

Paper Reference(s)

**6688**

# **Edexcel GCE**

## **Statistics S6**

### **Advanced/Advanced Subsidiary**

**Wednesday 25 May 2005 – Afternoon**

**Time: 1 hour 30 minutes**

**Materials required for examination**

Answer Book (AB16)

Graph Paper (ASG2)

Mathematical Formulae (Lilac)

**Items included with question papers**

Nil

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

#### **Instructions to Candidates**

---

In the boxes on the answer book, 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 name and signature.

Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

#### **Information for Candidates**

---

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

Full marks may be obtained for answers to ALL questions.

This paper has seven questions. The total for this question paper is 75.

#### **Advice to Candidates**

---

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

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

N21149A

This publication may only be reproduced in accordance with Edexcel Limited copyright policy.  
©2005 Edexcel Limited

1. Students claim that they work better when listening to music than when working in silence. To test this claim 10 students were selected at random. Each student was given 20 arithmetic calculations to carry out in silence and 20 similar calculations to carry whilst listening to music. The number of correct answers was recorded in each case and the results are summarised in the table.

Student	A	B	C	D	E	F	G	H	I	J
Silence	16	14	19	17	15	14	18	20	17	18
Music	18	15	17	19	16	16	19	18	20	19

Use a sign test, at the 5% significance level, to determine whether or not there is evidence that students carry out calculations more accurately when listening to music.

**(Total 6 marks)**

---

2. A training agency claims that the median number of delegates attending its English courses is 20. From past records the number of delegates who attended a random sample of 9 courses are given below.

30 12 15 21 27 17 16 14 22

Using a Wilcoxon signed-ranks test at the 5% level of significance level, test whether or not the training agency is justified in making its claim.

**(Total 7 marks)**

---

3. As part of her project, Jane wanted to determine whether or not girls and boys have different linguistic skills. She did this by measuring the time taken by 25 boys and 25 girls to translate a given passage of French prose into English. She analysed her data using a Wilcoxon rank sum test and obtained a T value of 704.

Stating your hypothesis clearly and using a 5% level of significance, test whether or not there is evidence of a difference in the times taken by boys and girls to translate French.

**(Total 7 marks)**

---

4. An experiment was carried out to establish whether or not there were any differences in yields of wheat between 4 different areas of the country. The yield of wheat,  $x$  tones/hectare, was recorded for each of 6 randomly selected farms in each of the 4 areas.

The total yields for each area were

48.13 43.13 46.50 53.65.

Given that  $\sum x^2 = 1543.9043$ , analyse the data to determine whether or not there are any differences in the mean yields of wheat between these areas.

**(Total 11 marks)**

---

5. A company produces sweatshirts and on completion each one is checked for faults. To monitor the quality of the production, the company takes random samples of size 50 at regular intervals and records the number of sweatshirts with faults. At a time when production was assumed to be satisfactory the following data were obtained.

Sample	1	2	3	4	5	6	7	8	9	10
Number of faults	4	2	3	1	6	3	2	4	2	5

- (a) Estimate the proportion of sweatshirts with faults. (1)
- (b) Draw a fraction defective control chart using a 95% upper warning limit and a 99% upper action limit. (7)
- (c) Explain why lower limits are not needed. (2)
- (d) Plot the data given in the table on your control chart. (2)
- (e) Comment on the state of production in the company. (1)

**(Total 13 marks)**

6. Electronics components are manufactured using three machine operators and four machines that produce similar items. An experiment was carried out to assess whether or not variation in the quality of components produced was caused by differences between machines or differences between operators. The quality of components produced during the experiment was assessed by a quality control inspector. The results, based on a scale of 0 to 100, are shown in the table below.

	Machine				
Operator	A	B	C	D	Totals
1	55	90	91	66	302
2	63	81	93	60	297
3	61	79	87	60	287
Totals	179	250	271	186	886

- (a) Write down the name of the experimental design that has been used to obtain these data. (1)

An analysis of the data found the total sum of squares to be 2255.67.

- (b) Carry out an analysis of variance and test whether or not there are differences between  
 (i) machines, (ii) operators. (14)

**(Total 15 marks)**

7. An engineer measured the abrasion loss,  $y$  mg, and the hardness,  $x$  kg mm<sup>-2</sup>, of 20 erosion shields used in steam turbines. The results are summarized below.

$$\sum x = 14061, \quad \sum y = 9297, \quad \sum xy = 6493576, \quad S_{xx} = 13998.95, \quad S_{yy} = 164592.55$$

It is assumed that a suitable model relating  $x$  to  $y$  is given by

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

where  $\alpha$  and  $\beta$  are constants, the  $\varepsilon_i$  are residual errors and  $n$  is the number of pairs of observations.

- (a) Calculate an estimate for  $\alpha$  and an estimate for  $\beta$ , giving your values to 2 decimal places. (5)

- (b) Evaluate a 95% confidence interval for  $\beta$ . (6)

The engineer had assumed that the value of the regression coefficient would be  $-3.00$ .

- (c) Comment on this assumption. (2)

- (d) Explain what assumption is made about  $\varepsilon_i$  and how you would use the residuals to verify that the above model is a valid one. (3)

**(Total 16 marks)**

---

**TOTAL FOR PAPER: 75 MARKS**

**END**