Mark Scheme (Results)

Summer 2022

Pearson Edexcel International Advanced Level
In Statistics S3 (WST03) Paper 01

Question Number	Scheme	Marks
1 (a)	You would assign an average rank between the tied ranks	B1
		(1)
(b)	Rank for total tournaments 1 3 4 6 8 9 2 5 10 7	M1
	$\sum d^2 = 0+1+1+4+9+9+25+9+1+9$ [$= 68$]	M1
	$r_s = 1 - \dfrac{6 \times '68'}{10\left(10^2 - 1\right)}$	dM1
	$= 0.5878\ldots$ awrt 0.588	A1
		(4)
(c)	$H_0 : \rho = 0$, $H_1 : \rho > 0$	B1
	Critical Value = 0.5636 or CR … 0.5636	B1
	Reject H_0 or significant or lies in the critical region	dM1
	There is sufficient evidence of a positive correlation between rank and total tournaments won	A1
		(4)
(d)	2.5% and $r_s = 0.6485$ or CR … 0.6485	B1
		(1)
	Notes	**Total 10**
(a)	**B1**	for an appropriate explanation of how to deal with tied ranks. Ignore any comments regarding PMCC Do not allow add 0.5 to both ranks
(b)	**M1**	attempt to rank total tournaments (at least four correct) Condone reversed ranks
	M1	finding the difference between players rank and each of their total tournaments ranks and evaluating $\sum d^2$ May be implied by 68
	dM1	dependent on 1st M1. Using $1 - \dfrac{6\sum d^2}{10(99)}$ with their $\sum d^2$ (you will need to check their $\sum d^2$ if no value shown)
	A1	awrt 0.588 Allow $\frac{97}{165}$
(c)	**B1**	both hypotheses correct. Must be in terms of ρ. Must be attached to H_0 and H_1 If r_s is negative in part (b) then allow $H_1 : \rho < 0$
	B1	critical value of 0.5636 If r_s is negative in part (b) then allow –0.5636
	dM1	dependent on 2nd B1. A correct statement ft their part (b) and their CV– no context needed but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.
	A1	correct conclusion which is rejecting H_0, which must mention **rank** and **total tournaments**. No hypotheses is A0.
		NB If they have used $H_1 : \rho < 0$ then the maximum they can score is B1B1dM1A0
(d)	**B1**	for 2.5% and a correct critical value of 0.6485

Question Number	Scheme	Marks
2 (a)	$\bar{x} = \left[\dfrac{7690}{100}\right] = 76.9$	B1
	$s_x^{\,2} = \dfrac{669.24}{99} = 6.76$	M1 A1
		(3)
(b)	$H_0 : \mu_x = \mu_y \qquad H_1 : \mu_x \neq \mu_y$	B1
	$Z = \dfrac{"76.9"- 75.9}{\sqrt{\dfrac{"6.76"}{100} + \dfrac{2.2^2}{80}}} = 2.793...$ awrt ± 2.79	M1 M1 A1
	2 tailed critical value $z = \pm 2.5758$	B1
	Reject H_0/Significant/In the critical region	M1
	There is sufficient evidence to suggest that the mean <u>water temperature</u> after 4 hours for brand <u>A</u> is different to brand <u>B</u>	A1ft
		(7)
(c)	(It is reasonable) since both samples are (reasonably) large	B1
		(1)
	Notes	**Total 11**
(a)	**B1** for 76.9	
	M1 for use of $\dfrac{1}{n-1}\sum\left(x - \bar{x}\right)^2$ oe	
	A1 for 6.76	
(b)	**B1** for both hypotheses correct. Must be attached to H_0 and H_1 Allow equivalent hypotheses. Must be in terms of μ Allow any letter for the subscripts	
	M1 for a correct method to find the standard error. Follow through their values from (a)	
	M1 an attempt at $\pm\dfrac{a-b}{\sqrt{\dfrac{c}{100} + \dfrac{d^2}{80}}}$ with at least 3 of a, b, c or d correct.	
	A1 awrt ± 2.79	
	B1 $z = $ awrt ± 2.5758 seen (Allow $z = $ awrt ± 2.3263 if a one tailed test is used)	
	M1 a correct statement consistent with their CV and Z value – need not be contextual but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.	
	A1ft This mark is dependent on the 2nd M mark being awarded. A correct contextual statement ft their CV and their Z value	
(c)	**B1** a correct explanation, which makes reference to both samples. e.g. Do not allow the sample is large enough	

Question Number	Scheme	Marks
3 (a)	$\left[\dfrac{26.624 + 28.976}{2}\right] = 27.8$	B1
		(1)
(b)	$28.976 - 26.624 = 2 \times 1.96 \times \dfrac{\sigma}{\sqrt{25}}$ or $26.624 = \,'27.8'-1.96 \times \dfrac{\sigma}{\sqrt{25}}$ or $28.976 = \,'27.8'+1.96 \times \dfrac{\sigma}{\sqrt{25}}$	M1 B1
	$\sigma = 3\,*$	A1* cso
		(3)
(c)	$2 \times z \times \dfrac{3}{\sqrt{25}} = 2.1$ So $z = 1.75$	M1 A1
	$P(Z > '1.75') = P(Z < -'1.75') = 1 - '0.9599' = '0.0401'$	M1 A1ft
	Confidence level $= 100 \times (1 - 2 \times '0.0401') = 91.98\%$	M1 A1
		(6)
(d)	$2 \times 1.96 \times \dfrac{3}{\sqrt{n}} < 1.5$	M1
	$\sqrt{n} > \dfrac{6 \times 1.96}{1.5}$	dM1
	$\sqrt{n} > $ awrt 7.84 So $n = 62$	A1 A1
		(4)
	Notes	**Total 14**
(a)	**B1**	for 27.8
(b)	**M1**	for $28.976 - 26.624 = 2 \times z$ value $\times \dfrac{\sigma}{\sqrt{25}}$ or $26.624 = \,'27.8'- z$ value $\times \dfrac{\sigma}{\sqrt{25}}$ or $28.976 = \,'27.8'- z$ value $\times \dfrac{\sigma}{\sqrt{25}}$ where $1.5 < z < 2.4$
	B1	awrt 1.96
	A1* cso	answer is given so no incorrect working must be seen
(c)	**M1**	for $2 \times z \times \dfrac{3}{\sqrt{25}} = 2.1$
	A1	for $z = 1.75$
	M1	for $1 - p$, where p is a probability
	A1ft	for 0.0401 or ft their z value (Allow 0.04)
	M1	for $100 \times (1 - 2 \times 0.0401)$ ft their $P(Z < -1.75)$
	A1	awrt 92.0 (allow 92)
(d)	**M1**	for $2 \times z$ value $\times \dfrac{3}{\sqrt{n}} < 1.5$ oe z value must either be correct or consistent with part (b) Allow $\leqslant$ or $=$ Condone $>$ or $\geqslant$
	dM1	Dependent on previous M mark. Correct rearrangement to get $\sqrt{n} > ...$ or $n > ...$ oe Allow $\geqslant$ or $=$ Condone $<$ or $\leqslant$
	A1	awrt 7.84 may be implied by awrt 61.5
	A1	for $n = 62$

Question Number	Scheme	Marks
4 (a)	[Continuous] uniform on the interval [0, 7]	B1
		(1)
(b)	mean = 3.5	B1
	standard deviation = $\sqrt{\dfrac{(7-0)^2}{12}}$	M1
	$= \dfrac{7}{\sqrt{12}} = 2.0207...$ awrt 2.02	A1
		(3)
(c)	By the CLT $\bar{T} \sim N\left(3.5, \ \dfrac{49}{552}\right)$	M1
	$P\left(3.4 < \bar{T} < 3.6\right) = P\left(\dfrac{3.4 - "3.5"}{"\sqrt{\dfrac{49}{552}}"} < Z < \dfrac{3.6 - "3.5"}{"\sqrt{\dfrac{49}{552}}"}\right) = \left[P\left(-0.34 < Z < 0.34\right)\right]$	M1 A1
	$= 0.6331 - (1 - 0.6331)$ (Calculator gives 0.6314…)	M1
	$= 0.2662$ (Calculator gives 0.2628…) awrt 0.263 to 0.266	A1
		(5)
(d)	Large/ independent/ random sample allows use of CLT	B1
		(1)
	Notes	**Total 10**
(a)	**B1**	For the correct distribution stated (need uniform and correct interval) Allow U[0, 7] A fully correct pdf implies B1 e.g. $f(x) = \begin{cases} \dfrac{1}{7} & 0 \leqslant x \leqslant 7 \\ 0 & \text{otherwise} \end{cases}$
(b)	**B1**	For 3.5
	M1	For a correct method for finding the standard deviation
	A1	awrt 2.02 (Allow $\frac{7}{\sqrt{12}}$ or $\frac{7\sqrt{3}}{6}$ oe)
(c)	**M1**	For writing or using $N\left(3.5, \ \dfrac{49}{552}\right)$ oe Allow $N\left(3.5, \ \dfrac{2.02^2}{46}\right)$ or ft from part (b) e.g. if Po(7) given in part (a) allow $N\left(7, \ \dfrac{7}{46}\right)$
	M1	For standardising using either 3.4 or 3.6 and their mean and standard deviation
	A1	For a fully correct expression for either 3.4 or 3.6. May be implied by $\pm$ awrt 0.34
	M1	For $p - (1 - p)$ or $2(p - 0.5)$ oe
	A1	awrt 0.263 to 0.266
(d)	**B1**	Any suitable assumption

Question Number		Scheme	Marks
5 (a)		It is not a statistic as it involves <u>unknown</u> [population parameters]	B1
			(1)
(b)		An estimator for μ is unbiased if its <u>expected</u> value is equal to μ	B1
			(1)
(c)		$E(U_1) = 3E(X_1) - 2E(X_2)$ or $E(U_2) = \dfrac{1}{4}\left(E(X_1) + 3E(X_2)\right)$	M1
		$E(U_1) = 3\mu - 2\mu = \mu$ (therefore unbiased)	A1cso
		$E(U_2) = \dfrac{1}{4}(\mu + 3\mu) = \mu$ (therefore unbiased)	A1cso
			(3)
(d)		$\text{Var}(U_1) = 9\text{Var}(X_1) + 4\text{Var}(X_2)$ or $\text{Var}(U_2) = \dfrac{1}{16}\text{Var}(X_1) + \dfrac{9}{16}\text{Var}(X_2)$	M1
		$\left[\text{Var}(U_1) = \right]13\sigma^2$	A1
		$\left[\text{Var}(U_2) = \right]\dfrac{5}{8}\sigma^2$	A1
		As $\text{Var}(U_1) > \text{Var}(U_2)$ $\quad U_2$ is the most efficient estimator for μ	A1
			(4)
		Notes	**Total 9**
(a)	**B1**	for a correct explanation, must include unknown	
(b)	**B1**	for a correct explanation that refers to expected *X*. Allow $\mu - E(X) = 0$, but bias = 0 is B0	
(c)	**M1**	for use of $aE(X_1) + bE(X_2)$ May be implied by $3\mu - 2\mu$ or $\dfrac{1}{4}(\mu + 3\mu)$	
	A1cso	for a correct solution for $E(U_1)$ with no incorrect working Condone missing notation. Condone missing subscripts	
	A1cso	for a correct solution for $E(U_2)$ with no incorrect working seen Condone missing notation. Condone missing subscripts	
(d)	**M1**	for use of $a^2\text{Var}(X_1) + b^2\text{Var}(X_2)$	
	A1	Allow $9\sigma^2 + 4\sigma^2$	
	A1	Allow $\dfrac{1}{16}\sigma^2 + \dfrac{9}{16}\sigma^2$ or $\dfrac{5}{8}\sigma^2$ oe	
	A1	for U_2 with a correct reason	
		NB It is possible to score M1 A0 A0 A1 if $\text{Var}(U_1)$ and $\text{Var}(U_2)$ are correct	

Question Number	Scheme	Marks
6 (a)	$M \sim \text{N}(80,100)$ $\qquad$ $W \sim \text{N}(69,25)$ $X = M_1 + M_2 + M_3 + M_4 + M_5 + M_6 + W_1 + W_2 + W_3$	
	$X \sim \text{N}(687, 675)$	M1 A1
	$\text{P}(X > 700) = \text{P}\left(Z > \dfrac{700 - 687}{\sqrt{675}} \right) = \text{P}(Z > 0.500...)$	M1
	$\left(= 1 - 0.6915 \right) = 0.3085$ (Calculator gives 0.3084)	A1
		(4)
(b)	Let Y = Number of men in the lift	
	$Y \sim \text{N}(80x, 100x)$	M1
	$\text{P}(Y > 700) = \text{P}\left(Z > \dfrac{700 - 80x}{10\sqrt{x}} \right) < 0.025$	M1
	$\dfrac{700 - 80x}{10\sqrt{x}} > 1.96$	B1
	$80x + 19.6\sqrt{x} - 700[< 0]$ $\qquad$ $\mid$ $\qquad$ $6400x^2 - 112384.16x + 490000[> 0]$	M1
	Solving leading to $\sqrt{x} < 2.838...$ $\qquad$ $\mid$ $\qquad$ Solving leading to $x < 8.05...$	M1
	So $c = 8$ (people)	A1
		(6)
	Notes	**Total 10**
(a)	**B1** $\quad$ for setting up normal distribution with mean 687	
	B1 $\quad$ for a correct variance (675) or for standard deviation ($15\sqrt{3}$)	
	M1 $\quad$ for standardising with 700, 687 and their standard deviation	
	A1 $\quad$ for answer between 0.308 – 0.309	
(b)	**M1** $\quad$ for setting up normal distribution with mean $80x$ and variance $100x$ (may be implied by use of sd = $10\sqrt{x}$) Allow any letter	
	M1 $\quad$ for standardising with 700, their mean and their standard deviation (if not stated then these must be correct)	
	B1 $\quad$ for an equation or inequality set = to 1.96 (Allow – 1.96)	
	M1 $\quad$ for a correct 3TQ ft their mean and standard deviation	
	M1 $\quad$ for an attempt to solve their 3TQ with either $\sqrt{x} < ...$ or $x < ...$ Allow = instead of < Condone > or $\geqslant$ If the answer is incorrect then we must see use of the quadratic formula/completing the square (Allow one error)	
	A1 $\quad$ cao	

Question Number	Scheme					Marks
7 (a)	H_0 : The observed distribution can be modelled by a discrete uniform distribution H_1 : The observed distribution cannot be modelled by a discrete uniform distribution					B1 (1)
(bi)		Observed	Expected	$\dfrac{(O-E)^2}{E}$	$\dfrac{O^2}{E}$		B1 M1
		$x+6$	x	$\dfrac{36}{x}$	$\dfrac{(x+6)^2}{x}$		
		$x-8$	x	$\dfrac{64}{x}$	$\dfrac{(x-8)^2}{x}$		
		$x+8$	x	$\dfrac{64}{x}$	$\dfrac{(x+8)^2}{x}$		
		$x-5$	x	$\dfrac{25}{x}$	$\dfrac{(x-5)^2}{x}$		
		$x+4$	x	$\dfrac{16}{x}$	$\dfrac{(x+4)^2}{x}$		
		$x-5$	x	$\dfrac{25}{x}$	$\dfrac{(x-5)^2}{x}$		
		Total $= 6x$	Total $= 6x$	Total $= \dfrac{230}{x}$	Total $= \dfrac{6x^2+230}{x}$		

$X^2 = \sum \dfrac{(O-E)^2}{E}$ or $\sum \dfrac{O^2}{E} - 6x$; $\dfrac{230}{x}$ or $\dfrac{6x^2+230}{x} - 6x$ — **M1 ; A1**

$v = 6-1 = 5$; $c_5^2(0.05) = 11.070 \Rightarrow$ CR: $X^2 \ldots 11.070$ — **B1 ; B1**

Do not reject H_0 if $\dfrac{230}{x}$,, $'11.070'$ or $\dfrac{6x^2+230}{x} - 6x$,, $'11.070'$ — **M1**

$x \ldots 20.7768\ldots$ So $x = 21$ — **A1 (8)**

| (bii) | Hence the die was rolled "21" $\times$ 6 $=$ 126 times | | | | | M1 A1
 (2) |

Notes						Total 11

(a)	**B1**	for both hypotheses correct Allow H_0 : the die is not biased H_1 : the die is biased
(bi)	**B1**	for expected frequency $= x$
	M1	for one correct $\dfrac{(O-E)^2}{E}$ or $\dfrac{O^2}{E}$ ft their expected frequency
	M1	for an attempt at X^2 ft their values (At least 4 of these need to be seen and added)
	A1	for either $\dfrac{230}{x}$ or $\dfrac{6x^2+230}{x} - 6x$
	B1	for $v = 6-1 = 5$ May be implied by a correct critical value
	B1	for a correct critical value ft their DOF (**NB** common error is $v = 4$ so $c_4^2(0.05) = 9.488$)
	M1	for either $\dfrac{230}{x}$,, their CV or $\dfrac{6x^2+230}{x} - 6x$,, their CV Allow $<$ rather than ,,
	A1	for $x = 21$ provided the previous M mark has been awarded
(bii)	**M1**	for their 21×6 Allow $6 \times x$ or the answer to $6 \times$ their value for x
	A1	cao