General Certificate of Education
Advanced Subsidiary Examination
June 2011

Mathematics MS/SS1B

Unit Statistics 1B

Statistics

Unit Statistics 1B

Friday 20 May 2011 1.30 pm to 3.00 pm

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.
- Unit Statistics 1B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

1 The number of matches in each of a sample of 85 boxes is summarised in the table.

Number of matches	**Number of boxes**
Less than 239	1
239–243	1
244–246	2
247	3
248	4
249	6
250	10
251	13
252	16
253	20
254	5
255–259	3
More than 259	1
Total	**85**

(a) For these data:

(i) state the modal value; *(1 mark)*

(ii) determine values for the median and the interquartile range. *(3 marks)*

(b) Given that, on investigation, the 2 extreme values in the above table are 227 and 271 :

(i) calculate the range; *(1 mark)*

(ii) calculate estimates of the mean and the standard deviation. *(4 marks)*

(c) For the numbers of matches in the 85 boxes, suggest, with a reason, the most appropriate measure of spread. *(2 marks)*

2 The diameter, D millimetres, of an American pool ball may be modelled by a normal random variable with mean 57.15 and standard deviation 0.04.

(a) Determine:

(i) $P(D < 57.2)$; *(3 marks)*

(ii) $P(57.1 < D < 57.2)$. *(2 marks)*

(b) A box contains 16 of these pool balls. Given that the balls may be regarded as a random sample, determine the probability that:

(i) all 16 balls have diameters less than 57.2 mm; *(2 marks)*

(ii) the mean diameter of the 16 balls is greater than 57.16 mm. *(4 marks)*

---

3 (a) During a particular summer holiday, Rick worked in a fish and chip shop at a seaside resort.

He suspected that the shop's takings, £y, on a weekday were dependent upon the forecast of that day's maximum temperature, x °C, in the resort, made at 6.00 pm on the previous day.

To investigate this suspicion, he recorded values of x and y for a random sample of 7 weekdays during July.

x	23	18	27	19	25	20	22
y	4290	3188	5106	3829	5057	4264	4485

(i) Calculate the equation of the least squares regression line of y on x. *(4 marks)*

(ii) Estimate the shop's takings on a weekday during July when the maximum temperature was forecast to be 24 °C. *(2 marks)*

(iii) Explain why your equation may not be suitable for estimating the shop's takings on a weekday during February. *(1 mark)*

(iv) Describe, in the context of this question, a variable other than the maximum temperature, x, that may affect y. *(1 mark)*

(b) Seren, who also worked in the fish and chip shop, investigated the possible linear relationship between the shop's takings, £z, recorded in £000s, and **each** of two other explanatory variables, v and w.

(i) She calculated correctly that the regression line of z on v had a z-intercept of -1 and a gradient of 0.15.

Draw this line, for values of v from 0 to 40, on **Figure 1** on page 4.

Turn over ▶

(ii) She also calculated correctly that the regression line of z on w had a z-intercept of 5 and a gradient of -0.40.

Draw this line, for values of w from 0 to 10, on **Figure 2** below. *(3 marks)*

Figure 1

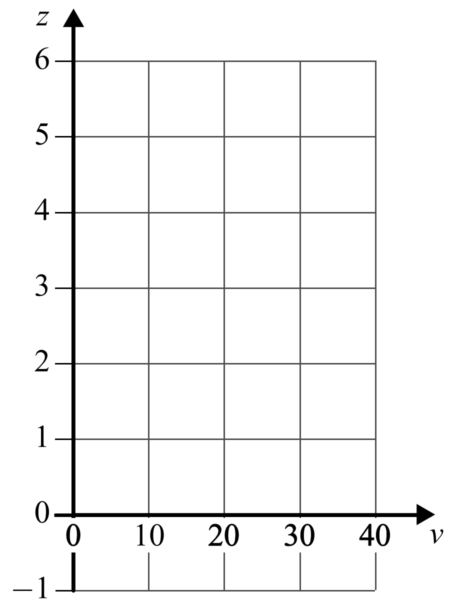
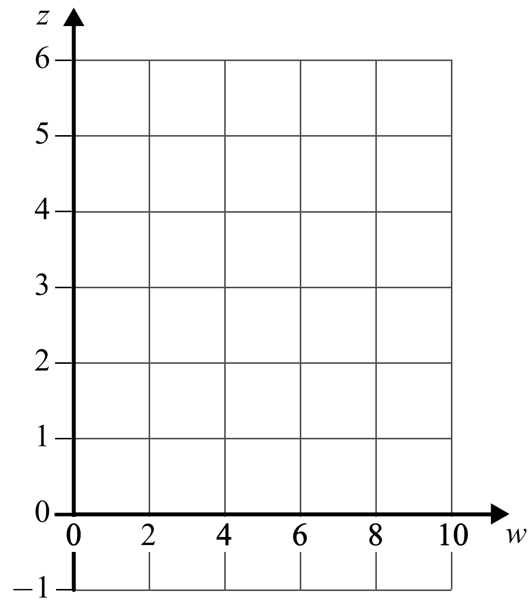


Figure 2



4 Rice that can be cooked in microwave ovens is sold in packets which the manufacturer claims contain a mean weight of more than 250 grams of rice.

The weight of rice in a packet may be modelled by a normal distribution.

A consumer organisation's researcher weighed the contents, x grams, of each of a random sample of 50 packets. Her summarised results are:

$$\bar{x} = 251.1 \quad \text{and} \quad \sum (x - \bar{x})^2 = 184.5$$

(a) Show that, correct to two decimal places, $s = 1.94$, where s^2 denotes the unbiased estimate of the population variance. *(1 mark)*

(b) (i) Construct a 96% confidence interval for the mean weight of rice in a packet, giving the limits to one decimal place. *(4 marks)*

(ii) Hence comment on the manufacturer's claim. *(2 marks)*

(c) The statement '250 grams' is printed on each packet.

Explain, with reference to the values of $\bar{x}$ and s, why the consumer organisation may consider this statement to be dubious. *(2 marks)*

---

5 (a) Emma visits her local supermarket every Thursday to do her weekly shopping.

The event that she buys orange juice is denoted by J, and the event that she buys bottled water is denoted by W. At each visit, Emma may buy neither, or one, or both of these items.

(i) Complete the table of probabilities, printed below, for these events, where J' and W' denote the events 'not J' and 'not W' respectively. *(3 marks)*

(ii) Hence, or otherwise, find the probability that, on any given Thursday, Emma buys either orange juice or bottled water but not both. *(2 marks)*

(iii) Show that:

(A) the events J and W are **not** mutually exclusive;

(B) the events J and W are **not** independent. *(3 marks)*

(b) Rhys visits the supermarket every Saturday to do his weekly shopping. Items that he may buy are milk, cheese and yogurt.

The probability, $\mathrm{P}(M)$, that he buys milk on any given Saturday is 0.85.
The probability, $\mathrm{P}(C)$, that he buys cheese on any given Saturday is 0.60.
The probability, $\mathrm{P}(Y)$, that he buys yogurt on any given Saturday is 0.55.
The events M, C and Y may be assumed to be independent.

Calculate the probability that, on any given Saturday, Rhys buys:

(i) none of the 3 items; *(2 marks)*

(ii) exactly 2 of the 3 items. *(3 marks)*

	J	J'	**Total**
W			0.65
W'	0.15		
Total		0.30	1.00

---

6 An amateur tennis club purchases tennis balls that have been used previously in professional tournaments.

The probability that each such ball fails a standard bounce test is 0.15.

The club purchases boxes each containing 10 of these tennis balls. Assume that the 10 balls in any box represent a random sample.

(a) Determine the probability that the number of balls in a box which fail the bounce test is:

(i) at most 2; *(1 mark)*

(ii) at least 2; *(2 marks)*

(iii) more than 1 but fewer than 5. *(3 marks)*

(b) Determine the probability that, in **5 boxes**, the total number of balls which fail the bounce test is:

(i) more than 5; *(2 marks)*

(ii) at least 5 but at most 10. *(3 marks)*

7 (a) Three airport management trainees, Ryan, Sunil and Tim, were each instructed to select a random sample of 12 suitcases from those waiting to be loaded onto aircraft.

Each trainee also had to measure the volume, x, and the weight, y, of each of the 12 suitcases in his sample, and then calculate the value of the product moment correlation coefficient, r, between x and y.

- Ryan obtained a value of -0.843.
- Sunil obtained a value of $+0.007$.

Explain why neither of these two values is likely to be correct. *(2 marks)*

(b) Peggy, a supervisor with many years' experience, measured the volume, x cubic feet, and the weight, y pounds, of each suitcase in a random sample of 6 suitcases, and then obtained a value of 0.612 for r.

- Ryan and Sunil each claimed that Peggy's value was different from their values because she had measured the volumes in cubic feet and the weights in pounds, whereas they had measured the volumes in cubic metres and the weights in kilograms.
- Tim claimed that Peggy's value was almost exactly half his calculated value because she had used a sample of size 6 whereas he had used one of size 12.

Explain why neither of these two claims is valid. *(2 marks)*

(c) Quentin, a manager, recorded the volumes, v, and the weights, w, of a random sample of 8 suitcases as follows.

v	28.1	19.7	46.4	23.6	31.1	17.5	35.8	13.8
w	14.9	12.1	21.1	18.0	19.8	19.2	16.2	14.7

(i) Calculate the value of r between v and w. *(3 marks)*

(ii) Interpret your value in the context of this question. *(2 marks)*

END OF QUESTIONS