MEI STRUCTURED MATHEMATICS   **4766**

Statistics 1

MEI STATISTICS   **G241**

Statistics 1 (Z1)

Wednesday   **24 MAY 2006**   Afternoon   1 hour 30 minutes

Additional materials:
    8 page answer booklet
    Graph paper
    MEI Examination Formulae and Tables (MF2)

TIME   1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- You are permitted to use a graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets [ ] at the end of each question or part question.
- 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.

This question paper consists of 5 printed pages and 3 blank pages.

Section A (36 marks)

1 Every day, George attempts the quiz in a national newspaper. The quiz always consists of 7 questions. In the first 25 days of January, the numbers of questions George answers correctly each day are summarised in the table below.

Number correct	1	2	3	4	5	6	7
Frequency	1	2	3	3	4	7	5

   (i) Draw a vertical line chart to illustrate the data. [2]

  (ii) State the type of skewness shown by your diagram. [1]

 (iii) Calculate the mean and the mean squared deviation of the data. [3]

 (iv) How many correct answers would George need to average over the next 6 days if he is to achieve an average of 5 correct answers for all 31 days of January? [2]

2 Isobel plays football for a local team. Sometimes her parents attend matches to watch her play.

-    A is the event that Isobel's parents watch a match.

-    B is the event that Isobel scores in a match.

You are given that $P(B \mid A) = \frac{3}{7}$ and $P(A) = \frac{7}{10}$.

 (i) Calculate $P(A \cap B)$. [2]

The probability that Isobel does not score and her parents do not attend is 0.1.

 (ii) Draw a Venn diagram showing the events A and B, and mark in the probability corresponding to each of the regions of your diagram. [2]

 (iii) Are events A and B independent? Give a reason for your answer. [2]

 (iv) By comparing $P(B \mid A)$ with $P(B)$, explain why Isobel should ask her parents not to attend. [2]

3 The score, X, obtained on a given throw of a biased, four-faced die is given by the probability distribution

$$P(X = r) = kr(8 - r) \text{ for } r = 1, 2, 3, 4.$$

 (i) Show that $k = \frac{1}{50}$. [2]

 (ii) Calculate $E(X)$ and $Var(X)$. [5]

4 Peter and Esther visit a restaurant for a three-course meal. On the menu there are 4 starters, 5 main courses and 3 sweets. Peter and Esther each order a starter, a main course and a sweet.

  (i) Calculate the number of ways in which Peter may choose his three-course meal. [2]

 (ii) Suppose that Peter and Esther choose different dishes from each other.

    (*A*) Show that the number of possible combinations of starters is 6. [1]

    (*B*) Calculate the number of possible combinations of 6 dishes for both meals. [1]

(iii) Suppose instead that Peter and Esther choose their dishes independently.

    (*A*) Write down the probability that they choose the same main course. [1]

    (*B*) Find the probability that they choose different dishes from each other for every course. [2]

5 Douglas plays darts, and the probability that he hits the number he is aiming at is 0.87 for any particular dart.

Douglas aims a set of three darts at the number 20; the number of times he is successful can be modelled by $B(3, 0.87)$.

  (i) Calculate the probability that Douglas hits 20 twice. [3]

 (ii) Douglas aims fifty sets of 3 darts at the number 20. Find the expected number of sets for which Douglas hits 20 twice. [1]

(iii) Douglas aims four sets of 3 darts at the number 20. Calculate the probability that he hits 20 twice for two sets out of the four. [2]
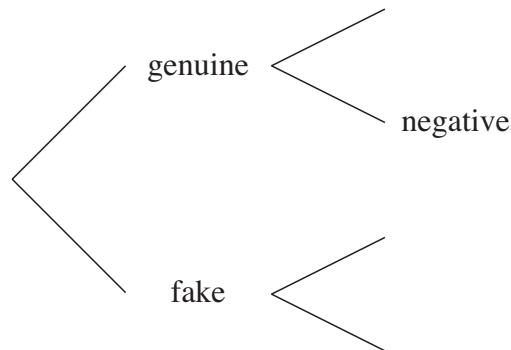
Section B (36 marks)

6 It has been estimated that 90% of paintings offered for sale at a particular auction house are genuine, and that the other 10% are fakes. The auction house has a test to determine whether or not a given painting is genuine. If this test gives a positive result, it suggests that the painting is genuine. A negative result suggests that the painting is a fake.

If a painting is genuine, the probability that the test result is positive is 0.95.

If a painting is a fake, the probability that the test result is positive is 0.2.

(i) Copy and complete the probability tree diagram below, to illustrate the information above.

[2]

Calculate the probabilities of the following events.

(ii) The test gives a positive result. [2]

(iii) The test gives a correct result. [2]

(iv) The painting is genuine, given a positive result. [3]

(v) The painting is a fake, given a negative result. [3]

A second test is more accurate, but very expensive. The auction house has a policy of only using this second test on those paintings with a negative result on the original test.

(vi) Using your answers to parts **(iv)** and **(v)**, explain why the auction house has this policy. [2]
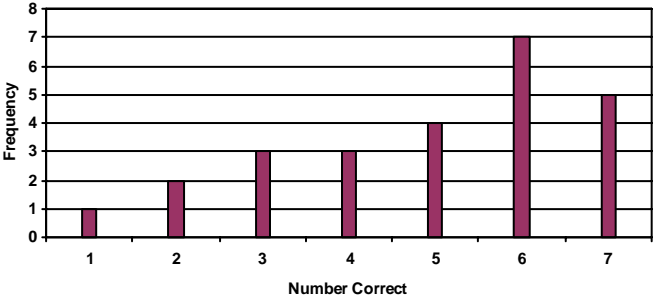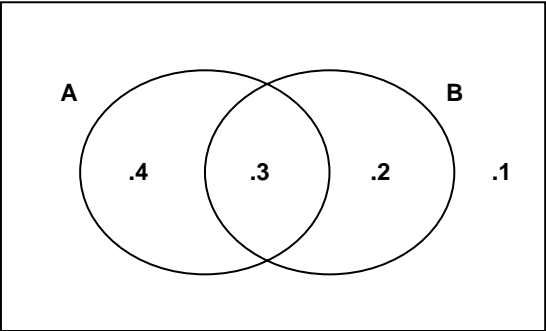
The probability that the second test gives a correct result is 0.96 whether the painting is genuine or a fake.

(vii) Three paintings are independently offered for sale at the auction house. Calculate the probability that all three paintings are genuine, are judged to be fakes in the first test, but are judged to be genuine in the second test. [4]
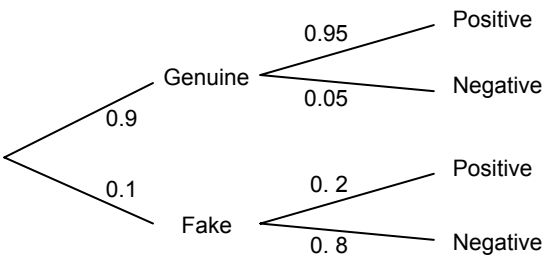
7   A geologist splits rocks to look for fossils. On average 10% of the rocks selected from a particular area do in fact contain fossils.

The geologist selects a random sample of 20 rocks from this area.

 (i) Find the probability that

   (*A*) exactly one of the rocks contains fossils, [3]

   (*B*) at least one of the rocks contains fossils. [3]

 (ii) A random sample of *n* rocks is selected from this area. The geologist wants to have a probability of 0.8 or greater of finding fossils in at least one of the *n* rocks. Find the least possible value of *n*. [3]

 (iii) The geologist explores a new area in which it is claimed that less than 10% of rocks contain fossils. In order to investigate the claim, a random sample of 30 rocks from this area is selected, and the number which contain fossils is recorded. A hypothesis test is carried out at the 5% level.

   (*A*) Write down suitable hypotheses for the test. [3]

   (*B*) Show that the critical region consists only of the value 0. [4]

   (*C*) In fact, 2 of the 30 rocks in the sample contain fossils. Complete the test, stating your conclusions clearly. [2]

Q1 **(i)**		G1 Labelled linear scales G1 Height of lines	**2**
(ii)	Negative (skewness)	B1	**1**
(iii)	Σfx = 123 so mean = 123/25 = 4.92 o.e. $S_{xx} = 681 - \dfrac{123^2}{25} = 75.84$ M.s.d = $\dfrac{75.84}{25} = 3.034$	B1 M1 for S_{xx} attempted A1 FT their 4.92	**3**
(iv)	Total for 25 days is 123 and totals for 31 days is 155. Hence total for next 6 days is 32 and so mean = 5.33	M1 31 x 5 – 25xtheir 4.92 A1 FT their 123	**2**
		TOTAL	**8**
Q2 **(i)**	$P(A \cap B) = P(A)P(B\mid A) = \dfrac{7}{10} \times \dfrac{3}{7}$ $\rightarrow P(A \cap B) = 0.3$ o.e.	M1 Product of these fractions A1	**2**
(ii)		B1FT either 0.4 or 0.2 in correct place B1FT all correct and labelled	**2**
(iii)	$P(B\mid A) \neq P(B)$, 3/7 ≠ 0.5 Unequal so not independent	E1 Correct comparison E1*dep* for 'not independent'	**2**
(iv)	3/7 < 0.5 so Isobel is less likely to score when her parents attend	E1 for comparison E1*dep*	**2**
		TOTAL	**8**

Q3 (i)	P(X = 1) = 7k, P(X = 2) = 12k, P(X = 3) = 15k, P(X = 4) = 16k 50k = 1 so k = 1/50	M1 for addition of four multiples of k A1 **ANSWER GIVEN**	**2**
(ii)	E(X) = 1 x 7k + 2 x 12k + 3 x 15k + 4 x 16k = 140k = 2.8 OR E(X) = 1 x $^7/_{50}$ + 2 x $^{12}/_{50}$ + 3 x $^{15}/_{50}$ + 4 x $^{16}/_{50}$ = $^{140}/_{50}$ = 2.8 oe Var(X) = 1 x 7k + 4 x 12k + 9 x 15k + 16 x 16k -7.84 = 1.08 OR Var(X) = 1 x $^7/_{50}$ + 4 x $^{12}/_{50}$ + 9 x $^{15}/_{50}$ + 16 x $^{16}/_{50}$ -7.84 = 8.92 – 7.84 = 1.08	M1 for Σxp (at least 3 terms correct) A1 CAO M1 $\Sigma x^2 p$ (at least 3 terms correct) M1*dep* for – their E(X)² NB provided Var(X) > 0 A1 FT their E(X)	**5**
		TOTAL	**7**
Q4 (i)	4 x 5 x 3 = 60	M1 for 4 x 5 x 3 A1 CAO	**2**
(ii)	(**A**) $\binom{4}{2}$ = 6 (**B**) $\binom{4}{2}\binom{5}{2}\binom{3}{2}$ = 180	B1 **ANSWER GIVEN** B1 CAO	**2**
(iii)	(**A**) 1/5 (**B**) $\frac{3}{4} \times \frac{4}{5} \times \frac{2}{3} = \frac{2}{5}$	B1 CAO M1 for $\frac{3}{4} \times \frac{4}{5} \times \frac{2}{3}$ A1	**3**
		TOTAL	**7**
Q5 (i)	$P(X=2)=\binom{3}{2}\times 0.87^2 \times 0.13 = 0.2952$	M1 0.87^2 x 0.13 M1 $\binom{3}{2}$ x $p^2 q$ **with p+q=1** A1 CAO	**3**
(ii)	In 50 throws expect 50 (0.2952) = 14.76 times	B1 FT	**1**
(iii)	P (two 20's twice) = $\binom{4}{2} \times 0.2952^2 \times 0.7048^2 = 0.2597$	M1 $0.2952^2 \times 0.7048^2$ A1 FT their 0.2952	**2**
		TOTAL	**6**

Q6 (i)		G1 for left hand set of branches fully correct including labels and probabilities G1 for right hand set of branches fully correct	2
(ii)	P (test is positive) = (0.9)(0.95) + (0.1)(0.2) = 0.875	M1 Two correct pairs added A1 CAO	2
(iii)	P (test is correct) = (0.9)(0.95) + (0.1)(0.8) = 0.935	M1 Two correct pairs added A1 CAO	2
(iv)	P (Genuine|Positive) = 0.855/0.875 = 0.977	M1 Numerator M1 Denominator A1 CAO	3
(v)	P (Fake|Negative) = 0.08/0.125 = 0.64	M1 Numerator M1 Denominator A1 CAO	3
(vi)	EITHER: A positive test means that the painting is almost certain to be genuine so no need for a further test. However, more than a third of those paintings with a negative result are genuine so a further test is needed. NOTE: Allow sensible alternative answers	E1FT E1FT	2
(vii)	P (all 3 genuine) = $(0.9 \times 0.05 \times 0.96)^3$ $= (0.045 \times 0.96)^3$ $= (0.0432)^3$ $= 0.0000806$	M1 for 0.9 x 0.05 (=0.045) M1 for complete correct triple product M1*indep* for cubing A1 CAO	4
		TOTAL	**18**

Q7 **(i)**	$X \sim B(20, 0.1)$ (A) $P(X = 1) = \binom{20}{1} \times 0.1 \times 0.9^{19} = 0.2702$ OR from tables $0.3917 - 0.1216 = 0.2701$ (B) $P(X \geq 1) = 1 - 0.1216 = 0.8784$	M1 0.1×0.9^{19} M1 $\binom{20}{1} \times pq^{19}$ A1 CAO OR: M2 for 0.3917 – 0.1216 A1 CAO M1 P(X=0) *provided that P(X≥1)=1–P(X≤1) not seen* M1 1-P(X=0) A1 CAO	3 3
(ii)	EITHER: $1 - 0.9^n \geq 0.8$ $0.9^n \leq 0.2$ Minimum n = 16 OR (using trial and improvement): Trial with 0.9^{15} or 0.9^{16} or 0.9^{17} $1 - 0.9^{15} = 0.7941 < 0.8$ and $1 - 0.9^{16} = 0.8147 > 0.8$ Minimum n = 16 NOTE: n = 16 unsupported scores SC1 only	M1 for 0.9^n M1 for inequality A1 CAO M1 M1 A1 CAO	3
(iii)	(A) Let p = probability of a randomly selected rock containing a fossil (for population) H_0: $p = 0.1$ H_1: $p < 0.1$ (B) Let $X \sim B(30, 0.1)$ $P(X \leq 0) = 0.0424 < 5\%$ $P(X \leq 1) = 0.0424 + 0.1413 = 0.1837 > 5\%$ So critical region consists only of 0. (C) 2 does not lie in the critical region. So there is insufficient evidence to reject the null hypothesis and we conclude that it seems that 10% of rocks in this area contain fossils.	B1 for definition of p B1 for H_0 B1 for H_1 M1 for attempt to find $P(X \leq 0)$ or $P(X \leq 1)$ using binomial M1 for both attempted M1 for comparison of either of the above with 5% A1 for critical region dep on both comparisons **(NB Answer given)** M1 for comparison A1 for conclusion **in context**	3 4 2
		TOTAL	**18**