Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Bronze Level B2

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 7 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
74	70	63	56	49	41

1. (*a*) Write down the conditions under which the Poisson distribution can be used as an approximation to the binomial distribution.

(2)

The probability of any one letter being delivered to the wrong house is 0.01.
On a randomly selected day Peter delivers 1000 letters.

(*b*) Using a Poisson approximation, find the probability that Peter delivers at least 4 letters to the wrong house.

Give your answer to 4 decimal places.

(3)

---

2. The number of defects per metre in a roll of cloth has a Poisson distribution with mean 0.25.

Find the probability that

(*a*) a randomly chosen metre of cloth has 1 defect,

(2)

(*b*) the total number of defects in a randomly chosen 6 metre length of cloth is more than 2.

(3)

A tailor buys 300 metres of cloth.

(*c*) Using a suitable approximation find the probability that the tailor's cloth will contain less than 90 defects.

(5)

---

3. A random variable X has the distribution $B(12, p)$.

(*a*) Given that $p = 0.25$, find

(i) $P(X < 5)$,

(ii) $P(X \geq 7)$.

(3)

(*b*) Given that $P(X = 0) = 0.05$, find the value of p to 3 decimal places.

(3)

(*c*) Given that the variance of X is 1.92, find the possible values of p.

(4)

---

4. The number of houses sold by an estate agent follows a Poisson distribution, with a mean of 2 per week.

(a) Find the probability that in the next four weeks the estate agent sells

(i) exactly 3 houses,

(ii) more than 5 houses.

(5)

The estate agent monitors sales in periods of 4 weeks.

(b) Find the probability that in the next twelve of those 4 week periods there are exactly nine periods in which more than 5 houses are sold.

(3)

The estate agent will receive a bonus if he sells more than 25 houses in the next 10 weeks.

(c) Use a suitable approximation to estimate the probability that the estate agent receives a bonus.

(6)

---

5. An administrator makes errors in her typing randomly at a rate of 3 errors every 1000 words.

(a) In a document of 2000 words find the probability that the administrator makes 4 or more errors.

(3)

The administrator is given an 8000 word report to type and she is told that the report will only be accepted if there are 20 or fewer errors.

(b) Use a suitable approximation to calculate the probability that the report is accepted.

(7)

---

6. A call centre agent handles telephone calls at a rate of 18 per hour.

(*a*) Give two reasons to support the use of a Poisson distribution as a suitable model for the number of calls per hour handled by the agent.

(2)

(*b*) Find the probability that in any randomly selected 15 minute interval the agent handles

(i) exactly 5 calls,

(ii) more than 8 calls.

(5)

The agent received some training to increase the number of calls handled per hour. During a randomly selected 30 minute interval after the training the agent handles 14 calls.

(*c*) Test, at the 5% level of significance, whether or not there is evidence to support the suggestion that the rate at which the agent handles calls has increased. State your hypotheses clearly.

(6)

---

7. The queuing time in minutes, X, of a customer at a post office is modelled by the probability density function

$$\mathrm{f}(x) = \begin{cases} kx(81 - x^2) & 0 \leq x \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Show that $k = \dfrac{4}{6561}$.

(3)

Using integration, find

(*b*) the mean queuing time of a customer,

(4)

(*c*) the probability that a customer will queue for more than 5 minutes.

(3)

Three independent customers shop at the post office.

(*d*) Find the probability that at least 2 of the customers queue for more than 5 minutes.

(3)

TOTAL FOR PAPER: 75 MARKS

END

Question Number	Scheme	Marks
1(a)	n large	B1
	p small	B1
		(2)
(b)	Let X be the random variable the number of letters delivered to the wrong house	
	X~B(1000,0.01)	
	Po(10)	B1
	$P(X \geq 4) = 1 - P(X \leq 3)$	M1
	$= 1 - 0.0103$	
	$= 0.9897$	A1
		(3)
		Total 5

2(a)	$P(X = 1) = 0.25e^{-0.25} = 0.1947$ awrt 0.195	M1A1
		(2)
2(b)	$X \sim$ Po(1.5)	B1
	$P(X > 2) = 1 - P(X \leq 2)$	M1
	$= 1 - 0.8088$	
	$= 0.1912$ awrt 0.191	A1
		(3)
2(c)	[$\lambda = 300 \times 0.25 = 75$]	
	$X \sim$ N(75,75)	B1 B1
	$P(X < 90) = P(X \leq (89.5 - 75)/\sqrt{75})$	M1M1
	$= P(Z \leq 1.6743..)$	
	= awrt 0.953 or 0.952	A1
		(5)
		[10]

Question Number	Scheme		Marks
3(a) (i)	$P(X < 5) = 0.8424$	awrt 0.842	B1
(ii)	$P(X \geq 7) = 1 - P(X \leq 6)$		M1
	$= 1 - 0.9857$		
	$= 0.0143$	awrt 0.0143	A1
			(3)
(b)	$P(X = 0) = (1 - p)^{12}$		
	$(1 - p)^{12} = 0.05$		M1
	$(1 - p) = \sqrt[12]{0.05}$		M1
	$p = 0.221$	awrt 0.221	A1
			(3)
(c)	Variance $= 12p(1 - p)$		
	$12p(1 - p) = 1.92$		M1
	$12p - 12p^2 = 1.92$		
	$12p^2 - 12p + 1.92 = 0$ or $p^2 - p + 0.16 = 0$ $25p^2 - 25p + 4 = 0$		
	$p = \dfrac{12 \pm \sqrt{12^2 - 4 \times 12 \times 1.92}}{24}$ $(5p - 1)(5p - 4) = 0$		M1
	$p = 0.2$ or 0.8		A1,A1
			(4)
			Total 10

Question Number	Scheme	Marks
4(a)	Let X be the random variable the number of houses sold.	
	X ~Po(8)	B1
(i)	$P(X \leq 3) - P(X \leq 2) = 0.0424 - 0.0138$ or $\frac{e^{-8}8^3}{3!}$	M1
	$= 0.0286$ awrt 0.0286	A1
(ii)	$P(X > 5) = 1 - P(X \leq 5)$	M1
	$= 1 - 0.1912$	
	$= 0.8088$ awrt 0.809	A1
		(5)
(b)	Let Y be the random variable = the number of periods where more than 5 houses are sold	
	Y~ B(12,0.8088)	M1
	$P(Y = 9) = (0.8088)^9(1 - 0.8088)^3\frac{12!}{9!3!}$	M1
	$= 0.228$ awrt 0.228	A1
		(3)
(c)	N(20,20)	M1A1
	$P(X > 25) = 1 - P\left(Z \leq \frac{25.5 - 20}{\sqrt{20}}\right)$	M1,M1, A1
	$= 1 - P(Z \leq 1.23)$	
	$= 1 - 0.8907$	
	$= 0.1093 / 0.1094$ awrt 0.109	A1
		(6)
		Total 14

Question Number	Scheme	Marks
5. (a)	X = the number of errors in 2000 words so X ~Po(6)	B1
	P($X \geq 4$) = 1 - P($X \leq 3$)	M1
	= 1 - 0.1512 = 0.8488 awrt 0.849	A1 (3)
(b)	Y = the number of errors in 8000 words. Y ~ Po(24) so use a <u>Normal</u> approx	M1
	$Y \approx\sim \text{N}(24, \sqrt{24}^2)$	A1
	Require P($Y \leq 20$) = $\text{P}\left(Z < \frac{20.5 - 24}{\sqrt{24}}\right)$	M1 M1 A1
	= P(Z < - 0.714…)	M1
	= 1 – 0.7611	A1 (7)
	= 0.2389 awrt (0.237~0.239)	**(10 marks)**

Question Number	Scheme	Marks
6. (a)	Calls occur singly — any two of the 3 only need calls once	B1
	Calls occur at a constant rate	
	Calls occur independently or randomly.	B1 (2)
(b) (i)	X~ Po(4.5) — used or seen in (i) or (ii)	M1
	$P(X = 5) = P(X \le 5) - P(X \le 4)$	M1
	$= 0.7029 - 0.5321$	A1 (3)
	$= 0.1708$	
(ii)	$P(X > 8) = 1 - P(X \le 8)$	M1
	$= 1 - 0.9597$	A1 (2)
	$= 0.0403$	
(c)	$H_0 : \lambda = 9$ ($\lambda = 18$) — may use λ or μ	B1
	$H_1 : \lambda > 9$ ($\lambda > 18$)	B1
	$X \sim \text{Po}(9)$ — may be implied	M1
	$P(X \ge 14) = 1 - P(X \le 13) = 1 - 0.9261 = 0.0739$ | [$P(X \ge 14) = 1 - 0.9261 = 0.0739$] $P(X \ge 15) = 1 - 0.9585 = 0.0415$ CR $X \ge 15$ — att $P(X \ge 14)$ $P(X \ge 15)$; awrt 0.0739	A1
	$0.0739 > 0.05$ | $14 \le 15$	M1
	Accept H_0. or it is not significant or a correct statement in context from their values	A1 (6)
	There is insufficient evidence to say that the <u>number of calls per hour</u> handled by the agent has <u>increased</u>.	**(13 marks)**

Question Number	Scheme	Marks
7.		
(a)	$\int_0^9 k(81x - x^3)\,\mathrm{d}x = 1$	**M1**
	$k\left[\frac{81}{2}x^2 - \frac{1}{4}x^4\right]_0^9 = 1$	**M1**
	$k\left(\frac{6561}{2} - \frac{6561}{4}\right) = 1$	**A1 cso**
	$k = \frac{4}{6561} = **\text{ag}**$	**(3)**
(b)	$\mathrm{E}(X) = \int_0^9 k\,x^2(81 - x^2)\,\mathrm{d}x$	
	$= k\left[\frac{81}{2}x^3 - \frac{x^5}{5}\right]_0^9$	**M1 A1**
	$= k(19683 - 11809.8)$	**dM1**
	$= 4.8$	**A1 cao**
		(4)
(c)	P(X > 5) $= \int_5^9 k(81x - x^3)$	**M1**
	$= k\left[\frac{81}{2}x^2 - \frac{1}{4}x^4\right]_5^9$	
	$= k\left(\frac{6561}{4} - 856.25\right) = \text{awrt } 0.478 \text{ or } \frac{3136}{6561}$	**M1d**
		A1
		(3)
		[13]
(d)	P(At least 2 queue for more than 5 mins) $= 3(1 - 0.478)(0.478)^2 + 0.478^3$	**M1 A1 ft**
	$= 0.467$	**A1**
		(3)
		[13]

Examiner reports

Question 1

This question was accessible to the majority of candidates, with many gaining full marks. Responses to part (a) reflected some misunderstanding in interpreting the question. This was shown by candidates who gave a list of ‘conditions for a Poisson distribution to be used’ rather than how the Poisson could be used ‘as an approximation to the binomial distribution’. Common errors seen in part (a) included $n > 30$, $p < 0.5$ and p is low. Part (b) was generally well answered and a high proportion of candidates correctly used Po (10) as the approximation to B(1000, 0.01). Common errors usually involved interpretation of inequalities e.g. using $\text{P}(X \geq 4) = 1 - \text{P}(X \leq 4)$ or $1 - \text{P}(X \leq 5)$, or finding $\text{P}(X \leq 3)$.

Question 2

This was a relatively straightforward question, for which many candidates were awarded full marks.

In part (a), nearly all candidates correctly identified the Poisson distribution, and were generally able to accurately use its probability function However, although a small minority were able to write down a correct expression for $\text{P}(X = 1)$, they were unable to accurately evaluate it on their calculators.

The responses to part (b) were usually correct, with the main errors being writing $\text{P}(X > 2)$ as $1 - \text{P}(X \leq 1)$ instead of $1 - \text{P}(X \leq 2)$ and reading the tables incorrectly.

The normal approximation to a Poisson distribution was familiar to virtually all candidates in part (c). The main errors were:

(i) standardising with 90.5 or 90 instead of 89.5, causing a loss of 3 marks;
(ii) using an incorrect value of 56.25 for the variance.

Question 3

Part (a)(i) again tested candidates’ ability to handle inequalities which they often found challenging. This was shown by incorrect answers such as $\text{P}(X \leq 5) = 0.9456$ or those finding $1 - \text{P}(X \leq 5)$. Part (b) challenged a significant number of candidates although the majority who attempted this question were able to state $(1 - p)^{12} = 0.05$. The final solution to this part, however, often proved beyond the ability of many candidates. Successful candidates were adept at finding roots using calculators or using logs to get the solution whereas those less successful made several attempts before admitting defeat.

Common errors seen involved using a mix of common and natural logs or having written correctly $\log(1 - p) = \dfrac{\log 0.05}{12}$ then not being able to write an expression for ‘$1 - p = \ldots$...’.
A high proportion of candidates answered part (c) confidently and successfully, with the majority of candidates gaining at least one mark for writing Variance = $12p(1 - p)$ or for $12pq = 1.92$ and attempting to solve their quadratic. Errors were quite often due to writing/using the quadratic formula incorrectly due to errors in basic arithmetical calculations.

Question 4

A high proportion of the candidates attempted all parts of this question successfully. Those less successful in part (a) either used Po(2) instead of Po(8), or used Po(8) but found the answer, using Poisson tables, to $P(X \le 3)$ instead of $P(X = 3)$. A common error seen in part (a)(ii) was to write and use $P(X > 5) = 1 - P(X \le 4)$.

In part (b) many candidates were able to find $P(Y = 9)$ using their answer to (a)(ii) but, for a small minority, incorrect calculations included calculating $(0.8088)^9$ or using Po(λ) with $\lambda = 8$, 9 or 24. Overall, many exemplary answers were evident for the solution to part (c) reflecting sound preparation on this topic. Marks were lost occasionally for using an incorrect, or no, continuity correction or for finding an incorrect area.

Question 5

Part (a) was answered well with the majority of candidates gaining full marks.

Part (b) was also a good source of marks for a large majority of the candidates. Common errors included using 23.9…for variance and 19.5 instead of 20.5. A sizeable minority of candidates used 21.5 after applying the continuity correction. A few candidates had correct working up to the very end when they failed to find the correct probability by not subtracting the tables' probability from 1.

Question 6

Part(a) was done well generally although a reference to 'calls' was not made by a few candidates. Weak candidates talked about the Poisson needing large numbers and others seemed to not understand what was required at all; writing 'quick and easy'. Part (b) was done correctly by the majority of candidates. A few did not use Po(4.5) in (i) and a number used $P(X > 8) = 1 - P(X \le 7)$ in (ii).In part (c) weaker candidates did not use λ for their hypotheses nor did they use Po(9). Some hypotheses had $\lambda = 3.5$ and $\lambda \ge 3.5$. Two tail tests were often suggested. Most candidates got to 0.0739 and only a few candidates used the critical value route. Only the able candidates got the interpretation of the significance test correct. Weak candidates generally only considered whether it was significant or not, with mixed success. They rarely managed to interpret correctly in context.

Question 7

Many competent and exemplary responses were seen here showing that candidates were well prepared for this type of question and a high percentage gained full marks on parts (a) to (c). In part (a) the majority of candidates realised that they had to find F(X) with only a small minority neglecting to put the integral = 1.

Common errors in part (b) included writing $E(X) = \int_0^9 x\mathrm{f}(x)\,\mathrm{d}x$ and then finding $\int_0^9 \mathrm{f}(x)\,\mathrm{d}x$ or when multiplying $x\mathrm{f}(x)$ making the very basic error of omitting to multiply the second term by x. In part (c) the most common error when using $\int_5^9 k(81x - x^3)\,\mathrm{d}x$ was to use a lower limit of 6 rather than 5. A small minority of candidates who used $P(X > 5) = 1 - P(X \le 5)$ found $P(X \le 5)$ then forgot to find $1 - P(X \le 5)$.

Part (d) was perhaps the most challenging part of a question in the paper. There were many exemplary responses but also a high proportion of incorrect attempts at using the binomial. The most common error was to swap the p and $1 - p$ over. Candidates who used a 'common sense' approach and listed the possibilities were generally successful.

Statistics for S2 Practice Paper Bronze 2

				Mean average scored by candidates achieving grade:							
Qu	**Max Score**	**Modal score**	**Mean %**	**ALL**	**A***	**A**	**B**	**C**	**D**	**E**	**U**
1	5	5	91.0	4.54	4.76	4.75	4.46	4.14	3.59	3.29	1.74
2	10	10	88.9	8.89	9.77	9.63	9.30	8.92	8.22	6.96	4.38
3	10	10	88.0	8.78	9.70	9.59	8.78	7.74	7.02	6.21	3.64
4	14		83.3	11.66	13.65	13.31	12.47	11.17	9.43	7.27	3.96
5	10		81.5	8.15		9.37	8.70	7.84	6.63	5.17	2.44
6	13		78.3	10.18		11.74	10.68	9.70	8.16	6.40	3.70
7	13		79.9	10.39	12.41	11.65	10.16	8.88	6.39	5.40	3.01
	75		**83.5**	**62.59**		**70.04**	**64.55**	**58.39**	**49.44**	**40.70**	**22.87**