Paper Reference(s)

6685

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Tuesday 28 May 2002 – 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 S3), the paper reference (6685), 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. 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 hotel has 160 rooms of which 20 are classified as De-luxe, 40 Premier and 100 as Standard. The manager wants to obtain information about room usage in the hotel by taking a 10% sample of the rooms.

(*a*) Suggest a suitable sampling method.

(1)

(*b*) Explain in detail how the manager should obtain the sample.

(4)

2. A random sample of 100 classical CDs produced by a record company had a mean playing time of 70.6 minutes and a standard deviation of 9.1 minutes. An independent random sample of 120 CDs produced by a different company had a mean playing time of 67.2 minutes with a standard deviation of 8.4 minutes.

(*a*) Using a 1% level of significance, test whether or not there is a difference in the mean playing times of the CDs produced by these two companies. State your hypotheses clearly.

(8)

(*b*) State an assumption you made in carrying out the test in part (*a*).

(1)

3. The weights of a group of males are normally distributed with mean 80 kg and standard deviation 2.6 kg. A random sample of 10 of these males is selected.

(*a*) Write down the distribution of $\overline{M}$, the mean weight, in kg, of this sample.

(2)

(*b*) Find P($\overline{M} < 78.5$).

(3)

The weights of a group of females are normally distributed with mean 59 kg and standard deviation 1.9 kg. A random sample of 6 of the males and 4 of the females enters a lift that can carry a maximum load of 730 kg.

(*c*) Find the probability that the maximum load will be exceeded when these 10 people enter the lift.

(5)

4. At the end of a season an athletics coach graded a random sample of ten athletes according to their performances throughout the season and their dedication to training. The results, expressed as percentages, are shown in the table below.

Athlete	Performance	Dedication
A	86	72
B	60	69
C	78	59
D	56	68
E	80	80
F	66	84
G	31	65
H	59	55
I	73	79
J	49	53

(*a*) Calculate the Spearman rank correlation coefficient between performance and dedication.

(5)

(*b*) Stating clearly your hypotheses and using a 10% level of significance, interpret your rank correlation coefficient.

(5)

(*c*) Give a reason to support the use of the rank correlation coefficient rather than the product moment correlation coefficient with these data.

(1)

5. The manager of a leisure centre collected data on the usage of the facilities in the centre by its members. A random sample from her records is summarised below.

Facility	Male	Female
Pool	40	68
Jacuzzi	26	33
Gym	52	31

Making your method clear, test whether or not there is any evidence of an association between gender and use of the club facilities. State your hypotheses clearly and use a 5% level of significance.

(11)

6. Data were collected on the number of female puppies born in 200 litters of size 8. It was decided to test whether or not a binomial model with parameters $n = 8$ and $p = 0.5$ is a suitable model for these data. The following table shows the observed frequencies and the expected frequencies, to 2 decimal places, obtained in order to carry out this test.

Number of females	Observed number of litters	Expected number of litters
0	1	0.78
1	9	6.25
2	27	21.88
3	46	R
4	49	S
5	35	T
6	26	21.88
7	5	6.25
8	2	0.78

(*a*) Find the values of *R*, *S* and *T*.

(4)

(*b*) Carry out the test to determine whether or not this binomial model is a suitable one. State your hypotheses clearly and use a 5% level of significance.

(7)

An alternative test might have involved estimating *p* rather than assuming $p = 0.5$.

(*c*) Explain how this would have affected the test.

(1)

7. The weights of tubs of margarine are known to be normally distributed. A random sample of 10 tubs of margarine were weighed, to the nearest gram, and the results were as follows.

498 502 500 496 509 504 511 497 506 499

(*a*) Find unbiased estimates of the mean and the variance of the population from which this sample was taken.

(5)

Given that the population standard deviation is 5.0 g,

(*b*) estimate limits, to 2 decimal places, between which 90% of the weights of the tubs lie,

(2)

(*c*) find a 95% confidence interval for the mean weight of the tubs.

(5)

A second random sample of 15 tubs was found to have a mean weight of 501.9 g.

(*d*) Stating your hypotheses clearly and using a 1% level of significance, test whether or not the mean weight of these tubs is greater than 500 g.

(5)

END