Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced Level

Friday 24 May 2013 – 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 bag contains a large number of 1p, 2p and 5p coins.

50% are 1p coins

20% are 2p coins

30% are 5p coins

A random sample of 3 coins is chosen from the bag.

(*a*) List all the possible samples of size 3 with median 5p.

(2)

(*b*) Find the probability that the median value of the sample is 5p.

(4)

(*c*) Find the sampling distribution of the median of samples of size 3.

(5)

---

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. An online shop sells a computer game at an average rate of 1 per day.

(*a*) Find the probability that the shop sells more than 10 games in a 7 day period.

(3)

Once every 7 days the shop has games delivered before it opens.

(*b*) Find the least number of games the shop should have in stock immediately after a delivery so that the probability of running out of the game before the next delivery is less than 0.05.

(3)

In an attempt to increase sales of the computer game, the price is reduced for six months. A random sample of 28 days is taken from these six months. In the sample of 28 days, 36 computer games are sold.

(*c*) Using a suitable approximation and a 5% level of significance, test whether or not the average rate of sales per day has increased during these six months. State your hypotheses clearly.

(7)

---

4. A continuous random variable X is uniformly distributed over the interval $[b, 4b]$ where b is a constant.

(*a*) Write down $\mathrm{E}(X)$.

(1)

(*b*) Use integration to show that $\mathrm{Var}(X) = \frac{3b^2}{4}$.

(3)

(*c*) Find $\mathrm{Var}(3 - 2X)$.

(2)

Given that $b = 1$, find

(*d*) the cumulative distribution function of X, $\mathrm{F}(x)$, for all values of x,

(2)

(*e*) the median of X.

(1)

---

5. The continuous random variable X has a cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0, & x < 1, \\ \dfrac{x^3}{10} + \dfrac{3x^2}{10} + ax + b, & 1 \leq x \leq 2, \\ 1, & x > 2, \end{cases}$$

where a and b are constants.

(*a*) Find the value of a and the value of b.

(4)

(*b*) Show that $\mathrm{f}(x) = \dfrac{3}{10}(x^2 + 2x - 2), \quad 1 \leq x \leq 2.$

(1)

(*c*) Use integration to find $\mathrm{E}(X)$.

(4)

(*d*) Show that the lower quartile of X lies between 1.425 and 1.435.

(3)

---

6. In a manufacturing process 25% of articles are thought to be defective. Articles are produced in batches of 20.

(*a*) A batch is selected at random. Using a 5% significance level, find the critical region for a two tailed test that the probability of an article chosen at random being defective is 0.25.

You should state the probability in each tail, which should be as close as possible to 0.025.

(5)

The manufacturer changes the production process to try to reduce the number of defective articles. She then chooses a batch at random and discovers there are 3 defective articles.

(*b*) Test at the 5% level of significance whether or not there is evidence that the changes to the process have reduced the percentage of defective articles. State your hypotheses clearly.

(5)

---

7. A telesales operator is selling a magazine. Each day he chooses a number of people to telephone. The probability that each person he telephones buys the magazine is 0.1.

(*a*) Suggest a suitable distribution to model the number of people who buy the magazine from the telesales operator each day.

(1)

(*b*) On Monday, the telesales operator telephones 10 people. Find the probability that he sells at least 4 magazines.

(3)

(*c*) Calculate the least number of people he needs to telephone on Tuesday, so that the probability of selling at least 1 magazine, on that day, is greater than 0.95.

(3)

A call centre also sells the magazine. The probability that a telephone call made by the call centre sells a magazine is 0.05. The call centre telephones 100 people every hour.

(*d*) Using a suitable approximation, find the probability that more than 10 people telephoned by the call centre buy a magazine in a randomly chosen hour.

(3)

TOTAL FOR PAPER: 75 MARKS

END

Question Number	Scheme	Marks
1(a)	(5,5,5) or (1,5,5) or (2,5,5)	B1
	(5,5,5) (5,5,1) (5,1,5) (1,5,5) (5,5,2) (5,2,5) (2,5,5) or (5,5,5) **and** (5,5,1) ($\times$ 3) **and** (5,5,2) ($\times$ 3)	B1 (2)
1(b)	(5,5,5) $\left(\frac{3}{10}\right)^3 = \frac{27}{1000} = 0.027$	B1
	(5,5,1) $3\times\frac{1}{2}\times\left(\frac{3}{10}\right)^2 = \frac{135}{1000}$ *or* $\frac{27}{200} = 0.135$	M1
	(5,5,2) $3\times\frac{1}{5}\times\left(\frac{3}{10}\right)^2 = \frac{54}{1000} = \frac{27}{500} = 0.054$	
	$\text{P}(M=5) = \left(\frac{3}{10}\right)^3 + 3\times\frac{1}{2}\times\left(\frac{3}{10}\right)^2 + 3\times\frac{1}{5}\times\left(\frac{3}{10}\right)^2 = \frac{27}{125} = 0.216$ oe	A1A1 (4)
1(c)	$\text{P}(M=1) = (0.5)^3 + 3(0.5)^2(0.2) + 3(0.5)^2(0.3)$	M1
	$= 0.5$	A1
	$\text{P}(M=2) = \left(\frac{1}{5}\right)^3 + 3\times\left(\frac{1}{5}\right)^2\times\frac{1}{2} + 3\times\left(\frac{1}{5}\right)^2\times\frac{3}{10} + 6\times\frac{1}{2}\times\frac{1}{5}\times\frac{3}{10}$	M1
	$= 0.284$ or $\frac{71}{250}$ oe	A1
	<table><tr><td>*m*</td><td>1</td><td>2</td><td>5</td></tr><tr><td>P(*M* = *m*)</td><td>0.5</td><td>0.284</td><td>0.216</td></tr></table>	A1 (5) Total 11 marks

	Notes	
1(a)	1st B1 for two of the given triples, any order 2nd B1 for all 7 cases. no incorrect extras	
1(b)	B1 $\left(\frac{3}{10}\right)^3$ *or* 0.027 oe. This can be a single term in a summation M1 either "3" $\times\frac{1}{2}\times\left(\frac{3}{10}\right)^2$ **or** "3" $\times\frac{1}{5}\times\left(\frac{3}{10}\right)^2$ oe. May omit the 3 × or have another positive integer in place of the 3. These may be seen as a single term in a summation A1 $\left(\frac{3}{10}\right)^3 + 3\times\frac{1}{2}\times\left(\frac{3}{10}\right)^2 + 3\times\frac{1}{5}\times\left(\frac{3}{10}\right)^2$ oe A1 0.216 oe	
1(c)	1st M1 correct calculation for P(*M* = 1) **or** P(*M* = 2), working must be shown and **not** implied by a correct answer. 1st A1 either P(*M* = 1) **or** P(*M* = 2) correct 2nd M1 correct calculation for both P(*M* = 1) **and** P(*M* = 2), **or** their probabilities adding up to 1, but do not allow probabilities of 0.5, 0.2 and 0.3 2nd A1 both P(*M* = 1) **and** P(*M* = 2) correct 3rd A1dep on both M marks awarded. All three values written down with their correct probabilities. They must be in part (c) but they do not need to be in a table. **NB** A fully correct table with no working will get M0 A0 M1 A1 A0.	

Question Number	Scheme	Marks

2(a)	$P(X = 1) = 0.25e^{-0.25} = 0.1947$ awrt 0.195	M1A1 (2)
2(b)	$X \sim Po(1.5)$ $P(X > 2) = 1 - P(X \leq 2)$ $= 1 - 0.8088$ $= 0.1912$ awrt 0.191	B1 M1 A1 (3)
2(c)	$[\lambda = 300 \times 0.25 = 75]$ $X \sim N(75,75)$ $P(X < 90) = P(X \leq \frac{89.5-75}{\sqrt{75}})$ $= P(Z \leq 1.6743..)$ = awrt 0.953 or 0.952	 B1 B1 M1M1 A1 (5) Total 10 marks
	Notes	
2(a) **2(b)** **2(c)**	M1 $0.25e^{-0.25}$ o.e B1 stating or using Po(1.5) M1 stating or using $1 - P(X \leq 2)$ 1st B1 for normal approximation and correct mean 2nd B1 Var (X) = 75or sd = $\sqrt{75}$ or awrt 8.66 (may be given if correct in standardisation formula) 1st M1 using either 89.5 or 88.5 2nd M1 Standardising using their mean and their sd, using [89.5, 88.5 or 89] and for finding correct area **NB** use of Poisson gives an answer of 0.9498 and gains no marks	

Question Number	Scheme	Marks
3(a)	$X \sim \mathrm{Po}(7)$ $P(X > 10) = 1 - P(X \le 10)$ $= 1 - 0.9015$ $= 0.0985$ awrt 0.0985	B1 M1 A1 (3)
3(b)	$P(X > d) < 0.05$ **Or** $P(X \ge d) < 0.05$ $P(X \le d) > 0.95$ $P(X < d) > 0.95$ $P(X \le 11) = 0.9467$ $P(X < 12) = 0.9467$ $P(X \le 12) = 0.9730$ $P(X < 13) = 0.9730$ Least number of games =12 Least number of games 13	 M1 A1 A1 (3)
3(c)	H_0: $\lambda = 1$, $(\mu = 28)$ H_1: $\lambda > 1\ (\mu > 28)$ $Y \sim \mathrm{Po}(28)$ approximated by N(28,28) $P(Y \ge 36) = P(Z \ge \frac{35.5-28}{\sqrt{28}})$ | $1.6449 = \frac{x - 0.5 - 28}{\sqrt{28}}$ $= P(Z \ge 1.42)$ $= 0.0778$ or $1.42 < 1.6449$ | CR $X \ge 37.2$ $0.0778 > 0.05$ so do not reject H_0/not significant. Not in CR There is no evidence that the average **rate** of **sales** per day has **increased**.	B1 B1 M1M1 A1 M1 A1cso (7) Total 13 marks
	Notes	
3(a)	B1 stating or using Po(7) M1 stating or using $1 - P(X \le 10)$	
3(b)	M1 using or writing $P(X > d) < 0.05$ or $P(X < d) > 0.95$ (condone $\ge$ instead of $>$ and $\le$ instead of $<$) May be implied by correct answer. Different letters may be used. 1st A1 $P(X \le 12)$/ $P(X < 13)$ = awrt 0.973 **or** $P(X \le 11)$ / $P(X < 12)$ = awrt 0.947 May be implied by a correct answer 2nd A1 12 or 13 **NB** An answer of 12/13 on its own with no working gains M1A1A1	
3(c)	1st B1 both hypotheses correct using λ or μ, **and** 1 or 28 2nd B1 for writing or using a normal approximation with correct mean and Var (may be given if sd correct in standardisation formula) 1st M1 for use of a continuity correction 35.5 or 36.5 or $x \pm 0.5$ 2nd M1 Standardising using their mean and their sd. If they have not written down a mean and sd then these need to be correct here to award the mark. They must use [35.5, 36.5, 36, x or $x \pm 0.5$] For CR must have = awrt 1.64 or 1.65 1st A1 awrt 0.0778 or 0.9222 or the statement 1.42 < awrt 1.65/1.64 or CR $X \ge 37.2$ / $X > 37.2$ 3rd M1 a correct conclusion for their probability. May be implied by a correct contextual conclusion. NB Non contextual contradicting statements gets M0 2nd A1 a correct contextual conclusion for their hypotheses and a fully correct solution with no errors seen. Need the words **"rate/average number"**, **"sales"** and **"increased"**oe NB If found $P(X = 36)$ they can get B1B10M0A0M0A0	

Question Number	Scheme	Marks

4(a)	E(X) = $\frac{5b}{2}$	B1 (1)
4(b)	Var(X) = E(X^2) - (E(X))^2 $= \int_b^{4b} \frac{x^2}{3b} dx - (\frac{5b}{2})^2$ $= \left[\frac{x^3}{9b}\right]_b^{4b} - \frac{25b^2}{4}$ $= \frac{63b^3}{9b} - \frac{25b^2}{4}$ $= \frac{3b^2}{4}$	M1 M1d A1cso (3)
4(c)	Var(3 - 2X) = 4Var(X) $= 3b^2$	M1 A1 (2)
4(d)	$F(x) = \begin{cases} 0 & x < 1 \\ \frac{x-1}{3} & 1 \le x \le 4 \\ 1 & x > 4 \end{cases}$	B1B1 (2)
4(e)	$\frac{x-1}{3} = 0.5$ so $x = 2.5$	B1 (1)
Alt 4(b)	Var(X) $= \int_a^b \frac{(x-\bar{x})^2}{b-a} dx$ $= \int_b^{4b} \frac{4x^2-20bx+25b^2}{12b} dx$ $= \left[\frac{\frac{4x^3}{3}-10bx^2+25b^2x}{12b}\right]_b^{4b}$ $= \frac{9b^3}{12b}$ $= \frac{3b^2}{4}$	Total 9 marks M1 M1 A1cso(3)
	Notes	
4(b)	**NB remember the answer is given (AG) so they must show their working** 1st M1 for using $\int \frac{x^2}{3b}$ dx - (their (a))^2 limits not needed and condone missing dx. NB need not use the letter x but if they use b instead do not award if they cancel down to $\frac{b}{3}$ NB Check they have subtracted (their(a))^2 2nd M1 dependent on previous M being awarded. For some correct integration $x^n \to x^{n+1}$ and correct limits substituted at some point. condone 4b^3 instead of (4b)^3 A1 for correct solution with no incorrect working seen.	
4(c)	M1 for writing or using 4Var(X)	
4(d)	1st B1 top **and** bottom line. Allow use of $\le$ instead of < and $\ge$ instead of > 2nd B1 middle row. Allow use of < instead of $\le$	

Question Number	Scheme	Marks
5(a)	F(1) = 0, $\frac{4}{10} + a + b = 0$	M1 A1

	$a = -\frac{3}{5}$ *or* $b = \frac{1}{5}$ F(2) = 1, $2 + 2a + b = 1$ Solving gives $a = -\frac{3}{5}, b = \frac{1}{5}$ **Alt** F(2) - F(1) = 1, $2 + 2a + b - \frac{4}{10} - a - b = 1$ $a = -\frac{3}{5}$ F(2) =1 **or** F(1) = 0 $2 - \frac{6}{5} + b = 1$ **or** $\frac{4}{10} - \frac{3}{5} + b = 0$ $b = \frac{1}{5}$	 M1 A1 (4) M1 A1 M1 A1 (4)
5(b)	Differentiating cdf gives f(x) = $\frac{3}{10}x^2 + \frac{6}{10}x + a$, $1 \le x \le 2$ $= \frac{3}{10}(x^2 + 2x - 2)$	B1 cso (1)
5(c)	$\mathrm{E}(X) = \int_1^2 \frac{3}{10}(x^3 + 2x^2 - 2x)\mathrm{d}x$ $= \frac{3}{10}\left[\frac{1}{4}x^4 + \frac{2}{3}x^3 - x^2\right]_1^2$ $= \frac{13}{8}$	M1 M1d A1 A1 (4)
5(d)	F(1.425) = 0.24355, F(1.435) = 0.25227 0.25 lies between F(1.425)and F(1.435) hence result.	M1A1 A1 (3)
	Notes	Total 12 marks
5(a)	1st M1 using F(1) = 0. Clear attempt to form a linear equation for *a* and *b* 1stA1 either *a* = -0.6 or *b* = 0.2 Previous M must be awarded 2nd M1 using F(2) = 1. Clear attempt to form a second linear equation for *a* and *b* 2nd A1 if 1st A1 awarded then both *a* and *b* must be correct otherwise award if either *a* = -0.6 or *b* = 0.2 **alt** 1st M1 F(2) - F(1) = 1. Leading to a value for *a:* 1st A1 *a* = -0.6 2nd M1 using F(2) = 1 or F(1) = 0. Leading to a value for *b:* 2nd A1 *b* = 0.2 NB correct values for *a* and *b* with no working scores no marks.	
5(b)	B1 They must differentiate and then factorise. cso	
5(c)	1st M1 for clear attempt to use *x*f(*x*) with an intention of integrating (Integral sign enough) Ignore limits. Must substitute in f(*x*) or "their f(*x*)". 2nd M1d dependent on previous M being awarded for some correct integration… at least one correct term with the correct coefficient. 1st A1 for fully correct (possibly unsimplified) integration. Ignore limits 2nd A1 Accept 1.63 and 1.625 or some other exact equivalent	
5(d)	M1 expression showing substitution of 1.425 or 1.435 into F(*x*) [**or** into F(*x*) – 0.25] [or putting their F(*x*) = 0.25 and attempting to solve leading to *x* = ...] May be implied by either pair of the correct answers as given below for the 1st A1 1st A1 awrt 0.244 and awrt 0.252 [**or** awrt -0.00645 and awrt 0.00227] [or *x* = awrt 1.432] 2nd A1 0.25 lies between F(1.425)and F(1.435) [**or** change in sign therefore root between] [**or** "1.432" lies between 1.425 and 1.435 therefore root between]. Statement must be true for their method	

Question Number	Scheme	Marks
6(a)	$X \sim \mathrm{B}(20, 0.25)$ $\mathrm{P}(X \ge 10) = 1 - 0.9861 = 0.0139$ $\mathrm{P}(X \le 1) = 0.0243$	M1 A1 A1

	$(0 \leq) X \leq 1 \cup 10 \leq X (\leq 20)$	A1A1 (5)
6(b)	$H_0: p = 0.25$ $H_1: p < 0.25$ $X \sim B(20, 0.25)$ $P(X \leq 3) = 0.2252$ or CR $X \leq 1$ Insufficient evidence to reject H_0, Accept H_0, Not significant. 3 does not lie in the Critical region. No evidence that the **changes** to the process have **reduced** the **percentage** of **defective articles (oe)**	 B1 M1A1 M1d A1cso (5) Total 10 marks
	Notes	
6(a)	M1 using B(20,0.25) may be implied by a correct CR (allow written as a probability statement) 1st A1 awrt 0.0139 2nd A1 awrt 0.0243 3rd A1 $X \leq 1$ or $0 \leq X \leq 1$ or [0,1] or 0,1 or equivalent statements 4th A1 $X \geq 10$ or $10 \leq X \leq 20$ or 10,11,12,13,14,15,16,17,18,19,20 or [10,20] or equivalent statements **NB** These two A marks must be for statements with *X* (any letter) only – not in probability statements and **SC** for CR written as $1 \geq X \geq 10$ gets A1 A0	
6(b)	B1 both hypotheses with *p* 1st M1 using B(20, 0.25) and finding $P(X \leq 3)$ or $P(X \geq 4)$ may be implied by a correct CR 1st A1 0.2252 (allow 0.7748) if not using CR or CR $X \leq 1$ or $X < 2$ 2nd M1dependent on previous M being awarded. A correct statement (do not allow if there are contradicting non contextual statements) A1cso Conclusion must contain the words **changes/new process oe**, **reduced oe number/percentage oe** , and **defective articles/defectives**. There must be no incorrect working seen.	

Question Number	Scheme	Marks
7(a)	Distribution $X \sim B(n,0.1)$	B1 (1)
7(b)	$Y \sim B(10,0.1)$ $P(Y \geq 4) = 1 - P(Y \leq 3)$ $= 1 - 0.9872$ $= 0.0128$	B1 M1 A1 (3)
7(c)	$0.9^n < 0.05$ or $1 - (0.9)^n > 0.95$ $n > 28.4$ $n = 29$ *alternative* B(28,0.1): P(0) = 0.0523 B(29,0.1): P(0) = 0.0471 $n = 29$	M1 A1 A1 M1 A1 A1cao (3)
7(d)	$C \sim Po(5)$ $P(C > 10) = 1 - P(C \leq 10)$ $= 1 - 0.9863$ $= 0.0137$	B1 M1 A1 (3) Total marks 10
	Notes	
7(a) **7(b)**	B1 for "binomial" or B(… B1 writing or using B(10,0.1) M1 writing or using $1 - P(Y \leq 3)$ A1 awrt 0.0128	
7(c)	M1 $(0.9)^n < 0.05$,oe, or $(0.9)^n = 0.05$,oe, or $(0.9)^n > 0.05$, oe, or seeing 0.0523 or seeing 0.0471 1st A1 [P(0)] = 0.0471 or getting awrt 28.4 May be implied by correct answer. 2nd A1 cao $n = 29$ should not come from incorrect working. **NB** An answer of 29 on its own with no working gains M1A1A1	
7(d)	B1 writing or using Po(5) M1 writing or using $1 - P(C \leq 10)$ A1 awrt 0.0137	