Paper Reference(s)

# 6684/01

# Edexcel GCE

## Statistics S2

## Advanced /Advanced Subsidiary

## Tuesday 24 June 2014 – 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 6 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.** Patients arrive at a hospital accident and emergency department at random at a rate of 6 per hour.

(*a*) Find the probability that, during any 90 minute period, the number of patients arriving at the hospital accident and emergency department is

(i) exactly 7,

(ii) at least 10.

**(5)**

A patient arrives at 11.30 a.m.

(*b*) Find the probability that the next patient arrives before 11.45 a.m.

**(3)**

---

**2.** The length of time, in minutes, that a customer queues in a Post Office is a random variable, $T$, with probability density function

$$\mathrm{f}(t) = \begin{cases} c\left(81 - t^2\right) & 0 \le t \le 9 \\ 0 & \text{otherwise} \end{cases}$$

where $c$ is a constant.

(*a*) Show that the value of $c$ is $\frac{1}{486}$.

**(4)**

(*b*) Show that the cumulative distribution function F($t$) is given by

$$\mathrm{F}(t) = \begin{cases} 0 & t < 0 \\ \frac{t}{6} - \frac{t^3}{1458} & 0 \le t \le 9 \\ 1 & t > 9 \end{cases}$$

**(2)**

(*c*) Find the probability that a customer will queue for longer than 3 minutes.

**(2)**

A customer has been queuing for 3 minutes.

(*d*) Find the probability that this customer will be queuing for at least 7 minutes.

**(3)**

Three customers are selected at random.

(*e*) Find the probability that exactly 2 of them had to queue for longer than 3 minutes.

**(3)**

---

**3.** A company claims that it receives emails at a mean rate of 2 every 5 minutes.

(*a*) Give two reasons why a Poisson distribution could be a suitable model for the number of emails received.

**(2)**

(*b*) Using a 5% level of significance, find the critical region for a two-tailed test of the hypothesis that the mean number of emails received in a 10 minute period is 4. The probability of rejection in each tail should be as close as possible to 0.025.

**(2)**

(*c*) Find the actual level of significance of this test.

**(2)**

To test this claim, the number of emails received in a random 10 minute period was recorded.

During this period 8 emails were received.

(*d*) Comment on the company's claim in the light of this value. Justify your answer.

**(2)**

During a randomly selected 15 minutes of play in the Wimbledon Men's Tennis Tournament final, 2 emails were received by the company.

(*e*) Test, at the 10% level of significance, whether or not the mean rate of emails received by the company during the Wimbledon Men's Tennis Tournament final is lower than the mean rate received at other times. State your hypotheses clearly.

**(5)**

---

**4.** A cadet fires shots at a target at distances ranging from 25 m to 90 m. The probability of hitting the target with a single shot is $p$. When firing from a distance $d$ m, $p = \frac{3}{200}(90 - d)$.

Each shot is fired independently.

The cadet fires 10 shots from a distance of 40 m.

(*a*) (i) Find the probability that exactly 6 shots hit the target.

(ii) Find the probability that at least 8 shots hit the target.

**(5)**

The cadet fires 20 shots from a distance of $x$ m.

(*b*) Find, to the nearest integer, the value of $x$ if the cadet has an 80% chance of hitting the target at least once.

**(4)**

The cadet fires 100 shots from 25 m.

(*c*) Using a suitable approximation, estimate the probability that at least 95 of these shots hit the target.

**(5)**

---

**5.** (*a*) State the conditions under which the normal distribution may be used as an approximation to the binomial distribution.

**(2)**

A company sells seeds and claims that 55% of its pea seeds germinate.

(*b*) Write down a reason why the company should not justify their claim by testing all the pea seeds they produce.

**(1)**

To test the company's claim, a random sample of 220 pea seeds was planted.

(*c*) State the hypotheses for a two-tailed test of the company's claim.

**(1)**

Given that 135 of the 220 pea seeds germinated,

(*d*) use a normal approximation to test, at the 5% level of significance, whether or not the company's claim is justified.

**(7)**

---

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

$$\mathrm{f}(x) = \begin{cases} \dfrac{2x}{9} & 0 \le x \le 1 \\ \dfrac{2}{9} & 1 < x < 4 \\ \dfrac{2}{3} - \dfrac{x}{9} & 4 \le x \le 6 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Find E($X$).

**(4)**

(*b*) Find the cumulative distribution function F($x$) for all values of $x$.

**(6)**

(*c*) Find the median of $X$.

**(3)**

(*d*) Describe the skewness. Give a reason for your answer.

**(2)**

---

**TOTAL FOR PAPER: 75 MARKS**

**END**

| Question Number | Scheme | Marks |
|---|---|---|
| **1.** | | |
| **(a)** | Po(9) | B1 |
| **(i)** | $P(X \le 7) - P(X \le 6)$ = 0.3239 – 0.2068 $\quad \frac{e^{-9}9^7}{7!}$ | M1 |
| | = 0.1171 | A1 |
| **(ii)** | $P(X \ge 10) = 1 - P(X \le 9)$ | M1 |
| | = 1 – 0.5874 | |
| | = 0.4126 | A1 |
| | | (5) |
| **(b)** | Po(1.5) | B1 |
| | P(next patient before 11:45) = 1- P(0) | M1 |
| | = 1 - $e^{-1.5}$ | |
| | = 0.7769 | A1 |
| | | (3) |
| | | [8] |
| | **Notes** | |
| **(a) (i)** | B1 Po(9) written or used in either (i) or (ii) | |
| | M1 writing $P(X \le 7) - P(X \le 6)$ or $\frac{e^{-\lambda}\lambda^7}{7!}$ | |
| | This may be implied by 0.3239 – 0.2068 | |
| | A1 awrt 0.117 | |
| **(ii)** | M1 writing $1 - P(X \le 9)$ | |
| | This may be implied by 1 – 0.5874. | |
| | A1 awrt 0.413 | |
| **(b)** | B1 Po(1.5) written or used | |
| | M1 writing or using 1 – P(0) or 1 - $e^{-\lambda}$. | |
| | This may be implied by 1 – 0.2231 | |
| | A1 awrt 0.777 | |

| Question Number | Scheme | Marks |
|---|---|---|
| **2.** **(a)** | $\int_0^9 c\left(81-t^2\right)\mathrm{d}t = 1$<br>$c\left[81t-\frac{t^3}{3}\right]_0^9 = 1$<br>$c\left[81\times 9-\frac{9^3}{3}\right] = 1$<br>$486c = 1$<br>$c = \frac{1}{486}$ | M1<br>A1<br>M1d<br>A1cso<br>(4) |
| **(b)** | $\mathrm{F}(t) = \frac{1}{486}\int_0^t 81-x^2\mathrm{d}x$<br>$= \frac{1}{486}\left[81t-\frac{x^3}{3}\right]_0^t$<br>$= \frac{t}{6}-\frac{t^3}{1458}$<br>$\mathrm{F}(t) = \begin{cases} 0 & t<0 \\ \frac{t}{6}-\frac{t^3}{1458} & 0\le t\le 9 \\ 1 & t>9 \end{cases}$ | M1<br>A1cso<br>(2) |
| **(c)** | $\mathrm{P}(T>3) = 1-\left(\frac{3}{6}-\frac{3^3}{1458}\right)$<br>$= \frac{14}{27}$ or awrt 0.519 | M1<br>A1<br>(2) |
| **(d)** | $\mathrm{P}(T>7 \mid T>3) = \frac{0.068587}{0.5185}$<br>$= \frac{25}{189}$ or awrt 0.132 | M1A1ft<br>A1<br>(3) |
| **(e)** | ${}^3C_2\left(0.5185\right)^2\left(1-0.5185\right) = \frac{2548}{6561}$ or awrt 0.388/ 0.387 | M1A1ftA1<br>(3)<br>[14] |

| | Notes | |
|---|---|---|
| **(a)** | 1st M1 Attempting to integrate, For attempt $x^n \rightarrow x^{n+1}$ and $c$ must remain as $c$ or 1/486. Ignore limits<br>1st A1 Correct integration. Ignore limits.<br>2nd M1 dependent on previous M being awarded.<br>Putting = 1 and substitution of 9 as a limit seen. Need at least one intermediate step before getting 486<br>or substitution of 1/486 and 9 seen and leading to an answer of 1<br>A1 $c = \frac{1}{486}$ cso or if verifying, the statement $c = \frac{1}{486}$ | |
| **(b)** | M1 Attempting to integrate with correct limits **or** $\int f(t)dt$ + C and F(0) = 0 or F(9) = 1.<br>Subst in $c$ at some point<br>A1 F($t$) must be stated and cso. Condone use of < instead of ≤ etc. | |
| **(c)** | M1 using or writing 1 – F(3) or $\frac{1}{486}\int_3^9 81 - x^2 dx$ or 1 – P($X \leq 3$)<br>A1 awrt 0.519 | |
| **(d)** | M1 $\frac{a\ probability}{their\ (c)}$<br>where 0 < *a probability* < *their (c)* < 1. If *a probability* ≥ *their (c)*, give M0.<br>A1ft $\frac{\frac{50}{729}}{their\ (c)}$ or $\frac{awrt 0.0686}{their\ (c)}$<br>A1 $\frac{25}{189}$ or awrt 0.132 | |
| **(e)** | M1 Allow $(their\ '0.5185')^2 (1 - their\ '0.5185')$<br>A1ft Allow ${}^3C_2$ (*their* '0.5185')$^2$(1 – *their* '0.5185')<br>A1 awrt 0.388 or 0.387 | |

| Question Number | Scheme | Marks |
|---|---|---|
| **3.** **(a)** | Any two of<br>• Emails are independent/occur at random<br>• Emails occur singly<br>• Emails occur at a constant rate | B1B1d (2) |
| **(b)** | $X \sim$ Po(4)<br>$P(X = 0) = 0.0183$<br>$P(X \geq 9) = 0.0214$<br>CR $X = 0$; $X \geq 9$ | B1B1 (2) |
| **(c)** | 0.0183 + 0.0214 = 0.0397 or 3.97% | M1A1 (2) |
| **(d)** | 8 is not in the critical region **or** $P(X \geq 8) = 0.0511$<br>therefore there is evidence that the company's **claim** is true | M1<br>A1ft (2) |
| **(e)** | $H_0 : \lambda = 6$ (or $\lambda = 2$) $H_1 : \lambda < 6$ (or $\lambda = 2$) allow $\lambda$ or $\mu$<br>Po(6)<br>$P(X \leq 2) = 0.0620$ CR $X \leq 2$<br><br>0.0620 < 0.10<br>Reject $H_0$ or Significant.<br>There is evidence at the 10% level of significance that the mean **rate/number/amount** of **emails** received **is lower/ has decreased/is less**.<br>Or **fewer emails** are received | B1<br>M1<br>A1<br><br>M1 dep.<br>A1 cso<br>(5)<br>[13] |
| | **Notes** | |
| **(a)** | B1 any correct statement with context of emails in<br>B1d Dependent on previous B1. Any correct statement, need not have context<br>**SC** for 2 correct statements without context B1 B0 | |
| **(b)** | B1 $X = 0$ or $X \leq 0$ Allow any letter.<br>B1 $X \geq 9$ or $X > 8$ Allow any letter.<br>**SC** if write correct CR's as probability statements award B1 B0<br>For these 2 marks ignore any union sign ($\cup$) or intersection sign ($\cap$) | |
| **(c)** | M1 adding their probabilities of 'their' critical regions if sum gives a probability less than 1 or award if a correct answer given<br>A1 awrt 0.0397 | |
| **(d)** | M1 correct reason ft their CR. Do not allow non-contextual contradictions.<br>A1 correct conclusion for their CR. Allow conclusion in context of **emails** are **received** at a rate of 2 every 5 mins | |
| **(e)** | B1 both hypotheses correct, must have $\lambda$ or $\mu$ and either 2 or 6.<br>M1 using Po(6) may be implied by correct answer.<br>A1 0.062 or $X \leq 2$<br>M1 dependent on previous method being awarded. Do not allow conflicting non-contextual statements. Follow through their hypotheses. | |

| Question Number | Scheme | Marks |
|---|---|---|
| **4. (a)** | $X$ is the random variable the Number of successes, $X \sim$ B(10, 0.75) | B1 |
| **(i)** | P ($X = 6$) = $(0.75)^6 (0.25)^4 \, {}^{10}C_6$ or $\mathrm{P}(X \le 6) - \mathrm{P}(X \le 5)$ | M1 |
| | = 0.145998 awrt 0.146 | A1 |
| **(ii)** | Using $X \sim$ B(10, 0.75) | |
| | $\mathrm{P}(X \ge 8) = \mathrm{P}(X = 8) + \mathrm{P}(X = 9) + \mathrm{P}(X = 10)$ | M1 |
| | $= (0.75)^8 (0.25)^2 \, {}^{10}C_8 + (0.75)^9 (0.25)^1 \, {}^{10}C_9 + (0.75)^{10}$ | |
| | = 0.52559 awrt 0.526 | A1 |
| | Or | |
| | Using $Y \sim$ B(10, 0.25) and $\mathrm{P}(Y \le 2) = 0.5256$ | (5) |
| **(b)** | $1 - \mathrm{P}(0) = 0.8$ or $\mathrm{P}(0) = 0.2$ | M1 |
| | $(1-p)^{20} = 0.2$ | |
| | $1 - p = 0.9227$ | |
| | $p = 0.0773$ | A1 |
| | $\frac{3}{200}(90 - x) = 0.0773$ | M1 |
| | $x = 84.84$ | |
| | $x = 85$ | A1cao (4) |
| **(c)** | $X$ – successes ~B(100, 0.975) | B1 |
| | $Y$ – not successes ~B(100, 0.025) | |
| | $Y$~Po(2.5) | M1A1 |
| | $\mathrm{P}(Y \le 5) = 0.958$ | M1A1 (5) |
| | **Notes** | [14] |
| **(a)** | B1 writing or using $p = 0.75$ or $p = 0.25$ anywhere in (a)(i) or (a)(ii) | |
| **(i)** | M1 writing or using $(p)^6 (1-p)^4 \, {}^{10}C_6$ **or** writing for $p$ =0.75, $\mathrm{P}(X \le 6) - (X \le 5)$ **or** for $p$ = 0.25, P($X \le 4$) – P($X \le 3$) **or** correct answer. | |
| **(ii)** | M1 writing B(10, 0.75) and writing or using $\mathrm{P}(X = 8) + \mathrm{P}(X = 9) + \mathrm{P}(X = 10)$ oe **or** writing B(10, 0.25) and writing or using $\mathrm{P}(Y \le 2)$. | |
| | Using correct Binomial must be shown by $(0.75)^n (0.25)^{10-n}$ or a correct answer. | |
| **(b)** | M1 for writing or using 1 – P(0) = 0.8 or P(0) = 0.2 or $(1-p)^{20} = 0.2$. Allow any inequality sign. | |
| | A1 awrt 0.0773 or awrt 0.923. | |
| | M1 subst in $\frac{3}{200}(90 - x)$ for $p$ **NB** this may be substituted in earlier for $p$. | |
| | Allow for $\frac{3}{200}(90 - x) = k$ where $0 < k < 1$ $k \neq 0.8$ or 0.2 Allow any inequality sign | |
| | A1 condone $x \ge 85$. Do not allow x ≤ 85. | |
| **(c)** | B1 writing or using 0.975 or 0.025, may be implied by Po(2.5) | |
| | M1 using Po approximation | |
| | A1 Po(2.5) | |
| | M1 writing or using $\mathrm{P}(Y \le 5)$ | |
| | A1 awrt 0.958 | |
| | **SC use of normal approximation** can get B1 M0A0M1A0 | |
| | B1 writing or using 0.975 or 0.025 implied by normal with mean 97.5 or answer of 0.973 | |
| | M1 for awrt 0.973 | |

| Question Number | Scheme | Marks |
|---|---|---|
| **5.(a)** | $n$ is large and $p$ close to 0.5 | B1B1 (2) |
| **(b)** | There would be no pea seeds left | B1 (1) |
| **(c)** | $H_0$: $p = 0.55$ $H_1$ : $p \neq 0.55$ | B1 (1) |
| **(d)** | $X$~N(121, 54.45) | B1 |
| | $P(X \geq 134.5) = P\left(Z \geq \frac{134.5-121}{\sqrt{54.45}}\right)$ or $\pm\frac{x-0.5-121}{\sqrt{54.45}} = 1.96$ | M1M1A1 |
| | $= P(Z \geq 1.8295..)$ | |
| | $= 1 - 0.9664$ | |
| | $= 0.0336/0.0337$ $x = 135.96$ | A1 |
| | Accept $H_0$ not in CR, not significant | M1 |
| | The **company's claim** is justified **or** **55**% of its pea **seeds germinate** | A1cso (7) |
| | **Alternative** | |
| | $X$~N(99, 54.45) | B1 |
| | $P(X \leq 85) = P\left(Z \leq \frac{85.5-99}{\sqrt{54.45}}\right)$ or $\pm\frac{x+0.5-99}{\sqrt{54.45}} = 1.96$ | M1 M1 A1 |
| | $= P(Z \geq 1.8295..)$ | |
| | $= 1 - 0.9664$ | |
| | $= 0.0336/0.0337$ $x = 107.5$ | |
| | Accept $H_0$ not in CR, not significant | M1 |
| | The **company's claim** is justified **or** **55**% of its pea **seeds germinate** | A1cso [11] |
| | **Notes** | |
| **(a)** | B1 accept $n > 50$ (or any number bigger than 50) | |
| | B1 $p$ close to 0.5 | |
| | NB Do not accept $np > 5$, $nq > 5$. | |
| **(b)** | Must have the idea of no peas left. They must mention either **pea** or **seeds**. | |
| **(c)** | B1 both hypotheses correct. Must use $p$ or $\pi$ and 0.55 oe. Accept the hypotheses in part (d). | |
| **(d)** | B1 correct mean and Var, may be seen in the standardiation formula as 121 and $\sqrt{54.45}$ or 7.38 to 2dp or implied by a correct answer | |
| | M1 for attempting a continuity correction (Method 1:135/85 $\pm$ 0.5 / Method 2:$x \pm 0.5$) | |
| | M1 for standardising using their mean and their standard deviation and using either Method 1 [134.5, 135, 135.5, 85, 85.5 or 84.5 accept $\pm z$.] Method 2 [ ($x \pm 0.5$) and equal to a $\pm$ $z$ value] | |
| | A1 correct $z$ value awrt $\pm$ 1.83 **or** $\pm\frac{134.5-121}{\sqrt{54.45}}$ $\left(\frac{85.5-99}{\sqrt{54.45}}\right)$ or $\pm\frac{x-0.5-121}{\sqrt{54.45}} = 1.96$ | |
| | $\left(\pm\frac{x+0.5-99}{\sqrt{54.45}} = 1.96\right)$ or(allow 1.6449 if 1 tail test in (c)) | |
| | A1 awrt 0.0336/0.0337 **or** awrt 136 (allow 126 if one tail test in (c)) or a comparison of awrt1.83 with 1.96 (1.6449) | |
| | M1 A correct statement. Accept $H_0$, oe if a 2-tailed test in (c), reject $H_0$, oe if a 1-tailed test in (c). Allow for a correct contextual statement. Do not allow contradictions of non-contextual statements. | |
| | A1 A correct contextual statement to include words in bold/underlined for a 2-tailed test. This is not a follow through mark. | |
| | **NB** if finding $P(X = 135)$ they can get B1 M1 M1 A0 A0 M0 A0 | |

| Question Number | Scheme | Marks |
|---|---|---|
| **6.** | | |
| **(a)** | $\mathrm{E}(X)=\int_0^1 \frac{2x^2}{9}\mathrm{d}x+\int_1^4 \frac{2x}{9}\mathrm{d}x+\int_4^6 \frac{2x}{3}-\frac{x^2}{9}\mathrm{d}x$ | M1 |
| | $=\left[\frac{2x^3}{27}\right]_0^1+\left[\frac{2x^2}{18}\right]_1^4+\left[\frac{x^2}{3}-\frac{x^3}{27}\right]_4^6$ | A1 |
| | $=\left[\frac{2}{27}\right]+\left[\frac{32}{18}-\frac{2}{18}\right]+\left[4-\frac{80}{27}\right]$ | M1d |
| | $=2\frac{7}{9}$ or awrt 2.78 | A1 |
| | | (4) |
| **(b)** | $\mathrm{F}(x)=\begin{cases} 0 & x<0 \\ \frac{x^2}{9} & 0\le x\le 1 \\ \frac{2x}{9}-\frac{1}{9} & 1<x<4 \\ \frac{2x}{3}-\frac{x^2}{18}-1 & 4\le x\le 6 \\ 1 & x>6 \end{cases}$ | B1<br>M1A1<br>M1 A1<br>B1 |
| | 1st M1 For $1<x<4$, $\mathrm{F}(x)=\int_1^x \frac{2}{9}\mathrm{d}x+\frac{1}{9}$ | |
| | 2nd M1 For $4\le x\le 6$ , $\mathrm{F}(x)=\int_4^x \frac{2}{3}-\frac{x}{9}\mathrm{d}x+\frac{7}{9}$ or use +C and F(6) =1 | |
| | | (6) |
| **(c)** | F($x$) = 0.5 | M1 |
| | $\frac{2m}{9}-\frac{1}{9}=0.5$ | A1ft |
| | $m$ = 2.75 | A1 |
| | | (3) |
| **(d)** | Median < mean therefore positive skew<br>**Or** Mean $\approx$ median therefore no skewness | M1A1cao<br>(2) |
| | | [15] |

| | Notes | |
|---|---|---|
| **(a)** | M1 using $\int xf(x)dx$ ignore limits. Must have at least one $x^n \rightarrow x^{n+1}$<br>They must add the 3 parts together. Do not allow division by 3.<br>A1 all integration correct; ignore limits<br>M1 dependent on previous M being awarded. Subst in correct limits – no need to see zero substituted.<br>A1 $2\frac{7}{9}$ oe or awrt 2.78 | |
| **(b)** | B1 for 2nd line- allow use of < instead of ≤<br>M1 For $1 < x < 4$, $F(x) = \int_1^x \frac{2}{9}dx + \frac{1}{9}$ . Limits are needed.<br>**or** use $F(x) = \int_1^x \frac{2}{9}dx +$ their F(1) need limits<br>**or** use "their $F(1)" = \int \frac{2}{9}dx + C$ and subst $x = 1$ into RHS<br>**or** use "their $F(4)" = \int \frac{2}{9}dx + C$ and subst $x = 4$ into RHS<br>A1 for 3rd line allow use of ≤ instead of <<br>M1 For $4 \le x \le 6$ , $F(x) = \int_4^x \frac{2}{3} - \frac{x}{9}dx + \frac{7}{9}$ . Limits are needed.<br>**or** use $F(x) = \int_4^x \frac{2}{3} - \frac{x}{9}dx +$ their F(4) . Limits are needed.<br>**or** use "their $F(4)" = \int \frac{2}{3} - \frac{x}{9}dx + C$ and subst $x = 4$ into RHS<br>**or** use $1 = \int \frac{2}{3} - \frac{x}{9}dx + C$ and subst $x = 6$ into RHS<br>A1 for 4th line allow use of < instead of ≤<br>B1 for first and last line - allow use of ≤ instead of < and ≥ instead of > and "otherwise" for one of $x < 0$ and $x > 6$ | |
| **(c)** | M1 putting any one of their lines = 0.5<br>A1their 3rd line = 0.5<br>A1 2.75 | |
| **(d)** | M1 reason must match their values / a correctly shaped and labelled sketch.<br>Must compare the median and mean, ignore references to mode<br>A1 no ft Correct answer only from correct values of the mean and median or a correct and fully labelled sketch. | |