

APOLLO

Statistics and probability

HL

Mathematics: applications and interpretation

First assessment 2021

117 questions | mark schemes

1. $X \sim \text{Po}(8.8)$ **(M1)**

Note: Award **(M1)** for calculating the mean, 8.8, of the distribution

$P(X > 9) = P(X \geq 10)$ **OR** $P(X > 9) = 1 - P(X \leq 9)$ **(M1)**

$P(X > 9) = 0.386$ (0.386260...) **(M1)A1**

Note: Award **(M1)(M0)(M1)A0** for finding $P(X \geq 9) = 0.518$ (0.517719...)
OR $P(X \leq 9) = 0.614$ (0.613740...).

Total [4 marks]

8. (a)

t	1	2	3	4	5	6
$P(T=t)$	$\frac{1}{36}$ (0.027777...)	$\frac{3}{36}$ (0.083333...)	$\frac{5}{36}$ (0.138888...)	$\frac{7}{36}$ (0.194444...)	$\frac{9}{36}$ (0.25)	$\frac{11}{36}$ (0.305555...)

A2

Note: Award **A1** if three to five probabilities are correct.

[2 marks]

(b) (i) $\frac{32}{36} \left(\frac{8}{9}, 0.888888..., 88.9\% \right)$

(A1)

(ii) use of conditional probability
e.g. denominator of 32 **OR** denominator of 0.888888..., etc.
 $\frac{11}{32} (0.34375, 34.4\%)$

(M1)**A1****[3 marks]**

(c) $\frac{1 \times 1 + 3 \times 2 + 5 \times 3 + \dots + 11 \times 6}{36}$
 $= \frac{161}{36} \left(4\frac{17}{36}, 4.47, 4.47222... \right)$

(M1)**A1****[2 marks]****Total [7 marks]**

11. (a) Convenience **A1**
[1 mark]
- (b) H_0 : 1% of the toys produced are faulty **A1**
 H_1 : More than 1% are faulty **A1**
[2 marks]
- (c) $X \sim B(200, 0.01)$ **(M1)**
 $P(X \geq 4) = 0.142$ **A1**
- Note:** Any attempt using Normal approximation to find p -value is awarded **M0A0**.
- [2 marks]**
- (d) $14\% > 10\%$ **R1**
 so there is insufficient evidence to reject H_0 . **A1**
- Note:** Do not award **R0A1**. Accept “fail to reject H_0 ” or “accept H_0 ”.
- [2 marks]**
- Total [7 marks]**

14. (a) $158 \times 6 = 948(\text{g})$

(M1)A1

[2 marks]

(b) variance 6×13^2
SD = 31.8(g) ($13\sqrt{6}$, 31.8433...)

(M1)

A1

[2 marks]

(c) $X \sim N(948, 31.8433...^2)$
 $P(X > 1000) = 0.0512$ (0.0512350...)

(M1)A1

Note: Accept 0.0510 (0.0510014...) if 3 sf value 31.8 is used.

Award **(M1)A1FT** if the answer is correct for their SD, even if no working is shown. e.g. If the SD is 78 then accept 0.252.

[2 marks]

Total [6 marks]

3. (a) (let μ_c = population mean for chinchilla rabbits, μ_s = population mean for sable rabbits)

$$H_0 : \mu_c = \mu_s \quad \mathbf{A1}$$

$$H_1 : \mu_c > \mu_s \quad \mathbf{A1}$$

Note: Accept an equivalent statement in words, must include mean and reference to “**population** mean” / “mean for **all** chinchilla rabbits” for the first **A1** to be awarded. The terms “*on average*” and “*generally*” are also acceptable to indicate populations.
Do not accept an imprecise “*the means are equal*”.

[2 marks]

- (b) $p\text{-value} = 0.0408$ (0.0408065...) **A2**

Note: Award **A1** for an answer of 0.041565..., from “unpooled” settings on GDC.

[2 marks]

- (c) $0.0408 < 0.05$. **R1**

(there is sufficient evidence to) reject (or not accept) H_0 **A1**

(there is sufficient evidence to suggest that chinchilla rabbits are (generally) heavier than sable rabbits)

Note: Do not award **R0A1**. Accept ‘accept H_1 ’.

[2 marks]

Total [6 marks]

9. (a) $\left(\frac{74+97+91+86+112}{5}\right) = 92$ **A1**

[1 mark]

- (b) (i) H_0 : The data satisfies the model **A1**
 H_1 : The data does not satisfy the model **A1**

Note: Do not accept " H_0 : The same number of copies will be sold each day" but accept a similar statement if the word 'expect' or 'expected' is included. Similarly for H_1 .

(ii) 4 **A1**

(iii) $\chi^2_{\text{calc}} = 8.54$ (8.54347...) **OR** $p\text{-value} = 0.0736$ (0.0735802...) **A2**

$8.54 < 9.49$ **OR** $0.0736 > 0.05$ **R1**

therefore there is insufficient evidence to reject H_0 **A1**

(i.e. the data satisfies the model)

Note: Do not award **ROA1**. Accept "accept" or "do not reject" in place of "insufficient evidence to reject".

Award the **R1** for comparing their p -value with 0.05 or their χ^2 value with 9.49 and then **FT** their final conclusion.

[7 marks]

Total [8 marks]

1

10. (a) $\bar{x} = \frac{\sum x}{n} = \frac{2506}{30} = 83.5 \text{ (83.5333...)}$

A1**[1 mark]**

(b)
$$s_{n-1}^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} = \frac{209738 - \frac{2506^2}{30}}{29}$$

 $= 13.9 \text{ (13.9126...)}$

(M1)**A1****[2 marks]**

(c) (82.1, 84.9) (82.1405..., 84.9261...)

A2**[2 marks]**

- (d) 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]**Total [6 marks]**

13.

(a) transition matrix is

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	0	$\frac{1}{3}$	$\frac{1}{2}$	0	0	0
<i>B</i>	$\frac{1}{3}$	0	0	0	0	$\frac{1}{5}$
<i>C</i>	0	$\frac{2}{3}$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{5}$
<i>D</i>	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{2}{5}$
<i>E</i>	$\frac{1}{3}$	0	0	0	0	$\frac{1}{5}$
<i>F</i>	$\frac{1}{3}$	0	0	$\frac{1}{2}$	0	0

M1A1A1

Note: Allow the transposed matrix.

Award **M1** for a 6x6 matrix with all values between 0 and 1, and all columns (or rows if transposed) adding up to 1, award **A1** for one correct row (or column if transposed) and **A1** for all rows (or columns if transposed) correct.

[3 marks]

(b) attempting to raise the transition matrix to a large power **(M1)**

steady state vector is

(0.157)
(0.0868)
(0.256)
(0.241)
(0.0868)
0.173

(A1)

so percentage of time spent at vertex F is 17.3% **A1**

Note: Accept 17.2%.

[3 marks]

(c) the model assumes instantaneous travel from junction to junction, **R1**
and hence the answer obtained would be an overestimate **R1**

OR

the mouse may eat the sugar over time **R1**

and hence the probabilities would change **R1**

Note: Accept any other sensible answer.

[2 marks]

Total [7 marks]

.....

15. (a) $H_0 : m = 110$, $H_1 : m > 110$

A1

Note: Accept other appropriate variables for the mean.
Accept 22 in place of 110.

[1 mark]

(b) $P(X \geq 128) = 0.05024$

(M1)(A1)

$P(X \geq 129) = 0.04153$

(M1)

(probability of making a type I error is) 0.0415

A1

Note: If other probabilities are seen, the final **A1** cannot be awarded unless 0.0415 is clearly identified as the final answer.

[4 marks]

(c) $X \sim \text{Po}(110)$

$P(X \geq 126) = 0.072 > 0.05$ **OR** recognizing $126 < 129$ or ≤ 128

R1

so there is insufficient evidence to reject H_0

A1

(ie there is insufficient evidence to suggest that the number of coffees being sold has increased)

Note: Accept 'Accept H_0 '.
Do not award **R0A1**.

[2 marks]

Total [7 marks]

3. (a) (i) evidence of correct probability (M1)
e.g. sketch **OR** correct probability statement $P(X < 6.5)$

0.0151 A1

(ii) 0.0228 A1

Note: Answers should be given to 4 decimal place.

[3 marks]

- (b) (i) multiplying **their** probability by 1000 (M1)
451.7 A1

(ii) 510.5 A1

[3 marks]

Note: Answers should be given to 4 sf.

- (c) H_0 : stopping distances can be modelled by $N(6.76, 0.12^2)$
 H_1 : stopping distances cannot be modelled by $N(6.76, 0.12^2)$ A1A1

Note: Award **A1** for correct H_0 , including reference to the mean and standard deviation.
Award **A1** for the negation of their H_0 .

[2 marks]

- (d) 15.1 or 22.8 seen (M1)

0.0727 (0.0726542..., 7.27%) A2

[3 marks]

- (e) $0.05 < 0.0727$ R1
there is insufficient evidence to reject H_0 (or “accept H_0 ”) A1

Note: Do not award **R0A1**.

[2 marks]

Total [13 marks]

5. (a) finding T^3 **OR** use of tree diagram (M1)

$$T^3 = \begin{pmatrix} 0.65 & 0.525 \\ 0.35 & 0.475 \end{pmatrix}$$

the probability of sunny in three days' time is 0.65

A1

[2 marks]

- (b) attempt to find eigenvalues (M1)

Note: Any indication that $\det(T - \lambda I) = 0$ has been used is sufficient for the (M1).

$$\begin{vmatrix} 0.8 - \lambda & 0.3 \\ 0.2 & 0.7 - \lambda \end{vmatrix} = (0.8 - \lambda)(0.7 - \lambda) - 0.06 = 0$$

$$(\lambda^2 - 1.5\lambda + 0.5 = 0)$$

$$\lambda = 1, \lambda = 0.5$$

A1

attempt to find either eigenvector

(M1)

$$0.8x + 0.3y = x \Rightarrow -0.2x + 0.3y = 0 \text{ so an eigenvector is } \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

A1

$$0.8x + 0.3y = 0.5x \Rightarrow 0.3x + 0.3y = 0 \text{ so an eigenvector is } \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

A1

Note: Accept multiples of the stated eigenvectors.

[5 marks]

- (c) (i) $P = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}$ **OR** $P = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$ A1

Note: Examiners should be aware that different, correct, matrices P may be seen.

$$(ii) \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix} \text{ **OR** } D = \begin{pmatrix} 0.5 & 0 \\ 0 & 1 \end{pmatrix}$$

A1

Note: P and D must be consistent with each other.

[2 marks]

- (d) $0.5^n \rightarrow 0$ (M1)

$$D^n = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ **OR** } D^n = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

(A1)

Note: Award A1 only if their D^n corresponds to their P .

$$PD^n P^{-1} = \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix}$$

(M1)

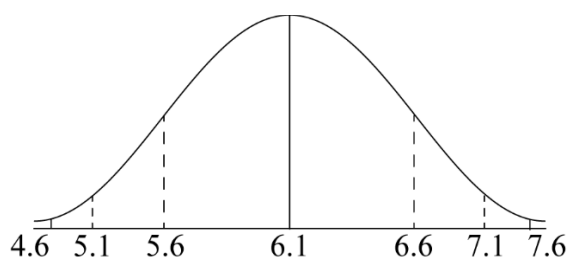
60 %

A1

[4 marks]

Total [13 marks]

2. (a)

**A1A1**

Note: Award **A1** for a normal curve with mean labelled 6.1 or μ , **A1** for indication of SD (0.5): marks on horizontal axis at 5.6 and/or 6.6 **OR** $\mu - 0.5$ and/or $\mu + 0.5$ on the correct side and approximately correct position.

[2 marks]

(b) $X \sim N(6.1, 0.5^2)$

 $P(5.5 < X < 6.5)$ **OR** labelled sketch of region**(M1)**

$= 0.673 \text{ (0.673074...)}$

A1**[2 marks]**

(c) $P(X < 5.3) = 0.0547992...$

(A1)

$0.0547992... \times 80$

(M1)

$= 4.38 \text{ (4.38393...)}$

A1**[3 marks]**

(d) (i) $Y \sim N(4.5, 0.45^2)$,

$P(Y > 4.62) = 0.394862...$

(A1)

use of binomial seen or implied

(M1)using $B(10, 0.394862...)$ **(M1)**

$0.0430 \text{ (0.0429664...)}$

A1

(ii) $np(1-p) = 2.39 \text{ (2.38946...)}$

A1**[5 marks]**

(e) $P(F \cap (W > 4.7)) = 0.5 \times 0.3284 (= 0.1642)$

(A1)attempt use of tree diagram **OR** use of $P(F | W > 4.7) = \frac{P(F \cap (W > 4.7))}{P(W > 4.7)}$ **(M1)**

$$\frac{0.5 \times 0.3284}{0.5 \times 0.9974 + 0.5 \times 0.3284}$$

(A1)

$= 0.248 \text{ (0.247669...)}$

A1**[4 marks]****Total [16 marks]**

4. (a) $T = \begin{pmatrix} 0.965 & 0.05 \\ 0.035 & 0.95 \end{pmatrix}$ **M1A1**

Note: Award **M1A1** for $T = \begin{pmatrix} 0.95 & 0.035 \\ 0.05 & 0.965 \end{pmatrix}$.

Award the **A1** for a transposed T if used correctly in part (b) i.e. preceded by 1×2 matrix (2100 3500) rather than followed by a 2×1 matrix.

[2 marks]

(b) $\begin{pmatrix} 0.965 & 0.05 \\ 0.035 & 0.95 \end{pmatrix}^2 \begin{pmatrix} 2100 \\ 3500 \end{pmatrix}$ **(M1)**
 $= \begin{pmatrix} 2294 \\ 3306 \end{pmatrix}$

so ratio is 2294 : 3306 (= 1147 : 1653, 0.693889...)

A1

[2 marks]

(c) to solve $Ax = \lambda x$:

$$\begin{vmatrix} 0.965 - \lambda & 0.05 \\ 0.035 & 0.95 - \lambda \end{vmatrix} = 0$$
 (M1)

$$(0.965 - \lambda)(0.95 - \lambda) - 0.05 \times 0.035 = 0$$

$$\lambda = 0.915 \quad \text{OR} \quad \lambda = 1$$
 (A1)

attempt to find eigenvectors for at least one eigenvalue

(M1)

when $\lambda = 0.915$, $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (or any real multiple) **(A1)**

when $\lambda = 1$, $x = \begin{pmatrix} 10 \\ 7 \end{pmatrix}$ (or any real multiple) **(A1)**

therefore $P = \begin{pmatrix} 1 & 10 \\ -1 & 7 \end{pmatrix}$ (accept integer valued multiples of their eigenvectors and

columns in either order)

A1

[6 marks]

continued...

Question 4 continued

$$(d) \quad P^{-1} = \begin{pmatrix} 1 & 10 \\ -1 & 7 \end{pmatrix}^{-1} = \frac{1}{17} \begin{pmatrix} 7 & -10 \\ 1 & 1 \end{pmatrix} \quad (A1)$$

Note: This mark is independent, and may be seen anywhere in part (d).

$$D = \begin{pmatrix} 0.915 & 0 \\ 0 & 1 \end{pmatrix} \quad (A1)$$

$$T^n = PD^nP^{-1} = \begin{pmatrix} 1 & 10 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} 0.915^n & 0 \\ 0 & 1^n \end{pmatrix} \frac{1}{17} \begin{pmatrix} 7 & -10 \\ 1 & 1 \end{pmatrix} \quad (M1)A1$$

Note: Award **(M1)A0** for finding $P^{-1}D^nP$ correctly.

$$\text{as } n \rightarrow \infty, D^n = \begin{pmatrix} 0.915^n & 0 \\ 0 & 1^n \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad R1$$

$$\text{so } T^n \rightarrow \frac{1}{17} \begin{pmatrix} 1 & 10 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 7 & -10 \\ 1 & 1 \end{pmatrix} \quad A1$$

$$= \begin{pmatrix} \frac{10}{17} & \frac{10}{17} \\ \frac{7}{17} & \frac{7}{17} \end{pmatrix} \quad AG$$

Note: The **AG** line must be seen for the final **A1** to be awarded.

[6 marks]

(e) **METHOD ONE**

$$\begin{pmatrix} \frac{10}{17} & \frac{10}{17} \\ \frac{7}{17} & \frac{7}{17} \end{pmatrix} \begin{pmatrix} 2100 \\ 3500 \end{pmatrix} = \begin{pmatrix} 3294 \\ 2306 \end{pmatrix} \quad (M1)$$

so ratio is 3294 : 2306 (1647 : 1153, 1.42844..., 0.700060...)

A1

METHOD TWO

long term ratio is the eigenvector associated with the largest eigenvalue
10 : 7

(M1)

A1

[2 marks]

Total [18 marks]

5. (a) $X_1 \sim \text{Po}(3.1)$

$P(X_1 = 4) = 0.173 \text{ (0.173349...)}$

A1**[1 mark]**

(b) (i) $X_2 \sim \text{Po}(3 \times 3.1) = \text{Po}(9.3)$

(M1)

$P(X_2 = 12) = 0.0799 \text{ (0.0798950...)}$

A1

(ii) $(P(X_1 > 0))^2 \times P(X_1 = 0)$

(M1)

$0.95495^2 \times 0.04505$

(A1)

$= 0.0411 \text{ (0.0410817...)}$

A1**[5 marks]**

(c) $P(X_1 = 0) = 0.04505$

(A1)

$X_1 \sim B(12, 0.04505)$

(M1)(A1)**Note:** Award **M1** for recognizing binomial probability, and **A1** for correct parameters.

$= 0.0133 \text{ (0.013283....)}$

A1**[4 marks]**(d) **METHOD ONE**

n	λ	$P(X \geq 30)$
...	...	...
10	24.1	0.136705
11	26.2	0.253384

(M1)(A1)(A1)**Note:** Award **M1** for evidence of a cumulative Poisson with $\lambda = 3.1 + 2.1n$, **A1** for 0.136705 and **A1** for 0.253384.

so require 12 magpies (including Bill)

A1**METHOD TWO**evidence of a cumulative Poisson with $\lambda = 3.1 + 2.1n$ **(M1)**sketch of curve and $y = 0.2$ **(A1)**

(intersect at) 10.5810...

(A1)rounding up gives $n = 11$

so require 12 magpies (including Bill)

A1**[4 marks]****Total [14 marks]**

2. (a) Use of χ^2 test for independence (M1)

H_0 : Staying (or leaving) the firm and interview rating are independent.

H_1 : Staying (or leaving) the firm and interview rating are not independent A1

Note: For H_1 accept '...are dependent' in place of '...not independent'.

$p\text{-value} = 0.487$ (0.487221...) A2

Note: Award **A1** for $\chi^2 = 1.438...$ if $p\text{-value}$ is omitted or incorrect.

$0.487 > 0.05$ R1

(the result is not significant at the 5% level)

insufficient evidence to reject the H_0 (or "accept H_0 ") A1

Note: Do not award **R0A1**.
The final **R1A1** can follow through from their incorrect $p\text{-value}$

[6 marks]

(b) $\frac{55}{91} \times 18 = 10.9$ (10.8791...) M1A1

Note: Award **A1** for anything that rounds to 10.9.

≈ 11 AG

[2 marks]

continued...

Question 2 continued

- (c) (i) there seems to be a difference between the two departments
- (A1)**

the international department manager seems to be less generous
than the national department manager **R1**

Note: The **A1** is for commenting there is a difference between the two departments
and the **R1** is for correctly commenting on the direction of the difference

[2 marks]

(ii)

	L	M	N	O	P	Q	R
Written assessment rank	1	2	3	4	5	6	7
Manager score rank	1	2.5	4.5	2.5	4.5	6	7

(M1)(A1)

Note: Award **(M1)** for an attempt to rank the data, and **(A1)** for correct ranks
for both variables. Accept either set of rankings in reverse.

$$r_s = 0.909 \text{ (0.909241.....)}$$

(M1)A1

Note: The **(M1)** is for calculating the PMCC for their ranks.

Note: If a final answer of 0.9107 is seen, from use of $1 - \frac{6\sum d^2}{n(n^2 - 1)}$, award **(M1)(A1)A1**.
Accept -0.909 if one set of ranks has been ordered in reverse.

[4 marks](iii) **EITHER**

there is a (strong) association between the written assessment mark and the
manager scores. **A1**

OR

there is a (strong) agreement in the rank order of the written assessment marks and
the rank order of the manager scores. **A1**

OR

there is a (strong linear) correlation between the rank order of the written assessment
marks and the rank order of the manager scores. **A1**

Note: Follow through on a value for their value of r_s in c(ii).

THEN

the written assessment is likely to be a valid measure (of the level
of employee performance)

R1**[2 marks]**

continued...

Question 2 continued

(d) (i) test-retest

A1

[1 mark]

(ii) $p\text{-value} = 0.00209$ (0.0020939...)

A2

$$0.00209 < 0.05$$

R1

(the result is significant at the 5% level)

(there is sufficient evidence to) reject H_0

A1

Note: Do not award **R0A1**. Accept “accept H_1 ”.
The final **R1A1** can follow through from their incorrect p -value.

[4 marks]

(iii) the test seems reliable

A1

Note: Follow through from their answer in part (d)(ii). Do not award if there is no conclusion in d(ii).

[1 mark]

(e) (i) 25

A1

[1 mark]

(ii) probability of significant result given no correlation is 0.05

(M1)

probability of at least one significant result in 25 tests is

$$1 - 0.95^{25}$$

(M1)(A1)

Note: Award **(M1)** for use of $1 - P(0)$ or the binomial distribution with any value of p .

[1 mark]

$$= 0.723 \text{ (0.722610...)}$$

A1

[4 marks]

(iii) (though the result is significant) it is very likely that one significant result would be achieved by chance, so it should be disregarded or further evidence sought

R1

[1 mark]

[Total 28 marks]

1. (a) (i) Any one from:
 increase sample size / increase response rate / repeat process
 check whether sample is representative
 test-retest participants or do a parallel test
 use a stratified sample
 use a random sample

R1

Note: Do not condone:
 Ask different types of doctor
 Ask for proof of income
 Ask for proof of being a doctor
 Remove anonymity
 Remove response K.

[1 mark]

- (ii) Any one from:
 non-random sampling means a subset of population might be responding
 self-reported happiness is not the same as happiness
 happiness is not a constant / cannot be quantified / is difficult to measure
 income might include external sources
 Juliet is only sampling doctors in her city
 correlation does not imply causation
 sample might be biased

R1

Note: Do not condone the following common but vague responses unless they make a clear link to validity:
 Sample size is too small
 Result is not generalizable
 There may be other variables Juliet is ignoring
 Sample might not be representative

[1 mark]

- (b) because the income is very different / implausible / clearly contrived

R1

Note: Answers must explicitly reference "income" to get credit.

[1 mark]

continued...

Question 1 continued

- (c) (i) (\$) 90 200 (M1)A1 [2 marks]
- (ii) $r = 0.558$ (0.557723...) A2 [2 marks]
- (d) (i) **EITHER**
only looking for change in one direction R1
OR
only looking for greater happiness with greater income R1
OR
only looking for evidence of positive correlation R1
- (ii) $H_0: \rho = 0; H_1: \rho > 0$ A1A1 [1 mark]

Note: Award **A1** for ρ seen (do not accept r), **A1** for both correct hypotheses, using **their** ρ or r . Accept an equivalent statement in words, however reference to “correlation for the population” or “association for the population” must be explicit for the first **A1** to be awarded.

Watch out for a null hypothesis in words similar to “Annual income is not associated with greater happiness”. This is effectively saying $\rho \leq 0$ and should not be condoned.

[2 marks]

continued...

Question 1 continued

- (iii) **METHOD 1 – using critical value of r**
 $0.558 > 0.549$ ($0.557723... > 0.549$) **R1**
 (therefore significant evidence of) a positive correlation **A1**

Note: Do not award **R0A1**.

METHOD 2 – using p -value

$$0.0469 < 0.05 \quad (0.0469463... < 0.05) \quad \text{A1}$$

Note: Follow through from their r -value from part (c)(ii).

(therefore significant evidence of) a positive correlation **A1**

Note: Do not award **A0A1**.

[2 marks]

- (e) (i) $a = 0.000126$ ($0.000125842...$), $b = 41.1$ ($41.1490...$) **A1**

[1 mark]

- (ii) **EITHER**
 the amount the happiness score increases for every \$1 increase in (annual) income **A1**
OR
 rate of change of happiness with respect to (annual) income **A1**

Note: Accept equivalent responses e.g. an increase of 1.26 in happiness for every \$10000 increase in salary.

[1 mark]

- (iii) $c = -2.06 \times 10^{-9}$ ($-2.06191... \times 10^{-9}$),
 $d = 7.05 \times 10^{-4}$ ($7.05272... \times 10^{-4}$),
 $e = 12.6$ ($12.5878...$) **A1**

[1 mark]

continued...

Question 1 continued

(iv) for quadratic model: $R^2 = 0.659$ (0.659145...) **A1**

for linear model: $R^2 = 0.311$ (0.311056...) **A1**

Note: Follow through from their r value from part (c)(ii).

[2 marks]

(v) **EITHER**

quadratic model is a better fit to the data / more accurate **A1**

OR

quadratic model explains a higher proportion of the variance **A1**

[1 mark]

(vi) **EITHER**

not valid, R^2 not a useful measure to compare models with different numbers of parameters **A1**

OR

not valid, quadratic model will always have a better fit than a linear model **A1**

Note: Accept any other sensible critique of the validity of the method. Do not accept any answers which focus on the conclusion rather than the method of model selection.

[1 mark]

continued...

Question 1 continued

(f) (i) (single sample) t -test**A1****[1 mark]**(ii) **EITHER** $H_0 : \mu = 80\,000$; $H_1 : \mu \neq 80\,000$ **A1****OR** H_0 : (sample is drawn from a population where) the population mean is \$80 000 H_1 : the population mean is not \$80 000**A1****Note:** Do not allow **FT** from an incorrect test in part (f)(i) other than a z -test.**[1 mark]**(iii) $p = 0.610$ (0.610322...)**A1****Note:** For a z -test follow through from part (f)(i), either 0.578 (from biased estimate of variance) or 0.598 (from unbiased estimate of variance). $0.610 > 0.05$ **R1****EITHER**

no (significant) evidence that mean differs from \$ 80 000

A1**OR**the sample could plausibly have been drawn from the quoted population **A1****Note:** Allow **R1FTA1FT** from an incorrect p -value, but the final **A1** must still be in the context of the original research question.**[3 marks]****[Total 24 marks]**

7. (a) 75

A1**[1 mark]**

(b) recognition that all entries add up to 120
 $a = 120 - 6 - 13 - 26 - b$ **OR** $a = 75 - b$

(M1)**A1****[2 marks]**

(c) (i) $\frac{6 \times 1 + 13 \times 2 + 26 \times 3 + (75 - b) \times 4 + b \times 5}{120} = 3.65$

(M1)(A1)

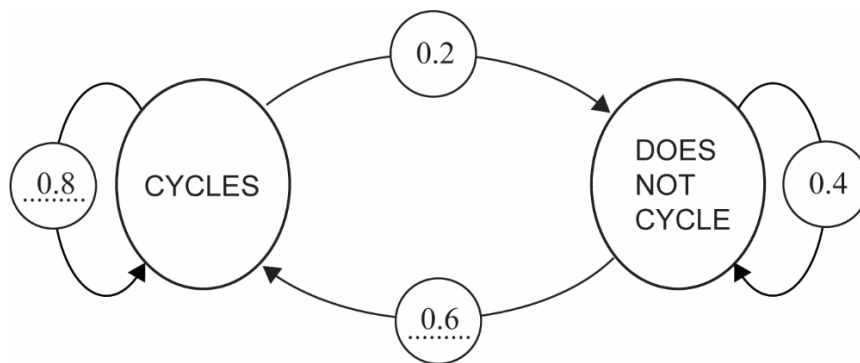
Note: Award **(M1)** for attempt to substitute into mean formula, LHS expression is sufficient for the **M** mark. Award **(A1)** for correct substitutions in one variable **OR** in two variables, followed by evidence of solving simultaneously with $a + b = 75$.

 $(b =) 28$ **A1**

(ii) 120 – their part (c)(i) seen (e.g. 92 indicated on graph)
 84

(M1)**A1****[5 marks]****Total: [8 marks]**

9. (a)



A1A1
[2 marks]

(b) $A = \begin{pmatrix} 0.8 & 0.6 \\ 0.2 & 0.4 \end{pmatrix}$

(A1)

$$A^{180} = \begin{pmatrix} 0.75 & 0.75 \\ 0.25 & 0.25 \end{pmatrix}$$

(M1)

0.75

A1
[3 marks]
Total: [5 marks]

12. (a) $m = -0.695$ ($-0.695383\dots$); $b = 4.63$ ($4.62974\dots$)

A1A1

[2 marks]

(b) $\ln x = -0.695(\ln 25) + 4.63$

M1

$\ln x = 2.39288\dots$

(A1)

$x = 10.9\%$

A1

[3 marks]

Total: [5 marks]

14. (a) let X be the random variable “the weight of a sack of potatoes”
 $P(X < 50)$
 $= 0.588 \text{ kg } (0.587929\dots)$

(M1)

A1

[2 marks]

- (b) $P(X < l) = 0.25$
 $49.2 \text{ kg } (49.1929\dots)$

(M1)

A1

[2 marks]

- (c) attempt to sum 10 independent random variables
 $Y = \sum_{i=1}^{10} X_i \sim N(498, 10 \times 0.9^2)$

(M1)

(A1)

$$P(Y > 500) = 0.241$$

A1

[3 marks]

Total: [7 marks]

3. (a) (i) Let X be the random variable “distance from O”.
 $X \sim N(10, 3^2)$
 $P(X < 13) = 0.841$ (0.841344...) (M1)A1
- (ii) $(P(X > 15) =) 0.0478$ (0.0477903) A1
 [3 marks]
- (b) $P(X > 15) \times P(X > 15)$ (M1)
 $= 0.00228$ (0.00228391...) A1
 [2 marks]
- (c) $1 - (0.8143)^3$ (M1)
 $= 0.460$ (0.460050...) A1
 [2 marks]
- (d) (i) let Y be the random variable “number of points scored”
 evidence of use of binomial distribution (M1)
 $Y \sim B(10, 0.539949...)$ (A1)
 $(E(Y) =) 10 \times 0.539949...$ (M1)
 $= 5.40$ A1
- (ii) $(P(Y \geq 5) =) 0.717$ (0.716650...) A1
- (iii) $P(5 \leq Y < 8)$ (M1)
 $= 0.628$ (0.627788...) A1
- Note:** Award **M1** for a correct probability statement or indication of correct lower and upper bounds, 5 and 7.
- (iv) $\frac{P(5 \leq Y < 8)}{P(Y \geq 5)} \left(= \frac{0.627788...}{0.716650...} \right)$ (M1)
 $= 0.876$ (0.876003...) A1
 [9 marks]
Total: [16 marks]

7. (a) (i) let X be the random variable “number of patients arriving in a minute”, such that $X \sim \text{Po}(m)$.
 $H_0 : m = 1.5$ **A1**
 $H_1 : m > 1.5$ **A1**

Note: Allow a value of 270 for m . Award at most **A0A1** if it is not clear that it is the population mean being referred to e.g
 H_0 : *The number of patients is equal to 1.5 every minute*
 H_1 : *The number of patients exceeds 1.5 every minute.*
Referring to the “expected” number of patients or the use of μ or λ is sufficient for **A1A1**.

- (ii) under H_0 let Y be the number of patients in 3 hours
 $Y \sim \text{Po}(270)$ **(A1)**
 $P(Y \geq 320) (= 1 - P(Y \leq 319)) = 0.00166$ (0.00165874) **(M1)A1**

since $0.00166 < 0.05$ **R1**
(reject H_0)
Loreto should employ more staff **A1**

[7 marks]

- (b) (i) H_0 : The probability of a patient waiting less than 20 minutes is 0.95 **A1**
 H_1 : The probability of a patient waiting less than 20 minutes is less than 0.95 **A1**

(ii) under H_0 let W be the number of patients waiting more than 20 minutes
 $W \sim B(150, 0.05)$ **(A1)**
 $P(W \geq 11) = 0.132$ (0.132215...) **(M1)A1**
since $0.132 > 0.1$ **R1**
(fail to reject H_0)
insufficient evidence to suggest they are not meeting their target **A1**

Note: Do not accept “they are meeting target” for the **A1**.
Accept use of $B(150, 0.95)$ and $P(W \leq 139)$ and any consistent use of a random variable, appropriate p -value and significance level.

[7 marks]**Total: [14 marks]**

1. (a) discrete

A1
[1 mark]

(b)
$$\frac{24 + 60 + 3k + 40 + 15 + 6}{88 + k} = 2$$

M1A1

Note: Award **M1** for substitution into the formula for the mean, award **A1** for a correct equation.

attempt to solve their equation

(M1)

$k = 31$

A1
[4 marks]

(c) systematic

A1
[1 mark]

Total [6 marks]

6. (a) $\frac{4}{18} \left(\frac{2}{9} \right)$ **A1**

[1 mark]

(b) $-3 \times \frac{1}{18} + (-1) \times \frac{4}{18} + 0 \times \frac{3}{18} + \dots + 5 \times \frac{7}{18}$ **(M1)**

Note: Award **(M1)** for their correct substitution into the formula for expected value.

$= 1.83 \left(\frac{33}{18}, 1.83333\dots \right)$ **A1**

[2 marks]

(c) $2 \times \frac{1}{18} \times \frac{3}{18}$ **(M1)(M1)**

Note: Award **(M1)** for $\frac{1}{18} \times \frac{3}{18}$, award **(M1)** for multiplying their product by 2.

$= \frac{1}{54} \left(\frac{6}{324}, 0.0185185\dots, 1.85\% \right)$ **A1**

[3 marks]

Total [6 marks]

9. (a) $s_{n-1} = \sqrt{\frac{10}{9}} \times 0.0196 = 0.02066\dots$

(M1)A1**[2 marks]**

(b) (1.463, 1.493)

(M1)A1

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

[2 marks]

- (c) 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 weight is less than 1.5 kg

R1**[1 mark]****Total [5 marks]**

- 10.** let T be the time to serve both customers and T_i the time to serve the i th customer

assuming independence of T_1 and T_2 **R1**

T is normally distributed and $T = T_1 + T_2$ **(M1)**

$$E(T) = 1.5 + 1.5 = 3$$
A1

$$\text{Var}(T) = 0.4^2 + 0.4^2 = 0.32$$
M1A1

$$P(T < 4) = 0.961$$
A1

Total [6 marks]

12. (Model A)

$$R = 3pe^{-0.5p}$$

M1

predicted values

p	R
1	1.8196
2	2.2073
3	2.0082

(A1)

$$SS_{res} = (1.8196 - 1.5)^2 + (2.2073 - 1.8)^2 + (2.0082 - 1.5)^2$$

(M1)

$$= 0.5263\dots$$

A1**(Model B)**

$$R = 2.5pe^{-0.6p}$$

predicted values

p	R
1	1.372
2	1.506
3	1.2397

(A1)

$$SS_{res} = 0.170576\dots$$

A1

chose model B

A1

Note: Method marks can be awarded if seen for either model A or model B.
Award final **A1** if it is a correct deduction from their calculated values for A and B.

Total [7 marks]

16. (a) $H_0 : m = 3, H_1 : m < 3$

A1

Note: Accept equivalent statements in words.

[1 mark]

(b) *(let X be the number of fish caught)*

$$P(X \leq 1 | m = 3) = 0.199$$

M1A1

[2 marks]

(c) $P(X \geq 2 | m = 2.5) (= 1 - P(X \leq 1 | m = 2.5))$

M1A1

Note: Award **M1** for using $m = 2.5$ to evaluate a probability, award **A1** for also having $X \geq 2$.

$$= 0.713$$

A1

[3 marks]

Total [6 marks]

17. (a) $P(X = 8)$ **(M1)**

Note: Award **(M1)** for evidence of recognizing binomial probability.

eg, $P(X = 8), X \sim B\left(20, \frac{6}{15}\right)$.

$$= 0.180 \text{ (0.179705...)}$$

A1**[2 marks]**(b) let x be the number of male students

recognize that probability of selecting a male is equal to $\frac{x}{80}$

(A1)

$$\left(\text{set up equation } {}^{20}C_8 \left(\frac{x}{80} \right)^8 \left(\frac{80-x}{80} \right)^{12} = \right) 0.153357$$

(M1)

number of male students = 37

(M1)A1

Note: Award **(M1)A0** for 27.

[4 marks]**Total [6 marks]**

2. (a) (i) $P(Y) = 0.8 \times 0.1 + 0.2 \times 0.3$ **M1**
 $= 0.14$ **A1**

(ii) $P(\text{Star} | Y) = \frac{0.8 \times 0.1}{0.14}$ **M1**

$= 0.571 \left(\frac{4}{7}, 0.571428... \right)$ **A1**

[4 marks]

(b) the colours of the sweets are distributed according to manufacturer specifications **A1**

[1 mark]

(c)

Colour	Brown	Red	Green	Orange	Yellow	Purple
Expected Frequency	12	20	16	16	8	8

A2

Note: Award **A2** for all 6 correct expected values,
A1 for 4 or 5 correct values, **A0** otherwise.

[2 marks]

(d) 5 **A1**

[1 mark]

(e) 0.469 (0.4688117...) **A2**

[2 marks]

(f) since $0.469 > 0.05$ **R1**

fail to reject the null hypothesis. There is insufficient evidence to reject the manufacturer's specifications

A1

Note: Award **R1** for a correct comparison of their correct p -value to the test level, award **A1** for the correct result from that comparison. Do not award **R0A1**.

[2 marks]

Total [12 marks]

6. (a) $\begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$

M1A1**[2 marks]**

(b) $\begin{vmatrix} 0.8 - \lambda & 0.1 \\ 0.2 & 0.9 - \lambda \end{vmatrix} = 0$

M1

$\lambda = 1$ and 0.7

A1

eigenvectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

(M1)A1

Note: Accept any scalar multiple of the eigenvectors.

[4 marks]

(c) **EITHER**

$P = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0.7 \end{pmatrix}$

A1A1**OR**

$P = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \quad D = \begin{pmatrix} 0.7 & 0 \\ 0 & 1 \end{pmatrix}$

A1A1**[2 marks]**

(d) $P^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$

A1

$\frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1200 \\ 1200 \end{pmatrix}$

M1A1

attempt to multiply matrices

M1

so in company A, after n years, $400(2 + 0.7^n)$

A1**[5 marks]**

(e) $400 \times 2 = 800$

A1**[1 mark]****Total [14 marks]**

1. (a) χ^2 (goodness of fit) **A1**
[1 mark]

(b) **EITHER**

because aim is to measure improvement

OR

because the students may be of different ability in the two schools

R1
[1 mark]

(c) (i) 0.1875 (accept 0.188, 0.19)

A1

(ii) 2.46

(M1)A1

Note: Award **(M1)A0** for 2.63.

[3 marks]

(d) H_0 : there has been no improvement

H_1 : there has been an improvement

A1

attempt at a one-tailed paired t -test

(M1)

p -value = 0.423

A1

there is no significant evidence that the students have improved

R1

Note: If the hypotheses are not stated award a maximum of **A0M1A1R0**.

[4 marks]

(e) (i) H_0 : there is no difference between the schools

H_1 : school B did better than school A

A1

one-tailed 2 sample t -test

(M1)

p -value = 0.0984

A1

0.0984 > 0.05 (not significant at the 5% level) so do not reject the null hypothesis

R1A1

Note: The final **A1** cannot be awarded following an incorrect reason.
The final **R1A1** can follow through from their incorrect p -value.
Award a maximum of **A1(M1)A0R1A1** for p -value = 0.0993.

(ii) sample too small for the central limit theorem to apply (and t -tests assume normal distribution)

R1
[6 marks]

continued...

Question 1 continued

- (f) (i) $H_0: \rho = 0$
 $H_1: \rho > 0$

A1

Note: Allow hypotheses to be expressed in words.

$$p\text{-value} = 0.00157$$

A1

(0.00157 < 0.01) there is a significant evidence of a (linear) correlation between effort and improvement (so it is reasonable to assume a linear relationship)

R1

- (ii) (gradient of line of regression =) 6.6

A1**[4 marks]**

- (g) H_0 : improvement and gender are independent

H_1 : improvement and gender are not independent

A1

choice of χ^2 test for independence

(M1)

groups first two columns as expected values in first column less than 5
 new observed table

M1

	$(f - p) < 0$	$0 \leq (f - p) < 2$	$(f - p) \geq 2$
Male	14	10	9
Female	11	14	8

(A1)

$$p\text{-value} = 0.581$$

A1

no significant evidence that gender and improvement are dependent

R1**[6 marks]**

- (h) *For example:*
 larger samples / include data from whole school
 take equal numbers of boys and girls in each sample
 have a similar range of abilities in each sample
 (if possible) have similar ranges of effort

R1R1

Note: Award **R1** for each reasonable suggestion to improve the validity of the test.

[2 marks]**Total [27 marks]**

3. (a) H_0 : The die is fair **OR** $P(\text{any number}) = \frac{1}{6}$ **OR** probabilities are equal

H_1 : The die is not fair **OR** $P(\text{any number}) \neq \frac{1}{6}$ **OR** probabilities are not equal **A1**

[1 mark]

- (b) 5 **A1**

[1 mark]

- (c) 10 **A1**

[1 mark]

- (d) (p -value =) 0.287 (0.28724163....) **A2**

[2 marks]

- (e) $0.287 > 0.05$ **R1**

EITHER

Insufficient evidence to reject the null hypothesis

A1

OR

Insufficient evidence to reject that the die is fair

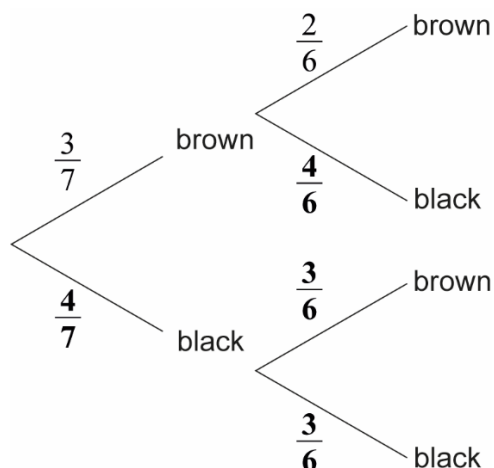
A1

Note: Do not award **R0A1**. Condone “accept the null hypothesis” or “the die is fair”.
Their conclusion must be consistent with their p -value and their hypothesis.

[2 marks]

[Total 7 marks]

4. (a)

**A1A1**

Note: Award **A1** for $\frac{4}{7}$ and $\frac{4}{6}$ correctly placed, **A1** for final two probabilities correct.

[2 marks]

(b) multiplying along branches and then adding outcomes

(M1)

$$\frac{3}{7} \times \frac{2}{6} + \frac{4}{7} \times \frac{3}{6}$$

$$= \frac{18}{42} \left(= \frac{3}{7} \approx 0.429 \text{ (42.9\%)} \right)$$

A1**[2 marks]**

(c) use of conditional probability formula

M1

$$\frac{\left(\frac{3}{7} \times \frac{2}{6} \right)}{\left(\frac{3}{7} \right)}$$

A1

$$= \frac{6}{18} \left(= \frac{1}{3} \right) \left(\frac{252}{756}, 0.333, 33.3\% \right)$$

A1**[3 marks]****[Total 7 marks]**

9. (a) let X be the weight of sugar in the bag

$$P(X < 950) = 0.308537... \approx 0.309$$

(M1)A1

[2 marks]

- (b) **METHOD 1**

let $\bar{X}$ be the mean weight of 5 bags of sugar

$$E(\bar{X}) = 1000$$

(A1)

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

(M1)

$$\text{Var}(\bar{X}) = \frac{100^2}{5} (= 2000)$$

(A1)

$$\bar{X} \sim N(1000, 2000)$$

$$P(\bar{X} > 950) = 0.868223... \approx 0.868 \text{ (86.8\%)}$$

A1

METHOD 2

let T be the total weight of 5 bags of sugar

$$E(T) = 5000$$

(A1)

use of $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2)$ for independent random variables

(M1)

$$\text{Var}(T) = 5 \times 100^2 (= 50000)$$

(A1)

$$T \sim N(5000, 50000)$$

$$P(T > 4750) = 0.868223... \approx 0.868 \text{ (86.8\%)}$$

A1

[4 marks]

[Total 6 marks]

15. (a) $X \sim \text{Po}(324)$

A1**Note:** Both distribution and mean must be seen for **A1** to be awarded.**[1 mark]**

(b) $P(X \leq 300)$
 $= 0.0946831... \approx 0.0947$

(M1)**A1****[2 marks]**

(c) (mean number of cars =) $4.5 \times 60 = 270$
 $P(X > 300 | \lambda = 270)$

(A1)**(M1)****Note:** Award **M1** for using $\lambda = 270$ to evaluate a probability.

$P(X \geq 301)$ **OR** $1 - P(X \leq 300)$
 $= 0.0334207... \approx 0.0334$

(M1)**A1****[4 marks]****[Total 7 marks]**

1. (a) $\left(\frac{17+25}{130}=\right) \frac{42}{130} \left(\frac{21}{65}, 0.323076\dots\right)$

A1**[1 mark]**

(b) $\left(\frac{17}{17+25}=\right) \frac{17}{42} (0.404761\dots)$

A1A1

Note: Award **A1** for correct numerator and **A1** for correct denominator.
Award **A1A0** for working of $\frac{17}{130}$ if followed by an incorrect answer.

[2 marks]

(c) $\frac{41}{130} \times \frac{40}{129}$

A1M1

Note: Award **A1** for two correct fractions seen, **M1** for multiplying their fractions.

$$= \frac{1640}{16770} \approx 0.0978 \left(0.0977936\dots, \frac{164}{1677}\right)$$

A1**[3 marks]****Total [6 marks]**

3. (a) $(E(X) =) 10 \times 0.8$ (M1)
 8 (people) A1
 [2 marks]
- (b) recognition of binomial probability (M1)
 0.0881 (0.0880803...) A1
 [2 marks]
- (c) 0.8 and 6 seen **OR** 0.2 and 3 seen (A1)
 attempt to use binomial probability (M1)
 0.121 (0.120873...) A1
 [3 marks]
Total [7 marks]

9. (a) $\bar{x} = 4.63$ (4.62686...) **A1**
[1 mark]

(b) $s_{n-1} = 1.098702$ **(A1)**
 $s_{n-1}^2 = 1.21$ (1.207146...) **A1**

Note: Award **A0A0** for an answer of 1.19 from biased estimate.

[2 marks]

(c) (i) $H_1 : \mu > 4.4$ **A1**

(ii) **METHOD 1**
 using a z -test **(M1)**
 $p = 0.0454992\dots$ **A1**
 $p < 0.05$ **R1**
 reject null hypothesis **A1**
 (therefore there is significant evidence that the IB HL math students know more digits of π than the population in general)

Note: Do not award **R0A1**. Allow **R1A1** for consistent conclusion following on from their p -value.

METHOD 2
 using a t -test **(M1)**
 $p = 0.0478584\dots$ **A1**
 $p < 0.05$ **R1**
 reject null hypothesis **A1**
 (therefore there is significant evidence that the IB HL math students know more digits of π than the population in general)

Note: Do not award **R0A1**. Allow **R1A1** for consistent conclusion following on from their p -value.

[5 marks]

Total [8 marks]

12. (a) $P(\text{Type I error}) = P(\text{stating female when male})$
 $= P(W_{\text{Male}} > 11.5)$
 $= 0.00135 \quad (0.00134996\dots)$

(M1)

A1

[2 marks]

(b) $P(\text{Type II error}) = P(\text{stating male when female})$
 $= P(W_{\text{Female}} < 11.5)$
 $= 0.309 \quad (0.308537\dots)$

(M1)

A1

[2 marks]

(c) attempt to use the total probability
 $P(\text{error}) = 0.9 \times 0.00134996\dots + 0.1 \times 0.308537\dots$
 $= 0.0321 \quad (0.0320687\dots)$

(M1)

A1

[2 marks]

Total [6 marks]

1. (a) (i) $\frac{370 + 472}{2}$ (M1)

Note: This (M1) can also be awarded for either a correct Q_3 or a correct Q_1 in part (a)(ii).

$Q_3 = 421$ A1

- (ii) their part (a)(i) – their Q_1 (clearly stated) (M1)

$IQR = (421 - 318) = 103$ A1

[4 marks]

- (b) $(Q_3 + 1.5(IQR) =) 421 + (1.5 \times 103)$ (M1)
 $= 575.5$

since $498 < 575.5$ R1

Netherlands is not an outlier A1

Note: The R1 is dependent on the (M1). Do not award R0A1.

[3 marks]

- (c) not appropriate (“no” is sufficient) A1
 as r is too close to zero / too weak a correlation R1

[2 marks]

- (d) (i) 6 A1

- (ii) 4.5 A1

- (iii) 4.5 A1

[3 marks]

- (e) (i) $r_s = 0.683$ (0.682646...) A2

- (ii) **EITHER**
 there is a (positive) association between the population size and the score A1

OR
 there is a (positive) linear correlation between the ranks of the population size and the ranks of the scores (when compared with the PMCC of 0.249) A1

[3 marks]

- (f) lowering the top score by 20 does not change its rank so r_s is unchanged R1

Note: Accept “this would not alter the rank” or “Netherlands still top rank” or similar. Condone any statement that clearly implies the ranks have not changed, for example: “The Netherlands still has the highest score.”

[1 mark]

[Total 16 marks]

3. (a) quota A1
[1 mark]
- (b) (i) $27.125 \approx 27.1$ (M1)A1
(ii) $8.29815... \approx 8.30$ A1
[3 marks]
- (c) (let μ be the national mean)
- $H_0: \mu = 25.2$
 $H_1: \mu > 25.2$ A1
- Note:** Accept hypotheses in words if they are clearly expressed and ‘population mean’ or ‘school mean’ is referred to. Do not accept $H_0: \mu = \mu_0$ unless μ_0 is explicitly defined as “national standard mark” or given as 25.2.
- recognizing t -test (M1)
 p -value = 0.279391... A1
- $0.279391... > 0.05$ R1
- Note:** The **R1** mark is for the comparison of their p -value with 0.05.
- insufficient evidence to reject the null hypothesis (that the mean for the school is 25.2) A1
- Note:** Award the final **A1** only if the **null** hypothesis is also correct (e.g. $\mu_0 = 25.2$ or (population) mean = 25.2) and the conclusion is consistent with both the direction of the inequality and the alternative hypothesis.
- [5 marks]
- (d) **EITHER**
the sampling process is not random R1
- For example:*
the school asked for volunteers
the students were selected from a single class
- OR**
the quota might not be representative of the student population R1
- For example:*
the school may have only 4 boys and 400 girls.
- Note:** Do not accept ‘the sample is too small’.
- [1 mark]
- (e) (i) $(28.1 \times 2 + 20 =) 76.2$ A1
- (ii) 8.4×2 (A1)
 $= 16.8$ A1
[3 marks]
- [Total 13 marks]

5. (a) (let T be the number of passengers who arrive)

$$(P(T > 72) =) P(T \geq 73) \quad \text{OR} \quad 1 - P(T \leq 72) \quad (\text{A1})$$

$$T \sim B(74, 0.9) \quad \text{OR} \quad n = 74 \quad (\text{M1})$$

$$= 0.00379 \quad (0.00379124...) \quad \text{A1}$$

Note: Using the distribution $B(74, 0.1)$, to work with the 10% that do not arrive for the flight, here and throughout this question, is a valid approach.

[3 marks]

(b) (i) 72×0.9 (M1)
64.8 A1

(ii) $n \times 0.9 = 72$ (M1)
80 A1

[4 marks]

- (c) **METHOD 1**

EITHER

when selling 74 tickets

	$T \leq 72$	$T = 73$	$T = 74$
Income minus compensation (I)	11100	10800	10500
Probability	0.9962...	0.003380...	0.0004110...

top row

A1A1

bottom row

A1A1

Note: Award **A1A1** for **each** row correct. Award **A1** for one correct entry and **A1** for the remaining entries correct.

$$E(I) = 11100 \times 0.9962... + 10800 \times 0.00338... + 10500 \times 0.000411 \approx 11099 \quad (\text{M1})\text{A1}$$

continued...

Question 5 continued

ORincome is $74 \times 150 = 11100$ **(A1)**

expected compensation is

$$0.003380... \times 300 + 0.0004110... \times 600 (= 1.26070...)$$

(M1)A1A1

Note: The **(M1)** is for an attempt to work out expected compensation by multiplying a probability for tickets sold by either 300 or 600.

expected income when selling 74 tickets is $11100 - 1.26070...$ **(M1)**

Note: Award **(M1)** for subtracting their expected compensation from 11100.

$$= 11098.73.. (= \$11099)$$

A1**THEN**

$$\text{income for 72 tickets} = 72 \times 150 = 10800$$

(A1)

$$\text{so expected gain} \approx 11099 - 10800 = \$299$$

A1**METHOD 2**for 74 tickets sold, let C be the compensation paid out

$$P(T = 73) = 0.00338014..., P(T = 74) = 0.000411098...$$

A1A1

$$E(C) = 0.003380... \times 300 + 0.0004110... \times 600 (= 1.26070...)$$

(M1)A1A1

$$\text{extra expected revenue} = 300 - 1.01404... - 0.246658... (300 - 1.26070...)$$

(A1)(M1)

Note: Award **A1** for the 300 and **M1** for the subtraction.

$$= \$299 \quad (\text{to the nearest dollar})$$

A1**METHOD 3**let D be the change in income when selling 74 tickets.

	$T \leq 72$	$T = 73$	$T = 74$
Change in income	300	0	-300

(A1)(A1)

Note: Award **A1** for one error, however award **A1A1** if there is no explicit mention that $T = 73$ would result in $D = 0$ and the other two are correct.

$$P(T \leq 73) = 0.9962..., P(T = 74) = 0.000411098...$$

A1A1

$$E(D) = 300 \times 0.9962... + 0 \times 0.003380... - 300 \times 0.0004110$$

(M1)A1A1

$$= \$299$$

A1**[8 marks]****[Total 15 marks]**

5. (a) (i) 0.02 A1

(ii) 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.

[2 marks]

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

Note: Award **M1** for an attempt to find eigenvalues. Any indication that $\det(\mathbf{M} - \lambda \mathbf{I}) = 0$ 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)

$\lambda = 1, 0.92 \begin{pmatrix} 23 \\ 25 \end{pmatrix}$ A1

[3 marks]

(c) $\begin{pmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ 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: This **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]

(d) (i) $\begin{pmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{pmatrix}^5 \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 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 (0.744311...) A1

(ii) $\begin{pmatrix} 0.25 \\ 0.75 \end{pmatrix}$ (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

[4 marks]

Total [12 marks]

1. (a) (i) mean = 4.23 (4.23333...) A1
- variance = 4.27 (4.26777...) A1
- [2 marks]
- (ii) mean is close to the variance A1
- [1 mark]

(b) *One of the following:*

- the number of bags sold each day is independent of any other day
 the sale of one bag is independent of any other bag sold
 the sales of bags of rice (each day) occur at a *constant mean* rate

A1

Note: Award **A1** for a correct answer in context. Any statement referring to independence must refer to either the independence of each bag sold or the independence of the number of bags sold each day. If the third option is seen, the statement must refer to a “constant mean” or “constant average”. Do not accept “the number of bags sold each day is constant”.

[1 mark]

- (c) attempt to find Poisson probabilities and multiply by 90 (M1)
- $a = 7.018$ A1
- $b = 17.498$ A1

EITHER

$$90 \times P(X \geq 8) = 90 \times (1 - P(X \leq 7)) \quad (M1)$$

$c = 5.755$ A1

OR

$$90 - 7.018 - 11.903 - 16.665 - 17.498 - 14.698 - 10.289 - 6.173 \quad (M1)$$

$c = 5.756$ A1

Note: Do not penalize the omission of clear a , b and c labelling as this will be penalized later if correct values are interchanged.

[5 marks]

(d) (i) 7

A1**[1 mark]**(ii) H_0 : The number of bags of rice sold each day follows a Poisson distribution with mean 4.2.**A1** H_1 : The number of bags of rice sold each day does not follow a Poisson distribution with mean 4.2.**A1**

Note: Award **A1A1** for **both** hypotheses correctly stated and in correct order. Award **A1A0** if reference to the data and/or “mean 4.2” is not included in the hypotheses, but otherwise correct.

evidence of attempting to group data to obtain the observed frequencies for ≤ 1 and ≥ 8 **(M1)** $p\text{-value} = 0.728 \text{ (0.728100...)}$ **A2** $0.728 \text{ (0.728100...) } > 0.05$ **R1**the result is not significant so there is no reason to reject H_0
(the number of bags sold each day follows a Poisson distribution)**A1**

Note: Do not award **R0A1**. The conclusion **MUST** follow through from their hypotheses. If no hypotheses are stated, the final **A1** can still be awarded for a correct conclusion as long as it is in context (e.g. therefore the data follows a Poisson distribution).

[7 marks]

(e) (i) **METHOD 1**evidence of multiplying 4.2×60 (seen anywhere)**M1**

$$H_0 : \mu = 252$$

$$H_1 : \mu > 252$$

A1**Note:** Accept $H_0 : \mu = 4.2$ and $H_1 : \mu > 4.2$ for the **A1**.

evidence of finding probabilities around critical region

(M1)**Note:** Award **(M1)** for any of these values seen:

$$P(X \geq 277) = 0.0630518... \quad \text{OR} \quad P(X \leq 276) = 0.936948...$$

$$P(X \geq 278) = 0.0558415... \quad \text{OR} \quad P(X \leq 277) = 0.944158...$$

$$P(X \geq 279) = 0.0493055... \quad \text{OR} \quad P(X \leq 278) = 0.950694....$$

critical value = 279

A1

$$282 \geq 279,$$

R1

the null hypothesis is rejected

A1

(the advertising increased the number of bags sold during the 60 days)

Note: Do not award **R0A1**. Accept statements referring to the advertising being effective for **A1** as long as the **R** mark is satisfied. For the **R1A1**, follow through within the part from their critical value.**METHOD 2**

evidence of dividing 282 by 60 (or 4.7 seen anywhere)

M1

$$H_0 : \mu = 4.2$$

$$H_1 : \mu > 4.2$$

A1

attempt to find critical value using central limit theorem

(M1)

$$\left(\text{e.g. sample standard deviation} = \sqrt{\frac{4.2}{60}}, \bar{X} \sim N\left(4.2, \sqrt{\frac{4.2}{60}}\right), \text{etc.}\right)$$

Note: Award **(M1)** for a p -value of 0.0293907... seen.

critical value = 4.63518...

A1

$$4.7 > 4.63518...$$

R1

the null hypothesis is rejected

A1

(the advertising increased the number of bags sold during the 60 days)

Note: Do not award **R0A1**. Accept statements referring to the advertising being effective for **A1** as long as the **R** mark is satisfied. For the **R1A1**, follow through within the part from their critical value.**[6 marks]**

$$(ii) \quad (P(X \geq 279 | \mu = 252) =) 0.0493 \quad (0.0493055...)$$

A1**Note:** If a candidate uses **METHOD 2** in part (e)(i), allow an **FT** answer of 0.05 for this part but only if the candidate has attempted to find a p -value.**[1 mark]**

(f) attempt to compare profit *difference* with cost of advertising

(M1)

Note: Award **(M1)** for evidence of candidate mathematically comparing a profit difference with the cost of the advertising.

EITHER

(comparing profit from 30 extra bags of rice with cost of advertising)

$14850 < 18000$

A1

OR

(comparing total profit with and without advertising)

$121590 < 124740$

A1

OR

(comparing increase of average daily profit with daily advertising cost)

$247.50 < 300$

A1

THEN

EITHER

Even though the number of bags of rice increased, the advertising is not worth it as the overall profit did not increase.

R1

OR

The advertising is worth it even though the cost is less than the increased profit, since the number of customers increased (possibly buying other products and/or returning in the future after advertising stops)

R1

Note: Follow through within the part for correct reasoning consistent with their comparison.

[3 marks]

[Total 27 marks]

1. (a) The favourite breakfast/berry (of adults) is independent of (their) income (level). **A1** [1 mark]

(b) $\chi^2 = 2.27$ (2.26821...) **A2** [2 marks]

(c) **EITHER**
 $2.27 < 7.78$ **OR** $2.27 < \text{critical value}$ **R1**
OR
 $0.687 > 0.1$ (using p -value)

THEN

(Do not reject H_0)

Insufficient evidence (at the 10% significance level) that the favourite berry depends on income level. **A1**

Note: Do not award **R0A1**. Accept " χ^2 " in place of their "2.27", provided an answer was seen in part (b). Their conclusion must be consistent with their χ^2 (or a correct p -value) and their hypothesis.

[2 marks]
Total [5 marks]

5. (a) $0.5 \times 0.1 + 0.4 \times 0.4 + 0.1 \times 0.5$ (M1)(M1)(M1)

Note: Award **M1** for 0.5×0.1 or 0.1×0.5 , **M1** for 0.4×0.4 , **M1** for adding three correct products.

0.26

A1
[4 marks]

(b) $0 = -8 \times 0.5 + 4 \times 0.4 + 0.1k$ (M1)(M1)

Note: Award **M1** for correct substitution into the formula for expected value, award **M1** for the expected value formula equated to zero.

$(k =) 24$ (points)

A1
[3 marks]
Total [7 marks]

16. (a) (let p be the probability of a student choosing healthy options)

$$H_0: p = 0.3$$

A1

$$H_1: p > 0.3$$

A1

Note: Award **A0A1** for correct hypotheses with μ in place of p .
Accept equivalent hypotheses in words.

[2 marks]

- (b) a type I error is rejecting H_0 when H_0 is true

(M1)

(let N = number of students choosing a healthy option)

$$N \sim B(80, 0.3)$$

$$P(31 \leq N \leq 80) \quad \text{OR} \quad P(N \geq 31) \quad \text{OR} \quad 1 - P(N \leq 30)$$

(M1)

Note: Do not accept the use of the Normal approximation.

$$0.0587 \quad (0.0587481\dots)$$

A1

[3 marks]

- (c) a type II error is accepting H_0 when H_0 is not true

(M1)

$$N \sim B(80, 0.4)$$

$$P(0 \leq N \leq 30) \quad \text{OR} \quad P(N \leq 30)$$

(M1)

$$0.369 \quad (0.368726\dots)$$

A1

[3 marks]

Total [8 marks]

1. (a) (i) ($m =$) 54(%) **A1**
- (ii) ($n =$) 14(%) **A1**
- (iii) ($p =$) 22(%) **A1**
- (iv) ($q =$) 10(%) **A1**

Note: Based on their n , follow through for parts (i) and (iii), but only if it does not contradict the given information. Follow through for part (iv) but only if the total is 100%.

[4 marks]

- (b) (i) $0.54 \left(\frac{54}{100}, \frac{27}{50}, 54\% \right)$ **A1**
- (ii) $\frac{54}{64} \left(0.844, \frac{27}{32}, 84.4\%, 0.84375 \right)$ **A1A1**

Note: Award **A1** for a correct denominator (0.64 or 64 seen), **A1** for the correct final answer.

[3 marks]

- (c) (i) recognizing Binomial distribution with correct parameters **(M1)**
 $X \sim B(10, 0.68)$
 $(P(X = 5) =) 0.123 \text{ (0.122940..., 12.3\%)}$ **A1**
- (ii) $1 - P(X \leq 3)$ **OR** $P(X \geq 4)$ **OR** $P(4 \leq X \leq 10)$ **(M1)**
 $0.984 \text{ (0.984497..., 98.4\%)}$ **A1**
- (iii) $(0.68)^9 \times 0.32$ **(M1)**
 recognition of two possible cases **(M1)**
 $2 \times ((0.68)^9 \times 0.32)$
 $0.0199 \text{ (0.0198957..., 1.99\%)}$ **A1**

[7 marks]

- (d) **EITHER**
 the probability is not constant **A1**
OR
 the events are not independent **A1**
OR
 the events should be modelled by the hypergeometric distribution instead **A1**

[1 mark]**Total [15 marks]**

6. (a) wood layer, $W \sim N(7, 0.3^2)$; plastic, $P \sim N(3, 0.16^2)$
 door: $X = W + P$
 $E(X) = 10$ (mm) (A1)
 $\text{Var}(X) = \text{Var}(W) + \text{Var}(P) = 0.1156$ (mm²) (M1)(A1)
 recognizing the distribution is Normal, with their mean and variance (M1)
 $X \sim N(10, 0.34^2)$
 $P(X < 9.5) = 0.0707$ (0.07070125...) A1
 [5 marks]
- (b) $E(T) = 80$ (A1)
 $\text{Var}(T) (= 0.1156 \times 8) = 0.9248$ (M1)(A1)
 $T \sim N(80, 0.9248)$
 $P(T > 82) = 0.0188$ (0.0187753...) A1
 [4 marks]
- (c) (i) 6.93 mm (6.93428...) A1
- (ii) $(s_{n-1}^2) = 0.404$ (A1)
 $(s_{n-1}^2) = 0.163$ mm² (0.162928...) A1
 [3 marks]
- (d) $H_0: \mu_A = \mu_B$ and $H_1: \mu_A > \mu_B$ A1A1

Note: Award **A1** for use of μ or in words “population mean”, and **A1** for both correct equality in null hypothesis and correct inequality in alternative hypothesis. Accept an equivalent statement in words, must include mean and reference to “**population** mean” / “mean for **all** Machine B layers” for the first **A1** to be awarded.

use a two-sample t -test (M1)
 $p\text{-value} = 0.406975...$ A1
 since $0.406975... > 0.05$ **OR** $p\text{-value} > 0.05$ R1
 Do not reject H_0 (Insufficient evidence to support the employee’s claim) A1

Note: Accept a p -value of 0.415861... from use of 3sf values from part (c). Follow through within the question for the final **R1** and **A1** for their p -value provided $0 \leq p \leq 1$. Do not award **R0A1**.

[6 marks]
Total [18 marks]

1. (a) $\frac{12}{5} \left(\frac{144}{60}, 2.4 \right)$ (M1)A1

[2 marks]

(b) $\frac{3}{5} \left(\frac{144}{240}, 0.6 \right)$ A1

[1 mark]

(c) (i) $\text{wins} \sim B\left(4, \frac{3}{5}\right)$
 $P(\text{wins} = 0) = 0.0256$ A1

Note: Allow **FT** from use of their probability in part (b) but only when used with $n = 4$.

(ii) $\text{expected frequency} = 60 \times 0.0256$ (M1)
 $= 1.536$ A1

[3 marks]

(d) (i) H_0 : data follows a Binomial distribution with $n = 4$ A1

(ii) $(df = 4 - 1 - 1 =) 2$ A1

(iii) $p\text{-value} = 0.954$ (0.953872...) (M1)A1

(iv) $0.954 > 0.05$ R1
insufficient evidence to reject H_0 A1

Note: Condone “accept H_0 ”. Follow through from their p -value in part (d)(iii) if the reasoning is correct and correct conclusions are made. Do NOT award **R0A1**.

[6 marks]

(e) $\text{wins} \sim B\left(4, \frac{3}{5}\right)$ OR $1 - 0.0256$ (M1)

$P(\text{wins} \geq 1) = 0.974$ (0.9744) A1

[2 marks]

- (f) (i) there are 145 transitions that start with Argentina, (M1)

(of which 85 lead to Argentina winning,) so the probability is $\frac{85}{145}$ A1

$$= \frac{17}{29} \quad \text{AG}$$

(ii) $\begin{pmatrix} \frac{17}{29} & \frac{31}{47} \\ \frac{12}{29} & \frac{16}{47} \end{pmatrix} \left(= \begin{pmatrix} 0.586 & 0.660 \\ 0.414 & 0.340 \end{pmatrix} \right)$ A1A1

Note: Accept the transposed matrix as correct.

Award **A1** for $\frac{17}{29}$ placed in a leading diagonal.

Award **A1** for all other values correct and in correct position in the matrix.

[4 marks]

- (g) (i) write their matrix with λ subtracted from the leading diagonal (M1)
equate determinant to zero (M1)

$$\det \begin{pmatrix} \frac{17}{29} - \lambda & \frac{31}{47} \\ \frac{12}{29} & \frac{16}{47} - \lambda \end{pmatrix} = 0$$

$$\left(\frac{17}{29} - \lambda \right) \left(\frac{16}{47} - \lambda \right) - \frac{12}{29} \times \frac{31}{47} = 0 \quad \text{A1}$$

correct intermediate step

$$1363\lambda^2 - 1263\lambda - 100 = 0 \quad \text{AG}$$

Note: Do not award **A1** if there is no intermediate step leading from determinant to given answer.

Solving $T \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ for x and y may be seen and is a valid alternative method.

Accept working in the form $\det(\lambda I - T) = 0$.

(ii) $\lambda = 1, -\frac{100}{1363} \text{ } (-0.0733675\dots)$ A1

(iii) attempt to solve $T \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ (M1)

$$y = \frac{564}{899}x$$

eigenvector for $\lambda = 1$ is $\begin{pmatrix} 1 \\ \frac{564}{899} \end{pmatrix} \left(= \begin{pmatrix} 1 \\ 0.627 \end{pmatrix} \right)$ A1

eigenvector for $\lambda = -\frac{100}{1363}$ is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ A1

Note: Allow correct multiples of the eigenvectors.

If eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is stated without the second eigenvector, or any other working, then award **M0A0A1**.

[7 marks]

(h) **EITHER**

solution found using $\begin{pmatrix} 1 \\ \frac{564}{899} \end{pmatrix} \left(= \begin{pmatrix} 1 \\ 0.627 \end{pmatrix} \right)$

$$x + \frac{564}{899}x = 1 \quad (M1)$$

$$x = 0.614 \left(0.614490\dots, \frac{899}{1463} \right) \quad (A1)$$

OR

solution can be found from high power of transition matrix

$$\begin{pmatrix} \frac{17}{29} & \frac{31}{47} \\ \frac{12}{29} & \frac{16}{47} \end{pmatrix}^{50} = \begin{pmatrix} 0.614 & 0.614 \\ 0.386 & 0.386 \end{pmatrix} \quad (M1)$$

Note: Accept the transposed matrix if consistent with their answer to part (f)(ii).

$$\text{probability} = 0.614 \left(0.614490\dots, \frac{899}{1463} \right) \quad (A1)$$

THEN

$$P(3 \text{ wins}) = 0.614 \times 0.586^2 \left(= \frac{899}{1463} \times \left(\frac{17}{29} \right)^2 \right) \quad (M1)$$

$$= 0.211 \left(0.211162\dots, \frac{8959}{42427} \right) \quad A1$$

[4 marks]

[Total: 29 marks]

2. (a) $H_0 : \mu_b = \mu_m$ **A1**
 $H_1 : \mu_b > \mu_m$ **A1**

Note: Accept equivalent statements in words such as “the **mean** score of bilingual people equals the **mean** score of monolingual people”.

[2 marks]

- (b) 0.119 (0.119395...) **A2**
[2 marks]

- (c) $0.119395... > 0.05$ ($11.9395...\% > 5\%$) **R1**

(fail to reject H_0) there is insufficient evidence to suggest that bilingual people have better memory retention than monolingual people **A1**

Note: Do not award **R0A1**.
 The answer to part (c) MUST be consistent with **their** hypotheses and **their** p -value.

[2 marks]

[Total: 6 marks]

8. (a) H_0 : X and Y are not (linearly) correlated **OR** $\rho = 0$ **A1**
 H_1 : X and Y are (linearly) correlated **OR** $\rho \neq 0$ **A1**

Note: Accept “independent” or “not associated” in place of “not correlated”.
 If H_0 and H_1 are reversed, then award **A0A1**.

[2 marks]

- (b) (i) $r = 0.849$ (0.848886...) **A1**

- (ii) $p\text{-value} = 0.0325$ (0.0325277...) **A2**

Note: Award **A1** for $p\text{-value} = 0.033$ or $p\text{-value} = 0.03$.
 Award **FT** for $\rho > 0$ or $\rho < 0$ in part (a), $p\text{-value} = 0.0163$ (0.0162638...)
 or $p\text{-value} = 0.984$ (0.983736...)
 Award the full marks for seeing the values of r and $p\text{-value}$ from the
 markscheme when H_0 and H_1 are reversed in part (a).

[3 marks]

- (c) $0.0325 < 0.05$ **R1**
 (so we reject H_0 in favour of H_1)
 (there is sufficient evidence to suggest) X and Y are (linearly) correlated **A1**

Note: Their conclusion must be consistent with their $p\text{-value}$ and their hypotheses and
 it must be in context.

[2 marks]

[Total: 7 marks]

9. (a) attempt to find the difference between 75.7 and 67.3 (M1)

$$\frac{75.7 - 67.3}{2}$$

$$4.2 \text{ (km h}^{-1}\text{)}$$

A1

[2 marks]

- (b) **METHOD 1 (Comparing areas above and below the mean)**

P(67.3 < speed < 74) **OR** Normal CDF(67.3, 74, 67.3, 4.2) **OR** sketch of normal distribution with 67.3 and 74 labelled and shaded between (M1)

area of region between mean and q is at least 0.445 (0.444670...) A1

Hence no more than 0.375 (0.375329...) between mean and p R1

The region between p and q is not symmetrical AG

METHOD 2 (Comparing areas in the tails)

attempt to calculate probability that speed < p and speed > q with $q=74$ (M1)

$$P(\text{speed} < 74) = 0.944670\dots$$

$$P(\text{speed} < p) = (0.944670\dots - 0.82) = 0.124670\dots$$

$$P(\text{speed} > q) = (1 - 0.944670\dots) = 0.0553295\dots$$

A1

if $q \geq 74$, then $P(\text{speed} > q) \leq 0.0553295$ and $P(\text{speed} < p) \geq 0.124670$ so

$P(\text{speed} > q)$ will never equal $P(\text{speed} < p)$ R1

the region between p and q is not symmetrical AG

continued...

*Question 9 continued***METHOD 3 (Assumption of symmetry comparing speeds)**

attempt to calculate area below q assuming distribution is symmetrical **(M1)**

e.g. $P(\text{speed} < q) = 0.82 + \frac{1}{2} \times 0.18$ (0.91)

EITHER

$(q =) 72.9$ (72.9311...) **A1**

$72.9 < 74$ so 74 would not be in the region **R1**

the region between p and q is not symmetrical **AG**

OR

$P(\text{speed} < 74) = 0.945$ (0.944670...) **A1**

$0.945 > 0.91$ so 74 would not be in the region **R1**

the region between p and q is not symmetrical **AG**

METHOD 4 (Assumption of symmetry comparing areas)

attempt to calculate symmetrical area with 74 as a boundary **(M1)**

$P(60.6 < \text{speed} < 74)$ **OR** Normal CDF(60.6, 74, 67.3, 4.2) **OR**

$P(67.3 < \text{speed} < 74)$ **OR** Normal CDF(67.3, 74, 67.3, 4.2)

EITHER

0.889 (0.889340...) **A1**

$0.889 > 0.82$ so 74 would not be in the region **R1**

the region between p and q is not symmetrical **AG**

OR

0.445 (0.444670...) **A1**

$0.445 > 0.82 \div 2$ so 74 would not be in the region **R1**

the region between p and q is not symmetrical **AG**

[3 marks]

[Total: 5 marks]

11. Let $D = O - L - L$ (A1)

(mean =) $205 - 105 - 105$ ($= -5$) (A1)

manipulating variances (not standard deviations) (M1)

(variance =) $25 + 9 + 9$ ($= 43$) **OR** (SD =) $6.55743\dots$ (A1)

$D \sim N(205 - 105 - 105, 25 + 9 + 9)$

attempt to find the probability that $D > 0$ (M1)

$P(D > 0)$

$= 0.223$ ($0.222882\dots$) A1

Note: If $D = O - 2L$ is seen or implied, award at most (A0)A1(M0)(A0)M1A0.
--

[Total: 6 marks]

16. (a) attempt to find what the model predicts in terms of k (M1)

$$k, \frac{k}{4}, \frac{k}{25}$$

correct expression for sum of square residuals (A1)

$$(k - 42)^2 + \left(\frac{k}{4} - 11\right)^2 + \left(\frac{k}{25} - 1.5\right)^2$$

valid attempt to find c by expanding or recognizing the constant terms (M1)

$$c = 42^2 + 11^2 + 1.5^2$$

$$= 1887.25$$

A1

[4 marks]

- (b) valid method to find the k value at the minimum (M1)

$$k = \frac{89.62}{2 \times 1.0641} (= 42.1107...), \text{ graph, completing the square}$$

$$(\text{so least squares regression is}) \quad I = \frac{42.1}{d^2}$$

A1

[2 marks]

[Total: 6 marks]

2. (a) 11.0 (11.0212...)

A2

Note: Award **A1** for a final answer of 11 if no unrounded answer is seen.

[2 marks](b) **EITHER**

$$11.0 > 9.488 \text{ (} 11.0212... > 9.488 \text{)}$$

R1**OR**

$$0.0263 < 0.05 \text{ (} 0.0263264... < 0.05 \text{)}$$

R1**THEN****EITHER**(there is significant evidence to) reject H_0 **A1****OR**

(there is significant evidence that) the (food) quality and the type of meal are not independent

A1**Note:** Do not award **R0A1**.Award **R1** for $\chi^2_{\text{calc}} > \chi^2_{\text{crit}}$, provided the calculated value is explicitly seen in part (b).Accept " p -value < significance level" provided their p -value is seen and their p -value is between 0 and 1.**[2 marks]****Total [4 marks]**

3. (a) attempting to use $P(R \cap S) = P(R)P(S)$ (M1)
 $0.2 = 0.8(0.2 + x)$ (A1)
 $x = 0.05$ A1
[3 marks]
- (b) $x + 0.2 + 0.6 + y = 1$ (M1)
 $y = 0.15$ A1
[2 marks]
- (c) **METHOD 1**
attempting to apply $P(R' | S') = \frac{P(R' \cap S')}{P(S')}$ (M1)
 $\frac{0.15}{0.2}$
 $= \frac{3}{4}$ A1
- METHOD 2**
 $P(R' | S') = P(R')$ (because R, S are independent) (M1)
 $= 1 - 0.25 = 0.75$ A1
- Note:** *FT* from their values of x or y .
- [2 marks]
Total [7 marks]

5. (a) $X \sim N(4, 0.25^2)$

EITHER

correct probability expression

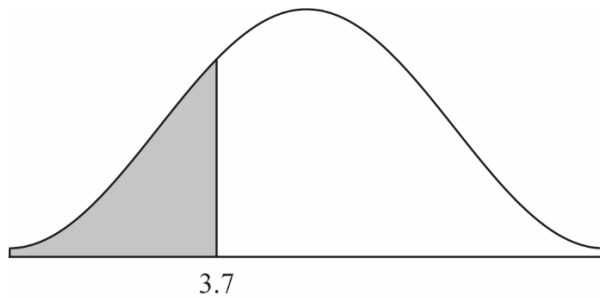
(M1)

$$P(X < 3.7)$$

Note: Accept a weak or strict inequality, and any label instead of X , e.g. length or L .

OR

normal curve with vertical line, left of mean, labelled 3.7, and shaded region **(M1)**



THEN

0.115 (0.115069..., 11.5%)

A1

Note: Award **M1A0** for 0.12 if no previous working.

[2 marks]

- (b) **EITHER**

Correct probability expression

(M1)

$$(P(X < k) = 0.7 \quad \text{OR} \quad P(X > k) = 0.3)$$

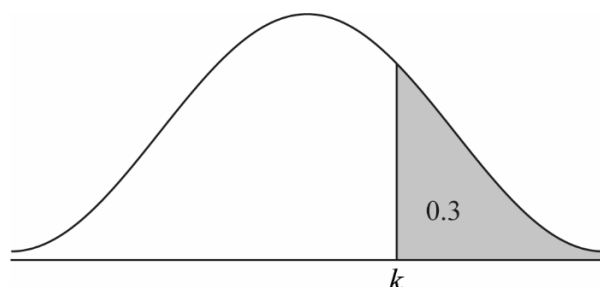
Note: Accept a weak or strict inequality, and any label instead of X e.g., length or L .

continued...

Question 5 continued

OR

normal curve with vertical line to the right of the mean and shaded region, correctly labelled either 0.3 or 0.7

(M1)**THEN**

$(k =) 4.13$ (4.13110...)

A1

Note: Award **M1A0** for 4.1 if no previous working.

[2 marks](c) **EITHER**

correct probability equation

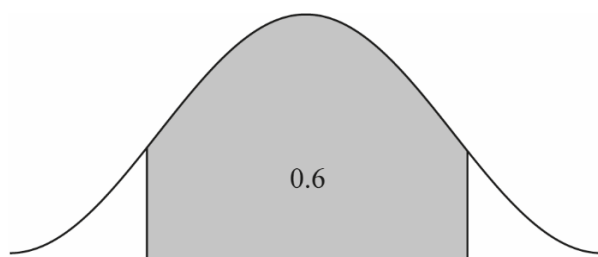
$P(\text{length} < 4 + m) = 0.8$ **OR** $P(\text{length} < 4 - m) = 0.2$

(M1)

Note: Accept any letter instead of "length" e.g., X or L .

OR

normal curve with vertical lines symmetrical about the mean line with a correct indication of an area of 0.6 or 0.2 or 0.8

(M1)**THEN**

0.210 (0.210405...)

A1

Note: Award **(M1)A0** for an answer of 3.7895 or 4.2105 seen without working.
Condone 0.21 seen and award **(M1)A1**.

[2 marks]**Total [6 marks]**

7. (a) $(56 \times 0.86) = 48.2$ (48.16)

A1

Note: Accept 48.

[1 mark]

(b) recognizing binomial distribution (may be seen in (a))

(M1)

e.g. $X \sim B(56, 0.86)$

$(P(X \geq 50) =) 0.316$

A2

[3 marks]

(c) $P(X \leq n) \geq 0.25$

$n = 46$

A2

[2 marks]

Total [6 marks]

9. (a)

Athlete	A	B	C	D	E	F	G	H
Age rank	7	6	3	5	4	2	8	1
Time rank	3.5	2	3.5	6	7	8	1	5

A1A1

Note: Award **A1** for each correct row.

[2 marks](b) $r_s = -0.671$ ($-0.670670\dots$)**A2**

Note: Only follow through from an incorrect table provided the ranks are all between 1 and 8.
Award **A1** for -0.67 **OR** for the omission of the negative sign, e.g. 0.671 ($0.670670\dots$) or 0.67

[2 marks]

(c) (A value of $r_s = -0.671$) indicates a negative correlation between a person's age and the best time they take to run 100m.

R1

Note: Condone any comment that includes "weak" or "strong" etc. Accept an interpretation in words, but only if there is a general link described and not a rule: "The older a person gets, the faster they *tend to* run".
Answer must be in context.

[1 mark](d) Award **R1** for any sensible reason:**R1**

The correlation, such that it is, is unlikely to be linear for this type of data.
Spearman's CC is less sensitive to outliers
Sung-Jin is not sure the data is drawn from a bivariate normal distribution
There are outliers/extreme data
Same time for two athletes with significantly different ages

[1 mark](e) (i) 0.264 ($0.263762\dots$)**A2**

Note: Award **A1** for 0.26 with no working. Given that the exact model is not specific in the question, accept correct r^2 values from other regression models: 0.631, 0.650, 0.759 and 0.256.

(ii) approximately 26% of the variability in the times taken can be explained by the runner's age.

R1**[3 marks]****Total [9 marks]**

10. (a) H_0 : there is no particular preference for any of the flavours **A1**
 H_1 : there is a particular preference for some of the flavours **A1**

Note: Accept equivalent statements such as “ H_0 : the population ratio of flavour preferences is 1:1:1:1” or “ H_0 : the population proportions are equal for each flavour” or “ H_0 : the data is drawn from a uniform distribution”.

[2 marks]

- (b) **EITHER** **A2**
 $p\text{-value} = 0.0629$ (0.0629034...) **R1**
 $0.0629 > 0.05$ **R1**
OR
 $\chi^2_{\text{calc}} = 7.30$ **A2**
 $7.30 < 7.82$ **R1**

Note: Award **A2** for either $p\text{-value} = 0.063$ or $\chi^2_{\text{calc}} = 7.3$ seen. Award **R1** for a correct comparison involve their $p\text{-value}$ or χ^2_{calc} , and follow through for their conclusion.

THEN

so there is insufficient evidence to reject H_0 , i.e.

there is no particular preference for any of the flavours.

A1

Note: Do not award **R0A1**.

[4 marks]**Total [6 marks]**

15. (a) (i) 224 g (224.25 g) **A1**
- (ii) [222.1, 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 $\bar{x}$.
Award **A1A0** for correct answers not given correct to 4 sf.

[3 marks]

- (b) **EITHER**
the (population) weight of granules of Apollo coffee is normally distributed. **R1**
OR
the readings are independent **R1**

[1 mark]

- (c) 226g 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]

Total [6 marks]

1. (a) $\frac{9.45 - 8.73}{1958 - 1708}$ (M1)

$= 0.00288 \left(\frac{9}{3125} \right)$ A1

[2 marks]

(b) (i) the (mean) yearly change in (mean annual) temperature A1

Note: Accept equivalent statements, e.g. “rate of change of temperature”.

(ii) °C / year **OR** degrees C per year A1

Note: Do not follow through from part (b)(i) into (b)(ii).

[2 marks]

(c) attempt to substitute point and gradient into appropriate formula (M1)

$$8.73 = 0.00288 \times 1708 + c \Rightarrow c = 3.81096\dots$$

or

$$9.45 = 0.00288 \times 1958 + c \Rightarrow c = 3.81096\dots$$

equation is $y = 0.00288x + 3.81$ A1

[2 marks]

(d) attempt to substitute 2000 into their part (c) (M1)

$$0.00288 \times 2000 + 3.81096\dots$$

$= 9.57$ (°C) (9.57096...) A1

[2 marks]

continued...

Question 1 continued

(e) (i) $y = 0.00256x + 4.46$ ($0.00255714...x + 4.46454...$) **(M1)A1**

Note: Award **(M1)A0** for answers that show the correct method, but are presented incorrectly (e.g. no “ $y =$ ” or truncated values etc.). Accept 4.465 as the correct answer to 4 sf.

(ii) 0.861 (0.861333...) **A1**

[3 marks]

(f) attempt to substitute 2000 into their part (e)(i) **(M1)**

$0.00255714... \times 2000 + 4.46454...$

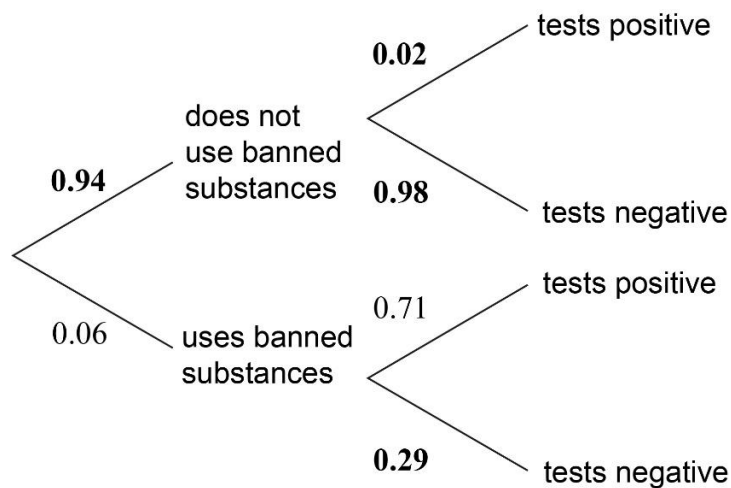
$= 9.58(^{\circ}\text{C})$ ($9.57882...(^{\circ}\text{C})$) **A1**

Note: Award **A1** for 9.57 from $0.00255714 \times 2000 + 4.46$.

[2 marks]

[Total: 13 marks]

3. (a)

**A1A1**

Note: Award **A1** for any one value correct, **A1** for other three values correct. Accept percentage responses as equivalent forms on **all** branches.

[2 marks](b) (i) multiplication of two probabilities along the tree diagram **(M1)**

$$0.94 \times 0.98$$

$$= 0.921 \text{ (0.9212, 92.1\%, 92.12\%)} \quad \mathbf{A1}$$

(ii) $(0.9212)^2$ **(A1)**

$$= 0.849 \text{ (0.848609..., 84.9\%, 84.8609...%)} \quad \mathbf{A1}$$

[4 marks]*continued...*

Question 3 continued

(c) (i) $0.94 \times 0.02 + 0.06 \times 0.29$ (A1)(M1)

Note: Award **A1** for two correct products from their tree diagram seen, **M1** for the addition of their two products.

0.0362 (3.62%) A1

(ii) multiplying their part (c)(i) by 1300

0.0362×1300 (M1)

47.1 (47.06) A1

[5 marks]

(d) $p = 0.02$ OR $p = 0.98$ (A1)

recognition of binomial probability with $n = 20$ (M1)

$P(X = 0)$ OR $P(X = 20)$ (M1)

0.668 (0.667607...) A1

Note: Award (A1)(M1)(M1)A0 for an answer of 0.667 .

$0.98^{20} = 0.668$ (0.667607...) is awarded full marks.

[4 marks]

(e) $P(X \geq 3)$ OR $P(X \leq 17)$ (M1)

0.00707 (0.00706869...) A1

Note: Award (M1)A0 for an answer of 0.00706. Award (M1)A0 for an answer of 0.0599 (0.0598989...), obtained from the use of $P(X \geq 2)$.

FT from their value of p in part (d)

[2 marks]

[Total: 17 marks]

5. (a) $(s_{n-1} =) 1.30243\dots$ **(M1)(A1)**
 $1.70 (1.69632)$ **A1**

Note: Award **(M1)A0A0** for a value of $(s_n =) 1.28934\dots$ or $(s_n^2 =) 1.6624$ seen.

[3 marks]

- (b) the variance and the mean are similar **R1**

Note: Do not accept a general statement “the variance and the mean are equal” unless their answer in part (a) is 1.76.

[1 mark]

- (c) (i) attempt to find $P(X = 4)$ under the null hypothesis $(= 0.0687830\dots)$ **(M1)**
multiplying by 50 **(M1)**
 $j = 3.44 (3.43915\dots)$ **A1**

- (ii) **EITHER**
attempt to find $P(X \geq 5)$ under the null hypothesis and multiply by 50 **(M1)**

OR

$$50 - (8.60 + 15.14 + 13.32 + 7.82 + 3.44) \quad (= 5.12 - 3.44) \quad \textbf{(M1)}$$

THEN

$$k = 1.68 (1.67925\dots) \quad \textbf{A1}$$

[5 marks]

- (d) there are expected frequencies less than 5 **A1**

[1 mark]

continued...

Question 5 continued

(e) 3

A1

[1 mark]

(f) 0.991 (0.991187)

(M1)A1

Note: Award **M1** for a table of observed and expected frequencies with columns for 4 and 5 or more combined.

[2 marks]

(g) $99\% > 5\%$

R1

EITHER

so there is insufficient evidence to reject H_0 .

A1

OR

we accept that the number of sightings follows a Poisson distribution

A1

Note: Do not award **R0A1**.
A p -value must be seen in part (f) to award **FT**.

[2 marks]

[Total: 15 marks]

5.

$$(a) \quad (T =) \begin{matrix} & \begin{matrix} (B) & (G) & (N) \end{matrix} \\ \begin{pmatrix} 0.945 & 0.015 & 0.02 \\ 0.05 & 0.965 & 0.03 \\ 0.005 & 0.02 & 0.95 \end{pmatrix} \end{matrix} \quad \text{M1A1A1}$$

Note: Accept the columns in any order. Accept the transpose of this matrix.

Award **M1** for a 3x3 matrix with all values between (but not including) 0 and 1, and all columns (or rows if transposed) adding up to 1, award **A1** for one correct row (or column if transposed) and **A1** for all rows (or columns if transposed) correct.

[3 marks]

$$(b) \quad (T^6 =) \begin{pmatrix} 0.72 & 0.077 & 0.098 \\ 0.24 & 0.83 & 0.16 \\ 0.035 & 0.098 & 0.74 \end{pmatrix} \quad \text{(M1)}$$

Note: Accept a transposed matrix.

multiplying their T^6 by a correct matrix of the initial populations (M1)

$$\begin{pmatrix} 0.72 & 0.077 & 0.098 \\ 0.24 & 0.83 & 0.16 \\ 0.035 & 0.098 & 0.74 \end{pmatrix} \begin{pmatrix} 26000 \\ 240000 \\ 50000 \end{pmatrix}$$

Note: Award this **M1** for a transposed T if used correctly in part (b) i.e. preceded by 1×3 matrix rather than followed by a 3×1 matrix.

$$= \begin{pmatrix} 42133 \\ 212205 \\ 61661 \end{pmatrix} \quad \text{(A1)}$$

so the expected population of the German side would be 212000 (212205) A1

Note: Award **M0M1A0A1** for an answer of 174000 (= 174031). This is the case when T^{30} has been used.

[4 marks]

continued...

Question 5 continued

$$(c) \quad (i) \quad \begin{pmatrix} 0.945 & 0.015 & 0.02 \\ 0.05 & 0.965 & 0.03 \\ 0.005 & 0.02 & 0.95 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

at least two of these three:

$$0.945u_1 + 0.015u_2 + 0.02u_3 = u_1$$

$$0.05u_1 + 0.965u_2 + 0.03u_3 = u_2$$

A1

$$0.005u_1 + 0.02u_2 + 0.95u_3 = u_3$$

and

$$u_1 + u_2 + u_3 = 1 \text{ (may be seen in part (c)(ii))}$$

A1

$$(ii) \quad (u) = \begin{pmatrix} 0.231 \\ 0.533 \\ 0.236 \end{pmatrix} \quad \left(u = \begin{pmatrix} 0.231155... \\ 0.532663... \\ 0.236180... \end{pmatrix} \right)$$

A1

Note: The **A1** in part (c)(ii) can be awarded independently of the working in part (c)(i).

[3 marks]

$$(d) \quad 0.532663... \times (26000 + 240000 + 50000) \\ = 168000 \text{ (168321...)}$$

(M1)**A1**

Note: Award **(M1)A1** for answers using T^n with n large that lead to a correct answer.
Award **(M0)A0** for answers that use T^n that lead to an incorrect answer.

[2 marks]

(e) Award **R1** for each appropriate reason. For example:

Movement unlikely to be constant

Total population for entire region likely to grow over time

Each power of the transition matrix takes five years; a relatively long time in terms of population movement.

There may be other/new external factors such as wars in other adjoining countries, leading to an influx of economic migrants.

R1R1

Note: Do not award **R1** for any response that shows a lack of understanding of the assumption that the total population remains constant.

[2 marks]**Total [14 marks]**

6. (a) slugs appear discretely / independently / randomly / at a constant (average) rate / mean is (approximately) equal to variance **R1R1**
[2 marks]
- (b) new ($m =$) 0.2×12 ($= 2.4$) (so $X \sim \text{Po}(2.4)$) **(A1)**
attempt to use a pdf (e.g. $P(X = 4)$) **(M1)**
 0.125 ($0.125408\dots$) **A1**
[3 marks]
- (c) $P(X < 3)$ **OR** $P(X \leq 2)$ **(A1)**
 0.570 ($0.569708\dots$) **A1**
[2 marks]
- (d) $P(X \geq 1) = 0.909282\dots$ **(A1)**
raising a probability to a power of 3 **(M1)**
 $0.909282\dots^3$
 $= 0.752$ ($0.751788\dots$) **A1**
- Note:** Award at most **(A1)(M1)(A0)** for a final answer of 0.751. Working may not be seen.
- [3 marks]
- (e) $H_0 : m = 2.4,$ **A1**
 $H_1 : m > 2.4$ **A1**
- Note:** The hypotheses may be written in words but must include reference to the mean (e.g. “number of snails” is not sufficient to award **A1**), and state clearly for H_1 that the mean increases.
- [2 marks]
- (f) **EITHER**
finding either $P(X \geq 7)$ or $P(X \geq 8)$ **(M1)**
 $(P(X \geq 7) =) 0.01160\dots$ **AND** $(P(X \geq 8) =) 0.00334\dots$ **A1**
- OR**
finding either $P(X \leq 7)$ or $P(X \leq 6)$ **(M1)**
 $(P(X \leq 7) =) 0.996661\dots$ **AND** $(P(X \leq 6) =) 0.988405\dots$ **A1**
- THEN**
so critical region is $X \geq 8$ **OR** $X > 7$ **A1**
- Note:** **(M1)A0A1** can be awarded for a correct answer that is unsupported.
- [3 marks]
- (g) $(0.75 \times 12 =) 9$ **(A1)**
 $P(X \leq 7 \mid m = 9)$ **(M1)**
 $= 0.324$ **A1**
[3 marks]
- Total [18 marks]**

2. (a) (i) because the (population) standard deviation(s) are unknown **A1**

Note: Ignore any references to sample size.

[1 mark]

- (ii) **EITHER**
he has no idea beforehand which way the difference would be if there is a difference **A1**

OR
he is only interested that there is a difference (not the direction) **A1**

[1 mark]

- (b) (i) **EITHER**
 $H_0: \mu_F = \mu_G$; $H_1: \mu_F \neq \mu_G$ **A1**

OR
 $H_0: \mu_D = 0$; $H_1: \mu_D \neq 0$ **A1**

Note: Accept an equivalent statement in words, must include mean and reference to “population mean” / “mean for all those taking the French exam” etc. for the first **A1** to be awarded. The terms “on average” and “generally” are also acceptable to indicate populations. Do not accept an imprecise “the means are equal”.

Do not accept “There is (no) (significant) evidence of a difference between μ_F and μ_G ” for either hypothesis or “There is (no) significant difference between marks in French and German”.

[1 mark]

- (ii) Generate a third column giving French mark – German mark or German mark – French mark. **(M1)**
e.g.

French mark	German mark	Difference
42	39	3
65	66	–1
82	71	11
...	...	...

$p\text{-value} = 0.153$. **A1**

[2 marks]

- (iii) The p -value gives the probability of seeing the observed difference in means (or a larger difference) assuming H_0 to be true. **A1**

Note: Do not accept “the probability that the data occurs by chance” or similar.

[1 mark]

- (iv) because $0.153 > 0.05$ **R1**
EITHER
 there is not (significant evidence of) a difference between the
 (population) means **A1**
OR
 fail to reject H_0 (accept “accept H_0 ”) **A1**

Note: Do not award **R0A1**.
 Remember to **FT** from part (b)(ii).
 Do not award the final **A1** if the null hypothesis in part (b)(i) is logically wrong (i.e. if the null and the alternative have been reversed or are nonsense) but this can be awarded if part (b)(i) is just poorly communicated.

[2 marks]

- (c) (i) $H_0: \rho = 0$; $H_1: \rho > 0$ **A1**

Note: Condone $H_0: \rho \leq 0$.

[1 mark]

- (ii) $p\text{-value} = 0.00286$ **A2**
 $0.00286 < 0.05$ **R1**
 he should conclude that the two sets of marks are (generally)
 positively correlated **A1**

Note: Allow **FT** from any test for correlation.
 Do not award **R0A1**.
 The final **R1A1** should follow through from their p -value.
 Do not award the final **A1** if the null hypothesis in part (c)(i) is wrong (i.e. if the null and the alternative have been reversed or are nonsense), but this can be awarded if part (c)(i) is just poorly communicated.
 The final conclusion must be in context.

[4 marks]

- (d) (i) the regression line of German on French is **(A1)**
 $\text{German} = 10.2393... + 0.737495... \text{French}$
EITHER
 substituting $\text{French} = 58$ into their regression line **(M1)**
OR
 sketch showing regression line and $x = 58$ **(M1)**
THEN
 Paul's German mark = 53 **A1**

Note: Accept an answer of 53.0 (53.0140...) or 52.9 as integer results are not explicitly stated in the question.
 Regression lines may be written in terms of y and x .

[3 marks]

- (ii) recognizing need to use line French on German
 French = $4.04116... + 1.01122... \text{German}$
 putting French = 71, Sue's German mark = 66

(A1)
 A1

Note: Accept an answer of 66.2 (66.2158...) or 66.3 as integer results are not explicitly stated in the question.
 Although not required in the markscheme as presented, candidates may have considered French = 70.5 and French = 71.5; this is valid and will lead to the correct answer.
 If the line German on French is used in part (d)(ii) the answer is 63; award **A0A0**.

[2 marks]

- (e) (i) **EITHER**
 the maximum value of τ occurs when all pairs are concordant so $\max = +1$
 the minimum value of τ occurs when all pairs are discordant so $\min = -1$

A1

OR

when all concordant $C - D = \frac{n(n-1)}{2}$, and when all discordant $C - D = -\frac{n(n-1)}{2}$

A1

OR

when all concordant $C = \frac{n(n-1)}{2}$, $D = 0$ and when all discordant $C = 0$, $D = \frac{n(n-1)}{2}$

A1

THEN

hence the range is $[-1, +1]$

AG

Note: Accept an answer which is just based on $n = 6$.

[1 mark]

- (ii) $(53 - 76)(41 - 70) > 0$
 Hence concordant

A1
 AG

[1 mark]

- (iii) Evidence of a valid method, eg

M1 $P_1: C, D, C, C, C$ $P_2: C, C, C, D$ $P_3: D, C, D$ $P_4: D, C$ $P_5: D$ $P_6:$ **Note:** At least one pair beyond (P_1, P_2) needs to be compared to award **M1**.

any evidence (a statement or a list) that 15 pairs need to be considered

A1

$$C = 9, D = 6$$

A1using their stated C and D values in given formula with $n = 6$ **M1**

$$\frac{2(9-6)}{6(6-1)} \text{ OR } \frac{9-6}{15}$$

$$\tau = 0.2$$

AG**[4 marks]**

- (f) (i)
- H_0
- : There is no (underlying) association (or correlation) between the two sets of marks

 H_1 : There is an (underlying) association (or correlation) between the two sets of marks**A1****Note:** Do not accept independence in the hypotheses.**[1 mark]**

- (ii)
- τ
- does not lie in the critical region
- OR**
- $0.2 < 0.733$

R1**EITHER**

there is insufficient evidence to indicate that there is an association between the two sets of marks

A1**OR**fail to reject H_0 (accept “accept H_0 ”)**A1****Note:** Do not award **R0A1**.In this question the final **A1** mark can be awarded for “fail to reject H_0 ” or “accept H_0 ” even if the hypotheses in (f)(i) are the wrong way round as the critical region is given.**[2 marks]**

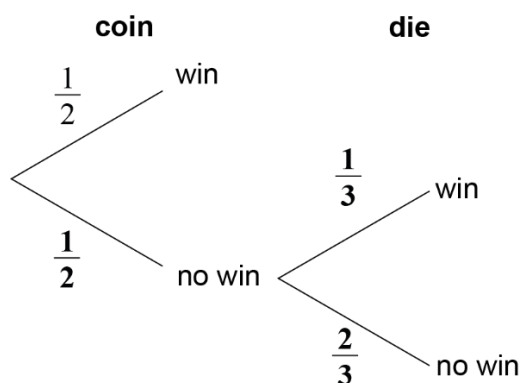
- (g) no

A1

because scaling the marks will not affect the concordances/ discordances

R1**Note:** Do not award **A1R0**.**[2 marks]****Total [29 marks]**

7. (a)

**A1A1**

Note: Award **A1** for completing first set of branches, **A1** for completing second set of branches.

[2 marks]

(b) attempt to multiply along the branches

(M1)

$$\frac{1}{2} \times \frac{2}{3}$$

$$= \frac{1}{3} \quad (= 0.333\dots)$$

A1**[2 marks]**(c) **EITHER**

$$\frac{\frac{1}{2}}{\frac{1}{2} + \left(\frac{1}{2} \times \frac{1}{3}\right)}$$

M1A1

Note: Award **M1** for recognizing conditional probability, **A1** for correct substitution.

OR

$$\frac{\frac{1}{2}}{1 - \frac{1}{3}}$$

M1A1

Note: Award **M1** for recognizing conditional probability, **A1** for correct substitution.

THEN

$$= \frac{3}{4}$$

A1**[3 marks]****[Total 7 marks]**

✓ ✗

10. (a) $M = 1000 \times t^{-0.6}$
 $a = 1000 (= 999.972...)$ $b = -0.600 (-0.599991...)$ **A1A1**

[2 marks]

(b) $y = -0.600x + 6.908$
 $c = -0.600$ $d = 6.908$ **A1A1**

Note: Long answer for c is 0.599991... and for d is 6.90772... If both answers are correct but not given to 3 decimal places award **A1A0**.

[2 marks]

(c) **METHOD 1 (starting with the result in part (b))**
 attempt to apply addition (or subtraction) log laws **M1**
 attempt to apply inverse log **M1**

Note: These **M1** marks can be applied in either order depending on the approach.

e.g. $\ln M = \ln t^{-0.600} + \ln e^{6.908}$ then $\ln M = \ln(e^{6.908} \times t^{-0.600})$

OR $\ln \frac{M}{t^{-0.600}} = 6.908$ then $\frac{M}{t^{-0.600}} = e^{6.908}$

$M = e^{6.908} t^{-0.600}$ **A1**

$(M = 1000.24 t^{-0.6})$

$M = 1000 t^{-0.6}$ and hence (close enough to be) equivalent **AG**

Note: The **AG** line (or something which approximates it) must be seen for the final **A1** to be awarded. If 3 sf answers are used from part (b), the coefficient is 1002; this can be condoned in the working, as it equals 1000 when rounded to 3 sf.

METHOD 2 (starting with the result in part (a))
 attempt to apply log **M1**

$\ln M = \ln(1000 \times t^{-0.6})$

attempt to apply addition (or subtraction) log laws **M1**

$\ln M = \ln 1000 + \ln t^{-0.6}$

$\ln M = 6.90775... - 0.6 \ln t$ **A1**

$(\ln M = y, \ln t = x)$

$y = -0.600x + 6.908$ and hence (close enough to be) equivalent **AG**

Note: The **AG** line (or something which approximates it) must be seen for the final **A1** to be awarded. Condone $b = -0.6$.

[3 marks]

[Total 7 marks]

15. (a) probability of non veg remaining non veg **A1**
[1 mark]
- (b) attempt to use $\det(A - \lambda I) = 0$ **(M1)**
- $$\begin{vmatrix} 0.8 - \lambda & 0.1 \\ 0.2 & 0.9 - \lambda \end{vmatrix} = 0$$
- $(0.8 - \lambda)(0.9 - \lambda) - 0.1 \times 0.2 = 0$ **(A1)**
- $\lambda = 1; \lambda = 0.7$ **A1**
[3 marks]
- (c) $-2a + b = 0$ **M1**
- $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ (accept any multiples of this answer) **A1**
- v_1 means that in the long term the ratio of veg to non-veg is 1:2
(in the long term one-third of students will be veg and two-thirds will not) **A1**
[3 marks]
[Total 7 marks]

16. (a) Let X be the random variable number of shots taken in a 12 minute period
 $X \sim \text{Po}(5)$ (A1)
 $P(X \leq 6) = 0.762$ (= 0.762183...) A1

[2 marks]

- (b) $P(\text{less than 4 shots} \cap \text{success at least once})$

METHOD 1

$$= P(\text{less than 4 shots}) - P(\text{less than 4 shots} \cap \text{zero success}) \quad (\text{M1})$$

Note: Might be communicated in Venn diagram.

$$\text{attempt to multiply by different powers of } 0.6 \quad (\text{M1})$$

$$= P(X \leq 3) - (P(X = 0) \times (0.6)^0 + P(X = 1) \times (0.6)^1 + P(X = 2) \times (0.6)^2 + P(X = 3) \times (0.6)^3)$$

$$= 0.414 \quad (= 0.413845...) \quad (\text{A1})$$

A1

METHOD 2

$$\text{attempt to multiply by different powers of } 0.4 \quad (\text{M1})$$

$$= P(X = 1) \times (0.4)^1 + P(X = 2) \times ((0.4)^2 + 2 \times 0.4 \times 0.6) + P(X = 3) \times ((0.4)^3 + 3 \times 0.4^2 \times 0.6 + 3 \times 0.4 \times 0.6^2)$$

(M1)(A1)

Note: Award **M1** for recognizing the six different cases, e.g. $2 \times 0.4 \times 0.6$ (etc.) or equivalent seen, **A1** for completely correct expression.

$$= 0.414 \quad (= 0.413845...) \quad \text{A1}$$

[4 marks]

[Total 6 marks]

2. (a) (i) 150 (cm) **A1**

(ii) attempt to substitute values in the mean formula with at least one mid-interval value multiplied by a corresponding frequency **(M1)**

(mean =) 176 (176.3) (cm) **A1**

[3 marks]

(b) 183 **OR** 168 seen **(A1)**

Note: These values may be seen in the working for part (c).

(IQR = 183 – 168 =) 15 (cm) **A1**

[2 marks]

(c) (upper bound =) $183 + 1.5 \times 15$ **OR** 205.5 seen **A1**

$205.5 > 204$ **OR** $204 - 183 < 22.5$ **OR** $204 - 22.5 < 183$ **R1**

Laszlo's height is not an outlier **A1**

Note: Do not award **R0A1**.

[3 marks]

(d) H_0 : The heights of the students can be modelled by $N(176, 13.5^2)$

H_1 : The heights of the students cannot be modelled by $N(176, 13.5^2)$ **A1A1**

Note: Award **A1** for each correct hypothesis that includes a reference to normal distribution with a mean of 176 and a standard deviation of 13.5 (or variance of 13.5^2). “Correlation”, “independence”, “association”, and “relationship” are incorrect.

Award at most **A0A1** for correctly worded hypotheses that include a reference to a normal distribution but omit the distribution's parameters in one or both hypotheses. Award **A0A1** for correct hypotheses that are reversed.

[2 marks]

- (e) (i) $h \sim N(176, 13.5^2)$
 attempt to find normal probability in either correct range **(M1)**
 $P(170 \leq h < 180)$ **OR** $P(h \geq 190)$
 recognition of multiplying either of their probabilities by 200 **(M1)**
 $0.288137... \times 200$ **OR** $0.149859... \times 200$
 $a = 57.6$ (57.6274...), $b = 30.0$ (29.9718...) **A1A1**
- (ii) $df = 4$ **(A1)**
 $(p =) 0.0166$ (=0.0166282...) **A1**
 comparing their p -value to 0.05 **R1**
 $0.0166 < 0.05$

Note: Accept p value of 0.0165 (= 0.0164693...) from using a and b to 3 sf.

(Reject H_0 There is sufficient evidence to say that) the data has not
 been drawn from the ($N(176, 13.5^2)$) distribution. **A1**

Note: Do not award **R0A1**.

The conclusion to part (e)(ii) **MUST** follow through from their hypotheses seen in part (d); if hypotheses are incorrect/reversed etc., the answer to part (e)(ii) must reflect this in order for the **A1** to be credited.

[8 marks]

[Total 18 marks]

7. (a) (let X be the random variable the weight of an individual in the city of Melba)
 $X \sim N(72, 10^2)$
 recognizing need to find $P(X > 85)$ (condone “86” for the **M1**) **(M1)**
 e.g. correct sketch of normal curve **OR** 0.0968 (= 0.0968005...) seen

let Y be the random variable the number of people more than 85 kg

attempt to use a binomial distribution **(M1)**
 $Y \sim B(10, 0.0968005...)$ **(A1)**

Note: This **(A1)** can be implied by the value 0.988580...

$(P(Y \geq 4) =) 0.0114 (= 0.0114196...)$ **A1**

[4 marks]

- (b) let W be the random variable the total weight of a sample of eight people
 $W \sim N(576, 8 \times 10^2)$

A1A1A1

Note: Award **A1** for normal distribution; **A1** correct mean; **A1** correct variance or SD ($SD = 28.2842\dots$).

[3 marks]

- (c) attempt to use inverse normal (or equivalent)
 $P(W > w) = 0.01$
 $(w =) 642 \text{ (kg)} \quad (641.799\dots)$

(M1)**A1****[2 marks]**

- (d) (i) *Any two correct assumptions identified,*
e.g.
 That Laetitia's clients are a random sample of the city's population
 That people take only one holiday a year
 That the choice of individual holidays is independent
 That Laetitia is her clients' only agent

A1A1

Note: Accept "assumes the proportion that takes a holiday abroad is 42%".

- (ii) $H_0 : p = 0.42$

A1

$$H_1 : p < 0.42$$

A1

- (iii) let Q be the random variable the number who go holiday abroad

$$Q \sim B(200, 0.42)$$

(A1)

$$(P(Q \leq 67) =) 0.00850 \quad (= 0.00849906\dots)$$

A1

$$0.00850 < 0.05$$

R1**EITHER**

there is evidence that Laetitia's claim is reasonable

A1**OR**

there is insufficient evidence to accept the newspaper's claim

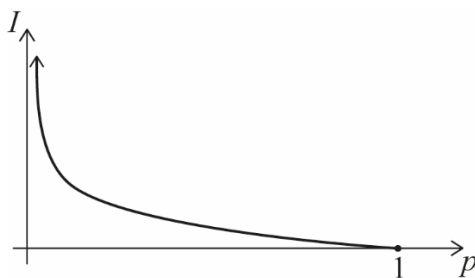
A1

Note: Follow through within this part, for correctly comparing and concluding with their **probability**, e.g. it is possible to award **A0A0R1A1**.

The conclusion to part (e)(iii) MUST follow through from their hypotheses seen in part (e)(ii); if hypotheses are incorrect/reversed etc., the answer to part (e)(iii) must reflect this in order for the **A1** to be credited.

[8 marks]**[Total 17 marks]**

2. (a) (i)

approximately correct shape (decreasing and convex) **A1**asymptotic behaviour to the I -axis labelled (e.g. arrow) / **A1**equation of asymptote ($p=0$) seen **A1** p -intercept labelled at $p=1$ and graph does not extend beyond **A1**

Note: Condone I and p being labelled as y and x .
Do not award second **A1** if y -intercept label is seen.

[3 marks]

(ii)
$$\frac{dI}{dp} = -\frac{1}{p}$$

A1A1

Note: Award **A1** for $\frac{dI}{dp}$ or equivalent (e.g. I') and **A1** for $-\frac{1}{p}$. Do not accept $\frac{dy}{dx}$ for the first **A1** unless followed by $-\frac{1}{x}$, which can earn **A1A0**.

(for $0 < p \leq 1$) we therefore have $\frac{dI}{dp} < 0$ **R1**hence the function is decreasing **AG****[3 marks]**(iii) Any plausible interpretation IN CONTEXT. **R1***For example:*

More information is gained from a rarer event.

Less information is gained from a more common event.

Information (gained) decreases as probability increases.

[1 mark]

(b) (i)
$$\frac{5}{10} \quad (= \frac{1}{2})$$

A1**[1 mark]**(ii) attempt to substitute $p = \text{their (b)(i)}$ into $I = -\ln p$ **(M1)**

$$= 0.693 \quad (0.693147\dots, -\ln(\frac{1}{2}))$$

A1**[2 marks]**

– 12 –

8823 – 7203M

$$(iii) \quad \frac{9}{10} \quad (A1)$$

$$= 0.105 \quad (0.105360..., -\ln\left(\frac{9}{10}\right)) \quad A1$$

[2 marks]

- (c) (i) attempt to substitute into the formula for $E(I)$ and recognise that $n=2$ (or two terms are needed) **M1**

$$E(I) = -\frac{1}{10} \ln\left(\frac{1}{10}\right) - \frac{9}{10} \ln\left(\frac{9}{10}\right) \quad A1$$

$$0.325 \quad (0.325082...) \quad AG$$

[2 marks]

$$(ii) \quad E(I) = -\frac{1}{2} \ln\left(\frac{1}{2}\right) - \frac{1}{2} \ln\left(\frac{1}{2}\right) \quad A1$$

$$0.693 \quad (0.693147..., \ln(2)) \quad A1$$

$$0.693 > 0.325 \quad AG$$

[2 marks]

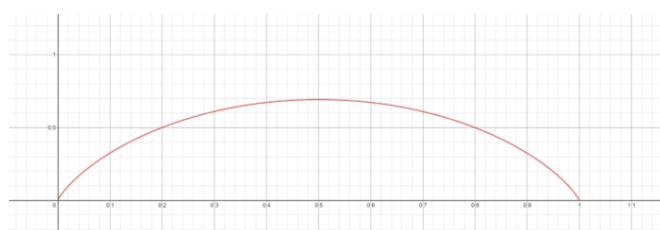
(d) (i) $(I) = -\ln(1-p)$ **A1**

[1 mark]

$$(ii) \quad (E(I)) = -p \ln p - (1-p) \ln(1-p) \quad A1$$

[1 mark]

(iii) attempt to use graphical method or calculus to maximize $E(I)$ **M1**



maximum occurs when $p = \frac{1}{2}$ **A1**

[2 marks]

(e) (i) $(x =) \frac{6}{12} \left(= \frac{1}{2} \right)$ **A1**

EITHER

for the scales to balance, the odd ball must be in the six balls not chosen **R1**

OR

for the scales to balance, all the balls chosen must be of equal weight and hence

$$\frac{11}{12} \times \frac{10}{11} \times \frac{9}{10} \times \dots \times \frac{6}{7}$$
R1

[2 marks]

(ii) **EITHER**

recognition that the sum of the probabilities on the third row of the table equals 1 **(M1)**
e.g. $x + 2y = 1$

OR

for one side to be heavier, the odd ball must be one of six balls chosen

$\left(\frac{6}{12} \right)$ and half the time this will result in left-side being heavier,

therefore $y = \frac{6}{12} \times \frac{1}{2}$ **(M1)**

$$y = \frac{1}{4}$$

A1

[2 marks]

(iii) $z = -\frac{1}{6} \ln \frac{1}{6} - \frac{1}{6} \ln \frac{1}{6} - \frac{2}{3} \ln \frac{2}{3}$ **(M1)**

$$= 0.868 \quad (0.867563\dots)$$

A1

[2 marks]

(iv) 4 balls on each side because that configuration has the largest $E(I)$ **R1**

Note: Award **R1** for giving a correct reason **AND** stating “4 balls on each side”

[1 mark]

[Total: 27 marks]

2. (a) (i) $r=0.995$ (0.994705...)

A2

Note: Award **A1** for 0.99.

- (ii) $m = 10.6t + 43.9$ (10.6032... t + 43.8780...)

A1A1

Note: Second **A1** is for the correct variables.

[4 marks]

- (b) **EITHER**

$$10.6032... \times 1.5$$

(M1)

OR

$$(10.6032...(t+1.5) + 43.8780...) - (10.6032...t + 43.8780...)$$

(M1)

THEN

$$15.9 \text{ (marks)} \text{ (15.9048...)}$$

A1

Note: Accept 16.

[2 marks]

- (c) *Accept any valid reason*

R1

e.g:

The students in the sample might not be of equal ability / she has not controlled for ability

She might have originally obtained close to full marks so an extra 15.9 would not be possible.

[1 mark]

[Total: 7 marks]

3. (a) (i) attempt to rearrange to isolate C (M1)
e.g., subtracting 32 or dividing the equation by 1.8

$$C = \frac{5}{9}(F - 32) \quad \left(C = \frac{F - 32}{1.8}, C = 0.556F - 17.8 \right) \quad \text{A1}$$

Note: If the answer is not written as an equation, award at most **M1A0**.

(ii) $C = \left(\frac{77 - 32}{1.8} = \right) 25 \text{ } (^{\circ}\text{C})$ A1

[3 marks]

(b) (i) $(1.8 \times 17 + 32 =) 62.6 \text{ } (^{\circ}\text{F})$ A1

(ii) recognizing that the “+32” does not affect the SD (M1)
 $(1.8 \times 9 =) 16.2 \text{ } (^{\circ}\text{F})$ A1

Note: Award **MOA0** for $1.8 \times 9 + 32 (= 48.2)$.

[3 marks]

[Total 6 marks]

12. (a) $40 = 100 - 100 \times 2^{-t}$ (A1)
 0.737 (hours) (0.736965...) A1

Note: Accept 44.2 minutes.

[2 marks]

- (b) values of P are 50, 75, 87.5, 93.75 (A1)

$$SS_{res} = (50 - 48)^2 + (75 - 74)^2 + (87.5 - 86)^2 + (93.75 - 91)^2 \quad (M1)(A1)$$

$$= 14.8 \quad (14.8125) \quad A1$$

[4 marks]

- (c) (i) The sum of the square residuals is smaller so it is a better fit R1

- (ii) Accept a valid argument in favour of model P or against the quadratic model. R1

e.g. any one of the following:

quadratic has no asymptote

quadratic will begin to go down

quadratic will become negative

quadratic might not go through (0, 0)

model P is the manufacturer's model

[2 marks]

[Total 8 marks]

17. (a) $\begin{pmatrix} 0.25 & 0.25 \\ 0.1 & 0.1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ (M1)

Note: Accept equivalent methods including only using one line of the matrix.

$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (or any multiple) A1

[2 marks]

(b) $D^n = \begin{pmatrix} 0.65^n & 0 \\ 0 & 1 \end{pmatrix}$ A1

[1 mark]

(c) $\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0.65^n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix}^{-1}$ (M1)

EITHER

multiplying by the initial state

(M1)

$$\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0.65^n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 7000 \\ 0 \end{pmatrix}$$

$\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0.65^n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2000 \\ 1000 \end{pmatrix}$ (A1)

$\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2000 \times 0.65^n \\ 1000 \end{pmatrix}$ (A1)

$\begin{pmatrix} 2000 \times 0.65^n + 5000 \\ -2000 \times 0.65^n + 2000 \end{pmatrix}$ (A1)

Note: Award A0 if either term in the matrix is incorrect.

OR

$$\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix}^{-1} = \frac{1}{7} \begin{pmatrix} 2 & -5 \\ 1 & 1 \end{pmatrix}$$

A1

$$\begin{pmatrix} 1 & 5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0.65^n & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.65^n & 5 \\ -0.65^n & 2 \end{pmatrix}$$

A1

Note: The preceding **A1** marks can be awarded independently.

$$\frac{1}{7} \begin{pmatrix} 5 + 2 \times 0.65^n & 5 - 5 \times 0.65^n \\ 2 - 2 \times 0.65^n & 2 + 5 \times 0.65^n \end{pmatrix}$$

A1

Note: Award **A0** if any term in the matrix is incorrect.

multiplying by the initial state

(M1)

$$\frac{1}{7} \begin{pmatrix} 5 + 2 \times 0.65^n & 5 - 5 \times 0.65^n \\ 2 - 2 \times 0.65^n & 2 + 5 \times 0.65^n \end{pmatrix} \begin{pmatrix} 7000 \\ 0 \end{pmatrix}$$

THEN

$$2000 - 2000 \times 0.65^n \quad (= 2000(1 - 0.65^n))$$

A1

Note: For the final **A1**, follow through within the question part from the bottom-left entry of their 2x2 matrix or the bottom entry of their 2x1 matrix but only if “in terms of n ”.

If any mistake in the matrices is seen that DOES NOT affect the correct final answer, do not award the penultimate **A1** mark.

[6 marks]**[Total 9 marks]**

2. (a) (i) recognition of binomial distribution (condone incorrect parameter) **(M1)**
 e.g. $M \sim B(20, 0.04)$ **OR** $P(M = 2) = \text{binpdf}(20, 0.04, 2)$
 $= 0.146$ (0.145799...) **A1**
- (ii) recognition the cumulative probability required **(M1)**
 e.g. $P(M \geq 3) = 1 - \text{bincdf}(20, 0.04, 2)$ **OR** $\text{bincdf}(20, 0.04, 3, 20)$
 $= 0.0439$ (0.0438627...) **A1**
[4 marks]
- (b) either one of two terms in expected value formula correct **(M1)**
 $50(20(0.96)) + 15(20(0.04))$
 $= 972$ (pesos) **A1**
[2 marks]
[Total: 6 marks]

4. (a) recognizing that only way to score 7 is to achieve a head and a 6 on die (M1)
e.g. $\frac{1}{6}$ and $\frac{1}{2}$ seen in an attempt to combine probabilities

$$\left(\frac{1}{6} \times \frac{1}{2} =\right) \frac{1}{12} \text{ (0.0833333...)} \quad \text{A1}$$

Note: Condone 0.0835 from the use of 0.167.

[2 marks]

- (b) there are two ways to score (e.g.) 5 (M1)
achieve a head and a 4 on die, or a tail and a 5 on die

$$\left(2\left(\frac{1}{6} \times \frac{1}{2}\right) =\right) \frac{2}{12} \left(\frac{1}{6}, 0.167, 0.16666...\right) \quad \text{A1}$$

Note: Award these marks for equivalent working for the 2, 3, 4 or 6 point scenarios.

Final Score	1	2	3	4	5	6	7
Probability	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$

A1

Note: Award **A1** for a completely correct table. Award at most **(M1)A1A0** if their follow-through answer from part (a) leads to a total probability not equal to 1.

[3 marks]

- (c) **EITHER** (M1)
multiplying at least two columns from their table

$$1 \times \frac{1}{12} + 2 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6} + 7 \times \frac{1}{12}$$

OR

recognizing the probabilities in the table are symmetric (M1)

OR (for HL markscheme and/or removed from SL at publication)

Considering the sum of two random variables (M1)

$$E(X + Y) = E(X) + E(Y) \quad (= 3.5 + 0.5)$$

THEN

(expected value =) 4 A1

Note: Accept 4.01 (4.00640...) from use of their 3 sf values from (b).
Award at most **M1A0** if their final answer is not in the range 1 – 7

[2 marks]

[Total: 7 marks]

9. (a) (i) let X be number of accidents per week $\Rightarrow X \sim \text{Po}(0.76)$

recognition cumulative Poisson distribution

$$P(X \geq 2) = 1 - P(X \leq 1)$$

$$= 0.177 \text{ (0.176907...)}$$

(M1)

A1

- (ii) let F be number of accidents per 4-week $\Rightarrow F \sim \text{Po}(3.04)$

$$P(Y = 3) = 0.224 \text{ (0.223982...)}$$

(M1)

A1

[4 marks]

- (b) binomial model

A1

with $p = 0.177$ **OR** $p = 1 - 0.177 (= 0.823)$ and $n = 8$

A1

[2 marks]

[Total: 6 marks]

14. (a) room temperature / the temperature below which the hot water will not cool **A1**

Note: Accept answers similar to “the temperature of the water in the cup after 25 minutes”.

[1 mark]

- (b) evidence of subtracting 25 from the temperature data **(M1)**
 $a = 244$ (243.920...) and $b = -1.03$ (-1.02965...) **A1A1**

Note: Award **M0A1A0** for both answers $a = 116.764...$ and $b = -0.375218...$ seen, from not subtracting the 25.

[3 marks]

- (c) $k = 61.1$ (61.0848...) and $c = 0.923$ (0.923029...) **A1**

Note: If not subtracting 25 was penalized in part (b), award **A1** for $k = 68.9023...$ and $c = 0.972506...$ seen in part (c).

[1 mark]

- (d) Award **R1** for any appropriate reason that supports Soo Min’s model
 e.g.

EITHER

compare r^2 values: 0.76573583... and 0.9233117... **R1**

OR

consider value at $t = 0$ / “water cannot reach a temperature more than 100 degrees”

R1

Note: If subtracting 25 was penalized in (b) then award **R1** for comparing r^2 values 0.904086... and 0.994522...

Award **R1** for correct reasoning seen for the use of comparison of the sum of square residuals.

[1 mark]

[Total: 8 marks]

5. (a) $\begin{pmatrix} 0.88 & 0.08 \\ 0.12 & 0.92 \end{pmatrix}$ **M1A1**

Note: Award **M1** for correct values used, **A1** if in correct positions.
Accept alternative consistent matrix (e.g. the transpose or diagonal elements exchanged) and follow through to eigenvectors and initial state vector.

[2 marks]

(b) 5 (seen) **(A1)**
 $\begin{pmatrix} 0.88 & 0.08 \\ 0.12 & 0.92 \end{pmatrix}^5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.596608 \\ 0.403392 \end{pmatrix}$ **OR** $\begin{pmatrix} 0.88 & 0.08 \\ 0.12 & 0.92 \end{pmatrix}^5 = \begin{pmatrix} 0.596608 & 0.268928 \\ 0.403392 & 0.731072 \end{pmatrix}$
(M1)
A1
 $P(\text{Friday evening}) = 0.403 \quad (0.403392)$

Note: Award **A0M1A0** for use of 4 (and resulting probability 0.354).

[3 marks]

(c) attempt to find $\det(A - \lambda I)$ **(M1)**
 $\begin{vmatrix} 0.88 - \lambda & 0.08 \\ 0.12 & 0.92 - \lambda \end{vmatrix}$ **OR** $(0.88 - \lambda)(0.92 - \lambda) - (0.12)(0.08)$
 $\lambda^2 - 1.8\lambda + 0.8$ **A1**

[2 marks]

(d) eigenvalues are 0.8 and 1 **(A1)**

Note: If no attempt is made to find eigenvectors, do not award **A1** for finding eigenvalues.

$$\begin{pmatrix} 0.88 & 0.08 \\ 0.12 & 0.92 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.8 \begin{pmatrix} x \\ y \end{pmatrix}$$

$$0.88x + 0.08y = 0.8x$$

$$\text{eigenvector} = \text{eg. } \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

A1

EITHER

$$\begin{pmatrix} 0.88 & 0.08 \\ 0.12 & 0.92 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \begin{pmatrix} x \\ y \end{pmatrix}$$

(M1)

$$0.88x + 0.08y = x$$

$$0.08y = 0.12x$$

OR

eigenvalue 1 gives

$$\begin{pmatrix} -0.12 & 0.08 \\ 0.12 & -0.08 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(M1)

$$-0.12x + 0.08y = 0$$

$$0.08y = 0.12x$$

Note: Award **M1** for an attempt to find the eigenvector with eigenvalue 1.

THEN

$$\text{eigenvector} = \text{eg. } \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

A1

Note: Award **A0A1M0A0** if only $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is seen and no eigenvalues are found.

[4 marks]

$$(e) \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0.8 \end{pmatrix}, P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \quad \text{OR} \quad D = \begin{pmatrix} 0.8 & 0 \\ 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad \textbf{A1A1}$$

Note: Award **A1** for one of **P** or **D** correct. Do not award the second **A1** unless **P** and **D** are consistent.

[2 marks](f) **EITHER**

$$\text{attempt to use } T^n = (PDP^{-1})^n = PD^nP^{-1}$$

M1

Note: Award **M1** for their **Dⁿ** seen.

limit of **Dⁿ** calculated**A1**

$$\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}^{-1}$$

Note: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ must be seen to award **A1**.

ORattempt to expand their **PDⁿP⁻¹** using explicit **P**, **P⁻¹****M1**

$$(T^n =) \frac{1}{5} \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.8^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$

$$(T^n =) \frac{1}{5} \begin{pmatrix} 2+3(0.8^n) & 2-2(0.8^n) \\ 3-3(0.8^n) & 3+2(0.8^n) \end{pmatrix}$$

A1

Note: Using this method, the limit of 0.8^n may be inferred and **M1A1** awarded.

THEN

0.6

A1

Note: Multiplication by initial condition $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ may be seen at any point as part of their method.

For an answer of 0.6 from incomplete methods award a maximum of **M1A0A0**, or if no working is seen, award **M0A0A1**.

[3 marks]**[Total: 16 marks]**

7. (a) $P(X < 340)$ **OR** labelled sketch of region **OR** calc syntax with correct bounds
 $= 0.159$ (0.158656...) (M1) A1 [2 marks]
- (b) recognizing endpoint occurs at either 0.975 or 0.025 (M1)
 $P(X < k) = 0.975$ **OR** $P(X < m) = 0.025$
 $330 < X < 370$ (330.400... < X < 369.599...) A1A1 [3 marks]
- (c) (i) recognizing mean of W is sum of individual means within wall (M1)
 $W = C_1 + C_2 + L$ may be seen
 $E(W) = 2E(C) + E(L)$
 $= 800$ A1
 recognizing **variance** of W is sum of individual **variances** within wall (M1)
 $\text{Var}(W) = 2\text{Var}(C) + \text{Var}(L)$ **OR** 225 seen (A1)
 $(\text{SD}(W) =) 15$ A1
- Note:** Award **M1A0A0** for an answer of 20.6 from using $\text{Var}(2C)$ in place of $2\text{Var}(C)$.
- (ii) recognizing that W is modelled by a normal distribution (M1)
 $(P(780 < W < 810) =) 0.656$ (0.656296...) A1
- Note:** The answer is 0.521 (0.520) from using $\text{SD} = 20.6$ ($5\sqrt{17}$). Follow through from part (c)(i) without working seen.
- [7 marks]
- (d) $810 = 350 + 350 + E(L)$ (or equivalent) (A1)
 $(E(L) =) 110$ A1
 $\text{Var}(W) = 2\text{Var}(C) + \text{Var}(L)$ **OR** $256 = 2(100) + \text{Var}(L)$ (A1)
 $(\text{SD}(L) =) 7.48$ (7.48331..., $\sqrt{56}$) A1 [4 marks]
- (e) 116 (116.298) A1
- Note:** Do not follow through from either a negative variance or a negative SD.
- [1 mark]
[Total: 17 marks]

1. (a) (i) $141 - 88$
 $= 53$ **A1**
AG
[1 mark]
- (ii) $(120 - 53 - 54 =) 13$ **A1**
[1 mark]
- (b) Restricting the size of the sample space to 150 ($54 + 96$) **(M1)**
 $= 0.36 \left(\frac{54}{150}, \frac{9}{25} \right)$ **A1**
[2 marks]
- (c) (i) H_0 : The development of cracks and the technique used are independent
 H_1 : The development of cracks and the technique used are not independent **A1**
- Note:** Condone equivalent statements such as ‘not dependent’ but do not accept “uncorrelated” or “not related” in place of “independent”.
- [1 mark]
- (ii) (p -value =) 0.0170 (0.0169864...) **(M1)A1**
[2 marks]
- (iii) $0.0170 < 0.05$ **R1**
- hence there is sufficient evidence to reject the null hypothesis that the development of cracks and the technique used are independent. **A1**
- Note:** Do not award **R0A1**.
- [2 marks]
- (d) $\frac{96 + 56}{240} \left(= \frac{152}{240} \right)$ **A1**
 $\frac{19}{30}$ **AG**
[1 mark]
- continued...

Question 1 continued.

(e) **EITHER**

the probability of each component developing cracks is independent of all the other components in the sample.

R1

Note: Do not accept the word “independence” on its own. Appropriate context must be seen.

OR

the development of cracks can be partitioned into two clear groups

R1

Note: Do not accept ‘the samples are representative of the population’.

[1 mark]

(f) 67 seen

(A1)

EITHER

attempt to find a probability ≤ 67 (condone strict inequality for **(M1)**)
 $(P(X \leq 67) =) 0.0549 \text{ (} 0.0549093 \dots \text{)}$

(M1)
A1

Note: Award **(A1)(M1)A0** for an unsupported $p = 0.0372$, from use of strict inequality.

$0.0549 > 0.05$

R1

OR

attempt to find the critical region

(M1)

critical region is $X \leq 66$

A1

$66 < 67$ or ‘67 is not in the critical region’

R1

THEN

EITHER

do not reject the null hypothesis (as there is insufficient evidence that the new technique reduces the number of cracks).

A1

OR

do not accept the alternative hypothesis (as there is insufficient evidence that the new technique reduces the number of cracks).

A1

continued...

Question 1 continued.

Note: Do not award **R0A1**.
Only follow through within part (f) for final **R1A1** if the **(M1)** has been awarded.
Do not condone “accept the null hypothesis”.

[5 marks]

- (g) (i) the test for a proportion is directional and so considers whether the new treatment reduces the number of components developing cracks.

R1

[1 mark]

- (ii) **EITHER**

there could be variation in the value of p chosen for the null hypothesis /
the value of p from the sample might not be a representative of
the current technique

R1

OR

the test in (f) does not treat minor and major cracks as different attributes /
the test in (c) does treat minor and major cracks as different attributes

R1

OR

the test in (f) has to make an additional assumption (for example ‘independence’) **R1**

[1 mark]

continued...

Question 1 continued.

(h) **EITHER**

let μ_1 be the mean length of time before cracks appear with the new technique
and μ_2 be the mean length with the current technique

$$H_0: \mu_1 = \mu_2$$

A1

$$H_1: \mu_1 > \mu_2$$

A1

Note: Award **A1A0** for correct hypotheses in which the two population means are not clearly defined (e.g. unsupported μ_1 and μ_2).

OR

H_0 : the POPULATION mean length of time before cracks appear is the same for both groups

A1

H_1 : the new technique increases the POPULATION mean length of time before cracks appear.

A1

OR

H_0 : the mean length of time before cracks appear in ALL components made with the new technique is the same as for ALL components made with the current technique.

A1

H_1 : the mean length of time before cracks appear in ALL components made with the new technique is greater than the mean for ALL components made with the current technique.

A1

Note: Award **A1A0** if “population” (or equivalent, such as “all”) is omitted from an otherwise correct answer.

THEN

recognition of the need to use of a two-sample test

(M1)

$$p\text{-value} = 0.0162 \text{ (0.0162328...)}$$

A2

Note: If not pooled, answer is 0.0164368... award **(M1)A2**.

$$0.0162 < 0.05$$

R1

reject the null hypothesis (**OR** accept the alternative hypothesis)
(there is sufficient evidence to that the new technique increases the mean length of time before the cracks appear)

A1

Note: Do not award **R0A1**.

Follow through within part (h) for the last **R1A1**, provided their p -value is between 0 and 1 inclusive.

[7 marks]

continued...

Question 1 continued.

(i) **EITHER**

(though statistically significant) the new technique only seems to increase the time before cracks appear by 1 hour out of 250, so it is not a significant increase (i.e. the effect size is small)

R1

OR

the minimum time (not mean time) before cracks appear should be considered given the context / An appropriate confidence interval should be considered, and not simply the mean.

R1

Note: If a not significant p -value was seen in part (h), do not award **R1** for an answer of “the result is not significant” in part (i).

[1 mark]

[Total 26 marks]

1. (a) $a = 70.35$ **OR** 70.4
 $b = 73.45$ **OR** 73.5
 $c = 5.45$

A1
A1
A1

Note: Follow through within the question from their (labelled) value of b , when finding c (applies both if b is incorrect or correct to a different level of accuracy).

[3 marks]

- (b) **EITHER**
 median/average dart score has increased
OR
 Darts players are showing (some) improvement
OR
 Spread in average dart score has decreased
OR
 Players have become more similar

R1
R1
R1
R1

Note: Answers without context “e.g. the median is higher in year 2” score **R0**.

[1 mark]

- (c) **EITHER**
 (difference in averages is approximately) normally distributed
OR
 the individuals' changes in mean are independent

R1
R1

Note: Do not award **R1** if there is also a reference to variances being pooled / equal variances or to independence between the two years or anything else which shows a misunderstanding of the test used.

[1 mark]

- (d) (i) let μ_d represent the mean of the differences from table 1
 (e.g. column 2-column 1)

$$H_0 : \mu_d = 0$$

A1

$$H_1 : \mu_d > 0$$

A1

Note: Condone an alternative hypothesis of $H_1 : \mu_d < 0$, provided it is consistent with their subsequent work (e.g. their “differences” are year 1 – year 2). Accept hypotheses communicated in words, e.g. “ H_0 : there has been no change in the mean of the players' averages”. The word “mean” must be seen in any word-based response.
 Allow $\mu_1 = \mu_2$ and $\mu_2 > \mu_1$.
 Allow a compound null hypothesis (i.e. $\mu_d \leq 0$ or equivalent).
 Condone $\mu = 0$ or $\mu_0 = 0$ for null hypothesis and equivalent for alternative hypothesis.

[2 marks]

- (ii) finding/entering differences in ‘means’ across both years
 p -value = 0.134 (0.133797...)

(M1)
A1

[2 marks]

- (iii) $0.134 > 0.1$,
(so do not reject H_0)
there is insufficient evidence to suggest that the darts players' averages have increased

R1

A1

Note: Do not award **R0A1**. Award **A0** if the conclusion is not in context or if the null hypothesis was logically incorrect (i.e. an error of confusing null and alternative rather than just an error in communication).

[2 marks]

- (e) (i) decrease the significance level

A2

Note: Award **A1** if type I error has been defined (e.g. As rejecting H_0 when H_0 is true) but correct answer not seen.

Increased sample size is incorrect.
If more than one answer is given, mark only the first one.

[2 marks]

- (ii) increase the sample size / increase the significance level

A2

Note: Condone "Collect more data".
If more than one answer is given, mark only the first one.

[2 marks]

(f)

Year 1 mean	Year 2 mean	Difference	<i>P</i>	<i>N</i>
68.1	70.1	2.0	6	
72.2	72.0	−0.2		1
65.8	65.0	−0.8		2
73.7	74.9	1.2	3	
69.8	68.0	−1.8		5
64.5	68.0	3.5	7	
69.2	70.6	1.4	4	
64.9	92.9	28	8	
		Total	28	8

- (i) $D = 4$
 $B = 5$
 $A = 6, C = 7$

A1

A1

A1

[3 marks]

- (ii) $(\sum N) = 8$

A1

[1 mark]

- (g) (i) (H_1 :) the (population's) median (three dart average) increased (from year 1 to year 2). **A1**
[1 mark]
- (ii) ($T =$) 8 **A1**
[1 mark]
- (iii) ($T =$) $8 > 5$ **R1**

Note: Condone $8 \geq 5$.

(so do not reject H_0 (allow 'accept H_0 '))
there is insufficient evidence to suggest that the median of the players' averages
have increased **A1**

Note: Do not award **R0A1**. Award **A0** if the conclusion is not in context.

Nearly anything better than "Do not reject H_0 " will be accepted as being in context.

[2 marks]

- (h) **EITHER** **R1**
she could repeat the test(s) the following year
- OR** **R1**
she could repeat the test(s) with a different sample (for the same years)
- OR** **R1**
apply test-retest

Note: Increasing sample size is NOT a test for reliability. Any reference to this
would score **R0**.

[1 mark]

Total [24 marks]

3. (a) recognizing need to find $P(W > 3.5)$ **OR** $1 - P(W < 3.5)$ (M1)

$$(P(W > 3.5) =) 0.994 \text{ (0.993790...)} \quad \text{A1}$$

Note: Award **M1A0** for an answer of 0.99 with no working.

[2 marks]

- (b) (i) appropriate use of the GDC to find either a or b seen in either (b)(i) or (b)(ii) (M1)
 $a = 3.99 \text{ (3.98737...)} \quad \text{A1}$

$$(ii) \quad b = 5.01 \text{ (5.01262...)} \quad \text{A1}$$

Note: The working for (M1) may be seen and awarded in part (b)(ii) if part (b)(i) has not been answered.

[3 marks]

- (c) multiplying two identical probabilities (M1)
 $P(W < 3.5) \times P(W < 3.5)$ **OR** $(1 - \text{their (a)}) \times (1 - \text{their (a)})$

Note: Accept a lower bound of zero for W , given the context.

$$(0.00620967)^2$$

$$0.0000386 \text{ (0.0000385601... OR } 3.86 \times 10^{-5} \text{)} \quad \text{A1}$$

Note: Award **M0A0** for an answer greater than 1 such as 3.86.
 Award **M1A0** for 3.86E-5.

[2 marks]

[Total: 7 marks]

15. (a) normal distribution A1
 Mean = 15 (cm) A1
 Standard deviation = 0.4 (cm) **OR** variance = 0.16 (cm²) A1
[3 marks]
- (b) $E(W) = 3$ (A1)
 $\text{Var}(W) = 0.4^2 + 0.25^2$ (A1)
 $= 0.223$ (0.2225) (A1)
 W is normally distributed (M1)
 $(P(W < 2.5) =) 0.145$ (0.144572...) A1

Note: Accept $P(0 < W < 2.5)$.

[5 marks]
[Total 8 marks]

1. (a) (i) (
- $Q_3 =$
-) 7.2

A2

Note: Award **A1A0** if the lower quartile, 6.5, is given as the answer.
Award **A1A0** for a correct ordered list of happiness scores,
when the correct Q_3 is not seen.

(ii) $Q_1 = 6.5$

(A1)

$$IQR = 7.2 - 6.5$$

$$= 0.7$$

A1

[4 marks]

(b) $Q_3 + 1.5 \times IQR$

(A1)

$$(7.2 + 1.5 \times 0.7 =) 8.25$$

A1

since $7.5 < 8.25$

R1

Switzerland is not an outlier

AG

Note: Do not award **A0A0R1**.

[3 marks]

(c) (i) $a = 3.5$

A1

(ii) $b = 8$

A1

(iii) $c = 3.5$

A1

[3 marks]

(d) (i) ($r_s =$) 0.164 (0.164134...)

A2

(ii) France rank (of sixth) is unchanged (so the r_s is unchanged)

R1

[3 marks]

(e) Because r_s is too close to zero and hence Jose's conclusion is not appropriate

R1

Note: Award **R1** for a comment on the value r_s and "not appropriate" as a conclusion.

Accept " r_s indicates a weak correlation of the ranks".Do not accept " r_s indicates a weak correlation of happiness score and country rank".

[1 mark]

[Total 14 marks]

1. (a) evidence of using the nearest neighbour algorithm **(M1)**
 HAED(CBH) **OR** $40 + 50 + 60 + \dots$
 $40 + 50 + 60 + 90 + 80 + 85$ **(A1)**
 upper bound = 405 **A1**

[3 marks]

- (b) (i) evidence of using Prim's algorithm to find MST **M1**
 AE, DE, CD, ... **OR** $50 + 60 + 90 + \dots$
 AE, DE, CD, BC **OR** $50 + 60 + 90 + 80$ **A1**
 $= 280$ **A1**

Note: Award at most **M1A0A1** for using Prim's from an incorrect starting point.
 Award **M0A0A1** for obtaining 280 using Kruskal's or if no working seen.

[3 marks]

- (ii) evidence of reconnecting H with two edges to MST **(M1)**
 $280 + 40 + 70$
 lower bound = 390 **A1**

[2 marks]

- (c) (i) 0.4 **A1**
[1 mark]
 (ii) 0.3 **A1**
[1 mark]
 (iii) 0.2

Note: **FT** from 0.6 – their p .

A1
[1 mark]

continued...

Question 1 continued.

(d) (i) $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

A1
[1 mark]

(ii) recognition to find T^4 or s_4

(M1)

$$T^4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.1025 \\ 0.1832 \\ 0.2162 \\ 0.2513 \\ 0.2468 \end{pmatrix}$$

0.216 (0.2162)

A1
[2 marks]

(iii) recognition to use T^3

(M1)

recognition of two cases DD or EE

(M1)

$$0.7T^3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.0749 \\ 0.1274 \\ 0.1526 \\ 0.1736 \\ 0.1715 \end{pmatrix} \quad \text{AND} \quad 0.3T^3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.0276 \\ 0.0558 \\ 0.0636 \\ 0.0777 \\ 0.0753 \end{pmatrix}$$

A1A1

$(0.1736 + 0.0753) = 0.249$ (0.2489)

A1
[5 marks]

continued...

Question 1 continued.

(e) (i) attempt to find steady state matrix

$$\begin{pmatrix} 0.104072\dots \\ 0.190045\dots \\ 0.226244\dots \\ 0.235294\dots \\ 0.244343\dots \end{pmatrix}$$

(M1)

Note: May be seen in part (e)(ii).

$$0.104 \ (0.104072\dots) \left(\frac{23}{221}\right)$$

A1

[2 marks]

(ii) attempt to multiply distances by probabilities from steady state matrix

(M1)

$$2 \begin{pmatrix} 40 & 85 & 160 & 100 & 70 \end{pmatrix} \begin{pmatrix} 0.104 \\ 0.190 \\ 0.226 \\ 0.235 \\ 0.244 \end{pmatrix}$$

(A1)

194.298...

194 (km)

A1

Note: Award (M1)(A1)A0 for 97.1, from omitting the 2.

[3 marks]

Total [24 marks]

2. (a) $2.88 \left(\frac{72}{25} \right)$ **A2**

Note: Award **M1A0** for an attempt to calculate mean by hand, leading to an incorrect answer e.g values and frequencies must be seen substituted into formula.

[2 marks]

- (b) using Poisson distribution with their part (a) as m **(M1)**

$$P(X = 0)$$

$$= 0.05613476 \dots$$
 (A1)

$$(0.05613476 \dots \times 50)$$

$$= 2.81 \text{ (2.80673...)} \quad \textbf{A1}$$

[3 marks]

- (c) to ensure the expected frequency is greater than 5 **R1**

Note: Accept equivalent reasoning, provided it includes “expected” and “5”.

[1mark]

- (d) (i) H_0 : the data can be modelled by a Poisson distribution **A1**

Note: Award **A0** if they include mean = 2.88. However, in this specific case it is possible to award **A1** in part (iv).

[1 mark]

- (ii) $df = 3$ **A1**

[1 mark]

- (iii) Evidence of combining columns for observed frequencies **(M1)**

e.g. 9, 12, 11, 10, 8

$$p\text{-value} = 0.843 \text{ (0.842640...)} \quad \textbf{A2}$$

Note: If their $df = 4$, then award **FT** for a p-value of 0.934587...

If their $df = 2$, then award **FT** for a p -value of 0.660837...

Award **M1A1A0** for an answer of 0.84 (2 sf value) or 0.842 (truncated value)

[3 marks]*continued...*

Question 2 continued.

- (iv) since $p\text{-value} > 0.05$ **R1**
 Insufficient evidence to reject H_0 , (data can be modelled by a Poisson) **A1**
[2 marks]

Note: Award **R1** for correctly comparing their p -value to 0.05, and **A1** for a consistent conclusion.
 Do not award final **A1** if the null hypothesis is incorrect.
 Do not award **R0A1**.
 Condone “accept the null hypothesis” or “do not accept the alternative hypothesis”.

- (e) use of t -test **(M1)**
 $p\text{-value} = 0.0263$ (0.0262738...) **A1**
 since $p\text{-value} < 0.05$ **R1**
 sufficient evidence to reject H_0 /
 there is evidence of a decrease in the population mean **AG**

Note: Award **M0A0R0** for use of the z -test ($p\text{-value} = 0.0170858\dots$)
 Award a maximum of **M1A0R1** for a p -value of 0.026 (only 2 sf)

[3 marks]

- (f) (i) must be positive, **A1**
 since $\frac{dL}{dt} > 0$ (and $L_\infty - L > 0$) **R1**

Note: Do not award **A1R0**.
 Accept equivalent reasoning in words, e.g. the growth is positive.

[2 marks]

- (ii) 0 **A1**
 Since $L \rightarrow L_\infty$ **R1**

Note: Accept equivalent reasoning in words, e.g. the fish will stop growing.

[2 marks]

continued...

Question 2 continued.

(iii) **EITHER**

L_{∞} , since the length decreased

A2

OR

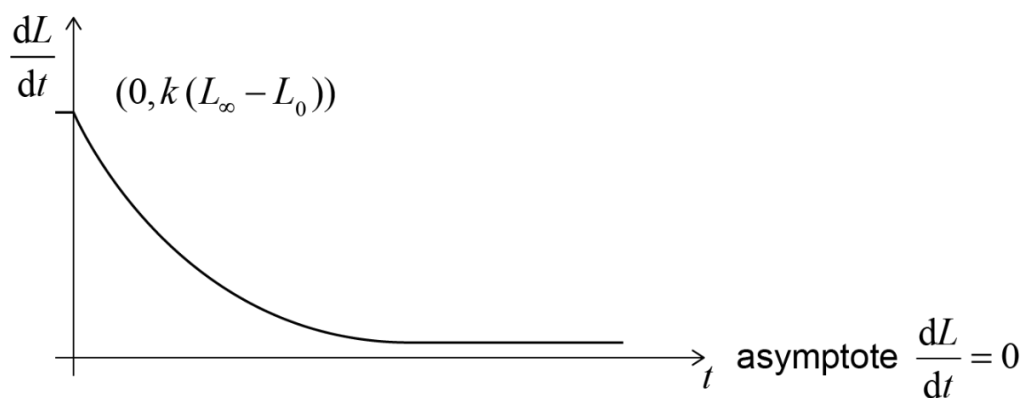
k , not growing as quickly, finite lifespan

A2

Note: Award **A1** for “ L , since the length decreased”

[2 marks]

(g)



Award **A1** for a concave-up curve which is positive and decreasing

Award **A1** for a $\frac{dL}{dt}$ -intercept at $k(L_{\infty} - L_0)$ (condone omission of t coordinate)

Award **A1** for a horizontal asymptote $\frac{dL}{dt} = 0$ (accept $y = 0$)

[3 marks]

continued...

Question 2 continued.

(h) attempt to separate variables

(M1)

$$\int \frac{1}{L_{\infty} - L} dL = \int k dt$$

$$-\ln|L_{\infty} - L| = kt + c$$

A1A1

Note: Award **A1** for correct integration of RHS including “+c” **A1** for correct integration of LHS.
Accept absence of absolute value, since $L_{\infty} - L > 0$.

attempt to remove ln using e from their integrated equation

(M1)

$$L_{\infty} - L = e^{-kt-c} = Ae^{-kt}$$

substitute initial conditions into their integrated equation, to find constant of integration (A1)

$$L_{\infty} - L_0 = e^{-c} (= A)$$

Note: The previous two marks may be seen in either order.

$$L = L_{\infty} - (L_{\infty} - L_0)e^{-kt}$$

A1

[6 marks]

Total [31 marks]

9. (a) Evidence of use of normal cdf with 45 **OR** sketch of bell curve with area above 45 labelled. (M1)

0.0338556... (A1)

recognition of binomial distribution with probability of their 0.0338556... (M1)

$X \sim B(180, 0.0338556...)$

$P(X \geq 10)$ **OR** $1 - P(X \leq 9)$

0.0869 (0.0869379...) A1
[4 marks]

- (b) defining event as being late and lateness marked correctly
 $0.0338556... \times (1 - 0.35) = 0.0220061...$ (A1)

Y is the number of occasions of being marked late

$Y \sim B(180, 0.0220061)$ **AND** $P(Y \geq 4)$ **OR** $P(Y > 3)$ **OR** $1 - P(Y \leq 3)$ (M1)

0.561 (0.561005...) A1
[3 marks]

[Total 7 marks]

3. (a) 4
- (b)

A1
[1 mark]

	A	B	C	D	E	F	G	H
Brett	1	4	2	3	5	6	7	8
Clarence	1	2	4	3	5	7	6	8

(A1)(A1)

Note: Award (A1) for each line in table

0.881 (0.880952...)

A2

Note: The ranks maybe seen within printed table.
Award A1 for 0.88 (2sf) answer.

For a common mistake:
Brett: 1 3 4 2 5 6 7 8 & Clarence correct
Award (A0)(A1)A2FT if 0.952 seen.

Note: The use of $r_s = 1 - \frac{6\sum d^2}{n(n-1)}$ i.e. $(r_s =) 1 - \frac{6(0^2 + 2^2 + 2^2 + 0^2 + 0^2 + 1^2 + 1^2 + 0^2)}{8(8^2 - 1)}$ may be seen which is not in the syllabus, but is a valid method because there are no ties. As they need the table to get the differences, there is no change to MS for this method.

[4 marks]

- (c) (since r_s is close to 1) there is a strong (positive) correlation between the **rankings** of the sheepdogs by Brett and the **rankings** of the sheepdogs by Clarence
(since r_s is close to 1) there is a strong (positive) association between Brett's **rankings** and Clarence's **rankings**.
since r_s is not equal to 1, Brett and Clarence do not always agree on the **rankings** of the sheepdogs.

A1

Note: For A1 to be awarded, a numerical value must be seen in part (b), and the conclusion consistent with that numerical value.
"Strong positive" is insufficient because it is not in terms of the **ranks**.
"Similar" in isolation is insufficient, however if they state "the order in which the sheepdogs were **ranked** by the judges is similar" would be awarded A1.
The names of judges do not need to be seen for the A mark.

[1 mark]

Total [6 marks]

6. (a) attempt to use formula $S_{n-1}^2 = \frac{n}{n-1} S_n^2$ (M1)

$$= \frac{5}{4} \times 3.2^2$$

$$= 12.8 \text{ (}^\circ\text{C}^2\text{)}$$

A1

[2 marks]

(b) indication of use of t -interval (M1)

[16.1267, 24.0733] from use of 3.2

from GDC [15.7, 24.5] ([15.6577..., 24.5423...]) A1

Note: Do not penalize use of round brackets, strong inequalities or using terms such as “upper” and “lower” to describe the endpoints of the interval.

[2 marks]

(c) 17 is within the confidence interval, so it is a plausible mean R1

[1 mark]

Total [5 marks]

13. let X be the number of worms in 1 square metre then $X \sim \text{Po}(1.2)$

(a) $P(X = 1) = 0.361$ (0.361433...)

A1
[1 mark]

(b) $P(X \geq 1) = 0.699$ (0.698805...)

A1
[1 mark]

(c) if Y is the number of worms in 5 square metres then $Y \sim \text{Po}(6)$

(A1)

$P(Y = 5) = 0.161$ (0.160623...)

A1
[2 marks]

(d) **EITHER**

attempting to find an expression in terms of $P(X = 1)$

(M1)

$P(X = 1)^5 = 0.00617$ (0.00616792...)

A1

OR

let V be the number of square metres containing exactly one worm

attempting to use the binomial with $P(X = 1)$

(M1)

then $V \sim B(5, P(X = 1))$

$P(V = 5) = 0.00617$ (0.00616792...)

A1

Note: Accept 0.00613 from 3sf **FT** from part (a).

[2 marks]

(e) let W be the number of square metres containing at least one worm.

attempting to use binomial

(M1)

then $W \sim B(5, P(X \geq 1))$

(A1)

$P(W = 3) = 0.310$ (0.309572...)

A1

Note: Accept 0.31 for final **A1**. Accept 0.309 from 3sf **FT** from part (b).

[3 marks]

Total [9 marks]

14. use the second line since it has a higher R^2

A1R1

Note: Do not award **R0A1**. The **R1** can be awarded independently within the question.

$$\ln w = 0.302 \ln d + 0.693$$

A1

attempt to exponentiate both sides

M1

$$w = e^{0.693} \times d^{0.302}$$

A1

$$w = 2.00d^{0.302} \quad \left(= 1.99970\dots d^{0.302} \right)$$

A1

Note: Accept $w = 2d^{0.302}$ for final **A1**.

Award **A0R0A1M1A1A0** for use of the first line

Total [6 marks]

Q3.

(a) H_1 : satisfaction level and the type of service interaction are not independent. **A1**

[1 mark]

(b) 4 **A1**

[1 mark]

(c) 3.32 (3.31906...) **A2**

Note: Award **A1A0** for a 2 sf answer of 3.3 as recognition of correct method.

[2 marks]

(d) (i) $3.319106... < 9.488$ OR $0.506 (0.505917...) > 0.05$ **R1**

No, he isn't correct (insufficient evidence to reject H_0) **A1**

Note: Accept a comparison in words (e.g. $\chi^2_{calc} < \chi^2_{crit}$ or ' p -value < sig level') for the **R** mark only if their χ^2_{calc} value or p -value is seen. Do not award **A1R0**.

(ii) insufficient evidence to say "satisfaction level and the type of service interaction are not independent". **A1**

Note: Answer to part d (ii) follows through from their answer to part d(i).

[3 marks]**Total [7 marks]**

Q4.

- (a)

4

A1
[1 mark]
(A1)
- (b)

identification of lower quartile as 24.28
24.70 – 24.28

0.42 (seconds)

A1
[2 marks]
- (c)

(i)

attempt to use correct formula to find boundaries for outliers
24.70 + 1.5 × 0.42 OR 24.28 – 1.5 × 0.42

25.33 and 23.65

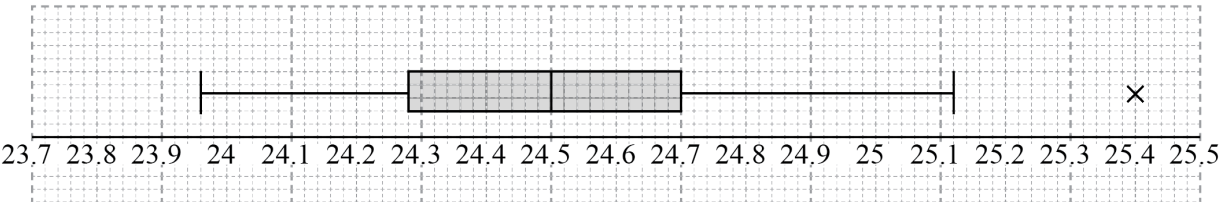
only 25.4 is an outlier

(M1)

A1

A1

(ii)



A1

Note: Award the **A1** only if both whiskers and the outlier are correctly marked.
Accept symbols other than a cross to represent the outlier.
The whiskers should not be drawn through the box.

[4 marks]
Total [7 marks]

Q6

- (a) use of formula for sum of an arithmetic sequence or first few terms seen **(M1)**

$$\frac{10}{2}(2 \times 6 + 9 \times 3) \text{ OR } 6 + 9 + 12 + 15 + \dots$$

sum of angles is 195° **(A1)**

$$p = \frac{195}{360} \left(0.542, \frac{13}{24}, 0.541666\dots \right) \quad \mathbf{A1}$$

[3 marks]

- (b) recognition of binomial distribution **(M1)**

$$n = 4, \quad p = \frac{195}{360} \quad \text{OR} \quad X \sim B\left(4, \frac{195}{360}\right) \quad \mathbf{(A1)}$$

let X be the number of times the arrow lands in white

$$P(X \geq 3) = 0.377 \quad (0.377450\dots) \quad \mathbf{A1}$$

Note: Accept 0.378 (0.377988...) from use of 0.542 from part (a).

[3 marks]

Total [6 marks]

Q7.

(a) stratified (sampling)

A1
[1 mark](b) (i) $r = 0.854$ (0.853899...)**A2****Note:** Award **(M1)A0** for a truncated value $r = 0.853$ or incorrectly rounded value eg. $r = 0.85$.(ii) it is valid (since $r > 0.8$)**A1****Note:** Follow through from their value of r , for the final **A1**.**[3 marks]**(c) $y = 0.266x + 2.40$ ($y = 0.265851...x + 2.39573...$)**A1****Note:** Accept 2.4.

(d) substitution of 17.2 into their regression line

[1 mark]
(M1)

$$y = 0.265851... \times 17.2 + 2.39573...$$

$$= 6.97(6.96837...)$$

A1**Note:** Accept $6.98(6.9752)$ from use of 3sf answers from part (c).**[2 marks]****Total [7 marks]**

- Q12.** (a) (i) rejecting the null hypothesis when it is true **A1**
OR
 rejecting the mean as 12.4 / min (124 / 10 mins) when it is true **A1**
- (ii) let X be the number of cars that arrive in 10 minutes
 attempt to find $P(X \geq 131)$ **(M1)**
 use of $Po(124)$ **(M1)**
 $P(X \geq 131) = 0.276$ (0.276397...) **A1**
[4 marks]
- (b) any valid reason **R1**
 for example:
 the arrival of cars are not independent (there might be a traffic light shortly before the intersection)
 the rate of arrival is not uniform (some times of day might be busier than others)

Note: Do not accept “mean and variance are not the same”.
--

[1 mark]**Total [5 marks]**

Q15. (a) $\begin{pmatrix} p-1 & 1-q \\ 1-p & q-1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ **(M1)**

$(p-1)x + (1-q)y = 0$ (or equivalent) **(A1)**

$\begin{pmatrix} 1-q \\ 1-p \end{pmatrix}$ (or any multiple) **A1**

[3 marks]

(b) $\frac{1-q}{2-p-q}$ **A1A1**

Note: Award **A1** for numerator, **A1** for denominator.
For an answer of $1-q$ only, award **A0A0**.

[2 marks]

(c) $\frac{4}{11}(0.364, 0.363636\dots)$ **A1**

[1 mark]

Total [6 marks]

1. (a) ($k =$) 5**A1**
[1 mark]

- (b) (i) attempt to substitute into the mean formula with a sum of at least two correct products **and** equating to 2.2 (or $2.2(p+27)$ if denominator not seen) or equating to “mean”. Equation may be seen in part b(ii). **(M1)**

$$2.2 = \frac{0 \times 3 + 1 \times 11 + 2 \times 7 + 3 \times 5 + 4 \times p + 5 \times 1}{3 + 11 + 7 + 5 + p + 1} \quad \text{OR} \quad 2.2 = \frac{45 + 4p}{p + 27}$$

A1(ii) ($p =$) 8**A1**

Note: Follow through from part (b)(i) only if their value of p is a whole number.

[3 marks]

- (c) finding the median ($= 2$), quartiles ($Q_1 = 1$ and $Q_3 = 4$), or interquartile range of the data $IQR = 3$ from the 2024 tournament **(A1)**

Note: Award **(A1)** for any correctly identified/labelled value referring to 2024 data as listed in the question. Accept “Range from 0 to 5” without reference to 2024.

Any two of the following statements:

R1R1The median number of goals per game **in both years** is 2.**In both years**, (the lower quartile is 1 and the upper quartile is 4) the IQR is 3.**Both** distributions are (roughly) skewed to the right

Note: Award **R1** for each correct statement. Each statement must reference values from both the box and whisker diagram **and** the data from the 2024 tournament to award **R1**. (e.g. “the median is 2” earns **R0**, “the median for both years is 2” earns **R1**.) If an incorrect statement is made, award at most **R1R0**.

[3 marks]

(d) B and C **A1A1**

Note: Award **A1** for each correct event, if two events are selected.
Award **A1A0** for two correct events and one additional incorrect event selected.
Award **A0A1** for an answer of “ A , D , and E ” from misreading the meaning of F' .
Award **A0A0** for any other combination of events selected.

[2 marks]

(e) $\frac{5}{17} \left(\frac{10}{34} \right)$

A1A1

Note: Award **A1** for correct numerator, **A1** for correct denominator. Award at most **A1A0** for a decimal answer (0.294117...) if no fraction is shown. Do not award **A1ft** for their denominator if their answer to part (a) or part (b)(ii) is not a whole number.

[2 marks]

(f) $\frac{1}{35} \times \frac{3}{34}$

(A1)(A1)(M1)

Note: Award **A1** for each correct probability, **M1** for multiplying their two probabilities. Do not award any marks if their answer to part (a) or part (b)(ii) is not a whole number.

0.00252 (0.00252100..., $\frac{3}{1190}$, 0.252%)

A1**[4 marks]****Total [15 marks]**

5. (a) 186 A1
[1 mark]
- (b) using midpoints within formula for mean M1

$$\frac{11 \times 103 + 14 \times 52 + 17 \times 31}{186}$$
12.8 (12.8387...) (mm) A1
[2 marks]
- (c) $\frac{42}{86}$ or $\frac{21}{43}$ seen (0.488372...) (A1)

$$\left(\frac{21}{43}\right)^2$$
method mark for squaring their probability (M1)

$$= 0.239 \text{ (0.238507...) } \left(= \frac{441}{1849} \right)$$
 A1
[3 marks]
- (d) choosing a chi-squared test for independence (M1)
 p -value = 0.127 (0.126687...) A1
 p -value > 0.05 therefore do not reject H_0 R1
- Note:** Accept p -value > significance level, as this is stated in the question.
- there is no significant evidence that length category and sex are not independent A1
- Note:** Do not award **R0A1**. Accept length category and sex are independent. Allow internal follow through in this part for a consistent comparison and conclusion from their p -value.
- [4 marks]
- (e) $H_0 : \phi = 0.45$; $H_1 : \phi > 0.45$ A1
choosing a binomial test (M1)
- Note:** can be implied by mention of Binomial, calculator notation, recognisable p -value
- under the null hypothesis, if X is the number of males then $X \sim B(186, 0.45)$
- p -value = $P(X \geq 100)$ (M1)
= 0.0101 (0.010103...) A1
since p -value < 0.05 we reject H_0
there is significant evidence that more than 45% of cane beetles are male. A1
- Note:** Do not award the final **A1** unless the correct p -value is seen.
- [5 marks]
- Total [15 marks]**

7. (a) (i) 0.4 A1
(ii) 0.6 A1
[2 marks]
- (b) (i) $0.4^4 = 0.0256$ (M1)A1
(ii) considering $M^4 = \begin{pmatrix} 0.5422 & 0.5341 \\ 0.4578 & 0.4659 \end{pmatrix}$ (M1)(A1)
required probability is therefore 0.542 A1
[5 marks]
- (c) attempting Mv with given eigenvector M1

$$\begin{pmatrix} 0.4 & p \\ 0.6 & 1-p \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
correct matrix multiplication

$$\begin{pmatrix} 0.4 - p \\ -0.4 + p \end{pmatrix}$$
 A1
writing as a multiple of given vector

$$= (0.4 - p) \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
 (A1)
therefore the eigenvalue is $0.4 - p$ A1
[4 marks]
- (d) consider eigenvector with eigenvalue 1 $Mv = v$ (M1)

$$\begin{pmatrix} 0.4 & p \\ 0.6 & 1-p \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

turn into equations:
 $0.4x + py = x$ OR $0.6x + (1-p)y = y$ A1
also use $x + y = 1$ A1
then $x + \frac{0.6x}{p} = 1$

$$x = \frac{p}{p + 0.6}$$
 A1
[4 marks]
- (e) graphical, inequality or equality approach to solve $\frac{p}{p + 0.6} \geq 0.6$ (M1)
 $p = 0.9$ A1
[2 marks]
Total [17 marks]

- Q6.** (a) (i) 0.45 **A1**
 (ii) attempt to find the expected value (at least two correct products) **(M1)**
 $0 \times 0.21 + 1 \times 0.34 + 2 \times 0.28 + 3 \times 0.17$
 1.41 (customers) **A1**

[3 marks]

- (b) $E(T) = 115$ **(A1)**
 1.41×115
 162 (162.15) (seconds) (2 minutes 42 seconds) **A1**

[2 marks]

- (c) $P(T > 180)$ **(M1)**
 $0.0101(0.0101318\dots)$ **A1**

Note: Award **M1A0** for an answer of 0.467.

[2 marks]

- (d) attempt at linear combination of two normal variables **(M1)**
 time to serve two people, $T_2 \sim N(230, 2 \times 28^2)$ **(A1)(A1)**
 $P(T_2 > 180) = 0.897(0.896649\dots)$ **A1**

[4 marks]

- (e) (i) (if no-one in the queue the waiting time is 0)
 $(0.21 \times 0) + 0.34 \times 0.0101318\dots + 0.28 \times 0.896649 + 0.17 \times 1$ **(M1)(A1)**
 $0.425(0.424506\dots)$ **A1**
 (ii) $(0.425 > 0.4)$ employ more staff **A1**

Note: Award if consistent with their probability from part (e) (i).

[4 marks]

[Total 15 marks]

1. (a)

$$(T =) \begin{pmatrix} 0.91 & 0.04 & 0.06 \\ 0.05 & 0.95 & 0.02 \\ 0.04 & 0.01 & 0.92 \end{pmatrix}$$

A1A1A1

Note: Award **A1** for each correct column.

[3 marks](b) recognizing the need to calculate T^5 **(M1)**

$$(T^5 =) \begin{pmatrix} 0.659 & 0.156 & 0.219 \\ 0.196 & 0.792 & 0.101 \\ 0.146 & 0.0514 & 0.680 \end{pmatrix} \left(\begin{pmatrix} 0.658687... & 0.156128... & 0.219028... \\ 0.195559... & 0.792469... & 0.101208... \\ 0.145753... & 0.0514021... & 0.679762... \end{pmatrix} \right) \quad \textbf{(A1)}$$

selects relevant cell (e.g. 0.219 seen separately)
21.9%

(M1)**A1**

Note: Final answer must be a percentage. Follow through from their T , including the transpose.

[4 marks](c) attempt to find steady state vector (e.g. raising T to a high power)**(M1)**

$$\begin{pmatrix} 0.342 \\ 0.432 \\ 0.225 \end{pmatrix} \left(\begin{pmatrix} 0.342342... \\ 0.432432... \\ 0.225225... \end{pmatrix} \right)$$

A1**[2 marks]**

Note: Award **M1A0** if they express steady state vector as

$$\begin{pmatrix} 0.342 & 0.342 & 0.342 \\ 0.432 & 0.432 & 0.432 \\ 0.225 & 0.225 & 0.225 \end{pmatrix}$$

(d) 34.2%

A1**[1 mark]**

- (e) require new steady state vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ such that $a \geq 0.40$ (seen anywhere) **(R1)**

any correct manipulation of Clover's entries in T (e.g. complementary increase and decrease of $T_{1,3}$ and $T_{3,3}$) **(A1)**

Note: This **A1** can be awarded independently of the **R1**.

$$(T \Rightarrow) \begin{pmatrix} 0.91 & 0.04 & 0.08 \\ 0.05 & 0.95 & 0.02 \\ 0.04 & 0.01 & 0.90 \end{pmatrix}$$

$$\begin{pmatrix} 0.403 \\ 0.459 \\ 0.138 \end{pmatrix}$$

when $C \rightarrow A$ is 13%

A1

Note: If candidates go on to state $13\% - 6\% = 7\%$, then ignore this and award **A1**. If their answer is just '7%' then award **A0**.

[3 marks]

- (f) since $0.459 > 0.432$
therefore Bodega will benefit

R1

A1

Note: Do not award **R0A1**.

[2 marks]

- (g) Accept any appropriate reason that shows engagement with the scenario.
e.g.
(Convergence to steady state vector takes a relatively long time and ...)

...during which time either of the three stores may notice a drop/influx of customers and change their advertising accordingly, thus changing T .

A1

...during this time, a fourth (etc.) store could open, impacting the model.

A1

...during this time, customers may move in and out of the region in a non-consistent rate. **A1**

Note: 'Customers may change their minds' or similar, award **A0**.

[1 mark]

- (h) attempt to apply $u_n = u_1 + (n-1)d$
 $= 600 + 9 \times 30$
 $= 870$

(M1)

A1

[2 marks]

(i) 720 seen anywhere A1
 $720 + 1.12 \times 30$ A1
 $= 753.6$ AG
[2 marks]

(j) attempts to sum geometric series with $r = 1.12$ (M1)
 $720 + 30(1.12 + 1.12^2 + \dots + 1.12^5)$ (A1)
 $= 933$ (933.455...) A1
[3 marks]

(k) **METHOD 1 (AS formula from booklet and adapted general formula from part (j))**

general formula for AS: $600 + 30(n-1)$ (A1)

attempt to adapt GS formula (from part (j)) to include “ $n-5$ ” (M1)

general formula for GS: $720 + 30 \left[\frac{1.12(1.12^{n-5} - 1)}{1.12 - 1} \right]$ (A1)

EITHER

attempts to solve $GS - AS = 500$ (M1)

$n = 17.5052\dots$ A1

Note: Other correct n values, depending on original expressions, are possible but only differing by an integer. **FT** within the question.

OR

creates list of both sequences

1080, 1530.87... **AND** 1110, 1661.77... (M1)

A1

Note: If list method is used, award **A1** for the cross-over values clearly stated.

THEN

$\Rightarrow$ week 18

A1

Note: Award at most **(A1)(M1)(A1)(M0)A0A1** for a correct answer with no working.

[6 marks]

– 10 –

2225 – 7308M

METHOD 2 (general formula from part (j) and adapted AS formula)

general formula for GS: $720 + 30 \left[\frac{1.12(1.12^n - 1)}{1.12 - 1} \right]$ **(A1)**

attempt to express AS formula for weeks after fifth week **(M1)**

general formula for AS: $720 + 30n$ **(A1)**

Note: The n values can be offset (e.g $n-1$ or $n-5$ etc.) provided the approach is consistent between GS and AS.

EITHER

attempts to solve GS – AS = 500 **(M1)**

$n = 12.5052...$ **A1**

Note: Other correct n values, depending on original expressions, are possible but only differing by an integer. **FT** within the question.

OR

creates list of both sequences **(M1)**

1080, 1530.87... **AND** 1110, 1661.77... **A1**

Note: If list method is used, award **A1** for the cross-over values clearly stated.

THEN

$\Rightarrow$ week 18 **A1**

Note: Award at most **(A1)(M1)(A1)(M0)A0A1** for a correct answer with no working.

[6 marks]**[Total: 29 marks]**

2. (a) $(P(X > 55) =) 0.208737\dots$ (A2)
 recognizing Binomial distribution (M1)
 $(P(Y \geq 5) =) 0.621 (0.621150\dots)$ A1

Note: Award **(A2)(M1)A0** for an answer of 0.622 (from use of 0.209).

[4 marks]

- (b) (variance $=) \frac{3.7^2}{25}$ OR (SD $=) \frac{3.7}{5}$ (A1)
 $\bar{X} \sim N\left(52.0, \frac{3.7^2}{25}\right)$

$$P(51 < \bar{X} < 53) = 0.823 (0.823417\dots)$$

A2
[3 marks]

- (c) $H_0 : \mu = 52.0$ A1
 $H_1 : \mu > 52.0$ A1

Note: Accept hypotheses in words if they are clearly expressed and 'population mean' is referred to. Do not accept $H_0 : \mu = \mu_0$ unless μ_0 is explicitly defined to be 52.0.

$$p\text{-value} = 0.105055\dots$$

A2

$$0.105 > 0.05$$

R1

therefore there is insufficient evidence to reject H_0

A1

(i.e. the population mean weight of eggs is 52.0 g)

Note: Follow through within the question for the **R1A1**, from their p -value seen. Award at most **A0A0A2R1A0** if hypotheses are omitted or incorrect. Award the final **A1** only if the null hypothesis is also correct and the conclusion is consistent with both the direction of the inequality and the alternative hypothesis.

[6 marks]

- (d) **EITHER**
 52.66 ± 1.5888 A1A1

Note: Award **A1** for each correct value. Award at most **A1A0** if the $\pm$ is incorrect or omitted.

OR

$$(51.1, 54.2) ((51.0712\dots, 54.2488\dots))$$

A1A1

Note: Award **A1** for each correct end point. Award at most **A1A0** if interval notation is incorrect or omitted.
 Allow $51.1 \leq \mu \leq 54.2$

[2 marks]

- (e) (i) paired t -test **A1**
[1 mark]
- (ii) **EITHER**
That the differences in weights are normally distributed. **A1**
OR
The weights of eggs from any two chickens are independent. **A1**
[1 mark]
- (f) $H_0 : \mu_d = 0$ **OR** $H_0 : \mu_2 - \mu_1 = 0$ **A1**
 $H_1 : \mu_d > 0$ **OR** $H_1 : \mu_2 - \mu_1 > 0$ **A1**
- p -value = 0.03377479... **A2**
 $0.0338 < 0.05$ **R1**
 (there is sufficient evidence that) the mean weight of the chickens' eggs has increased following the new feed **A1**
- Note:** The conclusion must be in context for the final **A1** to be awarded. Follow through within the question for the **R1A1**, from their p -value seen. Award at most **A0A0A2R1A0** if hypotheses are omitted or incorrect. Do not follow through from an incorrect part (e)(i).
- [6 marks]**
- (g) 2.39... (2.391555...) **A1**
[1 mark]
- (h) an attempt to find s_R^2 using the given expression and their part (g) **(M1)**
 $s_R^2 = 2.391555... \times 3.1789$
 $= 7.60$ **A1**
[2 marks]
- [Total: 26 marks]**
-

1. (a) (i) any sensible comment, eg:
 the sample will be more reliable / (likely to be) more representative / valid
 the sample mean is likely to be closer to the population mean

R1**[1 mark]**

Note: Do not accept “The sample will be (more) accurate, precise or exact”, or anything referring to outliers or SD only.

- (ii) more time consuming / expensive / more open to human error

R1**[1 mark]**

Note: Do not accept anything referring to outliers or SD only

- (b) 7.60 (7.597537...)

(M1)A1

Note: M1 for 7.60 seen anywhere, **A1** for correct final answer.

[2 marks]

- (c) any sensible comment, eg:

R1

the population standard deviation may be different from the sample

these s_{n-1} values are only estimates / approximations and the values are too close together.

standard deviation is not the only measure of spread

Note: Award **R0** if part (b) is not in range 7-7.65

[1 mark]

- (d) (i) (the population variances) are the same.

A1

Note: Must say ‘the same/equal’. ‘Similar’ or ‘close to each other’ is not correct.

[1 mark]

- (ii) any attempt to compare sample standard deviations or sample variances

M1

the sample variances/sample standard deviation/ s_{n-1} are similar, so

assumption is plausible. (Therefore Augustin should use a pooled t-test.)

R1

Note: Award **R0** if part (b) is not in range 7-8

[2 marks]*continued...*

Question 1 continued.

(e) (i) $H_0 : \mu_A = \mu_B$

A1

$H_1 : \mu_A < \mu_B$

A1

Note: Award **A1** for a correct null hypothesis and **A1** for choosing the correct tail for *their* alternative hypothesis. Accept correct hypotheses in words, but only if “population” is clear.

Award at most **A1A0** if μ_1 or μ_2 used without definition.

[2 marks]

(ii) 0.0445 (0.0444586...)

A2

Note: Accept anything in the range 0.044–0.045 due to rounding issues.

[2 marks]

(iii) $0.0445 < 0.05$

R1

there is (significant) evidence that the (population) mean of school B is higher than the (population) mean of school A **OR** Augustin’s belief is supported

A1

Note: Do not award **R0A1**. Do not award the final **A1** if the null hypothesis is incorrect or the conclusion is not in context.

[2 marks]

(f) (i) $H_0 : \rho = 0; H_1 : \rho \neq 0$

A1A1

Note: Hypotheses must involve ρ . Hypotheses in terms of r are not correct. In words ‘population correlation’ is equally valid.

$r > 0.576$ **OR** $0.876 > 0.576$

R1

Note: Accept $|r| > 0.576$

therefore reject H_0 , there is significant evidence of correlation between entry exam and final exam results.

A1

Note: Do not award **R0A1**. Do not award final **A1** for just ‘reject H_0 ’, there must be some interpretation.

Condone ‘there is evidence of correlation’. Of course, the final two marks can still be given if hypotheses were in terms of r .

[4 marks]

(ii) could use Spearman Rank.

A1

[1 mark]

continued...

Question 1 continued.

- (g) as entry exam results goes up by one, final exam result goes up by 0.37

A1

Note: Statement should clearly be in context. Anything vague about the gradient of the regression line is not acceptable.

[1 mark]

- (h) (i) **METHOD 1**

the exam board model predicts $\hat{y} = 0.37 \times 38.7 + 37.6 = 51.919$

M1

therefore the improvement is $52.5 - 51.919 = 0.581$

A1

so the answer is 0.6 to 1 decimal place.

AG

[2 marks]

- (ii) $H_0 : \mu_A = \mu_B; H_1 : \mu_A > \mu_B$

A1

so the p value is 0.0213 (0.0212715...)

A1

Note: First two **A1** marks are independent. Hypotheses must clearly be in terms of population means. Expressions involving $\mu_{(y-\bar{y})A}$ or μ_1 are acceptable.

since $p < 0.05$

R1

therefore reject H_0 , evidence of higher school value added in School A

A1

Note: 'Reject H_0 ' alone is not enough. Answer must be in context.

[4 marks]

- (i) school A outperforms school B is supported by the test in part (g)/better added value

A1

school B outperforms school A is supported by the test in part (e)/higher mean

A1

[2 marks]

Total [28 marks]

1. (a) **EITHER**
 the data points come from normal populations **A1**
OR
 the variances (standard deviations) of the two groups are (approximately) equal. **A1**
OR
 data points of each group are independent. **A1**
OR
 the sample is representative of the population **A1**
[1 mark]
- (b) (i) $\mu_s = \mu_p$ **OR** $\mu_1 = \mu_2$ **A1**
 (no difference in the mean (population) distance travelled using premium/standard fuel)
 (ii) $\mu_s < \mu_p$ **OR** $\mu_1 < \mu_2$ **A1**
 (mean (population) distance travelled using premium fuel is greater than using standard fuel)
[2 marks]
- (c) (i) 0.0264 (0.0263584 ...) **A2**

Note: Their p -value must be consistent with *their* H_1 .

[2 marks]

(ii) $0.0263584 \dots < 0.05$

R1

Note: Accept " $p < 0.05$ " or " p -value is less than the significance level", provided their p -value is shown in part (c)(i).

EITHER

(there is sufficient evidence that) the car can travel further using premium fuel. **A1**

OR

(there is sufficient evidence that) the car uses less premium fuel compared to standard fuel. **A1**

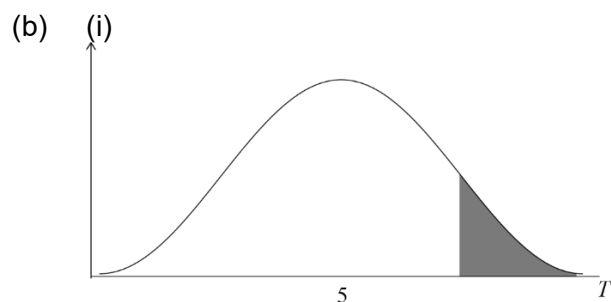
Note: Do not award **R0A1**. The conclusion must be in context to award **A1**; "reject H_0 " earns **A0**. Their conclusion must be consistent with *their* p -value and *their* hypothesis.

[2 marks]

Total [7 marks]

3. (a) 0.159 (0.158655...)

A2
[2 marks]



A1

Note: Accept any shaded region as shown provided it is to the right of 5.

(ii) Recognition to use the inverse normal function on GDC

(M1)

Note: Award **(M1)** for $P(X > k) = 0.15$ **OR** $P(X < k) = 0.85$ **OR** inverse normal (invNorm) seen.

6.04 (seconds) (6.03643...)

A1

[3 marks]
Total [5 marks]

9. (a) (for a 30 minute period,) mean = 17 (A1)
 recognize lower bound is 23 **OR** $P(\text{number of calls} > 22)$ (M1)
 0.0953 (0.0952720...) A1
[3 marks]
- (b) recognition of binomial distribution (M1)
 with $n = 10$ and $p = 0.0953\dots$ OR $p = 0.904\dots$ (A1)
 0.754 (0.754360...) A1
[3 marks]
Total [6 marks]

16. (a) $(q =) 0.3$

A1
[1 mark]

(b) 0.7

A1
[1 mark]

(c) 0.2×0.2 **OR** 0.1×0.1 **OR** 0.7×0.4
recognize 3 possible ways BB or CC or TT
 $0.2 \times 0.2 + 0.1 \times 0.1 + 0.7 \times 0.4$
0.33

(A1)
(M1)
A1
[3 marks]

(d) $\begin{pmatrix} 0.2 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.3 \\ 0.7 & 0.5 & 0.4 \end{pmatrix}$

A1
[1 mark]

(e) **EITHER**
recognize need to find large power of transition matrix (such that a steady state is found)

(M1)

OR
attempt to find eigenvector with eigenvalue 1

(M1)

e.g. $\begin{pmatrix} 0.2 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.3 \\ 0.7 & 0.5 & 0.4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ **OR** $0.2x + 0.4y + 0.3z = x$

eigenvector = $\begin{pmatrix} \frac{39}{68} \\ \frac{27}{68} \\ \frac{68}{68} \\ 1 \end{pmatrix}$

THEN
0.201 (0.201492...)

A1
[2 marks]
Total [8 marks]

1. (a) (i) 4 A1
 (ii) 6 A1
 (iii) 3 A1
[3 marks]

- (b) $(1.5 \times 3) + 6$
 $= 10.5$ (A1)
 $(k =) 11$ A1
[2 marks]

- (c) at least one correct mid-interval value identified (A1)

Note: Mid-interval value may be seen in mean formula.

4.54 $(4.53650\dots, \frac{5716}{1260})$ A2
[3 marks]

- (d) the frequency of the number of letters (in each word in his essay) is **not** consistent with the English language. A1
[1 mark]

- (e) (i) 6 A1
 (ii) 20.3 (20.3131...) A2
 (iii) 0.00244 (0.00243537... 2.44×10^{-3}) A1
[4 marks]

- (f) **EITHER**
 $20.3 > 16.812$ R1
OR
 $0.00244 < 0.01$ R1
THEN
 Dr. Petrillo should reject the null hypothesis A1

Note: Do not award **R0A1**.

[2 marks]
Total [15 marks]

5. (a) convenience

A1

Note: Accept opportunity.

[1 mark]

(b) (i) $p > 0.6$

A1

(ii) recognition to use binomial distribution (with $n = 80$ and $p = 0.6$)

(M1)

recognize need to find $P(X \geq 54)$

(A1)

 p -value = 0.104 (0.103731 ...)

A1

since $0.104 > 0.05$,

R1

insufficient evidence to reject H_0

insufficient evidence that the new recipe is an improvement.

A1

Note: FT from their p -value for the final R1A1. Award A1 only if reference is specifically made to the context. Do not accept "the new recipe is not an improvement". Do not award R0A1.

[6 marks]

(c) (i)

m	$P(X \geq m)$
55	0.067464...
56	0.041747...

a region with a lower bound of 56 e.g. $[56, 80]$ OR $X \geq 56$

A2

(ii) attempt to find $P(X \leq 55)$ when $p = 0.65$

M1

0.793 (0.792833...)

A1

[4 marks]

(d) H_0 : buying snacks is independent of age H_1 : buying snacks is not independent of age

A1

Note: Award A1 for both hypotheses correct.

recognize need to combine cells as $fe < 5$

(M1)

new table is

Age, x	Will buy	Will not buy
$x < 20$	22	8
$20 \leq x < 30$	18	5
$30 \leq x$	14	13

(A1)

 p -value = 0.0957 (0.0956988 ...)

A2

 $0.0957 > 0.05$ insufficient evidence to reject H_0

A1

Note: Award at most A1M0A0A2(FT)A1(FT), if a candidate does not combine cells, obtaining a p -value of 0.123596 ...
Award A1 for a p -value = 0.096 (2 s.f.) or 0.0956 (truncated).
FT from their p -value for the final A1, providing H_0 is stated. Award A1 for "Accept H_0 ".

[6 marks]

Total [17 marks]

