Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Tuesday 19 January 2010 – Morning

Time: 1 hour 30 minutes

Materials required for examination
Mathematical Formulae (Pink or Green)

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 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. A continuous random variable X has cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < -2 \\ \dfrac{x+2}{6}, & -2 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

(*a*) Find $\mathrm{P}(X < 0)$.

(2)

(*b*) Find the probability density function $\mathrm{f}(x)$ of X.

(3)

(*c*) Write down the name of the distribution of X.

(1)

(*d*) Find the mean and the variance of X.

(3)

(*e*) Write down the value of $\mathrm{P}(X = 1)$.

(1)

---

3. A robot is programmed to build cars on a production line. The robot breaks down at random at a rate of once every 20 hours.

(*a*) Find the probability that it will work continuously for 5 hours without a breakdown.

(3)

Find the probability that, in an 8 hour period,

(*b*) the robot will break down at least once,

(3)

(*c*) there are exactly 2 breakdowns.

(2)

In a particular 8 hour period, the robot broke down twice.

(*d*) Write down the probability that the robot will break down in the following 8 hour period. Give a reason for your answer.

(2)

---

4. The continuous random variable X has probability density function f(x) given by

$$\mathrm{f}(x) = \begin{cases} k(x^2 - 2x + 2), & 0 < x \leq 3, \\ 3k, & 3 < x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

where k is a constant.

(*a*) Show that $k = \frac{1}{9}$.

(4)

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

(6)

(*c*) Find the mean of X.

(3)

(*d*) Show that the median of X lies between $x = 2.6$ and $x = 2.7$.

(4)

---

5. 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)

---

6. (*a*) Define the critical region of a test statistic.

(2)

A discrete random variable X has a Binomial distribution B(30, p). A single observation is used to test $H_0 : p = 0.3$ against $H_1 : p \neq 0.3$

(*b*) Using a 1% level of significance find the critical region of this test. You should state the probability of rejection in each tail which should be as close as possible to 0.005.

(5)

(*c*) Write down the actual significance level of the test.

(1)

The value of the observation was found to be 15.

(*d*) Comment on this finding in light of your critical region.

(2)

---

7. A bag contains a large number of coins. It contains only 1p and 2p coins in the ratio 1:3.

(*a*) Find the mean μ and the variance σ^2 of the values of this population of coins.

(3)

A random sample of size 3 is taken from the bag.

(*b*) List all the possible samples.

(2)

(*c*) Find the sampling distribution of the mean value of the samples.

(6)

TOTAL FOR PAPER: 75 MARKS

END

January 2010
6684 Statistics S2
Mark Scheme

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)	$P(X > 4)$ $= 1 - P(X \leq 4)$ $= 1 - 0.9974$ $= 0.0026$	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]**

Notes

Q1 (a) **1st B1** for binomial
2nd B1 for 20 and 0.05 o.e
These must be in part (a)

(b) **M1** for finding $(p)^{20}$ $0< p <1$ this working needs to be seen if answer incorrect to gain the M1
A1 awrt 0.358 or 0.359.

(c) **M1** for writing $1 - P(X \leq 4)$
or $1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)]$
or $1 - 0.9974$
or $1 - 0.9568$
A1 awrt 0.0026 or 2.6×10^{-3}, do not accept a fraction e.g. 26/10000

(d) **1st B1** for 1
2nd B1 for 0.95

NB In parts b, c and d correct answers with no working gain full marks

Question Number		Scheme	Marks
Q2	(a)	$P(X < 0) = F(0)$	M1
		$= \frac{2}{6} = \frac{1}{3}$	A1 (2)
	(b)	$f(x) = \frac{dF(x)}{dx}$	M1
		$f(x) = \begin{cases} \frac{1}{6} & -2 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$	A1 B1 (3)
	(c)	Continuous Uniform (Rectangular) distribution	B1 (1)
	(d)	Mean = 1	B1
		Variance is $\frac{(4--2)^2}{12} = 3$	M1 A1 (3)
	(e)	$P(X = 1) = 0$	B1 (1)
			Total [10]
Q2	(a)	**Notes** **M1** for attempting to find F(0) by a correct method eg subst 0 into F(x) or $\int_{-2}^{0} \frac{1}{6} dx$ Do **NOT** award M1 for $\int_{-2}^{0} \frac{x+2}{6} dx$ or $\frac{1}{2} \times \frac{1}{3} \times 2$ both of which give the correct answer by using F(x) as the pdf **A1** 1/3 o.e or awrt 0.333 Correct answer only with no incorrect working gets M1 A1	
	(b)	**M1** for attempting to differentiate F(x). (for attempt it must have no xs in) **A1** for the first line. Condone < signs **B1** for the second line. – They must have 0 $x < -2$ and $x > 4$ only.	
	(c)	**B1** must have "continuous" and "uniform" **or** "Rectangular"	
	(d)	**B1** for mean = 1 **M1** for attempt to use $\frac{[\pm(b-a)]^2}{12}$, they must subst in values and not just quote the formula, or using $\int_{-2}^{4} x^2 (their\ f(x)) - (their\ mean)^2$, including limits. Must get x^3 **when they integrate.** **A1** cao .	
	(e)	**B1** cao	

Question Number	Scheme	Marks
Q3 (a)	$Y \sim \text{Po}(0.25)$	B1
	$\text{P}(Y=0) = e^{-0.25}$	M1
	$= 0.7788$	A1
		(3)
(b)	$X \sim \text{Po}(0.4)$	B1
	$\text{P(Robot will break down)} = 1 - \text{P}(X = 0)$	
	$= 1 - e^{-0.4}$	M1
	$= 1 - 0.067032$	
	$= 0.3297$	A1
		(3)
(c)	$\text{P}(X = 2) = \frac{e^{-0.4}(0.4)^2}{2}$	M1
	$= 0.0536$	A1
		(2)
(d)	0.3297 or answer to part (b) as Poisson events are <u>independent</u>	B1ft B1 dep
		(2)
		Total [10]

<u>Notes</u>

Q3 (a) **B1** for seeing or using Po(0.25)

M1 for finding P(Y=0) either by e^{-a}, where a is positive (a needn't equal their λ) or using tables if their value of λ is in them

Beware common Binomial error using, $p = 0.05$ gives 0.7738 but scores B0 M0 A0

A1 awrt 0.779

(b) **B1** for **stating or a clear use of** Po(0.4) in part (b) or (c)

M1 for writing or finding 1 – P(X =0)

A1 awrt 0.33

(c) **M1** for finding P(X =2) e.g $\frac{e^{-\lambda}\lambda^2}{2!}$ with their value of λ in

or if their λ is in the table for writing P($X \leq 2$) - P($X \leq 1$)

A1 awrt 0.0536

(d) **1st B1** their answer to part(b) correct to 2 sf or awrt 0.33

2nd B1 need the word independent. This is dependent on them gaining the first B1

SC

Use of Binomial.

Mark parts a and b as scheme. They could get (a) B0,M0,A0 (b) B0 M1 A0

In part c allow M1 for $^nC_2$ $(p)^2(1-p)^{n-2}$ with "their n" and "their p". They could get (c) M1,A0

DO NOT GIVE for $p(x \leq 2) - p(x \leq 1)$

In (d) they can get the first B1 only. They could get (d) B1B0

Question Number	Scheme	Marks
Q4 (a)	$\int_0^3 k(x^2-2x+2)\mathrm{d}x+\int_3^4 3k\mathrm{d}x \quad =1$	M1
	$k\left[\frac{1}{3}x^3-x^2+2x\right]_0^3+[3kx]_3^4 \quad (=1)$ or $k\left[\frac{1}{3}x^3-x^2+2x\right]_0^3+3k \quad (=1)$	A1 M1 dep
	$9k \quad =1$	
	$k \quad =\frac{1}{9}$ **given** cso	A1
		(4)
(b)	For $0<x\le 3$, $\mathrm{F}(x) \quad =\int_0^x \frac{1}{9}(t^2-2t+2)\mathrm{d}t$	M1
	$=\frac{1}{9}\left(\frac{1}{3}x^3-x^2+2x\right)$	A1
	For $3<x\le 4$, $\mathrm{F}(x) \quad =\int_3^x 3k\mathrm{d}t+\frac{2}{3}$	M1
	$=\frac{x}{3}-\frac{1}{3}$	A1
	$\mathrm{F}(x)=\begin{cases}0 & x\le 0\\ \frac{1}{27}(x^3-3x^2+6x) & 0<x\le 3\\ \frac{x}{3}-\frac{1}{3} & 3<x\le 4\\ 1 & x>4\end{cases}$	B1 ft B1
		(6)
(c)	$\mathrm{E}(X) \quad =\int_0^3 \frac{x}{9}(x^2-2x+2)\mathrm{d}t+\int_3^4 \frac{x}{3}\mathrm{d}x$	M1
	$=\frac{1}{9}\left[\frac{1}{4}x^4-\frac{2}{3}x^3+x^2\right]_0^3+\left[\frac{1}{6}x^2\right]_3^4$	A1
	$=\frac{29}{12}$ or 2.416 or awrt 2.42	A1
		(3)
(d)	F(m) = 0.5	M1
	$\mathrm{F}(2.6)=\frac{1}{27}(2.6^3-3\times 2.6^2+6\times 2.6)=$ awrt 0.48	M1
	$\mathrm{F}(2.7)=\frac{1}{27}(2.7^3-3\times 2.7^2+6\times 2.7)=$ awrt 0.52	A1
	Hence median lies between 2.6 and 2.7	A1 dA
		(4)
		Total [17]

		Notes	
Q4	(a)	**1st M1** attempting to integrate **at least one** part (at least one $x^n \rightarrow x^{n+1}$) (ignore limits) **1st A1 Correct integration.** Limits not needed. **2nd M1** dependent on the previous M being awarded. Adding the two answers together, putting equal to 1 and have the correct limits. **2nd A1** cso	
	(b)	**1st M1** Att to integrate $\frac{1}{9}\left(t^2 - 2t + 2\right)$ (at least one $x^n \rightarrow x^{n+1}$). Ignore limits for method mark **1st A1** $\frac{1}{9}\left(\frac{x^3}{3} - x^2 + 2x\right)$ allow use of *t*. Must have used/implied use of limit of 0. This must be on its own without anything else added **2nd M1** attempting to find $\int_3^x 3k + \ldots$ (must get $3kt$ or $3kx$) and they must use the correct limits and add $\int_0^3 \frac{1}{9}\left(t^2 - 2t + 2\right)$ or $\frac{2}{3}$ or use + C and use F(4) = 1 **2nd A1** $\frac{x}{3} - \frac{1}{3}$ must be correct **1st B1** middle pair followed through from their answers. condone them using < or $\leq$ incorrectly they do not need to match up **2nd B1** end pairs. condone them using < or $\leq$. They do not need to match up NB if they show no working and just write down the distribution. If it is correct they get full marks. If it is incorrect then they cannot get marks for any incorrect part. So if $0<x\leq 3$ is correct they can get M1 A1 otherwise M0 A0. If $3<x\leq 4$ is correct they can get M1 A1 otherwise M0 A0. you cannot award B1ft if they show no working unless the middle parts are correct.	
	(c)	**1st M1** attempting to use integral of *x* f(*x*) on one part **1st A1** Correct Integration for both parts added together. Ignore limits. **2nd A1** cao or awrt 2.42	
	(d)	**1st M1** for using F(*X*) = 0.5. This may be implied by subst into F(*X*) and comparing answers with 0.5. **2nd M1** for substituting both 2.6 and 2.7 into "their F(*X*)" – 0.5 or "their F(*X*)" **1st A1** awrt 0.48 and 0.52 if using "their F(*X*)" . and awrt -0.02 and 0.02 or if using "their F(X)" 0.5 Other values possible. You may need to check their values for their **correct** equation **NB these last two marks are B1 B1 on ePEN but mark as M1 A1** **2nd A1** for conclusion but only award if it follows from their numbers. Dependent on previous A mark being awarded **SC** using calculators **M1** for sign of a suitable equation **M1 A1** for awrt 2.66 provided equation is correct **A1** correct comment	

Question Number	Scheme	Marks
Q5 (a)	$X \sim \text{Po}(10)$	B1
	$\text{P}(X < 9) = \text{P}(X \leq 8)$	M1
	$= 0.3328$	A1
		(3)
(b)	$Y \sim \text{Po}(40)$	
	Y is approximately N(40,40)	M1 A1
	$\text{P}(Y > 50) = 1 - \text{P}(Y \leq 50)$	
	$= 1 - \text{P}\left(Z < \frac{50.5 - 40}{\sqrt{40}}\right)$	M1
		M1
	$= 1 - \text{P}(Z < 1.660..)$	A1
	$= 1 - 0.9515$	
	$= 0.0485$	A1
		(6)
		Total [9]
	N.B. Calculator gives 0.048437. Poisson gives 0.0526 (but scores nothing)	
	Notes	
Q5 (a)	**B1** for using Po(10)	
	M1 for attempting to find $\text{P}(X \leq 8)$: useful values $\text{P}(X \leq 9)$ is 0.4579(M0), using Po(6) gives 0.8472, (M1).	
	A1 awrt 0.333 but do not accept $\frac{1}{3}$	
(b)	**1st M1** for identifying the normal approximation	
	1st A1 for [mean = 40] and [sd = $\sqrt{40}$ or var = 40]	
	NB These two marks are B1 M1 on ePEN	
	These first two marks may be given if the following are seen in the standardisation formula : 40 and $\sqrt{40}$ or awrt 6.32	
	2nd M1 for attempting a continuity correction (50 or 30 $\pm$ 0.5 is acceptable)	
	3rd M1 for standardising using their mean and their standard deviation and using either 49.5, 50 or 50.5. (29.5, 30, 30.5) accept $\pm$	
	2nd A1 correct z value awrt $\pm$1.66 **or** this may be awarded if see $\pm\frac{50.5 - 40}{\sqrt{40}}$ or $\pm\frac{29.5 - 40}{\sqrt{40}}$	
	3rd A1 awrt 3 sig fig in range 0.0484 – 0.0485	

Question Number		Scheme	Marks
Q6	(a)	The set of values of the test statistic for which the null hypothesis is rejected in a hypothesis test.	B1 B1 (2)
	(b)	X~B(30,0.3) $P(X \leq 3) = 0.0093$ $P(X \leq 2) = 0.0021$ $P(X \geq 16) = 1 - 0.9936 = 0.0064$ $P(X \geq 17) = 1 - 0.9979 = 0.0021$ Critical region is $(0 \leq)x \leq 2$ or $16 \leq x (\leq 30)$	M1 A1 A1 A1A1 (5)
	(c)	Actual significance level 0.0021+0.0064=0.0085 or 0.85%	B1 (1)
	(d)	15 (it) is not in the critical region not significant No significant evidence of a change in $p = 0.3$ accept H_0, (reject H_1) $P(x \geq 15) = 0.0169$	Bft 2, 1, 0 (2) **Total [10]**

Notes

Q6 (a)

1st B1 for "**values/ numbers**"
2nd B1 for "reject the null hypothesis" o.e or the test is significant

(b)

M1 for using B(30,0.3)
1st A1 $P(x \leq 2) = 0.0021$
2nd A1 0.0064

3rd A1 for $(X) \leq 2$ or $(X) < 3$ They get **A0** if they write **P**($X \leq 2$/ $X < 3$)
4th A1 $(X) \geq 16$ or $(X) > 15$ They get **A0** if they write **P**($X \geq 16$ $X > 15$
NB these are B1 B1 but mark as A1 A1

$16 \leq X \leq 2$ etc is accepted
To describe the critical regions they can use any letter or no letter at all. It does not have to be X.

(c) **B1** correct answer only

(d) Follow through 15 and their critical region
B1 for any one of the 5 correct statements up to a maximum of B2
– **B1** for any incorrect statements

Question Number	Scheme	Marks
Q7 (a)	$\mu = 1\times\frac{1}{4}+2\times\frac{3}{4}=\frac{7}{4}$ or $1\frac{3}{4}$ or 1.75	B1
	$\sigma^2 = 1^2\times\frac{1}{4}+2^2\times\frac{3}{4}-\left(\frac{7}{4}\right)^2$	M1
	$=\frac{3}{16}$ or 0.1875	A1 (3)
(b)	(1,1,1) , (1,1,2) any order, (1,2,2) any order, (2,2,2)	B1
	(1,2,1) (2,1,1) (2,1,2) (2,2,1) all 8 cases considered. May be implied by 3 * (1,1,2) and 3*(1,2,2)	B1 (2)
(c)	$\bar{x}$: 1 ; $\frac{4}{3}$; $\frac{5}{3}$; 2 $P(\bar{X}=\bar{x})$: $\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$; $3\times\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{9}{64}$; $3\times\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{27}{64}$; $\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{27}{64}$	B1 M1 A1 M1 A1A1 (6)
		Total [11]

Notes

Q7 (a) B1 1.75 oe
M1 for using $\sum\left(x^2 p\right)-\mu^2$
A1 0.1875 oe

(b) ignore repeats

(c) 1st B1 4 correct means (allow repeats)
1st M1 for p^3 for either of the ends
1st A1 for 1/64or awrt 0.016 **and** 27/64 or awrt 0.422
2nd M1 $3\times p^2(1-p)$ for either of the middle two $0 < p < 1$
May be awarded for finding the probability of the 3 samples with mean of either 4/3 or 5/3 .
2nd A1 for 9/64 (or 3/64 three times) and 27/64 (or 9/64 three times) accept awrt 3dp.
3rd A1 fully correct table, accept awrt 3dp.