

WST03/01: Statistics S3

Question Number	Scheme	Marks
Q2	<p style="text-align: center;">[$P \sim N(90,9)$ and $J \sim N(91,12)$]</p> <p>(a) $(J - P) \sim N(1, 21)$ $P(J < P) = P(J - P < 0)$ $= P\left(Z < \frac{0-1}{\sqrt{21}}\right)$ $= P(Z < -0.2182\dots)$ $= 1 - 0.5871 = 0.4129$ calculator (0.4136....) awrt (0.413 ~ 0.414)</p> <p>(b) $X = (J_1 + J_2 + \dots + J_{60}) - (P_1 + P_2 + \dots + P_{60})$ $E(X) = 60 \times 91 - 60 \times 90 = 60$ [stated as $E(X) = 60$ or $X \sim N(60, \dots)$] $\text{Var}(X) = 60 \times 9 + 60 \times 12 = 1260$ $P(X > 120) = P\left(Z > \frac{120 - 60}{\sqrt{1260}}\right)$ $= P(Z > 1.69030\dots)$ $= 1 - 0.9545 = 0.0455$ awrt (0.0455)</p>	<p>M1, A1</p> <p>dM1</p> <p>A1</p> <p>(4)</p> <p>M1</p> <p>B1</p> <p>A1</p> <p>M1</p> <p>A1</p> <p>(5)</p> <p>9</p>
Use of means	<p>(a) 1st M1 for attempting $J - P$ and $E(J - P)$ or $P - J$ and $E(P - J)$ 1st A1 for variance of 21 (Accept $9 + 12$). Ignore any slip in μ here. 2nd dM1 for attempting the correct probability and standardising with their mean and sd. This mark is dependent on previous M so if $J - P$ (or $P - J$) is not being used score M0 If their method is not crystal clear then they must be attempting $P(Z < -ve \text{ value})$ or $P(Z > +ve \text{ value})$ i.e. their probability <u>after</u> standardisation should lead to a prob. < 0.5 so e.g. $P(J - P < 0)$ leading to 0.5871 is M0A0 unless the M1 is clearly earned. 2nd A1 for awrt 0.413 or 0.414</p> <p style="text-align: center;">The first 3 marks may be implied by a correct answer</p> <p>(b) 1st M1 for a clear attempt to identify a correct form for X. This may be implied by correct variance of 1260 B1 for $E(X) = 60$. Can be awarded even if they are using $X = 60J - 60P$. Allow $P - J$ and -60 1st A1 for a correct variance. If 1260 is given the M1 is scored by implication. 2nd M1 for attempting a correct probability and standardising with 120 and their 60 and 1260 If the answer is incorrect a full <u>expression</u> must be seen following through their values for M1 e.g. $P\left(Z > \frac{120 - \text{their } 60}{\sqrt{\text{their variance}}}\right)$. If using -60, should get $P\left(Z < \frac{-120 - -60}{\sqrt{\text{their variance}}}\right)$</p> <p>Attempt to use $\bar{J} - \bar{P}$ for 1st M1, $E(\bar{J} - \bar{P}) = 1$ for B1 and $\text{Var}(\bar{J} - \bar{P}) = 0.35$ for A1 Then 2nd M1 for standardisation with 2, and their 1 and 0.35</p>	

WST03/01: Statistics S3

Question Number	Scheme		Marks
Q3	(a)	$E \sim N(0, 0.5^2)$ or $X \sim N(w, 0.5^2)$ $P(E < 0.6) = P\left(Z < \frac{0.6}{0.5}\right)$ or $P(X - w < 0.6) = P\left(Z < \frac{0.6}{0.5}\right)$ $= P(Z < 1.2)$ $= 2 \times 0.8849 - 1 = 0.7698$ awrt 0.770	M1 A1 (2)
	(b)	$\bar{E} \sim N\left(0, \frac{1}{64}\right)$ or $\bar{X} \sim N\left(w, \frac{0.5^2}{16}\right)$ $P(\bar{E} < 0.3) = P\left(Z < \frac{0.3}{\frac{1}{8}}\right)$ or $P(\bar{X} - w < 0.3) = P\left(Z < \frac{0.3}{\frac{1}{8}}\right)$ $= P(Z < 2.4)$ $= 2 \times 0.9918 - 1 = 0.9836$ awrt 0.984	M1 M1, A1 A1 (4)
	(c)	$35.6 \pm 2.3263 \times \frac{1}{8}$ (35.3, 35.9)	M1 B1 A1, A1 (4) 10
	(a)	1 st M1 for identifying a correct probability (they must have the 0.6) and attempting to standardise. Need . This mark can be given for 0.8849 - 0.1151 seen as final answer. 1 st A1 for awrt 0.770. NB an answer of 0.3849 or 0.8849 scores M0A0 (since it implies no) M1 may be implied by a correct answer	
	(b)	1 st M1 for a correct attempt to define \bar{E} or \bar{X} but must attempt $\frac{\sigma^2}{n}$. Condone labelling as E or X This mark may be implied by standardisation in the next line. 2 nd M1 for identifying a correct probability statement using \bar{E} or \bar{X} . Must have 0.3 and 1 st A1 for correct standardisation as printed or better 2 nd A1 for awrt 0.984 The M marks may be implied by a correct answer.	
Sum of 16, not means		1 st M1 for correct attempt at suitable sum distribution with correct variance ($= 16 \times \frac{1}{4}$) 2 nd M1 for identifying a correct probability. Must have 4.8 and 1 st A1 for correct standardisation i.e. need to see $\frac{4.8}{\sqrt{4}}$ or better	
	(c)	M1 for $35.6 \pm z \times \frac{0.5}{\sqrt{16}}$ B1 for 2.3263 or better. Use of 2.33 will lose this mark but can still score $\frac{3}{4}$ 1 st A1 for awrt 35.3 2 nd A1 for awrt 35.9	

WST03/01: Statistics S3

Question Number	Scheme	Marks																																
Q4	<table border="1"> <tr> <td>Distance rank</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> </tr> <tr> <td>Depth rank</td> <td>1</td> <td>2</td> <td>4</td> <td>3</td> <td>6</td> <td>7</td> <td>5</td> </tr> <tr> <td>d</td> <td>0</td> <td>0</td> <td>1</td> <td>1</td> <td>1</td> <td>1</td> <td>2</td> </tr> <tr> <td>d^2</td> <td>0</td> <td>0</td> <td>1</td> <td>1</td> <td>1</td> <td>1</td> <td>4</td> </tr> </table> <p> $\sum d^2 = 8$ $r_s = 1 - \frac{6 \times 8}{7 \times 48}$ $= \frac{6}{7} = 0.857142$ </p>	Distance rank	1	2	3	4	5	6	7	Depth rank	1	2	4	3	6	7	5	$ d $	0	0	1	1	1	1	2	d^2	0	0	1	1	1	1	4	<p>M1</p> <p>M1</p> <p>M1A1</p> <p>M1</p> <p>A1</p> <p>(6)</p>
Distance rank	1	2	3	4	5	6	7																											
Depth rank	1	2	4	3	6	7	5																											
$ d $	0	0	1	1	1	1	2																											
d^2	0	0	1	1	1	1	4																											
(a)	<p>$H_0 : \rho = 0, H_1 : \rho > 0$</p> <p>Critical value at 1% level is 0.8929</p> <p>$r_s < 0.8929$ so not significant evidence to reject H_0.</p> <p>The researcher's claim is not correct (at 1% level).</p> <p><u>or</u> insufficient evidence for researcher's claim</p> <p><u>or</u> there is insufficient evidence that water gets deeper further from inner bank.</p> <p><u>or</u> no (positive) correlation between depth of water and distance from inner bank</p>	<p>B1</p> <p>B1</p> <p>M1</p> <p>A1ft</p> <p>(4)</p>																																
(a)	<p>1st M1 for an attempt to rank the depths against the distances</p> <p>2nd M1 for attempting d for their ranks. Must be using ranks.</p> <p>3rd M1 for attempting $\sum d^2$ (must be using ranks)</p> <p>1st A1 for sum of 8 (or 104 for reverse ranking)</p> <p>4th M1 for use of the correct formula with their $\sum d^2$. If answer is not correct an expression is required.</p> <p>2nd A1 for awrt (\pm) 0.857. Sign should correspond to ranking (so use of 104 should get -0.857)</p>																																	
(b)	<p>1st B1 for both hypotheses in terms of ρ, H_1 must be one tail and compatible with their ranking</p> <p>2nd B1 for cv of 0.8929 (accept \pm)</p> <p>M1 for a correct statement relating their r_s with their cv but cv must be such that $cv < 1$</p> <p>A1ft for a correct contextualised comment. Must mention "researcher" and "claim" <u>or</u> "distance (from bank)" and "depth (of water)"</p> <p>Follow through their r_s and their cv (provided it is $cv < 1$)</p> <p>Use of "association" is A0</p>																																	
		10																																

WST03/01: Statistics S3

Question Number	Scheme					Marks																												
Q5	<table border="1"> <thead> <tr> <th>Finances</th> <th>Worse</th> <th>Same</th> <th>Better</th> <th></th> </tr> </thead> <tbody> <tr> <td>Income</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Under £15 000</td> <td>10.54</td> <td>10.54</td> <td>12.92</td> <td>34</td> </tr> <tr> <td>£15 000 and above</td> <td>20.46</td> <td>20.46</td> <td>25.08</td> <td>66</td> </tr> <tr> <td></td> <td>31</td> <td>31</td> <td>38</td> <td>100</td> </tr> </tbody> </table>					Finances	Worse	Same	Better		Income					Under £15 000	10.54	10.54	12.92	34	£15 000 and above	20.46	20.46	25.08	66		31	31	38	100	M1 A1			
	Finances	Worse	Same	Better																														
	Income																																	
	Under £15 000	10.54	10.54	12.92	34																													
	£15 000 and above	20.46	20.46	25.08	66																													
		31	31	38	100																													
	H_0 : State of finances and income are independent (not associated) H_1 : State of finances and income are not independent (associated)					B1																												
	<table border="1"> <thead> <tr> <th>O_i</th> <th>E_i</th> <th>$\frac{(O_i - E_i)^2}{E_i}$</th> <th>$\frac{O_i^2}{E_i}$</th> </tr> </thead> <tbody> <tr> <td>14</td> <td>10.54</td> <td>1.1358....</td> <td>18.59..</td> </tr> <tr> <td>11</td> <td>10.54</td> <td>0.0200....</td> <td>11.48..</td> </tr> <tr> <td>9</td> <td>12.92</td> <td>1.1893...</td> <td>6.269..</td> </tr> <tr> <td>17</td> <td>20.46</td> <td>0.5851...</td> <td>14.12..</td> </tr> <tr> <td>20</td> <td>20.46</td> <td>0.0103...</td> <td>19.55..</td> </tr> <tr> <td>29</td> <td>25.08</td> <td>0.6126...</td> <td>33.53..</td> </tr> </tbody> </table>					O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$	14	10.54	1.1358....	18.59..	11	10.54	0.0200....	11.48..	9	12.92	1.1893...	6.269..	17	20.46	0.5851...	14.12..	20	20.46	0.0103...	19.55..	29	25.08	0.6126...	33.53..	M1 A1
	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$																														
	14	10.54	1.1358....	18.59..																														
11	10.54	0.0200....	11.48..																															
9	12.92	1.1893...	6.269..																															
17	20.46	0.5851...	14.12..																															
20	20.46	0.0103...	19.55..																															
29	25.08	0.6126...	33.53..																															
$\sum \frac{(O_i - E_i)^2}{E_i} = 3.553... \text{ or } \sum \frac{O_i^2}{E_i} - 100 = 103.553... - 100 = 3.553... \text{ (awrt 3.55)}$					A1																													
$v = (3 - 1)(2 - 1) = 2$ cv is 5.991					B1 B1																													
$3.553 < 5.991$ so insufficient evidence to reject H_0 or not significant There is no evidence of association between state of finances and income.					M1 A1																													
10																																		
1 st M1	for some use of $\frac{\text{Row Total} \times \text{Col. Total}}{\text{Grand Total}}$. May be implied by correct E_i																																	
1 st A1	for all expected frequencies correct																																	
B1	for both hypotheses. Must mention "state" or "finances" and "income" at least once Use of "relationship" or "correlation" or "connection" is B0																																	
2 nd M1	for at least two correct terms (as in 3 rd or 4 th column) or correct expressions with their E_i																																	
2 nd A1	for all correct terms. May be implied by a correct answer. (2 dp or better - allow eg 1.13...)																																	
3 rd M1	for a correct statement linking their test statistic and their cv. Must be χ^2 not normal.																																	
4 th A1	for a correct comment in context - must mention "state" or "finances" and "income" condone "relationship" or "connection" here but not "correlation". No follow through. e.g. "There is no evidence of a relationship between finances and income"																																	

WST03/01: Statistics S3

Question Number	Scheme						Marks	
Q6	Distance from centre of site (m)	0-1	1-2	2-4	4-6	6-9	9-12	
	$b-a$	1	1	2	2	3	3	M1
	No of artefacts	22	15	44	37	52	58	A1
	$P(a \leq X < b)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$	A1
	$228 \times P(a \leq X < b)$	19	19	38	38	57	57	A1
	Class	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$			
	0-1	22	19	$\frac{9}{19} = 0.4736\dots$	25.57...			M1
	1-2	15	19	$\frac{16}{19} = 0.8421\dots$	11.84...			
	2-4	44	38	$\frac{36}{38} = 0.9473\dots$	50.94...			
	4-6	37	38	$\frac{1}{38} = 0.0263\dots$	36.02...			
6-9	52	57	$\frac{25}{57} = 0.4385\dots$	47.43...			A1	
9-12	58	57	$\frac{1}{57} = 0.0175\dots$	59.01...			B1	
H_0 : <u>continuous uniform</u> distribution <u>is</u> a good fit H_1 : <u>continuous uniform</u> distribution <u>is not</u> a good fit							B1	
$\sum \frac{(O_i - E_i)^2}{E_i} = \frac{313}{114} = 2.75$ <u>or</u> $\sum \frac{O_i^2}{E_i} - 228 = 230.745\dots - 228 = \dots$ (awrt 2.75)							dM1A1	
$\nu = 6 - 1 = 5$							B1	
$\chi^2_5(0.05) = 11.070$ (ft their ν i.e. $\chi^2_\nu(0.05)$)							B1ft	
2.75 < 11.070, insufficient evidence to reject H_0							M1	
Continuous uniform distribution is a suitable model							A1	
							12	
1 st M1	for calculation of at least 3 widths and attempting proportions/probs. <u>or</u> for 1:2:3 ratio seen							
1 st A1	for correct probabilities							
2 nd A1	for all correct expected frequencies							
2 nd M1	for attempting $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$, at least 3 correct expressions or values.							
Follow through their E_i provided they are not all = 38								
3 rd A1	for a correct set of calcs - 3 rd or 4 th column. (2 dp or better and allow e.g. 0.94...)							
3 rd dM1	dependent on 2nd M1 for attempting a correct sum or calculation (must see at least 3 terms and +)							
4 th M1	The first three Ms and As can be implied by a test statistic of awrt 2.75 for a correct statement based on their test statistic (> 1) and their cv (> 3.8) Contradictory statements score M0 e.g. "significant" do not reject H_0 .							
5 th A1	for a correct comment suggesting that continuous uniform model is suitable. No ft							

WST03/01: Statistics S3

Question Number	Scheme	Marks
Q7	(a) Label full time staff 1-6000, part time staff 1-4000 Use random numbers to select Simple random sample of 120 full time staff and 80 part time staff	M1 M1 A1 (3)
	(b) Enables estimation of statistics / errors for each strata <u>or</u> “reduce variability” <u>or</u> “more representative” <u>or</u> “reflects population structure” NOT “more accurate”	B1 (1)
	(c) $H_0: \mu_f = \mu_p, \quad H_1: \mu_f \neq \mu_p$ (accept μ_1, μ_2) $\text{s.e.} = \sqrt{\frac{21}{80} + \frac{19}{80}}, \quad z = \frac{52 - 50}{\sqrt{\frac{21}{80} + \frac{19}{80}}} = (2\sqrt{2})$ $= 2.828\dots$ (awrt 2.83)	B1 M1,M1 A1
	Two tailed critical value $z = 2.5758$ (or prob of awrt 0.002 (<0.005) or 0.004 (<0.01)) [2.828 > 2.5758 so] significant evidence to reject H_0 There is evidence of a difference in policy awareness between full time and part time staff	B1 dM1 A1ft (7)
	(d) Can use mean full time and mean part time ~ Normal	B1 B1 (2)
	(e) Have assumed $s^2 = \sigma^2$ or variance of sample = variance of population	B1 (1)
	(f) $2.53 < 2.5758$, not significant <u>or</u> do not reject H_0 So there is insufficient evidence of a difference in mean awareness	M1 A1ft (2)
	(g) Training course has closed the gap between full time staff and part time staff’s mean awareness of company policy.	B1 (1)
		17
	(a) 1 st M1 for attempt at labelling full-time and part-time staff. One set of correct numbers. 2 nd M1 for mentioning use of random numbers 1 st A1 for s.r.s. of 120 full-time and 80 part-time	
	(c) 1 st M1 for attempt at s.e. - condone one number wrong . NB correct s.e. = $\sqrt{\frac{1}{2}}$ 2 nd M1 for using their s.e. in correct formula for test statistic. Must be $\frac{\pm(52 - 50)}{\sqrt{\frac{p}{q} + \frac{r}{s}}}$ 3 rd dM1 dep. on 2nd M1 for a correct statement based on their normal cv and their test statistic 2 nd A1 for correct comment in context. Must mention “scores” or “ policy awareness” and types of “staff”. Award A0 for a one-tailed comment. Allow ft	
	(d) 1 st B1 for mention of mean(s) <u>or</u> use of \bar{X} , provided \bar{X} clearly refers to full-time or part-time 2 nd B1 for stating that distribution can be assumed normal e.g. “mean score of the test is normally distributed” gets B1B1	
	(f) M1 for correct statement (may be implied by correct contextualised comment) A1 for correct contextualised comment. Accept “no difference in mean scores”. Allow ft	
	(g) B1 for correct comment in context that implies training was effective. This must be supported by their (c) and (f). Condone one-tailed comment here.	