MEI

Mathematics in Education and Industry

MEI STRUCTURED MATHEMATICS

STATISTICS 1, S1

Practice Paper S1-B

Additional materials: Answer booklet/paper
Graph paper
MEI Examination formulae and tables (MF12)

TIME 1 hour 30 minutes

INSTRUCTIONS

- Write your Name on each sheet of paper used or the front of the booklet used.
- Answer **all** the questions.
- You **may** use a graphical calculator in this paper.

INFORMATION

- The number of marks is given in brackets [] at the end of each question or part-question.
- You are advised that you may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- Final answers should be given to a degree of accuracy appropriate to the context.
- The total number of marks for this paper is **72**.

Section A (36 marks)

1 In a certain region of the country the percentages of blood donors in each of four different blood groups are as follows.

Group	O	A	B	AB
Percentage	45%	43%	9%	3%

(i) Two of the region's donors are chosen at random. Find the probability that

(*A*) at least one has blood group O, [2]

(*B*) both have the same blood group. [3]

(ii) Show that at least 8 donors would have to be chosen, at random, to be 99% sure of finding at least one donor of blood type O. [3]

2 A company which hires out equipment by the day has three mowers. The number, X, of mowers which are hired on any one day has the following probability distribution.

$$\mathrm{P}(X = r) = k\left(\frac{1}{2}\right)^r \quad \text{for } r = 0, 1, 2 \text{ and } 3.$$

(i) Show that $k = \frac{8}{15}$. [2]

(ii) Sketch the probability distribution of X. [2]

(iii) Calculate the expectation and variance of X. [4]

3 Next September I intend to buy a new car. Its registration plate will be of the form

HW 55 MSD

where HW is the local area code for the Isle of Wight, 55 represents the second half of the year 2005, and the last three letters are chosen at random.

Both parts of the question refer to the last three letters of the registration plate. You may assume that each of the 26 letters in the alphabet, of which 5 are vowels and 21 are consonants, is available for each of the random choices.

(i) Find the probability that the random letters on the plate are MSD, appearing in that order. [2]

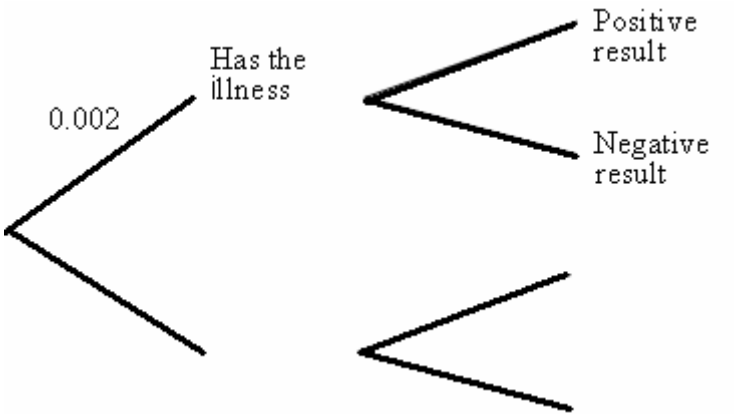
(ii) Find the probability that the letters are M, S, D in any order. [2]

4 New-born babies are tested for a mild illness which affects 1 in 500 babies. The result of a test is either positive or negative. A positive result suggests that the baby has the illness. However, the test is not perfect.

- For babies with the illness, the probability of a positive result is 0.99
- For babies without the illness, the probability of a negative result is 0.95

A new-born baby is chosen at random and tested for the illness.

(i) Copy and complete this probability tree diagram to illustrate the situation. [2]



(ii) Find the probability that the result is positive. [3]

(iii) Given that the result of a test is positive, show that the conditional probability that the baby has the illness is 0.038 (correct to 3 decimal places). [3]

5 The magazine *Nearly Eighteen* has a web site, on which it recently ran a pop trivia quiz with ten questions. The results of the first 1000 entries were analysed. Summary statistics for the numbers of questions answered correctly, x, and associated frequencies, f, are as follows:

$$n = 1000 \qquad \Sigma xf = 6100 \qquad \Sigma x^2 f = 42260$$

(i) Show that the mean score is 6.1. [1]

(ii) Calculate the mean square deviation and the variance, s^2, of the data. Comment on the relative size of the two answers. [3]

Each question in the quiz was in fact of the multiple choice variety, with four possible answers. Three points are awarded for a question answered correctly and one point is deducted for a question which is not answered correctly.

(iii) Show that, if x questions are answered correctly, the number of points, y, is given by $y = 4x - 10$. [1]

(iv) Hence find the mean and variance of the points scored. [3]

Section B (36 marks)

6 A motoring magazine carried out a survey of the value of 60 petrol-driven cars that were five years old. In the survey, the value of each car was expressed as a percentage of its value when new. The results of the survey are summarised in the following table.

Percentage of original value (x)	Number of cars
$15 \leq x < 20$	4
$20 \leq x < 25$	12
$25 \leq x < 30$	18
$30 \leq x < 35$	13
$35 \leq x < 40$	6
$40 \leq x < 45$	5
$45 \leq x < 55$	2

(i) Draw a histogram on graph paper to illustrate the data. [4]

(ii) Carefully describe the shape of the distribution. [2]

(iii) Calculate an estimate of the median of the data. [2]

(iv) Use your calculator to find estimates of the mean and standard deviation of the data, giving your answers correct to 2 decimal places. [4]

(v) Hence identify any outliers, explaining your method. [4]

A similar survey of 60 five-year old diesel-driven cars produced a mean of 34.2% and a standard deviation of 11.7%.

(vi) Use these statistics to compare the values of petrol and diesel cars, five years after they were purchased as new. [2]

7 A motoring organisation reports that the proportion of drivers who would fail a basic sight test is 1 in 6.

For parts **(i)** to **(iii)** you may assume that the report is correct.

(i) Write down the value of p, the probability that a driver chosen at random would **pass** the sight test. [1]

(ii) A random sample of 30 drivers is taken.

(A) How many would be expected to pass the test ?

(B) Find the probability that exactly this number pass the test. [4]

(iii) A random sample of n drivers is taken.

(A) Find the probability that all pass the sight test when $n = 13$.

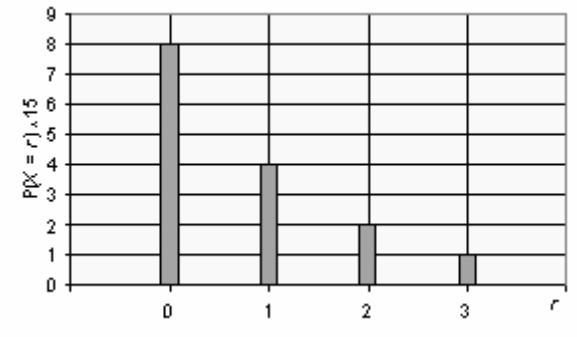
(B) Find the smallest value of n such that the probability of all drivers passing is less than 5% . [4]

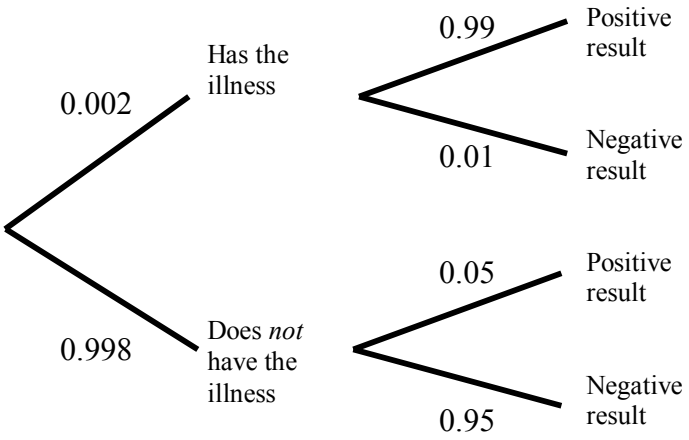
A journalist wishes to test the accuracy of the motoring organisation's report by checking the sight of a random sample of 15 drivers.

(iv) Write down suitable hypotheses for this test in terms of p. [2]

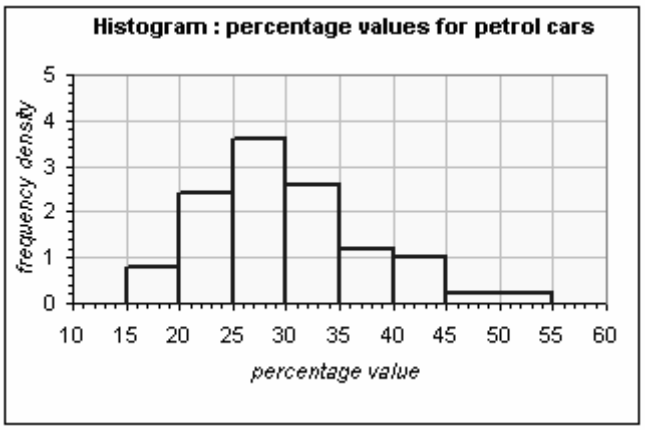
(v) Find the critical region for the test at the 10% significance level and illustrate it on a number line. [5]

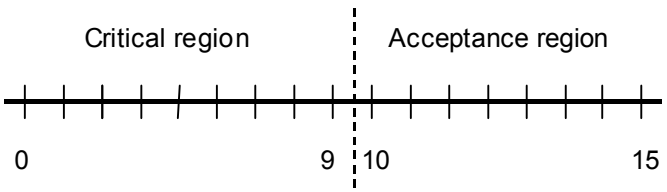
(vi) Find the minimum sample size for which the upper tail of the critical region would not be empty. [2]

Qu		Answer	Mark	Comment
Section A				
1	**(i)**	(*A*) P(at least one has blood group O) $= 1 - 0.55^2 = 0.698$ (to 3 s.f.)	M1 A1 **2**	For calculation
		(*B*) P(both have same blood group) $= 0.45^2 + 0.43^2 + 0.09^2 + 0.03^2$ $= 0.396$ (to 3.s.f.)	M1 M1 A1 **3**	For at least 2 squares For sum of 4 squares
	(ii)	For $n = 7$, P(at least one has blood group O) $= 1 - 0.55^7 = 0.985$ (to 3 s.f.) $< 99\%$ For $n = 8$, P(at least one has blood group O) $= 1 - 0.55^8 = 0.992$ (to 3 s.f.) $> 99\%$	M1 A1 E1 **3**	Attempt at least one evaluation & comparison Both evaluations correct Comparisons
2	**(i)**	r: 0, 1, 2, 3 P($X = r$): k, $\frac{1}{2}k$, $\frac{1}{4}k$, $\frac{1}{8}k$ $k\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}\right) = 1 \Rightarrow \frac{15}{8}k = 1 \Rightarrow k = \frac{8}{15}$	M1 A1 **2**	For forming equation For solution with fractions
	(ii)		G1 G1 **2**	For lines in proportion (*dependent*) For scaled axes
	(iii)	$E(X) = \Sigma r\, P(X = r)$ $= 0 \times \frac{8}{15} + 1 \times \frac{4}{15} + 2 \times \frac{2}{15} + 3 \times \frac{1}{15} = \frac{11}{15}$ $Var(X) = E(X^2) - [E(X)]^2$ $= 0 \times \frac{8}{15} + 1 \times \frac{4}{15} + 4 \times \frac{2}{15} + 9 \times \frac{1}{15} - \left(\frac{11}{15}\right)^2$ $= \frac{194}{225} = 0.862$ (to 3 s.f.)	M1 A1 M1 A1 **4**	For $E(X)$ For $E(X^2)$
3	**(i)**	P(letters on plate are MSD, in that order) $= \left(\frac{1}{26}\right)^3 = \frac{1}{17576} = 0.000057$ (to 2 s.f.)	M1 A1 **2**	For $\left(\frac{1}{26}\right)^3$
	(ii)	P(letters on plate are M, S, D in any order) $= 3! \times \left(\frac{1}{26}\right)^3 = \frac{6}{17576} = \frac{3}{8788} = 0.00034$ (to 2 s.f.)	M1 A1 **2**	For "3! × their part (i)"

4	**(i)**	Tree diagram:  	B1 B1 **2**	For 3 correct probabilities For 2 further correct probabilities
	(ii)	P(result is positive) $= 0.002 \times 0.99 + 0.998 \times 0.05$ $= 0.05188 = 0.052$ (to 2 s.f.)	M1 M1 A1 **3**	For one product For sum of 2 products
	(iii)	P(baby has the illness | positive test result) $= \dfrac{\text{P(baby has the illness and test result positive)}}{\text{P(test result positive)}}$ $= \dfrac{0.002 \times 0.99}{0.05188} = 0.038$ (3 d.p.)	M1 M1 A1 **3**	M1 for numerator M1 for quotient with their part (ii)
5	**(i)**	Mean $= \dfrac{6100}{1000} = 6.1$	B1 **1**	For mean
	(ii)	$msd = \dfrac{42260 - 1000 \times 6.1^2}{1000} = \dfrac{5050}{1000} = 5.05$ $variance = \dfrac{42260 - 1000 \times 6.1^2}{999} = \dfrac{5050}{999} = 5.055$ They are almost the same due to large n.	M1 A1 E1 **3**	For S_{xx} For both values For reference to n
	(iii)	Number of points $= 3x - (10 - x) = 4x - 10$	B1 **1**	For "$3x - (10 - x)$"
	(iv)	$\bar{y} = 4\bar{x} - 10 = 4 \times 6.1 - 10 = 14.4$ variance(y) $= 16 \times$ variance(x) $= 16 \times 5.055\ldots = 80.9$ (to 3 s.f.)	M1 A1 B1 **3**	For $\bar{y}$ For "16 × their variance(x)"

Total = 36

Qu		Answer	Mark	Comment
Section B				
6	**(i)**		G1 G1 G1 G1 **4**	For linear scaled axes For frequency density *or* equivalent *or* key For heights of first 6 bars (joined); all correctly positioned For size of 7th bar
	(ii)	Shape of distribution: It has positive skew.	B1 B1 **2**	For *skew* For *positive*
	(iii)	Find median by simple interpolation $= 25 + \frac{14.5}{18} \times 5 = 29$ *or* $25 + \frac{14}{18} \times 5 = 28.9$	M1 A1 **2**	For identifying interval containing median For precise value
	(iv)	Mid-interval points: 17.5, 22.5, 27.5, 32.5, 37.5, 42.5, 50 Mean $= \frac{1795}{60} = 29.92$ (to 2 d.p.) s.d. $= \sqrt{\frac{57112.5 - 60 \times 29.92^2}{59}} = 7.60$ (to 2 d.p.)	B1 B1 M1 A1 **4**	For mid-interval points For mean For variance ft their mean
	(v)	Mean – 2 s.d. = 29.92 – 2 x 7.60 = 14.71 (to 2 d.p.) Mean + 2 s.d. = 29.92 + 2 x 7.60 = 45.13 (to 2 d.p.) Hence the 2 percentages in range 45 to 55 could be outliers, since they may lie more than 2 s.d. from the mean.	M1 A1 M1 A1 **4**	For attempt at $\bar{x} \pm 2sd$ For both values For "2 outliers" *or equivalent*
	(vi)	On average cars with diesel engines held their value better than cars with petrol engines. Greater variation in the percentage values for cars with diesel engines compared to cars with petrol engines.	E1 E1 **2**	*or equivalent* *or equivalent*

7	**(i)**	$p = \frac{5}{6}$	B1 **1**	For probability
	(ii)(A)	Expected number to pass $= 30 \times \frac{5}{6} = 25$	B1 M1 **2**	For expected number For $(5/6)^r \times (1/6)^{30-r}$ For ${}^{30}C_r \times \ldots$
	(ii)(B)	P(exactly 25 pass) $= {}^{30}C_{25} \times \left(\frac{5}{6}\right)^{25} \times \left(\frac{1}{6}\right)^{5}$ $= 0.192$ (to 3 s.f.) $= 0.19$ (to 2 s.f.)	M1 A1 **2**	
	(iii)(A)	P(all pass sight test) $= 1 - 0.9065 = 0.0935$ [*using tables*] *or* $= \left(\frac{5}{6}\right)^{13} = 0.0935$ (to 3 s.f.)	M1 A1 **2**	For use of tables or working out
	(iii)(B)	Searching for appropriate n: for $n = 16$: P(all pass) $= 1 - 0.9459$ [*tables*] *or* $\left(\frac{5}{16}\right)^{16} = 0.0541$ for $n = 17$: P(all pass) $= 1 - 0.9549$ [*tables*] *or* $\left(\frac{5}{16}\right)^{17} = 0.0451$ hence smallest sample size is where $n =$ **17**	M1 A1 **2**	For attempt at search
	(iv)	H_0: $p = \frac{5}{6}$; H_1: $p \neq \frac{5}{6}$	B1,1 **2**	For hypotheses
	(v)	Using binomial tables for $n = 15$: $P(X \leq 9) =$ **0.0274** < 0.05, but $P(X \leq 10) =$ **0.0898** > 0.05 So lower tail of crit. reg. is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} $P(X \geq 15) = 1 - P(X < 14) = 1 - 0.9351$ $=$ **0.0649** > 0.05, So upper tail of crit. reg. is empty.  	M1 A1 M1 A1 G1 **5**	For at least one comparison For comparison For diagram
	(vi)	From part (iv), for $n = 16$: P(all pass) $= 0.0541 > 5\%$, and for $n = 17$: P(all pass) $= 0.0451 < 5\%$. Hence minimum sample size for which upper tail is not empty is 17.	M1 A1 **2**	For comparisons

Total = 36