GCE AS/A LEVEL

2305U20-1

S23-2305U20-1

FRIDAY, 19 MAY 2023 – AFTERNOON

FURTHER MATHEMATICS – AS unit 2
FURTHER STATISTICS A

1 hour 30 minutes

2305U201
01

ADDITIONAL MATERIALS

In addition to this examination paper, you will need:

- a WJEC pink 16-page answer booklet;
- a Formula Booklet;
- a calculator;
- statistical tables (RND/WJEC Publications).

INSTRUCTIONS TO CANDIDATES

Use black ink or black ball-point pen.

Answer **all** questions.

Sufficient working must be shown to demonstrate the **mathematical** method employed.

Answers without working may not gain full credit.

Unless the degree of accuracy is stated in the question, answers should be rounded appropriately.

INFORMATION FOR CANDIDATES

The maximum mark for this paper is 70.

The number of marks is given in brackets at the end of each question or part-question.

You are reminded of the necessity for good English and orderly presentation in your answers.

© WJEC CBAC Ltd.

stemup.co.uk

Additional Formulae for 2023

Laws of Logarithms

$$\log_a x + \log_a y \equiv \log_a (xy)$$

$$\log_a x - \log_a y \equiv \log_a \left(\frac{x}{y}\right)$$

$$k \log_a x \equiv \log_a \left(x^k\right)$$

Sequences

General term of an arithmetic progression:

$$u_n = a + (n-1)d$$

General term of a geometric progression:

$$u_n = ar^{n-1}$$

Mensuration

For a circle of radius, r, where an angle at the centre of θ radians subtends an arc of length s and encloses an associated sector of area A :

$$s = r\theta \qquad A = \frac{1}{2}r^2\theta$$

Calculus and Differential Equations

Differentiation

Function	Derivative
$f(x)g(x)$	$f'(x)g(x) + f(x)g'(x)$
$f\left(g(x)\right)$	$f'\left(g(x)\right)g'(x)$

Integration

Function	Integral
$f'\left(g(x)\right)g'(x)$	$f\left(g(x)\right) + c$

Area under a curve $= \int_a^b y \, \mathrm{d}x$

© WJEC CBAC Ltd.

stemup.co.uk

Reminder: Sufficient working must be shown to demonstrate the **mathematical** method employed.

1. The random variable X has mean 17 and variance 64. The independent random variable Y has mean 10 and variance 16. Find the value of

(a) E$(4Y - 2X + 1)$, [2]

(b) Var$(4Y - 5X + 3)$, [2]

(c) E(X^2Y). [3]

2. For a set of 30 pairs of observations of the variables x and y, it is known that $\sum x = 420$ and $\sum y = 240$. The least squares regression line of y on x passes through the point with coordinates (19, 20).

(a) Show that the equation of the regression line of y on x is $y = 2{\cdot}4x - 25{\cdot}6$ and use it to predict the value of y when $x = 26$. [6]

(b) State two reasons why your prediction in part (a) may not be reliable. [2]

3. It is known that the average lifetime of hair dryers from a certain manufacturer is 2 years. The lifetimes are exponentially distributed.

(a) Find the probability that the lifetime of a randomly selected hair dryer is between 1·8 and 2·5 years. [4]

(b) Given that 20 % of hair dryers have a lifetime of at least k years, find the value of k. [3]

(c) Jon buys his first hair dryer from the manufacturer today. He will replace his hair dryer with another from the same manufacturer immediately when it stops working. Find the probability that, in the next 5 years, Jon will have to replace more than 3 hair dryers. [3]

(d) State one assumption that you have made in part (c). [1]

TURN OVER

© WJEC CBAC Ltd.

stemup.co.uk

4. A continuous random variable X has cumulative distribution function F given by

$$F(x) = \begin{cases} 0 & \text{for} \quad x < 0, \\ \frac{1}{4}x & \text{for} \quad 0 \leqslant x \leqslant 2, \\ \frac{1}{480}x^4 + \frac{7}{15} & \text{for} \quad 2 < x \leqslant b, \\ 1 & \text{for} \quad x > b. \end{cases}$$

(a) Show that $b = 4$. [2]

(b) Find $P(X \leqslant 2{\cdot}5)$. [2]

(c) Write down the value of the lower quartile of X. [1]

(d) Find the value of the upper quartile of X. [3]

(e) Find, correct to three significant figures, the value of k that satisfies the equation $P(X > 3{\cdot}5) = P(X < k)$. [4]

© WJEC CBAC Ltd.

stemup.co.uk

5. (a) Give two circumstances where it may be more appropriate to use Spearman's rank correlation coefficient rather than Pearson's product moment correlation coefficient. [2]

(b) A farmer needs a new tractor. The tractor salesman selects 6 tractors at random to show the farmer. The farmer ranks these tractors, in order of preference, according to their ability to meet his needs on the farm. The tractor salesman makes a note of the price and power take-off (PTO) of the tractors.

Tractor	Farmer's rank	PTO (horsepower)	Price (£1000s)
A	1	77·5	80
B	6	87·9	45
C	5	53·0	47
D	4	41·0	53
E	2	112·0	60
F	3	90·0	61

Spearman's rank correlation coefficient between the farmer's ranks and the price is 0·9429.

(i) Test at the 5% significance level whether there is an association between the price of a tractor and the farmer's judgement of the ability of the tractor to meet his needs on the farm. [4]

(ii) Calculate Spearman's rank correlation coefficient between the farmer's rank and PTO. [4]

(iii) How should the tractor salesman interpret the results in (i) and (ii)? [2]

TURN OVER

© WJEC CBAC Ltd.

stemup.co.uk

6. A company has 20 boats to hire out. Payment is always taken in advance and all 20 boats are hired out each day.

A manager at the company notices that 10% of groups do not turn up to take the boats, despite having already paid to hire them. The manager wishes to investigate whether the numbers of boats that do not get taken each day can be modelled by the binomial distribution B(20, 0·1). The numbers of boats that were not taken for 110 randomly selected days are given below.

Number of boats not taken	0	1	2	3	4	5 or more
Frequency	10	35	29	25	8	3

(a) State suitable hypotheses to carry out a goodness of fit test. [1]

(b) Here is part of the table for a χ^2 goodness of fit test on the data.

Number of boats not taken	0	1	2	3	4	5 or more
Observed	10	35	29	25	8	3
Expected	f	29·72	g	20·91	9·88	4·75

(i) Calculate the values of f and g.

(ii) By completing the test, give the conclusion the manager should reach. [10]

The cost of hiring a boat is £15. Since demand is high and the proportion of groups that do not turn up is also relatively high, the manager decides to take payment for 22 boats each day. She would give £20 (a full refund and some compensation) to any group that has paid and turned up, but cannot take a boat out due to the overselling. Assume that the proportion of groups not turning up stays the same.

(c) (i) Suggest a binomial model that the manager could use for the number of groups arriving expecting to hire a boat.

(ii) Hence calculate the expected daily net income for the company following the manager's decision. [8]

(d) Is the manager justified in her decision? Give a reason for your answer. [1]

END OF PAPER

© WJEC CBAC Ltd.

stemup.co.uk

BLANK PAGE

© WJEC CBAC Ltd.

stemup.co.uk

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

1 (a)

$E(4Y - 2X + 1) = 4E(Y) - 2E(X) + 1$

$= 4 \times 10 - 2 \times 17 + 1$

$= 7$

1 (b)

$Var(4Y - 5X + 3) = 4^2Var(Y) + 5^2Var(X)$

$= 16 \times 16 + 25 \times 64$

$= 1856$

1 (c)

$E(X^2) = Var(X) + [E(X)]^2 = 64 + 17^2 = 353$

X and Y independent: $E(X^2Y) = E(X^2) \times E(Y)$

$E(X^2Y) = 353 \times 10 = 3530$

2 (a)

$\bar{x} = \frac{420}{30} = 14, \quad \bar{y} = \frac{240}{30} = 8$

The regression line passes through $(\bar{x}, \bar{y}) = (14, 8)$ and $(19, 20)$

$b = \frac{20 - 8}{19 - 14} = \frac{12}{5} = 2.4$

$y - 20 = 2.4(x - 19)$

$y = 2.4x - 45.6 + 20$

$y = 2.4x - 25.6$ as required

When $x = 26$: $y = 2.4 \times 26 - 25.6 = 36.8$

2 (b)

1. The relationship between x and y may not be linear (the correlation might be weak, so the points may not lie close to the line).
2. $x = 26$ may be outside the range of the observed x values

Question number

(extrapolation).

3 (a) Let X be the lifetime in years. $X \sim \text{Exp}(\lambda)$ with $\lambda = \frac{1}{2} = 0.5$

$P(1.8 \leqslant X \leqslant 2.5) = \int_{1.8}^{2.5} 0.5e^{-0.5x}\,dx$

$= [-e^{-0.5x}]_{1.8}^{2.5}$

$= e^{-0.9} - e^{-1.25} = 0.40656\ldots - 0.28650\ldots$

$= 0.12006\ldots = 0.120$ (3 s.f.)

3 (b) $P(X \geqslant k) = 0.2 \Rightarrow e^{-0.5k} = 0.2$

$-0.5k = \ln 0.2$

$k = -2 \ln 0.2 = 2 \ln 5 = 3.2188\ldots = 3.22$ years

3 (c) Let Y be the number of hair dryers replaced in 5 years.

Failures occur at rate 0.5 per year, so $Y \sim \text{Po}(0.5 \times 5) = \text{Po}(2.5)$

$P(Y > 3) = 1 - P(Y \leqslant 3)$

$= 1 - 0.75758$

$= 0.24242$ (0.242)

3 (d) The hair dryers still last 2 years on average (the quality of the hair dryers / manufacturing methods do not change in five years).

4 (a) $F(b) = 1$: $\frac{1}{480}b^4 + \frac{7}{15} = 1$

$\frac{1}{480}b^4 = \frac{8}{15} \Rightarrow b^4 = 256$

$b = \sqrt[4]{256} = 4$ $(b > 0)$ as required

4 (b) $P(X \leqslant 2.5) = F(2.5) = \frac{1}{480} \times 2.5^4 + \frac{7}{15}$

Question number

$= 0.0813802... + 0.4666666... = 0.548 \quad (\frac{1403}{2560})$

4 (c) $F(1) = \frac{1}{4} \times 1 = 0.25$, so the lower quartile = 1

4 (d) Upper quartile u: $\frac{1}{480}u^4 + \frac{7}{15} = 0.75$

$u^4 = 480 \times (0.75 - \frac{7}{15}) = 136$

$u = \sqrt[4]{136} = 3.4149... = 3.41$

4 (e) $P(X > 3.5) = 1 - F(3.5) = 1 - (\frac{1}{480} \times 3.5^4 + \frac{7}{15})$

$= 1 - 0.779296... = 0.220703...$

Since $0.125 < 0.2207... < 0.5$, k lies in $0 \leq k \leq 2$:

$P(X < k) = \frac{1}{4}k = 0.220703...$

$k = 0.882812... = 0.883$ (3 s.f.)

5 (a) 1. When the data are ordinal (given as ranks rather than measurements).

2. When the data are not from a bivariate normal distribution.

5 (b)(i) H_0: there is no association between the rank given by the farmer and the price of a tractor

H_1: there is an association between the rank given by the farmer and the price of a tractor

$n = 6$, 5% two-tailed critical value = 0.8286

Since $0.9429 > 0.8286$, there is sufficient evidence to reject H_0.

There is sufficient evidence to suggest that there is an association

Question number

between the rank given by the farmer and the price of a tractor.

5 (b)(ii) Rank PTO from highest (1) to lowest (6):

Tractor	A	B	C	D	E	F
Farmer	1	6	5	4	2	3
PTO rank	4	3	5	6	1	2
d	-3	3	0	-2	1	1
d^2	9	9	0	4	1	1

$\Sigma d^2 = 9 + 9 + 0 + 4 + 1 + 1 = 24$

$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 24}{6 \times 35} = \frac{11}{35} = 0.3143$ (4 d.p.)

5 (b)(iii) The farmer already prefers the more expensive tractors, so the salesman can concentrate on showing him the expensive ones. PTO is not strongly associated with the farmer's ranks, so it may not be worth focusing on. However the sample is small (6 tractors), so he should not rely too much on these results.

6 (a) H_0: the number of boats not taken each day can be modelled by B(20, 0.1)

H_1: the number of boats not taken each day cannot be modelled by B(20, 0.1)

6 (b)(i) $f = 110 \times P(X = 0) = 110 \times 0.9^{20} = 110 \times 0.121576... = 13.37$

$g = 110 \times P(X = 2) = 110 \times {}^{20}C_2 \times 0.1^2 \times 0.9^{18} = 110 \times 0.285180...$

$g = 31.37$

6 (b)(ii) Expected frequency for "5 or more" is less than 5, so combine with 4:

Boats	0	1	2	3	4 or more
Observed	10	35	29	25	11
Expected	13.37	29.72	31.37	20.91	14.63

Question number

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}$$

$$\chi^2 = \frac{(10 - 13.37)^2}{13.37} + \frac{(35 - 29.72)^2}{29.72} + \frac{(29 - 31.37)^2}{31.37}$$

$$+ \frac{(25 - 20.91)^2}{20.91} + \frac{(11 - 14.63)^2}{14.63}$$

$\chi^2 = 0.8494 + 0.9380 + 0.1791 + 0.8000 + 0.9007 = 3.667...$

Degrees of freedom $= 5 - 1 = 4$

5% critical value: $\chi^2_4 = 9.488$

Since $3.667 < 9.488$, there is insufficient evidence to reject H_0.

The manager should conclude that there is insufficient evidence to reject the suggestion that the number of boats not taken each day can be modelled by B(20, 0.1).

6 (c)(i) Let Y be the number of groups that turn up expecting to take a boat.

$Y \sim B(22, 0.9)$

6 (c)(ii) Income from 22 payments $= 22 \times £15 = £330$

Let S be the daily net income (£). Values of S: 330, 310, 290

$P(S = 330) = P(Y \leq 20) = 1 - 0.9^{22} - 22 \times 0.9^{21} \times 0.1 = 0.6608$

$P(S = 310) = P(Y = 21) = 22 \times 0.9^{21} \times 0.1 = 0.2407$

$P(S = 290) = P(Y = 22) = 0.9^{22} = 0.0985$

$E(S) = 330 \times 0.6608 + 310 \times 0.2407 + 290 \times 0.0985$

$E(S) = £321.25$

6 (d) Yes: the expected daily income of £321.25 is greater than the £15 × 20 = £300 the company receives when taking payment for 20 boats.