General Certificate of Education
January 2005
Advanced Level Examination

MATHEMATICS AND STATISTICS (SPECIFICATION B) **MBS7**
Unit Statistics 7

Monday 31 January 2005 Morning Session

In addition to this paper you will require:
- an 8-page answer book;
- the AQA booklet of formulae and statistical tables.

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 MBS7.
- 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 Horticultural tests were carried out on a particular variety of tomato plant grown in one of two types of growbag, A and B.

A random sample of 20 of the plants was grown in Type A growbags and, for each plant, the weight of tomatoes produced was recorded. The mean of these weights was 3.44 kilograms.

A second random sample of 25 of the plants was grown in Type B growbags and, for each plant, the weight of tomatoes produced was recorded. The mean of these weights was 2.76 kilograms.

For plants grown in either type of growbag, the weight of tomatoes produced may be assumed to be normally distributed with a standard deviation of 0.4 kilograms.

Stating null and alternative hypotheses, test, at the 5% level of significance, the claim that, for this variety of tomato plant, the mean weight of tomatoes from plants grown in Type A growbags is larger than that of tomatoes from plants grown in Type B growbags by 0.5 kilograms per plant. *(8 marks)*

2 A machine produces plastic bottles in batches of 1000.

(a) Prior to a minor overhaul of the machine, the mean number of faulty bottles was 8 per batch. Following the overhaul, the machine produces 3 faulty bottles in the first batch.

Using a Poisson approximation, investigate, at the 5% level of significance, whether the overhaul has reduced the mean number of faulty bottles per batch. *(4 marks)*

(b) A major overhaul of the machine is undertaken.

The table below shows the number of faulty bottles found in each of 250 batches after this overhaul.

Number of faulty bottles	0	1	2	3	4	5	6
Number of batches	41	57	74	35	28	12	3

(i) Show that the mean number of faulty bottles per batch is 2. *(1 mark)*

(ii) Hence, using a χ^2 goodness of fit test and the 1% level of significance, test the hypothesis that a Poisson distribution models the number of faulty bottles per batch after the major overhaul. *(10 marks)*

3 As part of an investigation into the relationship between the depth, y metres, of a river and the distance, x metres, from the riverbank, a point was chosen on the riverbank one kilometre upstream from the mouth of the river.

At this point, measurements were made at various positions along a line at right angles to the riverbank, with the following results.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Distance, x	5	10	15	20	25	30	35
Depth, y	2.5	3.6	5.8	10.2	11.6	13.6	15.7

(a) Determine the equation of the least squares regression line, $y = \hat{\alpha} + \hat{\beta}x$. *(2 marks)*

(b) Assuming that the relationship between x and y may be modelled by

$$y = \alpha + \beta x + \varepsilon \qquad \text{where} \qquad \varepsilon \sim \mathrm{N}(0, \sigma^2):$$

(i) calculate an unbiased estimate of σ^2; *(3 marks)*

(ii) test, at the 5% level of significance, the hypothesis that $\beta = 0.5$. *(6 marks)*

(c) At the point chosen on the riverbank, the width of the river is 90 metres.

(i) Use your equation to estimate the depth of the river when $x = 45$. *(1 mark)*

(ii) Explain why the value that you calculated in part (c)(i) might be regarded as an estimate of the maximum depth of the river at this point. *(1 mark)*

(iii) Give **one** statistical reason and **one** practical reason why your estimate in part (c)(i) is **unlikely** to be accurate. *(2 marks)*

TURN OVER FOR THE NEXT QUESTION

4 Trains stop at a particular station for a minimum of 1 minute.

The time, T minutes, **in excess of 1 minute**, that trains stop at the station may be modelled by an exponential distribution with mean 2.

(a) State the probability that a train stops at the station for at least 1 minute. *(1 mark)*

(b) Calculate the probability that a train stops at the station for more than 5 minutes. *(3 marks)*

(c) A train has been stopped at the station for 3 minutes.

Calculate the probability that it stops at the station for **a total of** less than 5 minutes. *(4 marks)*

(d) A random sample of five trains each stop at the station for more than 5 minutes.

Indicate why this casts doubt on the above model. *(2 marks)*

5 Each morning, Jack's journey to work consists of four parts: car, transfer, rail and walk. The random variables denoting the durations, in minutes, of the four parts of his journey are C, T, R and W respectively.

The mean and standard deviation of the duration of each part of his journey are shown in the table.

Part of journey	**Mean** (minutes)	**Standard deviation** (minutes)
Car	20	3
Transfer	10	3
Rail	75	10
Walk	10	2

The duration of each part of his journey may be assumed to be normally distributed and independent of other parts.

(a) Calculate the probability that, on a particular day, Jack's journey to work takes less than 2 hours. *(5 marks)*

(b) (i) Show that, for any day, the probability that the duration of Jack's rail journey is more than two-thirds of the **total duration** of his journey is equivalent to:

$$\mathrm{P}((R - 2(C + T + W)) > 0)$$ *(2 marks)*

(ii) Hence determine the value of this probability. *(5 marks)*

END OF QUESTIONS