

**ADVANCED GCE
MATHEMATICS (MEI)**

Statistics 4

4769

Candidates answer on the answer booklet.

OCR supplied materials:

- 8 page answer booklet (sent with general stationery)
- MEI Examination Formulae and Tables (MF2)

Other materials required:

- Scientific or graphical calculator

**Thursday 26 May 2011
Morning**

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 any **three** 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.

Option 1: Estimation

- 1** The random variable X has the Normal distribution with mean 0 and variance θ , so that its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta}, \quad -\infty < x < \infty,$$

where θ ($\theta > 0$) is unknown. A random sample of n observations from X is denoted by X_1, X_2, \dots, X_n .

- (i) Find $\hat{\theta}$, the maximum likelihood estimator of θ . [14]
- (ii) Show that $\hat{\theta}$ is an unbiased estimator of θ . [4]
- (iii) In large samples, the variance of $\hat{\theta}$ may be estimated by $\frac{2\hat{\theta}^2}{n}$. Use this and the results of parts (i) and (ii) to find an approximate 95% confidence interval for θ in the case when $n = 100$ and $\sum X_i^2 = 1000$. [6]

Option 2: Generating Functions

- 2** The random variable X has the χ_n^2 distribution. This distribution has moment generating function $M(\theta) = (1 - 2\theta)^{-\frac{1}{2}n}$, where $\theta < \frac{1}{2}$.

- (i) Verify the expression for $M(\theta)$ quoted above for the cases $n = 2$ and $n = 4$, given that the probability density functions of X in these cases are as follows. [10]

$$n = 2: \quad f(x) = \frac{1}{2}e^{-\frac{1}{2}x} \quad (x > 0)$$

$$n = 4: \quad f(x) = \frac{1}{4}xe^{-\frac{1}{2}x} \quad (x > 0)$$

- (ii) For the general case, use $M(\theta)$ to find the mean and variance of X in terms of n . [7]
- (iii) Y_1, Y_2, \dots, Y_k are independent random variables, each with the χ_1^2 distribution. Show that $W = \sum_{i=1}^k Y_i$ has the χ_k^2 distribution. [4]

- (iv) Use the Central Limit Theorem to find an approximation for $P(W < 118.5)$ for the case $k = 100$. [3]

Option 3: Inference

- 3 (i) Explain the meaning of the following terms in the context of hypothesis testing: Type I error, Type II error, operating characteristic, power. [8]

- (ii) A market research organisation is designing a sample survey to investigate whether expenditure on everyday food items has increased in 2011 compared with 2010. For one of the populations being studied, the random variable X is used to model weekly expenditure, in £, on these items in 2011, where X is Normally distributed with mean μ and variance σ^2 . As the corresponding mean value in 2010 was 94, the hypotheses to be examined are

$$H_0: \mu = 94,$$

$$H_1: \mu > 94.$$

By comparison with the corresponding 2010 value, σ^2 is assumed to be 25.

The following criteria for the survey are laid down.

- If in fact $\mu = 94$, the probability of concluding that $\mu > 94$ must be only 2%
- If in fact $\mu = 97$, the probability of concluding that $\mu > 94$ must be 95%

A random sample of size n is to be taken and the usual Normal test based on \bar{X} is to be used, with a critical value of c such that H_0 is rejected if the value of \bar{X} exceeds c . Find c and the smallest value of n that is required. [13]

- (iii) Sketch the power function of an ideal test for examining the hypotheses in part (ii). [3]

Option 4: Design and Analysis of Experiments

- 4 (a) Provide an example of an experimental situation where there is one factor of primary interest and where a suitable experimental design would be

(i) randomised blocks,

(ii) a Latin square.

In each case, explain carefully why the design is suitable and why the other design would not be appropriate. [12]

- (b) An industrial experiment to compare four treatments for increasing the tensile strength of steel is carried out according to a completely randomised design. For various reasons, it is not possible to use the same number of replicates for each treatment. The increases, in a suitable unit of tensile strength, are as follows.

Treatment A	Treatment B	Treatment C	Treatment D
10.1	21.1	9.2	22.6
21.2	20.3	8.8	17.4
11.6	16.0	15.2	23.1
13.6		15.0	19.2
		12.4	

[The sum of these data items is 256.8 and the sum of their squares is 4471.92.]

Construct the usual one-way analysis of variance table. Carry out the appropriate test, using a 5% significance level. [12]

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE



Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.