Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 14 January 2011 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	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 disease occurs in 3% of a population.

(*a*) State any assumptions that are required to model the number of people with the disease in a random sample of size n as a binomial distribution.

(2)

(*b*) Using this model, find the probability of exactly 2 people having the disease in a random sample of 10 people.

(3)

(*c*) Find the mean and variance of the number of people with the disease in a random sample of 100 people.

(2)

A doctor tests a random sample of 100 patients for the disease. He decides to offer all patients a vaccination to protect them from the disease if more than 5 of the sample have the disease.

(*d*) Using a suitable approximation, find the probability that the doctor will offer all patients a vaccination.

(3)

---

2. A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.

State your hypotheses clearly.

(6)

---

3. The continuous random variable X is uniformly distributed over the interval $[-1, 3]$.
Find

(*a*) $\mathrm{E}(X)$

(1)

(*b*) $\mathrm{Var}(X)$

(2)

(*c*) $\mathrm{E}(X^2)$

(2)

(*d*) $\mathrm{P}(X < 1.4)$

(1)

A total of 40 observations of X are made.

(*e*) Find the probability that at least 10 of these observations are negative.

(5)

---

4. Richard regularly travels to work on a ferry. Over a long period of time, Richard has found that the ferry is late on average 2 times every week. The company buys a new ferry to improve the service. In the 4-week period after the new ferry is launched, Richard finds the ferry is late 3 times and claims the service has improved. Assuming that the number of times the ferry is late has a Poisson distribution, test Richard's claim at the 5% level of significance. State your hypotheses clearly.

(6)

---

5. A continuous random variable X has the probability density function $f(x)$ shown in Figure 1.

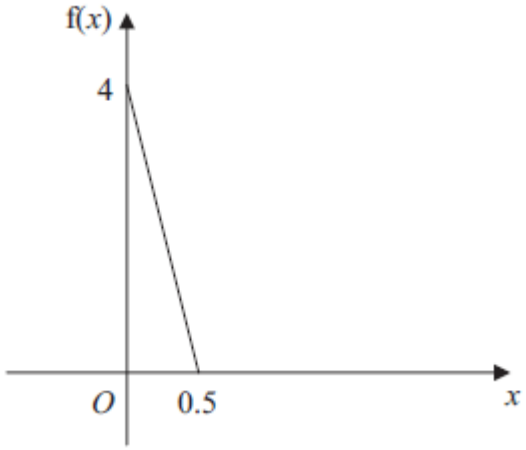

Figure 1

(*a*) Show that $f(x) = 4 - 8x$ for $0 \leq x \leq .0.5$ and specify $f(x)$ for all real values of x.

(4)

(*b*) Find the cumulative distribution function $F(x)$.

(4)

(*c*) Find the median of X.

(3)

(*d*) Write down the mode of X.

(1)

(*e*) State, with a reason, the skewness of X.

(1)

---

6. 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)

---

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

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

(*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

January 2011
Statistics S2 6684
Mark Scheme

Question Number	Scheme	Marks
1. **(a)**	Occurrences of the disease are independent The probability of catching the disease remains constant.	B1 B1 (2)
(b)	$X \sim \text{Bin}(10,0.03)$ $\text{P}(X=2)=\frac{10\times 9}{2}(0.03)^2(0.97)^8=0.0317$	B1 M1A1 (3)
(c)	$\text{E}(X)=100\times 0.03=3$ $\text{Var}(X)=100\times 0.03\times 0.97=2.91$	B1cao B1cao (2)
(d)	$\lambda=100\times 0.03=3$ $Y\sim \text{Po}(3)$ $\text{P}(Y>5) \quad =1-\text{P}(Y\le 5)$ $=1-0.9161$ $=0.0839$	B1 (use of) dM1 A1 (3) **[10]**
	Notes	
(a)	B1 independent B1 probability remains constant. One of these must have the context of disease. No context only one correct B0B0 If only one mark awarded give the first B1 SC if they are both correct without context award B1B0	
(b)	B1 for writing or using B(10,0.03) M1 for writing or using $(p)^2\,(1-p)^8\ \frac{10!}{2!8!}$ allow ${}^{10}\text{C}_2, \binom{10}{2}$ etc Allow P(X≤2) – P(X≤1) A1 awrt 0.0317	
(d)	B1 for using Poisson. Any mean. Common values which imply Poisson used are 0.9665 and 0.8153 dM1 for writing or using $1-\text{P}(X\le 5)$ - use of binomial gets M0. This is dependent on them being awarded the previous B mark. A1 awrt 0.0839 **SC: Use of Normal in (d)** Can get B0 M1 A0.- for M1 we must see $1-\text{P}(X\le 5)$ or $1-\text{P}(X\le 5.5)$ oe or get awrt 0.071	

Question Number	Scheme	Marks
2.	$H_0 : p = 0.2 \quad H_1 : p > 0.2$ Under H_0, $X \sim \text{Bin}(10,0.2)$ $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.8791 = 0.1209$ OR $P(X \leq 4) = 0.9672$, $P(X \geq 5) = 0.0328$, CR $X \geq 5$ 0.1209>0.05. Insufficient evidence to reject H_0 so teacher's claim is supported.	B1 B1 M1 A1 M1A1ft **[6]**

Notes

B1 for both H_0 and H_1 correct. Must use p or π (pi)
B1 for writing or using Bin(10,0.2)
M1 for finding or writing $1 - P(X \leq 3)$ or $P(X \leq 4) = 0.9672$
$P(X \geq 5) = 0.0328$ oe or a correct critical region
A1 awrt 0.121 or CR $X \geq 5$
M1 need $p<0.5$ and:
correct statement using their Probability and 0.05 if one tail test or
correct statement using their Probability and 0.025 if two tail test (condone a comparison with 0.05 instead of 0.025 for a two tail test).
Do not allow non-contextual conflicting statements eg "significant" and "accept H_0"
A1ft correct contextual statement followed through from "their prob".
Either a comment on whether the teacher's claim was correct or on whether the student was guessing the answers.

NB if a correct contextual statement only is given for their probability then award M1 A1

If $p>0.5$
They may compare with 0.95 (one tail method) or 0.975 (two tail method)
Probability is 0.8791.

Question Number	Scheme	Marks
3. **(a)**	$\mathrm{E}(X)=\dfrac{3-1}{2}=1$	B1 cao (1)
(b)	$\mathrm{Var}(X)=\dfrac{(3+1)^2}{12}=\dfrac{4}{3}$ oe	M1A1 (2)
(c)	$\mathrm{E}(X^2)=\dfrac{4}{3}+1,\ =\dfrac{7}{3}$ oe	M1,A1 (2)
(d)	$\mathrm{P}(X<1.4)=0.6$	B1 cao (1)
(e)	$\mathrm{P}(X<0)=0.25$ Y is number of values less than 0 $Y\sim \mathrm{Bin}(40,0.25)$ $\mathrm{P}(Y\geq 10)=1-\mathrm{P}(Y\leq 9)$ $=1-0.4395=0.5605$	B1 M1A1 M1 A1 (5) **[11]**
	Notes	
(b)	M1 $\dfrac{(3-1)^2}{12}$ *or* $\dfrac{(3+1)^2}{12}$ *or* $\dfrac{(3--1)^2}{12}$ A1 awrt 1.33	
(c)	M1 "their(b)" + ["their (a)"]^2 or $\int_{-1}^{3}\dfrac{x^2}{4}\mathrm{d}x$ A1 awrt 2.33	
(e)	B1 For writing or using the probability of a negative = 0.25 M1 Writing or use of B(40, p) A1 Writing or use of B(40, 0.25) M1 Writing or using $1-\mathrm{P}(Y\leq 9)$ A1 awrt 0.561 or 0.560	

Question Number	Scheme	Marks
4.	H_0: $\lambda = 8$ or $\mu = 2$ H_1: $\lambda < 8$ or $\mu < 2$ Under H_0, $X \sim \text{Po}(8)$ $P(X \leq 3) = 0.0424$ CR $X \leq 3$ $0.0424 < 0.05$, Reject H_0. Richard's claim is supported.	B1 B1 M1 A1 M1A1ft **[6]**
	Notes	
	B1 for H_0 correct. Must use λ or μ and 8 or 2 B1 for H_1 correct. Must use λ or μ and 8 or 2 M1 for writing or using Po(8) – may be implied by correct CR A1 awrt 0.0424 or CR $X \leq 3$ M1 need p<0.5 and: correct statement using their Probability and 0.05 if one tail test or correct statement using their Probability and 0.025 if two tail test (condone a comparison with 0.05 instead of 0.025 for a two tail test). Do not allow non-contextual conflicting statements eg "significant" and "accept H_0" A1ft correct contextual statement followed through from "their prob". Either a comment on whether Richard's claim was correct or on whether the service has improved. NB if a correct contextual statement only is given for their probability then award M1 A1 p>0.5 They may compare with 0.95 (one tail method) or 0.975 (two tail method) Probability is 0.9576	

Question Number	Scheme	Marks
5. **(a)**	$m = -\frac{4}{0.5} = -8$ $\mathrm{f}(x) = 4 - 8x\,(*)$ $\mathrm{f}(x) = \begin{cases} -8x+4 & 0 \le x \le 0.5 \\ 0 & \textit{otherwise} \end{cases}$	M1 A1cso B1 B1 (4)
(b)	$\mathrm{F}(x) = \int_0^x (-8x+4)\mathrm{d}x$ $= \left[-4x^2+4x\right]_0^x$ $\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ -4x^2+4x & 0 \le x \le 0.5 \\ 1 & x > 0.5 \end{cases}$	M1 M1 A1 B1 (4)
(c)	$-4x^2+4x = 0.5$ $x = \frac{1}{4}(2-\sqrt{2}) = 0.146$	M1 M1A1 (3)
(d)	$x = 0$	B1 (1)
(e)	Positive Skew as mode<median	B1ft (1) **[13]**

Question Number	Scheme	Marks
	Notes	
(a)	M1 for $\pm\frac{4}{0.5}$ or attempt at gradient A1cso for proceeding to given expression with no incorrect working seen B1 for top line. Must have f(x) and { and more than one line. Condone use of <. B1 for 0 otherwise and no other parts.	
(b)	M1 attempting to integrate (at least one $x^n \rightarrow x^{n+1}$) (ignore limits) M1 correct limits used or +C and either F(0) = 0 or F(0.5) = 1, may be implied by seeing $4x - 4x^2$ A1 middle line. May write $4x - 4x^2$ B1 top and bottom line	
(c)	M1 Their F(x) = 0.5 M1 attempting to solve – either correct use of quadratic formula or correct completion of the square A1 awrt 0.146 or $\frac{2-\sqrt{2}}{4}$ o.e	
(d)	B1 for 0	
(e)	B1 ft their mode and median. Need direction and correct corresponding reason OR B1 positive skew from tail on right hand side in diagram	

Question Number	Scheme	Marks
6.		
(a)	$X \sim \text{Po}(2.5)$	M1A1 (2)
(b)	Cars arrive at the toll booth independently/randomly Cars arrive one at a time The rate of arrival at a toll booth remains constant at 2.5 per minute	 B1 B1 (2)
(c)(i)	$\text{P}(X = 0) = \text{e}^{-2.5} = 0.0821$	B1 (1)
(c)(ii)	$\text{P}(X > 3) = 1 - \text{P}(X \leq 3)$ $= 0.2424$	M1 A1 (2)
(d)	Use of Po(10) $1 - 0.0487 = 0.9513$ $m = 15$	M1 M1 A1 cao (3)
(e)	$Y \sim \text{N}(25, 25)$ $\text{P}(X < 15) = \text{P}(Y \leq 14.5)$ $= \text{P}\left(Z \leq \frac{14.5 - 25}{5}\right)$ $= \text{P}(Z \leq -2.1)$ $= 0.01786$	B1B1 M1 M1 A1 A1 (6) **[16]**

Question Number	Scheme	Marks
	Notes	
(a)	M1 Poisson A1 2.5	
(b)	Any two of the statements or equivalent. At least one must be in context. Need words that imply "cars arrive" or "rate of arrival." **SC** no context but 2 correct reasons B1B0 No context but 1 correct reason B0B0	
(c) (i)	B1 awrt 0.0821	
(ii)	M1 for writing or finding 1 - P($X \leq 3$)	
(d)	A1 awrt 0.242 M1 writing or using Po(10) M1 for 1- 0.0487 or 0.9513 seen or implied by correct value for *m*	
(e)	B1 use of normal B1 using or seeing mean and variance of 25 These first two marks may be given if the following are seen in the correct places in the standardisation formula : 25 and $\sqrt{25}$ or 5 M1 for attempting a continuity correction (14 $\pm$ 0.5) or (15 $\pm$ 0.5) M1 for standardising using their mean and their standard deviation and using [14.5, 14, 13.5, 15 or 15.5] accept $\pm$ z. A1 correct z value ± 2.1 **or** $\pm\frac{14.5-25}{5}$, A1 awrt 0.0179 NB use of calculator gets full marks if the answer is awrt 0.0179.	

Question Number	Scheme	Marks
7. **(a)**	$\int_0^9 k(81x - x^3)\mathrm{d}x = 1$ $k\left[\frac{81}{2}x^2 - \frac{1}{4}x^4\right]_0^9 = 1$ $k(\frac{6561}{2} - \frac{6561}{4}) = 1$ $k = \frac{4}{6561}$ **ag**	M1 M1 A1 cso (3)
(b)	$\mathrm{E}(X) = \int_0^9 kx^2(81 - x^2)\,\mathrm{d}x$ $= k\left[\frac{81}{3}x^3 - \frac{x^5}{5}\right]_0^9$ $= k(19683 - 11809.8)$ $= 4.8$	M1A1 dM1 A1 cao (4)
(c)	$\mathrm{P}(X > 5) = \int_5^9 k(81x - x^3)$ $= k\left[\frac{81}{2}x^2 - \frac{1}{4}x^4\right]_5^9$ $= k\left(\frac{6561}{4} - 856.25\right) =$ awrt 0.478 or $\frac{3136}{6561}$	M1 M1d A1 (3)
(d)	P(At least 2 queue for more than 5 mins) $= 3(1\text{-}0.478)(0.478)^2 + 0.478^3$ $= 0.467$	M1A1ft A1 (3) **[13]**

Question Number	Scheme	Marks
	Notes	
(a)	M1 putting integral = 1 ignore limits. =1 must appear at least once in the working. M1 attempting to integrate **at least one** part must have correct power of x (ignore limits) A1cso subst of at least 9. Allow 1/1640.25	
(b)	M1 attempt to use xf(x) and attempt to multiply out bracket and attempt at integration – must have x^3 and x^5 terms (ignore limits) A1 correct integration (ignore limits) dM1 substituting correct limits (need not explicitly see 0). Dependent on having been awarded the first M1.	
(c)	M1 attempting to integrate **at least one** part must have correct power of x (ignore limits) M1 dep on previous M being awarded, substituting correct limits [may use $1 - \int_0^5 k(81x - x^3)$ with limits 0 and 5]	
(d)	M1 $3(1-p)p^2 + p^3$ o r $1-(1-p)^3 - 3(1-p)^2 p$ 3 not needed A1 for $\mathbf{3}(1-p)p^2 + p^3$ $1-(1-p)^3 - \mathbf{3}(1-p)^2 p$ where p is their solution to part (c) A1 awrt 0.467	