General Certificate of Education
Advanced Subsidiary Examination
January 2013

Mathematics MS/SS1B

Unit Statistics 1B

Statistics

Unit Statistics 1B

Friday 18 January 2013 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 each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- 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.
- You do not necessarily need to use all the space provided.

1 Bob, a church warden, decides to investigate the lifetime of a particular manufacturer's brand of beeswax candle. Each candle is 30 cm in length.

From a box containing a large number of such candles, he selects one candle at random. He lights the candle and, after it has burned continuously for x hours, he records its length, y cm, to the nearest centimetre. His results are shown in the table.

x	5	10	15	20	25	30	35	40	45
y	27	25	21	19	16	11	9	5	2

(a) State the value that you would **expect** for a in the equation of the least squares regression line, $y = a + bx$. *(1 mark)*

(b) (i) Calculate the equation of the least squares regression line, $y = a + bx$. *(4 marks)*

(ii) Interpret the value that you obtain for b. *(2 marks)*

(iii) It is claimed by the candle manufacturer that the total length of time that such candles are likely to burn for is more than 50 hours.

Comment on this claim, giving a numerical justification for your answer. *(2 marks)*

---

2 The volume of *Everwhite* toothpaste in a pump-action dispenser may be modelled by a normal distribution with a mean of 106 ml and a standard deviation of 2.5 ml.

Determine the probability that the volume of *Everwhite* in a randomly selected dispenser is:

(a) less than 110 ml; *(3 marks)*

(b) more than 100 ml; *(2 marks)*

(c) between 104 ml and 108 ml; *(3 marks)*

(d) **not** exactly 106 ml. *(1 mark)*

3 *Stopoff* owns a chain of hotels. Guests are presented with the bills for their stays when they check out.

(a) Assume that the number of bills that contain errors may be modelled by a binomial distribution with parameters n and p, where $p = 0.30$.

Determine the probability that, in a random sample of 40 bills:

(i) at most 10 bills contain errors;

(ii) at least 15 bills contain errors;

(iii) exactly 12 bills contain errors. *(6 marks)*

(b) Calculate the mean and the variance for **each** of the distributions B(16, 0.20) and B(16, 0.125). *(3 marks)*

(c) Stan, who is a travelling salesperson, always uses *Stopoff* hotels. He holds one of its diamond customer cards and so should qualify for special customer care. However, he regularly finds errors in his bills when he checks out.

Each month, during a 12-month period, Stan stayed in *Stopoff* hotels on exactly 16 occasions. He recorded, each month, the number of occasions on which his bill contained errors. His recorded values were as follows.

2 1 4 3 1 3 0 3 1 0 5 1

(i) Calculate the mean and the variance of these 12 values. *(2 marks)*

(ii) Hence state with reasons which, if either, of the distributions B(16, 0.20) and B(16, 0.125) is likely to provide a satisfactory model for these 12 values. *(3 marks)*

4 Ashok is a work-experience student with an organisation that offers two separate professional examination papers, I and II.

For each of a random sample of 12 students, A to L, he records the mark, x per cent, achieved on Paper I, and the mark, y per cent, achieved on Paper II.

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
x	34	46	53	62	67	72	60	54	70	71	82	85
y	61	66	72	78	88	81	49	60	54	44	49	36

(a) (i) Calculate the value of the product moment correlation coefficient, r, between x and y. *(3 marks)*

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

(b) (i) Give **two** possible advantages of plotting data on a graph before calculating the value of a product moment correlation coefficient. *(2 marks)*

(ii) Complete the plotting of Ashok's data on the scatter diagram on page 5. *(2 marks)*

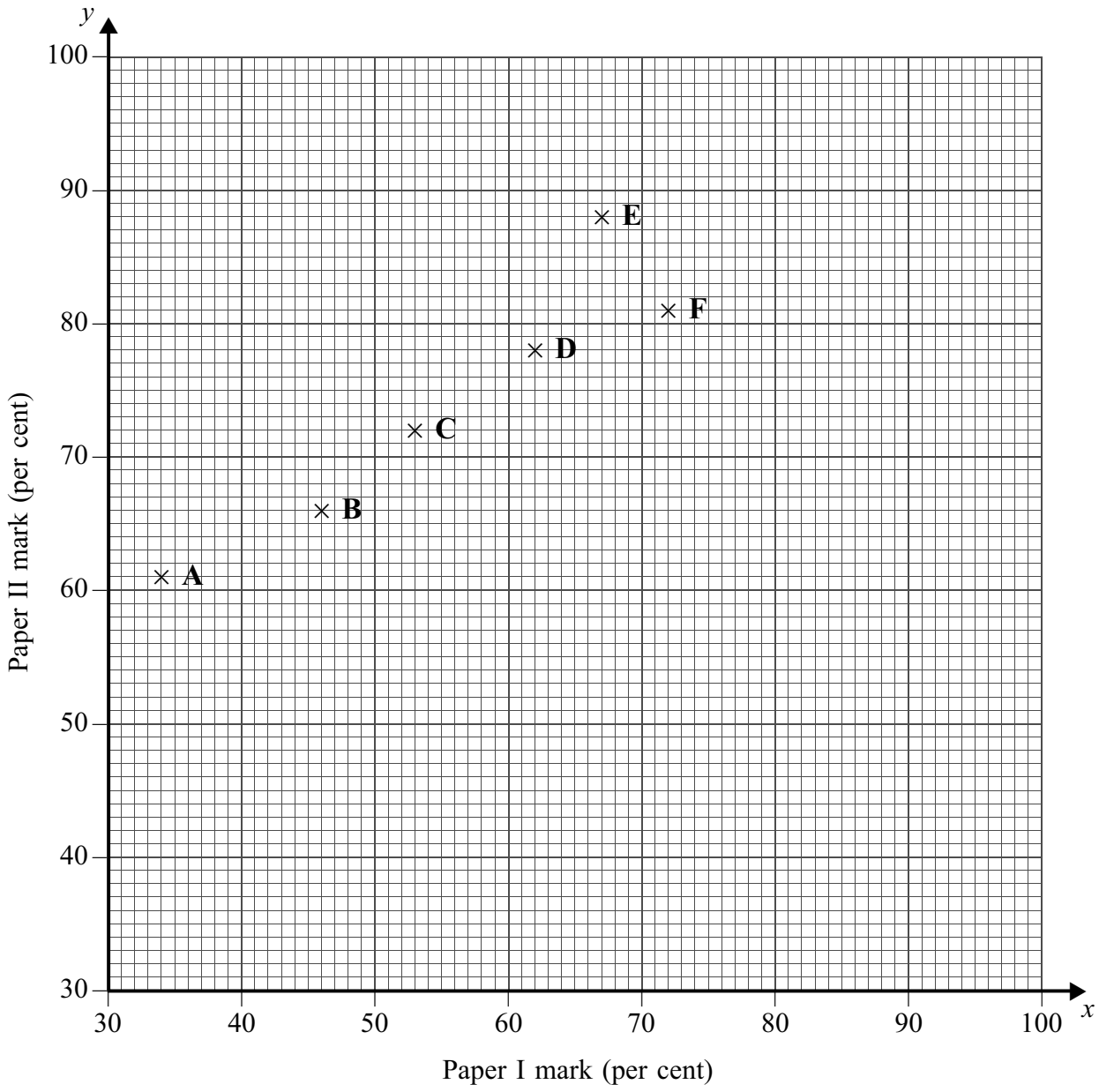
(iii) State what is now revealed by the scatter diagram. *(1 mark)*

(c) Ashok subsequently discovers that students A to F have a more scientific background than students G to L.

With reference to your scatter diagram, estimate the value of the product moment correlation coefficient for **each** of the two groups of students. You are **not** expected to calculate the two values. *(2 marks)*

	G	H	I	J	K	L
x	60	54	70	71	82	85
y	49	60	54	44	49	36

Examination Marks



5 Roger is an active retired lecturer. Each day after breakfast, he decides whether the weather for that day is going to be fine (F), dull (D) or wet (W). He then decides on only one of four activities for the day: cycling (C), gardening (G), shopping (S) or relaxing (R). His decisions from day to day may be assumed to be independent.

The table shows Roger's probabilities for each combination of weather and activity.

		Weather		
		Fine (F)	**Dull (D)**	**Wet (W)**
Activity	**Cycling (C)**	0.30	0.10	0
	Gardening (G)	0.25	0.05	0
	Shopping (S)	0	0.10	0.05
	Relaxing (R)	0	0.05	0.10

(a) Find the probability that, on a particular day, Roger decided:

(i) that it was going to be fine and that he would go cycling;

(ii) on either gardening or shopping;

(iii) to go cycling, given that he had decided that it was going to be fine;

(iv) **not** to relax, given that he had decided that it was going to be dull;

(v) that it was going to be fine, given that he did **not** go cycling. *(9 marks)*

(b) Calculate the probability that, on a particular Saturday and Sunday, Roger decided that it was going to be fine and decided on the same activity for both days.

(3 marks)

6 (a) The length of one-metre galvanised-steel straps used in house building may be modelled by a normal distribution with a mean of 1005 mm and a standard deviation of 15 mm.

The straps are supplied to house builders in packs of 12, and the straps in a pack may be assumed to be a random sample.

Determine the probability that the **mean** length of straps in a pack is less than one metre. *(4 marks)*

(b) Tania, a purchasing officer for a nationwide house builder, measures the **thickness**, x millimetres, of each of a random sample of 24 galvanised-steel straps supplied by a manufacturer. She then calculates correctly that the value of $\bar{x}$ is 4.65 mm.

(i) Assuming that the thickness, X mm, of such a strap may be modelled by the distribution $N(\mu, 0.15^2)$, construct a 99% confidence interval for μ. *(4 marks)*

(ii) Hence comment on the manufacturer's specification that the mean thickness of such straps is greater than 4.5 mm. *(2 marks)*

---

7 A machine, which cuts bread dough for loaves, can be adjusted to cut dough to any specified set weight. For any set weight, μ grams, the actual weights of cut dough are known to be approximately normally distributed with a mean of μ grams and a fixed standard deviation of σ grams.

It is also known that the machine cuts dough to within 10 grams of any set weight.

(a) Estimate, with justification, a value for σ. *(2 marks)*

(b) The machine is set to cut dough to a weight of 415 grams.

As a training exercise, Sunita, the quality control manager, asked Dev, a recently employed trainee, to record the weight of each of a random sample of 15 such pieces of dough selected from the machine's output. She then asked him to calculate the mean and the standard deviation of his 15 recorded weights.

Dev subsequently reported to Sunita that, for his sample, the mean was 391 grams and the standard deviation was 95.5 grams.

Advise Sunita on whether or not **each** of Dev's values is likely to be correct. Give numerical support for your answers. *(3 marks)*

(c) Maria, an experienced quality control officer, recorded the weight, y grams, of each of a random sample of 10 pieces of dough selected from the machine's output when it was set to cut dough to a weight of 820 grams. Her summarised results were as follows.

$$\sum y = 8210.0 \quad \text{and} \quad \sum (y - \bar{y})^2 = 110.00$$

Explain, with numerical justifications, why **both** of these values are likely to be correct. *(4 marks)*