OCR
Oxford Cambridge and RSA

Practice Paper – Set 3

A Level Further Mathematics B (MEI)

Y422/01 Statistics Major

MARK SCHEME

Duration: 2 hours 15 minutes

MAXIMUM MARK 120

This document consists of 16 pages

Text Instructions

1. Annotations and abbreviations

Annotation in scoris	Meaning
✓and ✗	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0, 1
SC	Special case
^	Omission sign
MR	Misread
Highlighting	
Other abbreviations in mark scheme	**Meaning**
E1	Mark for explaining a result or establishing a given result
dep*	Mark dependent on a previous mark, indicated by *
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working
AG	Answer given
awrt	Anything which rounds to
BC	By Calculator
DR	This indicates that the instruction **In this question you must show detailed reasoning** appears in the question.

2. Subject-specific Marking Instructions for A Level Further Mathematics B (MEI)

a Annotations should be used whenever appropriate during your marking. The A, M and B annotations must be used on your standardisation scripts for responses that are not awarded either 0 or full marks. It is vital that you annotate standardisation scripts fully to show how the marks have been awarded. For subsequent marking you must make it clear how you have arrived at the mark you have awarded.

b An element of professional judgement is required in the marking of any written paper. Remember that the mark scheme is designed to assist in marking incorrect solutions. Correct solutions leading to correct answers are awarded full marks but work must not be judged on the answer alone, and answers that are given in the question, especially, must be validly obtained; key steps in the working must always be looked at and anything unfamiliar must be investigated thoroughly. Correct but unfamiliar or unexpected methods are often signalled by a correct result following an apparently incorrect method. Such work must be carefully assessed. When a candidate adopts a method which does not correspond to the mark scheme, escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.
If you are in any doubt whatsoever you should contact your Team Leader.

c The following types of marks are available.

M
A suitable method has been selected and *applied* in a manner which shows that the method is essentially understood. Method marks are not usually lost for numerical errors, algebraic slips or errors in units. However, it is not usually sufficient for a candidate just to indicate an intention of using some method or just to quote a formula; the formula or idea must be applied to the specific problem in hand, e.g. by substituting the relevant quantities into the formula. In some cases the nature of the errors allowed for the award of an M mark may be specified.

A
Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. Accuracy marks cannot be given unless the associated Method mark is earned (or implied). Therefore M0 A1 cannot ever be awarded.

B
Mark for a correct result or statement independent of Method marks.

E
A given result is to be established or a result has to be explained. This usually requires more working or explanation than the establishment of an unknown result.

Unless otherwise indicated, marks once gained cannot subsequently be lost, e.g. wrong working following a correct form of answer is ignored. Sometimes this is reinforced in the mark scheme by the abbreviation isw. However, this would not apply to a case where a candidate passes through the correct answer as part of a wrong argument.

d When a part of a question has two or more 'method' steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. (The notation 'dep*' is used to indicate that a particular mark is dependent on an earlier, asterisked, mark in the scheme.) Of course, in practice it may happen that when a candidate has once gone wrong in a part of a question, the work from there on is worthless so that no more marks can sensibly be given. On the other hand, when two or more steps are successfully run together by the candidate, the earlier marks are implied and full credit must be given.

e The abbreviation FT implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only – differences in notation are of course permitted. A (accuracy) marks are not given for answers obtained from incorrect working. When A or B marks are awarded for work at an intermediate stage of a solution, there may be various alternatives that are equally acceptable. In such cases, what is acceptable will be detailed in the mark scheme. If this is not the case, please escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.
Sometimes the answer to one part of a question is used in a later part of the same question. In this case, A marks will often be 'follow through'. In such cases you must ensure that you refer back to the answer of the previous part question even if this is not shown within the image zone. You may find it easier to mark follow through questions candidate-by-candidate rather than question-by-question.

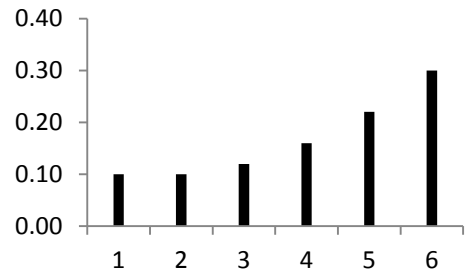
f Unless units are specifically requested, there is no penalty for wrong or missing units as long as the answer is numerically correct and expressed either in SI or in the units of the question. (e.g. lengths will be assumed to be in metres unless in a particular question all the lengths are in km, when this would be assumed to be the unspecified unit.) We are usually quite flexible about the accuracy to which the final answer is expressed; over-specification is usually only penalised where the scheme explicitly says so. When a value is given in the paper only accept an answer correct to at least as many significant figures as the given value. This rule should be applied to each case. When a value is not given in the paper accept any answer that agrees with the correct value to 2 s.f. Follow through should be used so that only one mark is lost for each distinct accuracy error, except for errors due to premature approximation which should be penalised only once in the examination. There is no penalty for using a wrong value for *g*. E marks will be lost except when results agree to the accuracy required in the question.

g Rules for replaced work: if a candidate attempts a question more than once, and indicates which attempt he/she wishes to be marked, then examiners should do as the candidate requests; if there are two or more attempts at a question which have not been crossed out, examiners should mark what appears to be the last (complete) attempt and ignore the others. NB Follow these maths-specific instructions rather than those in the assessor handbook.

h For a genuine misreading (of numbers or symbols) which is such that the object and the difficulty of the question remain unaltered, mark according to the scheme but following through from the candidate's data. A penalty is then applied; 1 mark is generally appropriate, though this may differ for some units. This is achieved by withholding one A mark in the question. Marks designated as cao may be awarded as long as there are no other errors. E marks are lost unless, by chance, the given results are established by equivalent working. 'Fresh starts' will not affect an earlier decision about a misread. Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

i If a calculator is used, some answers may be obtained with little or no working visible. Allow full marks for correct answers (provided, of course, that there is nothing in the wording of the question specifying that analytical methods are required). Where an answer is wrong but there is some evidence of method, allow appropriate method marks. Wrong answers with no supporting method score zero. If in doubt, consult your Team Leader.

j If in any case the scheme operates with considerable unfairness consult your Team Leader.

Question			Answer	Marks	AOs	Guidance	
1	(a)		$5k + 5k + 6k + 8k + 11k + 15k = 1$ $50k = 1$ so $k = 0.02$	**M1** **A1** **[2]**	**3.1a** **1.1**	Equating sum of probabilities to 1 **AG**	
1	(b)			**B1** **B1** **[2]**	**1.1** **1.1**	For heights For axes and labels	
1	(c)		Negative skew	**E1** **[1]**	**1.1**		
1	(d)		E(X) = 4.2 Var(X) = 2.8	**B1** **B1** **[2]**	**1.1a** **1.1**	**BC** **BC**	

Question			Answer	Marks	AOs	Guidance	
2	(a)		Positive tests occur randomly and occurrences are independent with constant probability 0.0005	E1	2.4	For partial explanation of binomial	
			The number of positive tests out of 1000 tests is being counted, so a binomial distribution is appropriate	E1	2.4	For full explanation of binomial	
			Because n is large and p is small the Poisson approximation to the binomial distribution is also appropriate	E1	2.4	For explanation of Poisson	
				[3]			
2	(b)		B(1000, 0.0005) or Po(0.5) used	M1	3.3	May be implied by a correct answer	Both distributions give the same answers correct to 4 dp
			$P(X = 2) = 0.0758$	A1	1.1	BC	
			$P(X > 2) = 0.0144$	A1	1.1	BC	
				[3]			
2	(c)		$P(>1500) = 0.9995^{1500}$	M1	3.3		
			$= 0.472$	A1	1.1		(0.472 277 96…)
				[2]			
3	(a)		The sample size is large so by the central limit theorem	E1	2.4	For both ‘large sample’ oe and CLT	
			the sample mean is approximately Normally distributed	E1	2.4	For full explanation	
				[2]			
3	(b)		$143.4593 < \mu < 147.1407$	B1	1.1	Accept rounded to 2 or 3dp	
				[1]			
3	(c)		z-value = 1.96	B1	3.4		
			z-value × SE = 1.96 × 0.93915	B1	1.1		
				[2]			
3	(d)		If the parent population is Normally distributed				

Question			Answer	Marks	AOs	Guidance	
			then the *t* distribution can be used to give a confidence interval	**E1**	**2.2a**	For mention of *t* distribution and Normality requirement	
			If the parent distribution is not known then a confidence interval cannot be calculated	**E1**	**3.2b**	For full explanation	
				[2]			

Question			Answer	Marks	AOs	Guidance	
4	(a)		Mean length of 10 stones : $N\left(598, \frac{3.4^2}{10}\right)$	M1	3.3	For Normal and correct mean	
			i.e. N(598, 1.156)	A1	1.1	For correct variance	
			P(Mean length > 600) = 0.0314	B1	3.4	**BC**	
				[3]			
4	(b)		Total length involves 5 stones and 4 gaps	M1	3.1b	For realising there are only 4 gaps	
			mean = $5 \times 598 + 4 \times 16$	M1	3.3	Method for mean	
			variance = $5 \times 3.4^2 + 4 \times 1.6^2$	M1	1.1	Method for variance	
			so distribution is N(3054, 68.04)	A1	1.1	Correct distribution	
			P(Total length < 3060) = 0.7665	A1	3.4	**BC**	
				[5]			

Question			Answer	Marks	AOs	Guidance	
5	(a)		e.g. It would be very costly to enrol all of the patients on a course e.g. If it was not effective it would have unnecessarily wasted their time	E1 [1]	2.2b	Allow other sensible correct responses	
5	(b)		The parent population of differences must be Normally distributed	E1 E1 [2]	1.2 1.2	For mention of Normal distribution For full answer including 'differences' and 'population'	
5	(c)		**DR** 3.6 ± 2.262 $\times \frac{3.4383}{\sqrt{10}}$ 3.6 ± 2.46 or (1.14, 6.06)	 B1 M1 M1 A1 [4]	 1.1 3.3 1.1 3.4	 For use of t_9 critical value 	
5	(d)		It does because the confidence interval does not contain zero	E1 [1]	3.5a		
5	(e)		Sample mean = 89.7(5) $2.35 = 2.262 \times \frac{SD}{\sqrt{10}}$ Sample SD = 3.28(5)	B1 M1 A1 [3]	1.1 3.1b 1.1		
5	(f)		In repeated sampling, 95% of confidence intervals constructed in this way will contain the true population mean	E1 E1	2.2b 2.2b	For idea of 95% proportion in repeated sampling For complete answer	

Question			Answer	Marks	AOs	Guidance	
				[2]			

Question			Answer	Marks	AOs	Guidance	
6	(a)		The data is not random on random … … since specific speeds were chosen	E1 E1 [2]	3.5a 2.4	For stating 'not random on random' For reference to speeds being chosen	
6	(b)		Speed is the independent variable and the independent variable is (normally) plotted on the horizontal axis	E1 [1]	2.4	For referring to the usual plotting convention	
6	(c)		At 56 mph estimated consumption = 52.1 mpg At 90 mph estimated consumption = 39.3 mpg	B1 B1 [2]	1.1 1.1		
6	(d)		Because the points lie very close to the line, and it is interpolation, the first prediction is very reliable The second prediction is less reliable because it involves extrapolation	E1 E1 [2]	2.2a 3.2b		
6	(e)		Residual = 50.9 – (–0.378 × 60 + 73.3) = 0.28	M1 A1 [2]	1.1 1.1	Allow this mark if subtraction reversed **cao**	

Question			Answer	Marks	AOs	Guidance	
7	(a)		The scatter diagram could be said to be very roughly elliptical so the distribution may be bivariate Normal	**E1** **E1** **[2]**	**3.5a** **2.4**	For 'roughly elliptical' shape identified For distribution being bivariate Normal	
7	(b)		$S_{lw} = 90942 - \frac{1}{18} \times 1713 \times 948$ (= 724) $S_{ll} = 165817 - \frac{1}{18} \times 1713^2$ (= 2796.5) $S_{ww} = 50408 - \frac{1}{18} \times 948^2$ (= 480) $r = \frac{S_{lw}}{\sqrt{S_{ll} S_{ww}}} = \frac{724}{\sqrt{2796.5 \times 480}}$ = 0.6249	**M1** **M1** **M1** **A1** **[4]**	**1.1a** **1.1** **3.3** **1.1**	For either S_{ll} or S_{ww} For general form including square root	The three **M1** marks in this part can be implied by a correct numerical answer; numerical evaluations of the expressions are not required for the award of these marks
7	(c)		H_0: $\rho = 0$, H_1: $\rho > 0$ where ρ is the population pmcc between *l* and *w* For n = 18, the 5% critical value is 0.4000 Since 0.6249 > 0.4000 the result is significant, so there is sufficient evidence to reject H_0 There is sufficient evidence at the 5% level to suggest that there is positive correlation between length and width of bay leaves	**B1** **B1** **B1** **M1** **A1** **[5]**	**3.3** **2.5** **3.4** **1.1** **2.2b**	For both hypotheses For defining ρ For correct critical value For comparison and conclusion **FT** for conclusion in words	

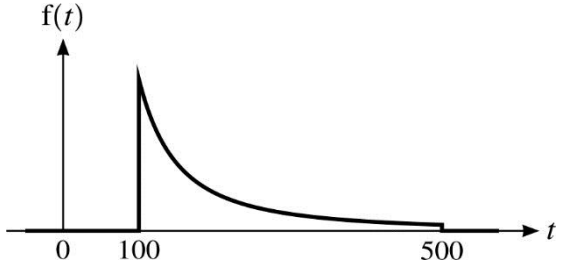
Question			Answer	Marks	AOs	Guidance	
8	(a)		Sample mean = 2.5	B1	1.1	BC	
			Sample variance = 2.555	B1	1.1	BC	
			The variance is close to the mean so a Poisson distribution could be a suitable model	E1	2.2b		
				[3]			
8	(b)		Cell C5 = 0.2138	B1	3.4	BC	
			Cell D5 = 25.6516	B1	2.2a	FT their C5 × 120	
			Cell E5 = $\frac{(28-25.6516)^2}{25.6516}$	M1	1.1a	soi	
			= 0.2150	A1	1.1	FT correct use of their D5	
				[4]			
8	(c)		Because otherwise some expected frequencies would be less than 5, so too small for the test to be valid	E1	3.5b	For 'less than 5 so invalid'	Must refer to expected, not observed, frequency
				[1]			
8	(d)		H_0: Poisson model is a good fit H_1: Poisson model is not a good fit	B1	2.5	For both hypotheses	
			X^2 = 10.86	B1	1.1		
			Refer to χ^2_5	B1	3.4	For degrees of freedom = 5 soi	
			Critical value at 5% level = 11.07	B1	1.1		
			10.86 < 11.07 so result is not significant	M1	1.1	For comparison with critical value	
			There is insufficient evidence to suggest that the Poisson model is not a good fit.	A1	2.2b	For conclusion in context	
				[6]			

Question			Answer	Marks	AOs	Guidance	
9	(a)		E(T) = 11 + 6 + 42 + 8 = 67	B1	1.1		
			Var of uniform on [2, 10] is $\frac{1}{12}(10-2)^2 = \frac{16}{3}$ oe	B1	1.2		
			$\text{Var}(T) = 1.2^2 + \frac{16}{3} + 3^2$	M1	1.1a		
			= 15.77	A1	1.1		(15.773 333…)
				[4]			
9	(b)		$P(T < 60) = \frac{30}{1000} = 0.030$	B1	1.1		
			$P(T \geq 68) = 1 - \frac{576}{1000}$	M1	1.1a		
			= 0.424	A1	1.1		
				[3]			
9	(c)		P(W < 60) = 0.039	B1	3.3	**BC FT** their mean and variance	
			P($W \geq$ 68) = 0.401	B1	1.1	**BC FT** their mean and variance	
				[2]			
9	(d)		The probabilities would not be expected to agree because the distribution of T is not Normal	E1	3.2a	For statement that T is not Normal	
			Adding independent Normal distributions for cycling time and train time (and a constant walking time) does give a Normal distribution, but adding the uniformly distributed waiting time makes the total not Normal	E1	3.2b	For identifying the uniform distribution as the reason for non-Normality	Accept More simulations will bring the simulated probabilities closer to the actual distribution of T.
				[2]			
9	(e)		The cumulative frequencies in column H could be used e.g. to estimate the median and quartiles and	E1	2.2b	For relevant reference to data in col. H	Other valid methods possible
			check whether they are equally spaced	E1	2.2b	For suitable suggested use	

Question			Answer	Marks	AOs	Guidance	
				[2]			

Question			Answer	Marks	AOs	Guidance	
10	(a)		A *t* test should be carried out	**B1**	**3.3**		
			The Normal probability plot is roughly straight and the Kolmogorov-Smirnov *p*-value is not too low …	**E1**	**1.1**	Both aspect to be referred to	
			… so there is no reason to doubt that the data comes from a Normal distribution (as required for a *t* test)	**E1**	**2.2b**	For conclusion about Normality	
				[3]			
10	(b)		**DR**				
			H_0: $\mu = 441$, H_1: $\mu \neq 441$,	**B1**	**1.1a**	If hypotheses are stated in words only, 'population' must be included	
			where μ is the population mean exhaust gas temperature	**B1**	**1.2**	For definition in context	
			Sample mean = 438.825	**B1**	**1.1**	**BC**	
			Sample SD = 3.802	**B1**	**1.1**	**BC** (or sample variance = 14.452 95…)	
			Test statistic is $\frac{438.8-441}{3.802/\sqrt{12}}$	**M1**	**3.3**	**FT** their mean and/or sd (or var)	
			= –1.982	**A1**	**1.1**		
			Refer to t_{11}	**M1**	**3.4**	Must be evidence of using 11 df	
			Critical value (2-tailed) at 5% level is 2.201	**A1**	**1.1**		
			–1.982 > –2.201 so not significant (do not reject H_0)	**M1**	**2.2b**	Or 1.982 < 2.201	Accept use of *p*-values
			Insufficient evidence to suggest that the mean exhaust gas temperature is different from 441°C	**E1**	**3.5a**	**FT** their comparison, but only if 11 df	
				[10]			

Question			Answer	Marks	AOs	Guidance	
11	(a)		$q-\frac{p}{a}=0$	M1	3.1a	Use of F(a) = 0	
			$q-\frac{p}{b}=1$	M1	1.1	Use of F(b) = 1	
			$\frac{p}{a}-\frac{p}{b}=1$	M1	3.1a	Eliminating q	
			$bp-ap=ab \Rightarrow p(b-a)=ab$	M1	2.1	Solving for p	
			$p=\frac{ab}{b-a}$	A1	1.1	**AG**	
				[5]			
11	(b)		p = 125 and q = 1.25	B1	1.1	For both	
			$\mathrm{P}(150 \le T \le 200)=\left(1.25-\frac{125}{200}\right)-\left(1.25-\frac{125}{150}\right)$	M1	1.1a	Use of F(200) – F(150)	
			= 0.625 – 0.41667 = 0.208	A1	1.1		
				[3]			
11	(c)	(i)	$\mathrm{f}(t)=125t^{-2}$	M1	3.1a	For differentiation	
			$\mathrm{E}(T)=\int_{100}^{500}\frac{125}{t}\,\mathrm{d}t$	M1	1.1		
			= 201.2	A1	1.1	**BC**	$(125\ln 5 = 201.1797...)$
				[3]			
11	(c)	(ii)	$\mathrm{E}(T^2)=\int_{100}^{500}125\,\mathrm{d}t=50\,000$	M1	1.1a	**BC**	
			$\mathrm{Var}(T)=50\,000-(201.17...)^2=9526.7...$	M1	1.1		
			SD(T) = 97.6	A1	1.1		(97.604 879…)
				[3]			

Question			Answer	Marks	AOs	Guidance	
11	(d)			B1 B1 [2]	1.1 1.1	For main part of graph For axes and for part where f(t) = 0	
11	(e)		Because the mode is at the highest point of the graph of the pdf.	E1 [1]	2.2a		

BLANK PAGE