

Paper Reference(s)

6670

Edexcel GCE

Statistics S3

Advanced Level

Specimen Paper

Time: 1 hour 30 minutes

Materials required for examination papers

Answer Book (AB16)
Mathematical Formulae (Lilac)
Graph Paper (ASG2)

Items included with question

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 S3), the paper reference (6670), your surname, other name and signature.

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 six 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.

1. The 240 members of a bowling club are listed alphabetically in the club's membership book. The committee wishes to select a sample of 30 members to fill in a questionnaire about the facilities the club offers.

(a) Explain how the committee could use a table of random numbers to take a systematic sample. (3)

(b) Give one advantage of this method over taking a simple random sample. (1)

2. The weights of pears, P grams, are normally distributed with a mean of 110 and a standard deviation of 8. Geoff buys a bag of 16 pears.

(a) Write down the distribution of \bar{P} , the mean weight of the 16 pears. (2)

(b) Find $P(110 < \bar{P} < 113)$. (3)

3. The three tasks most frequently carried out in a garage are A , B and C . For each of the tasks the times, in minutes, taken by the garage mechanics are assumed to be normally distributed with means and standard deviations given in the following table.

Task	Mean	Standard deviation
A	225	38
B	165	23
C	185	27

Assuming that the times for the three tasks are independent, calculate the probability that

(a) the total time taken by a single randomly chosen mechanic to carry out all three tasks lies between 533 and 655 minutes, (5)

(b) a randomly chosen mechanic takes longer to carry out task B than task C . (5)

4. At the end of a season a league of eight ice hockey clubs produced the following table showing the position of each club in the league and the average attendances (in hundreds) at home matches.

Club	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>
Position	1	2	3	4	5	6	7	8
Average	37	38	19	27	34	26	22	32

- (a) Calculate the Spearman rank correlation coefficient between position in the league and average home attendance. (5)
- (b) Stating clearly your hypotheses and using a 5% two-tailed test, interpret your rank correlation coefficient. (4)

Many sets of data include tied ranks.

- (c) Explain briefly how tied ranks can be dealt with. (2)
-

5. For a six-sided die it is assumed that each of the sides has an equal chance of landing uppermost when the die is rolled.

- (a) Write down the probability function for the random variable X , the number showing on the uppermost side after the die has been rolled. (2)
- (b) State the name of the distribution. (1)

A student wishing to check the above assumption rolled the die 300 times and for the sides 1 to 6, obtained the frequencies 41, 49, 52, 58, 37 and 63 respectively.

- (c) Analyse these data and comment on whether or not the assumption is valid for this die. Use a 5% level of significance and state your hypotheses clearly. (8)
-

6. A sociologist was studying the smoking habits of adults. A random sample of 300 adult smokers from a low income group and an independent random sample of 400 adult smokers from a high income group were asked what their weekly expenditure on tobacco was. The results are summarised below.

	<i>N</i>	mean	s.d.
Low income group	300	£6.40	£6.69
High income group	400	£7.42	£8.13

- (a) Using a 5% significance level, test whether or not the two groups differ in the mean amounts spent on tobacco. (9)

- (b) Explain briefly the importance of the central limit theorem in this example. (2)
-

7. A survey in a college was commissioned to investigate whether or not there was any association between gender and passing a driving test. A group of 50 male and 50 female students were asked whether they passed or failed their driving test at the first attempt. All the students asked had taken the test. The results were as follows.

	Pass	Fail
Male	23	27
Female	32	18

- Stating your hypotheses clearly test, at the 10% level, whether or not there is any evidence of an association between gender and passing a driving test at the first attempt. (11)
-

8. Observations have been made over many years of T , the noon temperature in °C, on 21st March at *Sunnymere*. The records for a random sample of 12 years are given below.

5.2, 3.1, 10.6, 12.4, 4.6, 8.7, 2.5, 15.3, -1.5, 1.8, 13.2, 9.3.

- (a) Find unbiased estimates of the mean and variance of T . (5)

Over the years, the standard deviation of T has been found to be 5.1.

- (b) Assuming a normal distribution find a 90% confidence interval for the mean of T . (5)

A meteorologist claims that the mean temperature at noon in *Sunnymere* on 21st March is 4 °C.

- (c) Use your interval to comment on the meteorologist's claim. (2)
-

END