6687/01

Edexcel GCE

Statistics

Unit S5

Advanced Subsidiary / Advanced

Time: 1 hour 30 minutes

Materials required for the examination	Items included with these question papers
Answer Book (AB04) Graph Paper (GP02) Mathematical Formulae	Nil

Candidates may use any calculator EXCEPT **those with a facility for symbolic algebra, differentiation and/or integration. Thus candidates may** NOT **use calculators such as Texas TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.**

Instructions to Candidates

Full marks may be obtained for answers to ALL questions.

In the boxes on the Answer Book provided, write the name of the Examining Body (Edexcel), your Centre Number, Candidate Number, the Unit Title (Statistics S5), the Paper Reference (6687), your surname, other names and signature.

Information for Candidates

A booklet 'Mathematical Formulae including Statistical Formulae and Tables' is provided.

Values from the Statistical Tables should be quoted in full. The answer to each part of a question which requires the use of tables or a calculator should be given to three significant figures, unless otherwise specified.

This paper has 7 questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly numbered.

You must show sufficient working to make your methods clear to the Examiner. Answers without working will gain no credit

N1234 R6081

Edexcel
Success through qualifications

1. Customers arrive at a Post Office queue, at random, at a rate of 3.4 customers per 5 minute period.

(*a*) Calculate the probability that no customer arrives in the next minute. **(2 marks)**

(*b*) Calculate the probability that the time interval between the arrivals of the first and second customers is less than 1 minute. **(2 marks)**

(*c*) Calculate the probability that the time interval between the arrivals of the second and third customers is between 1 and 2 minutes. **(3 marks)**

---

2. A component part for motor bikes is manufactured at three different factories A, B and C and then delivered to the main assembly line. Factory A supplies 48% of the total number of this component, factory B supplies 36% and factory C 16%. Of those components manufactured at factory A, 3% are faulty. The corresponding figures for B and C are 2% and 4% respectively.

(*a*) Find the probability that a component selected at random from the main assembly line is faulty. **(2 marks)**

(*b*) Given that a component is found to be faulty, find the probability that it was manufactured at factory C. **(2 marks)**

A large box of components has been delivered from one of the factories. Four components are selected at random from the box and none of them are faulty.

(*c*) Find the probability that the box came from factory B. **(4 marks)**

---

3. A bored shop assistant is playing with a large tin of coloured drawing pins. It is known that 15% of the pins are coloured red. The assistant selects a pin at random, notes the colour and returns it to the tin. She repeats this process until a red pin is selected. The random variable X represents the number of pins selected.

(*a*) State a distribution that could be used to model X. **(1 mark)**

(*b*) Find the probability that the 4th pin selected is the first one which is red. **(2 marks)**

(*c*) Find the probability that the number of pins selected is fewer than 9. **(3 marks)**

The shop assistant decides to estimate the proportion of yellow pins in the tin. She selects a pin at random, notes the colour and then returns it, repeating this process until a yellow one is selected. The random variable Y represents the number of pins selected. She repeats this experiment 50 times and finds that on 7 occasions the number of pins selected was more than 8.

(*d*) Use this information to estimate the proportion of yellow pins in the tin. **(4 marks)**

---

4. Three balls a, b and c are placed at random in boxes A, B and C so that each box contains one ball. The random variable X is defined by

$$X = 1, \quad \text{if ball } a \text{ is in box } A,$$
$$X = 0, \quad \text{otherwise.}$$

(*a*) Show that the probability generating function of X is given by $\mathrm{G}_X(t) = \dfrac{t+2}{3}$. **(2 marks)**

A second random variable Y has probability generating function $\mathrm{G}_Y(t) = \dfrac{2t+3}{5}$.

Given that X and Y are independent,

(*b*) show that the probability generating function of $Z = X + Y$ is given by

$$\mathrm{G}_Z(t) = \frac{2t^2 + 7t + 6}{15}.$$ **(2 marks)**

(*c*) Find the mean and variance of Z. **(6 marks)**

---

5 . Ahmed is playing a board game with a single fair die. In the final stages of the game he has to obtain two sixes in order to finish. The random variable X represents the number of rolls of the die needed for Ahmed to finish the game.

(*a*) Find $\mathrm{P}(X = 5)$. **(2 marks)**

(*b*) Find $\mathrm{P}(X \leq 15 \mid \text{the first six was obtained on the 10th throw})$. **(4 marks)**

(*c*) Write down $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **(2 marks)**

Ahmed plays this game 80 times and records the value of the random variable X on each occasion.

(*d*) Estimate the probability that the sample mean $\overline{X}$ is more than 14. **(4 marks)**

---

6. The continuous random variable Y has an exponential distribution with mean $\frac{1}{3}$.

(*a*) Write down the moment generating function for Y. **(2 marks)**

The continuous random variable X has probability density function, f(x), given by

$$\mathrm{f}(x) = \begin{cases} 9x\mathrm{e}^{-3x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(*b*) Show that the moment generating function of X is $\dfrac{9}{(3-t)^2}$.

[You may assume that for $t < 3$ $\lim\limits_{x \to \infty} \left(x\mathrm{e}^{-(3-t)x}\right) = 0$.] **(6 marks)**

(*c*) Explain the connection between the distributions of X and Y. **(2 marks)**

(*d*) Hence, or otherwise, find the mean and variance of X. **(4 marks)**

---

7. A quality control manager wishes to devise acceptance criteria for a large batch of components delivered to a factory. He first considers using scheme A as detailed below.

> **Scheme *A*.** Examine a random sample of 20 components and accept the batch if the number of defective components is no more than 3, otherwise reject the whole batch.

(*a*) Find the probability of accepting a batch where the proportion of defectives is:
(i) 5%,
(ii) 15%. **(3 marks)**

The manager requires the probability of accepting a batch where 15% of the components are defective to be below 0.5. At the same time he requires the probability of accepting a batch with 5% of defective components to be greater than 0.95. The manager considers scheme B.

> **Scheme *B*.** Examine a random sample of 25 components and accept the batch if the number of defective components is no more than 3, otherwise reject the whole batch.

The manager finds that the probability of accepting a batch with 5% of defective components is 0.9659.

(*b*) Find the probability of accepting a batch containing 15% of defective components using scheme B. **(3 marks)**

A statistical colleague suggests that the manager considers scheme C.

> **Scheme C.** First examine a random sample of 20 components and accept the batch if the number of defectives is less than 3, and reject the batch if the number of defectives is greater than 3. If exactly 3 defectives are found then take a further sample of 10 components. Accept the batch if there are no defectives found otherwise reject the whole batch.

(*c*) Show that scheme C satisfies the manager's requirements. **(4 marks)**

The manager wishes to use the scheme that satisfies his requirements and, over the long term, involves examining the smallest number of components.

(*d*) Determine, giving your reasons, which scheme the manager should use. **(4 marks)**

END

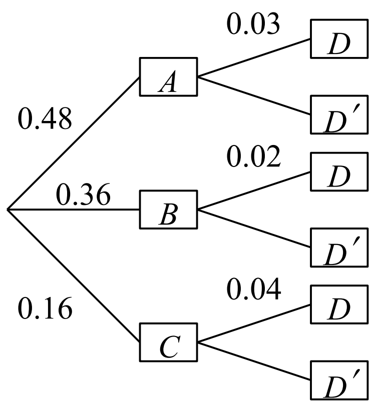
EDEXCEL FOUNDATION

Stewart House 32 Russell Square London WC1B 5DN
Telephone 0171-393 4444 Fax 0171-753 4558

Specimen Paper

Advanced Subsidiary/Advanced Level
General Certificate of Education

Subject **STATISTICS** Paper No. **S5**

Question number	Scheme	Marks	
1.	X = number of customers arriving in 1 minute $\quad X \sim \text{Po}(0.68)$		
(a)	$\text{P}(X = 0) = \text{e}^{-0.68} = \underline{0.5066}\ldots$	M1A1	(2)
(b)	T = time taken between arrivals $\quad T \sim \text{Exp}(0.68)$		
	$\text{P}(T < 1) = \int_0^1 0.68\text{e}^{-0.68t}\,\text{d}t$	M1	
	$= \left[-\text{e}^{-0.68t}\right]_0^1 = 1 - \text{e}^{-0.68} = 0.49338\ldots \quad (\text{AWRT } 0.493)$	A1	(2)
(c)	$\text{P}(1 < T < 2) = \left[-\text{e}^{-0.68t}\right]_1^2, \quad = -\text{e}^{-1.36} + \text{e}^{-0.68}$	M1, M1	
	$= 0.24995\ldots \quad \text{AWRT } \underline{0.250}$	A1	(3)
2.	 Let D = a defective component		
(a)	$\text{P}(D) = 0.48 \times 0.03 + 0.36 \times 0.02 + 0.16 \times 0.04$	M1	
	$= 0.028$ or $\underline{\dfrac{7}{250}}$	A1	(2)

Turn Over

(b)	$P(C \mid D) = \frac{P(C \cap D)}{P(D)} = \frac{0.16 \times 0.04}{0.028}$	M1	
	$= \underline{0.229}$ (AWRT)	A1	(2)
(c)	Let G = all components are free of faults		
	$P(G) = 0.48 \times (0.97)^4 + 0.36 \times (0.98)^4 + 0.16 \times (0.94)^4$	M1	
	$= 0.89288 \ldots$ AWRT $\underline{0.893}$	A1	
	$P(B \mid G) = \frac{P(B \cap G)}{P(G)} = \frac{0.36 \times (0.98)^4}{0.89288\ldots} = 0.371885\ldots$	M1	
	AWRT $\underline{0.372}$	A1	(4)
3. (a)	Geometric distribution (with parameter 0.15) or G(0.15)	B1	(1)
(b)	$P(X = 4) = 0.15 \times (0.85)^3, = 0.09211\ldots$ AWRT $\underline{0.0921}$	M1, A1	(2)
(c)	$P(X \leq 8) = 1 - P(X \geq 9), = 1 - (0.85)^8, = 0.727509\ldots$ AWRT $\underline{0.728}$	M1, A1, A1	(3)
(d)	$Y \sim G(p)$ Estimate of $P(Y > 8) = \frac{7}{50}$	M1	
	i.e. $(1 - \hat{p})^8 = \frac{7}{50} \Rightarrow 1 - \hat{p} = \sqrt[8]{\frac{7}{50}}$	M1 M1	
	i.e. $\hat{p} = 1 - 0.7821\ldots$ i.e. $\hat{p}$ = AWRT $\underline{0.218}$	A1	(4)

Question number	Scheme	Marks	
4. (a)	$G_X(t) = P(X=1).t^1 + P(X=0).t^\circ, = \frac{2}{6}t + \frac{4}{6}t^\circ = \frac{t+2}{3}$	M1A1	(2)
(b)	$G_Z(t) = G_X(t) \times G_Y(t)$	M1	
	$= \left(\frac{t+2}{3}\right) \times \frac{(2t+3)}{5}$		
	$= \frac{2t^2+7t+6}{15}$	A1	(2)
(c)	$[E(Z) = E(X) + E(Y)$ and $Var(Z) = Var(X) + Var(Y)$ acceptable$]$		
	$G'(t) = \frac{1}{15}(4t+7);\ E(Z) = G'(1) = \frac{11}{15}$	M1A1 √	
	$G''(t) = \frac{1}{15}(4)$	M1A1	
	$Var(Z) = G''(1) + G'(1) - [G'(1)]^2, = \frac{4}{15} + \frac{11}{15} - \left(\frac{11}{15}\right)^2$ = AWRT $\underline{0.462}$	M1A1	(6)
5. (a)	$P(X=5) = \binom{4}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right), = 0.06430...$ AWRT $\underline{0.0643}$	M1,A1	(2)
(b)	$P(X \le 15 \mid \text{1st 6 on 10th throw})$ Let $Y \sim G\left(\frac{1}{6}\right)$	M1 (Use of geom)	
	$= P(Y \le 5) = 1 - P(Y \ge 6) = 1 - \left(\frac{5}{6}\right)^5 = 1 - 0.4018...$	M1A1	
	= AWRT $\underline{0.598}$	A1	(4)
(c)	$E(X) = \frac{2}{\frac{1}{6}} = \underline{12}$ $\quad Var(X) = \frac{2\left(\frac{5}{6}\right)}{\left(\frac{1}{6}\right)^2} = \underline{60}$	B1B1	(2)
(d)	By CLT $\bar{X} \approx\sim N\left(12, \frac{60}{80}\right)$	M1	
	$P(\bar{X} > 14) \approx P\left(Z > \frac{14-12}{\sqrt{\frac{3}{4}}}\right) = P(Z > 2.309...)$	M1A1	
	= AWRT $\underline{0.010 \sim 0.011}$	A1	(4)

Question number	Scheme	Marks
6. **(a)**	$Y \sim \text{Exp}(3)$ $\quad\quad \mathrm{M}_Y(t) = \underline{\dfrac{3}{3-t}}$	M1A1 (2)
(b)	$\mathrm{M}_X(t) = \int_0^\infty \mathrm{e}^{tx} 9x\mathrm{e}^{-3x}\mathrm{d}x$	M1
	$= 9\int_0^\infty x\mathrm{e}^{-(3-t)x}\mathrm{d}x \quad\quad = -9\int_0^\infty x\mathrm{d}\left(\dfrac{\mathrm{e}^{-(3-t)x}}{3-t}\right)$	M1
	$= \dfrac{-9}{(3-t)}\left[\left[x\mathrm{e}^{-(3-t)x}\right]_0^\infty - \int_0^\infty \mathrm{e}^{-(3-t)x}\mathrm{d}x\right]$	M1A1
	$= \dfrac{9}{(3-t)}\left[-\dfrac{\mathrm{e}^{-(3-t)x}}{(3-t)}\right]_0^\infty$	M1
	$= \dfrac{9}{(3-t)}\left[0 - -\dfrac{1}{3-t}\right] = \underline{\dfrac{9}{(3-t)^2}}$	A1 (6)
(c)	$\mathrm{M}_Z(t) = \{\mathrm{M}_Y(t)\}^2$ suggests $\quad X = Y_1 + Y_2$	M1A1 (2)
(d)	$\mathrm{E}(Y) = \dfrac{1}{3} \Rightarrow \mathrm{E}(X) = \underline{\dfrac{2}{3}}$	B1
	$\mathrm{Var}(Y) = \dfrac{1}{9} \Rightarrow, \mathrm{Var}(X) = \underline{\dfrac{2}{9}}$	B1M1A1√ (4)

Question number	Scheme	Marks
7. (a)	X = number of defectives $\qquad X \sim \mathrm{B}(20, p)$	M1
	$\mathrm{P}(X \le 3 \mid p = 0.05) = \underline{0.9841} \qquad \mathrm{P}(X \le 3 \mid p = 0.15) = \underline{0.6477}$	A1A1 (3)
(b)	$Y \sim \mathrm{B}(25, 0.15)$	
	$\mathrm{P}(Y \le 3) = 0.85^{25} + 25(0.85)^{24}\,0.15 + \dfrac{25\times24}{2}(0.85)^{23}.\,(0.15)^2 + \dfrac{25\times24\times23}{3!}(0.85)^{22}\times(0.15)^3$	M1A1
	$= 0.47112\ldots \qquad$ AWRT $\underline{0.471}$	A1 (3)
(c)	P(Accept) $= \mathrm{P}(X \le 2) + \mathrm{P}(X = 3) \times \mathrm{P}(\mathrm{B}(10, p) = 0)$	M1
	p=0.05 $\quad = 0.9245 + (0.9841 - 0.9245) \times 0.5987$	
	$= 0.96018\ldots \qquad$ AWRT $\underline{0.960}$	A1
	p=0.15 $\quad = 0.4049 + (0.6477 - 0.4049) \times 0.1969$ (	M1
	$= 0.45270\ldots \qquad$ AWRT $\underline{0.453}$	A1 (4)
(d)	Scheme B examines 25	B1
	Scheme C $\quad$ 20 $\quad$ 30	
	Prob $\quad \mathrm{P}(X \ne 3) \quad \mathrm{P}(X = 3) \quad \Rightarrow$ Expected number $=20.596$	M1
	p=0.05 $\quad$ (0.0596) $\quad$ OR	
	p=0.15 $\quad$ (0.2428) = 22.428	A1
	$\therefore$ $\underline{\text{choose } C}$	A1 (4)