Paper Reference(s)

6688

Edexcel GCE

Statistics S6

Advanced/Advanced Subsidiary

Friday 28 May 2004 – Afternoon

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Answer Book (AB16)	Nil
Graph Paper (ASG2)	
Mathematical Formulae (Lilac)	

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.

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.

N17576A

1. (*a*) Explain what is meant by a randomised block design.

(2)

A randomised block design has 4 treatments and 5 blocks.

(*b*) Find the number of degrees of freedom associated with the residual sum of squares.

(1)

(*c*) Write down the critical value of F that would be used in testing the null hypothesis that there is no difference between the treatments at the 5% significant level.

(2)

---

2. As a result of an experiment a student collected 10 pairs of (x, y) observations and calculated the equation of the regression line to be

$$y = 0.543 + 0.631x.$$

Test, at the 5% significance level, whether or not the population regression coefficient is greater then 0.6. You may assume that $S_{xx} = 2.4137$ and that the residual sum of squares is 0.145.

(6)

---

3. A nutritional expert randomly allocated a sample of 15 female swimmers of comparably ability into 3 groups. Those within group A were given a vitamin supplement, those within group B were given a diet of healthy foods and those in group C were instructed to eat their usual diet. Subsequently the nutritional expert recorded the time taken, in seconds, for each swimmer to swim a length of a pool using the same stoke. The results are as follows:

A	B	C
29	27	26
32	26	24
27	30	23
34	33	27
31	28	22

Test whether or not diet has an effect on the performance of female swimmers.

(9)

---

4. It is believed that athletes run faster on an indoor track than on an outdoor track. To test this belief, the times taken to run 200 m on both an indoor and outdoor track were recorded for a random sample of 25 athletes. Analysis of the results using the Wilcoxon signed-ranks test gave $S = 106$.

Test, at the 5% significance level, whether or not there is evidence that athletes run faster on an indoor track.

(11)

---

5. A machine is set to dispense crisps into packets containing a nominal 50 g of crisps. The machine is known to deliver this weight of crisps with a standard deviation of 2.4 g.

(*a*) State the target value of this process.

(1)

(*b*) Explain why upper and lower action and warning limits should be used to monitor this process.

(2)

(*c*) Calculate the 95% warning limits and the 99% action limits for the mean based on a sample of size 10.

(4)

(*d*) Draw a control chart for the mean.

(2)

At a later date, four random samples each of 10 packets of crisps were selected. The mean weight, in g, of each sample was recorded. The results, in sequence, were as follows.

48.4, 49.6, 50.9, 53.2

(*e*) Plot these means on the chart. Comment on the state of the process.

(2)

(*f*) Explain why it is advisable to plot the control chart for the standard deviation as well as the control chart for the mean.

(1)

---

6. It is suggested by a car manufacturer that the median fuel consumption of their cars with 2 litre engines is 30 miles per gallon (mpg). A random sample of 10 cars gave the following fuel consumptions.

Car	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Fuel comsumption (mpg)	31.97	28.18	30.25	31.90	29.51	33.03	30.01	31.48	33.02	30.89

(*a*) Use a Sign Test to assess whether or not there is evidence that the median fuel consumption is 30 mpg.

(6)

It is claimed that a modification to the engine has improved the fuel consumption of the car. In order to test this claim, a second random sample of 5 modified cars was taken and the fuel consumptions, in mpg, were as follows.

35.47, 34.67, 33.48, 33.01, 29.80

(*b*) Use an appropriate non-parametric method to test, at the 5% significance level, whether or not the modification to the engine has improved the fuel consumption.

(7)

7. A pupil carried out an experiment to see if there was a relationship between the length of a pendulum, x cm and the time taken, t seconds, to complete ten swings.
The results are recorded below.

x	10	20	30	40	50	60	70	80	90
t	6.3	8.6	10.9	11.8	17.1	14.7	16.3	17.7	18.5

The pupil thought that a model relating t and x might be of the form

$$t = \alpha + \beta x + \varepsilon.$$

(*a*) Calculate an estimate for α and an estimate for β.

(3)

(You may assume $S_{xx} = 6000$, $S_{tt} = 150.36$, $S_{tx} = 898$, $\bar{x} = 50$, $\bar{t} = 13.544$.)

(*b*) Find a 95% confidence interval for the regression coefficient β.

(8)

(*c*) Find the value of r and the value of s in the following table of residuals.

x	10	20	30	40	50	60	70	80	90
$t - \hat{t}$	–1.26	–0.45	r	–0.25	3.56	–0.34	s	–0.33	–1.03

(3)

(*d*) Draw on graph paper a residual plot.

(3)

(*e*) Use your residual plot to decide whether or not the regression model is acceptable.

(2)

END