Surname	Other na	mes
Pearson Edexcel International Advanced Level	Centre Number	Candidate Number
Statistics : Advanced/Advance	<del></del>	
Monday 26 June 2017 – Aft Time: 1 hour 30 minutes	ernoon	Paper Reference WST02/01

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 quide 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 ▶



P50716A
©2017 Pearson Education Ltd.



1.	At a particular junction on a train line, signal failures are known to occur randomly at a rate of 1 every 4 days.
	(a) Find the probability that there are no signal failures on a randomly selected day. (2)
	(b) Find the probability that there is at least 1 signal failure on each of the next 3 days. (2)
	(c) Find the probability that in a randomly selected 7-day week, there are exactly 5 days with no signal failures.
	Repair works are carried out on the line. After these repair works, the number, <i>f</i> , of signal failures in a 32-day period is recorded.
	A test is carried out, at the 5% level of significance, to determine whether or not there has been a decrease in the rate of signal failures following the repair works.
	(d) State the hypotheses for this test. (1)
	(e) Find the largest value of $f$ for which the null hypothesis should be rejected. (3)

Question 1 continued		Leave
		01
		Q1
	(Total 11 marks)	



2. <i>Crispy-crisps</i> produces packets of crisps. During a promotion, a prize is plat of the packets. No more than 1 prize is placed in any packet. A box contains 6 crisps.						
		(i) Write down a suitable distribution to model the number of prizes found in a box.				
		(ii) Write down one assumption required for the model. (2)				
	(b)	Find the probability that in 2 randomly selected boxes, only 1 box contains exactly 1				
		prize. (3)				
	(c)	Find the probability that a randomly selected box contains at least 2 prizes.  (2)				
	Nel	ha buys 80 boxes of crisps.				
	(d)	Using a normal approximation, find the probability that no more than 30 of the boxes contain at least 2 prizes.				
		(5)				

	Le
	bl
Question 2 continued	
	1



	Leave blank
Question 2 continued	

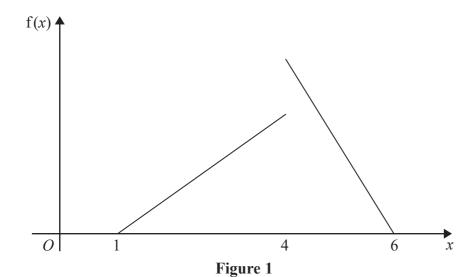
	blank
Question 2 continued	
	02
	Q2
(Total 12 marks)	
· · · · · · · · · · · · · · · · · · ·	



3. The random variable X has probability density function given by

$$f(x) = \begin{cases} ax + b & 1 \le x < 4 \\ \frac{3}{2} - \frac{1}{4}x & 4 \le x \le 6 \\ 0 & \text{otherwise} \end{cases}$$

as shown in Figure 1, where a and b are constants.



(a) Show that the median of *X* is 4

(2)

(b) Find the value of a and the value of b

**(5)** 

(c) Specify fully the cumulative distribution function of X

(5)

	Le   bla
uestion 3 continued	



Question 3 continued	

Question 3 continued		blank
		Q3
	otal 12 marks)	



**4.** In a large population, past records show that 1 in 200 adults has a particular allergy.

In a random sample of 700 adults selected from the population, estimate

- (a) (i) the mean number of adults with the allergy,
  - (ii) the standard deviation of the number of adults with the allergy. Give your answer to 3 decimal places.

**(3)** 

A doctor claims that the past records are out of date and the proportion of adults with the allergy is higher than the records indicate.

A random sample of 500 adults is taken from the population and 5 are found to have the allergy.

A test of the doctor's claim is to be carried out at the 5% level of significance.

- (b) (i) State the hypotheses for this test.
  - (ii) Using a suitable approximation, carry out the test.

**(6)** 

It is also claimed that 30% of those with the allergy take medication for it daily.

To test this claim, a random sample of n people with the allergy is taken. The random variable Y represents the number of people in the sample who take medication for the allergy daily.

A two-tailed test, at the 1% level of significance, is carried out to see if the proportion differs from 30%

The critical region for the test is Y = 0 or  $Y \ge w$ 

(c)	Find the	smallest	nossible	value	of $n$	and the	e corresr	onding	value	of u
(6)	rina me	Smanest	possible	varue	or $n$	and m	e corresp	onamg	varue	01 n

**(4)** 



	Leave blank
Question 4 continued	



Question 4 continued	blank
	Q4
	\ \frac{\fir}}}}}}}{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\fir}}}}}}}{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\fir}}}}}}}{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\fir}}}}}}}{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac{\frac}}}}}}}{\frac{\frac{\frac{\frac{\frac{\frac{\fir}}}}}}{\fra
(Total 13 marks)	



5. A call centre records the length of time, T minutes, its customers wait before being connected to an agent. The random variable T has a cumulative distribution function given by

$$F(t) = \begin{cases} 0 & t < 0 \\ 0.3t - 0.004t^3 & 0 \le t \le 5 \\ 1 & t > 5 \end{cases}$$

(a) Find the proportion of customers waiting more than 4 minutes to be connected to an agent.

**(2)** 

(b) Given that a customer waits more than 2 minutes to be connected to an agent, find the probability that the customer waits more than 4 minutes.

**(3)** 

(c) Show that the upper quartile lies between 2.7 and 2.8 minutes.

**(2)** 

(d) Find the mean length of time a customer waits to be connected to an agent.

**(4)** 



16

Question 5 continued	Lo bl



Question 5 continued	blank
	Q5
(Total 11 marks)	
(Iotal II marks)	



6.	At a men's tennis tournament there are 3, 4 or 5 sets in a match.
	Over many years, data collected show that 50% of matches last for exactly 3 sets, 30% of matches last for exactly 4 sets and 20% of matches last for exactly 5 sets.
	A random sample of 3 tennis matches is taken. The number of sets in each match is recorded as $S_1$ , $S_2$ and $S_3$ respectively.
	The random variable $M$ represents the maximum value of $S_1$ , $S_2$ and $S_3$
	(a) List all the samples where $M \neq 5$
	(b) Find the sampling distribution of $M$ (4)
	(c) Write down the mode of $S_1$ and the mode of $M$
	(1)



Question 6 continued	Leave blank	
	<b>Q6</b>	
(Total 7 marks)		



- 7. The continuous random variable X is uniformly distributed over the interval [a, b]
  - (a) Find an expression, in terms of a and b, for E(3-2X)

**(2)** 

(b) Find  $P(X > \frac{1}{3}b + \frac{2}{3}a)$ 

**(2)** 

Given that E(X) = 0

(c) find an expression, in terms of b only, for  $E(3X^2)$ 

**(3)** 

Given also that the range of X is 18

(d) find Var(X)

**(2)** 

estion 7 continued	



euestion 7 continued	bla
	Q7
(Total 9 marks)	
TOTAL FOR PAPER: 75 MARKS	
END	