AQA

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

Monday 21 June 2004   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.

MBS7

Answer **all** questions.

1 The weight, in kilograms, of a particular make of paving brick may be assumed to be normally distributed with mean μ and variance σ^2.

A random sample of 10 bricks gave the following weights, in kilograms.

1.95   2.14   2.25   2.06   2.16   1.95   2.26   1.97   2.05   2.21

(a) Past experience suggests that $\sigma^2 = 0.01$. Verify that this value continues to be plausible by testing, at the 5% level of significance, the hypothesis that $\sigma^2 = 0.01$. *(7 marks)*

(b) Hence construct a 99% confidence interval for μ. *(5 marks)*

2 Between 8 am and 6 pm, the time, T minutes, between successive arrivals of vehicles for fuel at a village garage may be assumed to have an exponential distribution with a mean of 8.

(a) Write down the numerical value for the standard deviation of T. *(1 mark)*

(b) Calculate the probability that the time between successive arrivals of vehicles for fuel at the garage is between 5 minutes and 15 minutes. *(3 marks)*

(c) Between 6 pm and 9 pm, the time, S minutes, between successive arrivals of vehicles for fuel at the garage may be assumed to be independent of T and to have an exponential distribution with a mean of 15.

Calculate the probability that no vehicles arrive for fuel at the garage between 5.45 pm and 6.15 pm. *(5 marks)*

3 A manufacturer of 13-amp fuses states that 1.5 per cent of fuses produced are faulty.

(a) A wholesaler, who suspects that more than 1.5 per cent of fuses produced are faulty, tests a random sample of 100 fuses and finds that 4 fuses are faulty.

(i) Write down null and alternative hypotheses for investigating the wholesaler's suspicion. *(2 marks)*

(ii) Use a Poisson approximation and the 10% level of significance to investigate the wholesaler's suspicion. *(6 marks)*

(b) In view of the wholesaler's suspicion, the manufacturer decides to examine a random sample of 2000 fuses and finds that 36 fuses are faulty.

Use this information to investigate, again at the 10% level of significance, the wholesaler's suspicion. *(5 marks)*

4 The following scatter diagram illustrates the data collected as part of a study into the effect of temperature, $x\,°C$, on the pH value, y, of a particular liquid.

Scatter Diagram of pH versus Temperature
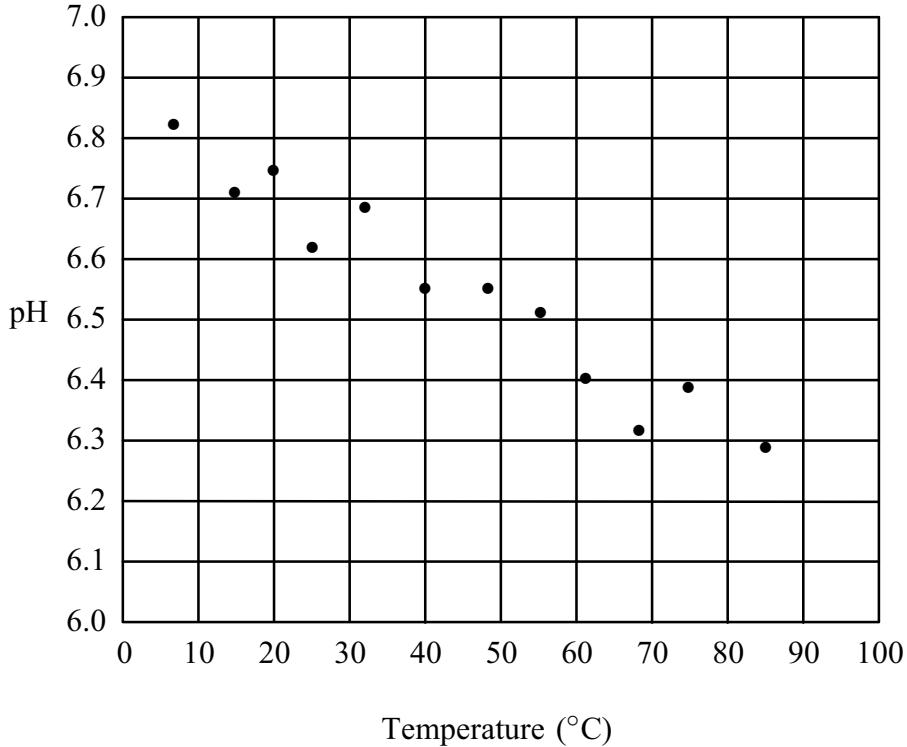
(a) Explain what this diagram reveals about the relationship between temperature and pH.

(2 marks)

(b) Given the following information, determine the equation of the least squares regression line, $y = \overset{\wedge}{\alpha} + \overset{\wedge}{\beta}x$.

$$\bar{x} = 44 \qquad \bar{y} = 6.55 \qquad S_{xx} = 6938 \qquad S_{xy} = -46.5$$

(2 marks)

(c) The relationship between x and y may be modelled by

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

   (i) Given that $n = 12$ and $S_{yy} = 0.3268$, calculate an unbiased estimate of σ^2.

(2 marks)

   (ii) Investigate, at the 1% level of significance, the hypothesis that $\beta < -0.005$.

(7 marks)

   (iii) Explain, in context, your conclusion to part (c)(ii).

(2 marks)

Turn over ▶

5 (a) The random variable X has a normal distribution with mean 200 and variance 100.

   (i) Show that $P(X < x) = P(Z < (0.1x - 20))$ where $Z \sim N(0, 1)$. *(1 mark)*

   (ii) **Table A** shows the probability that a value, x, lies within a particular interval.

Table A

Interval	Probability
$x < 180$	
$180 \leqslant x < 190$	
$190 \leqslant x < 200$	0.34134
$200 \leqslant x < 210$	
$210 \leqslant x < 220$	
$x \geqslant 220$	

   Copy and complete this table. *(2 marks)*

(b) **Table B** shows the frequency distribution of the times, X minutes, taken to complete a post round on a sample of 600 days.

Table B

Time to complete post round	Frequency
$x < 180$	19
$180 \leqslant x < 190$	74
$190 \leqslant x < 200$	193
$200 \leqslant x < 210$	218
$210 \leqslant x < 220$	85
$x \geqslant 220$	11

Using a goodness of fit test and the 5% level of significance, test the hypothesis that X may be modelled by a normal distribution with mean 200 and variance 100.
(Calculate expected frequencies to two decimal places.) *(8 marks)*

END OF QUESTIONS