OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE **4768/01**

MATHEMATICS (MEI)

Statistics 3

TUESDAY 15 JANUARY 2008 Morning

Time: 1 hour 30 minutes

Additional materials: Answer Booklet (8 pages)
Graph paper
MEI Examination Formulae and Tables (MF2)

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- Read each question carefully and make sure you know what you have to do before starting your answer.
- 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.
- The total number of marks for this paper is 72.
- 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.

This document consists of **4** printed pages.

1 **(a)** The time (in milliseconds) taken by my computer to perform a particular task is modelled by the random variable T. The probability that it takes more than t milliseconds to perform this task is given by the expression $\mathrm{P}(T > t) = \frac{k}{t^2}$ for $t \geqslant 1$, where k is a constant.

(i) Write down the cumulative distribution function of T and hence show that $k = 1$. [3]

(ii) Find the probability density function of T. [2]

(iii) Find the mean time for the task. [3]

(b) For a different task, the times (in milliseconds) taken by my computer on 10 randomly chosen occasions were as follows.

6.4 5.9 5.0 6.2 6.8 6.0 5.2 6.5 5.7 5.3

From past experience it is thought that the median time for this task is 5.4 milliseconds. Carry out a test at the 5% level of significance to investigate this, stating your hypotheses carefully. [10]

2 In the vegetable section of a local supermarket, leeks are on sale either loose (and unprepared) or prepared in packs of 4.

The weights of unprepared leeks are modelled by the random variable X which has the Normal distribution with mean 260 grams and standard deviation 24 grams. The prepared leeks have had 40% of their weight removed, so that their weights, Y, are modelled by $Y = 0.6X$.

(i) Find the probability that a randomly chosen unprepared leek weighs less than 300 grams. [3]

(ii) Find the probability that a randomly chosen prepared leek weighs more than 175 grams. [3]

(iii) Find the probability that the total weight of 4 randomly chosen prepared leeks in a pack is less than 600 grams. [3]

(iv) What total weight of prepared leeks in a randomly chosen pack of 4 is exceeded with probability 0.975? [3]

(v) Sandie is making soup. She uses 3 unprepared leeks and 2 onions. The weights of onions are modelled by the Normal distribution with mean 150 grams and standard deviation 18 grams. Find the probability that the total weight of her ingredients is more than 1000 grams. [3]

(vi) A large consignment of unprepared leeks is delivered to the supermarket. A random sample of 100 of them is taken. Their weights have sample mean 252.4 grams and sample standard deviation 24.6 grams. Find a 99% confidence interval for the true mean weight of the leeks in this consignment. [3]

3 Engineers in charge of a chemical plant need to monitor the temperature inside a reaction chamber. Past experience has shown that when functioning correctly the temperature inside the chamber can be modelled by a Normal distribution with mean 380 °C. The engineers are concerned that the mean operating temperature may have fallen. They decide to test the mean using the following random sample of 12 recent temperature readings.

374.0	378.1	363.0	357.0	377.9	388.4
379.6	372.4	362.4	377.3	385.2	370.6

(i) Give three reasons why a t test would be appropriate. [3]

(ii) Carry out the test using a 5% significance level. State your hypotheses and conclusion carefully. [9]

(iii) Find a 95% confidence interval for the true mean temperature in the reaction chamber. [4]

(iv) Describe briefly one advantage and one disadvantage of having a 99% confidence interval instead of a 95% confidence interval. [2]

4 **(a)** In Germany, towards the end of the nineteenth century, a study was undertaken into the distribution of the sexes in families of various sizes. The table shows some data about the numbers of girls in 500 families, each with 5 children. It is thought that the binomial distribution B(5, p) should model these data.

Number of girls	Number of families
0	32
1	110
2	154
3	125
4	63
5	16

(i) Use this information to calculate an estimate for the mean number of girls per family of 5 children. Hence show that 0.45 can be taken as an estimate of p. [3]

(ii) Investigate at a 5% significance level whether the binomial model with p estimated as 0.45 fits the data. Comment on your findings and also on the extent to which the conditions for a binomial model are likely to be met. [12]

(b) A researcher wishes to select 50 families from the 500 in part **(a)** for further study. Suggest what sort of sample she might choose and describe how she should go about choosing it. [3]

4768 Statistics 3

Q1 (a)	$\mathrm{P}(T > t) = \frac{k}{t^2}$, $t \geq 1$,			
(i)	$\mathrm{F}(t) = \mathrm{P}(T < t) = 1 - \mathrm{P}(T > t)$ $\therefore \mathrm{F}(t) = 1 - \frac{k}{t^2}$ $\mathrm{F}(1) = 0$ $\therefore 1 - \frac{k}{1^2} = 0$ $\therefore k = 1$	M1 M1 A1	Use of 1 – P(…). Beware: answer given.	3
(ii)	$\mathrm{f}(t) = \frac{\mathrm{d\,F}(t)}{\mathrm{d}t}$ $= \frac{2}{t^3}$	M1 A1	Attempt to differentiate c's cdf. (For $t \geq 1$, but condone absence of this.) Ft c's cdf provided answer sensible.	2
(iii)	$\mu = \int_1^\infty t\mathrm{f}(t)\mathrm{d}t = \int_1^\infty \frac{2}{t^2}\mathrm{d}t$ $= \left[\frac{-2}{t}\right]_1^\infty$ $= 0 - (-2) = 2$	M1 A1 A1	Correct form of integral for the mean, with correct limits. Ft c's pdf. Correctly integrated. Ft c's pdf. Correct use of limits leading to correct value. Ft c's pdf provided answer sensible.	3
(b)	H_0: $m = 5.4$ H_1: $m \neq 5.4$ where m is the population median time for the task.	B1 B1	Both hypotheses. Hypotheses in words only must include "population". For adequate verbal definition.	

Times	− 5.4	Rank of |diff|
6.4	1.0	8
5.9	0.5	5
5.0	−0.4	4
6.2	0.8	7
6.8	1.4	10
6.0	0.6	6
5.2	−0.2	2
6.5	1.1	9
5.7	0.3	3
5.3	−0.1	1

(table above)	M1 M1 A1	for subtracting 5.4. for ranks. FT if ranks wrong.	
$W_- = 1 + 2 + 4 = 7$ (or $W_+ = 3+5+6+7+8+9+10 = 48$)	B1		
Refer to tables of Wilcoxon single sample (/paired) statistic for $n = 10$.	M1	No ft from here if wrong.	
Lower (or upper if 48 used) double-tailed 5% point is 8 (or 47 if 48 used).	A1	i.e. a 2-tail test. No ft from here if wrong.	
Result is significant.	A1	ft only c's test statistic.	
Seems that the median time is no longer as previously thought.	A1	ft only c's test statistic.	10

Q2	$X \sim \text{N}(260,\ \sigma = 24)$		When a candidate's answers suggest that (s)he appears to have neglected to use the difference columns of the Normal distribution tables penalise the first occurrence only.	
(i)	$\text{P}(X < 300) = \text{P}\left(Z < \frac{300-260}{24} = 1.6667\right)$ $= 0.9522$	M1 A1 A1	For standardising. Award once, here or elsewhere.	3
(ii)	$Y \sim \text{N}(260 \times 0.6 = 156,$ $24^2 \times 0.6^2 = 207.36$ $\text{P}(Y > 175) = \text{P}\left(Z > \frac{175-156}{14.4} = 1.3194\right)$ $= 1 - 0.9063 = 0.0937$	B1 B1 A1	Mean. Variance. Accept sd (= 14.4). c.a.o.	3
(iii)	$Y_1 + Y_2 + Y_3 + Y_4 \sim \text{N}(624,$ $829.44)$ $\text{P}(\text{this} < 600) = \text{P}\left(Z < \frac{600-624}{28.8} = -0.8333\right)$ $= 1 - 0.7976 = 0.2024$	B1 B1 A1	Mean. Ft mean of (ii). Variance. Accept sd (= 28.8). Ft variance of (ii). c.a.o.	3
(iv)	Require w such that $0.975 = \text{P}(\text{above} > w) = \text{P}\left(Z > \frac{w-624}{28.8}\right)$ $= \text{P}(Z > -1.96)$ $\therefore w - 624 = 28.8 \times -1.96 \Rightarrow w = 567.5(52)$	M1 B1 A1	Formulation of requirement. − 1.96 Ft parameters of (iii).	3
(v)	$On \sim \text{N}(150,\ \sigma = 18)$ $X_1 + X_2 + X_3 + On_1 + On_2 \sim \text{N}(1080,$ $2376)$ $\text{P}(\text{this} > 1000) = \text{P}\left(Z > \frac{1000-1080}{48.744} = -1.6412\right)$ $= 0.9496$	 B1 B1 A1	 Mean. Variance. Accept sd (= 48.744). c.a.o.	3
(vi)	Given $\bar{x} = 252.4 \quad s_{n-1} = 24.6$ CI is given by $252.4 \pm 2.576 \times \frac{24.6}{\sqrt{100}}$ $= 252.4 \pm 6.33(6) = (246.0(63), 258.7(36))$	M1 B1 A1	Correct use of 252.4 and $24.6/\sqrt{100}$. For 2.576. c.a.o. Must be expressed as an interval.	3
				18

Q3				
(i)	A t test should be used because the sample is small, the population variance is unknown, the background population is Normal	 E1 E1 E1		3
(ii)	H$_0$: $\mu = 380$ H$_1$: $\mu < 380$ where μ is the mean temperature in the chamber.	B1 B1	Both hypotheses. Hypotheses in words only must include "population". For adequate verbal definition. Allow absence of "population" if correct notation μ is used, but do NOT allow "$\bar{X} = \ldots$" or similar unless $\bar{X}$ is clearly and explicitly stated to be a population mean.	
	$\bar{x} = 373.825$ $s_{n-1} = 9.368$	B1	$s_n = 8.969$ but do NOT allow this here or in construction of test statistic, but FT from there.	
	Test statistic is $\dfrac{373.825 - 380}{\frac{9.368}{\sqrt{12}}}$	M1	Allow c's $\bar{x}$ and/or s_{n-1}. Allow alternative: 380 + (c's – 1.796) $\times \frac{9.368}{\sqrt{12}}$ (= 375.143) for subsequent comparison with $\bar{x}$. (Or $\bar{x}$ – (c's –1.796) $\times \frac{9.368}{\sqrt{12}}$ (= 378.681) for comparison with 380.)	
	= –2.283(359).	A1	c.a.o. but ft from here in any case if wrong. Use of 380 – $\bar{x}$ scores M1A0, but ft.	
	Refer to t_{11}. Single-tailed 5% point is –1.796.	M1 A1	No ft from here if wrong. Must be minus 1.796 unless absolute values are being compared. No ft from here if wrong.	
	Significant. Seems mean temperature in the chamber has fallen.	A1 A1	ft only c's test statistic. ft only c's test statistic.	9
(iii)	CI is given by 373.825 ± 2·201 $\times \frac{9.368}{\sqrt{12}}$ = 373.825 ± 5.952= (367.87(3), 379.77(7))	 M1 B1 M1 A1	c.a.o. Must be expressed as an interval. ZERO/4 if not same distribution as test. Same wrong distribution scores maximum M1B0M1A0. Recovery to t_{11} is OK.	4
(iv)	Advantage: greater certainty. Disadvantage: less precision.	E1 E1	Or equivalents.	2 18

Q4				
(a) (i)	$\bar{x} = \frac{1125}{500} = 2.25$ For binomial E(X) = $n \times p$ $\therefore \hat{p} = \frac{2.25}{5} = 0.45$	B1 M1 A1	 Use of mean of binomial distribution. May be implicit. Beware: answer given.	3
(ii)	see table below			

f_o	32	110	154	125	63	16
f_e (calc)	25.164	102.944	168.455	137.827	56.384	9.226
f_e (tables)	25.15	102.95	168.45	137.85	56.35	9.25

		M1 A1	Calculation of expected frequencies. All correct.	
	X^2 = 1.8571 + 0.4836 + 1.2404 + 1.1938 + 0.7763 + 4.9737	M1	Or using tables: 1.8657 + 0.4828 + 1.2396 + 1.1978 + 0.7848 + 4.9257	
	= 10.52(49)	A1	c.a.o. Or using tables: 10.49(64)	
	Refer to χ_4^2.	M1	Allow correct df (= cells – 2) from wrongly grouped or ungrouped table, and FT. Otherwise, no FT if wrong.	
	Upper 5% point is 9.488. Significant. Suggests binomial model does not fit.	A1 A1 A1	No ft from here if wrong. ft only c's test statistic. ft only c's test statistic.	
	The model appears to overestimate in the middle and to underestimate at the tails. The biggest discrepancy is at X = 5.	E1 E1	Accept also any other sensible comment e.g. at 2.5% significance, the result would NOT have been significant.	
	A binomial model assumes all trials are independent with a constant probability of "success". It seems unlikely that there will be independence within families and/or that p will be the same for all families.	E2	(E2, 1, 0) Any sensible comment which addresses independence and constant p.	12
(b)	She should try to choose a simple random sample which would involve establishing a sampling frame and using some form of random number generator.	E1 E1 E1	Allow sensible discussion of practical limitations of choosing a random sample. Allow other sensible suggestions. E.g systematic sample - choosing every tenth family; stratified sample - by the number of girls in a family.	3
				18

4768: Statistics 3

General Comments

Once again the general standard of many of the scripts seen was pleasing. However, the work of equally many candidates showed carelessness and a lack of thought, together with an apparent failure to read the question properly. As in the past, the quality of the comments, interpretations and explanations was patchy, and usually less good than the rest of the work.

Invariably all four questions were attempted. Marks for Questions 2 and 3 were found to be somewhat higher on average than Questions 1 and 4. There was no evidence to suggest that candidates found themselves short of time at the end.

As in the past the examiners found themselves having to cope with sloppy notation from candidates who should know better. One further general point worth making is that when the conclusion to a hypothesis test turns out to be "Accept H_1" then it is not correct to say "there is no evidence for H_0."

Comments on Individual Questions

1) **Continuous random variables; Wilcoxon single sample test; times for a computer to perform various tasks.**

(a)(i) Answers to this opening part were very disappointing. It was felt that candidates rushed into it, apparently believing that the given expression was the p.d.f. This suggested either that they had not read the question properly in the first place or that they had a poor understanding of the relationship between probability and the c.d.f. Consequently many candidates were unable to show $k = 1$ without a fudge of some kind. Difficulties here usually had implications for the next two parts too.

(ii) Many candidates were able to indicate that they expected to differentiate something but often it was not well done. Frequently the outcome was a negative p.d.f. which seemed not to cause them concern. The notation often left the examiner wondering if the candidate knew which was the p.d.f. and which the c.d.f.

(iii) The integration was generally badly set up and badly carried out. The errors seen included using the wrong limits, obtaining a negative mean (and then the minus sign would be crossed out!) and substituting $t = 0$ into an expression of the form $1/t^n$.

(b) In contrast to part (a), this part was done well and successfully by very many candidates. There was just one widespread fault: the omission of the word "population" in the hypotheses. A noticeable minority of candidates appeared not to realise that a non-parametric test was required, and even after writing hypotheses involving the median they went ahead with a test for the mean.

2) **Combinations of Normal distributions; confidence interval for a population mean; weights of leeks.**

This question was very well answered with very many scoring full marks. Candidates seemed well prepared for it and understood what was expected. In many cases their answers were concise and to the point. Those who take the trouble to provide simple sketch graphs of the standard Normal distribution do much to enhance the quality of their responses and to guard against careless errors.

(i) This part was always answered correctly.

(ii) Except for a very occasional problem with the variance, this part, too, was almost always correct.

(iii) Usually the mean total weight was correct, but often the variance was not. Typically the error came about through a lack of proper understanding of the difference between $\text{Var}(4Y)$ (= $4^2\text{Var}(Y)$) and $\text{Var}(Y_1 + \ldots + Y_4)$ (= $\text{Var}(Y_1) + \ldots + \text{Var}(Y_4)$). Here the former was used when it should have been the latter. Even good candidates wrote $\text{Var}(4Y)$ when they subsequently worked out $\text{Var}(Y_1) + \ldots + \text{Var}(Y_4)$

(iv) Correct answers were not seen here as often. Candidates seemed to experience difficulty with the formulation of the requirement of this part. In fact an explicit statement of it in symbols was conspicuously missing and this, together with choosing the upper instead of the lower percentage point, seemed to contribute to their lack of success.

(v) There were many good answers to this part; the problems that did occur were the result of errors in the variance again.

(vi) The confidence interval was often obtained correctly, although quite a few candidates selected the wrong percentage point (usually from t_{100} instead of the Normal distribution).

3) **The *t* distribution: hypothesis test for the population mean; confidence interval for a population mean; temperatures in a reaction chamber.**

(i) Although most candidates scored some marks in this part it was relatively rare to find three correct, carefully expressed reasons for carrying out a *t* test. Furthermore it was important to specify clearly whether one was referring to the population or the sample.

(ii) The hypotheses were well expressed and were usually accompanied by a carefully worded definition of the symbol "μ".
In general candidates obtained correct values for the mean and sample variance, but there were a number who were a little less than careful about the accuracy and so the test statistic suffered slightly from premature approximation. Similarly the test was carried out and concluded correctly, the most common problem being the use of the wrong critical value (-2.201 instead of -1.796). Furthermore when the test is one-tailed, requiring the lower tail critical value and involving a negative test statistic, candidates are often less than clear and careful about the negative signs.

(iii) Most candidates showed that they were familiar with how to construct a confidence interval, and did so successfully. Unsurprisingly, there were a number who seemed to forget that they should still be using the t distribution.

(iv) Answers to this part were interesting to say the least. Many candidates responded much along the lines intended, but there were quite a few whose answers were, in effect, the opposite, suggesting that they had little or no understanding of the issue. Of particular interest, though not expected or intended, were the relatively few candidates who wrote in terms equivalent to a discussion of Type I and Type II errors.

4) **Chi-squared test of goodness of fit of a binomial model; Sampling; numbers of girls in families with 5 children.**

(a)(i) This part was almost always answered correctly, the given estimate of p being obtained convincingly.

(ii) By and large the expected frequencies and the value of the test statistic were calculated correctly, although some inconsistencies in rounding results were noticed. Quite a few used the wrong number of degrees of freedom, usually because they forgot to allow for the estimated parameter or thought that there was more than one, and hence their critical value was inappropriate. Following the conclusion of the test, most simply omitted to comment on their findings. Candidates were expected to undertake a brief discussion of what can be deduced by looking at the data in order to explain the outcome of the test. Furthermore most candidates did not even attempt to address the conditions of independence of trials and constant probability of outcome that are needed for a binomial model, and those who did attempt it failed to show any real understanding.

(b) Stratified sampling was by far the most popular choice among the candidates, but whatever the choice the description of how it should be done was often left wanting. A common shortcoming was a failure to appreciate that the sample was to be taken from the original 500 families of part (a).