General Certificate of Education
June 2004
Advanced Level Examination

MATHEMATICS AND STATISTICS   **MBS5**
(SPECIFICATION B)
Unit Statistics 5

Wednesday 16 June 2004   Afternoon Session

AQA

In addition to this paper you will require:
- a 12-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 45 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 MBS5.
- 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 80.
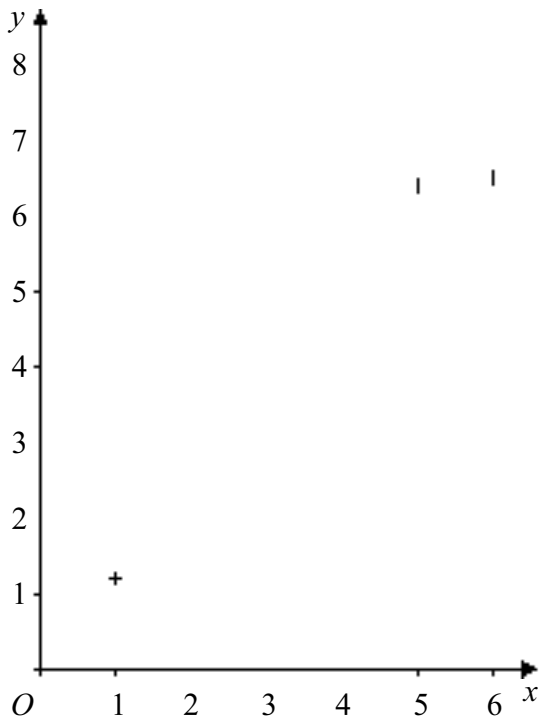- Mark allocations are shown in brackets.

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

P68292/0604/MBS5 6/6/

MBS5

Answer **all** questions.

1 The following two data sets are illustrated in the diagrams below.

Data Set A

x	1	2	3	4	5	6
y	1.2	3.5	5.2	6.0	6.4	6.5

Data Set B

x	1	2	3	4	5	6
y	1.7	4.0	3.4	6.4	5.6	7.6
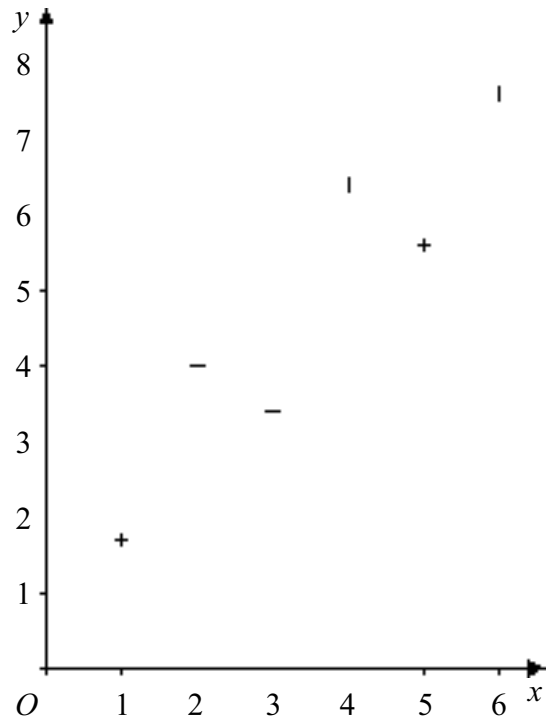
The product moment correlation coefficient for Data Set A is 0.924 correct to three significant figures.

(a) Calculate the product moment correlation coefficient for Data Set B.   *(3 marks)*

(b) Using Data Sets A and B as an example, explain why it is always advisable to draw a diagram, as well as calculating numerical measures, before interpreting bivariate data.
     (2 marks)

2 Priscilla supplies chilled food cabinets for use when displaying perishable foods in shops. She claims that the mean time to spoilage of yoghurt displayed in one of her cabinets is more than 200 hours.

To test this, Rio, a shopkeeper, placed 11 randomly selected tubs of fresh yoghurt in one of Priscilla's cabinets with a 'Not for Sale' sign. The times to spoilage, in hours, of the yoghurt in these tubs were as follows.

   206   213   197   215   202   200   209   217   195   203   204

 (a) Test Priscilla's claim using the 5% significance level. You may assume that the data is a sample from a normal distribution with standard deviation 6.0. *(7 marks)*

 (b) Discuss the problems, if any, which would have arisen in the interpretation of the results of the test if three of the tubs of yoghurt had accidentally been sold and so data were only available for the remaining eight tubs. *(2 marks)*

 (c) Define a Type I error in the context of this question. *(2 marks)*

 (d) Adam, another shopkeeper, intends to repeat Rio's data collection and analysis. State the probability of a Type I error occurring in the analysis of Adam's data when the mean time to spoilage of the tubs of fresh yoghurt displayed in Priscilla's cabinets is:

  (i) 200 hours;

  (ii) 202 hours. *(3 marks)*

TURN OVER FOR THE NEXT QUESTION

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

During train journeys, Ariane sells refreshments from a trolley to passengers at their seats. The following table shows for nine journeys the time, x minutes, she spent on the train and the value, £y, of her sales.

Journey	1	2	3	4	5	6	7	8	9
x	35	124	66	84	77	106	44	90	52
y	52	113	54	58	84	80	61	72	44

(a) Plot a scatter diagram of the data.   *(2 marks)*

(b) Calculate the equation of the regression line and draw the line on your scatter diagram.
                                                                                     (6 marks)

(c) The following table shows the residuals for some of the journeys.

Journey	1	2	3	4	5	6	7	8	9
Residual	7.6	15.1	−9.1	−15.9	14.3	−7.1	11.2		

   (i) Calculate the residuals for journeys 8 and 9.   *(3 marks)*

   (ii) Find the standard deviation of the nine residuals.   *(1 mark)*

(d) Desmond carries out the same duties as Ariane. His journey times and sales were recorded for nine journeys similar to those recorded for Ariane. The equation of the regression line for Desmond's journeys was $y = 50.2 + 0.290\,x$. The standard deviation of the residuals about this regression line was 6.2.

   (i) Draw Desmond's regression line on your scatter diagram.   *(2 marks)*

   (ii) Compare Desmond's sales performance with that of Ariane.   *(3 marks)*

4 A company manufactures components for the motor industry. The components are designed to have a length of 135.0 mm. Each day a technician takes a random sample of components and calculates a 95% confidence interval for the mean length.

(a) On a particular Monday the technician takes a sample of nine components. Their lengths, in millimetres, are as follows.

135.1   135.7   134.9   135.2   136.3   135.7   135.9   136.0   135.6

Assuming the lengths of components produced on this Monday may be modelled by a normal distribution with standard deviation 0.42, calculate a 95% confidence interval for the mean length. Give your answer to an appropriate degree of accuracy. *(6 marks)*

(b) On a particular Friday a random sample of 60 components had a mean length of 135.8 mm and a standard deviation of 3.9 mm. Calculate a 95% confidence interval for the mean length of components produced on this Friday. *(2 marks)*

(c) The company overhauls all machines when the confidence interval calculated does not contain 135.0 mm. State whether increasing the sample size, *n*, is likely to increase, decrease or leave unaltered:

(i) the width of the confidence interval; *(1 mark)*

(ii) the frequency of machine overhauls, explaining your answer. *(2 marks)*

(d) Using your results in parts (a) and (b) as an example, or otherwise, explain why it is possible that production on a day when the confidence interval does not contain 135.0 mm will sometimes be more satisfactory than production on a day when the confidence interval does contain 135.0 mm. *(2 marks)*

5 When Gordon and Louise play a board game, the probability that Gordon wins is 0.6 and the probability that Louise wins is 0.4. They agree to play a series of games. The winner of the series will be the first player to win three games. The result of each game is independent of the result of any other game.

Find the probability that:

(a) Gordon wins the series by three games to nil; *(2 marks)*

(b) more than three games are necessary to decide who wins the series; *(3 marks)*

(c) Louise wins the series by three games to one, given that she wins the first game; *(3 marks)*

(d) Louise wins the series, given that the final score is three games to two; *(2 marks)*

(e) Gordon wins the series, given that Louise wins the first game. *(4 marks)*

Turn over ▶

6 Each match in a 7-a-side football competition is scheduled to take 15 minutes (seven minutes each way plus one minute for half-time). However, due to delays, the actual time taken by a match may be modelled by a normal distribution with mean 19.5 minutes and standard deviation 1.2 minutes.

(a) Find the probability that a match will last between 16.5 and 18.0 minutes. *(5 marks)*

(b) (i) Find the probability that the mean length of six randomly selected matches will exceed 20 minutes. *(4 marks)*

(ii) Six matches are scheduled to be played on the same pitch, one after the other. Without further calculation, comment on the organisers' proposal that a total of two hours should be allocated for these six matches. *(2 marks)*

(c) For future competitions the organisers request that referees reduce delays to a minimum, so that matches finish on schedule. Following this request, it is observed that 80% of matches last less than 18 minutes and only 4% of matches last more than 19 minutes. Assuming that the lengths of matches may still be modelled by a normal distribution, find the mean and standard deviation. *(6 marks)*

END OF QUESTIONS

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE