OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 4767

Statistics 2

Wednesday **25 JANUARY 2006** Morning 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.

1 A roller-coaster ride has a safety system to detect faults on the track.

(i) State conditions for a Poisson distribution to be a suitable model for the number of faults occurring on a randomly selected day. [2]

Faults are detected at an average rate of 0.15 per day. You may assume that a Poisson distribution is a suitable model.

(ii) Find the probability that on a randomly chosen day there are

(A) no faults,

(B) at least 2 faults. [4]

(iii) Find the probability that, in a randomly chosen period of 30 days, there are at most 3 faults. [3]

There is also a separate safety system to detect faults on the roller-coaster train itself. Faults are detected by this system at an average rate of 0.05 per day, independently of the faults detected on the track. You may assume that a Poisson distribution is also suitable for modelling the number of faults detected on the train.

(iv) State the distribution of the total number of faults detected by the two systems in a period of 10 days. Find the probability that a total of 5 faults is detected in a period of 10 days. [4]

(v) The roller-coaster is operational for 200 days each year. Use a suitable approximating distribution to find the probability that a total of at least 50 faults is detected in 200 days. [5]

2 The drug EPO (erythropoetin) is taken by some athletes to improve their performance. This drug is in fact banned and blood samples taken from athletes are tested to measure their 'hematocrit level'. If the level is over 50 it is considered that the athlete is likely to have taken EPO and the result is described as 'positive'. The measured hematocrit level of each athlete varies over time, even if EPO has not been taken.

(i) For each athlete in a large population of innocent athletes, the variation in measured hematocrit level is described by the Normal distribution with mean 42.0 and standard deviation 3.0.

(A) Show that the probability that such an athlete tests positive for EPO in a randomly chosen test is 0.0038. [3]

(B) Find the probability that such an athlete tests positive on at least 1 of the 7 occasions during the year when hematocrit level is measured. (These occasions are spread at random through the year and all test results are assumed to be independent.) [3]

(C) It is standard policy to apply a penalty after testing positive. Comment briefly on this policy in the light of your answer to part **(i)**(B). [2]

(ii) Suppose that 1000 tests are carried out on innocent athletes whose variation in measured hematocrit level is as described in part **(i)**. It may be assumed that the probability of a positive result in each test is 0.0038, independently of all other test results.

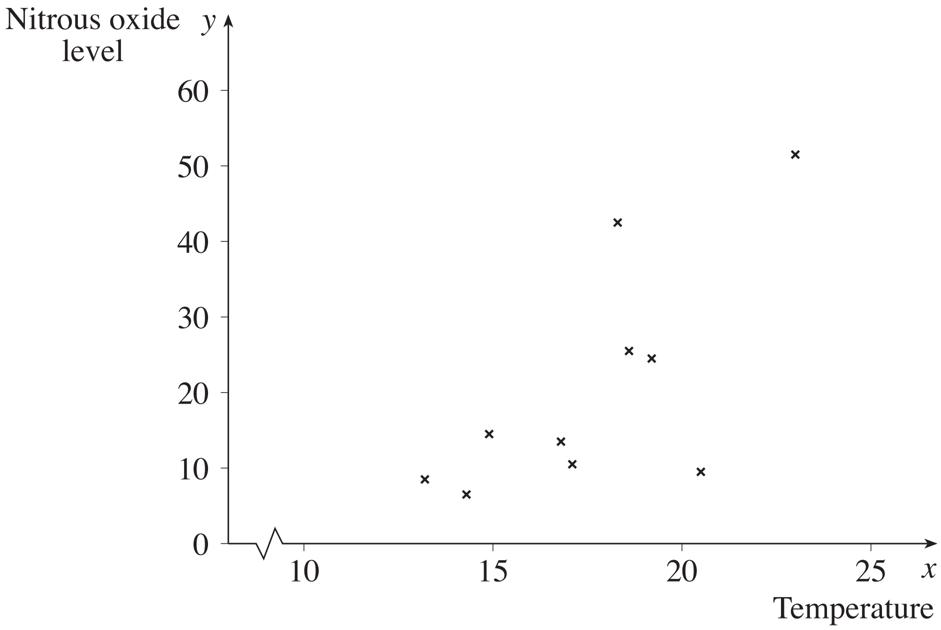
(A) State the exact distribution of the number of positive tests. [2]

(B) Use a suitable approximating distribution to find the probability that at least 10 tests are positive. [4]

(iii) Because of genetic factors, a particular innocent athlete has an abnormally high natural hematocrit level. This athlete's measured level is Normally distributed with mean 48.0 and standard deviation 2.0. The usual limit of 50 for a positive test is to be altered for this athlete to a higher value h. Find the value of h for which this athlete would test positive on average just once in 200 occasions. [4]

3 A researcher is investigating the relationship between temperature and levels of the air pollutant nitrous oxide at a particular site. The researcher believes that there will be a positive correlation between the daily maximum temperature, x, and nitrous oxide level, y. Data are collected for 10 randomly selected days. The data, measured in suitable units, are given in the table and illustrated on the scatter diagram.

x	13.3	17.2	16.9	18.7	18.4	19.3	23.1	15.0	20.6	14.4
y	9	11	14	26	43	25	52	15	10	7



(i) Calculate the value of Spearman's rank correlation coefficient for these data. [5]

(ii) Perform a hypothesis test at the 5% level to check the researcher's belief, stating your hypotheses clearly. [5]

(iii) It is suggested that it would be preferable to carry out a test based on the product moment correlation coefficient. State the distributional assumption required for such a test to be valid. Explain how a scatter diagram can be used to check whether the distributional assumption is likely to be valid and comment on the validity in this case. [3]

(iv) A statistician investigates data over a much longer period and finds that the assumptions for the use of the product moment correlation coefficient are in fact valid. Give the critical region for the test at the 1% level, based on a sample of 60 days. [2]

(v) In a different research project, into the correlation between daily temperature and ozone pollution levels, a positive correlation is found. It is argued that this shows that high temperatures cause increased ozone levels. Comment on this claim. [3]

4 The table summarises the usual method of travelling to school for 200 randomly selected pupils from primary and secondary schools in a city.

		Primary	Secondary
Method of travel	Bus	21	49
	Car	65	15
	Cycle or Walk	34	16

(i) Write down null and alternative hypotheses for a test to examine whether there is any association between method of travel and type of school. [1]

(ii) Calculate the expected frequency for primary school bus users. Calculate also the corresponding contribution to the test statistic for the usual χ^2 test. [4]

(iii) Given that the value of the test statistic for the usual χ^2 test is 42.64, carry out the test at the 5% level of significance, stating your conclusion clearly. [4]

The mean travel time for pupils who travel by bus is known to be 18.3 minutes. A survey is carried out to determine whether the mean travel time to school by car is different from 18.3 minutes. In the survey, 20 pupils who travel by car are selected at random. Their mean travel time is found to be 22.4 minutes.

(iv) Assuming that car travel times are Normally distributed with standard deviation 8.0 minutes, carry out a test at the 10% level, stating your hypotheses and conclusion clearly. [7]

(v) Comment on the suggestion that pupils should use a bus if they want to get to school quickly. [2]

Mark Scheme 4767
January 2006

Question 1

(i)	Faults are detected randomly and independently Uniform (mean) rate of occurrence	B1 B1	**2**
(ii)	(A) $P(X=0) = e^{-0.15}\frac{0.15^0}{0!} = 0.8607$ (B) $P(X \geq 2) = 1 - 0.8607 - e^{-0.15}\frac{0.15^1}{1!}$ $= 1 - 0.8607 - 0.1291 = 0.0102$	M1 for probability calc. M0 for tables unless interpolated A1 M1 A1	**4**
(iii)	$\lambda = 30 \times 0.15 = 4.5$ Using tables: $P(X \leq 3) = 0.3423$	B1 for mean (SOI) M1 attempt to find $P(X \leq 3)$ A1	**3**
(iv)	Poisson distribution with $\lambda = 10 \times (0.15 + 0.05) = 2$ $P(X=5) = e^{-2}\frac{2^5}{5!} = 0.0361$ (3 s.f.) or from tables $= 0.9834 - 0.9473 = 0.0361$	B1 for Poisson stated B1 for $\lambda = 2$ M1 for calculation or use of tables A1 FT	**4**
(v)	Mean no. of items in 200 days = $200 \times 0.2 = 40$ Using Normal approx. to the Poisson, $X \sim N(40,40)$: $P(X \geq 50) = P\left(Z > \frac{49.5 - 40}{\sqrt{40}}\right)$ $= P(Z > 1.502) = 1 - \Phi(1.502) = 1 - 0.9334$ $= 0.0666$ (3 s.f.)	B1 for Normal approx. (SOI) B1 for both parameters B1 for continuity corr. M1 for probability using correct tail A1 **cao, (**but FT wrong or omitted CC)	**5**
			18

Question 2

(i) (*A*)	$X \sim N(42,3^2)$ $P(X > 50.0) = P\left(Z > \frac{50.0-42.0}{3.0}\right)$ $= P(Z > 2.667)$ $= 1 - \Phi(2.667) = 1 - 0.9962$ $= 0.0038$	M1 for standardizing M1 for prob. calc. with correct tail A1 **NB answer given**	**3**
(i) (*B*)	P(not positive) = 0.9962 P(At least one is out of 7 is positive) $= 1 - 0.9962^7 = 1 - 0.9737$ $= 0.0263$	B1 for use of 0.9962 in binomial expression M1 for correct method A1 CAO	**3**
(i) (*C*)	If an innocent athlete is tested 7 times in a year there is a reasonable possibility (1 in 40 chance) of testing positive. Thus it is likely that a number of innocent athletes may come under suspicion and suffer a suspension so the penalty could be regarded as unfair. *Or* this is a necessary evil in the fight against performance enhancing drugs in sport.	E1 comment on their probability in (i) B E1 for sensible contextual conclusion consistent with first comment	**2**
(ii) (*A*)	B(1000, 0.0038)	B1 for B(,) or Binomial B1 *dep* for both parameters	**2**
(ii) (*B*)	A suitable approximating distribution is Poisson(3.8) P(at least 10 positive tests) $= P(X \geq 10) = 1 - P(X \leq 9)$ $= 1 - 0.9942$ $= 0.0058$ *NB Do not allow use of Normal approximation.*	B1 for Poisson soi B1FT *dep* for $\lambda = 3.8$ M1 for attempt to use $1 - P(X \leq 9)$ A1 FT	**4**
(iii)	P(not testing positive) = 0.995 From tables $z = \Phi^{-1}(0.995) = 2.576$ $\frac{h-48.0}{2.0} = 2.576$ $h = 48.0 + 2.576 \times 2.0 = 53.15$	B1 for 0.995 seen (or implied by 2.576) B1 for 2.576 (FT their 0.995) M1 for equation in *h* and positive z-value A1 CAO	**4**
			18

Question 3

(i)	Rank x: 1, 5, 4, 7, 6, 8, 10, 3, 9, 2 Rank y: 2, 4, 5, 8, 9, 7, 10, 6, 3, 1 d: -1, 1, -1, -1, -3, 1, 0, -3, 6, 1 d^2: 1, 1, 1, 1, 9, 1, 0, 9, 36, 1 $r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6 \times 60}{10 \times 99}$ = 0.636 (to 3 s.f.) [*allow* 0.64 to 2 s.f.]	M1 for ranking (allow all ranks reversed) M1 for d^2 A1 CAO for Σd^2 M1 for structure of r_s using their Σd^2 A1 f.t. for $\|r_s\| < 1$ NB No ranking scores zero	**5**
(ii)	H_0: no association between x and y H_1: positive association between x and y Looking for positive association (one-tail test): Critical value at 5% level is 0.5636 Since 0.636 > 0.5636, there is sufficient evidence to reject H_0, i.e. conclude that there appears to be positive association between temperature and nitrous oxide level.	B1 for H_0 B1 for H_1 NB H_0 H_1 <u>not</u> ito rho B1 for ± 0.5636 (FT their H_1) M1 for comparison with c.v., provided $\|r_s\| < 1$ A1 for conclusion in words f.t. their r_s and sensible cv	**5**
(iii)	Underlying distribution must be bivariate normal. If the distribution is bivariate normal then the scatter diagram will have an elliptical shape. This scatter diagram is not elliptical and so a PMCC test would not be valid. (Allow comment indicating that the sample is too small to draw a firm conclusion on ellipticity and so on validity)	B1 CAO for bivariate normal B1 indep for elliptical shape E1 dep for conclusion	**3**
(iv)	n=60, PMCC critical value is $r = 0.2997$ So the critical region is $r \geq 0.2997$	B1 B1 FT their sensible c.v.	**2**
(v)	Any three of the following: • Correlation does not imply causation; • There could be a third factor (causing the correlation between temperature and ozone level); • the claim could be true; • increased ozone could cause higher temperatures.	E1 E1 E1	**3**
			18

Question 4

(i)	H_0: no association between method of travel and type of school; H_1: some association between method of travel and type of school;..	B1 for both	**1**
(ii)	Expected frequency = 120/200 × 70 = 42 Contribution = $(21 - 42)^2 / 42$ = 10.5	M1 A1 M1 for valid attempt at $(O-E)^2/E$ A1 FT their 42 provided O = 21 (at least 1 dp)	**4**
(iii)	$X^2 = 42.64$ Refer to χ_2^2 Critical value at 5% level = 5.991 Result is significant There appears to be some association between method of travel and year group. NB if H_0 H_1 reversed, or 'correlation' mentioned, do not award first B1or final E1	B1 for 2 deg of f(seen) B1 CAO for cv B1 for significant (FT their c.v. provided consistent with their d.o.f. E1	**4**
(iv)	H_0: $\mu = 18.3$; H_1: $\mu \neq 18.3$ Where μ denotes the mean travel time by car for the whole population. Test statistic $z = \dfrac{22.4 - 18.3}{8.0/\sqrt{20}} = \dfrac{4.1}{1.789}$ = 2.292 10% level 2 tailed critical value of z is 1.645 2.292 > 1.645 so significant. There is evidence to reject H_0 It is reasonable to conclude that the mean travel time by car is different from that by bus.	B1 for both correct B1 for definition of μ M1 (standardizing sample mean) A1 for test statistic B1 for 1.645 M1 for comparison leading to a conclusion A1 for conclusion in words and context	**7**
(v)	The test suggests that students who travel by bus get to school more quickly. This may be due to their journeys being over a shorter distance. It may be due to bus lanes allowing buses to avoid congestion. It is possible that the test result was incorrect (ie implication of a Type I error). More investigation is needed before any firm conclusion can be reached.	E1, E1 for any two valid comments	**2** **18**

Question 4 chi squared calculations

H_0: no association between method of travel and type of school;
H_1: some association between method of travel and type of school;

Observed		Type of school: Year 6	Year 11	**Row totals**
Method of travel	Bus	21	49	**70**
	Car	65	15	**80**
	Cycle/Walk	34	16	**50**
Column totals		**120**	**80**	**200**

Expected		Type of school: Year 6	Year 11	**Row totals**
Method of travel	Bus	42	28	**70**
	Car	48	32	**80**
	Cycle/Walk	30	20	**50**
Column totals		**120**	**80**	**200**

Chi Squared Contribution		Type of school: Year 6	Year 11	**Row totals**
Method of travel	Bus	10.50	15.75	**26.25**
	Car	6.02	9.03	**15.05**
	Cycle/Walk	0.53	0.80	**1.33**
Column totals		**17.05**	**25.58**	**42.64**

4767: Statistics 2

General Comments

The majority of candidates were well prepared for this examination with a good overall standard seen. Candidates managed to answer all questions showing a good level of understanding of the topics concerned. A high level of competence in dealing with probability calculations using Binomial, Poisson and Normal distributions was seen; however, many were often unable to assess which distribution to use in any given situation. Candidates showed less understanding of how to phrase hypotheses appropriately for the different types of hypothesis test; where parameters were needed, candidates frequently used inappropriate symbols and in other cases used parameters when not required. It appeared that most candidates had adequate time to complete the paper; even those who calculated all expected frequencies and contributions to the Chi-squared test statistic in Question 4 when asked only for the contribution from one of the six cells.

Comments on Individual Questions

Section A

1) (i) This standard request rarely produced full marks for commenting upon the required conditions - "independence (of events)" and "uniform mean rate" of occurrence. Many candidates recognised the need for independence but struggled with the second condition. Common mistakes included "n large and p small" and "mean equals variance".

(ii)A This part of the question was well done with most candidates calculating the required probability using the Poisson probability function successfully.

(ii)B Many candidates introduced rounding errors when working to different levels of accuracy on $P(X = 0)$ and $P(X = 1)$, then provided only one significant figure accuracy in their answer.

(iii) Most candidates scored full marks.

(iv) Most candidates scored full marks, or lost just one for failing to "state the distribution" that was subsequently used despite this being requested in the question.

(v) Again, well-answered with many candidates obtaining full marks. Most recognised the need for a Normal approximation although some lost marks through using N(np, npq) rather than $N(\lambda, \lambda)$. Some candidates lost a mark for omitting the necessary continuity correction or applying an incorrect one. The majority used the correct tail.

2) (i)*A* This was well answered with most candidates successfully standardising and identifying the need to use $1 - \Phi(2.667)$. As the answer was given in the question, some candidates were penalised for not indicating that $\Phi(2.667) =$ 0.9962. Most candidates used Normal tables accurately, but some candidates failed to capitalise on the fact that the answer was provided, by using 1 - $\Phi(2.66)$ which gives 0.0039.

(i)*B* The question tested the use of $P(X \geq 1) = 1 - P(X = 0)$, with $P(X = 0)$ being found using 0.9962^7 (although a Poisson approximation could be used to give an acceptable answer). Many candidates appeared unable to understand what was required at all; some simply found $P(X = 1)$.

(i)*C* Few fully convincing comments were provided, with many simply restating their answer to (i) *B* rather than interpreting it. Candidates could obtain full marks for commenting on the magnitude of their answer and relating this to the fairness of the drug test.

(ii)*A* Many candidates failed to recognise that the exact distribution required was B(1000, 0.0038). Those realising a binomial distribution was needed generally gained both marks; both parameters were needed for full credit.

(ii)*B* This part was well answered with most candidates correctly using a Poisson approximation with mean 3.8, of which only a few used the incorrect $P(X \geq 10) = 1 - P(X \leq 10)$; a mistake often seen in such questions in previous years. Due to values of the parameters of the binomial distribution in this question, those using a Normal approximation were given no credit.

(iii) Many scored full marks for this question. Some candidates got off to a poor start by miscalculating 1 – 0.005 as 0.95; even so, credit could still be obtained for obtaining a corresponding *z*-value and using it to obtain a value in the right-hand tail of the Normal distribution. Some candidates used -2.576 and were thus penalised for working with the wrong tail.

3) (i) Many scored full marks on this part. Those failing to rank scores scored no marks; this happened with a significant proportion of candidates. A number of candidates made errors with their ranking but otherwise applied the correct expression for calculating Spearman's rank correlation coefficient. Other candidates were penalised for omitting the "1 -" from their expression.

(ii) Most candidates scored well on this part of the question; with marks for a critical value of 0.5636 and a comparison with their r_s from part (i) gained by most. The main reasons for loss of marks in this question were a failure to provide correct, contextual hypotheses and a failure to include a contextual conclusion. Some candidates who wrote their hypotheses solely in terms of ρ were penalised; although many candidates using hypotheses in terms of ρ also stated their hypotheses in words and could gain full credit. At this level, conclusions to hypothesis tests should end with a comment relating the findings back to the original context of the question.

(iii) An increasing number of candidates now seem comfortable with the idea of the need for an underlying distribution which is bivariate normal when carrying out a test for the product moment correlation coefficient, although the majority of candidates struggle to get this across. Many candidates knew to comment on the need for an elliptical spread of points on the scatter graph and went on to make a decisive comment on the appropriateness of the test with the points provided. A large number of candidates scored no marks on this part, for answers which explained about the need for the points to lie in a straight line together with general comments about correlation.

(iv) Most candidates identified a correct critical value of 0.2997 although some used the corresponding value from the Spearman's table or the 1% two-tailed test value. Few candidates gained the second mark for identifying the critical region as being $r \geq 0.2997$; the majority quoting the critical region as $r = 0.2997$.

(v) This part was reasonably attempted with many candidates gaining marks for appreciating that "correlation does not imply causation" and that a third factor could be involved. A number of candidates obtained full marks for providing a third relevant comment such as "the claim could be true" and "it could be that increased ozone could be the cause of high temperatures". Marks were awarded for comments of a statistical nature rather than lengthy essays on global warming.

4) (i) Well answered; although candidates who avoided context were penalised, as were those using "correlation" or parameters in their hypotheses.

(ii) Many candidates gained four straightforward marks here. Some candidates clearly failed to read the question carefully; as a result, they wasted time that the question was designed to save – even so, full credit could be obtained provided that the answer of 10.5 was seen in their working. Many candidates simply worked out the expected frequency of 42 without going on to find the contribution to the value of the chi-squared test statistic.

(iii) Generally well done, but many candidates lost marks for incorrect conclusions and for failing to comment in context; simply concluding that "there appears to be an association" was not enough to be awarded the final mark.

(iv) Although many candidates scored most of the available marks, this part was not well done. Most managed the mark for the hypotheses, which needed to be expressed in terms of μ, but failed to define μ as the mean travel time by car for the whole population. Candidates with a clear understanding of the difference between population mean and sample mean generally fared better. Many failed to use the correct distribution when standardising to find the test-statistic, or when finding the critical value(s) for the sample mean; many used the distribution for car travel times and not the distribution of the mean travel time for samples of size 20. Most gained marks for identifying the critical z-value of 1.645 and comparing it with their test-statistic. A number of candidates mixed up $\mu = 18.3$ with the observed sample mean of 22.4.

(v) Many candidates struggled to make comments related to the test in part (iv) or to factors which might have affected the outcome. Popular correct answers included comments on the fact that students might not all live the same distance from school, and that more investigation is needed. No credit was given to answers speculating about buses breaking down or general, environmental comments. Centres should encourage students to comment using statistical arguments.