

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

©2017 Pearson Education Ltd.

1/1/1



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)

DO NOT WRITE IN THIS AREA

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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 $P(Y \geq -3) = 0.45$

(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)

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

DO NOT WRITE IN THIS AREA

Handwritten response area with horizontal lines.

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) $P(X \leq 4)$

(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

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

6. The continuous random variable X has probability density function

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

(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

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

(d) Show that $c = -10$ (2)

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

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

7. A scientist wants to develop a model to describe the relationship between the average daily temperature, $x^\circ\text{C}$, and a household's daily energy consumption, $y\text{kWh}$, in winter.

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

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

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

- (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

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Lined writing area consisting of 30 horizontal lines.

(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)

Lined writing area for student response.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 8 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 8 is 8 marks)

TOTAL FOR SECTION B IS 40 MARKS
TOTAL FOR PAPER IS 80 MARKS