

General Certificate of Education
June 2005
Advanced Level Examination



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

MBS6

Tuesday 7 June 2005 Afternoon Session

In addition to this paper you will require:

- an 8-page answer book;
- the AQA booklet of formulae and statistical tables;
- an insert for use in Question 4 (enclosed).

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.
- Fill in the boxes at the top of the insert. Make sure that you attach the insert to your answer book.

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 A consumer research company carries out an investigation into the prices of a random sample of 11 'own brand' food items in standard sized packages from two large supermarkets, X and Y. The price, in pence, of each food item in each supermarket is given in the following table.

Food item	A	B	C	D	E	F	G	H	I	J	K
Price in supermarket X	46	242	198	63	51	22	198	35	80	59	121
Price in supermarket Y	45	226	218	63	44	20	219	40	76	55	113

- (a) Carry out a sign test, at the 10% significance level, to investigate whether or not the median price difference between 'own brand' food items in the two supermarkets differs from zero. *(5 marks)*
- (b) Given that the median price difference for 'own brand' food items is zero, state the probability of making a Type II error in the test that you carried out in part (a). *(1 mark)*
- 2 In a target golf contest, Anne, Brian and Gurwyn each hit a ball towards a ring. The probability that a ball finishes inside the ring is $\frac{3}{4}$ for Anne, $\frac{2}{3}$ for Brian and $\frac{4}{5}$ for Gurwyn.

These probabilities may be considered to be independent.

Find the probability that:

- (a) none of the three balls finishes inside the ring; *(2 marks)*
- (b) at least one of the three balls finishes inside the ring; *(1 mark)*
- (c) exactly two of the three balls finish inside the ring; *(3 marks)*
- (d) the ball hit by Gurwyn does not finish inside the ring, given that exactly two of the three balls finish inside the ring. *(2 marks)*

- 3 A new long-life battery is being launched. It is claimed that this new battery has a median lifetime of more than 250 hours.

A trial is carried out on 15 of the new batteries, and the measured lifetimes, in hours, of 9 of these batteries are given below.

222 228 235 242 248 251 259 260 275

The remaining 6 batteries had lifetimes in excess of 280 hours.

- (a) (i) A Wilcoxon signed-rank test is to be carried out using the given data to investigate the claim. Explain why it is **not** necessary to know the exact lifetimes of the 6 batteries that had lifetimes in excess of 280 hours in order to carry out this test. *(2 marks)*

- (ii) Carry out a Wilcoxon signed-rank test, at the 5% level of significance, to investigate the claim.

Interpret your conclusion in the context of the question. *(8 marks)*

- (b) State **two** assumptions that are necessary in order for the test in part (a)(ii) to be valid. *(2 marks)*

TURN OVER FOR THE NEXT QUESTION

4 [Figure 1, printed on the insert, is provided for use in this question.]

The data in the table below (repeated as **Figure 1** on the insert) were compiled for 17 randomly selected states in the USA.

It shows:

- the crime rate, x , measured as the number of offences reported per million population;
- the number, y , of males aged 16–24 years per thousand population;
- the unemployment rate, z , of males aged 16–24 years per thousand population.

Some of the rank values for x , y and z are also given in the table.

x	Rank x	y	Rank y	z	Rank z
79	11	154	3	81	13
164	3	146		116	1
58	15	145		77	16
197	1	139	10	114	2
123	5	144	8	100	5
68	13	124	17	84	
96	6	130	15	97	
156	4	135	13	108	3
86		159	1	96	7
71		143	9	83	12
167	2	128		102	4
85		137		92	
51		131		77	
66	14	138	11	77	
80	10	155	2	94	8
95	7	145		91	10
54	16	146		79	14

- (a) Complete **Figure 1** on the insert. (3 marks)
- (b) Hence, or otherwise, calculate:
- (i) the value of Spearman's rank correlation coefficient between x and y ; (3 marks)
- (ii) the value of Spearman's rank correlation coefficient between x and z . (3 marks)
- (c) (i) Comment, in context, on your calculated value from part (b)(i). (1 mark)
- (ii) Carry out a hypothesis test, at the 5% level of significance, to determine whether the value that you calculated in part (b)(ii) indicates an association between x and z . (4 marks)
- (d) Interpret your conclusion from part (c)(ii) and your comment from part (c)(i) in the context of the question. (2 marks)

- 5** A random sample of twenty school leavers, all of whom went to university and subsequently obtained jobs, is contacted and asked to state their starting salaries.

Eight of the students went to 'Top League' universities and the remaining students went to other universities.

Their starting salaries are given in the following table.

'Top League'	22 300	19 850	24 500	24 200	27 500	26 400	31 650	23 775				
Other	17 850	19 400	20 200	21 990	22 000	23 250	24 350	24 600	25 200	22 750	28 750	18 500

- (a) Carry out a Mann-Whitney U test, at the 5% level of significance, to investigate whether or not the starting salaries of those students who went to 'Top League' universities are higher than those of students who went to other universities. (12 marks)
- (b) For a test such as the one required in part (a), with two samples, one of size 8 and one of size 12, find the maximum value possible for the test statistic U .
- Assume that no two observations are equal. (2 marks)
- (c) (i) Explain how a matched pairs design could be used to investigate whether or not the starting salaries of those students who went to 'Top League' universities are higher than those of students who went to other universities. (3 marks)
- (ii) Give the name of a suitable distribution-free test that could be used to analyse the data from the matched pairs design that you suggested in part (c)(i). (1 mark)

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE