Friday 6 June 2014 – Afternoon

AS GCE MEI STATISTICS

G242/01 Statistics 2 (Z2)

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book G242/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

1 A dietician is investigating a claim that the dietary supplement 'red-yeast rice' can reduce cholesterol levels by 15 units on average. The dietician believes that the reduction will be greater than this. A group of volunteers is given the supplement for a period of eight weeks. The reduction in cholesterol levels of a sample of 10 of these volunteers is measured. The results are as follows.

16 24 27 22 11 13 26 21 18 20

(i) Stating any necessary assumptions, use a Wilcoxon test to examine, at the 5% significance level, whether these data support the dietician's belief. **[13]**

(ii) What further assumption is necessary for a test based on the t distribution to be appropriate? **[1]**

2 A medical researcher is investigating alternative methods for determining glucose levels in diabetes patients' blood. The most common method involves taking a small sample of blood; this is accurate but may be stressful for some patients. One alternative method involves measuring the glucose level in tear fluid. To assess the accuracy of this alternative method, the researcher records the difference in values given by the two methods for each of a random sample of 15 patients. The results, in suitable units, are as follows.

0.3 0.5 –0.2 0.8 0.3 –0.1 0.1 –0.2 0.4 –0.1 0.5 0.3 0.4 0.1 –0.7

(i) Given that the sample standard deviation is 0.374, use the t distribution to test, at the 5% significance level, the null hypothesis $H_0: \mu = 0$ against the alternative hypothesis $H_1: \mu \neq 0$, where μ represents the mean of the underlying population of differences. **[8]**

Another alternative method for determining blood glucose levels involves a test using ultrasound. The researcher records the difference between the value given by an ultrasound test and the value given by a blood test for a random sample of 12 patients.

(ii) Given that the sample mean is 0.19 and the sample standard deviation is 0.281, calculate a 95% confidence interval based on the t distribution for the mean difference in level given by these two methods. **[5]**

(iii) Given that underlying Normality holds, why is the t distribution required in parts **(i)** and **(ii)**? **[1]**

3 A sports equipment manufacturer produces 'yellow dot' and 'red dot' squash balls. The weights of each type of ball may be assumed to be Normally distributed. The manufacturer carries out regular checks of the production process. The weights, in grams, for a random sample of 10 yellow dot balls are as follows.

23.7 24.2 24.6 24.1 24.3 24.4 23.9 23.8 24.0 24.2

(i) Calculate the sample mean. **[1]**

The weight of a yellow dot ball should be 24 grams.

(ii) Given that the population standard deviation is 0.3 grams, use a test based on the Normal distribution, at the 5% significance level, to examine whether this sample provides evidence that the mean weight of yellow dot balls is not 24 grams. **[9]**

The weights of red dot balls are Normally distributed with mean 24.7 grams and standard deviation 0.4 grams. Red dot balls are sold in packs of 6. Balls in a pack may be regarded as a random sample.

(iii) Find the probability that the total weight of the 6 balls in a pack exceeds 150 grams. **[4]**

4 A seismologist is monitoring the global occurrence of earthquakes. Each week, she records the number of 'moderate strength' earthquakes that occur around the world. The following table summarises the results for a random sample of 200 weeks.

Number of earthquakes	0	1	2	3	4	5	6	$\geqslant 7$
Observed frequency	7	47	49	46	25	19	7	0

(i) The sample mean is 2.6. Calculate the sample standard deviation. Hence comment briefly on whether or not the Poisson distribution may provide a suitable model for these data. **[3]**

The seismologist decides to test the goodness of fit of a Poisson model. She uses the sample mean as an estimate for the mean of the underlying population to produce the following expected frequencies.

Number of earthquakes	0	1	2	3	4	5	6	$\geqslant 7$
Expected frequency	14.86	38.62	50.21	43.52	28.28	14.71	6.37	3.43

(ii) Show how the expected frequency of 28.28 for the number of weeks in which 4 earthquakes occur is calculated. **[3]**

(iii) Carry out the test of the goodness of fit of the Poisson model at the 5% level of significance. **[10]**

Question 5 begins on page 4

5 A car manufacturer decides to investigate whether there is an association between the age of customers and their 'brand loyalty' (ie whether they are returning customers or first time customers). A random sample of 150 customers is selected and classified as follows.

		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	25	12
	35 to 50	38	35
	Over 50	31	9

The following tables show some of the expected frequencies and contributions to the test statistic for use in part **(i)**.

Expected frequencies		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	23.187	13.813
	35 to 50	45.747	
	Over 50		14.933

Contributions to the test statistic		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	0.142	0.238
	35 to 50	1.312	
	Over 50		2.357

(i) Calculate the remaining expected frequencies and contributions. Carry out the test using a 5% level of significance. **[11]**

(ii) With reference to the contributions to the test statistic, comment briefly on brand loyalty in the different age groups. **[3]**

END OF QUESTION PAPER

Friday 6 June 2014 – Afternoon

AS GCE MEI STATISTICS

G242/01 Statistics 2 (Z2)

PRINTED ANSWER BOOK

Candidates answer on this Printed Answer Book.

OCR supplied materials:
- Question Paper G242/01 (inserted)
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

Candidate forename		Candidate surname	

Centre number						Candidate number				

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book**. Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

1 (i)

(answer space continued on next page)

1 (i)	(continued)
1 (ii)	

2 (i)	
2 (ii)	
	(answer space continued on next page)

2 (ii)	(continued)
2 (iii)	

PLEASE DO NOT WRITE IN THIS SPACE

3 (i)	
3 (ii)	
	(answer space continued on next page)

3 (ii)	(continued)
3 (iii)	

4 (i)	
4 (ii)	

4 (iii)

5 (i)

Expected frequencies		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	23.187	13.813
	35 to 50	45.747	
	Over 50		14.933

Contributions to the test statistic		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	0.142	0.238
	35 to 50	1.312	
	Over 50		2.357

(answer space continued on next page)

5 (i)	(continued)

5 (ii)	

Question			Answer	Marks	Guidance
1	(i)		Assume the underlying population (of reductions) is distributed symmetrically and that the sample is random. H_0: population median = 15 H_1: population median > 15 Actual differences 1 9 12 7 –4 –2 11 6 3 5 Associated ranks 1 8 10 7 4 2 9 6 3 5 $T^- = 4 + 2 = 6$ $\therefore T = 6$ From n = 10 tables – at the 5% level of significance in a one-tailed Wilcoxon single sample test, the critical value of T is 10 $6 < 10 \therefore$ the result is significant The evidence supports the dietician's belief	B1 B1 B1 B1 B1 M1* A1 B1 B1 M1 A1 M1dep* A1 **[13]**	Do not accept "not necessarily Normal" Condone "the distribution is symmetrical" Condone "the data is random" If population not stated then SC1 for H_0: median = 15, H_1:median >15 For ranking differences $T^+ = 49$ Use of n = 10 (may be implied by use of cv = 10) No further A marks from here if incorrect For sensible comparison leading to a conclusion For non-assertive conclusion in context.
1	(ii)		Normality of underlying population (of reductions)	B1 **[1]**	
2	(i)		Sample mean = 2.4 ÷ 15 = 0.16 $t = \frac{0.16 - 0}{0.374/\sqrt{15}} = 1.66$ (3.s.f.) 14 degrees of freedom At 5% level, critical value of t is 2.145 1.66 < 2.145 so the result is not significant Insufficient evidence to suggest that the mean difference is not zero	B1 M1 A1 B1 B1 M1 A1 A1 **[8]**	 CAO For ± 2.145. No further A marks from here if incorrect. For not significant o.e. For non-assertive conclusion in context. Allow "not enough evidence to suggest the tests are not equally accurate"

Question			Answer	Marks	Guidance
2	(ii)		$0.19 \pm 2.201 \times \frac{0.281}{\sqrt{12}}$ (0.0115, 0.369)	M1 B1 M1 A1 A1 **[5]**	Centred on 0.19 2.201 Structure using their "2.201" Alow (0.011, 0.369)
2	(iii)		Population variance unknown	B1 **[1]**	Or sample too small to justify use of Z
3	(i)		$241.2 \div 10 = 24.12$	B1 **[1]**	
3	(ii)		$H_0: \mu = 24$ & $H_1: \mu \neq 24$ Where μ represents the (population) mean weight $z = \frac{24.12 - 24}{0.3/\sqrt{10}} = 1.265$ At 5% level, critical value of z is 1.96 1.265 < 1.96 so the result is not significant. Insufficient evidence to suggest that the mean weight is not 24 grams.	B1 B1 B1 M1 A1 B1 M1 A1 A1 **[9]**	For hypotheses in terms of μ. If other symbols are used, population must be stated For defining μ. If numerator reversed then M1A0B1M1A0A0 max For ± 1.96. No further A marks from here if incorrect For sensible comparison based on the Normal distribution. For non-assertive conclusion in context
3	(iii)		mean = 24.7 + 24.7 + 24.7 + 24.7 + 24.7 + 24.7 variance = $0.4^2 + 0.4^2 + 0.4^2 + 0.4^2 + 0.4^2 + 0.4^2$ ie Total, $T \sim N(148.2, 0.96)$ $P(T > 150) = P(Z > 1.837\ldots)$ $= 1 - \Phi(1.837\ldots) = 1 - 0.9669 = 0.0331$	B1 B1 M1 A1 **[4]**	B1 for mean = 148.2 B1 for variance = 0.96 For structure of probability calculation using their mean and variance.

Question			Answer	Marks	Guidance
4	(i)		$\sqrt{\dfrac{1784 - 200 \times 2.6^2}{199}} = 1.473$ Variance = $1.473^2 = 2.169\ldots$	M1A1	For sample standard deviation.
			The mean is approximately equal to the variance so this does not discredit the use of a Poisson model.	E1 **[3]**	Allow suitable alternatives. FT their variance calculated from given data provided its value is stated.
4	(ii)		Using $X \sim$ Poisson(2.6)	M1*	Attempt at Poisson calculation
			$P(X = 4) = e^{-2.6} \times 2.6^4 \div 4! = 0.14142\ldots$	M1dep*	For $P(X = 4) \times 200$
			0.14142... × 200 = 28.284.. (= 28.28 A.G.)	A1 **[3]**	
4	(iii)		H_0: The Poisson model is suitable H_1: The Poisson model is not suitable	B1	No need to specify parameter
			Cells merged	B1	See table below
				M1	Attempt at $(f_o - f_e)^2 \div f_e$
				A1	Contributions (approx): 4.1574, 1.8183, 0.0292, 0.1413, 0.3804, 1.2511, 0.8 Allow A1 if one correct contribution seen.
			Test statistic = 8.578	A1	For correct total
			Degrees of freedom = 7 – 1 – 1 = 5	B1	
			Critical value = 11.07	B1	
			8.578 < 11.07	M1	FT their chi-squared CV and test statistic
			Result is not significant	A1	
			Insufficient evidence to suggest that the Poisson model is not a good fit to these data.	A1 **[10]**	

	0	1	2	3	4	5	**≥ 6**
Observed	7	47	49	46	25	19	**7**
Expected	14.86	38.62	50.21	43.52	28.28	14.71	**9.8**

<table>
<tr><th colspan="3">Question</th><th>Answer</th><th>Marks</th><th>Guidance</th></tr>
<tr><td>5</td><td>(i)</td><td></td><td>H₀: No association between age and brand loyalty
H₁: There is an association between age and brand loyalty</td><td>B1</td><td></td></tr>
<tr><td></td><td></td><td></td><td>

Expected frequencies		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	23.187	13.813
	35 to 50	45.747	**27.253**
	Over 50	**25.067**	14.933

</td><td>M1
A1</td><td>Any suitable method. May be implied by a correct expected frequency.
Both correct. Allow 25.066 and 27.254.</td></tr>
<tr><td></td><td></td><td></td><td>

Expected frequencies		Brand loyalty	
		Returning customer	First time customer
Age (in years)	Under 35	0.142	0.238
	35 to 50	1.312	**2.202**
	Over 50	**1.404**	2.357

</td><td>M1
A1</td><td>Allow 1.405 and 2.2015</td></tr>
<tr><td></td><td></td><td></td><td>$X^2 = 7.655$</td><td>A1</td><td>For TS rounding to 7.66</td></tr>
<tr><td></td><td></td><td></td><td>2 degrees of freedom</td><td>B1</td><td></td></tr>
<tr><td></td><td></td><td></td><td>Critical value for 5% significance level is 5.991</td><td>B1</td><td>No further A marks from here if incorrect.</td></tr>
<tr><td></td><td></td><td></td><td>As 7.655 > 5.991 the result is significant</td><td>M1 A1</td><td>FT their TS</td></tr>
<tr><td></td><td></td><td></td><td>There is evidence to suggest an association between age and brand loyalty.</td><td>A1

[11]</td><td></td></tr>
</table>

Question			Answer	Marks	Guidance
5	(ii)		The low contributions (of 0.142 and 0.238) for the Under 35 age group shows that the observed frequencies were as expected if there was no association, whereas	B1	For coherent comment with reference to contribution(s) for the Under 35 age group (reference may be implicit).
			the larger contributions for the other two age groups indicate that there were fewer returning customers than expected in the 35 to 50 age group…	B1	For coherent comment with reference to contribution(s) for either of the other two age groups (reference may be implicit).
			…and more returning customers than expected in the Over 50 age group	B1	Allow sensible alternatives. E.g. reference linking level of association between age and brand loyalty to size of contribution for individual cells.
				[3]	

For a description of how UMS marks are calculated see:
www.ocr.org.uk/learners/ums_results.html

Unit level raw mark and UMS grade boundaries June 2014 series

AS GCE / Advanced GCE / AS GCE Double Award / Advanced GCE Double Award

GCE Mathematics (MEI)

		Max Mark	a	b	c	d	e	u
4751/01 (C1) MEI Introduction to Advanced Mathematics	Raw	72	61	56	51	46	42	0
	UMS	100	80	70	60	50	40	0
4752/01 (C2) MEI Concepts for Advanced Mathematics	Raw	72	57	51	45	39	33	0
	UMS	100	80	70	60	50	40	0
4753/01 (C3) MEI Methods for Advanced Mathematics with Coursework: Written Paper	Raw	72	58	52	47	42	36	0
4753/02 (C3) MEI Methods for Advanced Mathematics with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4753/82 (C3) MEI Methods for Advanced Mathematics with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4753 (C3) MEI Methods for Advanced Mathematics with Coursework	UMS	100	80	70	60	50	40	0
4754/01 (C4) MEI Applications of Advanced Mathematics	Raw	90	68	61	54	47	41	0
	UMS	100	80	70	60	50	40	0
4755/01 (FP1) MEI Further Concepts for Advanced Mathematics	Raw	72	63	57	51	45	40	0
	UMS	100	80	70	60	50	40	0
4756/01 (FP2) MEI Further Methods for Advanced Mathematics	Raw	72	60	54	48	42	36	0
	UMS	100	80	70	60	50	40	0
4757/01 (FP3) MEI Further Applications of Advanced Mathematics	Raw	72	57	51	45	39	34	0
	UMS	100	80	70	60	50	40	0
4758/01 (DE) MEI Differential Equations with Coursework: Written Paper	Raw	72	63	56	50	44	37	0
4758/02 (DE) MEI Differential Equations with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4758/82 (DE) MEI Differential Equations with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4758 (DE) MEI Differential Equations with Coursework	UMS	100	80	70	60	50	40	0
4761/01 (M1) MEI Mechanics 1	Raw	72	57	49	41	34	27	0
	UMS	100	80	70	60	50	40	0
4762/01 (M2) MEI Mechanics 2	Raw	72	57	49	41	34	27	0
	UMS	100	80	70	60	50	40	0
4763/01 (M3) MEI Mechanics 3	Raw	72	55	48	42	36	30	0
	UMS	100	80	70	60	50	40	0
4764/01 (M4) MEI Mechanics 4	Raw	72	48	41	34	28	22	0
	UMS	100	80	70	60	50	40	0
4766/01 (S1) MEI Statistics 1	Raw	72	61	53	46	39	32	0
	UMS	100	80	70	60	50	40	0
4767/01 (S2) MEI Statistics 2	Raw	72	60	53	46	40	34	0
	UMS	100	80	70	60	50	40	0
4768/01 (S3) MEI Statistics 3	Raw	72	61	54	47	41	35	0
	UMS	100	80	70	60	50	40	0
4769/01 (S4) MEI Statistics 4	Raw	72	56	49	42	35	28	0
	UMS	100	80	70	60	50	40	0
4771/01 (D1) MEI Decision Mathematics 1	Raw	72	51	46	41	36	31	0
	UMS	100	80	70	60	50	40	0
4772/01 (D2) MEI Decision Mathematics 2	Raw	72	46	41	36	31	26	0
	UMS	100	80	70	60	50	40	0
4773/01 (DC) MEI Decision Mathematics Computation	Raw	72	46	40	34	29	24	0
	UMS	100	80	70	60	50	40	0
4776/01 (NM) MEI Numerical Methods with Coursework: Written Paper	Raw	72	54	48	43	38	32	0
4776/02 (NM) MEI Numerical Methods with Coursework: Coursework	Raw	18	14	12	10	8	7	0
4776/82 (NM) MEI Numerical Methods with Coursework: Carried Forward Coursework Mark	Raw	18	14	12	10	8	7	0
4776 (NM) MEI Numerical Methods with Coursework	UMS	100	80	70	60	50	40	0
4777/01 (NC) MEI Numerical Computation	Raw	72	55	47	39	32	25	0
	UMS	100	80	70	60	50	40	0
4798/01 (FPT) Further Pure Mathematics with Technology	Raw	72	57	49	41	33	26	0
	UMS	100	80	70	60	50	40	0

GCE Statistics (MEI)

		Max Mark	a	b	c	d	e	u
G241/01 (Z1) Statistics 1	Raw	72	61	53	46	39	32	0
	UMS	100	80	70	60	50	40	0
G242/01 (Z2) Statistics 2	Raw	72	55	48	41	34	27	0
	UMS	100	80	70	60	50	40	0
G243/01 (Z3) Statistics 3	Raw	72	56	48	41	34	27	0
	UMS	100	80	70	60	50	40	0