Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	8	3	/	0	1	Signature	

Paper Reference(s)

6683/01

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Friday 18 January 2013 – Afternoon

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
Total	

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
Nil

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 or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
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 for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 20 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

This publication may be reproduced only in accordance with Pearson Education Ltd copyright policy.
©2013 Pearson Education Ltd.

Printer's Log. No.
P41805A
W850/R6683/57570 5/5/5/6/

P 4 1 8 0 5 A 0 1 2 0

Turn over

PEARSON

Leave blank

1. A teacher asked a random sample of 10 students to record the number of hours of television, t, they watched in the week before their mock exam. She then calculated their grade, g, in their mock exam. The results are summarised as follows.

$$\sum t = 258 \qquad \sum t^2 = 8702 \qquad \sum g = 63.6 \qquad S_{gg} = 7.864 \qquad \sum gt = 1550.2$$

(a) Find S_{tt} and S_{gt}

(3)

(b) Calculate, to 3 significant figures, the product moment correlation coefficient between t and g.

(2)

The teacher also recorded the number of hours of revision, v, these 10 students completed during the week before their mock exam. The correlation coefficient between t and v was -0.753

(c) Describe, giving a reason, the nature of the correlation you would expect to find between v and g.

(2)

Leave blank

Question 1 continued

Q1

(Total 7 marks)

Leave blank

2. The discrete random variable X can take only the values 1, 2 and 3. For these values the cumulative distribution function is defined by

$$F(x) = \frac{x^3 + k}{40} \qquad x = 1, 2, 3$$

(a) Show that $k = 13$

(2)

(b) Find the probability distribution of X.

(4)

Given that $\text{Var}(X) = \frac{259}{320}$

(c) find the exact value of $\text{Var}(4X - 5)$.

(2)

Leave blank

Question 2 continued

Q2

(Total 8 marks)

Leave blank

3. A biologist is comparing the intervals (m seconds) between the mating calls of a certain species of tree frog and the surrounding temperature (t °C). The following results were obtained.

t °C	8	13	14	15	15	20	25	30
m secs	6.5	4.5	6	5	4	3	2	1

(You may use $\sum tm = 469.5$, $S_{tt} = 354$, $S_{mm} = 25.5$)

(a) Show that $S_{tm} = -90.5$

(4)

(b) Find the equation of the regression line of m on t giving your answer in the form $m = a + bt$.

(4)

(c) Use your regression line to estimate the time interval between mating calls when the surrounding temperature is 10 °C.

(1)

(d) Comment on the reliability of this estimate, giving a reason for your answer.

(1)

Leave blank

Question 3 continued

Q3

(Total 10 marks)

Leave blank

4. The length of time, L hours, that a phone will work before it needs charging is normally distributed with a mean of 100 hours and a standard deviation of 15 hours.

(a) Find P($L > 127$).

(3)

(b) Find the value of d such that P($L < d$) = 0.10

(3)

Alice is about to go on a 6 hour journey.
Given that it is 127 hours since Alice last charged her phone,

(c) find the probability that her phone will not need charging before her journey is completed.

(4)

Leave blank

Question 4 continued

Q4

(Total 10 marks)

Leave blank

5. A survey of 100 households gave the following results for weekly income £y.

Income y (£)	Mid-point	Frequency f
$0 \leqslant y < 200$	100	12
$200 \leqslant y < 240$	220	28
$240 \leqslant y < 320$	280	22
$320 \leqslant y < 400$	360	18
$400 \leqslant y < 600$	500	12
$600 \leqslant y < 800$	700	8

(You may use $\sum fy^2 = 12\,452\,800$)

A histogram was drawn and the class $200 \leqslant y < 240$ was represented by a rectangle of width 2 cm and height 7 cm.

(a) Calculate the width and the height of the rectangle representing the class $320 \leqslant y < 400$

(3)

(b) Use linear interpolation to estimate the median weekly income to the nearest pound.

(2)

(c) Estimate the mean and the standard deviation of the weekly income for these data.

(4)

One measure of skewness is $\dfrac{3(\text{mean} - \text{median})}{\text{standard deviation}}$.

(d) Use this measure to calculate the skewness for these data and describe its value.

(2)

Katie suggests using the random variable X which has a normal distribution with mean 320 and standard deviation 150 to model the weekly income for these data.

(e) Find $P(240 < X < 400)$.

(2)

(f) With reference to your calculations in parts (d) and (e) and the data in the table, comment on Katie's suggestion.

(2)

Leave blank

Question 5 continued

Leave blank

Question 5 continued

Leave blank

Question 5 continued

Q5

(Total 15 marks)

Leave blank

6. A fair blue die has faces numbered 1, 1, 3, 3, 5 and 5. The random variable B represents the score when the blue die is rolled.

(a) Write down the probability distribution for B.

(2)

(b) State the name of this probability distribution.

(1)

(c) Write down the value of E(B).

(1)

A second die is red and the random variable R represents the score when the red die is rolled.

The probability distribution of R is

r	2	4	6
P($R = r$)	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

(d) Find E(R).

(2)

(e) Find Var(R).

(3)

Tom invites Avisha to play a game with these dice.

Tom spins a fair coin with one side labelled 2 and the other side labelled 5. When Avisha sees the number showing on the coin she then chooses one of the dice and rolls it. If the number showing on the die is greater than the number showing on the coin, Avisha wins, otherwise Tom wins.

Avisha chooses the die which gives her the best chance of winning each time Tom spins the coin.

(f) Find the probability that Avisha wins the game, stating clearly which die she should use in each case.

(4)

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Q6

(Total 13 marks)

Leave blank

7. Given that

$$P(A) = 0.35 , \quad P(B) = 0.45 \quad \text{and} \quad P(A \cap B) = 0.13$$

find

(a) $P(A \cup B)$

(2)

(b) $P(A' \mid B')$

(2)

The event C has $P(C) = 0.20$

The events A and C are mutually exclusive and the events B and C are independent.

(c) Find $P(B \cap C)$

(2)

(d) Draw a Venn diagram to illustrate the events A, B and C and the probabilities for each region.

(4)

(e) Find $P([B \cup C]')$

(2)

Leave blank

Question 7 continued

Leave blank

Question 7 continued

Q7

(Total 12 marks)

TOTAL FOR PAPER: 75 MARKS

END