Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number	Candidate Number

Statistics S1

Advanced/Advanced Subsidiary

Wednesday 13 June 2018 – Morning
Time: 1 hour 30 minutes

Paper Reference
6683/01

You must have:
Mathematical Formulae and Statistical Tables (Pink)

Total Marks

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra 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). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- 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.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- 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 ▶

P51574A

1/1/1/1/

P 5 1 5 7 4 A 0 1 2 8

Pearson

Leave blank

1. The discrete random variable X has the following probability distribution

x	2	4	7	10
$P(X = x)$	a	b	0.1	c

where a, b and c are probabilities.

The cumulative distribution function of X is $F(x)$ and $F(3) = 0.2$ and $F(6) = 0.8$

(a) Find the value of a, the value of b and the value of c.

(3)

(b) Write down the value of $F(7)$.

(1)

Leave blank

Question 1 continued

Q1

(Total 4 marks)

Leave blank

2. The following grouped frequency distribution summarises the number of minutes, to the nearest minute, that a random sample of 100 motorists were delayed by roadworks on a stretch of motorway one Monday.

Delay (minutes)	Number of motorists (f)	Delay midpoint (x)
3–6	38	4.5
7–8	25	7.5
9–10	18	9.5
11–15	12	13
16–20	7	18

(You may use $\sum fx^2 = 8096.25$)

A histogram has been drawn to represent these data.

The bar representing a delay of (3–6) minutes has a width of 2 cm and a height of 9.5 cm.

(a) Calculate the width and the height of the bar representing a delay of (11–15) minutes.

(3)

(b) Use linear interpolation to estimate the median delay.

(2)

(c) Calculate an estimate of the mean delay.

(2)

(d) Calculate an estimate of the standard deviation of the delays.

(2)

One coefficient of skewness is given by $\dfrac{3(\text{mean} - \text{median})}{\text{standard deviation}}$

(e) Evaluate this coefficient for the above data, giving your answer to 2 significant figures.

(1)

On the following Friday, the coefficient of skewness for the delays on this stretch of motorway was – 0.22

(f) State, giving a reason, how the delays on this stretch of motorway on Friday are different from the delays on Monday.

(2)

Leave blank

Question 2 continued

Leave blank

Question 2 continued

Leave blank

Question 2 continued

(Total 12 marks)

Q2

Leave blank

3. The random variable Y has a normal distribution with mean μ and standard deviation σ
The $P(Y > 17) = 0.4$

Find

(a) $P(\mu < Y < 17)$

(1)

(b) $P(\mu - \sigma < Y < 17)$

(4)

Leave blank

Question 3 continued

Leave blank

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 3 continued

Q3

(Total 5 marks)

Leave blank

4. A bag contains 64 coloured beads. There are r red beads, y yellow beads and 1 green bead and $r + y + 1 = 64$

Two beads are selected at random, one at a time without replacement.

(a) Find the probability that the green bead is one of the beads selected.

(4)

The probability that both of the beads are red is $\frac{5}{84}$

(b) Show that r satisfies the equation $r^2 - r - 240 = 0$

(3)

(c) Hence show that the only possible value of r is 16

(2)

(d) Given that at least one of the beads is red, find the probability that they are both red.

(4)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued

Leave blank

Question 4 continued

Leave blank

Question 4 continued

(Total 13 marks)

Q4

Leave blank

5. The score when a spinner is spun is given by the discrete random variable X with the following probability distribution, where a and b are probabilities.

x	– 1	0	2	4	5
$P(X = x)$	b	a	a	a	b

(a) Explain why $E(X) = 2$

(1)

(b) Find a linear equation in a and b.

(1)

Given that $Var(X) = 7.1$

(c) find a second equation in a and b and simplify your answer.

(3)

(d) Solve your two equations to find the value of a and the value of b.

(3)

The discrete random variable $Y = 10 - 3X$

(e) Find (i) $E(Y)$

(ii) $Var(Y)$

(3)

The spinner is spun once.

(f) Find $P(Y > X)$.

(3)

Leave blank

Question 5 continued

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 5 continued

(Total 14 marks)

Q5

Leave blank

6. A group of climbers collected information about the height above sea level, h metres, and the air temperature, t °C, at the same time at 8 different points on the same mountain.

The data are summarised by

$$\sum h = 6370 \qquad \sum t = 61 \qquad \sum th = 31\,070 \qquad \sum t^2 = 693$$

(a) Show that $S_{th} = -17\,501.25$ and $S_{tt} = 227.875$

(3)

The product moment correlation coefficient for these data is –0.985

(b) State, giving a reason, whether or not this value supports the use of a regression equation to predict the air temperature at different heights on this mountain.

(1)

(c) Find the equation of the regression line of t on h, giving your answer in the form $t = a + bh$. Give the value of your coefficients to 3 significant figures.

(7)

(d) Give an interpretation of your value of a.

(1)

One of the climbers has just stopped for a short break before climbing the next 150 metres.

(e) Estimate the drop in temperature over this 150 metre climb.

(2)

Leave blank

Question 6 continued

Leave blank

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 6 continued

(Total 14 marks)

Q6

Leave blank

7. Farmer Adam grows potatoes. The weights of potatoes, in grams, grown by Adam are normally distributed with a mean of 140 g and a standard deviation of 40 g.

Adam cannot sell potatoes with a weight of less than 92 g.

(a) Find the percentage of potatoes that Adam grows but cannot sell.

(3)

The upper quartile of the weight of potatoes **sold** by Adam is q_3

(b) Find the probability that the weight of a randomly selected potato **grown** by Adam is more than q_3

(2)

(c) Find the lower quartile, q_1, of the weight of potatoes **sold** by Adam.

(5)

Betty selects a random sample of 3 potatoes **sold** by Adam.

(d) Find the probability that one weighs less than q_1, one weighs more than q_3 and one has a weight between q_1 and q_3

(3)

Leave blank

Question 7 continued

Leave blank

Question 7 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 7 continued

Leave blank

Question 7 continued

Q7

(Total 13 marks)

TOTAL FOR PAPER: 75 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA