Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 18 January 2013 – Afternoon

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. (*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. 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 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 continuous random variable X is uniformly distributed over the interval $[-4, 6]$.

(a) Write down the mean of X.

(1)

(b) Find $P(X \leq 2.4)$.

(2)

(c) Find $P(-3 < X - 5 < 3)$.

(2)

The continuous random variable Y is uniformly distributed over the interval $[a, 4a]$.

(d) Use integration to show that $E(Y^2) = 7a^2$.

(4)

(e) Find $\text{Var}(Y)$.

(2)

(f) Given that $P(X < \frac{8}{3}) = P(Y < \frac{8}{3})$, find the value of a.

(3)

---

5. The continuous random variable T is used to model the number of days, t, a mosquito survives after hatching.

The probability that the mosquito survives for more than t days is

$$\frac{225}{(t+15)^2}, \quad t \geq 0.$$

(a) Show that the cumulative distribution function of T is given by

$$F(t) = \begin{cases} 1 - \dfrac{225}{(t+15)^2}, & t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

(1)

(b) Find the probability that a randomly selected mosquito will die within 3 days of hatching.

(2)

(c) Given that a mosquito survives for 3 days, find the probability that it will survive for at least 5 more days.

(3)

A large number of mosquitoes hatch on the same day.

(d) Find the number of days after which only 10% of these mosquitoes are expected to survive.

(4)

---

6. (*a*) Explain what you understand by a hypothesis.

(1)

(*b*) Explain what you understand by a critical region.

(2)

Mrs George claims that 45% of voters would vote for her.

In an opinion poll of 20 randomly selected voters it was found that 5 would vote for her.

(*c*) Test at the 5% level of significance whether or not the opinion poll provides evidence to support Mrs George's claim.

(4)

In a second opinion poll of n randomly selected people it was found that no one would vote for Mrs George.

(*d*) Using a 1% level of significance, find the smallest value of n for which the hypothesis $H_0 : p = 0.45$ will be rejected in favour of $H_1 : p < 0.45$.

(3)

---

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

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

where a and b are constants.

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

(4)

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

(*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

January 2013
6684 Statistics S2
Mark Scheme

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	
	$X \sim B(1000,0.01)$	
	Po(10)	B1
	$P(X \geq 4) = 1 - P(X \leq 3)$	M1
	$= 1 - 0.0103$	
	$= 0.9897$	A1
		(3)
		Total 5

Notes

(a) B1 Accept n (the number of trials) large / high / big / $n > 50$ (accept any number larger than 50)

B1 Accept p (the probability) small / close to 0 / $p < 0.2$ (accept any number less than 0.2). Do not accept low.

These must appear in part (a).

(b) B1 writing or using Po(10)

M1 using a Poisson (λ need not equal 10) and for writing or using $1 - P(X \leq 3)$. (Do not accept writing $1 - P(X < 4)$ unless they have used $1 - P(X \leq 3)$).

A1 0.9897 cao must be 4 dp

NB

An awrt 0.990 on its own gains B0M0A0 unless there is evidence that Po(10) is used. In which case it gets B1M1A0

Using B(1000,0.01) gives 0.989927…. and gains B0M0A0

Question Number	Scheme	Marks
2 (a)	Let X be the random variable the number power cuts.	
	$X \sim \text{Po}(3)$	B1
(i)	$P(X = 7) = P(X \leq 7) - P(X \leq 6)$ or $\frac{e^{-3}3^7}{7!}$	M1
	$= 0.9881 - 0.9665$	
	$= 0.0216$ awrt 0.0216	A1
(ii)	$P(X \geq 4) = 1 - P(X \leq 3)$	M1
	$= 1 - 0.6472$	
	$= 0.3528$ awrt 0.353	A1
		(5)
(b)	$X \sim \text{Po}(30)$	
	N(30,30)	M1A1
	$P(X < 20) = P\left(Z < \frac{19.5-30}{\sqrt{30}}\right)$	M1M1 A1
	$= P(Z < -1.92)$	
	$= 1 - 0.9726$	
	$= 0.0274 - 0.0276$	A1
		(6)
	Notes	**Total 11**
(a)	B1 Writing or using Po(3) in either (i) or (ii)	
(i)	M1 writing or using $P(X \leq 7) - P(X \leq 6)$ or $\frac{e^{-\lambda}\lambda^7}{7!}$	
(ii)	M1 writing or using $1 - P(X \leq 3)$. (Do not accept writing $1 - P(X < 4)$ unless they have used $1 - P(X \leq 3)$).	
(b)	1st M1 for writing or using a normal approximation 1st A1 for correct mean and sd (may be given if correct in standardisation formula) 2nd M1 Standardising using their mean and their sd **and** using [18.5, 19, 19.5, 20 or 20.5] **and** for finding correct area by doing **1 – P(*Z* ≤ "their 1.92")** If they have not written down a mean and sd then these need to be correct here to award the mark 3rd M1 for attempting a continuity correction (19 ± 0.5) i.e. 18.5 or 19.5 **only**. 2nd A1 for $\pm\frac{19.5-30}{\sqrt{30}}$ or $\pm$ awrt 1.9 or better. 3rd A1 awrt 0.0274, 0.0275 or 0.0276 **SC using P(*X*< 20.5/19.5) – P(*X*< 19.5/18.5) can get M1A1 M0M1A0A0**	

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 $p^2 - p + 0.16 = 0$ or $25p^2 - 25p + 4 = 0$		
	$p = \frac{12 \pm \sqrt{12^2 - 4\times12\times1.92}}{24}$ or $(5p-1)(5p-4) = 0$		M1
	$p = 0.2$ or 0.8		A1,A1
			(4)
	Notes		**Total 10**
(a) (ii)	M1writing or using $1 - P(X \leq 6)$ Do not accept $1 - P(X < 7)$ unless $1 - P(X \leq 6)$ has been used		
(b)	1st M1 $(1-p)^n = 0.05$ 2nd M1 taking nth root. If they have used logs they need to get to a correct expression for $1 - p$ for their equation.		
(c)	1st M1 $12p(1 - p) = 1.92$ o.e. 2nd M1 solving a quadratic either by factorising / completing the square / or formula. Working must either be correct for **their** quadratic (they may use a quadratic from an incorrect rearrangement) **or** they must have written the appropriate formula down correctly and only made 1 error substituting into it. May be implied by a correct value of p. 1st A1 for 0.2 2nd A1 for 0.8		

Question Number	Scheme	Marks
4 (a)	Mean = 1	B1 **(1)**
(b)	$P(X \leq 2.4) = (2.4 - -4) \times \frac{1}{10}$	M1
	$= 0.64$ or $\frac{16}{25}$	A1
		(2)
(c)	$P(-3 < X - 5 < 3) = P(2 < X < 6)$	M1
	$= 0.4$	A1
		(2)
(d)	$\int_a^{4a} \frac{y^2}{4a-a} dy = \left[\frac{y^3}{9a}\right]_a^{4a}$	M1
		M1 dep
	$= \frac{64a^3 - a^3}{9a}$	A1
	$= 7a^2$ *AG	A1cso
		(4)
(e)	$\text{Var}(Y) = \frac{1}{12}(4a-a)^2$ or $\text{Var}(Y) = 7a^2 - \left(\frac{5}{2}a\right)^2$	M1
	$= \frac{3}{4}a^2$	A1cso
		(2)
(f)	$\frac{2}{3} = \frac{1}{3a}\left(\frac{8}{3} - a\right)$	M1
		A1
	$a = \frac{8}{9}$	A1
		(3)
		Total 14

Notes

(b) M1 $(2.4 - -4) \times \frac{1}{10}$ or $1 - (6 - 2.4) \times \frac{1}{10}$ o.e

(c) M1 finding $P(2 < X < 6)$ or $P(X > 2)$ or $1 - P(X < 2)$. May be implied by a correct answer if there is no incorrect working. Do not ignore subsequent incorrect working.
NB if they change the distribution to U[-9,1] then M1 is for finding $P(-3 < X < 1)$ or $P(X > -3)$ or $1 - P(X < -3)$. May be implied by a correct answer if there is no incorrect working. Do not ignore subsequent incorrect working.

(d) **NB remember the answer is given (AG) so they must show their working**

1st M1 writing or using $\int_a^{4a} y^2 f(y) dy$ with correct limits used at some point. Condone omission of dy. f(y) does not need to be correct.

2nd M1 dependent on previous M being awarded. Attempting to integrate at $y^n \rightarrow \frac{y^{n+1}}{n+1}$

1st A1 correct expression - the correct limits must be substituted.

2nd A1 cso

(e) (f)	M1 either use of $\frac{(b-a)^2}{12}$ or $E(Y^2) - [E(Y)]^2$:- they may use their part (d) for $E(Y^2)$ M1 using $\frac{1}{3a}\left(\frac{8}{3} - a\right)$ = a probability **or** $\frac{1}{3a}\left(4a - \frac{8}{3}\right)$ = a probability An answer of $\frac{8}{9}$ **with no incorrect working gains M1A1A1**	

Question Number	Scheme	Marks
5(a)	$P(T > t) = \dfrac{225}{(t+15)^2}$ $P(T \leq t) = 1 - P(T > t)$ $= 1 - \dfrac{225}{(t+15)^2}$ $F(t) = \begin{cases} 1 - \dfrac{225}{(t+15)^2} & t \geq 0 \\ 0 & \text{otherwise.} \end{cases}$	B1 **(1)**
(b)	$P(T < 3) = 1 - \dfrac{225}{(3+15)^2}$ $= \dfrac{11}{36}$ or 0.30555… awrt 0.306	M1 A1 **(2)**
(c)	$P(T > 8 \mid T > 3) = \dfrac{P(T > 8)}{P(T > 3)}$ $= \dfrac{\frac{225}{23^2}}{\frac{225}{18^2}}$ $= \dfrac{324}{529}$ or 0.612.. awrt 0.612 / 0.6125	M1 M1 A1 **(3)**
(d)	$1 - F(t) = 0.1$ $\dfrac{225}{(t+15)^2} = 0.1$ or $1 - \dfrac{225}{(t+15)^2} = 0.9$ $\dfrac{225}{0.1} = (t+15)^2$ $t = \sqrt{\dfrac{225}{0.1}} - 15$ t =32.4, also accept 32/33	M1 A1 M1 A1 **(4)** **Total 10**

Notes

(a) B1 The line $P(T \leq t) = 1 - P(T > t)$ or $F(t) = 1 - P(T > t)$ or both of the following statements $P(T > t) = \frac{225}{(t+15)^2}$ and $P(T \leq t)$ /$F(t) = 1 - \frac{225}{(t+15)^2}$ must be seen and no errors. Allow equivalent in words.
Condone use of < instead of ≤ or > instead of ≥ and vice versa.
The cdf must be given. Allow $t > 0$

(b) M1 substituting 3 into $F(t)$

(c) 1st M1 The conditional probability must,

- be a quotient and
- have $P(T > 3)$ or 'their numerical equivalent' for the denominator and
- have $P(T > 8)$ or $P(T > 5)$ or $P(T > 8 \cap T > 3)$ or $P(T > 5 \cap T > 3)$ or 'their numerical equivalent' for the numerator.

Allow ≥ in place of >

2nd M1 writing or using $P(T > 8)$ or $P(T \geq 8)$.
NB This is independent of the first M mark.

(d) 1st M1 writing or using $1 - F(t) = 0.1$ or $P(T \geq t) = 0.1$ May be implied by $\frac{225}{(t+15)^2} = 0.1$ o.e.

2nd M1 either square rooting or solving a quadratic either by factorising / completing the square / using the formula - must be correct for their quadratic.

A1 awrt 32.4 or 32 or 33. Do not accept $15\sqrt{10} - 15$

Question Number	Scheme	Marks
6(a)	A statement concerning a **population parameter**	B1
(b)	A critical region is the range / set of values / answers **or** a test statistic **or** region/area **or** values (where the test is significant)	B1
	that would lead to the rejection of H0 / acceptance of H_1	B1
		(3)
(c)	$H_0 : p = 0.45$ $H_1 : p < 0.45$ (or $p \neq 0.45$)	
	$X \sim B(20, 0.45)$	M1
	$P(X \leq 5) = 0.0553$ CR $X \leq 4$	A1
	Accept H_0. Not significant. 5 does not lie in the Critical region.	M1d
	There is no evidence that the proportion who voted for Mrs George is not 45% or there is evidence to support Mrs George's claim	A1cso
		(4)
(d)	B(8, 0.45): P(0) = 0.0084	M1
	B(7, 0.45): P(0) = 0.0152	A1
	Hence smallest value of n is **8**	B1
		(3)
	Alternative	
	$(0.55)^n < 0.01$	M1
	$n\log 0.55 < \log 0.01$	
	$n > 7.7\ldots$	A1
	Hence smallest value of n is **8**	B1cso
	Notes	**Total 10**
(a)	It must be a statement including the words **population parameter**.	
(c)	1st M1 using B(20, 0.45) and finding $P(X \leq 5)$ or $P(X \geq 6)$ Using the normal approximation to the binomial is M0 A1 0.0553 (allow 0.9447) if not using CR or CR $X \leq 4$ or $X < 5$ 2nd M1 dependent on previous M being awarded. A correct statement (do not allow if there are contradicting non contextual statements **nor** award if 2 probabilities are given which would result in different conclusions) A1cso Conclusion must contain the words **Mrs George**. There must be no incorrect working seen. If there are no hypotheses you cannot award this mark. **NB** A correct contextual statement on it's own will score M1 A1.	
(d)	M1 Attempt to find P(0) from B(n, 0.45) or $(0.55)^n < 0.01$ or $(0.55)^n = 0.01$ or $(0.55)^n > 0.01$ A1 P(0) = 0.0084 and P(0) = 0.0152 or getting 7.7 May be implied by correct answer. B1 cso. $n = 8$ should not come from incorrect working. NB An answer of 8 on its own with no working gains M1A1B1	

Question Number	Scheme	Marks
7(a)	$\int_0^5 a+bx \, \mathrm{d}x = 1$	M1
	$\left[ax + \frac{bx^2}{2}\right]_0^5 = 1$	A1
	$5a + \frac{25b}{2} = 1$	M1dep
	$10a + 25b = 2$	A1cso
		(4)
(b)	$\int_0^5 ax + bx^2 \, \mathrm{d}x = \frac{35}{12}$	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
	$10a + 25b = 2$	
	$a = 0.1 \;\; b = 0.04$	A1,A1
		(3)
(d)	$\int_0^m 0.1 + 0.04x \, \mathrm{d}x = 0.5$	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)
		Toal 15

Notes

(a)
1st M1 Attempting to integrate with correct limits or for an attempt to find area $0.5(a + b)h$ or Attempting to integrate and using F(5) = 1
1st A1 Correct integration or correct area
2nd M1 for using =1. This is dependent on the first M1 being awarded.
2nd A1 cso condone missing dx

(b)
M1 using or writing (limits not needed) $\int_0^5 ax + bx^2 \, \mathrm{d}x = \frac{35}{12}$
1st A1 correct integration
2nd A1 may be awarded for an unsimplified version $\frac{25a}{2} + \frac{125b}{3} = \frac{35}{12}$

(c)	M1 attempting to solve "their equations" simultaneously – either using rearranging and substitution **or** making one of the coefficients the 'same' (ignore sign) and either adding or subtracting. May be implied by correct values for a and b 1st A1 for 0.1 2nd A1 for 0.04
(d)	M1 writing or using $\int_0^m "their\ a"+"their\ b"x\ \mathrm{d}x = 0.5$: limits not needed 1st A1 correct integration for their "a" and "b" **NB** the correct equation simplifies to $m^2 + m - 25 = 0$ A1 3.09 only. If they have both roots then they must select 3.09
(e)	1st B1ft. They must compare their values for mean and median correctly. They only need to compare 2 of mean, median and mode. If they compare either the median or mean with the **mode only** then the value of the mode must be stated. They may draw a sketch that matches their values of 'a' and 'b' for $0 \leq x \leq 5$. It must not go below the x-axis This may be seen in part (a). 2nd B1 dependent f.t. on the previous B being awarded.