Mark Scheme (Final)

Summer 2015

Pearson Edexcel International A Level in Statistics 3 (WST03/01)

Edexcel and BTEC Qualifications

Pearson: helping people progress, everywhere

Summer 2015
Publications Code IA042726

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.

PEARSON 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)
   - d… or dep – dependent
   - indep – independent
   - dp decimal places
   - sf significant figures
   - ∗ The answer is printed on the paper or ag- answer given
   - ⎣ or d… 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. If a candidate makes more than one attempt at any question:
   - If all but one attempt is crossed out, mark the attempt which is NOT crossed out.
   - If either all attempts are crossed out or none are crossed out, mark all the attempts and score the highest single attempt.

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

June 2015
WST03 Statistics 3 Mark Scheme

Question Number	Scheme		Marks
1. (a)	$\{w\} = 018$ or 18	**018 or 18**	B1 [1]
(b)	$\{x\} = 18$	**18**	B1 [1]
(c)	$\{\text{prob} =\}\ 0$	**0**	B1 [1]
(d)	**Advantage:** Any one of: • Simple or easy to use also allow "quick" or "efficient" (o.e.) • It is suitable for large samples (or populations) • Gives a good spread of the data		B1
	Disadvantage: Any one of: • The alphabetical list is (probably) not random • Biased since the list is not (truly) random • Some combinations of names are not possible		B1 [2] **(Total 5)**
	Notes		
(d)	If no labels are given treat the 1st reason as an advantage and the 2nd as a disadvantage 1st B1: for advantage 2nd B1: for disadvantage "it requires a sampling frame" is 2nd B0 since the alphabetical list is given. Note: Do not score both B1 marks for opposing advantages and disadvantages.		

Question Number	Scheme	Marks
2. (a)		A	B	C	L	N	R	S	T	Y | Judge 1: 6, 3, 4, 9, 2, 8, 1, 5, 7 Judge 2: 8, 4, 5, 7, 3, 9, 1, 2, 6 or S, N, B, C, T, A, Y, R, L Judge 1: 1, 2, 3, 4, 5, 6, 7, 8, 9 Judge 2: 1, 3, 4, 5, 2, 8, 6, 9, 7	M1
	$\sum d^2 = 4+1+1+4+1+1+0+9+1$	M1
	or $0+1+1+1+9+4+1+1+4 = 22$ $\qquad \sum d^2 = 22$	A1
	$r_s = 1 - \frac{6(22)}{9(80)}$; $= 0.8166666...$ $\qquad \frac{49}{60}$ or awrt **0.817**	M1; A1
		[5]
(b)	$H_0: \rho = 0$, $H_1: \rho > 0$	B1
	Critical Value = 0.7833 <u>or</u> CR: $r_s \geqslant 0.7833$ $\qquad$ **0.7833**	B1
	Since $r_s = 0.8166...$ it lies in the CR, <u>or</u> reject H_0 (o.e.)	M1
	The two <u>judges</u> (or "they") are in <u>agreement</u> <u>or</u> there is a <u>positive correlation</u> between the ranks of the two <u>judges</u>.	A1ft
		[4] **(Total 9)**

Notes

(a)	1st M1 for an attempt to rank at least one row (at least 4 correct) 2nd M1 for an attempt at d^2 row (may be implied by sight of $\sum d^2 = 22$ or 221 for reverse ranks) 1st A1 for $\sum d^2 = 22$ (or 221 if reverse ranking is used) Can be implied by correct answer. 3rd M1 for use of the correct formula with their $\sum d^2$ (if it is clearly stated) If the answer is not correct then a correct expression is required
False Ranking	e.g Alphabetic ranking: Gives Judge 1: 7 5 2 3 8 1 9 6 4 Judge 2: 7 8 5 2 3 9 4 1 6 $\qquad \sum d^2 = 162$ and $r_s = -0.35$ **Scores:** M0(for ranking), M1(for attempt at d^2 row), A0, M1 (for use of their $\sum d^2$), A0 i.e. 2 out of 5 Can follow through their r_s in (b)
(b)	1st B1 for both hypotheses stated correctly in terms of ρ (allow ρ_s) H_1 must be compatible with ranking 2nd B1 for cv = 0.7833 (independent of their H_1 (no 2-tail value in tables) <u>but</u> compatible sign with their r_s) M1**:** for a correct statement (in words) relating their r_s with their critical value. e.g. "reject H_0 ","in critical region", "significant", "positive correlation" May be implied by a correct contextual comment.
|cv|>1	If their cv is |cv| > 1 (often from using normal tables) award M0A0 If |their r_s | > |their cv| then "significant" (o.e.) for M1 and "<u>judges</u> are in <u>agreement</u>" (o.e.) for A1ft If |their r_s | < |their cv| then "not significant" (o.e.) for M1 and "<u>judges</u> don't <u>agree</u>" (o.e.) for A1ft A1ft**:** for a correct follow through conclusion in context. "positive correlation" alone scores M1 A0 For reverse ranking should still say " judges <u>are</u> in agreement"

Question Number	Scheme	Marks
3. (a)	$\hat{\lambda} = \dfrac{0(47)+1(57)+2(46)+3(35)+4(9)+5(6)}{200} = \dfrac{320}{200} = 1.6$ * Full exp' or at least 2 products and 320/200 seen	B1 *
		[1]
(b)	$r = 200 \times \dfrac{e^{-1.6}(1.6)^2}{2!}$ $\{= 51.68550861...\}$ Using $r = 200 \times \dfrac{e^{-1.6}(1.6)^2}{2!}$	M1
	$s = 200 - (40.38 + 64.61 + \text{their } r + 27.57 + 11.03)$ $\{= 4.72449139...\}$ <u>or</u> their $r + s = 56.41$	M1
	$r = 51.68550861...$ and $s = 4.72449139...$ r = awrt **51.69** and s = awrt **4.72**	A1
		[3]
(c)	H_0: Poisson (distribution) is a suitable/ sensible (model) H_1: Poisson (distribution)is not a suitable/ sensible (model).	B1

Number of accidents	Observed	Expected	Combined Observed	Combined Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
0	47	40.38	47	40.38	1.0853	54.7053
1	57	64.61	57	64.61	0.8963	50.2863
2	46	51.69	46	51.69	0.6264	40.9364
3	35	27.57	35	27.57	2.0024	44.4324
4	9	11.03	15	15.75	0.0357	14.2857
$\geqslant 5$	6	4.72				
				Totals	4.6461	204.6461

	(table above)	M1
	$X^2 = \sum \dfrac{(O-E)^2}{E}$ or $\sum \dfrac{O^2}{E} - 200$;= 4.6461	M1;
	awrt **4.65**	A1
	$v = 5 - 1 - 1 = 3$ **3**	B1 ft
	$\chi^2_3(0.10) = 6.251 \Rightarrow$ CR: $X^2 \geqslant 6.251$ **6.251**	B1 ft
	[Since $X^2 = 4.6461$ does not lie in the CR, then there is insufficient evidence to reject H_0] The number of ***accidents*** per day can be modelled by a Poisson distribution <u>or</u> the ***supervisor's*** belief is correct.	A1 ft
		[7] **(Total 11)**

Notes

(b)	Note: Allow A1 for s = awrt 4.74 (found as a result of using expected values to full accuracy.)
(c)	1st B1: for <u>both</u> hypotheses and mentioning Poisson at least once. Allow Poisson is a "good fit/model" but <u>not</u> "good method" Inclusion of 1.6 for mean in hypotheses is B0 but condone in conclusion.
	1st M1: For an attempt to pool 4 accidents and $\geqslant 5$ accidents <u>or</u> pool when $E_i < 5$ No pooling is M0
	2nd M1: For an attempt at the test statistic, at least 2 correct expressions/values (to awrt 2 d.p.)
	1st A1: For awrt 4.65 (score M1M1A1 if awrt 4.65 seen)
No pooling	If no pooling can allow 2nd M1 if $X^2 = 5.33$ is seen
	2nd B1ft: For $n-1-1$ i.e. subtracting 2 from their n. 3rd B1ft: For a correct ft for their $\chi^2_k(0.10)$, where $k = n-1-1$ from their n. B1B1 may be implied by 6.251 (if pooling) or 7.779 for no pooling
	2nd A1ft**:** (***Dep. on the*** 2nd M1) For correct comment in context based on their test statistic and their critical value that mentions ***accidents*** or ***supervisor***. Condone mention of Po(1.6) in conclusion Score A0 for inconsistencies e.g. "significant" followed by "supervisor's belief is justified"
	Note: Full accuracy gives a combined expected frequency of 15.76, $\dfrac{(O-E)^2}{E} = 0.0366$, $\dfrac{O^2}{E} = 14.2766$, $X^2 = 4.64855...$ and p-value 0.199

Question Number	Scheme		Marks
4. (a)	Let X = weight of a sack of potatoes, $X \sim \mathrm{N}(25.6\ ,\ 0.24^2)$		
	So $D = X_1 - X_2 \sim \mathrm{N}(0\ ,\ 2(0.24)^2)$ or $D \sim \mathrm{N}(0\ ,\ 0.1152)$	Attempt at D and $D \sim \mathrm{N}(0, ..)$	M1
		$(0.24)^2 + (0.24)^2$; **0.1152**	A1; A1
	$\{\mathrm{P}(\lvert D\rvert > 0.5) =\}\ 2\,\mathrm{P}(D > 0.5)$	$2 \times \mathrm{P}(D > 0.5)$ can be implied	dM1
	$= 2 \times \mathrm{P}\left(Z > \dfrac{0.5}{\sqrt{0.1152}}\right)$		dM1
	$= 2 \times \mathrm{P}(Z > 1.4731...)$ <u>or</u> $= 2(1 - 0.9292)$		
	$= 0.1416$	**awrt 0.141** or **awrt 0.142**	A1
			[6]
(b)	Let Y = weight of an empty pallet, $Y \sim \mathrm{N}(20.0\ ,\ 0.32^2)$		
	So $T = X_1 + X_2 + \ ... \ + X_{30} + Y$		
	$T \sim \mathrm{N}(30(25.6) + 20,\ 30(0.24)^2 + 0.32^2)$	$30(25.6) + 20$ <u>or</u> **788**	B1
		$30(0.24)^2 + 0.32^2$	M1
	$T \sim \mathrm{N}(788, 1.8304)$	N and 1.8304 or awrt **1.83**	A1
	$\{\mathrm{P}(T > 785) =\}\ \mathrm{P}\left(Z > \dfrac{785 - 788}{\sqrt{1.8304}}\right)$		M1
	$= \mathrm{P}(Z > -2.2174...)$		
	$= 0.9868$	**awrt 0.987**	A1
			[5]
			(Total 11)

	Notes
(a)	1st M1: For clear definition of D and normal distribution with mean of 0 (Can be implied by 3rd M1)
	1st A1: for correct use of Var($X_1 - X_2$) formula
	2nd A1: for 0.1152
	2nd dM1: For realising need $2\times \mathrm{P}(D > 0.5)$ (Dependent on 1st M1 i.e. must be using suitable D)
	3rd dM1: Dep on 1st M1 for standardising with 0.5, 0 and their s.d.($\neq 0.24$)Must lead to $\mathrm{P}(Z > +\mathrm{ve})$ (o.e.)
	P(Z > 1.47) implies 1st M1 1st A1 2nd A1 and 3rd M1
	Correct answer only will score 6 out of 6
(b)	B1: For a mean of $30(25.6) + 20$. Can be implied by 788.
	1st M1: For $30(0.24)^2 + 0.32^2$. Can be implied by 1.8304 or awrt 1.83
	Allow M1 for swapping error i.e. $30 \times 0.32^2 + 0.24^2$ if the expression is seen
	1st A1: For normal and correct variance of 1.8304 or awrt 1.83.
	Normality may be implied by standardisation
	2nd M1: For standardising with 785 with their mean and st. dev..($\neq 0.24$) Must lead to $\mathrm{P}(Z > -\mathrm{ve})$ oe.
	2nd A1: awrt 0.987
	Correct answer only will score 5 out of 5
	Note: Calculator answers are (a) 0.14071... , (b) 0.98670...

Question Number	Scheme		Marks
5.	H_0 : Grades and gender are independent (or not associated) H_1 : Grades and gender are dependent (or associated)	"grades" and "gender" mentioned at least once.	B1 (1)
	(Observed table below)	An attempt to convert percentages to observed frequencies.	M1
		All observed frequencies are correct.	A1
	(Expected table below)	Some attempt at $\frac{\text{(Row Total)(Column Total)}}{\text{(Grand Total)}}$ Can be implied by a correct E_i	M1
		All expected frequencies are correct to nearest integer.	A1
	(Calculation table below)	At least 2 correct terms for $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ or correct expressions with their E_i. Accept 2 sf accuracy for the M1 mark.	M1
		All correct $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ terms to either 2 dp or better. Allow truncation. ($\Rightarrow$ by awrt 5.1 if 3rd M1 seen)	A1
	$X^2 = \sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 360$; = awrt 5.1	awrt **5.1**	A1 (7)
	$\nu = (3-1)(2-1) = 2$	($\nu =$) **2** (Can be implied by 5.991)	B1
	$\chi_2^2(0.05) = 5.991 \Rightarrow$ CR: $X^2 \geqslant 5.991$	For **5.991** only	B1
	Since $X^2 = 5.1$ does not lie in the CR, then there is insufficient evidence to reject H_0		M1
	Business Studies grades and gender are independent or There is no association between Business Studies grades and gender. Or Head of department's (belief) is correct		A1ft (4)
			[12] **(Total 12)**

Observed	Male	Female
Distinction	37	44
Merit	127	96
Unsatisfactory	36	20

Expected	Male	Female	Totals
Distinction	45	36	81
Merit	123.889	99.111	223
Unsatisfactory	31.111	24.889	56
Totals	200	160	360

Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
37	45	1.422	30.422
44	36	1.778	53.778
127	123.889	0.078	130.189
96	99.111	0.098	92.987
36	31.111	0.768	41.657
20	24.889	0.960	16.071
	Totals	5.104	365.104

	Notes
	Final M1: For a correct statement linking their test statistic and their critical value (> 3.8) Note: Contradictory statements score M0. E.g. "significant, do not reject H_0 ". Final A1ft: For a correct ft statement in context – must mention "grades" and "gender" or "sex" or "head of department" Condone "relationship" or "connection" here but **not** "correlation". e.g. "There is no evidence of a relationship between grades and gender"
5.10 only	Just seeing 5.10... only can imply 1st 3 Ms but loses 1st 3 As so can score 4 out of 7 (Qu says "show..")
	Note: Full accuracy gives $X^2 = 5.104356...$ and p-value 0.0779

<table>
<tr><th>Question Number</th><th>Scheme</th><th>Marks</th></tr>
<tr><td>5.</td><td><u>Mark Scheme for candidates who use percentages instead of observed values.</u>
H_0 : Grades and gender are independent (or not associated)
H_1 : Grades and gender are dependent (or associated)
"grades" and "gender" mentioned at least once.</td><td>B1</td></tr>
<tr><td></td><td>
<table>
<tr><th>Observed</th><th>Male</th><th>Female</th></tr>
<tr><td>Distinction</td><td>18.5</td><td>27.5</td></tr>
<tr><td>Merit</td><td>63.5</td><td>60.0</td></tr>
<tr><td>Unsatisfactory</td><td>18.0</td><td>12.5</td></tr>
</table>
These marks cannot be obtained.</td><td>M0 A0</td></tr>
<tr><td></td><td>
<table>
<tr><th>Expected</th><th>Male</th><th>Female</th><th>Totals</th></tr>
<tr><td>Distinction</td><td>23</td><td>23</td><td>46</td></tr>
<tr><td>Merit</td><td>61.75</td><td>61.75</td><td>123.5</td></tr>
<tr><td>Unsatisfactory</td><td>15.25</td><td>15.25</td><td>30.5</td></tr>
<tr><td>Totals</td><td>100</td><td>100</td><td>200</td></tr>
</table>
Some attempt at $\frac{\text{(Row Total)(Column Total)}}{\text{(Grand Total)}}$
Can be implied by one of these E_i 's
Expected frequencies are not correct.</td><td>M1
A0</td></tr>
<tr><td></td><td>
<table>
<tr><th>Observed</th><th>Expected</th><th>$\frac{(O-E)^2}{E}$</th><th>$\frac{O^2}{E}$</th></tr>
<tr><td>18.5</td><td>23</td><td>0.8804</td><td>14.8804</td></tr>
<tr><td>27.5</td><td>23</td><td>0.8804</td><td>32.8804</td></tr>
<tr><td>63.5</td><td>61.75</td><td>0.0496</td><td>65.2996</td></tr>
<tr><td>60.0</td><td>61.75</td><td>0.0496</td><td>58.2996</td></tr>
<tr><td>18.0</td><td>15.25</td><td>0.4959</td><td>21.2459</td></tr>
<tr><td>12.5</td><td>15.25</td><td>0.4959</td><td>10.2459</td></tr>
<tr><td></td><td>Totals</td><td>2.8518</td><td>202.8518</td></tr>
</table>
At least 2 "correct" terms for $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ or correct expressions with their E_i . Accept 2 sf accuracy for the M1 mark.
This mark cannot be obtained.</td><td>M1
A0</td></tr>
<tr><td></td><td>$X^2 = \sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E} - 360$;= 2.8518
This mark cannot be obtained.</td><td>A0</td></tr>
<tr><td></td><td>$\nu = (3-1)(2-1) = 2$ ($\nu =$) 2 (Can be implied by 5.991)</td><td>B1</td></tr>
<tr><td></td><td>$\chi^2_2(0.05) = 5.991 \Rightarrow$ CR: $X^2 \geqslant 5.991$ For 5.991 only</td><td>B1</td></tr>
<tr><td></td><td>Since $X^2 = 2.86$ does not lie in the CR, then there is insufficient evidence to reject H_0</td><td>M1</td></tr>
<tr><td></td><td>Not available since comes from incorrect working.</td><td>A0</td></tr>
<tr><td></td><td></td><td>[12]
(Total 12)</td></tr>
<tr><td></td><td>Notes</td><td></td></tr>
<tr><td></td><td>If a candidate uses percentages rather than observed values then they can obtain a maximum of 6 marks.
They can get B1 M0A0 M1A0 M1A0A0 B1B1M1A0.</td><td></td></tr>
</table>

Question Number	Scheme	Marks
6. (a)	$\left\{\hat{\mu} = \frac{\sum x}{n} = \frac{1570}{50} = \right\}$ $\bar{x} = 31.4$ $\bar{x} =$ **31.4**	B1 cao
	$\left\{\hat{\sigma}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \right\}$ $s_x^2 = \frac{49467.58 - 50(31.4)^2}{50-1}$	M1 A1ft
	$= 3.460816...$ awrt **3.46**	A1
		[4]
(b)	[Let Y = time taken to complete obstacle course in the afternoon.]	
	$H_0 : \mu_x = \mu_y$, $H_1 : \mu_x > \mu_y$	B1
	$(z =) \frac{"31.4" - 30.9}{\sqrt{\frac{"3.46"}{50} + \frac{3.03}{50}}}$	M1 A1ft
	$= 1.38781...$ awrt **1.39**	A1
	CR: $Z \geqslant 1.6449$ or probability = awrt 0.082 or awrt 0.083 **1.6449** or better	B1
	Since $z = 1.38781...$ does not lie in the CR, then there is insufficient evidence to reject H_0	M1
	Conclude that the <u>mean time</u> to complete the obstacle course is the same for the early <u>morning</u> and late <u>afternoon</u>.	A1
		[7]
(c)	$\bar{X}$ <u>and</u> $\bar{Y}$ are both approx. <u>normal</u>ly distributed <u>or</u> $\bar{X} - \bar{Y}$ normal (Condone $\bar{x}$ and $\bar{y}$)	B1
		[1]
(d)	Have assumed $s^2 \simeq \sigma^2$ or variance of sample $\simeq$ variance of population	B1
		[1]
		(Total 13)

	Notes
(a)	B1: 31.4 cao Allow 31 minutes, 24 seconds. 1st M1: A correct expression for either s or s^2 (ignore label) 1st A1ft: A correct expression for s^2 with their ft $\bar{x}$. 3rd A1: awrt 3.46 (Correct answer scores 3 out of 3)
(b)	1st B1: Both hypotheses stated correctly, with some indication of which μ is which. Eg: μ_M , μ_A 1st M1: For an attempt at $\frac{a-b}{\sqrt{\frac{c}{50} + \frac{d}{50}}}$ with at least 3 of a, b, c or d correct. Allow $\pm$ 1st A1ft: for $\pm \frac{\text{their } 31.4 - 30.9}{\sqrt{\frac{\text{their } 3.46}{50} + \frac{3.03}{50}}}$ Allow $D = \bar{x} - \bar{y}$ $1.64 \sim 1.65 = \frac{D - 0}{\sqrt{\frac{"3.46"}{50} + \frac{3.03}{50}}}$ [SE = 0.360277..] 2nd A1: for awrt 1.39 (possibly $\pm$)(Allow for CV D = awrt 0.593) (NB d = 0.5) Correct answer scores M1A1ftA1 <u>but</u> $0 - (31.4 - 30.9) \rightarrow -1.39$ loses this 2nd A mark 2nd B1: Critical value of 1.6449 or better (seen). Allow for probability = awrt 0.082 or awrt 0.083 Note: p-values are 0.0823 (tables) and 0.0826 (calculator). 2nd M1: For a correct statement linking their test statistic and their critical value. Note: Contradictory statements score M0. E.g. "significant, do not reject H_0 ". 3rd A1: For a correct statement in context that accepts H_0 (no ft) Condone "no difference in mean times" Must mention "<u>mean time</u>", "<u>morning</u>" and "<u>afternoon</u>" or "<u>both</u> <u>times</u> <u>of day</u>"
(c)	B1 E.g. $\bar{X} \sim N(...)$ need both. Allow in words e.g "sample means are normally distributed"
(d)	B1 condone only mentioning "x" or "y" <u>but</u> watch out for $s_x = s_y$ or $\sigma_x = \sigma_y$ which scores B0

Question Number	Scheme	Marks
7.	Let X = score on a die	
(a)	$\text{E}(S) = 3.5$, $\text{Var}(S) = \frac{35}{12}$ $\text{E}(S) =$ **3.5**	B1
	$\text{Var}(S) = \frac{35}{12}$ or awrt **2.92**	B1
		[2]
(b)	So, $\bar{S} \sim \text{N}\left(\text{"3.5"}, \frac{\text{"}\left(\frac{35}{12}\right)\text{"}}{40}\right)$ or $\bar{S} \sim \text{N}\left(\text{"3.5"}, \frac{7}{96}\right)$	B1ft
	$\text{P}\left(\bar{S} < 3\right) = \text{P}\left(Z < \frac{3 - \text{"3.5"}}{\sqrt{\frac{7}{96}}}\right)$ $\{= \text{P}(Z < -1.85164...)\}$	M1
	$\{= 1 - 0.9678\} = 0.0322$ **0.032 to 0.0322**	A1
		[3] **(Total 5)**

	Notes
(a)	2nd B1 allow awrt 2.92
(b)	B1ft for $\bar{S} \sim \text{N}\left(\text{"3.5"}, \frac{\text{"}\left(\frac{35}{12}\right)\text{"}}{40}\right)$ seen or implied. Follow through their E(S) and their Var(S) NB $\frac{7}{96} = 0.07291\dot{6}$ accept awrt 0.0729 M1 for an attempt to standardise with 3, their mean (>3) and $\sqrt{\frac{\text{their Var}(S)}{40}}$. Must lead to P($Z <$ −ve) A1 for 0.032 ~ 0.0322
ALT ΣS	B1ft for $\sum S \sim \text{N}\left(140, \frac{350}{3}\right)$ where 140 is $40 \times$ their E(S) and variance is $40 \times$ their Var(S) M1 for $\text{P}\left(Z < \frac{120 - \text{"140"}}{\sqrt{\frac{350}{3}}}\right)$ or $\text{P}\left(Z < \frac{119.5 - \text{"140"}}{\sqrt{\frac{350}{3}}}\right)$ {= P($Z <$ −1.8979...)} A1 for 0.032~0.0322 or (with continuity correction) 0.0287 (tables) or 0.0289 (calculator).

Question Number	Scheme	Marks
8. (a)	$\left\{\bar{x} = \dfrac{29.74 + 31.86}{2}\right\} \Rightarrow \bar{x} = 30.8$ $\bar{x} = 30.8$ This can be implied. See note.	B1
	$"1.96"\left(\dfrac{\sigma}{\sqrt{n}}\right) = 31.86 - 30.8$ or $2("1.96")\left(\dfrac{\sigma}{\sqrt{n}}\right) = 31.86 - 29.74$	M1
	$\text{SE}_{\bar{x}} = \dfrac{31.86 - 30.8}{1.96} = 0.540816... = 0.54\ (2\text{dp})$ awrt **0.54**	A1
		[3]
(b)	A 90% CI for μ is $\bar{x} \pm 1.6449\left(\dfrac{\sigma}{\sqrt{n}}\right)$	B1
	$= 30.8 \pm 1.6449(0.54)$ $(\text{their } \bar{x}) \pm (\text{their } z)(\text{their SE}_{\bar{x}} \text{ from (a)})$	M1
	$= (29.91,\ 31.69)$ (awrt **29.9**, awrt **31.7**)	A1
		[3]
(c)	Let X = number of confidence intervals containing μ or Y = number of confidence intervals not containing μ	
	So $X \sim \text{Bin}(4, 0.9)$ or $Y \sim \text{Bin}(4, 0.1)$	M1
	$\text{P}(X \geqslant 3)$ or $\text{P}(Y \leqslant 1) = {}^4C_3(0.9)^3(0.1) + (0.9)^4$ ${}^4C_3(0.9)^3(0.1) + (0.9)^4$	A1 oe
	$= 0.2916 + 0.6561 = 0.9477$ **0.9477** or **0.948**	A1
		[3] **(Total 9)**

	Notes
(a)	B1: $\bar{x} = 30.8$ may be implied by $1.96\left(\dfrac{\sigma}{\sqrt{n}}\right) = [31.86 - 30.8] = 1.06$ <u>or</u> $2(1.96)\left(\dfrac{\sigma}{\sqrt{n}}\right) = 31.86 - 29.74$
	M1: A correct equation for either a width or a half-width involving a z-value $1.5 \leqslant z \leqslant 2$ Eg: $"\text{their } z"\left(\dfrac{\sigma}{\sqrt{n}}\right) = 31.86 - "30.8"$ ft their $\bar{x}$ <u>or</u> $2("\text{their } z")\left(\dfrac{\sigma}{\sqrt{n}}\right) = 31.86 - 29.74$ or $"\text{their } z"(\text{SE}_{\bar{x}}) = 31.86 - "30.8"$ <u>or</u> $2("\text{their } z")(\text{SE}_{\bar{x}}) = 31.86 - 29.74$ are fine for M1.
	A1: 0.54 or awrt 0.54 Must be seen as final answer to (a) NB $\dfrac{53}{98}$ as final answer is A0 Condone $\bar{x} \pm 1.96\sigma = ...$ for B1 and M1 but A0 even if they say "σ = standard error = 0.54" Otherwise answer only of 0.54 scores 3 out of 3
(b)	B1 for use of 1.6449 or better in an attempt at a CI formula. Need at least 1.6449 (their SE) M1 for attempt at CI ft their values and provided $1 \leqslant z \leqslant 1.7$
(c)	M1: States or applies either $X \sim \text{Bin}(4, 0.9)$ <u>or</u> $Y \sim \text{Bin}(4, 0.1)$ Condone M1 for $0.9^4 + 0.9^3 \times 0.1$ (o.e.) 1st A1: ${}^4C_3(0.9)^3(0.1) + (0.9)^4$ or $(0.9)^4 + {}^4C_1(0.1)(0.9)^3$ oe 2nd A1: 0.9477 or 0.948

G. B. Attwood 30/05/15

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom