OXFORD
INTERNATIONAL
AQA EXAMINATIONS

INTERNATIONAL A-LEVEL
FURTHER MATHEMATICS
FM04

(9665/FM04) Unit FS2 Statistics

Mark scheme

January 2022

Version: 1.0 Final

221xFM04/MS

Mark schemes are prepared by the Lead Assessment Writer and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation events which all associates participate in and is the scheme which was used by them in this examination. The standardisation process ensures that the mark scheme covers the students' responses to questions and that every associate understands and applies it in the same correct way. As preparation for standardisation each associate analyses a number of students' scripts. Alternative answers not already covered by the mark scheme are discussed and legislated for. If, after the standardisation process, associates encounter unusual answers which have not been raised they are required to refer these to the Lead Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of students' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

Further copies of this mark scheme are available from oxfordaqaexams.org.uk

Key to mark scheme abbreviations

M	Mark is for method
m	Mark is dependent on one or more M marks and is for method
A	Mark is dependent on M or m marks and is for accuracy
B	Mark is independent of M or m marks and is for method and accuracy
E	Mark is for explanation
√ **or ft**	Follow through from previous incorrect result
CAO	Correct answer only
CSO	Correct solution only
AWFW	Anything which falls within
AWRT	Anything which rounds to
ACF	Any correct form
AG	Answer given
SC	Special case
oe	Or equivalent
A2, 1	2 or 1 (or 0) accuracy marks
–*x* EE	Deduct *x* marks for each error
NMS	No method shown
PI	Possibly implied
SCA	Substantially correct approach
sf	Significant figure(s)
dp	Decimal place(s)

Q	Answer	Marks	Comments
1(a)	$z_{\text{crit}} = 2.326$	B1	Condone 2.3263
		1	

Q	Answer	Marks	Comments
1(b)	$z = \dfrac{\lvert 10.1 - 9.5 \rvert}{\sqrt{\dfrac{4.2}{100} + \dfrac{4.8}{120}}}$	M1	**M1**: Correct numerator (positive or negative) **PI** by correct [±] z
		M1	**M1**: Correct denominator **PI** by correct [±] z
	$= 2.095[29..]$	A1	Exact value is $\dfrac{6\sqrt{205}}{41}$ Allow negative value $p = 0.01807$ If **M0 M0** awarded, allow **SC1** for use of t-statistic with value [±] 2.083
	$2.095 < 2.326$ Therefore do not reject H_0	B1ft	**ft** their critical value and their test statistic with the corresponding correct conclusion
		4	

Q	Answer	Marks	Comments
1(c)	Sample is large enough to approximate with a normal distribution	B1	Any correct assumption, such as correct reference to central limit theorem
		1	

	Question 1 Total	**6**	

Q	Answer	Marks	Comments
2	H_0: There is no association between location of a supporter and their predicted winning team H_1: There is an association between location of a supporter and their predicted winning team	**B1**	Both H_0 and H_1 correct
	<table><tr><th></th><th>Brazil</th><th>Germany</th><th>Other</th></tr><tr><td>Host country</td><td>**127.2**</td><td>**89.4**</td><td>**83.4**</td></tr><tr><td>Outside the host country</td><td>**84.8**</td><td>**59.6**</td><td>**55.6**</td></tr></table>	**M1** **A1**	At least three correct All correct
	$\sum \frac{(O-E)^2}{E} = \frac{(140-"127.2")^2}{"127.2"} + \frac{(101-"89.4")^2}{"89.4"} + \frac{(59-"83.4")^2}{"83.4"} + \frac{(72-"84.8")^2}{"84.8"} + \frac{(48-"59.6")^2}{"59.6"} + \frac{(80-"55.6")^2}{"55.6"}$	**m1**	**PI**, six terms with at least three correct
	$= 24.8295121...$	**A1**	**AWRT** 24.8, $p-$value of 4.06×10^{-6}
	$\text{dof} = \nu = (3-1)(2-1) = 2$	**B1**	**PI**
	$\chi^2(0.99) = 9.210$ $24.82... > 9.210$, Therefore reject H_0	**B1**	Comparison may use p value, $0.01 > 4.06\times10^{-6}$
	Sufficient evidence to suggest there is an association between the location of a supporter inside or outside the host country and their predicted winning team	**A1ft**	Final statement must be consistent with their test statistic
		8	

	Question 2 Total	8	

Q	Answer	Marks	Comments

3(a)	$z = 1.96$ $1.96 \times \frac{\sqrt{6.25}}{\sqrt{n}} = \frac{1.4}{2}$ $n = 49$	**B1** **M1** **A1**	**PI** Set up correct equation with their z value
		3	

Q	Answer	Marks	Comments
3(b)(i)	$(5.6, 7.0)$	**B1**	Condone 7 instead of 7.0
		1	

Q	Answer	Marks	Comments
3(b)(ii)	5.4 is not in the confidence interval, therefore reject H_0	**B1ft**	**ft** with their interval
		1	

Q	Answer	Marks	Comments
3(c)	$P(4.7 < \overline{X} < 6.1 \mid \mu = 5.7)$ $\frac{\sqrt{6.25}}{\sqrt{49}}$ or $\frac{5}{14}$ or $0.357[14\ldots]$ $P(4.7 < \overline{X} < 6.1 \mid \mu = 5.7) = 0.866[08\ldots]$ $1 - 0.866...$ $= 0.134$	**M1** **B1** **A1** **A1**	**oe** **PI** Use of standard error **PI** by use of standardisation of z value or $z_1 = -2.8$ or $z_2 = 1.12$ or correct answer **CAO**
		4	

	Question 3 Total	**9**	

Q	Answer	Marks	Comments
4(a)(i)	$\mathrm{E}(\bar{X})\left[=\frac{1}{n}n\mu\right]=\mu$ $\mathrm{E}(\bar{Y})\left[=\frac{1}{m}m\mu\right]=\mu$	**M1**	Both expectations equal to μ seen or used
	$\mathrm{E}(T)=\frac{n}{n+m}\times\mu+\frac{m}{n+m}\times\mu=\mu$	**A1**	
	T is an unbiased estimator of μ	**E1**	
		3	

Q	Answer	Marks	Comments
4(a)(ii)	$\mathrm{Var}(\bar{X})=\frac{\sigma^2}{n}$ and $\mathrm{Var}(\bar{Y})=\frac{\sigma^2}{m}$	**M1**	Both variances seen or used
	$\mathrm{Var}(T)=\left(\frac{n}{n+m}\right)^2\times\frac{\sigma^2}{n}+\left(\frac{m}{n+m}\right)^2\times\frac{\sigma^2}{m}$		Must see this line of working
	$\mathrm{Var}(T)=\frac{\sigma^2}{n+m}$	**A1**	**AG**
		2	

Q	Answer	Marks	Comments
4(a)(iii)	$\mathrm{Var}(T)\to 0$ as $n\to\infty$ and $m\to\infty$	**M1**	**oe** $\mathrm{Var}(\bar{X})\to 0$ as $n\to\infty$ $\mathrm{Var}(\bar{Y})\to 0$ as $m\to\infty$
	Estimator is consistent for either sample	**A1**	**AG** Conclusion must be stated
		2	

Q	Answer	Marks	Comments
4(b)(i)	$\mathrm{E}\left(S_x^2\right)=\sigma^2, \quad \mathrm{E}\left(S_y^2\right)=\sigma^2$	**B1**	**B1** for either expectation seen or used
	$E(V)=\dfrac{n-1}{n+m-2}\,\sigma^2+k\sigma^2$		
	$\dfrac{n-1}{n+m-2}\,\sigma^2+k\sigma^2=\sigma^2$	**M1**	Must see intermediate working
	$k\sigma^2=\sigma^2-\dfrac{n-1}{n+m-2}\,\sigma^2$		
	$k=\dfrac{m-1}{n+m-2}$	**A1**	**AG** Be convinced
		3	

Q	Answer	Marks	Comments
4(b)(ii)	Pooled [estimate of population] variance	**B1**	**oe** combined variance, composite variance
		1	

Q	Answer	Marks	Comments
4(b)(iii)	Test for the difference of two means [when the variances are unknown and may be assumed equal]	**B1**	Allow t-test [with 2 samples]
		1	

	Question 4 Total	**12**	

Q	Answer	Marks	Comments
5	$H_0 : \mu_{2021} = \mu_{2020}$ $H_1 : \mu_{2021} > \mu_{2020}$	**B1**	Both hypotheses correct
	<table><tr><th>Line</th><th>Difference %</th></tr><tr><td>A</td><td>+2.7</td></tr><tr><td>B</td><td>+2.8</td></tr><tr><td>C</td><td>–0.2</td></tr><tr><td>D</td><td>+4.2</td></tr><tr><td>E</td><td>–2.1</td></tr><tr><td>F</td><td>–1.7</td></tr><tr><td>G</td><td>+1.4</td></tr><tr><td>H</td><td>+3.1</td></tr></table>	**B1**	Allow 1 mistake Allow negative of table values
	dof $= \nu = 7$	**B1**	**PI**
	$\left[\text{critical value, } t_7 =\right]$ 1.415	**B1**	
	$\bar{d} = 1.275$	**B1**	
	$s^2 = \frac{1}{8-1}\left(\sum d^2 - 8\bar{d}^2\right)$	**M1**	**PI** Attempt at variance formula; allow one slip. Implied by correct answer $\sum d^2 = 51.68$
	$= 5.525$	**A1**	Accept $s =$ **AWRT** 2.35
	$t = \frac{\bar{d}}{s/\sqrt{8}} = \frac{1.275}{\sqrt{\frac{5.525}{8}}}$	**M1**	Using their mean and variance
	$t = +1.534\left[22\ldots\right]$	**A1**	p-value 0.0844
	$1.534 > 1.415,$ therefore reject H_0	**A1ft**	Comparison may use p value $0.10 > 0.0844$
	Sufficient evidence to suggest the percentage of trains that arrive at stations on time has increased	**E1**	Correct conclusion based on their comparison
		11	

	Question 5 Total	**11**	

Q	Answer	Marks	Comments
6	$H_0 : \sigma_A^2 = \sigma_B^2$ $H_1 : \sigma_A^2 < \sigma_B^2$	B1	Both hypotheses needed Correct direction of inequality needed for H_1 Allow hypotheses in terms of σ instead of σ^2
	dof $v_A = 11, v_B = 11$	B1	PI
	$F_{11,11}$ at 99% $= 4.462$	B1	
	$s_A^2 = \frac{1}{12-1}\left(19498 - \frac{482^2}{12}\right)$ $[= 12.51\ldots]$ $s_B^2 = \frac{1}{12-1}\left(19531 - \frac{477^2}{12}\right)$ $[= 51.84\ldots]$	M1	oe $\frac{413}{33}$ oe $\frac{2281}{44}$
	$\frac{s_B^2}{s_A^2} = \frac{51.84\ldots}{12.51\ldots}$	M1	oe $\frac{6843}{1652}$ condone $\frac{s_A^2}{s_B^2} = \frac{12.51\ldots}{51.84\ldots}$
	$= 4.142$	A1	Correct value $p-$value of 8.3×10^{-3}
	$4.462 > 4.142$ Do not reject H_0	B1ft	$8.3\times10^{-3} < 0.01$ Allow 'accept H_0'
	Insufficient evidence to suggest less variance in the number of faulty chips in each box for Company A	E1	oe allow 'manufacturing process is better' or support for the company's claim is not justified
		8	

	Question 6 Total	8	

Q	Answer	Marks	Comments
7(a)	$\mathrm{M}_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right) = (1-p)\mathrm{e}^0 + p\mathrm{e}^t$	**M1**	
	$= (1-p) + p\mathrm{e}^t$	**A1**	If **M0** awarded, **SC1** for $(1-p) + p\mathrm{e}^t$
		2	

Q	Answer	Marks	Comments
7(b)	$M_S(t) = \left((1-p) + p\mathrm{e}^t\right)^n$	**B1**	
		1	

Q	Answer	Marks	Comments
7(c)(i)	$M_S'(t) = np\mathrm{e}^t\left((1-p) + p\mathrm{e}^t\right)^{n-1}$	**M1**	Allow one slip for differentiating the correct answer from **part (b)** If **part (b)** scored 0, then **M1** for correct first derivative from their **part (b)**
	$M_S'(0) = np\mathrm{e}^0\left((1-p) + p\mathrm{e}^0\right)^{n-1}$	**m1**	Their first derivative evaluated at $t = 0$
	$\mathrm{E}(S) = M_S'(0) = np$	**A1**	**CSO**
		3	

Q	Answer	Marks	Comments
7(c)(ii)	$M_S''(t) = n(n-1)p^2\mathrm{e}^{2t}\left((1-p) + p\mathrm{e}^t\right)^{n-2} + np\mathrm{e}^t\left((1-p) + p\mathrm{e}^t\right)^{n-1}$	**M1**	**M1** Correct second derivative, allow one slip
	$M_S''(0) = np(1-p) + n^2p^2$	**A1**	Correct second derivative evaluated at $t = 0$
	$\mathrm{Var}(S) = M_S''(0) - \left(M_S'(0)\right)^2$	**M1**	**M1** for use of variance formula
	$= np(1-p)$	**A1**	**ACF, CSO**
		4	

	Question 7 Total	**10**	

Q	Answer	Marks	Comments
8(a)(i)	$\overline{X} = 2.1$ $p = \frac{2.1}{6} = 0.35$	**B1**	**AG**
		1	

Q	Answer	Marks	Comments
8(a)(ii)	94.20 0.74	**B1** **B1**	
		2	

Q	Answer	Marks	Comments
8(a)(iii)	H_0: Binomial model is a suitable distribution H_1: Binomial model is not a suitable distribution	**B1**	
	dof $= \nu = 4$	**B1**	**PI**
	$\chi^2(0.975) = 11.143$	**B1**	
	$\sum \frac{(O-E)^2}{E} = \frac{(40-30.17)^2}{30.17} + \frac{(103-97.46)^2}{97.46} + \frac{(113-131.20)^2}{131.20} + \frac{(83-\text{“}94.20\text{”})^2}{\text{“}94.20\text{”}} + \frac{(45-38.04)^2}{38.04} + \frac{(16-\text{"}8.93\text{"})^2}{\text{"}8.93\text{"}}$ $= 14.25$	**M1** **m1** **A1**	Attempt to calculate statistic Combines last two categories **AWFW** [14.2, 14.3] p-value is 0.0141
	$14.25 > 11.143$, therefore reject H_0		Comparison may use p value, $0.025 > 0.0141$
	Insufficient evidence to suggest binomial model is a suitable distribution	**E1**	Both conclusion and comparison required
		7	

Q	Answer	Marks	Comments
8(b)(i)	$\left[400\times e^{-2.1}\times\frac{2.1^0}{0!}=\right]$ 48.98	B1	
	$\left[400\times e^{-2.1}\times\frac{2.1^5}{5!}=\right]$ 16.67	B1	
		2	

Q	Answer	Marks	Comments
8(b)(ii)	dof $=\nu=7-1-1=5$	B1	**PI** Final 2 categories are not combined
	$\chi^2(0.975)=12.833$	B1	
	$8.41<12.833$, do not reject H_0		Allow 'accept H_0'
	or	E1	Conclusion must be correct
	sufficient evidence to suggest Poisson model is a suitable distribution		
		3	

Q	Answer	Marks	Comments
8(c)	The Poisson model is better		
	and		
	The Poisson model was not rejected [whereas the binomial model was rejected]	B1ft	Valid reason must be given. Allow explanation of lower χ^2 test statistic value for their choice of model
		1	

	Question 8 Total	16	