edexcel

Mark Scheme (Results)

January 2015

Pearson Edexcel International A Level in Statistics 2
(WST02/01)

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the world's leading learning company. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information, please visit our website at www.edexcel.com.

Our website subject pages hold useful resources, support material and live feeds from our subject advisors giving you access to a portal of information. If you have any subject specific questions about this specification that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

www.edexcel.com/contactus

Pearson: helping people progress, everywhere

Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

January 2015
Publications Code IA040682

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

EDEXCEL IAL MATHEMATICS

General Instructions for Marking

1. The total number of marks for the paper is 75.

2. The Edexcel Mathematics mark schemes use the following types of marks:

- **M** marks: method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
- **A** marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
- **B** marks are unconditional accuracy marks (independent of M marks)
- Marks should not be subdivided.

3. Abbreviations

These are some of the traditional marking abbreviations that will appear in the mark schemes.

- bod – benefit of doubt
- ft – follow through
- the symbol $\surd$ will be used for correct ft
- cao – correct answer only
- cso - correct solution only. There must be no errors in this part of the question to obtain this mark
- isw – ignore subsequent working
- awrt – answers which round to
- SC: special case
- oe – or equivalent (and appropriate)
- dep – dependent
- indep – independent
- dp decimal places
- sf significant figures
- $*$ The answer is printed on the paper
- $\lfloor$ The second mark is dependent on gaining the first mark

4. All A marks are 'correct answer only' (cao.), unless shown, for example, as A1 ft to indicate that previous wrong working is to be followed through. After a misread however, the subsequent A marks affected are treated as A ft, but manifestly absurd answers should never be awarded A marks.

5. For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.

6. Ignore wrong working or incorrect statements following a correct answer.

January 2015
WST02 Statistics S2
Mark Scheme

Question Number	Scheme	Marks
1(a)	X~Po(3.2) $P(X = 3) = \frac{e^{-3.2}3.2^3}{3!}$ $= 0.2226$ awrt 0.223	B1 M1 A1 (3)
(b)	Y~Po(1.6) $P(Y \geq 1) = 1 - P(Y = 0)$ $= 1 - e^{-1.6}$ $= 0.7981$ awrt 0.798	B1 M1 A1 (3)
(c)	X~Po(0.8) $\frac{P(X = 1)\times P(X = 3)}{P(Y = 4)} = \frac{\left(e^{-0.8}\times 0.8\right)\times\left(\frac{e^{-0.8}0.8^3}{3!}\right)}{\frac{e^{-1.6}1.6^4}{4!}}$ $= \frac{0.3594\times 0.0383}{0.05513}$ $= 0.25$	M1 M1 M1 A1 A1 (5)
(d)	A~Po(72) approximated by N(72,72) $\frac{5000}{60} = 83.33$ $P(A \geq 84) = P\left(Z \geq \frac{83.5 - 72}{\sqrt{72}}\right)$ $= P(Z \geq 1.355\ldots)$ $= 0.0869$ awrt 0.087/0.088 Notes	B1 M1 M1 M1 A1 (5)
(a)	B1 for writing or using Po(3.2) M1 $\frac{e^{-\lambda}\lambda^3}{3!}$	
(b)	B1 for writing or using Po(1.6) M1 $1 - P(Y = 0)$ or $1 - e^{-\lambda}$	
(c)	1st M1 using Po(0.8) with X=1 or X=3 (may be implied by 0.359... or 0.0383...) 2nd M1 $\left(e^{-\lambda}\times\lambda\right)\times\left(\frac{e^{-\lambda}\lambda^3}{3!}\right)$ (consistent lambda) awrt 0.0138 implies 1st 2 M marks 3rd M1 correct use of conditional probability with denominator $= \frac{e^{-1.6}1.6^4}{4!}$ 1st A1 fully correct expression 2nd A1 0.25 (allow awrt 0.250)	
(d)	B1 Writing or using N(72,72) 1st M1 for exact fraction **or** awrt 83.3 (may be implied by 84) (Note: Use of N(4320,4320) can score B1 and 1st M1) 2nd M1 Using 84 +/– 0.5 3rd M1 standardising using 82.5, 83, 83.$\dot{3}$ (awrt 83.3), 83.5, 83.8, 84 or 84.5, 'their mean' **and** 'their sd'	

Question Number	Scheme	Marks
2(a)	$P(X > 4) = 1 - F(4)$ $= 1 - \frac{3}{5}$ $= \frac{2}{5}$ oe	M1 A1 (2)
(b)	1	B1 (1)
(c)	$f(x) = \frac{dF(x)}{dx} = \frac{1}{5}$ $f(x) = \begin{cases} \frac{1}{5} & 1 \le x \le 6 \\ 0 & \text{otherwise} \end{cases}$	M1 A1 (2)
(d)	$E(X) = 3.5$	B1 (1)
(e)	Variance = $\frac{(6-1)^2}{12}$ or $\int_1^6 \frac{1}{5}x^2\ dx - (3.5)^2$ $= \frac{25}{12}$ awrt 2.08	M1 A1 (2)
(f)	$E(X^2) = Var(X) + [E(X)]^2$ $= \frac{25}{12} + 3.5^2$ or $\int_1^6 \frac{1}{5}x^2\ dx$ or $\int_1^6 \frac{1}{5}(3x^2+1)\ dx$ $= \frac{43}{3}$ $E(3X^2 + 1) = 3\ E(X^2) + 1$ $= \left[\frac{3x^3}{15} + \frac{x}{5}\right]_1^6$ $= 44$ $= 44$	M1 dM1 A1cao (3)
	Notes	
(a)	M1 writing or using 1 – F(4) oe	
(c)	M1 for differentiating to get 1/5 A1 both lines correct with ranges	
(e)	M1 $\frac{(6-1)^2}{12}$ or $\int_1^6 \frac{1}{5}x^2\ dx$ – 'their 3.5'^{2}	
(f)	1st M1 "their Var(X)" + ["their E(X)"]^2 (which must follow from the 1st method in (e)) **or** $\int_1^6 \frac{1}{5}x^2\ dx$ **and** integrating $x^n \to \frac{x^{n+1}}{n+1}$ (may be seen in (e)) **or** writing $\int_1^6 \frac{1}{5}(3x^2+1)\ dx$ (May be implied by $\frac{43}{3}$ seen) 2nd M1 (dependent on previous M1) **using** 3 × 'their E(X^2)' + 1 **or** $\int_1^6 \frac{1}{5}(3x^2+1)\ dx$ and integrating $x^n \to \frac{x^{n+1}}{n+1}$	

<table>
<tr><th>Question Number</th><th>Scheme</th><th>Marks</th></tr>
<tr><td>3(a)</td><td>(A random variable) that is a function of a (random) sample involving no unknown quantities/parameters
or
A quantity calculated solely from a random sample</td><td>B1
(1)</td></tr>
<tr><td>(b)</td><td>If all possible samples are chosen from a population;
then the values of a statistic and the associated probabilities is a sampling distribution
or a probability distribution of a statistic</td><td>B1
(1)</td></tr>
<tr><td>(c)</td><td>$\text{Mean} = 100\times\frac{4}{7}+200\times\frac{3}{7}$
$=\frac{1000}{7}$ awrt 143
$\text{Variance} = 100^2\times\frac{4}{7}+200^2\times\frac{3}{7}-\left(\frac{1000}{7}\right)^2$
$=\frac{120000}{49}$ awrt 2450 (to 3sf)</td><td>B1
M1
A1
(3)</td></tr>
<tr><td>(d)</td><td>(100,100,100)
(100,100,200) (100,200,100) (200,100,100) or 3 x (100,100,200)
(100,200,200) (200,100,200) (200,200,100) or 3 x (100,200,200)
(200,200,200)</td><td>B2
(2)</td></tr>
<tr><td>(e)</td><td>(100,100,100) $\left(\frac{4}{7}\right)^3=\frac{64}{343}$ awrt 0.187
(200,200,200) $\left(\frac{3}{7}\right)^3=\frac{27}{343}$ awrt 0.0787
(100,100,200) $3\times\left(\frac{4}{7}\right)^2\times\left(\frac{3}{7}\right)=\frac{144}{343}$ awrt 0.420 (allow 0.42)
(100,200,200) $3\times\left(\frac{4}{7}\right)\times\left(\frac{3}{7}\right)^2=\frac{108}{343}$ awrt 0.315
<table>
<tr><td>m</td><td>100</td><td>$400/3$
awrt 133</td><td>$500/3$
awrt 167</td><td>200</td></tr>
<tr><td>$P(M=m)$</td><td>$\frac{64}{343}$ or
awrt 0.187</td><td>$\frac{144}{343}$ or
awrt 0.420
(allow 0.42)</td><td>$\frac{108}{343}$ or
awrt 0.315</td><td>$\frac{27}{343}$ or
awrt 0.0787</td></tr>
</table>
</td><td>B1 both
M1
A1
A1
(4)</td></tr>
</table>

Question Number	Scheme	Marks
(a)	Notes B1 for a definition which includes each of the following 3 aspects A **function**[1] of a (random) **sample**[2] involving **no unknown quantities/parameters**[3] 1. function/quantity/calculation/value/random variable 2. sample/observations/data 3. no unknown parameters/no unknown values/solely (from a sample)	
(b)	B1 requires all underlined words: <u>All</u> <u>values</u> of a <u>statistic</u> with their associated <u>probabilities</u> **or** <u>probability distribution</u> of a <u>statistic</u>	
(c)	M1 $100^2 \times \frac{4}{7} + 200^2 \times \frac{3}{7} - (\text{their mean})^2$	
(d)	B1 any 2 of (100,100,100), (100,100,200) any order, (100,200,200) any order, (200,200,200) B1 all correct, allow 3 × (100,100,200) and 3 × (100,200,200) and (100,100,100) and (200,200,200) Note: Allow other notation for 100 and 200 e.g. Small and Large	
(e)	B1 Both probabilities for (100,100,100) and (200,200,200) correct M1 $3 \times p^2 \times (1-p)$ A1 either correct A1 all means correct **and** all probabilities correct (table not required but means must be associated with correct probabilities)	

Question Number	Scheme	Marks
4(a)	$X \sim \text{Po}(6)$ $\text{P}(5 \leq X < 7) = \text{P}(X \leq 6) - \text{P}(X \leq 4)$ or $\frac{e^{-6}6^5}{5!} + \frac{e^{-6}6^6}{6!}$ $= 0.6063 - 0.2851$ $= 0.3212$ awrt 0.321	M1 M1 A1 (3)
(b)	$H_0: \lambda = 9$ $H_1: \lambda < 9$ $X \sim \text{Po}(9)$ therefore $\text{P}(X \leq 4) = 0.05496...$ or CR $X \leq 3$ Insufficient evidence to reject H_0 **or** Not Significant **or** 4 does not lie in the critical region. There is no evidence that the mean number of accidents at the crossroads has reduced/decreased.	B1 B1 dM1 A1cso (4)
(a) **(b)**	**Notes** M1 writing or using Po(6) M1 either $\text{P}(X \leq 6) - \text{P}(X \leq 4)$ or $\frac{e^{-\lambda}\lambda^5}{5!} + \frac{e^{-\lambda}\lambda^6}{6!}$ 1st B1 both hypotheses correct (λ or μ) allow 0.5 instead of 9 2nd B1 either awrt 0.055 **or** critical region $X \leq 3$ dM1 for a correct comment (dependent on previous B1) Contradictory non-contextual statements such as "not significant" so "reject H_0" score M0 (May be implied by a correct contextual statement) A1 cso requires correct contextual conclusion with underlined words and all previous marks in (b) to be scored.	

Question Number	Scheme	Marks
5(a)	$\int_{-1}^{2} k(x^2+a)\mathrm{d}x + \int_{2}^{3} 3k\ \mathrm{d}x = 1$	M1
	$\left[k\left(\frac{x^3}{3}+ax\right)\right]_{-1}^{2} + \left[3kx\right]_{2}^{3} = 1$	dM1
	$k\left(\frac{8}{3}+2a+\frac{1}{3}+a\right)+9k-6k=1$	A1
	$6k+3ak=1$	
	$\int_{-1}^{2} k(x^3+ax)\mathrm{d}x + \int_{2}^{3} 3kx\ \mathrm{d}x \left[=\frac{17}{12}\right]$	M1
	$\left[k\left(\frac{x^4}{4}+\frac{ax^2}{2}\right)\right]_{-1}^{2} + \left[\frac{3kx^2}{2}\right]_{2}^{3} = \frac{17}{12}$	dM1
	$k\left(4+2a-\frac{1}{4}-\frac{a}{2}\right)+\frac{27k}{2}-6k=\frac{17}{12}$	A1
	$\frac{45k}{4}+\frac{3ak}{2}=\frac{17}{12}$ $135k+18ak=17$	
	$99k = 11$	ddM1
	$a=1, k=\frac{1}{9}$	A1
		(8)
(b)	2	B1
		(1)
(a)	Notes 1st M1 writing or using $\int_{-1}^{2} k(x^2+a)\mathrm{d}x + \int_{2}^{3} 3k\ \mathrm{d}x = 1$ ignore limits 2nd dM1 attempting to integrate at least one $x^n \to \frac{x^{n+1}}{n+1}$ **and** sight of correct limits (dependent on previous M1) 1st A1 a correct equation – need not be simplified 3rd M1 $\int_{-1}^{2} k(x^3+ax)\mathrm{d}x + \int_{2}^{3} 3kx\ \mathrm{d}x$ ignore limits 4th dM1 setting $=\frac{17}{12}$ **and** attempting to integrate at least one $x^n \to \frac{x^{n+1}}{n+1}$ **and** sight of correct limits (dependent on previous M1) 2nd A1 a correct equation – need not be simplified 5th ddM1 attempting to solve two simultaneous equations in a and k by eliminating 1 variable (dependent on 1st and 3rd M1s) 3rd A1 both a and k correct	

Question Number	Scheme	Marks
6. (a)	$P(X = 5) = {}^{20}C_5(0.3)^5(0.7)^{15}$ or $0.4164 - 0.2375$ $= 0.17886\ldots$ awrt 0.179	M1 A1 (2)
(b)	Mean = 6 sd = $\sqrt{20\times0.7\times0.3}$ $= 2.049\ldots$ awrt 2.05	B1 M1 A1 (3)
(c)	$H_0: p = 0.3$ $H_1: p > 0.3$ $X \sim B(20,0.3)$ $P(X \geq 8) = 0.2277$ or $P(X \geq 10) = 0.0480$, so CR $X \geq 10$ Insufficient evidence to reject H_0 **or** Not Significant **or** 8 does not lie in the critical region. There is no evidence to support the Director (of Studies') belief/There is no evidence that the proportion of parents that do not support the new curriculum is greater than 30%	B1 M1 A1 dM1 A1cso (5)
(d)	$X \sim B(2n, 0.25)$ $X \sim B(8, 0.25)$ $P(X \geq 4) = 0.1138$ $X \sim B(10, 0.25)$ $P(X \geq 5) = 0.0781$ $2n = 10$ $n = 5$	M1 A1 A1 (3)
	Notes	
(a)	M1 ${}^{20}C_5(p)^5(1-p)^{15}$ or using $P(X \leq 5) - P(X \leq 4)$	
(b)	M1 use of $20\times0.7\times0.3$ (with or without the square root)	
(c)	B1 both hypotheses correct (p or π) M1 using $X \sim B(20,0.3)$ (may be implied by 0.7723, 0.2277, 0.8867 or 0.1133) A1 awrt 0.228 **or** CR $X \geq 10$ dM1 a correct comment (dependent on previous M1) A1 cso requires correct contextual conclusion with underlined words and all previous marks in (c) to be scored.	
(d)	M1 for 0.1138 or 0.0781 or 0.8862 or 0.9219 seen 1st A1 B(10, 0.25) selected (may be implied by $n = 10$ or $2n = 10$ or $n = 5$) An answer of 5 with no incorrect working seen scores 3 out of 3 Special Case: Use of a normal approximation, M1 for $\frac{(n-0.5)-\frac{n}{2}}{\sqrt{\frac{3}{8}n}} = z$ with $1.28 \leq z \leq 1.29$, 1st A1 for n=4.2/4.3, 2nd A1 for n=5	

Question Number	Scheme	Marks
7.	$Y \sim \text{N}\left(\frac{n}{5}, \frac{4n}{25}\right)$	B1
	$\text{P}(Y \geq 30) = \text{P}\left(Z > \frac{29.5 - n/5}{\frac{2}{5}\sqrt{n}}\right)$	M1 M1A1
	$\frac{29.5 - n/5}{\frac{2}{5}\sqrt{n}} = 2$	B1
	$n + 4\sqrt{n} - 147.5 = 0$ or $0.04n^2 - 12.44n + 870.25 = 0$	dM1
	$\sqrt{n} = 10.3...$ $n = 106.26\ldots$ or $n = 204.73\ldots$	A1
	$n = 106$	A1 cao (8)
	Notes 1st B1 writing or using $\text{N}\left(\frac{n}{5}, \frac{4n}{25}\right)$ 1st M1 writing or using 30 +/– 0.5 2nd M1 standardising using 29, 29.5, 30 or 30.5 and their mean and their sd 1st A1 fully correct standardisation (allow +/-) 2nd B1 for z = +/– 2 or awrt 2.00 must be compatible with their standardisation 3rd dM1 (dependent on 2nd M1) getting quadratic equation **and** solving leading to a value of $\sqrt{n}$ or n 2nd A1 awrt 10.3 **or** awrt (106 **or** 107 **or** 204 **or** 205) 3rd A1 for 106 only (must reject other solutions if stated) Note: $\frac{29.5 - n/5}{\frac{2}{5}\sqrt{n}} = -2$ leading to an answer of 106 may score B1M1M1A1B0M1A1A1	