

General Certificate of Education
January 2004
Advanced Level Examination



**MATHEMATICS AND STATISTICS
(SPECIFICATION B)
Unit Statistics 6**

MBS6

Wednesday 14 January 2004 Morning Session

In addition to this paper you will require:

- an 8-page answer book;
- the AQA booklet of formulae and statistical tables;
- one sheet of graph paper for use in Question 3;
- a ruler.

You may use a graphics calculator.

Time allowed: 1 hour 15 minutes

Instructions

- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MBS6.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information

- The maximum mark for this paper is 60.
- Mark allocations are shown in brackets.

Advice

- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

- 1 Following an intensive advertising campaign for a particular brand of detergent, a market research company asked 21 households how many packets of this brand they had purchased during the campaign period.

The numbers of packets purchased are given below.

3 9 20 0 12 16 46 23 34 2 15 11 13 14 2 17 1 20 9 0 28

- (a) The median number of packets of this brand purchased over a similar time period before the campaign was 11.

Carry out a sign test, at the 10% level of significance, to investigate whether the median number of packets purchased was greater during the campaign. State the null and alternative hypotheses used. *(6 marks)*

- (b) Give **one** reason why the sign test, rather than the Wilcoxon signed-rank test, was used in part (a). *(1 mark)*

- 2 On a visit to a theme park, the probabilities that a visitor will go on the Monorail, the Big Mountain ride and the Fearsome ride are 0.65, 0.52 and 0.38 respectively.

The probability that a visitor will go on the Big Mountain ride and the Fearsome ride is 0.25.

The probability that a visitor will go on the Monorail and the Big Mountain ride is 0.30.

The probability that a visitor will go on the Fearsome ride, given that they go on the Monorail, is 0.46.

Find the probability that a randomly selected visitor to this theme park will go on:

- (a) the Monorail and the Fearsome ride; *(2 marks)*
- (b) the Monorail or the Big Mountain ride or both; *(2 marks)*
- (c) neither the Big Mountain ride nor the Fearsome ride; *(3 marks)*
- (d) the Monorail, given that they go on the Fearsome ride. *(3 marks)*

3 [A sheet of graph paper is provided for use in this question.]

A government water department in the USA keeps records of significant floods. The number of deaths, x , and the approximate cost at today's prices, y million dollars, of damage to property for the 11 significant floods between 1913 and 1999 are given in the following table.

x	y
467	2430
150	470
15	250
47	300
24	180
517	3200
69	480
290	1050
210	680
95	380
362	1390

- (a) Plot a scatter diagram to illustrate the above data. (3 marks)
- (b) Calculate the value of Spearman's rank correlation coefficient between x and y . (5 marks)
- (c) Carry out a hypothesis test, at the 1% level of significance, to determine whether your value calculated in part (b) indicates a positive association between x and y .

Interpret your conclusion in the context of the question. (4 marks)

- (d) With reference to your scatter diagram, give a reason why the product moment correlation coefficient is not an appropriate measure of association for the data given. (1 mark)

- 4 In order to determine whether two tests, Test 1 and Test 2, are equally effective in evaluating applicants for a job, the twenty questions from each test are combined in a random order to produce one intermixed test paper containing forty questions.

This combined paper is given to 12 applicants. For each applicant, the marks on the two separate sets of twenty questions are then totalled to give a Test 1 score and a Test 2 score.

These scores are given in the table below.

Applicant	A	B	C	D	E	F	G	H	I	J	K	L
Test 1 score	68	65	85	55	38	78	45	80	79	77	66	61
Test 2 score	64	74	82	60	13	68	49	88	82	78	68	68

- (a) Calculate values of the mean and the standard deviation of the scores for each of Test 1 and Test 2. *(3 marks)*
- (b) Carry out a Wilcoxon signed-rank test, at the 5% level of significance, to investigate whether there is any difference in the median scores for the two tests. *(9 marks)*
- (c) Find the value of the product moment correlation coefficient between the scores for the two tests. *(3 marks)*
- (d) With reference to your findings in parts (a), (b) and (c), compare the performance of the two tests. *(2 marks)*
- (e) It is also suggested that the combined paper is too long for applicants and that Test 1 and Test 2 should be given to two separate groups of applicants. The scores of the group taking Test 1 could then be compared with the scores of the group taking Test 2.

Give **two** reasons why using the combined paper as described above is preferable to this suggestion when investigating whether Test 1 and Test 2 are equally effective.

(2 marks)

5 The Mann-Whitney U test is used to test the following hypotheses:

H_0 : the two populations have identical distributions

H_1 : the two populations do not have identical distributions

(a) Random samples, each of size 5, are taken from two independent populations.

The complete set of 10 observations is ranked in ascending order. No two observations are equal. The sum of the ranks for one of the samples is denoted by T .

Find the minimum and the maximum possible values for T . (4 marks)

(b) Two suppliers for a material used in the construction industry are being considered.

Six samples are obtained from supplier A and seven from supplier B. The material from each supplier is measured and is put into rank order. Rank 1 indicates the lowest value for the material.

The total, T_A , of the ranks for supplier A is 29 and the total, T_B , of the ranks for supplier B is 62.

- (i) Carry out a Mann-Whitney U test, at the 5% level of significance, to determine whether there is evidence of a difference between the two suppliers. (5 marks)
- (ii) Explain the meaning of a Type II error in the context of the test carried out in part (b)(i). (2 marks)

END OF QUESTIONS