OCR
Oxford Cambridge and RSA

Friday 13 June 2025 – Afternoon

A Level Further Mathematics B (MEI)

Y422/01 Statistics Major

Time allowed: 2 hours 15 minutes

187041866O

You must have:
- the Printed Answer Booklet
- the Formulae Booklet for Further Mathematics B (MEI)
- a scientific or graphical calculator

QP

INSTRUCTIONS
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Fill in the boxes on the front of the Printed Answer Booklet.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give your final answers to a degree of accuracy that is appropriate to the context.
- Do **not** send this Question Paper for marking. Keep it in the centre or recycle it.

INFORMATION
- The total mark for this paper is **120**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **12** pages.

ADVICE
- Read each question carefully before you start your answer.

© OCR 2025 [H/508/5594]
DC (ST/FC) 351650/5

OCR is an exempt Charity

stemup.co.uk

Section A (34 marks)

1 A fair eight sided dice has its sides labelled 1, 2, …, 8. The random variable X represents the score when the dice is rolled once.

(a) Find $P(X > 4)$. **[1]**

(b) Find each of the following.

- $E(X)$
- $Var(X)$ **[3]**

The dice is rolled 3 times.

(c) Determine the exact probability that one of the scores is less than 4 and the other two are greater than 4. **[3]**

2 At a toy factory, wooden blocks of approximate heights 20 mm, 30 mm and 50 mm are made in red, yellow and green respectively. The heights of the blocks in mm are modelled by independent random variables which are Normally distributed with means and standard deviations as shown in the table.

Colour	Mean	Standard deviation
Red	20	0.8
Yellow	30	0.9
Green	50	1.2

In parts **(a)**, **(b)** and **(c)**, the blocks are selected randomly and independently of one another.

(a) Find the probability that the height of a red block is less than 19 mm. **[1]**

(b) A tower is made of 15 blocks stacked on top of each other consisting of 5 red blocks, 5 yellow blocks and 5 green blocks.

Determine the probability that the tower is at least 495 mm high. **[4]**

(c) Determine the probability that a tower made of 3 red blocks will be at least 1 mm higher than a tower made of 2 yellow blocks. **[3]**

© OCR 2025

stemup.co.uk

3 Potatoes are sold in sacks of nominal weight 25 kg. As part of a trading standards check, a random sample of 50 sacks from a large batch is selected and software is used to produce a 95% confidence interval for the mean weight of potatoes per sack. An extract from the software is shown below.

Sample Mean	24.878
Standard Deviation	0.5664
Standard Error	0.0801
Sample Size	50
Confidence Level	0.95
Interval	24.878 ± 0.157

(a) Show how the standard error was calculated. **[1]**

(b) State the confidence interval in the form $a < \mu < b$. **[1]**

(c) Explain whether the confidence interval suggests that the mean weight of potatoes per sack is different from 25 kg. **[1]**

(d) Determine whether your conclusion to part **(c)** would have been different if the confidence level had been 90%. **[4]**

4 Gabi is practising basketball. On each attempt to score a basket, the probability that Gabi scores is modelled as 0.4. It is assumed that each attempt is independent of previous attempts. The number of attempts which Gabi takes in order to score a basket for the first time is denoted by the random variable X.

(a) **(i)** Calculate $P(X = 5)$. **[1]**

(ii) Calculate $P(X > 5)$. **[1]**

(b) Determine the probability that X is within one standard deviation of its mean. **[5]**

(c) Determine the probability that Gabi scores at least 5 baskets in 20 attempts. **[2]**

(d) Determine the probability that Gabi scores the 5th basket on their 20th attempt. **[2]**

(e) Comment on the validity of the assumption that on each attempt to score a basket, the probability that Gabi scores is 0.4, independent of previous attempts. **[1]**

© OCR 2025

stemup.co.uk

BLANK PAGE

© OCR 2025

stemup.co.uk

Section B (86 marks)

5 The random variable X has a continuous uniform distribution over $[-4, 2]$.

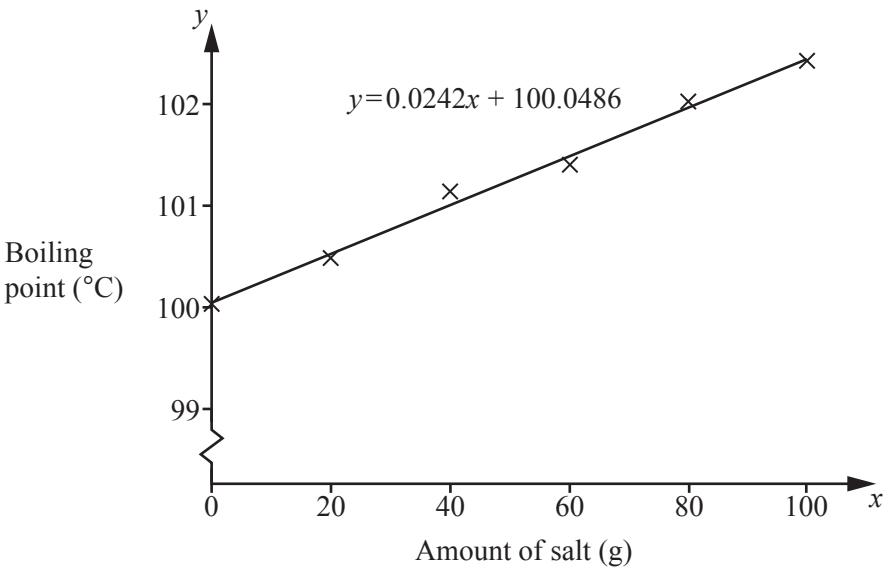
(a) Find the variance of X. **[2]**

(b) Determine the probability that if two independent values of X are taken, one is less than zero and the other is greater than zero. **[2]**

6 A student is investigating the link between the amount of salt dissolved in water and the boiling point of the water. The student dissolves fixed amounts of salt, x g, into water and then measures the boiling point, y °C, of the water.

The data are shown in the table together with a scatter diagram to illustrate them. The equation of the regression line of y on x is also shown.

Amount of salt, x g	0	20	40	60	80	100
Boiling point, y °C	100.03	100.49	101.15	101.41	102.04	102.44



(a) State what type of variable 'Boiling point' is. **[1]**

(b) Explain why 'Amount of salt' has been plotted on the horizontal axis. **[1]**

(c) Calculate an estimate of the boiling point with the following amounts of salt.

- 50 g
- 200 g

[2]

(d) Comment on the reliability of each of the estimates. **[2]**

(e) Determine the residual for the data point with 60 g of salt. **[2]**

stemup.co.uk

7 A student is investigating the relationship between different electricity generation methods and cost of electricity in a particular country. The student first checks whether there is any correlation between the cost per unit of electricity, x euros, and the amount of electricity being generated by wind, y GW. The data from 30 observations are summarised as follows.

$n = 30$ $\quad \Sigma x = 2.219$ $\quad \Sigma y = 357.7$ $\quad \Sigma x^2 = 0.2368$ $\quad \Sigma y^2 = 4648$ $\quad \Sigma xy = 25.01$

(a) In this question you must show detailed reasoning.

Determine the product moment correlation coefficient. **[4]**

(b) Carry out a hypothesis test at the 5% level to investigate whether there is any correlation between the cost per unit of electricity and the amount of electricity generated by wind. **[5]**

Table 7.1 shows the values of the product moment correlation coefficients between 5 variables. These variables are the cost per unit of electricity in euros and the amount of electricity generated in GW by four different types of generation, for a very large sample of observations.

For example, the correlation coefficient between the amount of electricity generated by wind and the amount of electricity generated by gas is −0.59. The 'storage' category consists of battery and other types of electricity taken from stored reserves. All of the correlations are significant at the 10% level.

Table 7.1

	Cost	Gas	Wind	Solar	Storage
Cost	1.00				
Gas	0.42	1.00			
Wind	−0.26	−0.59	1.00		
Solar	−0.18	−0.21	−0.07	1.00	
Storage	0.12	0.35	−0.38	−0.30	1.00

Table 7.2 shows standard guidelines for effect sizes.

Table 7.2

Product moment correlation coefficient	**Effect size**
$0.1 \leqslant r < 0.3$	Small
$0.3 \leqslant r < 0.5$	Medium
$r \geqslant 0.5$	Large

© OCR 2025

stemup.co.uk

The student analyses these data for effect size.

(c) Explain how the very large sample size relates to the interpretation of the correlation coefficient of −0.07 between the amount of electricity generated by wind and the amount of electricity generated by solar, as shown in **Table 7.1**. **[2]**

(d) Comment briefly on what the student might infer from these tables regarding different electricity generation methods and cost of electricity. **[2]**

8 A teacher is investigating whether students run more quickly round an athletics track if they run in a clockwise direction rather than in an anti-clockwise direction. She thinks that students will run more quickly on average in a clockwise direction.

The teacher selects 40 students at random and records how long it takes for each of them to run one lap of the track in a clockwise direction. The next day she records how long it takes for each of these students to run one lap in an anti-clockwise direction.

(a) Explain why the teacher does **not** get the students to run the second lap immediately after they have completed their first lap. **[1]**

(b) Explain why the teacher notes the names of each student, so that she knows which student took which time on each day, rather than just noting the 40 times on each day. **[1]**

The differences in times, d seconds (anti-clockwise time ─ clockwise time), are summarised as follows:

$\sum d = 72.2, \quad \sum d^2 = 1510$

(c) In this question you must show detailed reasoning.

Carry out a hypothesis test at the 5% significance level to investigate whether there is any evidence to suggest that on average students run more quickly round the track in a clockwise direction. **[10]**

(d) Comment briefly on any possible problem with the teacher's method. **[1]**

9 A consultant is modelling the number of trades per minute of a particular share in a stock exchange. The number of trades per minute in a random sample of 200 minutes is shown in the table.

Number of trades per minute	0	1	2	3	4	5	6	7	⩾8
Frequency	23	42	46	45	20	15	7	2	0

The consultant thinks that the number of trades per minute may be modelled by a Poisson distribution.

(a) Use the data in the table to estimate the value of the parameter λ for an appropriate Poisson model. **[2]**

The consultant decides to carry out a chi-squared goodness of fit test to investigate further. The screenshot below shows part of a spreadsheet to assess the goodness of fit of the distribution Po(λ), using the value of λ estimated from the data.

	A	B	C	D	E
1	Number of trades per minute	Frequency	Poisson probability	Expected frequency	Chi-squared contribution
2	0	23	0.0907	18.1436	1.2999
3	1	42			
4	2	46			0.7484
5	3	45	0.2090	41.8028	0.2445
6	4	20	0.1254	25.0817	1.0296
7	5	15	0.0602	12.0392	0.7281
8	⩾6	9	0.0357	7.1345	0.4878
9					

(b) Calculate the missing values in each of the following cells.

- C3
- D3
- E3

[4]

(c) Explain why the frequencies for 6, 7 and ⩾ 8 trades have been combined into the single category of ⩾ 6 trades, as shown in the spreadsheet. **[1]**

(d) In this question you must show detailed reasoning.

Carry out the test at the 5% significance level. **[6]**

© OCR 2025

stemup.co.uk

10 Two independent random variables, X and Y, have distributions B(6, 0.4) and B(3, 0.4) respectively. The random variable $T = X - (Y_1 + Y_2)$ where Y_1 and Y_2 are two independent values of Y.

(a) Determine P($T = 5$). **[4]**

A spreadsheet has been constructed to simulate the situation, as shown below. There are 100 simulated values of each of X, Y_1, Y_2 and T.

Cells G2 and G3 in the spreadsheet show the number of values of T which are greater than 0, and less than or equal to 0, respectively. Cell G5 shows the mean value of the 100 simulated values of T.

	A	B	C	D	E	F	G
1	X	Y_1	Y_2	$T=X-(Y_1+Y_2)$			
2	3	2	1	0		Number > 0	42
3	2	0	0	2		Number ⩽ 0	58
4	3	1	0	2			
5	1	3	2	−4		Mean value of T	0.25
6	4	2	1	1			
7							
8							
99	4	1	1	2			
100	2	2	2	−2			
101	3	0	3	0			
102							
103							

(b) Use the information in the spreadsheet to write down an estimate of P($T > 0$). **[1]**

As shown in the spreadsheet, the mean value of the 100 simulated values of T is 0.25.

(c) In this question you must show detailed reasoning.

Use the Central Limit Theorem to determine an estimate of the probability that the mean of 100 independent values of T is greater than 0.25. **[8]**

stemup.co.uk

11 The function F(x) is defined as follows.

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ \frac{ax+b}{9x+15} - \frac{4}{3} & 0 \leqslant x \leqslant 5 \\ 1 & x > 5 \end{cases}$$

where a and b are constants.

You are given that F(1) = 0.5 and F(2.5) = 0.8.

(a) Show that $a = 24$ and $b = 20$. **[3]**

(b) (i) Show that F(0) = 0. **[1]**

(ii) Show that F(x) is a non-decreasing function. **[3]**

(c) (i) State another property that is required for F(x) to be a cumulative distribution function. **[1]**

(ii) Show that the property in part **(c)(i)** is satisfied. **[1]**

You are given that F(x) is the cumulative distribution function of the continuous random variable X.

(d) Determine the upper quartile of X. Give your answer as an exact fraction. **[3]**

(e) Find E(X). **[2]**

(f) Determine the mode of X by considering the gradient of the probability density function of X. **[2]**

12 The random variable X takes the value 1 with probability p and the value 0 with probability $1 - p$. You are given that $\mathrm{E}(X) = p$ and $\mathrm{Var}(X) = p - p^2$.

The random variable $Y \sim B(64, 0.1)$ has mean μ and standard deviation σ.

Use the results given above to **prove** the following in either order.

- $\mu = 6.4$
- $\sigma = 2.4$ **[6]**

END OF QUESTION PAPER

stemup.co.uk

BLANK PAGE

© OCR 2025

stemup.co.uk

OCR
Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact The OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of Cambridge University Press & Assessment, which is itself a department of the University of Cambridge.

© OCR 2025

stemup.co.uk

OCR
Oxford Cambridge and RSA

Friday 13 June 2025 – Afternoon

A Level Further Mathematics B (MEI)

Y422/01 Statistics Major

Printed Answer Booklet

Time allowed: 2 hours 15 minutes

AB

Y422201

You must have:
- Question Paper Y422/01 (inside this document)
- the Formulae Booklet for Further Mathematics B (MEI)
- a scientific or graphical calculator

2045487966

Please write clearly in black ink. **Do not write in the barcodes.**

Centre number | | | | | | Candidate number | | | | |

First name(s) StemUp Model Solution

Last name

INSTRUCTIONS
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give your final answers to a degree of accuracy that is appropriate to the context.

INFORMATION
- This document has **20** pages.

ADVICE
- Read each question carefully before you start your answer.

© OCR 2025 [H/508/5594]
DC (ST) 351651/1

OCR is an exempt Charity

stemup.co.uk

Section A (34 marks)

1(a)

$P(X > 4) = P(X = 5, 6, 7 \text{ or } 8) = 4/8 = 1/2$

1(b)

X is discrete uniform on 1, 2, ..., n with $n = 8$

$E(X) = (n + 1)/2 = 9/2$

$\text{Var}(X) = (n^2 - 1)/12 = (64 - 1)/12 = 63/12$

$\text{E}(X) =$ 4.5

$\text{Var}(X) =$ 21/4 = 5.25

1(c)

$P(\text{score} < 4) = P(1, 2 \text{ or } 3) = 3/8$

$P(\text{score} > 4) = 1/2$

The low score can be on any one of the 3 rolls, so

$P = 3 \times (3/8) \times (1/2)^2$

$= 3 \times 3/8 \times 1/4$

$= 9/32$

© OCR 2025

2(a)

$R \sim N(20, 0.8^2)$, $P(R < 19) = 0.1056$ (3 s.f. 0.106)

2(b)

Let H = total height of the tower

$E(H) = 5 \times 20 + 5 \times 30 + 5 \times 50 = 500$

$\text{Var}(H) = 5 \times 0.8^2 + 5 \times 0.9^2 + 5 \times 1.2^2$

$= 5 \times (0.64 + 0.81 + 1.44) = 14.45$

$H \sim N(500, 14.45)$

$P(H \geqslant 495) = 0.9058$

$= 0.906$ (3 s.f.)

2(c)

$D = (R_1 + R_2 + R_3) - (Y_1 + Y_2)$

$E(D) = 3 \times 20 - 2 \times 30 = 0$

$\text{Var}(D) = 3 \times 0.8^2 + 2 \times 0.9^2 = 1.92 + 1.62 = 3.54$

$D \sim N(0, 3.54)$

$P(D \geqslant 1) = 0.2975$

$= 0.298$ (3 s.f.)

© OCR 2025

stemup.co.uk

3(a)

Standard error = $s/\sqrt{n} = 0.5664/\sqrt{50} = 0.0801$

3(b)

$24.878 - 0.157 < \mu < 24.878 + 0.157$

$24.721 < \mu < 25.035$

3(c)

No. The interval contains 25, so it does not suggest that the mean weight of potatoes per sack is different from 25 kg.

3(d)

For a 90% interval, $z = 1.645$

Interval: $24.878 \pm 1.645 \times 0.0801$

$= 24.878 \pm 0.1318$

$24.746 < \mu < 25.010$

This interval still contains 25 (just), so the conclusion would be the same: no suggestion that the mean weight per sack is different from 25 kg.

© OCR 2025

4(a)(i)

$X \sim Geo(0.4)$: $P(X = 5) = 0.6^4 \times 0.4 = 0.05184$

4(a)(ii)

$P(X > 5) = P(\text{first 5 attempts all miss}) = 0.6^5 = 0.07776$

4(b)

$E(X) = 1/p = 1/0.4 = 2.5$

$Var(X) = (1 - p)/p^2 = 0.6/0.4^2 = 3.75$

$sd = \sqrt{3.75} = 1.9365...$

$2.5 - 1.9365 = 0.5635, \quad 2.5 + 1.9365 = 4.4365$

Need $P(0.5635 < X < 4.4365) = P(1 \leq X \leq 4)$

$= 1 - P(X > 4) = 1 - 0.6^4$

$= 1 - 0.1296 = 0.8704$

4(c)

Y = number of baskets in 20 attempts, $Y \sim B(20, 0.4)$

$P(Y \geq 5) = 1 - P(Y \leq 4) = 1 - 0.05095$

$= 0.949$ (3 s.f.)

4(d)

Need exactly 4 baskets in the first 19 attempts, then a basket

$P = P(W = 4) \times 0.4$ where $W \sim B(19, 0.4)$

$= 0.046654... \times 0.4 = 0.0187$ (3 s.f.)

4(e)

It may not be valid: Gabi may get tired as the practice goes on, or improve with practice, so the probability may change and attempts may not be independent.

© OCR 2025

stemup.co.uk

Section B (86 marks)

5(a)

$X \sim U[-4, 2]$, so $a = -4$, $b = 2$

$\text{Var}(X) = (b - a)^2/12$

$= (2 - (-4))^2/12 = 36/12$

$= 3$

5(b)

$P(X < 0) = 4/6 = 2/3$, $P(X > 0) = 2/6 = 1/3$

Either the first is below zero and the second above, or

the other way round:

$P = 2 \times (2/3) \times (1/3)$

$= 4/9$

© OCR 2025

6(a)

Boiling point is a random (dependent / response) variable.

6(b)

Amount of salt is the independent (controlled) variable, as it is fixed by the student, not random, so it goes on the horizontal axis.

6(c)

50 g: $y = 0.0242 \times 50 + 100.0486 = 101.2586$

Boiling point ≈ 101.26 °C

200 g: $y = 0.0242 \times 200 + 100.0486 = 104.8886$

Boiling point ≈ 104.89 °C

6(d)

50 g: very reliable, since it is interpolation (within the range 0 to 100 g) and the points lie very close to the line.

200 g: unreliable, since it is extrapolation far beyond the data range, so the linear relationship may not continue.

6(e)

Predicted y at $x = 60$: $0.0242 \times 60 + 100.0486 = 101.5006$

Residual = observed − predicted = $101.41 - 101.5006$

$= -0.0906$

© OCR 2025

stemup.co.uk

7(a)

$S_{xy} = \Sigma xy - (\Sigma x \Sigma y)/n = 25.01 - (2.219 \times 357.7)/30$

$= 25.01 - 26.45788... = -1.44788...$

$S_{xx} = \Sigma x^2 - (\Sigma x)^2/n = 0.2368 - 2.219^2/30$

$= 0.2368 - 0.16413... = 0.0726680...$

$S_{yy} = \Sigma y^2 - (\Sigma y)^2/n = 4648 - 357.7^2/30$

$= 4648 - 4264.976... = 383.0237...$

$r = S_{xy} / \sqrt{S_{xx} S_{yy}}$

$= -1.44788... / \sqrt{0.0726680... \times 383.0237...}$

$= -1.44788... / \sqrt{27.8336...}$

$= -1.44788... / 5.27575...$

$r = -0.2744$ (4 d.p.)

7(b)

H_0: $\rho = 0$, H_1: $\rho \neq 0$

where ρ is the population product moment correlation coefficient between the cost per unit of electricity and the amount of electricity generated by wind.

Two-tailed test at 5% level, $n = 30$:

critical value = 0.3610

$|-0.2744| = 0.2744 < 0.3610$

so the result is not significant; do not reject H_0.

(answer space continued on next page)

© OCR 2025

7(b) **(continued)**

There is insufficient evidence at the 5% level to suggest that there is correlation between the cost per unit of electricity and the amount of electricity generated by wind.

7(c)

With a very large sample the critical value for r is very small, so even r = −0.07 is significant at the 10% level. However, r = −0.07 is very close to zero (below 0.1, so not even a small effect size), so although it is significant the correlation between wind and solar generation is so weak that it is of very little practical importance.

7(d)

Correlations with cost: gas 0.42 is a medium effect size; wind −0.26, solar −0.18 and storage 0.12 are all small. So the student might infer that the more electricity is generated by gas, the higher the cost of electricity tends to be (positive correlation), while more wind or solar generation is only weakly associated with lower cost. (Correlation does not imply causation.)

© OCR 2025

stemup.co.uk

8(a)

The students would be tired after their first lap, so they would probably run the second lap more slowly, which would bias the comparison.

8(b)

So that the data are paired: she can find the difference in times for each student, which removes the differences in fitness/ability between individual students.

8(c)

H_0: $\mu = 0$, H_1: $\mu > 0$

where μ is the population mean difference in lap times (anti-clockwise time − clockwise time) for students.

$n = 40$, $\bar{d} = 72.2/40 = 1.805$

$s^2 = (1/39)(\Sigma d^2 - (\Sigma d)^2/40)$

$= (1/39)(1510 - 72.2^2/40)$

$= (1/39) \times 1379.679 = 35.376...$

$s = 5.9478...$

$n = 40$ is large, so by the Central Limit Theorem, under H_0

$\bar{D} \sim N(0, 35.376.../40)$ approximately (s used for σ).

Test statistic $z = (1.805 - 0)/\sqrt{35.376.../40}$

$= 1.805/0.94043... = 1.919$

One-tailed test at 5% level: critical value = 1.645

(answer space continued on next page)

© OCR 2025

8(c)	**(continued)**
	1.919 > 1.645, so the result is significant; reject H_0.
	There is sufficient evidence at the 5% level to suggest that
	on average students run more quickly round the track in a
	clockwise direction (so the teacher is correct).
8(d)	The laps were run on different days, so conditions (eg the
	weather or wind) may have been different on the second day,
	which could affect the times.

DO NOT WRITE IN THIS SPACE

© OCR 2025

stemup.co.uk

9(a)

$\lambda = \Sigma fx/\Sigma f$

$= (0\times23 + 1\times42 + 2\times46 + 3\times45 + 4\times20 + 5\times15 + 6\times7 + 7\times2)/200$

$= (0 + 42 + 92 + 135 + 80 + 75 + 42 + 14)/200$

$= 480/200 = 2.4$

9(b)

$X \sim Po(2.4)$

C3: $P(X = 1) = e^{-2.4} \times 2.4 = 0.21772...$

D3: expected $= 200 \times 0.21772... = 43.5446$

E3: $(O - E)^2/E = (42 - 43.5446)^2/43.5446$

$= 0.0548$

C3 = 0.2177

D3 = 43.5446

E3 = 0.0548

9(c)

The expected frequencies for 6, 7 and ≥ 8 trades on their own would be less than 5, too small for the chi-squared test to be valid, so they are combined into one category.

© OCR 2025

9(d)

H_0: Po(2.4) is a good model for the number of trades per minute

H_1: Po(2.4) is not a good model for the number of trades per minute

$X^2 = \Sigma(O - E)^2/E$

$= 1.2999 + 0.0548 + 0.7484 + 0.2445 + 1.0296 + 0.7281 + 0.4878$

$= 4.5931$

Number of cells = 7

Degrees of freedom = 7 − 1 − 1 = 5

(1 for the total, 1 for estimating λ from the data)

Critical value for χ^2_5 at 5% level = 11.07

4.5931 < 11.07, so the result is not significant;

do not reject H_0.

There is insufficient evidence at the 5% level to suggest that the Poisson(2.4) model is not a good fit for the number of trades per minute.

© OCR 2025

stemup.co.uk

10(a)

Let $S = Y_1 + Y_2 \sim B(6, 0.4)$ (independent, same p)

$T = X - S = S$ when $(X = S, S = 0)$ or $(X = 6, S = 1)$

$P(X = S) = 6 \times 0.4^S \times 0.6 = 0.036864$

$P(X = 6) = 0.4^6 = 0.004096$

$P(S = 0) = 0.6^6 = 0.046656$

$P(S = 1) = 6 \times 0.4 \times 0.6^S = 0.186624$

$P(T = S) = 0.036864 \times 0.046656 + 0.004096 \times 0.186624$

$= 0.0017199... + 0.0007644...$

$= 0.00248$ (3 s.f.)

10(b)

42 of the 100 simulated values are greater than 0, so

$P(T > 0) \approx 42/100 = 0.42$

© OCR 2025

10(c)

$E(X) = 6 \times 0.4 = 2.4, \quad E(Y) = 3 \times 0.4 = 1.2$

$Var(X) = 6 \times 0.4 \times 0.6 = 1.44$

$Var(Y) = 3 \times 0.4 \times 0.6 = 0.72$

$E(T) = E(X) - E(Y_1) - E(Y_2) = 2.4 - 1.2 - 1.2 = 0$

$Var(T) = Var(X) + Var(Y_1) + Var(Y_2)$

$= 1.44 + 0.72 + 0.72 = 2.88$

Let $\bar{T}$ be the mean of 100 independent values of T.

$E(\bar{T}) = 0, \quad Var(\bar{T}) = 2.88/100 = 0.0288$

By the Central Limit Theorem ($n = 100$ is large),

$\bar{T} \sim N(0, 0.0288)$ approximately.

T only takes integer values, so the sum of 100 values is an integer and $\bar{T}$ goes up in steps of 0.01.

$\bar{T} > 0.25$ means $\bar{T} \geq 0.26$, so with a continuity correction use 0.255:

$P(\bar{T} > 0.25) \approx P(\bar{T} > 0.255)$

$= P(Z > 0.255/\sqrt{0.0288}) = P(Z > 1.5026)$

$= 0.0665$ (3 s.f.)

© OCR 2025

stemup.co.uk

11(a)

$F(1) = (a + b)/24 - 4/3 = 0.5 \Rightarrow (a + b)/24 = 11/6$

$\Rightarrow a + b = 44 \quad (1)$

$F(2.5) = (2.5a + b)/37.5 - 4/3 = 0.8 \Rightarrow (2.5a + b)/37.5 = 32/15$

$\Rightarrow 2.5a + b = 80 \quad (2)$

$(2) - (1)$: $1.5a = 36 \Rightarrow a = 24$

$b = 44 - 24 = 20$

So $a = 24$ and $b = 20$, as required.

11(b)(i)

$F(0) = (24 \times 0 + 20)/(9 \times 0 + 15) - 4/3 = 20/15 - 4/3$

$= 4/3 - 4/3 = 0$, as required.

11(b)(ii)

$F(x) = (24x + 20)/(9x + 15) - 4/3$ for $0 \leq x \leq 5$

Quotient rule:

$F'(x) = [24(9x + 15) - 9(24x + 20)]/(9x + 15)^2$

$= (216x + 360 - 216x - 180)/(9x + 15)^2$

$= 180/(9x + 15)^2$

Since $(9x + 15)^2 > 0$, $F'(x) > 0$ for $0 \leq x \leq 5$, so F is increasing on [0, 5]; it is constant (0 and 1) elsewhere.

So $F(x)$ is a non-decreasing function.

11(c)(i)

$F(5) = 1$ (F must reach 1 at the upper end of the range)

© OCR 2025

11(c)(ii)

$F(5) = (24 \times 5 + 20)/(9 \times 5 + 15) - 4/3 = 140/60 - 4/3$

$= 7/3 - 4/3 = 1$, so the property is satisfied.

11(d)

Upper quartile q: $F(q) = 3/4$

$(24q + 20)/(9q + 15) - 4/3 = 3/4$

$(24q + 20)/(9q + 15) = 25/12$

$12(24q + 20) = 25(9q + 15)$

$288q + 240 = 225q + 375$

$63q = 135 \Rightarrow q = 15/7$

11(e)

$f(x) = F'(x) = 180/(9x + 15)^2$, $0 \leq x \leq 5$

$E(X) = \int x\, f(x)\, dx$ from 0 to 5 $= \int 180x/(9x + 15)^2\, dx$

Let $u = 9x + 15$, $x = (u - 15)/9$, $dx = du/9$, u from 15 to 60:

$E(X) = (20/9) \int (1/u - 15/u^2)\, du$ from 15 to 60

$= (20/9)[\ln u + 15/u]$ from 15 to 60 $= (20/9)(\ln 4 - 3/4)$

$= 1.414$ (4 s.f.)

11(f)

$f(x) = 180(9x + 15)^{-2}$

$f'(x) = -2 \times 9 \times 180(9x + 15)^{-3} = -3240/(9x + 15)^3$

For $0 \leq x \leq 5$, $(9x + 15)^3 > 0$, so $f'(x) < 0$.

So $f(x)$ is decreasing over the whole range, and its greatest value is at $x = 0$.

Mode of $X = 0$

© OCR 2025

stemup.co.uk

12

$Y \sim B(64, 0.1)$ counts successes in 64 independent trials.

Let $X_i = 1$ if trial i is a success, 0 if not $(i = 1, ..., 64)$,

so each X_i is as given with $p = 0.1$:

$E(X_i) = p = 0.1$

$Var(X_i) = p - p^2 = 0.1 - 0.01 = 0.09$

Y = number of successes in 64 trials

$Y = X_1 + X_2 + ... + X_{64}$

$\mu = E(Y) = E(X_1) + E(X_2) + ... + E(X_{64})$

$= 64 \times E(X_i) = 64 \times 0.1$

$= 6.4$, as required.

Since the X_i are independent,

$Var(Y) = Var(X_1) + Var(X_2) + ... + Var(X_{64})$

$= 64 \times Var(X_i) = 64 \times 0.09 = 5.76$

$\sigma = \sqrt{5.76} = 2.4$, as required.

© OCR 2025

EXTRA ANSWER SPACE

If you need extra space use these lined pages. You must write the question numbers clearly in the margin.

© OCR 2025

stemup.co.uk

OCR
Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact The OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of Cambridge University Press & Assessment, which is itself a department of the University of Cambridge.

© OCR 2025

stemup.co.uk