E]**

Solution:

a $P(X \leq c_1) \leq 0.05; P(X \leq 3 | \lambda = 8) = 0.0424 \Rightarrow X \leq 3$

$P(X \geq c_2) \leq 0.05; P(X \geq 14 | \lambda = 8) = 0.0342$

$P(X \geq 13 | \lambda = 8) = 0.0638 \Rightarrow X \geq 13$

$\therefore$ critical region is $\{X \leq 3\} \cup \{X \geq 13\}$

b **i** $P(4 \leq X \leq 12 | \lambda = 10) = P(X \leq 12) - P(X \leq 3)$

$= 0.7916 - 0.0103$

$= 0.7813$

ii $\text{Power} = 1 - 0.7813 = 0.2187$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 9

Question:

The length X mm of a spring made by a machine is normally distributed $N(\mu, \sigma^2)$.
A random sample of 20 springs is selected and their lengths measured in millimetres.
Using this sample, the unbiased estimates of μ and σ^2 are $\bar{x} = 100.6$ $s^2 = 1.5$
Stating your hypotheses clearly test, at the 10% level of significance,
a whether or not the variance of the lengths of springs is different from 0.9,
b whether or not the mean length of the springs is greater than 100 mm. **[E]**

Solution:

a $H_0 : \sigma^2 = 0.9$ $H_1 : \sigma^2 \neq 0.9$
$\nu = 19$
CR (Lower tail 10.117)
Upper tail 30.144
Test statistic $= \frac{19 \times 1.5}{0.9} = 31.6666$, significant
There is sufficient evidence that the *variance* of the length of spring is *different from 0.9*

b $H_0 : \mu = 100$ $H_1 : \mu > 100$
$t_{19} = 1.328$ is the critical value
$$t = \frac{100.6 - 100}{\sqrt{\frac{1.5}{20}}} = 2.19$$
Significant. The mean *length of spring* is *greater than 100*

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 10

Question:

A town council is concerned that the mean price of renting two bedroom flats in the town has exceeded £650 per month. A random sample of eight two bedroom flats gave the following results, £x, per month.

705, 640, 560, 680, 800, 620, 580, 760

[You may assume $\sum x = 5345, \sum x^2 = 3\,621\,025$.]

a Find a 90% confidence interval for the mean price of renting a two bedroom flat.

b State an assumption that is required for the validity of your interval in part **a**.

c Comment on whether or not the town council is justified in being concerned.

Give a reason for your answer. **[E]**

Solution:

a $\bar{x} = 668.125 \quad s = 84.425$

$t_7(5\%) = 1.895$

$$\text{Confidence limits} = 668.125 \pm \frac{1.895 \times 84.425}{\sqrt{8}} = 611.6 \text{ and } 724.7$$

Confidence interval = (612, 725)

b Normal distribution

c £650 is within the confidence interval. No need to worry.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 11

Question:

A technician is trying to estimate the area μ^2 of a metal square. The independent random variables X_1 and X_2 are each distributed $N(\mu, \sigma^2)$ and represent two measurements of the sides of the square. Two estimators of the area, A_1 and A_2, are proposed where $A_1 = X_1X_2$ and $A_2 = \left(\frac{X_1 + X_2}{2}\right)^2$.

[You may assume that if X_1 and X_2 are independent random variables then $E(X_1X_2) = E(X_1)E(X_2)$]

a Find $E(A_1)$ and show that $E(A_2) = \mu^2 + \frac{\sigma^2}{2}$.

b Find the bias of each of these estimators.

The technician is told that $Var(A_1) = \sigma^4 + 2\mu^2\sigma^2$ and $Var(A_2) = \frac{1}{2}\sigma^4 + 2\mu^2\sigma^2$.

The technician decided to use A_1 as the estimator for μ^2.

c Suggest a possible reason for this decision.

A statistician suggests taking a random sample of n measurements of sides of the square and finding the mean $\bar{X}$. He knows that $E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$ and $Var(\bar{X}^2) = \frac{2\sigma^4}{n^2} + \frac{4\sigma^2\mu^2}{n}$.

d Explain whether or not $\bar{X}^2$ is a consistent estimator of μ^2. **[E]**

Solution:

a $E(A_1) = E(X_1)E(X_2) = \mu^2$

$A_2 = \bar{X}^2, \bar{X} \sim N\left(\mu, \frac{\sigma^2}{2}\right) \therefore E(\bar{X}^2) = E(A_2) = \mu^2 + \frac{\sigma^2}{2}$

alternative for A_2

$$E(A_2) = E\left[\left(\frac{X_1 + X_2}{2}\right)^2\right]$$
$$= E\left[\frac{X_1^2 + 2X_1X_2 + X_2^2}{4}\right]$$
$$= E\left[\frac{X_1^2}{4}\right] + E\left[\frac{X_1X_2}{2}\right] + E\left[\frac{X_2^2}{4}\right]$$
$$= \frac{1}{4}E\left(X_1^2\right) + \frac{1}{2}E\left(X_1X_2\right) + \frac{1}{4}E\left(X_2^2\right)$$

but $Var(X) = E(X^2) - \mu^2$

so $\mu^2 + Var X = E(X^2)$

$\therefore E\left(X_1\right)^2 = \mu^2 + \sigma^2$ and $E\left(X_2^2\right) = \mu^2 + \sigma^2$

$$\therefore E\left(A_2\right) = \frac{1}{4}\left(\mu^2 + \sigma^2\right) + \frac{1}{2}\mu^2 + \frac{1}{4}\left(\mu^2 + \sigma^2\right)$$
$$= \mu^2 + \frac{\sigma^2}{2}$$

b A_1 is unbiased, bias for A_2 is $\frac{\sigma^2}{2}$

c Used A_1 since it is unbiased

d $E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$; as $n \to \infty, E(\bar{X}^2) \to \mu^2$

$Var(\bar{X}^2) = \frac{2\sigma^4}{n^2} + \frac{4\sigma^2\mu^2}{n}$; as $n \to \infty, Var(\bar{X}^2) \to 0$

$\bar{X}^2$ is a consistent estimator of μ^2

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 12

Question:

A random sample of 15 tomatoes is taken and the weight x grams of each tomato is found. The results are summarised by $\sum x = 208$ and $\sum x^2 = 2962$.

a Assuming that the weights of the tomatoes are normally distributed, calculate the 90% confidence interval for the variance σ^2 of the weights of the tomatoes.

b State, with a reason, whether or not the confidence interval supports the assertion $\sigma^2 = 3$. **[E]**

Solution:

a $s^2 = \dfrac{2962 - 15 \times \left(\dfrac{208}{15}\right)^2}{14} = 5.55$

$\dfrac{14 \times 5.55}{23.685} < \sigma^2 < \dfrac{14 \times 5.55}{6.571}$

$3.28 < \sigma^2 < 11.83$

b Since 9 lies in the interval, yes, it supports the assertion.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 13

Question:

a Define
  i a type I error,
  ii a type II error.

A small aviary, that leaves the eggs with the parent birds, rears chicks at an average rate of 5 per year. In order to increase the number of chicks reared per year it is decided to remove the eggs from the aviary as soon as they are laid and put them in an incubator. At the end of the first year of using an incubator 7 chicks had been successfully reared.

b Assuming that the number of chicks reared per year follows a Poisson distribution test, at the 5% significance level, whether or not there is evidence of an increase in the number of chicks reared per year. State your hypotheses clearly.

c Calculate the probability of the type I error for this test.

d Given that the true average number of chicks reared per year when the eggs are hatched in an incubator is 8, calculate the probability of a type II error. **[E]**

Solution:

a Type I – H_0 rejected when it is true

Type II – H_0 is accepted when it is false

b $H_0 : \lambda = 5, H_1 : \lambda > 5$

$P(X \geq 7 \mid \lambda = 5) = 1 - 0.7622 = 0.2378 > 0.05$

No evidence of an increase in the number of chicks reared per year.

c $P(X \geq c \mid \lambda = 5) < 0.05$

$P(X \geq 9) = 0.0681, P(X \geq 10) = 0.0318, c = 10$

$P(\text{Type I Error}) = 0.0318$

d $\lambda = 8$

$P(X \leq 9 \mid \lambda = 8) = 0.7166$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 14

Question:

a Explain briefly what you understand by
  i an unbiased estimator,
  ii a consistent estimator
  of an unknown population parameter θ.

From a binomial population, in which the proportion of successes is p, 3 samples of size n are taken. The number of successes X_1, X_2, and X_3 are recorded and used to estimate p.

b Determine the bias, if any, of each of the following estimators of p.

$$\hat{p}_1 = \frac{X_1 + X_2 + X_3}{3n},$$

$$\hat{p}_2 = \frac{X_1 + 3X_2 + X_3}{6n},$$

$$\hat{p}_3 = \frac{2X_1 + 3X_2 + X_3}{6n}.$$

c Find the variance of each of these estimators.

d State, giving a reason, which of the three estimators for p is
  i the best estimator,
  ii the worst estimator. **[E]**

Solution:

a i $\mathrm{E}(\hat{\theta}) = \theta$

ii $\mathrm{E}(\hat{\theta}) = \theta$ or $\mathrm{E}(\hat{\theta}) \to \theta$
and $\mathrm{Var}(\hat{\theta}) \to 0$ as $n \to \infty$ where n is the sample size

b $\mathrm{E}(\hat{p}_1) = p$, $\therefore$ Bias $= 0$

$\mathrm{E}(\hat{p}_2) = \frac{5p}{6}$, $\therefore$ Bias $= -\frac{1}{6}p$

$\mathrm{E}(\hat{p}_3) = p$, $\therefore$ Bias $= 0$

c $\mathrm{Var}(\hat{p}_1) = \frac{1}{9n^2}\{npq + npq + npq\}$

$= \frac{pq}{3n}$ or $\frac{12pq}{36n}$

$\mathrm{Var}(\hat{p}_2) = \frac{1}{36n^2}\{npq + 9npq + npq\} = \frac{11pq}{36n}$

$\mathrm{Var}(\hat{p}_3) = \frac{1}{36n^2}\{4npq + 9npq + npq\} = \frac{7pq}{18n}$ or $\frac{14pq}{36n}$

d i $\hat{p}_1$; unbiased and smallest variance

ii $\hat{p}_2$; biased

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 15

Question:

Define

a a type I error,

b the size of a test.

Jane claims that she can read Alan's mind. To test this claim Alan randomly chooses a card with one of 4 symbols on it. He then concentrates on the symbol. Jane then attempts to read Alan's mind by stating what symbol she thinks is on the card. The experiment is carried out 8 times and the number of times, X, that Jane is correct is recorded.

The probability of Jane stating the correct symbol is denoted by p.

To test the hypothesis $H_0: p = 0.25$ against $H_1: p > 0.25$, a critical region of $X > 6$ is used.

c Find the size of this test.

d Show that the power function of this test is $8p^7 - 7p^8$.

Given that $p = 0.3$, calculate

e the power of this test,

f the probability of a type II error.

g Suggest two ways in which you might reduce the probability of a type II error. **[E]**

Solution:

a A Type I error occurs when H_0 is rejected when it is in fact true.

b The size of a test is the probability of a type I error.

c $X \sim B(8, 0.25)$

$$\text{Size} = P(X > 6) = 1 - P(X \le 6 \mid n = 8, p = 0.25)$$
$$= 1 - 0.9996 = 0.0004$$

d $\text{Power} = P(X > 6 \mid p = p, n = 8)$

$$= P(X = 7) + P(X = 8) = \frac{8}{7!1!} p^7(1 - p) + p^8$$
$$= 8p^7 - 8p^8 + p^8 = 8p^7 - 7p^8$$

e $\text{Power} = 8 \times 0.3^7 - 7 \times 0.3^8$

$$= 0.00129$$

f $\text{Power} = 1 - P(\text{Type II})$

$$\therefore P(\text{Type II}) = 1 - \text{power}$$
$$= 1 - 0.00129$$
$$= 0.9987$$

g Increase the probability of a Type I error, e.g. increase the significance level of the test. Increase the value of Po

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 16

Question:

Rolls of cloth delivered to a factory contain defects at an average rate of λ per metre. A quality assurance manager selects a random sample of 15 metres of cloth from each delivery to test whether or not there is evidence that $\lambda > 0.3$. The criterion that the manager uses for rejecting the hypothesis that $\lambda = 0.3$ is that there are 9 or more defects in the sample.

a Find the size of the test.

Table 1 gives some values, to 2 decimal places, of the power function of this test.

λ	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Power	0.15	0.34	r	0.72	0.85	0.92	0.96

Table 1

b Find the value of r.

The manager would like to design a test, of whether or not $\lambda > 0.3$, that uses a smaller length of cloth. He chooses a length of 10 m and requires the probability of a type I error to be less than 10%.

c Find the criterion to reject the hypothesis that $\lambda = 0.3$ which makes the test as powerful as possible.

d Hence state the size of this second test.

Table 2 gives some values, to 2 decimal places, of the power function for the test in part **c**.

λ	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Power	0.21	0.38	0.55	0.70	s	0.88	0.93

Table 2

e Find the value of s.

f Using the same axes, on graph paper draw the graphs of the power functions of these two tests.

g **i** State the value of λ where the graphs cross.

ii Explain the significance of λ being greater than this value.

The cost of wrongly rejecting a delivery of cloth with $\lambda = 0.3$ is low. Deliveries of cloth with $\lambda > 0.7$ are unusual.

h Suggest, giving your reasons, which the test manager should adopt. **[E]**

Solution:

a $X_1 =$ no. of defects in 15 m $\quad X_1 \sim \text{Po}(4.5)$

Size $= \text{P}(X_2 \geq 9) = 1 - 0.9597 = 0.0403$

b $r = \text{P}(X_2 \geq 9 \mid X_2 \sim \text{Po}(9)) = 1 - 0.4557 = 0.54(43)$

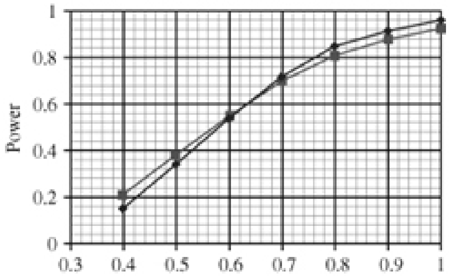
c $Y_1 =$ No of defects in 10 m $\quad Y_1 \sim \text{Po}(3)$

$\text{P}(Y_1 \geq c) < 0.10 \quad Y_1 \geq 6$

d Size $= \text{P}(Y_1 \geq 6) = 1 - \text{P}(Y_1 \leq 5) = 1 - 0.9161 = 0.0839$

e $s = 1 - \text{P}(Y_2 \leq 5) = 1 - 0.1912 = 0.8088$

f



g **i** 0.62 to 0.67

ii Test I more powerful

h Test 2 Has a higher P(Type I error) but cost of this is low.

Test 2 is more powerful for $\lambda < 0.7$ and $\lambda > 0.7$ is rare.

Adopt test 2

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 17

Question:

The number of tornadoes per year to hit a particular town follows a Poisson distribution with mean λ. A weatherman claims that due to climate changes the mean number of tornadoes per year has decreased. He records the number of tornadoes x to hit the town last year.

To test the hypotheses $H_0 : \lambda = 7$ and $H_1 : \lambda < 7$, a critical region of $x \leq 3$ is used.

a Find, in terms λ the power function of this test.

b Find the size of this test.

c Find the probability of a type II error when $\lambda = 4$. **[E]**

Solution:

a Power $= P(X \leq 3 \quad \lambda = 3)$

$$= e^{-\lambda} + e^{-\lambda}\lambda + \frac{e^{-\lambda}\lambda^2}{2} + \frac{e^{-\lambda}\lambda^3}{6}$$

$$= \frac{e^{-\lambda}}{6}(6 + 6\lambda + 3\lambda^2 + \lambda^3)$$

b CR is $X \leq 3$

Size $= P(X \leq 3 \quad \lambda = 7)$

$= 0.0818$

c P(Type II error) $= 1 -$ power

$$= 1 - \frac{e^{-4}}{6}(6 + 6\times 4 + 3\times 4^2 + 4^3)$$

$= 0.5665..$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 18

Question:

A nutritionist studied the levels of cholesterol, X mg/cm^3, of male students at a large college. She assumed that X was distributed $N(\mu, \sigma^2)$ and examined a random sample of 25 male students. Using this sample she obtained unbiased estimates of μ and σ^2 as

$\hat{\mu} = 1.68 \quad \hat{\sigma}^2 = 1.79$

a Find a 95% confidence interval for μ.

b Obtain a 95% confidence interval for σ^2.

A cholesterol reading of more than 2.5 mg/cm^3 is regarded as high.

c Use appropriate confidence limits from parts **a** and **b** to find the lowest estimate of the proportion of male students in the college with high cholesterol. **[E]**

Solution:

a 95% confidence interval for μ is

$$1.68 \pm t_{24}(2.5\%)\sqrt{\frac{1.79}{25}} = 1.68 \pm 2.064\sqrt{\frac{1.79}{25}} = (1.13, 2.23)$$

b 95% confidence interval for σ^2 is

$$12.401 < \frac{24 \times 1.79}{\sigma^2} < 39.364$$

$\sigma^2 > 1.09, \sigma^2 < 3.46$

$\therefore$ confidence interval on σ^2 is (1.09, 3.46)

c Require $P(X > 2.5) = P\left(Z > \frac{2.5 - \mu}{\sigma}\right)$ to be as small as possible OR

$\frac{2.5 - \mu}{\sigma}$ to be as large as possible; both imply lowest σ and μ.

$$\frac{2.5 - 1.13}{\sqrt{1.09}} = 1.31$$

$P(Z > 1.31) = 1 - 0.9049 = 0.0951$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 19

Question:

A random sample of three independent variables X_1, X_2 and X_3 is taken from a distribution with mean μ and variance σ^2.

a Show that $\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3$ is an unbiased estimator for μ.

An unbiased estimator for μ is given by $\hat{\mu} = aX_1 + bX_2$ where a and b are constants.

b Show that $\text{Var}(\hat{\mu}) = (2a^2 - 2a + 1)\sigma^2$.

c Hence determine the value of a and the value of b for which $\hat{\mu}$ has minimum variance. **[E]**

Solution:

a $\text{E}\left(\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3\right) = \frac{2}{3}\mu - \frac{1}{2}\mu + \frac{5}{6}\mu = \mu$

$\text{E}(Y) = \mu \Rightarrow \text{unbiased}$

b

$$\text{E}(aX_1 + bX_2) = a\mu + b\mu = \mu$$
$$a + b = 1$$
$$\text{Var}(aX_1 + bX_2) = a^2\sigma^2 + b^2\sigma^2$$
$$= a^2\sigma^2 + (1-a)^2\sigma^2$$
$$= (2a^2 - 2a + 1)\sigma^2$$

c Minimum value when $(4a - 2)\sigma^2 = 0$ (from differentiation)

$\Rightarrow 4a - 2 = 0$

$a = \frac{1}{2}, b = \frac{1}{2}$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 20

Question:

A supervisor wishes to check the typing speed of a new typist. On 10 randomly selected occasions, the supervisor records the time taken for the new typist to type 100 words. The results, in seconds, are given below.

110	125	130	126	128	127	118
120	122	125				

The supervisor assumes that the time taken to type 100 words is normally distributed.

a Calculate a 95% confidence interval for

i the mean,

ii the variance

of the population of times taken by this typist to type 100 words.

The supervisor requires the average time needed to type 100 words to be no more than 130 seconds and the standard deviation to be no more than 4 seconds.

b Comment on whether or not the supervisor should be concerned about the speed of the new typist. **[E]**

Solution:

a $\bar{x} = 123.1$

$s = 5.87745\ldots$

(NB: $\Sigma x = 1231$; $\Sigma x^2 = 151\ 847$)

i 95% confidence interval is given by

$$123.1 \pm 2.262 \times \frac{5.87745\ldots}{\sqrt{10}}$$

i.e. (118.8958 …, 127.30418 …)

ii 95% confidence interval is given by

$$\frac{9 \times 5.87745\ldots^2}{19.023} < \sigma^2 < \frac{9 \times 5.87745\ldots^2}{2.700}$$

i.e. (16.34336 …, 115.14806 ….)

b 130 is just above confidence interval

16 is just below confidence interval

Thus supervisor should be concerned about the speed of the new typist since both their average speed is two slow and the variability of the time is too large

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 21

Question:

A machine is filling bottles of milk. A random sample of 16 bottles was taken and the volume of milk in each bottle was measured and recorded. The volume of milk in a bottle is normally distributed and the unbiased estimate of the variance, s^2, of the volume of milk in a bottle is 0.003.

a Find a 95% confidence interval for the variance of the population of volumes of milk from which the sample was taken.

The machine should fill bottles so that the standard deviation of the volumes is equal to 0.07.

b Comment on this with reference to your 95% confidence interval. **[E]**

Solution:

a $\text{Confidence interval} = \left(\frac{15\times 0.003}{27.488}, \frac{15\times 0.003}{6.262}\right)$

$= (0.00164, 0.00719)$

b $0.07^2 = 0.0049$

0.0049 is within the 95% confidence interval.

There is no evidence to reject the idea that the standard deviation of the volumes is not 0.07 or the machine is working well.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 22

Question:

A butter packing machine cuts butter into blocks. The weight of a block of butter is normally distributed with a mean weight of 250 g and a standard deviation of 4 g.
A random sample of 15 blocks is taken to monitor any change in the mean weight of the blocks of butter.

a Find the critical region of a suitable test using a 2% level of significance.

b Assuming the mean weight of a block of butter has increased to 254 g, find the probability of a type II error. **[E]**

Solution:

a $\dfrac{\bar{X}-250}{\frac{4}{\sqrt{15}}} > 2.3263$ or $\dfrac{\bar{X}-250}{\frac{4}{\sqrt{15}}} < -2.3263$

$\bar{X} > 252.40\ldots$ or $\bar{X} < 247.6\ldots$

b $\mathrm{P}(\bar{X} > 252.4 / \mu = 254) - \mathrm{P}(\bar{X} < 247.6 / \mu = 254)$

$= \mathrm{P}\left(Z > \dfrac{252.4-254}{\frac{4}{\sqrt{15}}}\right) - \mathrm{P}\left(Z < \dfrac{247.6-254}{\frac{4}{\sqrt{15}}}\right)$

$= \mathrm{P}(Z > -1.5492) - \mathrm{P}(Z < -6.20)$

$= (1-0.9394)-(0)$

$= 0.0606$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 23

Question:

A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted.

a Write down suitable hypotheses to carry out this test.

b Find the probability of making a type I error.

The table below gives the value of the probability of the type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

c Calculate the value of r and the value of s.

d Calculate the power of the test for $p = 0.2$ and $p = 0.4$

e Comment, giving your reasons, on the suitability of this test procedure. **[E]**

Solution:

a $H_0 : p = 0.35 \quad H_1 : p \neq 0.35$

b Let $X =$ Number cured then $X \sim B(20, 0.35)$

$\alpha = P(\text{Type I error}) = P(x \leq 3) + P(x \geq 11)$ given $p = 0.35$

$= 0.0444 + 0.0532$

$= 0.0976$

c $\beta = P(\text{Type II error}) = P(4 \leq \times \leq 10)$

p	0.2	0.3	0.4	0.5
β	0.5880	0.8758	0.8565	0.5868

d Power $= 1 - \beta$

0.4120 0.1435

e Not a good procedure.
Better further away from 0.35 or
This is not a very powerful test (power $= 1 - \beta$)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 24

Question:

A doctor wishes to study the level of blood glucose in males. The level of blood glucose is normally distributed. The doctor measured the blood glucose of 10 randomly selected male students from a school. The results, in mmol/litre, are given below.

4.7	3.6	3.8	4.7	4.1
2.2	3.6	4.0	4.4	5.0

a Calculate a 95% confidence interval for the mean.

b Calculate a 95% confidence interval for the variance.

A blood glucose reading of more than 7 mmol/litre is counted as high.

c Use appropriate confidence limits from parts **a** and **b** to find the highest estimate of the proportion of male students in the school with a high blood glucose level. **[E]**

Solution:

$\bar{x} = 4.01$

$s = 0.7992\ldots$

a $4.01 \pm t_9(2.5\%)\dfrac{0.7992..}{\sqrt{10}} = 4.01 \pm 2.262\dfrac{0.7992..}{\sqrt{10}}$

$= 4.5816\ldots$ and $3.4383\ldots$

i.e. (3.4383, 4.5816.)

b $2.700 < \dfrac{9 \times 0.7992\ldots^2}{s^2} < 19.023$

$\sigma^2 < 2.13, \sigma^2 > 0.302$

i.e. (0.302, 2.13)

c $\text{P}(X > 7) = \text{P}\left(Z > \dfrac{7-\mu}{\sigma}\right)$ needs to be as high as possible

Therefore μ and σ must be as big as possible

propertion with high blood glucose level $= \text{P}\left(Z > \dfrac{7-4.581}{\sqrt{2.13}}\right)$

$= 1 - 0.9515$

$= 0.0485$

$= 4.85\%$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 1

Question:

The random variable X has an F distribution with 10 and 12 degrees of freedom. Find a and b such that $\mathrm{P}(a < X < b) = 0.90$. **[E]**

Solution:

$$F_{10,12}(5\%) = 2.75 \therefore b = 2.75$$

$$a = \frac{1}{F_{12,10}(5\%)} = \frac{1}{2.91} = 0.344$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 2

Question:

A doctor believes that the span of a person's dominant hand is greater than that of the weaker hand. To test this theory, the doctor measures the spans of the dominant and weaker hands of a random sample of 8 people. He subtracts the span of the weaker hand from that of the dominant hand. The spans, in millimetres, are summarised in the table below.

	Dominant hand	Weaker hand
A	202	195
B	251	249
C	215	218
D	235	234
E	210	211
F	195	197
G	191	181
H	230	225

Test, at the 5% significance level, the doctor's belief. **[E]**

Solution:

d: 7 2 −3 1 −1 −2 10 5

$\Sigma d = 19; \Sigma d^2 = 193$

$\therefore \bar{d} = \frac{19}{8} = 2.375; S_d^2 = \frac{1}{7}(193 - \frac{19^2}{8}) = 21.125$

$H_0 : \mu_D = 0; H_1 : \mu_D > 0$

$t = \frac{2.375 - 0}{\sqrt{\frac{21.125}{8}}} = 1.4615\ldots$

$\nu = 7 \Rightarrow$ critical region : $t > 1.895$

Since 1.4915 … is *not* in the critical region there is insufficient evidence to reject H_0 and we conclude that there is insufficient evidence to support the doctors' belief.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 3

Question:

The times, x seconds, taken by the competitors in the 100 m freestyle events at a school swimming gala are recorded. The following statistics are obtained from the data.

	Number of competitors	Sample mean $\bar{x}$	$\sum x^2$
Girls	8	83.10	55 746
Boys	7	88.90	56 130

Following the gala a proud parent claims that girls are faster swimmers than boys. Assuming that the times taken by the competitors are two independent random samples from normal distributions,

a test, at the 10% level of significance, whether or not the variances of the two distributions are the same. State your hypotheses clearly.

b Stating your hypotheses clearly, test the parent's claim. Use a 5% level of significance. **[E]**

Solution:

a $H_0 : \sigma_G^2 = \sigma_B^2, H_1 : \sigma_G^2 \neq \sigma_B^2,$

$s_B^2 = \frac{1}{6}(56\ 130 - 7 \times 88.9^2) = \frac{807.53}{6} = 134.6$

$s_G^2 = \frac{1}{7}(55\ 746 - 8 \times 83.1^2) = \frac{501.12}{7} = 71.58$

$\frac{s_B^2}{s_G^2} = 1.880\ldots$

critical value $F_{6,7} = 3.87$

not significant, variances are the same

b $H_0 : \mu_B = \mu_G, H_1 : \mu_B > \mu_G$

pooled estimate of variance $s^2 = \frac{6 \times 134.6 + 7 \times 71.58}{13} = 100.666153\ldots$

test statistic $t = \frac{88.9 - 83.1}{s\sqrt{\frac{1}{7} + \frac{1}{8}}} = 1.1169$

critical value $t_{13}(5\%) = 1.771$

Insufficient evidence to support parent's claim

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 4

Question:

Two methods of extracting juice from an orange are to be compared. Eight oranges are halved. One half of each orange is chosen at random and allocated to Method A and the other half is allocated to Method B. The amounts of juice extracted, in ml, are given in the table.

	Orange							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Method A	29	30	26	25	26	22	23	28
Method B	27	25	28	24	23	26	22	25

One statistician suggests performing a two-sample t-test to investigate whether or not there is a difference between the mean amounts of juice extracted by the two methods.

a Stating your hypotheses clearly and using a 5% significance level, carry out this test. (You may assume $\bar{x}_A = 26.125, s_A^2 = 7.84$, $\bar{x}_B = 25, s_B^2 = 4$ and $\sigma_A^2 = \sigma_B^2$)

Another statistician suggests analysing these data using a paired t-test.

b Using a 5% significance level, carry out this test.

c State which of these two tests you consider to be more appropriate. Give a reason for your choice. **[E]**

Solution:

a $$s_p^2 = \frac{7\times 7.84 + 7\times 4}{7+7} = 5.92$$

$$s_p = 2.433105$$

$$\mathrm{H}_0 : \mu_\mathrm{A} = \mu_\mathrm{B}, \mathrm{H}_1 : \mu_\mathrm{A} \neq \mu_\mathrm{B}$$

$$t = \frac{26.125 - 25}{2.43\sqrt{\frac{1}{8} + \frac{1}{8}}} = 0.92474$$

$$t_{14}(2.5\%) = 2.145$$

Insufficient evidence to reject H_0

Conclude that there is no difference in the means.

b $$d = 2, 5, -2, 1, 3, -4, 1, 3$$

$$\overline{d} = \frac{9}{8} = 1.125$$

$$s_d^2 = \frac{69 - 8\times 1.125^2}{7} = 8.410714$$

$$\mathrm{H}_0 : \delta = 0, \mathrm{H}_1 : \delta \neq 0$$

$$t = \frac{1.125}{\sqrt{\frac{8.41}{8}}} = 1.0972$$

$$t_7(2.5\%) = 2.365$$

There is no significant evidence of a difference between method A and method B.

c Paired sample as they are two measurements on the same orange

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 5

Question:

The random variable X has an F-distribution with 8 and 12 degrees of freedom.

Find $\mathrm{P}\left(\frac{1}{5.67} < X < 2.85\right)$. **[E]**

Solution:

$$\mathrm{P}(X > 2.85) = 0.05$$

$$\mathrm{P}\left(X < \frac{1}{5.67}\right) = 0.01$$

$$\therefore \mathrm{P}\left(\frac{1}{5.67} < X < 2.85\right) = 1 - 0.05 - 0.01$$

$$= 0.94$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 6

Question:

A grocer receives deliveries of cauliflowers from two different growers, A and B. The grocer takes random samples of cauliflowers from those supplied by each grower. He measures the weight x, in grams, of each cauliflower. The results are summarised in the table below.

	Sample size	$\sum x$	$\sum x^2$
A	11	6600	3 960 540
B	13	9815	7410 579

a Show, at the 10% significance level, that the variances of the populations from which the samples are drawn can be assumed to be equal by testing the hypothesis $H_0 : \sigma_A^2 = \sigma_B^2$ against hypothesis $H_1 : \sigma_A^2 \neq \sigma_B^2$.
(You may assume that the two samples come from normal populations.)
The grocer believes that the mean weight of cauliflowers provided by B is at least 150 g more than the mean weight of cauliflowers provided by A.

b Use a 5% significance level to test the grocer's belief.

c Justify your choice of test. **[E]**

Solution:

a $S_A^2 = \frac{1}{10}\{3\,960\,540 - \frac{6600^2}{11}\} = 54.0$

$S_B^2 = \frac{1}{12}\{7\,410\,579 - \frac{9815^2}{13}\} = 21.1\dot{6}$

$H_0 : \sigma_A^2 = \sigma_B^2; H_1 : \sigma_A^2 \neq \sigma_B^2$

critical region: $F_{10,12} > 2.75$

$\frac{S_A^2}{S_B^2} = \frac{54.0}{21.1\dot{6}} = 2.55118\ldots$

Since $2.55118\ldots$ is not in the critical region we can assume that the variances are equal.

b $H_0 : \mu_B = \mu_A + 150; H_1 : \mu_B > \mu_A + 150$

CR: $t_{22}(0.05) > 1.717$

$S_p^2 = \frac{10 \times 54.0 + 12 \times 21.1\dot{6}}{22} = 36.0909$

$t = \frac{755 - 600 - 150}{\sqrt{36.0909\ldots\left(\frac{1}{11} + \frac{1}{13}\right)}} = 2.03157$

Since $2.03\ldots$ is in the critical region we reject H_0 and conclude that the mean weight of cauliflowers from B exceeds that from A by at least 50g.

c Samples from normal populations
Equal variances
Independent samples

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 7

Question:

The random variable X has a χ^2-distribution with 9 degrees of freedom.

a Find $P(2.088 < X < 19.023)$.

The random variable Y follows an F-distribution with 12 and 5 degrees of freedom.

b Find the upper and lower 5% critical values for Y. **[E]**

Solution:

a

$$P(X > 19.023) = 0.025$$
$$P(X > 2.088) = 0.990$$
$$P(2.088 < X < 19.023) = 0.990 - 0.025$$
$$= 0.965$$

b

$$\text{Upper critical value of } F_{12,5} = 4.68$$
$$\text{Lower critical value of } F_{12,5} = \frac{1}{F_{5,12}}$$
$$= \frac{1}{3.11}$$
$$= 0.3215\ldots$$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 8

Question:

A group of 10 technology students is assessed by coursework and a written examination. The marks, given as percentages, are given in the table below.

Student	Coursework	Written exam.
1	65	61
2	73	76
3	62	65
4	81	77
5	78	72
6	74	71
7	68	72
8	59	42
9	76	69
10	70	63

a Use a suitable t-test to determine whether or not the coursework marks are significantly higher than the written examination marks. Use a 5% level of significance.

b State an assumption about the distribution of marks that is needed to make the above test valid. **[E]**

Solution:

a $d =$ coursework $-$ written: $4, -3, -3, 4, 6, 3, -4, 17, 7, 7$

$$\bar{d} = \frac{38}{10} = 3.8,\ s_d^2 = \frac{498 - 10\bar{d}^2}{9} = 39.2\dot{8}$$

test statistic: $t = \dfrac{3.8}{\frac{s_d}{\sqrt{10}}} = 1.917\ldots$

$H_0 : \mu_d = 0 \quad H_1 : \mu_d > 0$

$t_9(5\%)$ c.v. is 1.833;

$\therefore$ significant $-$ there is evidence coursework marks are higher

b The difference between the marks follows a normal distribution.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 9

Question:

An engineer decided to investigate whether or not the strength of rope was affected by water. A random sample of 9 pieces of rope was taken and each piece was cut in half. One half of each piece was soaked in water over night, and then each piece of rope was tested to find its strength. The results, in coded units, are given in the table below

Rope number	1	2	3	4	5	6	7	8	9
Dry rope	9.7	8.5	6.3	8.3	7.2	5.4	6.8	8.1	5.9
Wet rope	9.1	9.5	8.2	9.7	8.5	4.9	8.4	8.7	7.7

Assuming that the strength of rope follows a normal distribution, test whether or not there is any difference between the mean strengths of dry and wet rope. State your hypotheses clearly and use a 1% level of significance. **[E]**

Solution:

$D = \text{dry} - \text{wet} \qquad \text{H}_0 : \mu_D = 0, \text{H}_1 : \mu_D \neq 0$

$d : 0.6, -1, -1.9, -1.4, -1.3, 0.5, -1.6, -0.6, -1.8$

$\bar{d} : -\frac{8.5}{9} = -0.9\dot{4}, \quad s_d^2 = \frac{15.03 - 9 \times (\bar{d})^2}{8} = 0.87527\ldots$

$t = \frac{-0.9\dot{4}}{\frac{s_d}{\sqrt{9}}} = \text{awrt} -3.03$

t_8 2-tail 1% critical value $= 3.355$

Not significant – insufficient evidence of a difference between mean strength

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 10

Question:

An educational researcher is testing the effectiveness of a new method of teaching a topic in Mathematics. A random sample of 10 children were taught by the new method and a second random sample of 9 children, of similar age and ability, were taught by the conventional method. At the end of the teaching, the same test was given to both groups of children.

The marks obtained by the two groups are summarised in the table below.

	New method	Conventional method
Mean ($\bar{x}$)	82.3	78.2
Standard deviation (s)	3.5	5.7
Number of students (n)	10	9

a Stating your hypotheses clearly and using a 5% level of significance, investigate whether or not

i the variance of the marks of children taught by the conventional method is greater than that of children taught by the new method,

ii the mean score of children taught by the conventional method is lower than the mean score of those taught by the new method.

[In each case you should give full details of the calculation of the test statistics.]

b State any assumptions you made in order to carry out these tests.

c Find a 95% confidence interval for the common variance of the marks of the two groups. **[E]**

Solution:

a **i** $H_0: \sigma_C^2 = \sigma_N^2, H_1: \sigma_C^2 > \sigma_N^2$

$\frac{s_C^2}{s_N^2} = \frac{5.7^2}{3.5^2} = 2.652\ldots$; $F_{8,9}$ (5%) critical value $= 3.23$

Not significant so do not reject H_0.

There is insufficient evidence that variance using conventional method is greater

ii $H_0: \mu_N = \mu_C, H_1: \mu_N > \mu_C$

$s^2 = \frac{8 \times 5.7^2 + 9 \times 3.5^2}{17} = \frac{370.17}{17} = 21.774\ldots$

Test statistic $t = \frac{82.3 - 78.2}{\sqrt{21.774\ldots\left(\frac{1}{9}+\frac{1}{10}\right)}} = 1.9122\ldots$

t_{17} (5%) 1-tail critical value $= 1.740$

Significant $-$ reject H_0.

There is evidence that new style leads to an increase in mean

b Assumed population of marks obtained were normally distributed

c Unbiased estimate of common variance is s^2 in **ii**

$7.564 < \frac{17s^2}{\sigma^2} < 30.191$

$\sigma^2 > \frac{17 \times 21.774\ldots}{30.191} = 12.3$ (1 d.p.)

$\sigma^2 < \frac{17 \times 21.774\ldots}{7.564} = 48.9$ (1 d.p.)

Confidence interval on σ^2 is (12.3, 48.9)

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 11

Question:

Brickland and Goodbrick are two manufacturers of bricks. The lengths of the bricks produced by each manufacturer can be assumed to be normally distributed. A random sample of 20 bricks is taken from Brickland and the length, x mm, of each brick is recorded. The mean of this sample is 207.1 mm and the variance is 3.2 mm^2.

a Calculate the 98% confidence interval for the mean length of brick from Brickland.

A random sample of 10 bricks is selected from those manufactured by Goodbrick. The length of each brick, y mm, is recorded. The results are summarised as follows.

$\sum y = 2046.2 \quad \sum y^2 = 418\,785.4$

The variances of the length of brick for each manufacturer are assumed to be the same.

b Find a 90% confidence interval for the value by which the mean length of brick made by Brickland exceeds the mean length of brick made by Goodbrick. **[E]**

Solution:

a Confidence interval is given by

$$\bar{x} \pm t_{19} \times \frac{s}{\sqrt{n}}$$

i.e. $207.1 \pm 2.539 \times \sqrt{\frac{3.2}{20}}$

i.e. 207.1 ± 1.0156

i.e. $(206.08\ldots, 208.1156)$

b $\bar{x}_G = \frac{2046.2}{10} = 204.62$

$$S_p^2 = \frac{19 \times 3.2 + 9 \times 10.217\dot{3}}{28}$$

$= 5.45557\ldots$

Confidence interval is given by

$$\bar{x}_B - \bar{x}_G \pm t_{21} \times \sqrt{5.45557\left(\frac{1}{20} + \frac{1}{10}\right)}$$

i.e. $(207.1 - 204.62) \pm 1.701\sqrt{5.45557\left(\frac{1}{20} + \frac{1}{10}\right)}$

i.e. 2.48 ± 1.53875

i.e. $(0.94125, 4.0187)$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 12

Question:

The weights, in grams, of apples are assumed to follow a normal distribution.
The weights of apples sold by a supermarket have variance σ_s^2. A random sample of 4 apples from the supermarket had weights

114, 110, 119, 123.

a Find a 95% confidence interval for σ_s^2.

The weights of apples sold on a market stall have variance σ_M^2. A second random sample of 7 apples was taken from the market stall. The sample variance s_M^2 of the apples was 318.8.

b Stating your hypotheses clearly test, at the 1% level of significance, whether or not there is evidence that $\sigma_M^2 > \sigma_s^2$.

Solution:

a $\left(\bar{x} = \frac{466}{4} = 116.5\right)$ $s_x^2 = \frac{54\,386 - 4\bar{x}^2}{3} = 32.\dot{3}$ or $\frac{97}{3}$

$0.216 < \frac{3s_x^2}{\sigma^2} < 9.348$

$10.376\ldots < \sigma^2 < 449.07\ldots$ so confidence interval is (10.376, 449.07)

b $\mathrm{H}_0 : \sigma_M^2 = \sigma_s^2$ $\quad \mathrm{H}_1 : \sigma_M^2 > \sigma_s^2$

$\frac{S_M^2}{S_S^2} = \frac{318.8}{32.\dot{3}} = 9.859\ldots$

$\mathrm{F}_{6,3}$(1% c.v.) = 27.91

$9.15 < 27.91$, insufficient evidence of an increase in variance

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 13

Question:

As part of an investigation into the effectiveness of solar heating, a pair of houses was identified where the mean weekly fuel consumption was the same. One of the houses was then fitted with solar heating and the other was not. Following the fitting of the solar heating, a random sample of 9 weeks was taken and the table below shows the weekly fuel consumption for each house.

Week	1	2	3	4	5	6	7	8	9
Without solar heating	19	19	18	14	6	7	5	31	43
With solar heating	13	22	11	16	14	1	0	20	38

Units of fuel used per week

a Stating your hypotheses clearly, test, at the 5% level of significance, whether or not there is evidence that the solar heating reduces the mean weekly fuel consumption.

b State an assumption about weekly fuel consumption that is required to carry out this test. **[E]**

Solution:

a $H_0 : \mu_d = 0$

$H_1 : \mu_d > 0$

where d = without solar heating − with solar heating

$d = 6 \quad -3 \quad 7 \quad -2 \quad -8 \quad 6 \quad 5 \quad 11 \quad 5$

$\bar{d} = 3$

$s_d = 6$

$n_d = 9$

$\therefore$ test statistic $= \dfrac{(3-0)}{\left(\dfrac{6}{\sqrt{9}}\right)}$

t.s. $= 1.5$

critical value $= t_8(5\%) = 1.860$

so critical region: $t > 1.860$

Test statistic not in critical region so accept H_0. Conclude there is insufficient evidence that solar heating reduces mean weekly fuel consumption.

b The differences are normally distributed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 14

Question:

A large number of students are split into two groups A and B. The students sit the same test but under different conditions. Group A has music playing in the room during the test, and group B has no music playing during the test. Small samples are then taken from each group and their marks recorded. The marks are normally distributed.
The marks are as follows:

Sample from Group A	42	40	35	37	34	43	42	44	49
Sample from Group B	40	44	38	47	38	37	33		

a Stating your hypotheses clearly, and using a 10% level of significance, test whether or not there is evidence of a difference between the variances of the marks of the two groups.

b State clearly an assumption you have made to enable you to carry out the test in part **a**.

c Use a two-tailed test, with a 5% level of significance, to determine if the playing of music during the test has made any difference in the mean marks of the two groups. State your hypotheses clearly.

d Write down what you can conclude about the effect of music on a student's performance during the test. **[E]**

Solution:

a $H_1: \sigma_A^2 = \sigma_B^2 \quad H_0: \sigma_A^2 \neq \sigma_B^2$

$s_A^2 = 22.5 \quad s_B^2 = 21.6$

$\frac{s_A^2}{s_B^2} = 1.04$

$F_{(8,6)} = 4.15$

$1.04 < 4.15$ do not reject H_0. The variances are the same.

b Assume the samples are selected at random (independent)

c $s_p^2 = \frac{8(22.5)+6(21.62)}{14} = 22.12$

$H_0: \mu_A = \mu_B \quad H_1: \mu_A \neq \mu_B$

$t = \frac{40.667 - 39.57}{\sqrt{22.12}\sqrt{\frac{1}{9}+\frac{1}{7}}} = 0.462$

Critical value $= t_{14}(2.5\%) = 2.145$

$0.462 < 2.145$ No evidence to reject H_0

The means are the same.

d Music has no effect on performance

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 15

Question:

A company undertakes investigations to compare fuel consumption x, in miles per gallon, of two different cars the *Relaxant* and the *Elegane*, with a view to purchasing a number of cars. A random sample of 13 *Relaxants* and an independent random sample of 7 *Eleganes* were taken and the following statistics calculated.

Car	Sample size n	Sample mean $\bar{x}$	Sample variance s^2
Relaxant	13	32.31	14.48
Elegane	7	28.43	35.79

The company assumes that fuel consumption for each make of car follows a normal distribution.

a Stating your hypotheses clearly test, at the 10% level of significance, whether or not the two distributions have the same variance.

b Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is a difference in mean fuel consumption between the two types of car.

c Explain the importance of the conclusion to the test in part **a** in justifying the use of the test in part **b**.

d State two factors which might be considered when undertaking an investigation into fuel consumption of two models of car to ensure that a fair comparison is made. **[E]**

Solution:

a $H_0: \sigma_R^2 = \sigma_F^2 \quad H_1: \sigma_R^2 \neq \sigma_F^2$

$F_{6,12}(5\%)_{1\text{tail}}\,\text{cv} = 3.00, \quad \frac{s_F^2}{s_R^2} = \frac{35.79}{14.48} = 2.4716\ldots$

Not significant so do not reject H_0

Insufficient evidence to suspect $\sigma_R^2 \neq \sigma_F^2$

b $H_0: \mu_R = \mu_F \quad H_1: \mu_R \neq \mu_F$

$$s^2 = \frac{6 \times 35.79 + 12 \times 14.48}{18} = 21.58\dot{3}$$

$$t = \frac{32.31 - 28.43}{s\sqrt{\frac{1}{13} + \frac{1}{7}}} = 1.78146\ldots$$

$t_{18}(5\%)_{2\text{tail}}\,\text{cv} = 2.101$

$\therefore$ Not significant

Insufficient evidence of difference in mean performance

c Test in **b** requires $\sigma_1^2 = \sigma_2^2$

d for example, same type of driving
same roads and journey length
same weather
same driver

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Review Exercise 2
Exercise A, Question 16

Question:

A beach is divided into two areas A and B. A random sample of pebbles is taken from each of the two areas and the length of each pebble is measured. A sample of size 26 is taken from area A and the unbiased estimate for the population variance is $s_A^2 = 0.495\text{ mm}^2$. A sample of size 25 is taken from area B and the unbiased estimate for the population variance is $s_B^2 = 1.04\text{ mm}^2$.

a Stating your hypotheses clearly test, at the 10% significance level, whether or not there is a difference in variability of pebble length between area A and area B.

b State the assumption you have made about the populations of pebble lengths in order to carry out the test. **[E]**

Solution:

a $\mathrm{H_0}: \sigma_A^2 = \sigma_B^2, \mathrm{H_1}: \sigma_A^2 \neq \sigma_B^2$

critical value $F_{24,25} = 1.96$

$\frac{s_B^2}{s_A^2} = 2.10$

Since 2.10 is in the critical region we reject $\mathrm{H_0}$ and conclude there is evidence that the two variances are different.

b The populations of pebble lengths are normal.