Mark Scheme (Unused)

January 2022

Pearson Edexcel International A Level

In Statistics S3 (WST03) Paper 01

Question Number	Scheme	Marks
1 (a)	Number the 1200 students (1 – 1200)	B1
	Use a random starting point between 1 and 20	B1
	Select every 20th person on the list	B1
		(3)
(b)(i)	They only need to generate one random number	B1
		(1)
(b)(ii)	It is not random as the list is ordered alphabetically **or** not all combinations of sampling units are possible	M1
	e.g. unlikely siblings would be selected	A1
		(2)
(c)	Number of Y9 students = $\frac{200}{1200} \times 60$ [= 10]	M1
	The stratified sample gives a better proportion or is more representative oe	A1
		(2)
	Notes	**Total 8**

1 (a)	**B1**	numbering the students (Allow 0 – 1199).
	B1	using a random starting point. Must be between 1 and 20 (Allow 0 – 19).
	B1	selecting every 20th person.
(b)(i)	**B1**	a suitable comment.
(b)(ii)	**M1**	a suitable comment.
	A1	a suitable example.
(c)	**M1**	a suitable calculation to find the number of Y9 students e.g. $\frac{200}{1200} \times 60$
	A1	a correct explanation.

Question Number	Scheme	Marks
2 (a)	Use of $\bar{x} \pm z \times \frac{1.9}{\sqrt{10}}$; $z = 1.96$	M1;B1
	(52.54…, 54.897…) awrt 52.5 and 54.9	A1 A1
		(4)
(b)	Use of $1.5 > 2 \times z \times \frac{1.9}{\sqrt{n}}$ oe ; $z = 2.5758$ (or better)	M1;B1
	$1.5 > \frac{9.78804}{\sqrt{n}}$	dM1
	$n > 42.58...$ So $n = 43$	A1
		(4)

	Notes		Total 8
2 (a)	**M1**	for use of correct expression with 1.9, 10 and $1 < z < 3$	
	B1	for $z = 1.96$	
	A1	for awrt 52.5	
	A1	for awrt 54.9	
(b)	**M1**	use of $z \times \frac{1.9}{\sqrt{n}}$ in a correct inequality with 0.75 or 1.5 and $2 < z < 3$ (allow written as an equation)	
	B1	for $z = 2.5758$ (or better)	
	dM1	dependent on 1st M1, for solving a correct inequality for the width of the 99% CI (allow an equation rather than an inequality)	
	A1	cao	

Question Number	Scheme	Marks
3 (a)	**Driver**: *A*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*, *J* **Rank FQL**: 1, 5, 3, 2, 6, 4, 8, 9, 10, 7 **FP**: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	M1
	$\sum d^2 = 0+9+0+4+1+4+1+1+1+9$ [=30]	M1
	$r_s = 1 - \frac{6(30)}{10(99)}$	dM1
	$= 0.8181818...$ awrt 0.818	A1
		(4)
(b)	$H_0 : \rho = 0$, $H_1 : \rho > 0$	B1
	Critical Value r_s = 0.7455 or CR: $r_s \geqslant 0.7455$	B1
	Reject H_0 or significant or lies in the critical region	M1
	There is sufficient evidence of a positive correlation between fastest qualifying **lap time** and **finishing position** for these Formula One racing drivers	A1
		(4)

		Notes	Total 8
3 (a)	**M1**	attempt to rank fastest qualifying lap (at least four correct).	
	M1	finding the difference between each of the ranks and evaluating $\sum d^2$	
	dM1	dependent on 1st M1. Using $1 - \frac{6\sum d^2}{10(99)}$ with their $\sum d^2$	
	A1	$\frac{9}{11}$ or awrt 0.818	
(b)	**B1**	both hypotheses correct. Must be in terms of ρ. Must be attached to H_0 and H_1	
	B1	critical value of 0.7455	
	M1	A correct statement comparing their CV with their r_s - no context needed but do not allow contradicting non contextual comments.	
	A1	correct conclusion which is rejecting H_0, which must mention **lap time** and **finishing position**.	

Question Number	Scheme	Marks
4	H_0 : There is no association between type of property and the time taken to sell it H_1 : There is an association between type of property and the time taken to sell it	B1

Expected	**Bungalow**	**Flat**	**House**	**Total**
Within 3 months	10.496	31.488	40.016	(82)
More than 3 months	5.504	16.512	20.984	(43)
Total	(16)	(48)	(61)	(125)

M1 A1

Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
7	10.496	1.1644…	4.6684…
29	31.488	0.1965…	26.7085…
46	40.016	0.8948…	52.8788…
9	5.504	2.2205…	14.7165…
19	16.512	0.3748…	21.8628…
15	20.984	1.7064…	10.7224…
	Totals	6.557…	131.557…

dM1 A1

Scheme	Marks
$\left[X^2 = \right]\sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E} - 125$	dM1
$= 6.557...$ awrt 6.56	A1
$\nu = (2-1)(3-1) = 2$	B1
$\chi^2_2(0.05) = 5.991 \Rightarrow$ CR: $X^2 \geqslant 5.991$	B1
[in the CR/significant/Reject H_0] There is sufficient evidence to suggest that there is **an association** between **type** of property and the **time** taken to sell it.	A1
	(10)

	Notes		**Total 10**
4	**B1**	Both hypotheses correct. Must mention "type of property" **and** "time taken" at least once. (may be written in terms of independence)	
	M1	Some attempt at $\frac{\text{(Row Total)(Column Total)}}{\text{(Grand Total)}}$ Can be implied by at least one correct E_i to 1dp	
	A1	All expected frequencies correct	
	dM1	Dependent on 1st M1 for 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.	
	A1	At least 3 correct $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$ terms to 2dp or better. Allow truncated answers.	
	dM1	Dependent on 2nd M1 For applying either $\sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E} - 125$	
	A1	awrt 6.56	
	B1	$\nu = 2$ This mark can be implied by a correct critical value of 5.991	
	B1	5.991	
	A1	Dependent on the 3rd M1 and 3rd B1. A correct contextualised conclusion which is rejecting H_0 Must mention **type** and **time.** Contradictory statements score A0. e.g. "significant, do not reject H_0 ".Condone "relationship" or "connection" here but **not** "correlation".	

Question Number	Scheme	Marks
5 (a)(i)	$\left[\bar{x} = \frac{3610}{50} \Rightarrow\right]$ $\bar{x} = 72.2$ $\quad s_x^{\ 2} = \frac{260955.6 - 50(72.2)^2}{50-1} = 6.4$	B1; M1 A1
5(a)(ii)	$\left[\bar{y} = \frac{2585}{50} \Rightarrow\right]$ $\bar{y} = 51.7$ $\quad s_y^{\ 2} = \frac{133757.2 - 50(51.7)^2}{50-1} = 2.3$	B1 A1
		(5)
(b)	$H_0 : \mu_x - \mu_y = 20$ $H_1 : \mu_x - \mu_y > 20$	B1
	$z = \frac{'72.2' - '51.7' - 20}{\sqrt{\frac{'6.4'}{50} + \frac{'2.3'}{50}}}$	M1 M1
	$= 1.1986...$ awrt 1.20	A1
	One tailed c.v. $Z = 1.6449$ **or** CR: $Z \ldots 1.6449$	B1
	Not in CR/Not significant/Do not reject H_0	M1
	No significant evidence to support **Tammy's belief**	A1
		(7)
(c)	Since the sample is **large** the **CLT** applies.	M1
	No need to assume (the weights) are normally distributed.	A1
		(2)
(d)	Assumed that $s^2 = \sigma^2$	B1
		(1)
	Notes	**Total 15**

5 (a)(i)	**B1**	$\bar{x} = 72.2$
	M1	A correct method for finding an unbiased estimate of the variance e.g. $\frac{\sum x^2 - n(\bar{x})^2}{n-1}$ (May be seen in (i) or (ii))
	A1	6.4
5(a)(ii)	**B1**	$\bar{y} = 51.7$
	A1	2.3
(b)	**B1**	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ
	M1	For correct standard error. Follow through their values from (a)
	M1	An attempt at $\frac{a-b-20}{\sqrt{\frac{c}{50}+\frac{d}{50}}}$ with at least 2 of *a*, *b*, *c* or *d* correct. Allow ±
	A1	awrt 1.20 Allow 1.2 if no incorrect working shown
	B1	1.6449 or better (seen)
	M1	A correct statement – need not be contextual but do not allow contradicting non contextual comments.
	A1	A correct contextual statement. Allow the **difference** in mean weights is **not greater than 20 kg**
(c)	**M1**	A suitable comment that mentions large and CLT
	A1	A correct answer, context not required.
(d)	**B1**	for the assumption that sample variance = population variance

Question Number	Scheme	Marks
6 (a)	$\frac{0\times1+1\times10+2\times23+3\times15+4\times19+5\times9+6\times3}{80}=3$ *	B1
		(1)
(b)	$r=e^{-3}\times80=3.983$ $s=\frac{e^{-3}\times3^5}{5!}\times80$;$=8.066$	M1 ; A1
	$t=80-\left(r+11.949+17.923+17.923+13.443+s\right)$; $=6.713$	M1 ; A1
		(4)
(c)	H_0: Poisson (distribution) is a reasonable/suitable/ sensible (model) H_1: Poisson (distribution) is not a /reasonable/suitable/ sensible (model).	B1
	(table below)	M1
	$X^2=\sum\frac{(O-E)^2}{E}$ or $\sum\frac{O^2}{E}-80$	M1
	$=7.901...$ awrt 7.90	A1
	$v=6-1-1=4$	B1
	$\chi^2_4(0.10)=7.779 \Rightarrow$ CR: $X^2 \geqslant 7.779$	B1
	[since $X^2=7.90$ does lie in CR, then there is sufficient evidence to reject H_0]	
	Sufficient evidence to say that Poisson is not a reasonable model	A1
		(7)
	Notes	**Total 12**

Number of emails	Combined Observed	Combined Expected	$\frac{(O-E)^2}{E}$	$\frac{O^2}{E}$
$\leqslant 1$	11	15.932	1.5267…	7.5947…
2	23	17.923	1.4381…	29.5151…
3	15	17.923	0.4767…	12.5537…
4	19	13.443	2.2971…	26.8541…
5	9	8.065	0.1083…	10.0433…
$\geqslant 6$	3	6.714	2.0544…	1.3404…
		Totals	7.901…	87.901…

6 (a)	**B1**	For a correct method to shown that the mean is 3
(b)	**M1**	Use of $\frac{e^{-\lambda}\times\lambda^r}{r!}\times80$ or May be implied by a correct answer for either r or s
	A1	$r=3.983$ **and** $s=8.066$ (allow $r=3.984$ **and** $s=8.064$ as these come from tables)
	M1	A correct method that ensures that expected totals = 80
	A1	$t=6.713$ (allow $t=6.714$ if tables used)
(c)	**B1**	Both hypotheses correct. Must mention Poisson at least once.
	M1	Combining 0 emails and 1 email. Must have both observed and expected frequencies
	M1	An attempt at the test statistic, at least 2 correct expressions/values (to awrt 2dp)
	A1	awrt 7.90 Accept 7.9 if no incorrect working seen
	B1	$v=4$ This mark can be implied by a correct critical value of 7.779
	B1	7.779
	A1	A correct conclusion based on their X^2 value and their χ^2 critical value

Question Number	Scheme	Marks
7 (a)	Let X represent $B_1 + B_2 - C_1$	
	$X \sim \text{N}(0.268, 0.015633)$ awrt 0.0156	M1 A1
	$\text{P}(X < 0) = \text{P}\left(Z < \frac{0 - 0.268}{\sqrt{"0.015633"}}(= -2.14)\right)$	M1
	$(= 1 - 0.9838) = 0.0162$	A1
		(4)
(b)	Let Y represent $2.5B_1 + 3C_1 + 3C_2$	
	$Y \sim \text{N}(6.918, 0.071478)$ awrt 6.92, 0.0715	M1 A1
	$\text{P}(Y > 7) = \text{P}\left(Z > \frac{7 - "6.918"}{\sqrt{"0.071478"}}(= 0.31)\right)$	M1
	$(= 1 - 0.6217) = 0.3783$ (Calculator gives 0.3795…) 0.378 – 0.380	A1
		(4)
(c)	Mean = $2.94w$	B1
	Standard deviation = $0.084\sqrt{5}\ w \quad (= 0.188w)$	B1
		(2)
(d)	$\frac{6 - 2.94w}{0.084\sqrt{5}\ w} \leqslant -1.2816$	M1;B1
	$-1.2816 \times 0.084\sqrt{5}\ w + 2.94w \geqslant 6$	dM1
	$w \geqslant 2.22\ldots$ So $w = 2.23$	A1
		(4)
	Notes	**Total 14**

7 (a)	**M1**	for setting up normal distribution with mean 0.268
	A1	for a correct expression for variance (= 0.015633) or for standard deviation (= 0.125…)
	M1	for standardising with 0, 0.268 and their standard deviation
	A1	awrt 0.0162 (Allow awrt 0.0160 as this comes from a calculator)
(b)	**M1**	for setting up normal distribution with mean awrt 6.92
	A1	for a correct expression for variance (= 0.071478) or for standard deviation (= 0.267…)
	M1	for standardising with 7, 0.071478 and their standard deviation
	A1	for answer between 0.378 – 3.80
(c)	**B1**	for $2.94w$
	B1	for $0.084\sqrt{5}w$ or awrt $0.188w$
(d)	**M1**	for standardising using their mean and their standard deviation = z where $1 < \lvert z \rvert < 1.5$
	B1	for -1.28
	dM1	dependent on M1, for solving their inequality
	A1	awrt (£)2.23