Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Tuesday 17 January 2012 – Morning

Time: 1 hour 30 minutes

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
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

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
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 for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
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 not gain full credit.

1. The time in minutes that Elaine takes to checkout at her local supermarket follows a continuous uniform distribution defined over the interval [3, 9].

Find

(*a*) Elaine's expected checkout time,

(1)

(*b*) the variance of the time taken to checkout at the supermarket,

(2)

(*c*) the probability that Elaine will take more than 7 minutes to checkout.

(2)

Given that Elaine has already spent 4 minutes at the checkout,

(*d*) find the probability that she will take a total of less than 6 minutes to checkout.

(3)

---

2. David claims that the weather forecasts produced by local radio are no better than those achieved by tossing a fair coin and predicting rain if a head is obtained or no rain if a tail is obtained. He records the weather for 30 randomly selected days. The local radio forecast is correct on 21 of these days.

Test David's claim at the 5% level of significance.

State your hypotheses clearly.

(7)

---

3. The probability of a telesales representative making a sale on a customer call is 0.15.

Find the probability that

(*a*) no sales are made in 10 calls,

(2)

(*b*) more than 3 sales are made in 20 calls.

(2)

Representatives are required to achieve a mean of at least 5 sales each day.

(*c*) Find the least number of calls each day a representative should make to achieve this requirement.

(2)

(*d*) Calculate the least number of calls that need to be made by a representative for the probability of at least 1 sale to exceed 0.95.

(3)

---

4. 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)

---

5. The probability of an electrical component being defective is 0.075.
The component is supplied in boxes of 120.

(*a*) Using a suitable approximation, estimate the probability that there are more than 3 defective components in a box.

(5)

A retailer buys 2 boxes of components.

(*b*) Estimate the probability that there are at least 4 defective components in each box.

(2)

---

6. A random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{2}, & 0 \le x < 1, \\ x - \frac{1}{2}, & 1 \le x \le k, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a positive constant.

(a) Sketch the graph of f(x).

(2)

(b) Show that $k = \frac{1}{2}(1 + \sqrt{5})$.

(4)

(c) Define fully the cumulative distribution function F(x).

(6)

(d) Find P($0.5 < X < 1.5$).

(2)

(e) Write down the median of X and the mode of X.

(2)

(f) Describe the skewness of the distribution of X. Give a reason for your answer.

(2)

---

7. (*a*) Explain briefly what you understand by

(i) a critical region of a test statistic,

(ii) the level of significance of a hypothesis test.

(2)

(*b*) An estate agent has been selling houses at a rate of 8 per month. She believes that the rate of sales will decrease in the next month.

(i) Using a 5% level of significance, find the critical region for a one tailed test of the hypothesis that the rate of sales will decrease from 8 per month.

(ii) Write down the actual significance level of the test in part (*b*)(i).

(3)

The estate agent is surprised to find that she actually sold 13 houses in the next month. She now claims that this is evidence of an increase in the rate of sales per month.

(*c*) Test the estate agent's claim at the 5% level of significance. State your hypotheses clearly.

(5)

TOTAL FOR PAPER: 75 MARKS

END

January 2012
6684 Statistics S2
Mark Scheme

Question Number	Scheme	Marks
1 (a)	$\mathrm{E}(X)=\frac{9+3}{2}=6$	B1 (1)
(b)	$\mathrm{Var}(X)=\frac{(9-3)^2}{12}=3$	M1A1 (2)
(c)	$\mathrm{P}(X>7)=(9-7)\times\frac{1}{6}=\frac{1}{3}$	M1A1 (2)
(d)	$\mathrm{P}(X<6 \mid X>4)=\frac{\mathrm{P}(4<X<6)}{\mathrm{P}(X>4)}$ $=\frac{\frac{2}{6}}{\frac{5}{6}}=\frac{2}{5}$	M1A1 A1 (3) **8**
	Notes	
(b)	M1 $\frac{(9-3)^2}{12}$ or $\frac{(9+3)^2}{12}$	
(c)	M1 $\frac{(9-7)}{6}$ or $1-\frac{(7-3)}{6}$ or $\int_7^9 \frac{1}{6}\,\mathrm{d}x$ or $1-\int_3^7 \frac{1}{6}\,\mathrm{d}x$ A1 Also acceptable $0.\dot{3}$, $0.3\dot{3}$ and awrt 0.333	
(d)	M1 $\frac{\mathrm{P}(4<X<6)}{P(X>4)}$ or $\frac{\mathrm{P}(X<6)}{P(X>4)}$ or $\frac{\frac{2}{6}}{\frac{5}{6}}$ or $\frac{\frac{3}{6}}{\frac{5}{6}}$ or $1-\frac{\mathrm{P}(X>6)}{P(X>4)}$ or $\frac{6-4}{9-4}$ or $\frac{3}{5}$ A1 $\frac{\mathrm{P}(4<X<6)}{P(X>4)}$ or $\frac{\frac{2}{6}}{\frac{5}{6}}$ or $1-\frac{\mathrm{P}(X>6)}{P(X>4)}$ or $\frac{6-4}{9-4}$ An answer of $\frac{2}{5}$ gains all 3 marks. NB $\leq$ and $\geq$ are accepted in the above formulae	

Question Number	Scheme	Marks
2	$H_0 : p = 0.5$	B1
	$H_1 : p > 0.5$	B1
	$X \sim B(30,0.5)$ Using correct Bin	M1
	$P(X \geq 21) = 1 - P(X \leq 20)$ or $P(X \leq 19) = 0.9506$ $P(X \geq 20) = 0.0494$	M1
	$= 1 - 0.9786$	
	$= 0.0214$ CR $X \geq 20$	A1
	so significant/reject H_0 /in Critical region	M1 dep
	Evidence to suggest **<u>David's claim is incorrect</u>** or The weather **<u>forecast</u>** produced by the local **<u>radio</u>** is better than those achieved by **<u>tossing/flipping a coin</u>**	A1 (7)
		7

Notes

1st B1 for $H_0 : p = 0.5$
2nd B1 for $H_1 : p > 0.5$
SC If both hypotheses are correct but a different letter to p is used they get B1 B0. If no letter is used they get B0 B0.

1st M1 writing or using B(30,0.5)
<u>One tail</u>
2nd M1 for writing or using 1 - P($X \leq 20$) or writing P($X \leq 19$) = 0.9506 or P($X \geq 20$) = 0.0494. May be implied by correct CR.or probability = 0.0214
A1 for 0.0214 or CR $X \geq 20$/ $X > 19$. **NB** P($X \leq 20$) = 0.9786 on its own scores M1A1
3rd M1 dependent on the 2nd M1 being awarded. For a correct statement based on the table below. Do not allow non-contextual conflicting statements eg "significant" and "accept H_0". **Ignore comparisons**.
2nd A1 for a correct contextualised statement. NB A correct contextual statement on its own scores M1A1.

	$0.05 < p < 0.95$	$p < 0.05$ or $p > 0.95$
3rd M1	not significant/ accept H_0/ Not in CR	significant/ reject H_0/ In CR
2nd A1	David's claim is correct weather **<u>forecast</u>** produced by the local **<u>radio</u>** is no better than those achieved by **<u>tossing/flipping a coin</u>**	David's claim incorrect weather **<u>forecast</u>** produced by the local **<u>radio</u>** is better than those achieved by **<u>tossing/flipping a coin</u>**

<u>Two tail</u>
1st M1 for writing or using 1 - P($X \leq 20$) or writing P($X \leq 20$)= 0.9786 or P($X \geq 21$) = 0.0214. May be implied by correct CR. or probability = 0.197
A1 for 0.0214 or CR $X \geq 21$/ $X > 20$. **NB** P($X \leq 20$) = 0.9786 on its own scores M1A1
3rd M1 dependent on the 2nd M1 being awarded . For a correct statement based on the table below. Do not allow non-contextual conflicting statements eg"significant" and "accept H_0" . **Ignore comparisons**.
2nd A1 for a correct contextualised statement. NB A correct contextual statement on its own scores M1A1.

	$0.025 < p < 0.975$	$p < 0.025$ or $p > 0.975$
3rd M1	not significant/ accept H_0/ Not in CR	significant/ reject H_0/ In CR
2nd A1	David's .claim is correct weather **<u>forecast</u>** produced by the local **<u>radio</u>** is no better than those achieved by **<u>tossing/flipping a coin</u>**	David's claim incorrect weather **<u>forecast</u>** produced by the local **<u>radio</u>** is better than those achieved by **<u>tossing/flipping a coin</u>**

Question	Scheme	Marks

Number		
3 (a)	$P(X = 0) = 0.85^{10}$ or from tables	M1
	$= 0.1969$ awrt 0.197	A1
		(2)
(b)	$P(X > 3) = 1 - P(X \leq 3)$	M1
	$= 1 - 0.6477$	
	$= 0.3523$ awrt 0.352	A1
		(2)
(c)	$n \times 0.15 = 5$	M1
	$n = 33$ or 34	A1
		(2)
(d)	1 - P(X = 0) > 0.95	M1
	$1 - (0.85)^n > 0.95$.	A1
	$0.85^n < 0.05$	
	$n > 18.4$	
	$n = 19$	A1
		(3)
		9

Notes

(a) M1 $(p)^{10}$ with $0 < p < 1$

(b) M1 writing or using 1 - P($X \leq 3$)

(c) M1 $np = 5$ $0 < p < 1$

(d) M1 writing or using 1 - P(X = 0) > 0.95 or P(X = 0) < 0.05 (also accepted are = or $\geq$ instead of > and = or $\leq$ instead of or <) P($X \leq 0$) is equivalent to P(X =0)
A1 writing or using 1 - $(0.85)^n > 0.95$ or $(0.85)^n < 0.05$ (also accepted are $\geq$ instead of > and $\leq$ instead of or <). Any value of n may be used
A1 cao

NB an answer of 18.4 gets M1 A1 A0

An answer of 19 gets M1 A1 A1 unless it follows from clearly incorrect working.

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

Notes

(b) 1st B1 Any one of the 3 statements - no context required. NB It must be a constant (mean) rate and not a constant probability or a constant mean.
2nd B1 A different statement with context of **<u>hits.</u>** NB random and independent are the same statement.
If only one mark awarded give the 1st B1. Never award B0 B1

(c) B1 writing or using Po(5)

M1 writing or using $\text{P}(X\leq 10) - \text{P}(X\leq 9)$ or $\frac{e^{-5}5^{10}}{10!}$

(d) B1 writing or using Po(10)
M1 writing or using 1- $\text{P}(X\leq 14)$

(e) 1st B1 for a normal approximation 2nd B1 for correct mean and sd (may be seen in standardisation formula
1st M1 for attempting a continuity correction (71 ± 0.5)
2nd M1 Standardising using their mean and their sd and using [69.5, 70, 70.5, 71 or 71.5] allow $\pm$ z
NB if they have not written down a mean and sd then they need to be correct in the standardisation to gain this mark.

1st A1 for $z = \pm$ awrt 2.9 or better. May be awarded for $\pm\frac{70.5-50}{\sqrt{50}}$

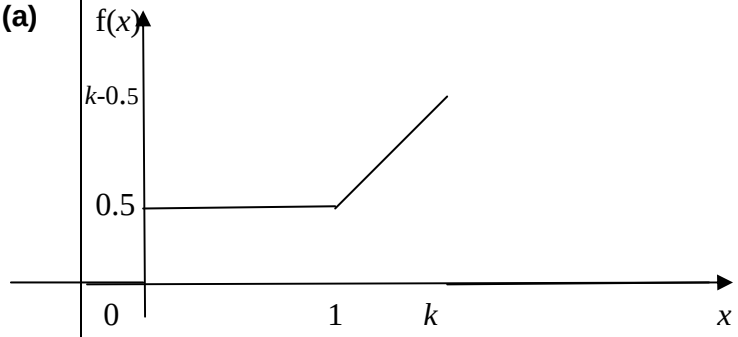
3rd M1 for 1 - tables value

SC using P(X< 70.5/71.5) – P(X<69.5/70.5) can get B1B1 M0M1A0 M0A0

Question Number	Scheme	Marks
5 (a)	$X \sim B(120,0.075)$	B1
	Approximated by Po(9)	M1A1
	$P(X > 3) = 1 - P(X \leq 3)$	M1
	$= 1 - 0.0212$	
	$= 0.9788$ awrt 0.979	A1
		(5)
(b)	P(At least 4 defective components in each box)	
	$= P(X > 3) \times P(X > 3)$	M1
	$= 0.9788^2$	
	$= 0.95804944$ awrt 0.958	A1
		(2)
		7
	Notes	
(a)	B1 Writing or use of B(120,0.075) may be implied by using Po(9) or N(9,8.325)	
	1st M1 writing or use of Poisson	
	1st A1 writing or use of Po(9)	
	2nd M1 for writing or using 1- P($X \leq 3$) or this may be implied by an awrt 0.972 using normal approximation.	
(b)	M1 ((their (a))^2 or 0.979^2 or 0.9788^2 or 0.98^2	

Question Number	Scheme	Marks

6 (a)



shape B1
labels B1

(2)

(b)

$\int_1^k \left(x-\frac{1}{2}\right)\mathrm{d}x = \frac{1}{2}$ M1

$\left[\frac{1}{2}x^2-\frac{1}{2}x\right]_1^k = \frac{1}{2}$

$k^2-k-1=0$ o.e. A1

$k=\frac{1}{2}\left(1+\sqrt{5}\right)$ M1A1 cso

(4)

(c)

$$F(x)=\begin{cases}0, & x<0\\ \frac{1}{2}x, & 0\le x<1\\ \frac{1}{2}x^2-\frac{1}{2}x+\frac{1}{2}, & 1\le x\le k\\ 1, & x>k\end{cases}$$

B1

M1A1A1B1

B1 1st and last

(6)

Note: Working for the M1A1A1

$\int_1^k x-\frac{1}{2}\mathrm{d}x + \mathrm{C} = \frac{1}{2}x^2-\frac{1}{2}x\ ;+\frac{1}{2}$ (M1A1;A1)

(d)

$\mathrm{P}(0.5<X<1.5) = \mathrm{F}(1.5)-\mathrm{F}(0.5)$

$= 0.875 - 0.25$ M1

$= 0.625$ A1

(2)

(e)

Median is $x=1$ B1

Mode is $x=k$ or $\frac{1}{2}(1+\sqrt{5})$ or awrt1.62 B1

(2)

(f)

Negative skew B1

Median<mode or from graph more values are to the right. B1d

(2)

18

(a)

Notes

1st B1 Correct shape with straight lines. Must all be above the *x*-axis

2nd B1 A fully correct graph with the labels 1, *k*, 0.5, *k* - 0.5 seen in the correct places.

Allow the use of $\frac{1}{2}(1+\sqrt{5})$ /awrt 1.62 instead of *k*.

(b)	1st M1 $\int_1^k x-\frac{1}{2}\,\mathrm{d}x = 0.5$ or $\int_1^k x-\frac{1}{2}\,\mathrm{d}x + 0.5 = 1$ ignore limits or $\int_1^k x-\frac{1}{2}\,\mathrm{d}x + \int_1^k \frac{1}{2}\,\mathrm{d}x = 1$ or $\frac{1}{2}(k-0.5+0.5)(k-1) = 0.5$ or any correct method of finding the area 1st A1 for a quadratic equation in the form $a(k^2 - k - 1) = 0$ or $ak^2 - ak = a$. where a is a constant. 2nd M1 correct method for solving a quadratic of the form $ak^2 - bk + c = 0$ where $a,b,c \neq 0$. There must be at least one correct step before the final answer. Allow substituting in k into a quadratic of the form $ak^2 - bk + c = 0$. 2nd A1 cso for $k = \frac{1}{2}\left(1+\sqrt{5}\right)$	
(c)	1st B1 for second line. Do not penalise the use of < instead of $\leq$ and vice versa M1 for use of $\int_1^k x-\frac{1}{2}\,\mathrm{d}x$ + C ignore limits. For use they must have $x \rightarrow x^2$ 1st A1 correct integration $\frac{1}{2}x^2 - \frac{1}{2}x$ 2nd A1 C = $\frac{1}{2}$ NB M1A1A1 may be implied by correct 3rd line in F(x) 2nd B1 for 3rd line. Statement of the form $\frac{1}{2}x^2 - \frac{1}{2}x \pm C$. Do not penalise the use of < instead of $\leq$ and vice versa. Allow k or value of k. C may equal 0. 3rd B1 for first and last line. Do not penalise the use of $\leq$ instead of < and $\geq$ instead of >. Allow k or value of k	
(d)	M1 **Using** F(1.5) - F(0.5) . 1.5 must be put into the third line of the c.d.f. and 0.5 must be put into the second line of the c.d.f.. or $\int_{0.5}^{1} \frac{1}{2}x\,\mathrm{d}x + \int_1^{1.5} x-\frac{1}{2}\,\mathrm{d}x$ need to attempt integration, at least one $x^n \rightarrow x^{n+1}$ or seeing 0.25 + 0.375 or any correct method of finding the area.. (NB if they have not used + C or C = 0 they will get 0.125. This will get M1A0). An answer of 0.125 from an incorrect method gains M0 A0.	
(e)	If it is not clear which one is the mode and which one is the median assume the median is the first answer and mode the second.	
(f)	B1 negative/negative skew(ness). Do not allow negative correlation. B1 dependent on previous B mark being awarded. Reason must follow from their values or diagram.	

Question Number	Scheme	Marks
7 (a) (i)	The **range of values/region/area/set of values** of the test statistic that would lead you to **reject H_0**	B1
(a) (ii)	The probability of incorrectly rejecting H_0 or Probability of rejecting H_0 when H_0 is true	B1 (2)

(b) (i)	$X \sim$Po(8)		M1
	$P(X \leq 4) = 0.0996$ $P(X \leq 3) = 0.0424$		
	Critical region [0,3]		A1
(b) (ii)	awrt 0.0424		B1 (3)
(c)	$H_0 : \lambda = 8$ (or μ =8)		B1
	$H_1 : \lambda > 8$ (or μ >8)		
	$P(X \geq 13) = 1 - P(X \leq 12)$	or $P(X \leq 13) = 0.9658$ or $P(X \geq 14) = 0.0342$	M1
	$= 1 - 0.9362$		
	$= 0.0638$	CR $X \geq 14$	A1
	so insufficient evidence to reject H_0 /not significant/ not in critical region		M1 dep
	There in insufficient evidence of an increase/change in the **rate/number** of sales per month **or** the estate **agents** claim is incorrect		A1 (5)
			10

Notes

(a)(i) Allow accept H_1 instead of reject H_0. It must be clear which hypothesis gets rejected/accepted.

(ii) Allow equivalent wording.

(b) M1 Writing or using Po(8). May be implied by correct critical region.

A1 allow $0 \leq X \leq 3$ or CR ≤ 3 or $X \leq 3$. Any letter may be used but not $P(X \leq 3)$. This must be on its own.

(c) B1 both hypotheses correct. Must use λ or μ.

One tail

1st M1 for writing or using 1 - $P(X \leq 12)$ or writing $P(X \leq 13) = 0.9658$ or $P(X \geq 14) = 0.0342$. May be implied by correct CR.or probability = 0.0638

A1 for 0.0638 or $X \geq 14$. Allow $X > 13$. NB $P(X \leq 12) = 0.9362$ on its own scores M1A1

2nd M1 dependent on the 1st M1 being awarded. For a correct statement based on the table below. Do not allow non-contextual conflicting statements eg "not significant" and "reject H_0". **Ignore comparisons**.

2nd A1 for a correct contextualised statement. NB A correct contextual statement on its own scores M1A1.

	$0.05 < p < 0.95$	$p < 0.05$ or $p > 0.95$
2nd M1	not significant/ accept H_0/ Not in CR	significant/ reject H_0/ In CR
2nd A1	Insufficient evidence of an increase/change in the **rate/number** of sales per month	Sufficient evidence of an increase/change in the **rate/number** of sales per month

Two tail

1st M1 for writing or using 1 - $P(X \leq 12)$ or writing $P(X \leq 14) = 0.9827$ or $P(X \geq 15) = 0.0173$. May be implied by correct CR.or probability = 0.0638

A1 for 0.0638 or $X \geq 15$. Allow $X > 14$. NB $P(X \leq 12) = 0.9362$ on its own scores M1A1

2nd M1 dependent on the 1st M1 being awarded . For a correct statement based on the table below. Do not allow non-contextual conflicting statements eg "not significant" and "reject H_0".**Ignore comparisons**.

2nd A1 for a correct contextualised statement. NB A correct contextual statement on its own scores M1A1.

	$0.025 < p < 0.975$	$p < 0.025$ or $p > 0.975$
2nd M1	not significant/ accept H_0/ Not in CR	significant/ reject H_0/ In CR
2nd A1	Insufficient evidence of an increase/change in the **rate/number** of sales per month	Sufficient evidence of an increase/change in the **rate/number** of sales per month