

Write your name here

Surname

Other names

Pearson Edexcel
International
Advanced Level

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Statistics S1

Advanced/Advanced Subsidiary

Thursday 18 January 2018 – Afternoon

Time: 1 hour 30 minutes

Paper Reference

WST01/01

You must have:

Mathematical Formulae and Statistical Tables (Blue)

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 ►

P51426A

©2018 Pearson Education Ltd.

1/1/1/e2




Pearson

1. Two classes of students, class A and class B , sat a test.

Class A has 10 students. Class B has 15 students.

Each student achieved a score, x , on the test and their scores are summarised in the table below.

	n	$\sum x$	$\sum x^2$
Class A	10	770	59610
Class B	15	t	58035

The mean score for Class A is 77 and the mean score for Class B is 61

- (a) Find the value of t (1)

- (b) Calculate the variance of the test scores for each class. (3)

The highest score on the test was 95 and the lowest score was 45

These were each scored by students from the same class.

- (c) State, with a reason, which class you believe they were from. (1)

The two classes are combined into one group of 25 students.

- (d) (i) Find the mean test score for all 25 students.
(ii) Find the variance of the test scores for all 25 students. (4)

The teacher of class A later realises that he added up the test scores for his class incorrectly. Each student's test score in class A should be increased by 3

- (e) Without further calculations, state, with a reason, the effect this will have on
- (i) the variance of the test scores for class A
 - (ii) the mean test score for all 25 students
 - (iii) the variance of the test scores for all 25 students. (3)



Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

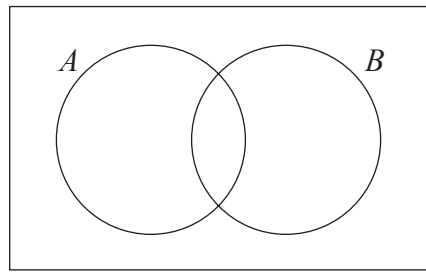
DO NOT WRITE IN THIS AREA

(Total 12 marks)

Q1



2. (a) Shade the region representing the event $A \cup B'$ on the Venn diagram below.



(1)

The two events C and D are mutually exclusive.

Given that $P(C) = \frac{1}{5}$ and $P(D) = \frac{3}{10}$

find

(b) (i) $P(C \cup D)$

(ii) $P(C|D)$

(3)

The two events F and G are independent.

Given that $P(F) = \frac{1}{6}$ and $P(F \cup G) = \frac{3}{8}$

find

(c) (i) $P(G)$

(ii) $P(F|G')$

(4)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total 8 marks)

Q2



Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



P 5 1 4 2 6 A 0 9 2 8

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Q3

(Total 8 marks)

Turn over



4. The discrete random variable X has probability distribution

x	-4	-3	1	2	5
$P(X=x)$	a	b	a	b	0.2

- (a) Find $E(X)$ in terms of a and b (1)

For this probability distribution, $\text{Var}(X) = E(X^2)$

- (b) (i) Write down the value of $E(X)$
(ii) Find the value of a and the value of b (5)

- (c) Find $\text{Var}(1 - 3X)$ (3)

Given that $Y = 1 - X$, find

- (d) (i) $P(Y < 0)$
(ii) the largest possible value of k such that $P(Y < k) = 0.2$ (4)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total 13 marks)

Q4



5. Franca is the manager of an accountancy firm. She is investigating the relationship between the salary, £ x , and the length of commute, y minutes, for employees at the firm. She collected this information from 9 randomly selected employees.

The salary of each employee was then coded using $w = \frac{x - 20\,000}{1000}$

The table shows the values of w and y for the 9 employees.

w	6	8	8	-1	25	15	3	-2	19
y	45	50	35	65	25	40	50	75	20

(You may use $\sum w = 81$ $\sum y = 405$ $\sum wy = 2490$ $S_{ww} = 660$ $S_{yy} = 2500$)

- (a) Calculate the salary of the employee with $w = -2$ (1)
- (b) Show that, to 3 significant figures, the value of the product moment correlation coefficient between w and y is -0.899 (3)
- (c) State, giving a reason, the value of the product moment correlation coefficient between x and y (1)

The least squares regression line of y on w is $y = 60.75 - 1.75w$

- (d) Find the equation of the least squares regression line of y on x giving your answer in the form $y = a + bx$ (3)
- (e) Estimate the length of commute for an employee with a salary of £21 000 (2)

Franca uses the regression line to estimate the length of commute for employees with salaries between £25 000 and £40 000

- (f) State, giving a reason, whether or not these estimates are reliable. (2)



Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



P 5 1 4 2 6 A 0 1 7 2 8

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Q5

(Total 12 marks)



- (e) Find $P(S = 5)$

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total 11 marks)

Q6



7. The weights, G , of a particular breed of gorilla are normally distributed with mean 180 kg and standard deviation 15 kg.

(a) Find the proportion of these gorillas whose weights exceed 174 kg. (2)

(b) Find, to 1 decimal place, the value of k such that $P(k < G < 174) = 0.3196$ (4)

The weights, B , of a particular breed of buffalo are normally distributed with mean 216 kg and standard deviation 30 kg.

Given that $P(G > w) = P(B < w) = p$

(c) (i) find the value of w
(ii) find the value of p (5)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Leave
blank

Question 7 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 7 continued

Handwriting practice area with 30 horizontal lines.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

--	--

TOTAL FOR PAPER: 75 MARKS

28

