AQA

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname ANSWERS

Forename(s)

Candidate signature

AS MATHEMATICS

Unit Statistics 1B

A

Wednesday 8 June 2016 Morning Time allowed: 1 hour 30 minutes

Materials

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

You may use a graphics calculator.

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.

Answer **all** questions.

Answer each question in the space provided for that question.

1 The table shows the heights, x cm, and the arm spans, y cm, of a random sample of 12 men aged between 21 years and 40 years.

x	152	166	154	159	179	167	155	168	174	182	161	163
y	143	154	151	153	168	160	146	163	170	175	155	158

(a) Calculate the value of the product moment correlation coefficient between x and y.

[3 marks]

(b) Interpret, in context, your value calculated in part **(a)**.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 1

1a) $r = 0.9591457841$

b) very strong positive linear correlation between heights and arm spans of men between 21 and 40 years.

2 A small chapel was open to visitors for 55 days during the summer of 2015. The table summarises the daily numbers of visitors.

Number of visitors	Number of days
20 or fewer	1 1
21	3 2
22	6 3
23	12 6
24	20 8
25	30 10
26	43 13
27	7
28	2
29	1
30 or more	2
Total	**55**

(a) For these data:

(i) state the modal value;

[1 mark]

(ii) find values for the median and the interquartile range.

[2 marks]

(b) Name one measure of average **and** one measure of spread that cannot be calculated exactly from the data in the table.

[2 marks]

(c) Reference to the raw data revealed that the 3 unknown exact values in the table were 13, 37 and 58.

Making use of this additional information, together with the data in the table, calculate the value of **each** of the two measures that you named in part **(b)**.

[3 marks]

QUESTION PART REFERENCE	**Answer space for question 2**
2ai)	26
ii)	Median → $\frac{55+1}{2}$ = 28th value
	= 25

QUESTION PART REFERENCE

Answer space for question 2

$IQR = UQ - LQ$

$UQ = \left(\frac{55+1}{4}\right) \times 3 = 42^{nd}$ value $\rightarrow 26$

$LQ = \frac{55+1}{4} = 14^{th}$ value $\rightarrow 24$

$IQR = 26 - 24 = \underline{2}$

b) cannot calculate mean or standard deviation / range

c) mean = $\underline{25.6}$

sd = $\underline{5.311064621}$ OR $\underline{5.262560752}$

range = $58 - 13 = \underline{45}$

3 The table shows, for a random sample of 500 patients attending a dental surgery, the patients' ages, in years, and the NHS charge bands for the patients' courses of treatment. Band 0 denotes the least expensive charge band and band 3 denotes the most expensive charge band.

		Charge band for course of treatment				
		Band 0	**Band 1**	**Band 2**	**Band 3**	**Total**
Age of patient (years)	**Under 19**	32	43	5	0	80
	Between 19 and 40	17	62	22	3	104
	Between 41 and 65	28	82	35	31	176
	66 or over	13	53	68	6	140
	Total	90	240	130	40	500

(a) Calculate, **to three decimal places**, the probability that a patient, selected at random from these 500 patients, was:

(i) aged between 41 and 65;

(ii) aged 66 or over and charged at band 2;

(iii) aged between 19 and 40 and charged **at most** at band 1;

(iv) aged 41 or over, given that the patient was charged at band 2;

(v) charged **at least** at band 2, given that the patient was **not** aged 66 or over.

[9 marks]

(b) Four patients at this dental surgery, **not** included in the above 500 patients, are selected at random.

Estimate, **to three significant figures**, the probability that two of these four patients are aged between 41 and 65 and are **not** charged at band 0, and the other two patients are aged 66 or over and are charged at either band 1 or band 2.

[5 marks]

QUESTION PART REFERENCE

Answer space for question 3

3ai) $P(41\text{-}65) = \frac{176}{500} = 0.352$

ii) $P(66 \text{ or over and band } 2) = \frac{68}{500} = 0.136$

iii) $P(19\text{-}40 \text{ and at most band } 1) = \frac{79}{500} = 0.158$

QUESTION PART REFERENCE

Answer space for question 3

iv) $P(41 \text{ or over} \mid \text{band 2}) = \frac{103}{130} = 0.7923076923$

$= 0.792$

v) $P(\text{at least band 2} \mid \text{not aged 66 or over}) = \frac{96}{360} = 0.267$

b) $P(41\text{-}65 \text{ and not band 0}) = \frac{150}{504}$

$P(66 \text{ or over and band 1 or 2}) = \frac{123}{504}$

METHOD MARK ONLY

0011, 0101, 0110, 1100, 1010, 1001

$\left(\frac{150}{504} \times \frac{149}{503} \times \frac{123}{502} \times \frac{122}{501}\right) \times 6$

$0.005260.. \times 6 = 0.03156121$

$= 0.0316$

$P(41\text{-}65 \text{ and not band 0}) = \frac{148}{500}$

$P(66 \text{ or over and band 1 or 2}) = \frac{121}{500}$

CORRECT ANSWER!

$\left(\frac{148}{500} \times \frac{148}{500} \times \frac{121}{500} \times \frac{121}{500}\right) \times 6$

$0.00513114 \times 6 = 0.030786860..$

$= 0.0308$

4 As part of her science project, a student found the mass, y grams, of a particular compound that dissolved in 100 ml of water at each of 12 different set temperatures, x °C. The results are shown in the table.

x	20	25	30	35	40	45	50	55	60	65	70	75
y	242	262	269	290	298	310	326	355	359	375	390	412

(a) Calculate the equation of the least squares regression line of y on x.

[4 marks]

(b) Interpret, in context, your value for the gradient of this regression line.

[2 marks]

(c) Use your equation to estimate the mass of the compound which will dissolve in 100 ml of water at 68 °C.

[1 mark]

(d) Given that the values of the 12 residuals for the regression line of y on x lie between -7 and $+9$, comment, with justification, on the likely accuracy of your estimate in part **(c)**.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 4

4a) $a = 181.3006993$

$b = 3.004195804$

$y = 181.3 + 3.004x$

b) for every additional 1°C, the mass increases by 3.004 g.

c) $x = 68$, $y = 181.3 + 3.004 \times 68$

$= 385.5853147$

d) the residuals are relatively small, therefore the regression line is reliable and the estimate will be fairly accurate.

5 Still mineral water is supplied in 1.5-litre bottles. The actual volume, X millilitres, in a bottle may be modelled by a normal distribution with mean $\mu = 1525$ and standard deviation $\sigma = 9.6$.

(a) Determine the probability that the volume of water in a randomly selected bottle is:

(i) less than 1540 ml;

(ii) more than 1535 ml;

(iii) between 1515 ml and 1540 ml;

(iv) not 1500 ml.

[7 marks]

(b) The supplier requires that only 10 per cent of bottles should contain more than 1535 ml of water.

Assuming that there has been no change in the value of σ, calculate the reduction in the value of μ in order to satisfy this requirement. Give your answer to one decimal place.

[4 marks]

(c) Sparkling spring water is supplied in packs of six 0.5-litre bottles. The actual volume in a bottle may be modelled by a normal distribution with mean 508.5 ml and standard deviation 3.5 ml.

Stating a necessary assumption, determine the probability that:

(i) the volume of water in **each** of the 6 bottles from a randomly selected pack is more than 505 ml;

(ii) the **mean** volume of water in the 6 bottles from a randomly selected pack is more than 505 ml.

[7 marks]

QUESTION PART REFERENCE

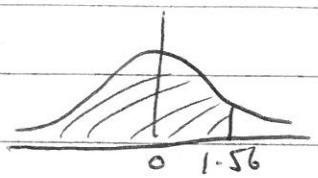
Answer space for question 5

5a) $X \sim N(1525, 9.6^2)$

i) $P(X < 1540) = P\left(z < \frac{1540 - 1525}{9.6}\right)$

$= P(z < 1.56)$

$= 0.94062$



QUESTION PART REFERENCE

Answer space for question 5

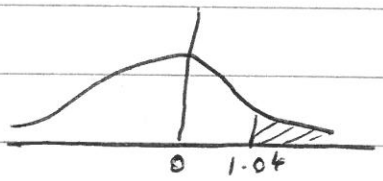
ii) $P(X > 1535) = P\left(z > \frac{1535 - 1525}{9.6}\right)$

$= P(z > 1.04)$

$= 1 - \Phi(1.04)$

$= 1 - 0.85083$

$= 0.14917$



iii) $P(1515 < X < 1540) = P\left(\frac{1515 - 1525}{9.6} < z < 1.56\right)$

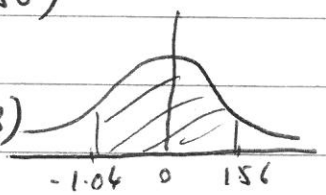
$= P(-1.04 < z < 1.56)$

$= \Phi(1.56) - \Phi(-1.04)$

$= 0.94062 - (1 - 0.85083)$

$= 0.94062 - 0.14917$

$= 0.79145$



iv) $P(X \neq 1500) = 1$

b) $P(X > 1535) = 0.1$

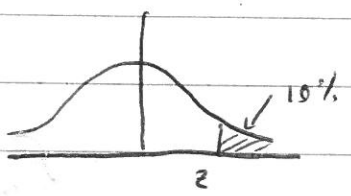
$\Phi(z) = 0.9$

$z = 1.2816$

$\frac{1535 - \mu}{9.6} = 1.2816$

$\mu = 1535 - 9.6 \times 1.2816$

$\mu = 1522.69664 = 1522.7$ ml



reduction in $\mu = 1525 - 1522.7$

$= 2.3$ ml

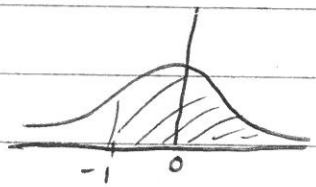
QUESTION PART REFERENCE

Answer space for question 5

c) i) assumptions → from correct population
it is a random sample

$X \sim N(508.5, 3.5^2)$

$P(X > 505) = P\left(z > \frac{505 - 508.5}{3.5}\right)$



$= P(z > -1)$

$= \Phi(1)$

$= 0.84134$

$P(\text{all } 6) = 0.84134^6$

$= 0.354673\ldots$

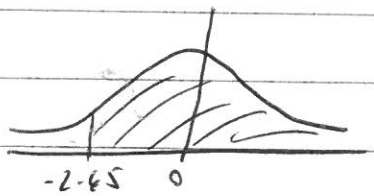
ii) $\bar{X} \sim N\left(508.5, \frac{3.5^2}{6}\right)$

$P(\bar{X} > 505) = P\left(\bar{z} > \frac{505 - 508.5}{\sqrt{\frac{3.5^2}{6}}}\right)$

$= P(\bar{z} > -2.45)$

$= \Phi(2.45)$

$= 0.99286$



6 The proportions of different colours of loom bands in a box of 10 000 loom bands are given in the table.

Colour	Blue	Green	Red	Orange	Yellow	White
Proportion	0.25	0.25	0.18	0.12	0.15	0.05

(a) A sample of 50 loom bands is selected at random from the box.

Use a binomial distribution with $n = 50$, together with relevant information from the table, to estimate the probability that this sample contains:

(i) exactly 4 **red** loom bands;

[2 marks]

(ii) at most 10 **yellow** loom bands;

[1 mark]

(iii) at least 30 **blue or green** loom bands;

[3 marks]

(iv) more than 35 but fewer than 45 loom bands that are **neither yellow nor white**.

[4 marks]

(b) The random variable R denotes the number of **red** loom bands in a random sample of **300** loom bands selected from the box.

Estimate values for the mean and the variance of R.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 6

6a) $X \sim B(50, p)$

i) $X \sim B(50, 0.18)$

$P(X=4) = {}^{50}C_4 \times 0.18^4 \times 0.82^{46} = 0.026231$

$X \sim B(50, 0.15)$

ii) $P(X \leq 10) = 0.8801$

$X \sim B(50, 0.5)$

iii) $P(X \geq 30) = 1 - P(X \leq 29)$

$= 1 - 0.8987$

$= 0.1013$

QUESTION PART REFERENCE

Answer space for question 6

iv) $X \sim B(50, 0.8)$ → $Y \sim B(50, 0.2)$

$P(35 < X < 45)$ → $P(6 \leqslant Y \leqslant 14)$

$36, 37, 38 \ldots, 44$ $\quad$ $14, 13, \ldots, 6$

$P(6 \leqslant Y \leqslant 14) = P(Y \leqslant 14) - P(Y \leqslant 5)$

$= 0.9393 - 0.0480$

$= \underline{0.8913}$

b) $R \sim B(300, 0.18)$

Mean $= np = 300 \times 0.18$

$= \underline{54}$

variance $= np(1-p)$

$= 54 \times 0.82$

$= \underline{44.28}$

7 Customers buying euros (€) at a travel agency must pay for them in pounds (£). The amounts paid, £x, by a sample of 40 customers were, in ascending order, as follows.

54.17	83.33	83.33	83.33	104.17	125.00	154.17	166.67	187.50	187.50
208.33	229.17	229.17	250.00	250.00	291.67	312.50	312.50	312.50	333.33
333.33	333.33	354.17	362.50	375.00	375.00	395.83	404.17	416.67	416.67
437.50	437.50	458.33	458.33	479.17	500.00	516.67	520.83	541.67	625.00

$\bar{x} = 317.50$ and $s = 146.30$

This sample of 40 customers may be regarded as a random sample.

(a) Construct a 99% confidence interval for the mean amount, in pounds, paid by customers buying euros at the travel agency. Give the limits to two decimal places.

[4 marks]

(b) The travel agency used an exchange rate of €1.20 to £1.00 for each of these 40 customers buying euros. There were no additional charges.

(i) Comment, with justification, on a claim that the mean number of euros bought by customers at the travel agency is 400.

(ii) Use the data in the table to comment on a claim that at most 25 per cent of customers at the travel agency buy fewer than €200.

[5 marks]

QUESTION PART REFERENCE

Answer space for question 7

7a) $\bar{x} = 317.50 \quad s = 146.30 \quad n = 40$

99% CI $\rightarrow \quad \Phi(z) = 0.995$

$z = 2.5758$

LB $\rightarrow \quad \bar{x} - z \times \frac{s}{\sqrt{n}} = 317.50 - 2.5758 \times \frac{146.30}{\sqrt{40}}$

$= £257.9164...$

UB $\rightarrow \quad \bar{x} + z \times \frac{s}{\sqrt{n}} = 317.50 + 2.5758 \times \frac{146.30}{\sqrt{40}}$

$= £377.08356...$

(£257.92, £377.08)

QUESTION PART REFERENCE	Answer space for question 7

b i) £257.92 × 1.2 = €309.504

£377.08 × 1.2 = €452.50

(309.50, 452.50)

Mean of €400 is within C.I so claim is justified

ii) €200 ÷ 1.20 = £166.67

number of customers below £166.67 is 7

$\frac{7}{40}$ is 17.5%

17.5% < 25% ∴ claim is justified.