OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS (MEI) 4768

Statistics 3

Wednesday 19 January 2011
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet. Please write clearly and in capital letters.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or 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 document consists of **4** pages. Any blank pages are indicated.

1   Each month the amount of electricity, measured in kilowatt-hours (kWh), used by a particular household is Normally distributed with mean 406 and standard deviation 12.

    (i) Find the probability that, in a randomly chosen month, less than 420 kWh is used. **[3]**

The charge for electricity used is 14.6 pence per kWh.

    (ii) Write down the distribution of the total charge for the amount of electricity used in any one month. Hence find the probability that, in a randomly chosen month, the total charge is more than £60. **[3]**

    (iii) The household receives a bill every three months. Assume that successive months may be regarded as independent of each other.

    Find the value of b such that the probability that a randomly chosen bill is less than £b is 0.99. **[4]**

In a different household, the amount of electricity used per month was Normally distributed with mean 432 kWh. This household buys a new washing machine that is claimed to be cheaper to run than the old one. Over the next six months the amounts of electricity used, in kWh, are as follows.

<div align="center">

404   433   420   423   413   440

</div>

    (iv) Treating this as a random sample, carry out an appropriate test, with a 5% significance level, to see if there is any evidence to suggest that the amount of electricity used per month by this household has decreased on average. **[9]**

2   (a)   (i) What is stratified sampling? Why would it be used? **[4]**

      (ii) A local authority official wishes to conduct a survey of households in the borough. He decides to select a stratified sample of 2000 households using Council Tax property bands as the strata. At the time of the survey there are 79 368 households in the borough. The table shows the numbers of households in the different tax bands.

Tax band	A – B	C – D	E – F	G – H
Number of households	32 298	33 211	9739	4120

      Calculate the number of households that the official should choose from each stratum in order to obtain his sample of 2000 households so that each stratum is represented proportionally. **[2]**

  (b)   (i) What assumption needs to be made when using a Wilcoxon single sample test? **[2]**

      (ii) As part of an investigation into trends in local authority spending, one of the categories of expenditure considered was 'Highways and the Environment'. For a random sample of 10 local authorities, the percentages of their total expenditure spent on Highways and the Environment in 1999 and then in 2009 are shown in the table.

Local authority	A	B	C	D	E	F	G	H	I	J
1999	9.60	8.40	8.67	9.32	9.89	9.35	7.91	8.08	9.61	8.55
2009	8.94	8.42	7.87	8.41	10.17	10.11	8.31	9.76	9.54	9.67

      Use a Wilcoxon test, with a significance level of 10%, to determine whether there appears to be any change to the average percentage of total expenditure spent on Highways and the Environment between 1999 and 2009. **[10]**

3 The masses, in kilograms, of a random sample of 100 chickens on sale in a large supermarket were recorded as follows.

Mass (m kg)	$m < 1.6$	$1.6 \leqslant m < 1.8$	$1.8 \leqslant m < 2.0$	$2.0 \leqslant m < 2.2$	$2.2 \leqslant m < 2.4$	$2.4 \leqslant m < 2.6$	$2.6 \leqslant m$
Frequency	2	8	30	42	11	5	2

   (i) Assuming that the first and last classes are the same width as the other classes, calculate an estimate of the sample mean and show that the corresponding estimate of the sample standard deviation is 0.2227 kg. **[3]**

A Normal distribution using the mean and standard deviation found in part **(i)** is to be fitted to these data. The expected frequencies for the classes are as follows.

Mass (m kg)	$m < 1.6$	$1.6 \leqslant m < 1.8$	$1.8 \leqslant m < 2.0$	$2.0 \leqslant m < 2.2$	$2.2 \leqslant m < 2.4$	$2.4 \leqslant m < 2.6$	$2.6 \leqslant m$
Expected frequency	2.17	10.92	f	33.85	19.22	5.13	0.68

  (ii) Use the Normal distribution to find f. **[3]**

  (iii) Carry out a goodness of fit test of this Normal model using a significance level of 5%. **[9]**

  (iv) Discuss the outcome of the test with reference to the contributions to the test statistic and to the possibility of other significance levels. **[3]**

4 A timber supplier cuts wooden fence posts from felled trees. The posts are of length $(k + X)$ cm where k is a constant and X is a random variable which has probability density function

$$\text{f}(x) = \begin{cases} 1 + x & -1 \leqslant x < 0, \\ 1 - x & 0 \leqslant x \leqslant 1, \\ 0 & \text{elsewhere.} \end{cases}$$

  (i) Sketch f(x). **[3]**

  (ii) Write down the value of E(X) and find Var(X). **[5]**

  (iii) Write down, in terms of k, the approximate distribution of $\overline{L}$, the mean length of a random sample of 50 fence posts. Justify your choice of distribution. **[4]**

  (iv) In a particular sample of 50 posts, the mean length is 90.06 cm. Find a 95% confidence interval for the true mean length of the fence posts. **[4]**

  (v) Explain whether it is reasonable to suppose that $k = 90$. **[1]**

GCE

Mathematics (MEI)

Advanced GCE

Unit **4768:** Statistics 3

Mark Scheme for January 2011

Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

Q1	$E \sim N(406, 12^2)$ 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)	$P(E < 420) = P\left(Z < \dfrac{420 - 406}{12} = 1.1666\right)$ $= 0.8783/4$	M1 A1 A1	For standardising. Award once, here or elsewhere. c.a.o.	3
(ii)	$C \sim N(406 \times 14.6 = 5927.6,$ $\quad\quad \sigma^2 = 12^2 \times 14.6^2 = 30695.04)$ P(this > 6000) = $P\left(Z > \dfrac{6000 - 5927.6}{175.2} = 0.4132\right) = 1 - 0.6602 = 0.3398$	 B1 B1 A1	Accept equivalent in £. Mean. Variance. Accept sd (= 175.2). Accept P($E > 6000/14.6$) o.e. c.a.o.	3
(iii)	$B = C_1 + C_2 + C_3 \sim N(17782.8,$ $\quad \sigma^2 = 175.2^2 + 175.2^2 + 175.2^2 = 92085.12)$ Require b s.t. $P(B < 100b) = 0.99$ $\therefore \dfrac{100b - 17782.8}{303.455} = 2.326$ $\therefore 100b = 17782.8 + 2.326 \times 303.455 = 18488.6...$ (p) $b = £184.89$	 B1 B1 B1 A1	Accept equivalent in £, or $E_1 + E_2 + E_3$. Mean. ft from (ii). Variance. Accept sd (= 303.455…). ft from (ii). Accept P($E_1 + E_2 + E_3 < 100b/14.6$) o.e. 2.326 seen. c.a.o. (Minimum 4 s.f. required in final answer.)	4
(iv)	$H_0: \mu = 432$ $H_1: \mu < 432$ where μ is the mean amount of electricity used. $\bar{x} = 422.16...$ $s_{n-1} = 13.075(4)$ Test statistic is $\dfrac{422.16 - 432}{\dfrac{13.075}{\sqrt{6}}}$ $= -1.842(13).$ Refer to t_5. Single-tailed 5% point is –2.015. Not significant. Insufficient evidence to suggest that the amount of electricity used has decreased on average.	B1 B1 B1 M1 A1 M1 A1 A1 A1	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 "$\bar{X} = ...$" or similar unless $\bar{X}$ is clearly and explicitly stated to be a <u>population</u> mean. $s_n = 11.936$ but do <u>NOT</u> allow this here or in construction of test statistic, but FT from there. Allow c's $\bar{x}$ and/or s_{n-1}. Allow alternative: 432 + (c's –2.015) × $13.075/\sqrt{6}$ (= 421.24) for subsequent comparison with $\bar{x}$. (Or $\bar{x}$ – (c's –2.015) × $13.075/\sqrt{6}$ (= 432.92) for comparison with 432.) c.a.o. but ft from here in any case if wrong. Use of $\mu - \bar{x}$ scores M1A0. No ft from here if wrong. P($t < -1.842(13)$) = 0.0624. Must be minus 2.015 unless absolute values are being compared. No ft from here if wrong. ft only c's test statistic. ft only c's test statistic. Conclusion in context to include "on average" o.e.	9
				19

Q2				
(a) (i)	There are identifiable subgroups or strata that might exhibit different characteristics. Each stratum is randomly sampled. Use it to obtain a representative sample. Can get information on the individual strata.	E1 E1 E1 E1		4
(ii)	For each stratum $\ldots \times \dfrac{2000}{79368}$ giving 813.9, 836.9, 245.4, 103.8 so 814, 837, 245, 104	M1 A1	All correct.	2
(b) (i)	The <u>population</u> (or underlying distribution) is assumed to be <u>symmetrical</u> about its <u>median</u>.	E2	E2, 1, 0. Award E1 for 2 out of 3 of the key features.	2
(ii)	$H_0: m = 0$ $H_1: m \neq 0$ where m is the population median difference for the percentages.	B1 B1	Both hypotheses. Hypotheses in words only must include "population". For adequate verbal definition.	

<table>
<tr><td>Diff</td><td>−0.66</td><td>0.02</td><td>−0.80</td><td>−0.91</td><td>0.28</td><td>0.76</td><td>0.40</td><td>1.68</td><td>−0.07</td><td>1.12</td></tr>
<tr><td>Rank</td><td>5</td><td>1</td><td>7</td><td>8</td><td>3</td><td>6</td><td>4</td><td>10</td><td>2</td><td>9</td></tr>
</table>

		M1 M1 A1 B1	For differences. ZERO (out of 8) in this section if paired differences not used. For ranks. ft from here if ranks wrong. (or $W_+ = 1 + 3 + 4 + 6 + 9 + 10 = 33$)	
	$W_- = 2 + 5 + 7 + 8 = 22$ Refer to tables of Wilcoxon paired (/single sample) statistic for $n = 10$. Lower (or upper if 33 used) 5% tail is 10 (or 45 if 33 used). Result is not significant. No evidence to suggest a change in spending on average.	M1 A1 A1 A1	No ft from here if wrong. i.e. a 2-tail test. No ft from here if wrong. ft only c's test statistic. ft only c's test statistic. Conclusion in context to include "on average" o.e.	10
				18

Q3				
(i)	Using mid- intervals 1.5, 1.7, etc $\bar{x} = \dfrac{205}{100} = 2.05$ $s = \sqrt{\dfrac{425.16 - 100 \times 2.05^2}{99}} = 0.2227(01...)$	M1 A1 E1	 Mean. s.d. Answer given; must show convincingly.	3
(ii)	$f = 100 \times P(1.8 \le M < 2.0)$ $= 100 \times P(-1.1226 \le z < -0.2245)$ $= 100 \times ((1 - 0.5888) - (1 - 0.8691))$ $= 100 \times (0.4112 - 0.1309) = 28.03$	M1 A1 A1	Probability $\times$ 100. Correct Normal probabilities. ft c's mean. Must show convincingly using Normal distribution. ft c's mean.	3
(iii)	H_0: The Normal model fits the data. H_1: The Normal model does not fit the data. $X^2 = 0.7294 + 0.1384 + 1.9623 + 3.5155 + 0.2437$ $= 6.589(3)$ Refer to χ_2^2. Upper 5% point is 5.991. Significant. Evidence suggests that the model does not fit the data.	B1 B1 M1 M1 A1 M1 A1 A1 A1	Ignore any reference to parameters. Merge first 2 and last 2 cells. Calculation of X^2. c.a.o. Allow correct df (= cells – 3) from wrongly grouped table and ft. Otherwise, no ft if wrong. $P(X^2 > 6.589) = 0.0371$. No ft from here if wrong. ft only c's test statistic. ft only c's test statistic. Conclusion in context.	9
(iv)	The model • overestimates in the 2.2 – 2.4 class, • underestimates in the 2 – 2.2 class. At lower significance levels the test would not have been significant.	E1 E1 E1		3
				18

Q4				
(i)		G1 G1 G1	One (straight) line segment correct. Second (straight) line segment correct. Fully labelled intercepts + no spurious other lines.	3
(ii)	$E(X) = 0$ (By symmetry.) $E(X^2) = \int_{-1}^{0} x^2(1+x)dx + \int_{0}^{1} x^2(1-x)dx$ $= \left[\frac{x^3}{3} + \frac{x^4}{4} \right]_{-1}^{0} + \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_{0}^{1}$ $= 0 - \left(\frac{-1}{3} + \frac{1}{4} \right) + \left(\frac{1}{3} - \frac{1}{4} \right) - 0$ $= \frac{1}{6}$ $\therefore Var(X) = \frac{1}{6}(-0^2) = \frac{1}{6}$	B1 M1 M1 M1 A1	One correct integral with limits (which may be implied subsequently). Second integral correct (with limits) or allow use of symmetry. Correctly integrated and attempt to use limits. c.a.o. Condone absence of explicit evidence of use of $Var(X) = E(X^2) - E(X)^2$.	5
(iii)	$\bar{L} \sim N\left(k, \frac{1}{300} \right)$ Normal distribution because of the Central Limit Theorem.	B1 B1 B1 E1	Normal. Mean. Variance. ft c's variance in (ii) (> 0) / 50. Any reference to the CLT.	4
(iv)	CI is given by $90.06 \pm$ 1.96 $\times \frac{1}{\sqrt{300}}$ $= 90.06 \pm 0.11316 = (89.947, 90.173)$	M1 B1 M1 A1	ft c's variance in (ii) (> 0) / 50. Must be expressed as an interval.	4
(v)	It is reasonable, because 90 lies within the interval found in (iv).	E1	Or equivalent.	1
				17

Chief Examiner's Introduction to Statistics Reports

Two general matters, to which some attention is also drawn in the individual subject reports, are worthy of mention in a general introduction, as they apply to all the Statistics modules.

First, advice was circulated several months ago concerning the issue of numerical accuracy of final answers, in particular to the practice of some candidates of gross over-specification in this regard. As an example, this would refer to the quotation of the value of a test statistic as, say, 2.18735693762 merely because this is the number that happened to appear on the candidate's calculator. This shows a complete lack of understanding of statistical practice and, indeed, of basic concepts of numeracy. In the current round of examinations, accuracy marks (but not method marks) were normally withheld in such cases. The earlier advice had explicitly stated that this would occur, and it will continue in future rounds. This is of course different from the desirable practice of retaining sufficient accuracy in intermediate calculations to avoid problems resulting from premature rounding.

Secondly, there are many references in the individual subject reports to the importance of securely stating hypotheses when conducting statistical tests. In future rounds of examinations, candidates will be expected to state their null and alternative hypotheses even if this is not explicitly asked for in the question. In many cases, this can sensibly and compactly be done in the usual notation of the subject, for example "$H_0: \mu = 25$; $H_1: \mu > 25$", but it would be expected that any parameters appearing in those statements are themselves briefly but adequately defined verbally. In the example, this might be achieved by adding "where μ is the population mean". There is no objection to hypotheses being stated verbally (for example "the null hypothesis is that the population mean is 25 (cm) and the alternative hypothesis is that it is greater"), but candidates must be careful to be precise in their wording (notably, explicit use of the word "population" will often be necessary for full marks to be awarded).

4768 Statistics 3

General Comments

There were 274 candidates from 38 centres (compared with January 2010: 280 from 41) for this sitting of the paper. Although there were several very competent scripts there was much work that was quite disappointing and poorly set out. Many candidates were unable to carry out basic tasks, not at all in keeping with what one might expect at this level.

Invariably all four questions were attempted. Marks for Question 1 were found to be higher on average than the other 3 questions. Question 4 seemed rushed at the end suggesting some candidates may have found themselves short of time.

Comments on Individual Questions

1) **(i)** Although intended as a gentle start to the paper, it was surprising how many failed to score full marks here. Two faults prevailed: the standardised value 1.166…became 1.666… either when transferred from the calculator or when looked up in the Normal tables.

   (ii) There was more success here. Provided candidates were careful with the variance of the monthly charge then the correct result would usually follow. However some candidates did get caught out when trying to convert pence-squared to pounds-squared.

   (iii) There were more problems in this part when trying to sort out the variance, occasionally made worse by premature approximation.

   (iv) There were many good solutions to this part. However, it was often the case that marks were lost as a result of hypotheses that were imprecisely specified and/or conclusions that were inadequate. Many neglected to express the final conclusion non-assertively, in context and including wording such as "on average" to refer to the mean. Quite a few candidates based their test on the differences from 432 kWh, the old mean – a strategy that works but introduces additional opportunities for making mistakes.

2) **(a)(i)** The syllabus topic "Sampling methods" remains consistently and conspicuously badly understood by candidates. The definition and subsequent discussion of stratified sampling was usually vague and woolly. Few explained coherently and concisely the idea of a population that divides up into identifiable subgroups. Any reference to the strata being sampled randomly was often omitted. So also was the phrase "representative sample" as a reason for the use of this method.
   It should be noted that a stratified sample does *not* need to be selected in proportion to the sizes of the strata. There may be very good reasons for not doing so, for example if some strata are much more variable than others (it would be sensible to take more observations in the more variable strata), or if some strata are much more expensive to sample than others (it may be necessary for budgetary reasons to restrict the sampling in the more expensive strata).

   (ii) This part was well answered, but perhaps not as well as expected.

   (b)(i) Given that the Wilcoxon test usually provides one of the more successful questions from the point of view of candidates, it was disappointing to discover how few could explain the circumstances under which it would be valid to use this test. A very common wrong answer was that the data (sic) should be Normal.

(ii) In most cases the calculation of the test statistic and the identification of the critical value were correct. Sometimes errors in the ranking arose through candidates misreading their table of differences. As in Question 1 (iv), solutions were let down badly by incorrect hypotheses and/or inadequate conclusions.
Some candidates stated their hypotheses in terms of "the difference of the medians" which is not necessarily the same as "the median of the differences" and which should be discouraged.

3) (i) Many candidates were not able to find the sample mean and standard deviation from grouped data. A large number of them had little, if any, idea about how to set out the calculations.

(ii) This was also badly answered. Many faked the answer by ignoring the instruction to use the Normal distribution, choosing instead to work out "100 – the sum of the given frequencies". Furthermore, among those who did what was intended, a wrong answer was likely to be taken forward without checking its feasibility.

(iii) The Chi-squared test was not carried out with the same competence as in the past. Errors abounded, notably the failure to merge classes at one end, at least, and the incorrect identification of the number of degrees of freedom resulting in an incorrect critical value. As in Questions 2 and 3 the hypotheses and conclusions were often expressed badly. A further common fault, mentioned in previous reports, involves statements such as "the data fits/follows the model."

(iv) The discussion of the outcome of the test rarely showed more than a superficial appreciation of what was going on. It was not uncommon for the wrong class to be identified as providing the largest contribution to the test statistic.

4) (i) The majority of sketches were considered to be adequate. One would like to think that students at this level could be relied on to label their axes "x" and "y".

(ii) Only a minority of candidates used the symmetry argument to "write down" the mean. As often as not, those who integrated got it wrong, largely because they could not apply the limits 0 and –1 correctly in the integral for the left hand portion. The same problem with limits occurred with the variance, which frequently turned out to be 0, or even, on occasion, negative.

(iii) Almost all named the required distribution correctly as Normal, and most gave the correct mean. Fewer candidates than usual could write down the correct variance and hardly any appeared to be aware of the Central Limit Theorem as the justification.

(iv) Most of the time there was evidence of some understanding of how to construct a confidence interval using the sample mean and the correct Normal percentage point. What was worrying was that many candidates seemed unable to make the connection between $\text{Var}(\overline{L})$ in the previous part and the standard deviation needed here: either a spurious extra $1/\sqrt{50}$ was introduced or they ignored $\text{Var}(\overline{L})$ completely.

(v) Most candidates appreciated that an appropriate response to this part depended on whether or not 90 was contained in the interval in part (iv).

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

14 – 19 Qualifications (General)
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance
programme your call may be recorded or monitored