Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Bronze Level B1

Time: 1 hour 30 minutes

Materials required for examination papers	**Items included with question**
Mathematical Formulae (Green)	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, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

Write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, initials and signature.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
There are 6 questions in this question paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

Suggested grade boundaries for this paper:

A*	A	B	C	D	E
75	70	64	58	50	42

1. A manufacturer supplies DVD players to retailers in batches of 20. It has 5% of the players returned because they are faulty.

(*a*) Write down a suitable model for the distribution of the number of faulty DVD players in a batch.

(2)

Find the probability that a batch contains

(*b*) no faulty DVD players,

(2)

(*c*) more than 4 faulty DVD players.

(2)

(*d*) Find the mean and variance of the number of faulty DVD players in a batch.

(2)

---

2. In a village, power cuts occur randomly at a rate of 3 per year.

(*a*) Find the probability that in any given year there will be

(i) exactly 7 power cuts,

(ii) at least 4 power cuts.

(5)

(*b*) Use a suitable approximation to find the probability that in the next 10 years the number of power cuts will be less than 20.

(6)

---

3. A website receives hits at a rate of 300 per hour.

(*a*) State a distribution that is suitable to model the number of hits obtained during a 1 minute interval.

(1)

(*b*) State two reasons for your answer to part (*a*).

(2)

Find the probability of

(*c*) 10 hits in a given minute,

(3)

(*d*) at least 15 hits in 2 minutes.

(3)

The website will go down if there are more than 70 hits in 10 minutes.

(*e*) Using a suitable approximation, find the probability that the website will go down in a particular 10 minute interval.

(7)

---

4. A café serves breakfast every morning. Customers arrive for breakfast at random at a rate of 1 every 6 minutes.

Find the probability that

(*a*) fewer than 9 customers arrive for breakfast on a Monday morning between 10 a.m. and 11 a.m.

(3)

The café serves breakfast every day between 8 a.m. and 12 noon.

(*b*) Using a suitable approximation, estimate the probability that more than 50 customers arrive for breakfast next Tuesday.

(6)

---

5. Cars arrive at a motorway toll booth at an average rate of 150 per hour.

(*a*) Suggest a suitable distribution to model the number of cars arriving at the toll booth, X, per minute.

(2)

(*b*) State clearly any assumptions you have made by suggesting this model.

(2)

Using your model,

(*c*) find the probability that in any given minute

(i) no cars arrive,

(ii) more than 3 cars arrive.

(3)

(*d*) In any given 4 minute period, find m such that $P(X > m) = 0.0487$

(3)

(*e*) Using a suitable approximation find the probability that fewer than 15 cars arrive in any given 10 minute period.

(6)

---

6. The continuous random variable X has the following probability density function

$$f(x) = \begin{cases} a + bx, & 0 \le x \le 5, \\ \\ 0, & \text{otherwise.} \end{cases}$$

where a and b are constants.

(*a*) Show that $10a + 25b = 2$.

(4)

Given that $E(X) = \dfrac{35}{12}$,

(*b*) find a second equation in a and b,

(3)

(*c*) hence find the value of a and the value of b.

(3)

(*d*) Find, to 3 significant figures, the median of X.

(3)

(*e*) Comment on the skewness. Give a reason for your answer.

(2)

TOTAL FOR PAPER: 75 MARKS

END

Question Number	Scheme	Marks
Q1 (a)	$X \sim B(20, 0.05)$	B1 B1 (2)
(b)	P(X = 0) = 0.95^{20} = 0.3584859… or 0.3585 using tables .	M1 A1 (2)
(c)	$\mathrm{P}(X > 4) \quad = 1 - \mathrm{P}(X \le 4)$ $= 1 - 0.9974$ $= 0.0026$	M1 A1 (2)
(d)	Mean $= 20 \times 0.05 = 1$ Variance $= 20 \times 0.05 \times 0.95 = 0.95$	B1 B1 (2) **Total [8]**

Question Number	Scheme	Marks
2 (a)	Let X be the random variable the number power cuts.	
	$X \sim \text{Po}(3)$	B1
(i)	$\text{P}(X = 7) = \text{P}(X \leq 7) - \text{P}(X \leq 6)$ or $\dfrac{e^{-3}3^7}{7!}$	M1
	$= 0.9881 - 0.9665$	
	$= 0.0216$ awrt 0.0216	A1
(ii)	$\text{P}(X \geq 4) = 1 - \text{P}(X \leq 3)$	M1
	$= 1 - 0.6472$	
	$= 0.3528$ awrt 0.353	A1
		(5)
(b)	$X \sim \text{Po}(30)$	
	N(30,30)	M1A1
	$\text{P}(X < 20) = \text{P}\left(Z < \dfrac{19.5 - 30}{\sqrt{30}}\right)$	M1M1 A1
	$= \text{P}(Z < -1.92)$	
	$= 1 - 0.9726$	
	$= 0.0274 - 0.0276$	A1
		(6)
		Total 11

Question Number	Scheme	Marks
3. (a)	Poisson	B1 (1)
(b)	Hits occur **singly** in time	
	Hits are **independent** **<u>or</u>** Hits occur **randomly**	B1 B1
	Hits occur at a **constant rate**	(2)
(c)	$X \sim \text{Po}(5)$	B1
	$P(X=10) = \text{P}(X \le 10) - \text{P}(X \le 9)$ or $\dfrac{e^{-5}5^{10}}{10!}$	M1
	$= 0.9863 - 0.9682$	
	$= 0.0181$ awrt 0.0181	A1 (3)
(d)	$X \sim \text{Po}(10)$	B1
	$\text{P}(X \ge 15) = 1 - \text{P}(X \le 14)$	M1
	$= 1 - 0.9165$	
	$= 0.0835$ awrt 0 .0835	A1 (3)
(e)	$X \sim \text{Po}(50)$	
	Approximated by N(50,50)	B1B1
	$\text{P}(X > 70) = \text{P}\left(Z > \dfrac{70.5 - 50}{\sqrt{50}}\right)$	M1M1
	$= \text{P}(Z > 2.899...)$	A1
	$= 1 - 0.9981$	M1
	$= 0.0019$ awrt 0.0019	A1 (7)
		(16 marks)

Question Number	Scheme	Marks
Q4 (a)	$X \sim \text{Po}(10)$ $\text{P}(X < 9) \quad = \text{P}(X \le 8)$ $= 0.3328$	B1 M1 A1 (3)
(b)	$Y \sim \text{Po}(40)$ Y is approximately N(40,40) $\text{P}(Y > 50) \quad = 1 - \text{P}(Y \le 50)$ $= 1 - \text{P}\left(Z < \frac{50.5 - 40}{\sqrt{40}}\right)$ $= 1 - \text{P}(Z < 1.660..)$ $= 1 - 0.9515$ $= 0.0485$ N.B. Calculator gives 0.048437. Poisson gives 0.0526 (but scores nothing)	M1 A1 M1 M1 A1 A1 (6) **Total [9]**
5. **(a)**	X~Po(2.5)	**M1A1** **(2)**
(b)	Cars arrive at the toll booth independently/randomly Cars arrive one at a time B1 The rate of arrival at a toll booth remains constant at 2.5 per minute	**B1** **B1** **(2)**
(c) (i)	$\text{P}(X = 0) = e^{-2.5} = 0.0821$	**B1** **(1)**
(c) (ii)	$\text{P}(X > 3) \quad = 1 - \text{P}(X \le 3)$ $= 0.2424$	**M1** **A1** **(2)**
(d)	Use of Po(10) $1 - 0.0487 = 0.9513$ $m = 15$	**M1** **M1** **A1 cao** **(3)**
(e)	$Y \sim \text{N}(25, 25)$ $\text{P}(X < 15) = \text{P}(Y \le 14.5)$ $= \text{P}\left(z \frac{14.5 - 25}{5}\right)$ $= \text{P}(Z \le -2.1)$ $= 0.01786$	**B1 B1** **M1** **M1** **A1** **A1** **(6)** **[16]**

Question Number	Scheme	Marks
6 (a)	$\int_0^5 a + bx \, dx = 1$	M1
	$\left[ax + \frac{bx^2}{2}\right]_0^5 = 1$	A1
	$5a + \frac{25b}{2} = 1$	M1dep
	$10a + 25b = 2$	A1cso
		(4)
(b)	$\int_0^5 ax + bx^2 \, dx = \frac{35}{12}$	M1
	$\left[\frac{ax^2}{2} + \frac{bx^3}{3}\right]_0^5 = \frac{35}{12}$	A1
	$\frac{25a}{2} + \frac{125b}{3} = \frac{35}{12}$	A1
	$30a + 100b = 7$	**(3)**
(c)	$30a + 100b = 7$	M1
	$10a + 25b = 2$	
	$a = 0.1 \quad b = 0.04$	A1,A1
		(3)
(d)	$\int_0^m 0.1 + 0.04x \, dx = 0.5$	M1
	$\left[0.1x + \frac{0.04x^2}{2}\right]_0^m = 0.5$	A1ft
	$0.1m + 0.02m^2 - 0.5 = 0$	
	$m = \frac{-0.1 \pm \sqrt{0.1^2 + 4 \times 0.02 \times 0.5}}{2 \times 0.02}$	
	m = 3.09, -8.09 therefore 3.09	A1
		(3)
(e)	mean < median (< mode)	B1ft
	negatively skewed	B1 dep ft
		(2)
		Total 15

Examiner reports

Question 1

This question proved to be a very good start to the paper for a large majority of candidates. In general parts (a), (b) and (d) were answered correctly. In part (c) the most common mistake was to use $P(X > 4) = 1 - P(X \leq 3)$.

Question 2

Overall this question was well answered and responses reflected good preparation and understanding in using a Poisson distribution and also using a normal approximation to a Poisson. A high percentage of candidates attempted both parts of (a) successfully, with the majority of candidates using Po(3) and getting at least one mark for (a)(i). Marks lost for part (a) were normally for finding $P(X \geq 4) = 1 - P(X \leq 4)$ or $P(X = 4)$, in (a)(ii) or occasionally for writing the answer to (a)(i) as 0.22. In part (b) candidates showed their ability to standardise correctly using a continuity correction to get a negative 'z' value, or in the case of candidates who used the symmetric properties of the distribution, the equivalent positive value. A minority of candidates lost marks through either using an incorrect continuity correction, i.e. 18.5 or 20.5, or none at all. Occasionally a candidate failed to find $1 - \Phi(1.92)$ although it was rare to see a final answer > 0.5.

Question 3

This question was well answered, with a high proportion of candidates gaining full marks. The choice of Poisson was recognised easily and, in most cases, justified with appropriate properties. One mark was lost frequently by candidates who did not write the answer in the context of 'hits'.

A large proportion of scripts showed excellent solutions to both parts (c) and (d). Errors were rare, but perhaps the most common was to write a final answer correct to two significant figures only. The inequalities were handled well by almost all candidates, although there were a tiny number who wrote, for example, $P(X = 10) = P(X \leq 11) - P(X \leq 10)$ in part (c) and $P(X \geq 15) = 1 - P(X \leq 15)$ in part (d).

A pleasing majority of candidates obtained full marks to part (e), with clear, confident and accurate responses. Most of these candidates achieved the perfect combination of a fully detailed method together with economy of expression.

Marks lost in this part were mainly due to using a value 69.5 instead of 70.5 in the standardisation or no continuity correction at all. Few candidates made no attempt at this part of the question.

Question 4

This question was accessible to the majority of candidates, with many gaining full marks. Most recognised the need to use a Poisson distribution in part (a) and translated the time of one hour successfully to a mean of 10. Common errors included using a mean of 6 or misinterpreting $P(X < 9)$ as $P(X \leq 9)$ or using $1 - P(X \leq 8)$. In part (b), a high percentage of candidates gained full marks for using a Normal approximation with correct working. Marks lost in this part were mainly due to using a 49.5 instead of 50.5 or no continuity correction at all. A small number of candidates wrote the distribution as B(240, 1/6) and translated this to N(40, 100/3).

Question 5

This question was answered well by a high proportion of candidates reflecting a good understanding of the Poisson distribution and also the use of the normal approximation to Poisson with many gaining full marks. In part (a) the main error was using Po(150).

In part (b) a minority of candidates failed to use a context when stating the conditions for any Poisson distribution or, if in context, failed to use words that implied “cars arrive” or “rate of arrival”. For part (c)(i) the most common error seen was a rounded answer of 0.082.

When finding the probability in part (c)(ii), a small minority of candidates calculated $P(X > 3) = 1 - P(X \leq 2)$ or found $P(X \leq 3)$.

In part (d), the most common error was for candidates to write $P(X > m) = 1 - P(X \leq m - 1)$ and, having successfully shown that $P(X \leq 15 \mid X \sim Po(10)) = 0.93150$, then write $m - 1 = 15$ so $m = 16$. The majority of candidates used a normal approximation successfully in part (e) and gained full marks.

Question 6

This question provided many difficulties for weaker candidates, particularly those weak at algebraic manipulation.

In part(a) nearly all candidates attempted to integrate the given expression with most placing the resultant expression equal to 1, substituting 5 and multiplying by 2 to provide the given equation.

Most candidates attempted part (b) successfully and achieved a correct equation. Many candidates then tried to rearrange the equation but often they did not posses the algebraic skills to do so accurately. Marks were awarded for the correct equation in this part as incorrect subsequent working was ignored. The candidates then unfortunately used their rearranged incorrect equation in part (c). All the candidates attempted to solve ‘their’ equations simultaneously and most of those with the correct equation attained the correct values for a and b. Part (d) of the question was not answered well by a large number of candidates. Those that had the correct values of a and b usually managed to integrate correctly but errors in substitution into the quadratic formula and/or in calculating the two possible values of m (the median) were not uncommon. In part (e) The majority of candidates compared ‘their’ median with the given mean correctly and most of them also stated the correct skew for ‘their’ values. Some candidates attempted to compare the mean or median (or both) with the mode with varying degrees of success. This comparison was only considered if ‘their’ mode was stated and, when given, these values varied anywhere between 0 and 5.

Statistics for S2 Practice Paper Bronze 1

				Mean average scored by candidates achieving grade:							
Qu	**Max Score**	**Modal score**	**Mean %**	**ALL**	**A***	**A**	**B**	**C**	**D**	**E**	**U**
1	8		94.5	7.56		7.76	7.48	7.24	6.71	6.23	4.08
2	11	11	90.0	9.89	10.50	10.54	9.66	9.00	7.73	6.48	3.33
3	16		88.1	14.10	15.30	14.82	13.79	12.09	10.41	9.42	3.38
4	9		85.3	7.68		8.22	7.66	6.12	4.94	3.49	1.64
5	16		86.1	13.78	15.28	14.71	13.56	12.69	9.70	7.44	4.87
6	15	15	82.0	12.31	14.16	13.89	12.36	10.65	8.18	6.97	3.30
	75		**87.1**	**65.32**		**69.94**	**64.51**	**57.79**	**47.67**	**40.03**	**20.60**