

# Mark Scheme (Results)

## Summer 2009

GCE

GCE Mathematics (6684/01)

June 2009  
6684 Statistics S2  
Mark Scheme

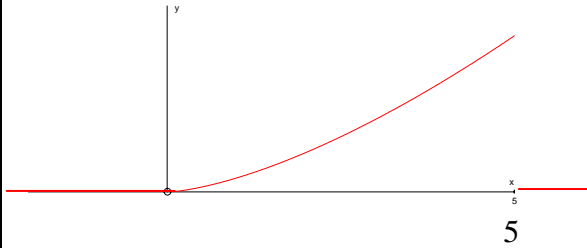
Question Number	Scheme	Marks
Q1 (a)	$[X \sim B(30, 0.15)]$  $P(X \leq 6) = 0.8474$	awrt 0.847  M1, A1 (2)
(b)	$Y \sim B(60, 0.15) \approx \text{Po}(9)$  $P(Y \leq 12) = 0.8758$	for using Po(9)  B1  awrt 0.876  M1, A1 (3)
[ N.B. normal approximation gives 0.897, exact binomial gives 0.894]		[5]
(a)	<p>M1 for a correct probability statement <math>P(X \leq 6)</math> or <math>P(X &lt; 7)</math> or <math>P(X=0) + P(X=1) + P(X=2) + P(X=4) + P(X=5) + P(X=6)</math>. (may be implied by long calculation) Correct answer gets M1 A1. allow 84.74%</p>	
(b)	<p>B1 may be implied by using Po(9). Common incorrect answer which implies this is 0.9261 M1 for a correct probability statement <math>P(X \leq 12)</math> or <math>P(X &lt; 13)</math> or <math>P(X=0) + P(X=1) + \dots + P(X=12)</math> (may be implied by long calculation) and attempt to evaluate this probability using their Poisson distribution.  Condone <math>P(X \leq 13) = 0.8758</math> for B1 M1 A1  Correct answer gets B1 M1 A1  Use of normal or exact binomial get B0 M0 A0</p>	

Question Number	Scheme	Marks
Q2	$H_0: \lambda = 2.5$ (or $\lambda = 5$ ) $H_1: \lambda < 2.5$ (or $\lambda < 5$ ) $\lambda$ or $\mu$  $X \sim \text{Po}(5)$  $P(X \leq 1) = 0.0404$ or    CR $X \leq 1$  [0.0404 < 0.05 ] this is significant    or    reject $H_0$ or    it is in the critical region  There is evidence of a <u>decrease</u> in the (mean) <u>number/rate</u> of <u>deformed blood cells</u>	B1B1  M1  A1  M1  A1    (6) [6]
	1 <sup>st</sup> B1 for $H_0$ must use lambda or mu; 5 or 2.5. 2 <sup>nd</sup> B1 for $H_1$ must use lambda or mu; 5 or 2.5 1 <sup>st</sup> M1 for use of Po(5) may be implied by probability( must be used not just seen) eg. $P(X = 1) = 0.0404 - \dots$ would score M1 A0 1 <sup>st</sup> A1 for 0.0404 seen or correct CR 2 <sup>nd</sup> M1 for a correct statement (this may be contextual) comparing their probability and 0.05 (or comparing 1 with their critical region). Do not allow conflicting statements. 2 <sup>nd</sup> A1 is not a follow through. Need the word decrease, number or rate and deformed blood cells for contextual mark.  If they have used $\neq$ in $H_1$ they could get B1 B0 M1 A1 M1A0 mark as above except they gain the 1 <sup>st</sup> A1 for $P(X \leq 1) = 0.0404$ or CR $X \leq 0$ 2 <sup>nd</sup> M1 for a correct statement (this may be contextual) comparing their probability and 0.025 (or comparing 1 with their critical region)  They may compare with 0.95 (one tail method) or 0.975 (one tail method) Probability is 0.9596.	

Question Number	Scheme	Marks
Q3 (a)	<p><i>A statistic</i> is a function of <math>X_1, X_2, \dots, X_n</math> that does not contain any unknown parameters</p> <p>The <u>probability</u> distribution of <math>Y</math> or the distribution of all possible values of <math>Y</math> (o.e.)</p> <p>Identify (ii) as not a statistic Since <u>it contains</u> unknown parameters <u><math>\mu</math> and <math>\sigma</math></u>.</p>	<p>B1 B1 (2)</p> <p>B1 (1)</p> <p>B1 dB1 (2)</p> <p>[5]</p>
(a)	<p>Examples of other acceptable wording:</p> <p>B1 e.g. is a function of the sample or the data / is a quantity calculated from the sample or the data / is a random variable calculated from the sample or the data</p> <p>B1 e.g. does not contain any unknown parameters/quantities contains only known parameters/quantities <u>only</u> contains values of the sample</p> <p><math>Y</math> is a function of <math>X_1, X_2, \dots, X_n</math> that does not contain any unknown parameters B1B1 is a function of the values of a sample with no unknowns B1B1 is a function of the sample values B1B0 is a function of all the data values B1B0 A random variable calculated from the sample B1B0 A random variable consisting of any function BOB0 A function of a value of the sample B1B0 A function of the sample which contains no other values/ parameters B1B0</p>	
(b)	<p>Examples of other acceptable wording</p> <p>All possible values of the statistic together with their associated probabilities</p>	
(c)	<p>1<sup>st</sup> B1 for selecting only (ii) 2<sup>nd</sup> B1 for a reason. This is dependent upon the first B1. Need to mention at least one of <math>\mu</math> (mean) or <math>\sigma</math> (standard deviation or variance) or unknown parameters. Examples since it contains <math>\mu</math> B1 since it contains <math>\sigma</math> B1 since it contains unknown parameters/quantities B1 since it contains unknowns B0</p>	

Question Number	Scheme	Marks
Q4 (a)	$X \sim B(20, 0.3)$ $P(X \leq 9) = 0.9520$ so $P(X \geq 10) = 0.0480$ Therefore the critical region is $\{X \leq 2\} \cup \{X \geq 10\}$ awrt (0.083 or 0.084)	M1 A1 A1 A1A1 (5)
(b)	$0.0355 + 0.0480 = 0.0835$	B1 (1)
(c)	11 is in the critical region there is evidence of a <u>change/ increase</u> in the <u>proportion/number</u> of <u>customers buying single tins</u>	B1ft B1ft (2)
(a)	M1 for B(20,0.3) seen or used 1 <sup>st</sup> A1 for 0.0355 2 <sup>nd</sup> A1 for 0.048 3 <sup>rd</sup> A1 for $(X) \leq 2$ or $(X) < 3$ or $[0,2]$ They get <b>A0</b> if they write $P(X \leq 2/ X < 3)$ 4 <sup>th</sup> A1 $(X) \geq 10$ or $(X) > 9$ or $[10,20]$ They get <b>A0</b> if they write $P(X \geq 10/ X > 9)$ $10 \leq X \leq 2$ 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.	[8]
(b)	B1 correct answer only	
(c)	1 <sup>st</sup> B1 for a correct statement about 11 and their critical region. 2 <sup>nd</sup> B1 for a correct comment in context consistent with their CR and the value 11	
	Alternative solution 1 <sup>st</sup> B0 $P(X \geq 11) = 1 - 0.9829 = 0.0171$ since no comment about the critical region 2 <sup>nd</sup> B1 a correct contextual statement.	

Question Number	Scheme	Marks
Q5 (a)	$X = \text{the number of errors in 2000 words}$ so $X \sim \text{Po}(6)$ $P(X \geq 4) = 1 - P(X \leq 3)$ $= 1 - 0.1512 = 0.8488$ awrt 0.849	B1 M1 A1 (3)
(b)	$Y = \text{the number of errors in 8000 words. } Y \sim \text{Po}(24)$ so use a <u>Normal</u> approx $Y \approx N(24, \sqrt{24}^2)$  Require $P(Y \leq 20) = P\left(Z < \frac{20.5 - 24}{\sqrt{24}}\right)$ $= P(Z < -0.714\dots)$ $= 1 - 0.7611$ $= 0.2389$ awrt (0.237~0.239)	M1 A1  M1 M1 A1 M1 A1 (7)
	[N.B. Exact Po gives 0.242 and no $\pm 0.5$ gives 0.207]	[10]
(a)	B1 for seeing or using Po(6) M1 for $1 - P(X \leq 3)$ or $1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$ A1 awrt 0.849  SC If B(2000, 0.003) is used and leads to awrt 0.849 allow B0 M1 A1 If no distribution indicated awrt 0.8488 scores B1M1A1 but any other awrt 0.849 scores B0M1A1	
(b)	1 <sup>st</sup> M1 for identifying the normal approximation 1 <sup>st</sup> A1 for [mean = 24] <b>and</b> [sd = $\sqrt{24}$ or var = 24]  These first two marks may be given if the following are seen in the standardisation formula : 24 $\sqrt{24}$ or awrt 4.90  2 <sup>nd</sup> M1 for attempting a continuity correction (20/ 28 $\pm$ 0.5 is acceptable) 3 <sup>rd</sup> M1 for standardising using their mean and their standard deviation. 2 <sup>nd</sup> A1 correct z value awrt $\pm 0.71$ <b>or</b> this may be awarded if see $\frac{20.5 - 24}{\sqrt{24}}$ or $\frac{27.5 - 24}{\sqrt{24}}$  4 <sup>th</sup> M1 for 1 - a probability from tables (must have an answer of < 0.5) 3 <sup>rd</sup> A1 answer awrt 3 sig fig in range 0.237 – 0.239	

Question Number	Scheme	Marks
Q6 (a) (b) (c) (d) (e) (f) (g)	$P(A > 3) = \frac{2}{5} = 0.4$ $(0.4)^3 = 0.064 \text{ or } \frac{8}{125}$ $f(y) = \frac{d}{dy}(F(y)) = \begin{cases} \frac{3y^2}{125} & 0 \leq y \leq 5 \\ 0 & \text{otherwise} \end{cases}$  <p>Shape of curve and start at (0,0)</p> <p>Point (5, 0) labelled and curve between 0 and 5 and pdf ≥ 0</p> <p>Mode = 5</p> $E(Y) = \int_0^5 \left( \frac{3y^3}{125} \right) dy = \left[ \frac{3y^4}{500} \right]_0^5 = \frac{15}{4} \text{ or } 3.75$ $P(Y > 3) = \begin{cases} \int_3^5 \frac{3y^2}{125} dy = 1 - \frac{27}{125} = \frac{98}{125} = 0.784 \\ \text{or } 1 - F(3) \end{cases}$	B1 (1) M1, A1 (2) M1A1 (2) B1 B1 (2) B1 (1) M1M1A1 (3) M1A1 (2) <b>[13]</b>
(a) (b) (c) (d) (e) (f) (g)	B1 correct answer only (cao). Do not ignore subsequent working M1 for cubing their answer to part (a) A1 cao M1 for attempt to differentiate the cdf. They must decrease the power by 1 A1 fully correct answer including 0 otherwise. Condone < signs B1 for shape. Must curve the correct way and start at (0,0). No need for y = 0 (patios) lines B1 for point (5,0) labelled and pdf only existing between 0 and 5, may have y=0 (patios) for other values B1 cao 1 <sup>st</sup> M1 for attempt to integrate their $yf(y) y^n \rightarrow y^{n+1}$ . 2 <sup>nd</sup> M1 for attempt to use correct limits A1 cao M1 for attempt to find $P(Y > 3)$ . e.g. writing $\int_3^5$ their $f(y)$ must have correct limits or writing $1 - F(3)$	

Question Number	Scheme	Marks
<p>Q7 (a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p> <p>(e)</p>	<p><math>E(X) = 2</math> (by symmetry)</p> <p><math>0 \leq x &lt; 2</math>, gradient = <math>\frac{1}{2} = \frac{1}{4}</math> and equation is <math>y = \frac{1}{4}x</math> so <math>a = \frac{1}{4}</math></p> <p><math>b - \frac{1}{4}x</math> passes through <math>(4, 0)</math> so <math>b = 1</math></p> $E(X^2) = \int_0^2 \left(\frac{1}{4}x^3\right) dx + \int_2^4 \left(x^2 - \frac{1}{4}x^3\right) dx$ $= \left[\frac{x^4}{16}\right]_0^2 + \left[\frac{x^3}{3} - \frac{x^4}{16}\right]_2^4$ $= 1 + \frac{64-8}{3} - \frac{256-16}{16} = 4\frac{2}{3} \text{ or } \frac{14}{3}$ <p><math>\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{14}{3} - 2^2</math>, <math>= \frac{2}{3}</math> (so <math>\sigma = \sqrt{\frac{2}{3}} = 0.816</math>) (*)</p> <p><math>P(X \leq q) = \int_0^q \frac{1}{4}x dx = \frac{1}{4}</math>, <math>\frac{q^2}{2} = 1</math> so <math>q = \sqrt{2} = 1.414</math> awrt 1.41</p> <p><math>2 - \sigma = 1.184</math> so <math>2 - \sigma</math>, <math>2 + \sigma</math> is wider than IQR, therefore greater than 0.5</p>	<p>B1 (1)</p> <p>B1</p> <p>B1 (2)</p> <p>M1M1</p> <p>A1</p> <p>M1A1</p> <p>M1 A1cso (7)</p> <p>M1A1,A1 (3)</p> <p>M1,A1 (2)</p> <p>[15]</p>
<p>(a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p> <p>(e)</p>	<p>B1 cao</p> <p>B1 for value of <math>a</math>. B1 for value of <math>b</math></p> <p>1<sup>st</sup> M1 for attempt at <math>\int ax^3</math> using their <math>a</math>. For attempt they need <math>x^4</math>. Ignore limits.</p> <p>2<sup>nd</sup> M1 for attempt at <math>\int bx^2 - ax^3</math> use their <math>a</math> and <math>b</math>. For attempt need to have either <math>x^3</math> or <math>x^4</math>. Ignore limits</p> <p>1<sup>st</sup> A1 correct integration for both parts</p> <p>3<sup>rd</sup> M1 for use of the correct limits on each part</p> <p>2<sup>nd</sup> A1 for either getting 1 and <math>3\frac{2}{3}</math> or awrt 3.67 somewhere or <math>4\frac{2}{3}</math> or awrt 4.67</p> <p>4<sup>th</sup> M1 for use of <math>E(X^2) - [E(X)]^2</math> must add both parts for <math>E(X^2)</math> and only have subtracted the mean<sup>2</sup> once. You must see this working</p> <p>3<sup>rd</sup> A1 <math>\sigma = \sqrt{\frac{2}{3}}</math> or <math>\sqrt{0.66667}</math> or better with no incorrect working seen.</p> <p>M1 for attempting to find LQ, integral of either part of <math>f(x)</math> with their 'a' and 'b' = 0.25</p> <p>Or their <math>F(x) = 0.25</math> i.e. <math>\frac{ax^2}{2} = 0.25</math> or <math>bx - \frac{ax^2}{2} + 4a - 2b = 0.25</math> with their <math>a</math> and <math>b</math></p> <p>If they add both parts of their <math>F(x)</math>, then they will get M0.</p> <p>1<sup>st</sup> A1 for a correct equation/expression using their 'a'</p> <p>2<sup>nd</sup> A1 for <math>\sqrt{2}</math> or awrt 1.41</p> <p>M1 for a reason based on their quartiles</p> <ul style="list-style-type: none"> <li>Possible reasons are <math>P(2 - \sigma &lt; X &lt; 2 + \sigma) = 0.6498</math> allow awrt 0.65</li> <li><math>1.184 &lt; LQ(1.414)</math></li> </ul> <p>A1 for correct answer <math>&gt; 0.5</math></p> <p>NB you must check the reason and award the method mark. A correct answer without a correct reason gets M0 A0</p>	



Question Number	Scheme	Marks
Q8 (a)	$X \sim \text{Po}(2) \quad P(X = 4) = \frac{e^{-2} \times 2^4}{4!} = 0.0902$ awrt 0.09	M1 A1 (2)
(b)	$Y \sim \text{Po}(8)$ $P(Y > 10) = 1 - P(Y \leq 10) = 1 - 0.8159 = 0.18411\dots$ awrt 0.184	B1 M1A1 (3)
(c)	$F = \text{no. of faults in a piece of cloth of length } x \quad F \sim \text{Po}(x \times \frac{2}{15})$ $e^{-\frac{2x}{15}} = 0.80$ $e^{-\frac{2}{15} \times 1.65} = 0.8025\dots, \quad e^{-\frac{2}{15} \times 1.75} = 0.791\dots$ These values are either side of 0.80 therefore $x = 1.7$ to 2 sf	M1A1 M1 A1 (4)
(d)	Expected number with no faults = $1200 \times 0.8 = 960$ Expected number with some faults = $1200 \times 0.2 = 240$ So expected profit = $960 \times 0.60 - 240 \times 1.50, \quad = \text{£}216$	M1 A1 M1, A1 (4) [13]
(a)	M1 for use of Po(2) may be implied A1 awrt 0.09	
(b)	B1 for Po(8) seen or used M1 for $1 - P(Y \leq 10)$ oe A1 awrt 0.184	
(c)	1 <sup>st</sup> M1 for forming a suitable Poisson distribution of the form $e^{-\lambda} = 0.8$ 1 <sup>st</sup> A1 for use of lambda as $\frac{2x}{15}$ (this may appear after taking logs) 2 <sup>nd</sup> M1 for attempt to consider a range of values that will prove 1.7 is correct <b>OR</b> for use of logs to show lambda = ... 2 <sup>nd</sup> A1 correct solution only. Either get 1.7 from using logs or stating values either side	
S.C	for $e^{-\frac{2}{15} \times 1.7} = 0.797\dots \approx 0.80 \quad \therefore x = 1.7$ to 2 sf allow 2 <sup>nd</sup> M1A0	
(d)	1 <sup>st</sup> M1 for one of the following $1200p$ or $1200(1-p)$ where $p = 0.8$ or $2/15$ . 1 <sup>st</sup> A1 for both expected values being correct or two correct expressions. 2 <sup>nd</sup> M1 for an attempt to find expected profit, must consider with and without faults 2 <sup>nd</sup> A1 correct answer only.	