

6688/01

Edexcel GCE

Statistics

Unit S6

Advanced Subsidiary / Advanced

Time: 1 hour 30 minutes

Materials required for the examination

Answer Book (AB04)
Graph Paper (GP02)
Mathematical Formulae

Items included with these question papers

Nil

Candidates may use any calculator EXCEPT those with a facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as Texas TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

Full marks may be obtained for answers to ALL questions.

In the boxes on the Answer Book provided, write the name of the Examining Body (Edexcel), your Centre Number, Candidate Number, the Unit Title (Statistics S6), the Paper Reference (6688), your surname, other names and signature.

Information for Candidates

A booklet 'Mathematical Formulae including Statistical Formulae and Tables' is provided.

Values from the Statistical Tables should be quoted in full. The answer to each part of a question which requires the use of tables or a calculator should be given to three significant figures, unless otherwise specified.

This paper has 7 questions.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly numbered.

You must show sufficient working to make your methods clear to the Examiner.
Answers without working will gain no credit

1. It is suggested by a motoring organisation that the median number of miles travelled per week by motorists is 250. A random sample of 20 motorists gave the following number of miles travelled in a particular week.

248 257 282 230 262 270 240 249 290 302
259 200 291 287 400 150 261 286 308 192

Stating clearly your hypotheses, use a sign test to assess whether or not there is evidence that the median number of miles has increased. **(6 marks)**

2. Some athletes have complained that the samples they have given for drug testing have not been analysed correctly and that they have been accused of having unduly high levels of certain substances when in fact that is not the case. The athletes' federation investigated the accuracy of the laboratory which tested the sample. Eight samples, each containing a known level of exactly 0.500 mg l^{-1} of a banned substance were sent to the laboratory to assess the level of the substance. The results are given, in mg l^{-1} , below.

0.485, 0.518, 0.460, 0.530, 0.560, 0.550, 0.490, 0.575

- (a) Test the athletes' claim, at the 5% level of significance, using the Wilcoxon signed-ranks test. State your hypotheses clearly. **(6 marks)**
- (b) Find the largest number of observations less than 0.500 mg l^{-1} that could lead to a significant result using the Wilcoxon test with just 8 observations and give an example to illustrate this. **(2 marks)**

2. In an attempt to reduce the number of errors made by machine operators some extra training was provided. When some of the operators had been trained, 6 trained and 13 untrained operators were given a simple task to complete and the number of errors each made was recorded.

| | |
|---------------------|---|
| Trained operators | 11, 9, 4, 7, 6, 2 |
| Untrained operators | 3, 17, 12, 13, 21, 29, 5, 1, 15, 19, 16, 14, 10 |

- (a) Stating your hypotheses clearly, use a suitable non-parametric method to test, at the 5% level of significance, whether or not the training is effective. **(8 marks)**
- (b) State any assumption you have made about the selection of the operators. **(1 mark)**

4. In a particular A level examination, the marks obtained by seven randomly chosen candidates from each of three schools were as follows.

| | | | | | | | | Sum of marks |
|-------------------|----|----|----|----|----|----|----|--------------|
| Abbotswood School | 45 | 36 | 72 | 13 | 65 | 34 | 56 | 321 |
| Brunswick School | 34 | 45 | 12 | 43 | 18 | 23 | 34 | 209 |
| Clarey School | 64 | 23 | 32 | 17 | 45 | 34 | 13 | 228 |

Test whether or not there are differences between the mean marks of the candidates from these three schools, using a 5% level of significance. You should state your hypotheses clearly. You may assume that the total sum of squares is 6301.81. **(9 marks)**



5. A machine is set to fill capsules with a particular drug. The distribution of the weights of the fillings is specified to be the normal distribution with mean 12.00 g and standard deviation 0.35 g.
- (a) Using graph paper and assuming that samples of size 10 will be taken, draw a control chart for the sample mean indicating 98% warning limits and 99% action limits. Give your limits to 2 decimal places. **(8 marks)**
- (b) On three separate occasions when the chart was in use, the following values of the mean of samples of size 10 were recorded:
- (i) 12.27,
(ii) 11.68,
(iii) 12.17.

In each case plot the point on your chart and in each case comment on the state of the filling process. **(4 marks)**

6. A process is carried out in a chemical plant such that the amount x of the chemical added to a mixture is varied and the concentration y of the final product is recorded. The results are shown in the table below.

| | | | | | | | | |
|---------------|-----|-----|-----|-----|-----|-----|-----|-----|
| $x(\text{g})$ | 10 | 12 | 14 | 16 | 18 | 20 | 25 | 30 |
| $y(\%)$ | 2.7 | 3.1 | 4.1 | 4.5 | 5.7 | 6.2 | 7.7 | 9.9 |

You may use $\sum x = 145$, $\sum x^2 = 2945$, $\sum y = 43.9$, $\sum y^2 = 282.19$, $\sum xy = 909.7$

- (a) Calculate S_{xx} , S_{yy} , S_{xy} . **(3 marks)**
- (b) Find the equation of the regression line of y on x in the form $y = \hat{\alpha} + \hat{\beta}x$. **(3 marks)**
- (c) Calculate a 95% confidence interval for β . **(7 marks)**
- (d) Give an interpretation of your interval. **(2 marks)**
-

7. Four different plant varieties, A , B , C and D are included in an experiment on the growth of lettuce. The experiment is laid out as shown below and the same number of plants is harvested from each of three plots. The weights, W kg, are summarised below.

| Block I | | Block II | | Block III | |
|---------|-----|----------|-----|-----------|-----|
| A | 2.9 | C | 2.8 | B | 2.6 |
| B | 3.0 | B | 3.4 | C | 3.1 |
| C | 3.1 | A | 2.6 | D | 3.4 |
| D | 3.4 | D | 3.2 | A | 2.5 |

- (a) Write down the name of the experimental design used. **(1 mark)**
- (b) Using the coding $V = 10(W - 3.0)$, analyse these data and test whether or not variety appears to affect the yield. Use a 5% level of significance and state your hypotheses clearly. **(15 marks)**

END

BLANK PAGE

Key to Marking Principles

The total number of marks for each paper is 75.

Method marks (M) are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.

Accuracy marks (A) can only be awarded if the relevant method marks (M) have been earned.

(B) marks are independent of method marks.

Method marks cannot be subdivided.

For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.

All marks are 'correct answer only' unless shown, for example, as $A1\checkmark$ to indicate that previous wrong working is to be followed through. After a misread the subsequent A marks affected are treated as $A\checkmark$ but manifestly absurd answers should never be awarded A marks.

The abbreviation awrt stands for 'anything which rounds to'.

The abbreviation -1ecoo stands for 'deduct one mark for every error or omission'.

The abbreviation cso stands for 'correct solution only'.

