

6688/01

Edexcel GCE

Statistics

Unit S6 Mock paper

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

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.

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 including Statistical Formulae and Tables' is provided.

Full marks may be obtained for answers to ALL questions.

This paper has 7 questions. Pages 6, 7 and 8 are blank.

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.

1. A quality assurance manager wishes to set up a control chart to monitor the variability of a process. At a time when the process was operating satisfactorily the manager obtained an estimate of the standard deviation of the process of 6.5mm. It is intended to use samples of size 5 to monitor the process.

CONTROL CHARTS VARIABILITY

For range charts multiply σ by the appropriate value of D.

For standard deviation charts multiply σ by the appropriate value of E.

To obtain an estimate of σ multiply the mean range by the appropriate value of b.

Normal distribution is assumed.

Sample size	D 0.999	D 0.975	D 0.025	D 0.001	E 0.999	E 0.975	E 0.025	E 0.001	b
2			3.170	4.654			2.24	3.29	0.8862
3	0.060	0.303	3.682	5.063	0.03	0.16	1.92	2.63	0.5908
4	0.199	0.595	3.984	5.309	0.09	0.27	1.76	2.33	0.4857
5	0.367	0.850	4.197	5.484	0.15	0.35	1.67	2.15	0.4299
6	0.535	1.066	4.361	5.619	0.20	0.41	1.60	2.03	0.3946
7	0.691	1.251	4.494	5.730	0.25	0.45	1.55	1.93	0.3698
8	0.835	1.410	4.605	5.823	0.29	0.49	1.51	1.86	0.3512
10	1.085	1.674	4.784	5.973	0.36	0.55	1.45	1.76	0.3249
12	1.293	1.884	4.925	6.096	0.41	0.59	1.41	1.69	0.3069

(a) Using the table provided above, evaluate warning and action limits for use on a range chart.

(3 marks)

(b) Explain how you would use a range chart in conjunction with a control chart for the mean to decide whether or not the process was operating satisfactorily.

(1 mark)

2. In a study of the teaching of a particular topic in a college, one group of students was taught by the traditional method and another group was taught by a new method. At the end of the term both groups were given the same test which was marked out of 25. The results for a random sample of students from each group were as follows.

Traditional method	14	9	17	15	13	12	8	6
New method	18	16	11	19	20	22		

Use a suitable non-parametric test, at the 5% level of significance, to test whether or not there is any difference between the marks obtained by the students being taught by these two methods. State your hypotheses clearly.

(7 marks)

3. Random samples of 100 components were taken at regular intervals from a production line that was considered to be working satisfactorily. For each sample the number of defective components was counted and the results from 8 samples are given below.

6 3 2 7 5 6 4 11

- (a) Use these data to calculate a 95% upper warning limit and a 99.8% upper action limit. **(5 marks)**
- (b) Explain why the lower warning and action limits were not needed. **(1 mark)**

The production manager suggested that the production line might not be working as satisfactorily as was first thought.

- (c) Without drawing the control chart, use the above data to comment on the manager's suggestion. **(1 mark)**

The manager then took a further 3 random samples of size 100 from the production line and found that all three had proportion defectives between the upper warning limit and upper action limit.

- (d) Find the probability of this happening. **(2 marks)**

4. A dietician wished to investigate whether or not there were any differences between three types of diet when used on men who were considered to be overweight. A group of 19 men was selected and each man was allocated at random to one of the diets. The percentage of body weight (to the nearest percent) lost by each of the men after being on the diet for several weeks is shown below.

Diet 1	Diet 2	Diet 3
5	1	1
7	5	-2
8	-2	3
-3	6	-1
1	4	0
	-1	2
	5	
	3	

- (a) Explain what is indicated by a negative value. **(1 mark)**
- (b) Test whether or not there are any differences between the mean percentage weight loss of the men using these three diets. State your hypotheses clearly and use a 5% level of significance. [You may assume that the Total Sum of Squares is 191.158.] **(9 marks)**

5. For patients suffering from a particular illness it is recommended that they exercise for 90 minutes each week. A random sample of 10 such patients recorded the number of minutes, to

the nearest minute, spent on exercise during a particular week. The results are shown below.

98 92 85 114 83 101 100 96 110 90

(a) Stating clearly your hypotheses, carry out

(i) a sign test,

(ii) a Wilcoxon signed-ranks test

to assess whether or not these patients were exercising for more than the recommended number of minutes. In both cases use a 5% level of significance. **(12 marks)**

(b) Comment on your results. **(1 mark)**

6. A landscape gardener tested three different lawn fertilisers F1, F2 and F3. The fertilisers were each used on lawns using turf from the same supplier, but the lawns were situated in different locations within a city. The growth rate of the grass in each lawn (in coded units) was recorded and summarised in the table below.

		Fertiliser			Total
		F1	F2	F3	
Location	1	3.1	3.0	2.7	8.8
	2	2.6	2.4	2.5	7.5
	3	2.9	2.1	2.3	7.3
	4	3.5	3.4	3.2	10.1
Total		12.1	10.9	10.7	33.7

$$(3.1^2 + 3.0^2 + \dots + 3.4^2 + 3.2^2 = 96.83)$$

(a) Write down the name of the experimental design that would have been used to obtain these results. **(1 mark)**

An analysis of the data found the Total Sum of Squares to be 2.189 and the Sum of Squares between Locations to be 1.689.

(b) Complete the analysis of variance and test whether or not there are differences between

(i) fertilisers,

(ii) locations.

State your hypotheses clearly and use a 5% level of significance. **(13 marks)**

(c) Comment on your results. **(1 mark)**

7. ~~The results of an experiment to investigate the amount of chemical compound, y grams, that dissolved in 100 grams of water at x °C are recorded below.~~

x	5	10	15	20	25	30	35	40	45	50	55	60
-----	---	----	----	----	----	----	----	----	----	----	----	----

y	9	11	13	17	21	24	27	29	31	35	38	42
-----	---	----	----	----	----	----	----	----	----	----	----	----

$$(\Sigma x^2 = 16\,250, \Sigma y^2 = 8\,641, \Sigma xy = 11\,795)$$

- (a) Calculate S_{xx} , S_{yy} and S_{xy} . **(3 marks)**
- (b) Find the equation of the regression line of y on x in the form $y = a + bx$. **(3 marks)**
- (c) Test, at the 5% level of significance, whether or not there is evidence that the gradient of the regression line is positive. **(8 marks)**
- (d) Explain how residuals can be used to decide on the suitability of a regression model. **(3 marks)**

END