General Certificate of Education
January 2004
Advanced Subsidiary Examination

MATHEMATICS AND STATISTICS　　　　　MBS3
(SPECIFICATION B)
Unit Statistics 3

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 4;
- 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 MBS3.
- 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.

　　　　　　　　　　　　　　　　　　　　　　MBS3

Answer **all** questions.

1 Following an intensive advertising campaign for a particular brand of detergent, a market research company asked 20 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 12 13 14 2 17 1 20 9 0

(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 A group of 50 female and 30 male students is questioned about their eating habits.

(a) Of the females, 40 per cent do not eat breakfast and, of the males, 10 per cent do not eat breakfast.

Find the probability that a student, selected at random from the group:

(i) does not eat breakfast; *(3 marks)*

(ii) is female, given that the student does not eat breakfast. *(2 marks)*

(b) Of all the students in the group, 10 per cent eat a hot breakfast, 20 per cent eat a hot lunch and 25 per cent either eat a hot breakfast or eat a hot lunch or eat both.

Find the probability that a student, selected at random from the group:

(i) eats both a hot breakfast and a hot lunch; *(2 marks)*

(ii) eats a hot breakfast, given that the student eats a hot lunch. *(2 marks)*

3 At Lake Tahoe, there are concerns about pollution caused by watercraft activity. Measurements, in micrograms per litre, of a volatile organic compound MTBE were made during weekends when there was considerable watercraft activity, and also midweek when there was little activity.

The results, which may be regarded as a random sample, are given in the following table.

Weekend	4.20	1.00	3.15	0.98	1.64	0.61	2.91	4.00	2.64	3.30
Midweek	1.52	0.42	1.75	0.45	0.30	1.14	2.20	0.68	0.46	

(a) Carry out a Mann-Whitney U test, using the 5% significance level, to investigate whether MTBE levels are higher during weekends. *(13 marks)*

(b) Explain, in the context of this question, the meaning of a Type II error. *(2 marks)*

TURN OVER FOR THE NEXT QUESTION

4 [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 numbers of deaths, x, and the approximate costs 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. *(6 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)*

5 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	55	85	55	38	80	45	73	80	77	66	61
Test 2 score	64	64	82	60	23	68	49	81	83	78	68	68

(a) Carry out a Wilcoxon signed-rank test, at the 5% level of significance, to investigate whether there is any difference in the average scores for the two tests. *(10 marks)*

(b) Find the value of the product moment correlation coefficient between the scores for the two tests. *(3 marks)*

(c) With reference to your findings in parts (a) and (b), compare the performance of the two tests. *(2 marks)*

END OF QUESTIONS