Write your name here

Surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number	Candidate Number

Further Mathematics

Advanced Subsidiary
Further Mathematics options
Paper 2G: Further Statistics 1 and Further Statistics 2

Sample Assessment Material for first teaching September 2017
Time: 1 hour 40 minutes

Paper Reference
8FM0/2G

You must have:
Mathematical Formulae and Statistical Tables, calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- There are **two** sections in this question paper. Answer **all** the questions in Section A and **all** the questions in Section B.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 8 questions in this question paper. The total mark for this paper is 80.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

S58533A

1/1/1/

S 5 8 5 3 3 A 0 1 3 0

Pearson

SECTION A

Answer ALL questions. Write your answers in the spaces provided.

1. A university foreign language department carried out a survey of prospective students to find out which of three languages they were most interested in studying.

A random sample of 150 prospective students gave the following results.

		Language		
		French	**Spanish**	**Mandarin**
Gender	**Male**	23	22	20
	Female	38	32	15

A test is carried out at the 1% level of significance to determine whether or not there is an association between gender and choice of language.

(a) State the null hypothesis for this test.
(1)

(b) Show that the expected frequency for females choosing Spanish is 30.6
(1)

(c) Calculate the test statistic for this test, stating the expected frequencies you have used.
(3)

(d) State whether or not the null hypothesis is rejected. Justify your answer.
(2)

(e) Explain whether or not the null hypothesis would be rejected if the test was carried out at the 10% level of significance.
(1)

Question 1 continued

Question 1 continued

Question 1 continued

(Total for Question 1 is 8 marks)

2. The discrete random variable X has probability distribution given by

x	-1	0	1	2	3
$P(X = x)$	c	a	a	b	c

The random variable $Y = 2 - 5X$

Given that $E(Y) = -4$ and $P(Y \geqslant -3) = 0.45$

(a) find the probability distribution of X.

(7)

Given also that $E(Y^2) = 75$

(b) find the exact value of $\text{Var}(X)$

(2)

(c) Find $P(Y > X)$

(2)

Question 2 continued

Question 2 continued

Question 2 continued

(Total for Question 2 is 11 marks)

3. Two car hire companies hire cars independently of each other.

Car Hire A hires cars at a rate of 2.6 cars per hour.

Car Hire B hires cars at a rate of 1.2 cars per hour.

(a) In a 1 hour period, find the probability that each company hires exactly 2 cars.
(2)

(b) In a 1 hour period, find the probability that the total number of cars hired by the two companies is 3
(2)

(c) In a 2 hour period, find the probability that the total number of cars hired by the two companies is less than 9
(2)

On average, 1 in 250 new cars produced at a factory has a defect.

In a random sample of 600 new cars produced at the factory,

(d) (i) find the mean of the number of cars with a defect,

(ii) find the variance of the number of cars with a defect.
(2)

(e) (i) Use a Poisson approximation to find the probability that no more than 4 of the cars in the sample have a defect.

(ii) Give a reason to support the use of a Poisson approximation.
(2)

Question 3 continued

Question 3 continued

Question 3 continued

(Total for Question 3 is 10 marks)

4. The discrete random variable X follows a Poisson distribution with mean 1.4

(a) Write down the value of

(i) $P(X = 1)$

(ii) $P(X \leqslant 4)$

(2)

The manager of a bank recorded the number of mortgages approved each week over a 40 week period.

Number of mortgages approved	0	1	2	3	4	5	6
Frequency	10	16	7	4	2	0	1

(b) Show that the mean number of mortgages approved over the 40 week period is 1.4

(1)

The bank manager believes that the Poisson distribution may be a good model for the number of mortgages approved each week.

She uses a Poisson distribution with a mean of 1.4 to calculate expected frequencies as follows.

Number of mortgages approved	0	1	2	3	4	5 or more
Expected frequency	9.86	r	9.67	4.51	1.58	s

(c) Find the value of r and the value of s giving your answers to 2 decimal places.

(2)

The bank manager will test, at the 5% level of significance, whether or not the data can be modelled by a Poisson distribution.

(d) Calculate the test statistic and state the conclusion for this test. State clearly the degrees of freedom and the hypotheses used in the test.

(6)

Question 4 continued

Question 4 continued

Question 4 continued

(Total for Question 4 is 11 marks)

TOTAL FOR SECTION A IS 40 MARKS

SECTION B

Answer ALL questions. Write your answers in the spaces provided.

5. In a gymnastics competition, two judges scored each of 8 competitors on the vault.

Competitor	A	B	C	D	E	F	G	H
Judge 1's scores	4.6	9.1	8.4	8.8	9.0	9.5	9.2	9.4
Judge 2's scores	7.8	8.8	8.6	8.5	9.1	9.6	9.0	9.3

(a) Calculate Spearman's rank correlation coefficient for these data.
(4)

(b) Stating your hypotheses clearly, test at the 1% level of significance, whether or not the two judges are generally in agreement.
(4)

(c) Give a reason to support the use of Spearman's rank correlation coefficient in this case.
(1)

The judges also scored the competitors on the beam.

Spearman's rank correlation coefficient for their ranks on the beam was found to be 0.952

(d) Compare the judges' ranks on the vault with their ranks on the beam.
(1)

Question 5 continued

(Total for Question 5 is 10 marks)

6. The continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{1}{18}(11 - 2x) & 1 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $P(X < 3)$

(2)

(b) State, giving a reason, whether the upper quartile of X is greater than 3, less than 3 or equal to 3

(1)

Given that $E(X) = \frac{9}{4}$

(c) use algebraic integration to find $\text{Var}(X)$

(3)

The cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{18}(11x - x^2 + c) & 1 \leqslant x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

(d) Show that $c = -10$

(2)

(e) Find the median of X, giving your answer to 3 significant figures.

(3)

Question 6 continued

Question 6 continued

Question 6 continued

(Total for Question 6 is 11 marks)

7. A scientist wants to develop a model to describe the relationship between the average daily temperature, x °C, and a household's daily energy consumption, y kWh, in winter.

A random sample of the average temperature and energy consumption are taken from 10 winter days and are summarised below.

$$\sum x = 12 \quad \sum x^2 = 24.76 \quad \sum y = 251 \quad \sum y^2 = 6341 \quad \sum xy = 284.8$$

$$S_{xx} = 10.36 \qquad S_{yy} = 40.9$$

(a) Find the product moment correlation coefficient between y and x.
(2)

(b) Find the equation of the regression line of y on x in the form $y = a + bx$
(3)

(c) Use your equation to estimate the daily energy consumption when the average daily temperature is 2 °C
(1)

(d) Calculate the residual sum of squares (RSS).
(2)

The table shows the residual for each value of x.

x	−0.4	−0.2	0.3	0.8	1.1	1.4	1.8	2.1	2.5	2.6
Residual	−0.63	−0.32	−0.52	−0.73	0.74	2.22	1.84	0.32	f	−1.88

(e) Find the value of f.
(2)

(f) By considering the signs of the residuals, explain whether or not the linear regression model is a suitable model for these data.
(1)

Question 7 continued

Question 7 continued

Question 7 continued

(Total for Question 7 is 11 marks)

8. The continuous random variable X is uniformly distributed over the interval $[-3, 5]$.

(a) Sketch the probability density function $f(x)$ of X.

(2)

(b) Find the value of k such that $P(X < 2[k - X]) = 0.25$

(3)

(c) Use algebraic integration to show that $E(X^3) = 17$

(3)

Question 8 continued

Question 8 continued

(Total for Question 8 is 8 marks)

TOTAL FOR SECTION B IS 40 MARKS

TOTAL FOR PAPER IS 80 MARKS

Paper 2 Option G

Further Statistics 1 Mark Scheme (Section A)

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AOs</th></tr>
<tr><td>1(a)</td><td>H0: There is no association between language and gender</td><td>B1</td><td>1.2</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(b)</td><td>$\frac{54\times 85}{150} = 30.6$ *</td><td>B1*cso</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(c)</td><td>
<table>
<tr><td colspan="2" rowspan="2">Expected frequencies</td><td colspan="3">Language</td></tr>
<tr><td>French</td><td>Spanish</td><td>Mandarin</td></tr>
<tr><td rowspan="2">Gender</td><td>Male</td><td>26.43…</td><td>23.4</td><td>15.16…</td></tr>
<tr><td>Female</td><td>34.56…</td><td>[30.6]</td><td>19.83…</td></tr>
</table>
</td><td>M1</td><td>2.1</td></tr>
<tr><td></td><td>$\chi^2 = \sum\frac{(O-E)^2}{E} = \frac{(23-26.43)^2}{26.43} + \ldots + \frac{(15-19.83)^2}{19.83}$</td><td>M1</td><td>1.1b</td></tr>
<tr><td></td><td>Awrt <u>3.6/3.7</u></td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(3)</td><td></td></tr>
<tr><td>(d)</td><td>Degrees of freedom (3 – 1)(2 – 1) → Critical value $\chi^2_{2,0.01} = 9.210$</td><td>M1</td><td>3.1b</td></tr>
<tr><td></td><td>As $\sum\frac{(O-E)^2}{E} < 9.210$, the null hypothesis is not rejected</td><td>A1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(2)</td><td></td></tr>
<tr><td>(e)</td><td>Still not rejected since $\sum\frac{(O-E)^2}{E} < \chi^2_{2,0.1} = 4.605$</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td colspan="4" align="right">(8 marks)</td></tr>
</table>

Notes:

(a)
B1: For correct hypothesis in context

(b)
B1*: For a correct calculation leading to the given answer and no errors seen

(c)
M1: For attempt at $\frac{\text{(Row Total)(Column Total)}}{\text{(Grand Total)}}$ to find expected frequencies

M1: For applying $\sum\frac{(O-E)^2}{E}$

A1: awrt 3.6 or 3.7

(d)
M1: For using degrees of freedom to set up a χ^2 model critical value
A1: For correct comparison and conclusion

(e)
A1ft: For correct conclusion with supporting reason

Question	Scheme	Marks	AOs
2(a)	$-4 = 2 - 5\text{E}(X)$	M1	3.1a
	$\text{E}(X) = 1.2$		
	$-1\times c + 0\times a + 1\times a + 2\times b + 3\times c = 1.2$	M1	1.1b
	$a + 2b + 2c = 1.2$ $\boxed{1}$		
	$\text{P}(Y \geqslant -3) = 0.45$ gives $\text{P}(2 - 5X \geqslant -3) = 0.45$ i.e. $\text{P}(X \leqslant 1) = 0.45$	M1	2.1
	$2a + c = 0.45$ $\boxed{2}$		
	$2a + b + 2c = 1$ $\boxed{3}$	M1	1.1b
	$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1.2 \\ 0.45 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 2 & -3 \\ -2 & -3 & 4 \end{pmatrix}\begin{pmatrix} 1.2 \\ 0.45 \\ 1 \end{pmatrix}$ <u>or</u>	M1	1.1b
	e.g. $\boxed{3} - \boxed{2} \Rightarrow b + c = 0.55$ sub. $2(b+c)$ into $\boxed{1} \Rightarrow a = 0.1$ etc		
	$a = 0.1$ $b = 0.3$ $c = 0.25$	A1 A1	1.1b 1.1b
		(7)	
(b)	$\text{Var}(Y) = 75 - (-4)^2$ <u>or</u> 59	M1	1.1a
	[$\text{Var}(Y) = 5^2\text{Var}(X)$ implies] $\text{Var}(X) = 2.36$	A1	1.2
		(2)	
(c)	$\text{P}(Y > X) = \text{P}(2 - 5X > X) \rightarrow \text{P}(X < \frac{1}{3})$	M1	3.1a
	$\text{P}(X < \frac{1}{3}) = a + c = 0.35$	A1ft	1.1b
		(2)	
		(11 marks)	

Notes:

(a)

M1: For using given information to find an expression for E(X) i.e. use of $\text{E}(Y) = 2 - 5\text{E}(X)$

M1: For use of $\sum x\text{P}(X = x) =$ '1.2'

M1: For use of $\text{P}(Y \geqslant -3) = 0.45$ to set up the argument for solving by forming an equation in a and c

M1: For use of $\sum \text{P}(X = x) = 1$

M1: For solving their 3 linear equations (matrix or elimination)

A1: For any 2 of a, b or c correct

A1: For all 3 correct values

Question 2 notes continued:	
Another method for part (a) is:	
M1:	For using given information to find the probability distribution for Y leading to an expression for E(Y)
M1:	For use of $\sum y\text{P}(Y = y) = -4$
M1:	For use of P($Y \geqslant -3$) = 0.45 to set up the argument for solving by forming an equation in a and c
M1:	For use of $\sum \text{P}(Y = y) = 1$
M1:	For solving their 3 linear equations (matrix or elimination)
A1:	For any 2 of a, b or c correct
A1:	For all 3 correct values
(b)	
M1:	For use of $\text{Var}(Y) = \text{E}(Y^2) - [\text{E}(Y)]^2$ (may be implied by a correct answer)
A1:	For use of $\text{Var}(aX) = a^2\,\text{Var}(X)$ to reach 2.36 or exact equivalent
(c)	
M1:	For rearranging to the form P($X < k$)
A1ft:	0.1' + '025' (provided their a and c and their $a + c$ are all probabilities)
Another method for part (c) is:	
M1:	For comparing distribution of X with distribution of Y to identify $X = -1$ and $X = 0$
A1ft:	'0.1' + '025' (provided their a and c and their $a + c$ are all probabilities)

Question	Scheme	Marks	AOs
3(a)	$X \sim \text{Po}(2.6)$ $\quad Y \sim \text{Po}(1.2)$		
	P(each hire 2 in 1 hour) $= \text{P}(X{=}2) \times \text{P}(Y{=}2) = 0.25104\ldots \times 0.21685\ldots$	M1	3.3
	$= 0.05444\ldots$ awrt **0.0544**	A1	1.1b
		(2)	
(b)	$W = X + Y \rightarrow \quad W \sim \text{Po}(3.8)$	M1	3.4
	$\text{P}(W = 3) = 0.20458\ldots$ awrt **0.205**	A1	1.1b
		(2)	
(c)	$T \sim \text{Po}((2.6{+}1.2){\times}2)$	M1	3.3
	$\text{P}(T < 9) = 0.64819\ldots$ awrt **0.648**	A1	1.1b
		(2)	
(d)	**(i)** Mean $= np =$ **2.4**	B1	1.1b
	(ii) Variance $= np(1-p) = 2.3904$ awrt **2.39**	B1	1.1b
		(2)	
(e)	**(i)** [$D \sim \text{Po}(2.4)$ $\quad \text{P}(D \leqslant 4)$] $= 0.9041\ldots$ awrt **0.904**	B1	1.1b
	(ii) Since n is large and p is small/mean is approximately equal to variance	B1	2.4
		(2)	
			(10 marks)

Notes:

(a)
M1: For $\text{P}(X{=}2) \times \text{P}(Y{=}2)$ from $X \sim \text{Po}(2.6)$ and $Y \sim \text{Po}(1.2)$ i.e. correct models (may be implied by correct answer)
A1: **awrt 0.0544**

(b)
M1: For combining Poisson distributions and use of Po('3.8') (may be implied by correct answer)
A1: **awrt 0.205**

(c)
M1: For setting up a new model and attempting mean of Poisson distribution (may be implied by correct answer)
A1: **awrt 0.648**

(d)(i)
B1: For **2.4**

(d)(ii)
B1: For **awrt 2.39**

(e)(i)
B1: For **awrt 0.904**

(e)(ii)
B1: For a correct explanation to support use of Poisson approximation in this case

Question	Scheme	Marks	AOs
4(a)	**(i)** P(X = 1) = 0.34523… awrt **0.345**	B1	1.1b
	(ii) P($X \leqslant$ 4) = 0.98575… awrt **0.986**	B1	1.1b
		(2)	
(b)	$\frac{(0\times10)+1\times16+2\times7+3\times4+4\times2+(5\times0)+6\times1}{40}=1.4$ *	B1*cso	1.1b
		(1)	
(c)	r = 40 × '0.34523…' s = 40 × '1 – 0.986…'	M1	3.4
	r = **13.81** s = **0.57**	A1ft	1.1b
		(2)	
(d)	H_0: The Poisson distribution is a suitable model H_1: The Poisson distribution is not a suitable model	B1	3.4
	[Cells are combined when expected frequencies < 5] So combine the last 3 cells	M1	2.1
	$\chi^2=\sum\frac{(O-E)^2}{E}=\frac{(10-9.86)^2}{9.86}+\ldots+\frac{(7-(4.51+1.58+0.57))^2}{(4.51+1.58+0.57)}$	M1	1.1b
	awrt **1.1**	A1	1.1b
	Degrees of freedom = 4 – 1 – 1 = 2	B1	3.1b
	(Do not reject H_0 since $1.10<\chi^2_{2,(0.05)}=5.991$). The number of mortgages approved each week follows a Poisson distribution	A1	3.5a
		(6)	
		(11 marks)	

Notes:

(a)(i)
B1: awrt 0.345

(a)(ii)
B1: awrt 0.986

(b)
B1*: For a fully correct calculation leading to given answer with no errors seen

(c)
M1: For attempt at r or s (may be implied by correct answers)
A1ft: For both values correct (follow through their answers to part (a))

(d)
B1: For both hypotheses correct (lambda should not be defined so correct use of the model)
M1: For understanding the need to combine cells before calculating the test statistic (may be implied)
M1: For attempt to find the test statistic using $\chi^2=\sum\frac{(O-E)^2}{E}$
A1: awrt 1.1
B1: For realising that there are 2 degrees of freedom leading to a critical value of $\chi_2^2(0.05)=5.991$
A1: Concluding that a Poisson model is suitable for the number of mortgages approved each week

Further Statistics 2 Mark Scheme (Section B)

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AOs</th></tr>
<tr><td>5(a)</td><td>
<table>
<tr><td>Competitor</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td></tr>
<tr><td>Judge 1's ranks</td><td>8</td><td>4</td><td>7</td><td>6</td><td>5</td><td>1</td><td>3</td><td>2</td></tr>
<tr><td>Judge 2's ranks</td><td>8</td><td>5</td><td>6</td><td>7</td><td>3</td><td>1</td><td>4</td><td>2</td></tr>
<tr><td>d^2</td><td>0</td><td>1</td><td>1</td><td>1</td><td>4</td><td>0</td><td>1</td><td>0</td></tr>
</table>
$\sum d^2 = 8$
$r_s = 1 - \frac{6 \times 8}{8(64-1)}$</td><td>M1
M1
dM1</td><td>1.1b
1.1b
1.1b</td></tr>
<tr><td></td><td>$r_s = 0.90476\ldots$ awrt <u>0.905</u></td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td></td><td>(4)</td><td></td></tr>
<tr><td>(b)</td><td>H0: $\rho_s = 0$ H1: $\rho_s > 0$</td><td>B1</td><td>2.5</td></tr>
<tr><td></td><td>Critical value $\rho_s = 0.8333$</td><td>B1</td><td>1.1b</td></tr>
<tr><td></td><td>$r_s = 0.905$ lies in the critical region/reject H0</td><td>M1</td><td>2.1</td></tr>
<tr><td></td><td>The two judges are in agreement.</td><td>A1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(4)</td><td></td></tr>
<tr><td>(c)</td><td>E.g. The data is unlikely to be from a bivariate normal distribution (competitor A)/The emphasis here is on the ranks and not the individual scores.</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td>(d)</td><td>Both show positive correlation, but the judges agree more on the beam (since 0.952 is closer to 1)</td><td>B1</td><td>2.2b</td></tr>
<tr><td></td><td></td><td>(1)</td><td></td></tr>
<tr><td colspan="4">(10 marks)</td></tr>
</table>

Notes:

(a)

M1: For an attempt to rank at least one row (at least four correct)

M1: For an attempt at d^2 row for their ranks

M1: Dependent on 1st M1 for use of $r_s = 1 - \frac{6 \times 8}{8(64-1)}$ with their $\sum d^2$

A1: For awrt 0.905

(b)

B1: Both hypotheses stated in terms of ρ_s

B1: For correct critical value

M1: For comparing their '0.905' with their '0.8333'

A1: For a correct contextual conclusion with no contradictions seen

(c)

B1: For a correct explanation to support the use of Spearman

(d)

B1: For a correct comparison of the correlation coefficients

Question	Scheme	Marks	AOs
6(a)	$P(X<3) = \int_1^3 \frac{1}{18}(11-2x)dx$ <u>or</u> area of trapezium	M1	1.1a
	$= \left[\frac{1}{18}(11x - x^2)\right]_1^3$		
	$= \frac{7}{9}$	A1	1.1b
		(2)	
(b)	Since $P(X<3) > 0.75$, the upper quartile is less than 3	B1ft	2.2a
		(1)	
(c)	$E(X^2) = \int_1^4 \frac{1}{18}x^2(11-2x)dx \left[= \frac{23}{4}\right]$	M1	1.1b
	$Var(X) = \frac{23}{4} - \left(\frac{9}{4}\right)^2$	M1	1.1b
	$= \frac{11}{16}$	A1	1.1b
		(3)	
(d)	$F(4) = 1 \rightarrow \frac{1}{18}(11(4) - 4^2 + c) = 1$ <u>or</u> $F(1) = 0 \rightarrow \frac{1}{18}(11(1) - 1^2 + c) = 0$	M1	2.1
	$c = -10$ *	A1*cso	1.1b
		(2)	
(e)	$F(m) = 0.5$	M1	1.2
	$\frac{1}{18}(11m - m^2 - 10) = 0.5 \rightarrow m^2 - 11m + 19 = 0$ and attempt to solve	M1	1.1b
	$m = \frac{11 \pm \sqrt{11^2 - 4(19)}}{2}$ [= 2.1458 or 8.8541…]		
	$m = 2.1458\ldots$ <u>**2.15**</u> (only)	A1	2.2a
		(3)	
		(11 marks)	

Notes:

(a)

M1: For integrating f(x) with correct limits **or** for finding area of trapezium

A1: For $\frac{7}{9}$ (allow awrt 0.778)

(b)

B1ft: For comparison of their (a) with 0.75 and concluding that the upper quartile is less than 3

(c)

M1: For an attempt to find $E(X^2)$

M1: For use of $Var(X) = E(X^2) - \left(\frac{9}{4}\right)^2$

A1: For $\frac{11}{16}$ (allow awrt 0.688)(M1 marks may be implied by a correct answer)

Question 6 notes continued:
(d) **M1:** For use of F(4) = 1 or F(1) = 0 **A1*cso:** For a fully correct solution leading to given answer with no errors seen
(e) **M1:** For use of F(m) = 0.5 **M1:** For setting up quadratic and attempt to solve **A1:** For 2.15 and rejecting the other solution

Question	Scheme	Marks	AOs
7(a)	$r = \dfrac{284.4 - \frac{251(12)}{10}}{\sqrt{10.36 \times 40.9}}$	M1	1.1b
	$r = -0.79671...$ awrt **$\underline{-0.797}$**	A1	1.1b
		(2)	
(b)	$b = \dfrac{'-16.4'}{10.36}$	M1	3.3
	$a = \dfrac{251}{10} - 'b'\dfrac{12}{10}$	M1	1.1b
	$y = 27.0 - 1.58x$	A1	1.1b
		(3)	
(c)	$y = [27.0 - 1.58(2)] = 23.84$ awrt **$\underline{23.8}$**	B1ft	3.4
		(1)	
(d)	$\text{RSS} = 40.9 - \dfrac{(-16.4)^2}{10.36}$	M1	1.1b
	$\text{RSS} = 14.938...$ awrt **$\underline{14.9}$**	A1	1.1b
		(2)	
(e)	$\sum\text{residuals} = 0 \rightarrow -0.63 + (-0.32) + \ldots + f + (-1.88) = 0$	M1	3.1a
	$f = \underline{\mathbf{-1.04}}$	A1	1.1b
		(2)	
(f)	The residuals should be randomly scattered above and below zero so linear model may not be appropriate	B1	3.5b
		(1)	
		(11 marks)	

Notes:

(a)
M1: For a complete correct method for finding r
A1: For awrt –0.797

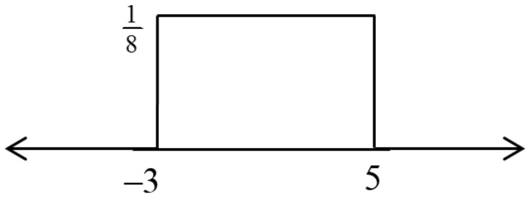
(b)
M1: For use of a correct model i.e. a correct expression for b (ft their S_{xy})
M1: For use of a correct model i.e. a correct (ft) expression for a
A1: For $y = 27.0 - 1.58x$ [a correct answer here can imply both method marks]

(c)
B1: For awrt 23.8 (evaluating their model found in part (b) with $x = 2$)

(d)
M1: For a correct expression for RSS
A1: For awrt 14.9

(e)
M1: For use of $\sum\text{residuals} = 0$ [Use of regression equation needs correct sign]
A1: For -1.04

(f)
B1: For identifying that the residuals are not randomly scattered above and below zero and concluding the linear regression model may not be appropriate

Question	Scheme	Marks	AOs
8(a)		B1 (shape) B1 (labels)	1.1b 1.1b
		(2)	
(b)	$P(X < 2(k-X)) = P(X < \frac{2}{3}k)$	M1	3.1a
	$\frac{\frac{2}{3}k-(-3)}{5-(-3)} = 0.25$	M1	1.1b
	$k = -\frac{3}{2}$	A1	1.1b
		(3)	
(c)	$E(X^3) = \int_{-3}^{5} \frac{1}{5-(-3)} x^3 dx$	M1	2.1
	$= \left[\frac{1}{32}x^4\right]_{-3}^{5} = \frac{1}{32}(5^4 - (-3)^4)$	dM1	1.1b
	$= 17$*	A1*cso	1.1b
		(3)	
			(8 marks)

Notes:

(a)
B1: For correct shape
B1: For correct labels

(b)
M1: For simplifying to $P(X < \frac{2}{3}k)$
M1: For equating probability expression to 0.25
A1: For $-\frac{3}{2}$

Another method for part (b) is:
M1: For understanding $2[k-x] = -1$ and $x = -1$
M1: For substitution and attempt to solve
A1: For $-\frac{3}{2}$

(c)
B1: For integrating x^3 f(x)
M1: For use of correct limits (dependent on previous M1)
A1*: For fully correct solution leading to the given answer with no errors seen