Confidence intervals [53 marks]

1. [Maximum mark: 5] SPM.1.AHL.TZ0.9

A manager wishes to check the mean mass of flour put into bags in his factory. He randomly samples 10 bags and finds the mean mass is 1.478 kg and the standard deviation of the sample is 0.0196 kg.

(a) Find s_{n-1} for this sample. [2]

Markscheme

$$s_{n-1} = \sqrt{\frac{10}{9}} \times 0.0196 = 0.02066\ldots$$ ***(M1)A1***

[2 marks]

(b) Find a 95 % confidence interval for the population mean, giving your answer to 4 significant figures. [2]

Markscheme

(1.463, 1.493) ***(M1)A1***

Note: If s_n used answer is (1.464, 1.492), award ***M1A0***.

[2 marks]

(c) The bags are labelled as being 1.5 kg mass. Comment on this claim with reference to your answer in part (b). [1]

Markscheme

95 % of the time these results would be produced by a population with mean of less than 1.5 kg, so it is likely the mean mass is less than 1.5 kg ***R1***

[1 mark]

2. [Maximum mark: 6] 23M.1.AHL.TZ2.15

A random sample of eight packets of Apollo coffee granules are selected from a supermarket shelf.

The weights of the coffee granules present in each packet are as follows:

$$222\,\text{g} \quad 226\,\text{g} \quad 221\,\text{g} \quad 228\,\text{g} \quad 227\,\text{g} \quad 225\,\text{g} \quad 222\,\text{g} \quad 223\,\text{g}$$

(a.i) Find an unbiased estimate for the mean weight of coffee granules in a packet of Apollo coffee. [1]

Markscheme

$224\,\text{g}\ (224.25\,\text{g})$ ***A1***

[1 mark]

(a.ii) Calculate a 95% confidence interval for the population mean. Give your answer to four significant figures. [2]

Markscheme

$[221.1,\ 226.4]$ ***A1A1***

Note: Award ***A1*** for each correct end of the interval. Accept open or closed (weak or strict) interval notation. Inequalities involving μ would also be accepted, but not involving $\overline{x}$.

Award ***A1A0*** for correct answers not given correct to 4 sf.

[2 marks]

(b) State one assumption you have made in order for your interval to be valid. [1]

Markscheme

EITHER

the (population) weight of granules of Apollo coffee is normally distributed. ***R1***

OR

the readings are independent ***R1***

[1 mark]

(c) The label of each packet has a description which includes the phrase: "contains 226 g of coffee granules".

Using your answer to part (a)(ii), briefly comment on the claim on the label. [2]

Markscheme

226 g lies within the confidence interval, ***R1***

so there is no evidence to dispute the claim on the label. ***A1***

Note: Do not award ***R0A1***.

[2 marks]

3. [Maximum mark: 6] 21M.1.AHL.TZ2.10

A manufacturer of chocolates produces them in individual packets, claiming to have an average of 85 chocolates per packet.

Talha bought 30 of these packets in order to check the manufacturer's claim.

Given that the number of individual chocolates is x, Talha found that, from his 30 packets, $\Sigma x = 2506$ and $\Sigma x^2 = 209\,738$.

(a) Find an unbiased estimate for the mean number (μ) of chocolates per packet. [1]

Markscheme

$\overline{x} = \frac{\Sigma x}{n} = \frac{2506}{30} = 83.5 \;\; (83.5333\ldots)$ ***A1***

[1 mark]

(b) Use the formula $s_{n-1}^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$ to determine an unbiased estimate for the variance of the number of chocolates per packet. [2]

Markscheme

$\left(s_{n-1}^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1} =\right) \frac{209738 - \frac{2506^2}{30}}{29}$ ***(M1)***

$= 13.9 \;\; (13.9126\ldots)$ ***A1***

[2 marks]

(c) Find a 95% confidence interval for μ. You may assume that all conditions for a confidence interval have been met. [2]

Markscheme

$(82.1,\ 84.9)$ $(82.1405\ldots,\ 84.9261\ldots)$ ***A2***

[2 marks]

(d) Suggest, with justification, a valid conclusion that Talha could make. [1]

Markscheme

85 is outside the confidence interval and therefore Talha would suggest that the manufacturer's claim is incorrect ***R1***

Note: The conclusion must refer back to the original claim.

Allow use of a two sided t-test giving a p-value rounding to $0.04 < 0.05$ and therefore Talha would suggest that the manufacturer's claims in incorrect.

[1 mark]

4. [Maximum mark: 12] 22M.2.AHL.TZ2.5

A geneticist uses a Markov chain model to investigate changes in a specific gene in a cell as it divides. Every time the cell divides, the gene may mutate between its normal state and other states.

The model is of the form

$$\begin{pmatrix} X_{n+1} \\ Z_{n+1} \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} X_n \\ Z_n \end{pmatrix}$$

where X_n is the probability of the gene being in its normal state after dividing for the $n\text{th}$ time, and Z_n is the probability of it being in another state after dividing for the $n\text{th}$ time, where $n \in \mathbb{N}$.

Matrix $\boldsymbol{M}$ is found to be $\begin{pmatrix} 0.94 & b \\ 0.06 & 0.98 \end{pmatrix}$.

(a.i) Write down the value of b. [1]

Markscheme

0.02 ***A1***

[1 mark]

(a.ii) What does b represent in this context? [1]

Markscheme

the probability of mutating from 'not normal state' to 'normal state' ***A1***

Note: The ***A1*** can only be awarded if it is clear that transformation is **from** the mutated state.

[1 mark]

(b) Find the eigenvalues of $\boldsymbol{M}$. [3]

Markscheme

$$\det\begin{pmatrix} 0.94-\lambda & 0.02 \\ 0.06 & 0.98-\lambda \end{pmatrix} = 0 \quad \textit{(M1)}$$

Note: Award ***M1*** for an attempt to find eigenvalues. Any indication that $\det(\boldsymbol{M} - \lambda\boldsymbol{I}) = 0$ has been used is sufficient for the ***(M1)***.

$(0.94-\lambda)(0.98-\lambda) - 0.0012 = 0$ **OR**
$\lambda^2 - 1.92\lambda + 0.92 = 0$ ***(A1)***

$\lambda = 1,\ 0.92\ \left(\frac{23}{25}\right)$ ***A1***

[3 marks]

(c) Find the eigenvectors of $\boldsymbol{M}$. [3]

Markscheme

$\begin{pmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ **OR**
$\begin{pmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0.92\begin{pmatrix} x \\ y \end{pmatrix}$ ***(M1)***

Note: Award ***M1*** can be awarded for attempting to find either eigenvector.

$0.02y - 0.06x = 0$ **OR** $0.02y + 0.02x = 0$

$\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ***A1A1***

Note: Accept any multiple of the given eigenvectors.

[3 marks]

The gene is in its normal state when $n = 0$. Calculate the probability of it being in its normal state

(d.i) when $n = 5$. [2]

Markscheme

$$\begin{pmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{pmatrix}^5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \textbf{ OR } \begin{pmatrix} 0.744 & 0.0852 \\ 0.256 & 0.915 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$ ***(M1)***

Note: Condone omission of the initial state vector for the ***M1***.

$0.744 \quad (0.744311\ldots)$ ***A1***

[2 marks]

(d.ii) in the long term. [2]

Markscheme

$$\begin{pmatrix} 0.25 \\ 0.75 \end{pmatrix}$$ ***(A1)***

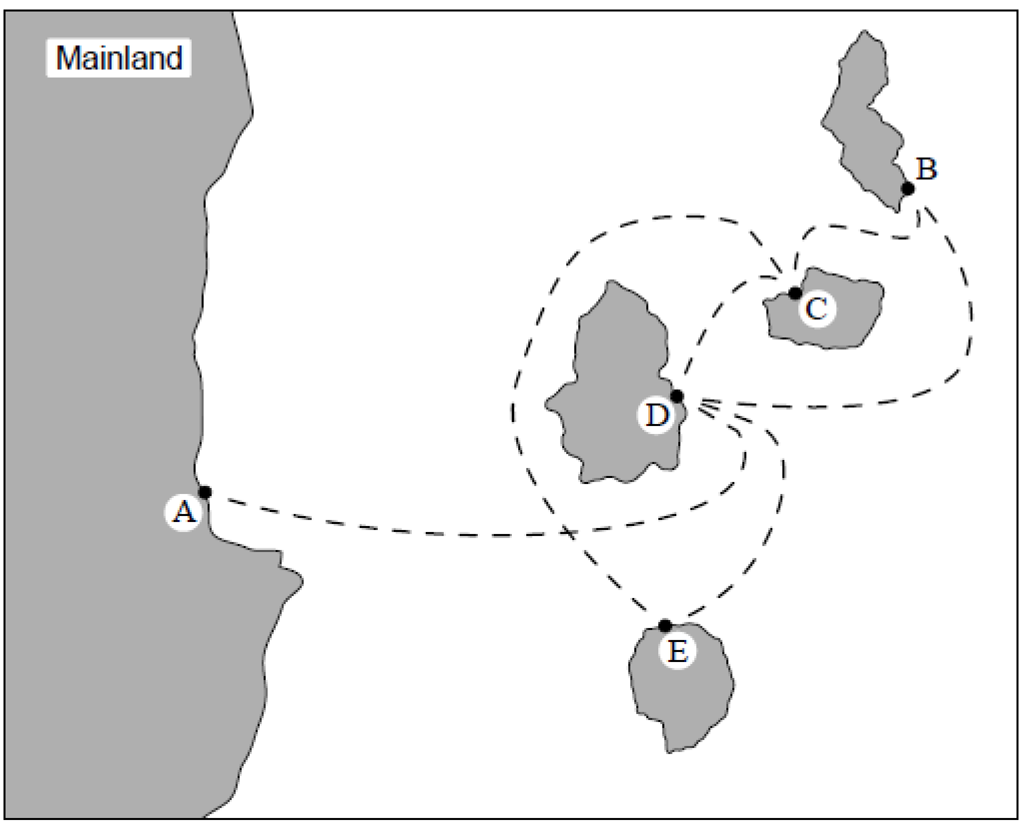
Note: Award ***A1*** for $\begin{pmatrix} 0.25 \\ 0.75 \end{pmatrix}$ **OR** $\begin{pmatrix} 0.25 & 0.25 \\ 0.75 & 0.75 \end{pmatrix}$ seen.

0.25 ***A1***

[2 marks]

5. [Maximum mark: 18] 23N.2.AHL.TZ0.5

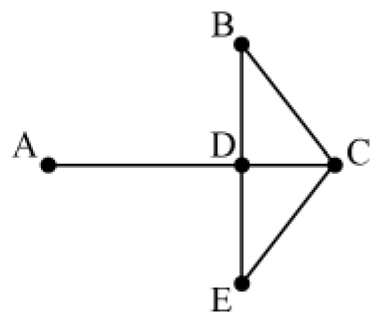
The following diagram is a map of a group of four islands and the closest mainland. Travel from the mainland and between the islands is by boat. The scheduled boat routes between the ports A, B, C, D and E are shown as dotted lines on the map.



Let the undirected graph G represent the boat routes between the ports A, B, C, D and E.

(a) Draw graph G. [1]

Markscheme



A1

[1 mark]

(b) Graph G can be represented by an adjacency matrix $\boldsymbol{P}$, where the rows and columns represent the ports in alphabetical order.

(b.i) Given that $\boldsymbol{P}^3 = \begin{pmatrix} 0 & 1 & 2 & 4 & 1 \\ 1 & 2 & 5 & a & 2 \\ 2 & 5 & 4 & 6 & 5 \\ 4 & a & 6 & 4 & 6 \\ 1 & 2 & 5 & 6 & 2 \end{pmatrix}$, find the value of a. [2]

Markscheme

$$\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$ *(M1)*

$$\boldsymbol{P}^3 = \begin{pmatrix} 0 & 1 & 2 & 4 & 1 \\ 1 & 2 & 5 & 6 & 2 \\ 2 & 5 & 4 & 6 & 5 \\ 4 & 6 & 6 & 4 & 6 \\ 1 & 2 & 5 & 6 & 2 \end{pmatrix}$$

$a = 6$ ***A1***

[2 marks]

(b.ii) Hence, write down the number of different ways that someone could start at port B and end at port C, using three boat route journeys. [1]

Markscheme

5 (routes) ***A1***

[1 mark]

(c) Find a possible Eulerian trail in G, starting at port A. [2]

Markscheme

A and C identified as start/finish points (in either order) ***(A1)***

for example: $\mathrm{A} - \mathrm{D} - \mathrm{E} - \mathrm{C} - \mathrm{D} - \mathrm{B} - \mathrm{C}$ ***A1***

[2 marks]

The cost of a journey on the different boat routes is given in the following table; all prices are given in USD. The cost of a journey is the same in either direction between two ports.

	A	B	C	D	E
A				10	
B			20	25	
C		20		50	45
D	10	25	50		30
E			45	30	

Sofia wants to make a trip where she travels on each of the boat routes at least once, beginning and ending at port A.

(d) Find the minimum cost of Sofia's trip. [3]

Markscheme

cost of their Eulerian trail A to C $(= 180)$ ***(A1)***

consider edges to get from C to A ***(M1)***

$235\,(\mathrm{USD})$ ***A1***

[3 marks]

The boat company decides to add an additional boat route to make it possible to travel on each boat route **exactly** once, starting and ending at the same port.

(e.i) Identify between which two ports the additional boat route should be added. [1]

Markscheme

A to C (or C to A) ***A1***

[1 mark]

(e.ii) Determine the cost of the additional boat route such that the overall cost of the trip is the same as your answer to part (d). [1]

Markscheme

best is CBDA

$55\,(\mathrm{USD})$ ***A1***

[1 mark]

The boat company plans to redesign which ports are connected by boat routes. Their aim is to have a single boat trip that visits all the islands and minimizes the total distance travelled, starting and finishing at the mainland, A.

The following table shows the distances in kilometres between the ports A, B, C, D and E.

	A	B	C	D	E
A		80	90	50	60
B	80		30	70	120
C	90	30		45	100
D	50	70	45		55
E	60	120	100	55	

(f.i) Use the nearest neighbour algorithm to find an upper bound for the minimum total distance. [3]

Markscheme

$\mathrm{A}-\mathrm{D}-\mathrm{C}-\mathrm{B}-\mathrm{E}-\mathrm{A}$ **OR** $50,\ 45,\ 30,\ 120,\ 60$ ***(A1)***

summing their 5 edges ***(M1)***

$50 + 45 + 30 + 120 + 60$

(upper bound $=$) $305\,(\text{km})$ ***A1***

[3 marks]

(f.ii) Use the deleted vertex algorithm on port A to find a lower bound for the minimum total distance. [4]

Markscheme

attempt to find MST without vertex A ***(M1)***

(MST $=$) 130 ***(A1)***

$130 + 50 + 60$ ***(M1)***

(lower bound $=$) $240\,(\text{km})$ ***A1***

[4 marks]

6. [Maximum mark: 6] EXN.1.AHL.TZ0.7

The position of a helicopter relative to a communications tower at the top of a mountain at time t (hours) can be described by the vector equation below.

$$\boldsymbol{r} = \begin{pmatrix} 20 \\ -25 \\ 0 \end{pmatrix} + t\begin{pmatrix} 4.2 \\ 5.8 \\ -0.5 \end{pmatrix}$$

The entries in the column vector give the displacements east and north from the communications tower and above/below the top of the mountain respectively, all measured in kilometres.

(a) Find the speed of the helicopter. [2]

Markscheme

* This sample question was produced by experienced DP mathematics senior examiners to aid teachers in preparing for external assessment in the new MAA course. There may be minor differences in formatting compared to formal exam papers.

$|v| = \sqrt{4.2^2 + 5.8^2 + 0.5^2}$ **(M1)**

$7.18\ (7.1784\ldots)\ \left(\text{kmh}^{-1}\right)$ **A1**

[2 marks]

(b) Find the distance of the helicopter from the communications tower at $t = 0$. [2]

Markscheme

$$\boldsymbol{r} = \begin{pmatrix} 20 \\ -25 \\ 0 \end{pmatrix}$$

$|\boldsymbol{r}| = \sqrt{20^2 + 25^2}$ **(M1)**

$= \sqrt{1025} = 32.0\ (32.0156\ldots)\ (\text{km})$ **A1**

[2 marks]

(c) Find the bearing on which the helicopter is travelling. [2]

Markscheme

Bearing is $\arctan\left(\frac{4.2}{5.8}\right)$ or 90°-$\arctan\left(\frac{5.8}{4.2}\right)$ **(M1)**

$035.9^\circ \quad (35.909\ldots)$ **A1**

[2 marks]