Surname	StemUp Model Solution	Centre Number	Candidate Number
First name(s)			2

GCE AS/A LEVEL

2305U20-1

S24-2305U20-1

FRIDAY, 17 MAY 2024 – 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 Formula Booklet;
- a calculator;
- statistical tables (RND/WJEC Publications).

INSTRUCTIONS TO CANDIDATES

Use black ink or black ball-point pen. Do not use gel pen or correction fluid.

You may use a pencil for graphs and diagrams only.

Write your name, centre number and candidate number in the spaces at the top of this page.

Answer **all** questions.

Write your answers in the spaces provided in this booklet. If you run out of space, use the additional page(s) at the back of the booklet, taking care to number the question(s) correctly.

For Examiner's use only		
Question	**Maximum Mark**	**Mark Awarded**
1	**14**	
2	**13**	
3	**12**	
4	**12**	
5	**12**	
6	**7**	
Total	**70**	

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.

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.

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

JUN242305U20101

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

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

1. Dave and Llinos like to go fishing. When they go fishing, on average, Dave catches 4·3 fish per day and Llinos catches 3·8 fish per day. A day of fishing is assumed to be 8 hours.

(a) (i) Calculate the probability that they will catch fewer than 2 fish in total on a randomly selected half-day of fishing. [4]

(ii) Justify any distribution you have used in answering (a)(i). [1]

(i) Total fish caught in a half-day, $F \sim Po((4.3 + 3.8) \times 0.5)$

$F \sim Po(4.05)$

$P(F < 2) = P(F = 0) + P(F = 1)$

$= \frac{e^{-4.05} \times 4.05^0}{0!} + \frac{e^{-4.05} \times 4.05^1}{1!}$

$= e^{-4.05}(1 + 4.05)$

$= 0.017422... \times 5.05$

$= 0.08798$

(ii) Fish are caught singly, independently and at random, at a constant average rate, and Dave and Llinos catch fish independently of each other, so a Poisson distribution (with the rates added) is suitable.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) On a randomly selected day, Dave starts fishing at 7 am. Given that Dave has not caught a fish by 11 am,

(i) find the expected time he catches his first fish,

(ii) calculate the probability that he will not catch a fish by 3 pm. [4]

(i) The Poisson process has no memory, so the time from 11 am to the first fish is exponential with mean $\frac{1}{4.3}$ day.

Expected time until the first fish = $\frac{1}{4.3} \times 8 = 1.860\ldots$ hours

= 1 hour 52 minutes

Expected time of first fish = 11 am + 1 h 52 min = 12:52 pm

(ii) 11 am to 3 pm is 4 hours = 0.5 day, so number of fish ~ Po(4.3 × 0.5) = Po(2.15)

P(no fish by 3 pm) = $P(X = 0) = e^{-4.3 \times 0.5}$

$= e^{-2.15}$

$= 0.1165$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(c) On average, only 2% of the fish that Llinos catches are trout. Over the course of a year, she catches 950 fish. Calculate the probability that at least 30 of these fish are trout. [3]

Let J be the number of trout she catches in a year.

$J \sim B(950, 0.02)$

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

$= 1 - 0.98891...$

$= 0.01109$

(d) State, with a reason, a distribution, including any parameters, that could approximate the distribution used in part (c). [2]

$Po(19) \quad (\lambda = np = 950 \times 0.02 = 19)$

A Poisson distribution is suitable since n is large (950) and p is small (0.02).

© WJEC CBAC Ltd.

stemup.co.uk

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

2. Emlyn aims to produce podcast episodes that are a standard length of time, which he calls the 'target time'. The time, X minutes, above or below the target time, which he calls the 'allowed time', can be modelled by the following cumulative distribution function.

$$F(x)=\begin{cases}0 & x<-2\\ \dfrac{x+2}{5} & -2\leqslant x<1\\ \dfrac{x^2-x+3}{5} & 1\leqslant x\leqslant 2\\ 1 & x>2\end{cases}$$

(a) Calculate the upper quartile for the 'allowed time'. [5]

$F(1) = \frac{1+2}{5} = 0.6 < 0.75$, so the upper quartile lies in $1 \leqslant x \leqslant 2$

$F(Q_3) = 0.75$: $\frac{x^2 - x + 3}{5} = 0.75$

$x^2 - x + 3 = 3.75$

$x^2 - x - 0.75 = 0$

$(x - 1.5)(x + 0.5) = 0$

$x = 1.5$ or $x = -0.5$

Reject $x = -0.5$ since it is not in the range $1 \leqslant x \leqslant 2$,

$\therefore$ upper quartile = 1.5 (minutes)

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) Find $f(x)$, the probability density function, for all values of x. [4]

$f(x) = F'(x) = \frac{d}{dx} F(x)$

$\frac{d}{dx}\left(\frac{x+2}{5}\right) = \frac{1}{5}$, $\frac{d}{dx}\left(\frac{x^2 - x + 3}{5}\right) = \frac{2x-1}{5}$

$$f(x) = \begin{cases} \frac{1}{5} & -2 \leq x < 1 \\ \frac{2x-1}{5} & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(c) (i) Calculate the mean 'allowed time'. [3]

(ii) Interpret your answer in context. [1]

(i) $E(X) = \int x\, f(x)\, dx = \int_{-2}^{1} \frac{x}{5}\, dx + \int_{1}^{2} \frac{2x^2 - x}{5}\, dx$

$= \left[\frac{x^2}{10}\right]_{-2}^{1} + \left[\frac{2x^3}{15} - \frac{x^2}{10}\right]_{1}^{2}$

$= \left(\frac{1}{10} - \frac{4}{10}\right) + \left(\left(\frac{16}{15} - \frac{4}{10}\right) - \left(\frac{2}{15} - \frac{1}{10}\right)\right)$

$= -\frac{3}{10} + \left(\frac{2}{3} - \frac{1}{30}\right)$

$= -\frac{3}{10} + \frac{19}{30}$

$= \frac{1}{3}$ (minutes)

(ii) On average, his podcast episodes are 20 seconds ($\frac{1}{3}$ minute) longer than the target time.

© WJEC CBAC Ltd.

stemup.co.uk

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

3. A company makes bags. The table below shows the number of bags sold on a random sample of 50 days. A manager believes that the number of bags sold per day can be modelled by the Poisson distribution with mean 2·2.

Number of bags sold	0	1	2	3	4	5 or more
Frequency	7	10	11	9	6	7

(a) Carry out a chi-squared goodness of fit test, using a 10% significance level. [11]

H_0: The number of bags sold per day can be modelled by a Poisson distribution with mean 2.2.

H_1: The number of bags sold per day cannot be modelled by a Poisson distribution with mean 2.2.

Expected frequency = 50 × P(X = r), where X ~ Po(2.2)

No. sold	0	1	2	3	4	5 or more
P(X = r)	0.1108	0.2438	0.2681	0.1966	0.1082	0.0725
Exp freq	5.54	12.19	13.41	9.83	5.41	3.62

Expected frequency for 5 or more is 3.62 < 5, so combine with 4:

No. sold	0	1	2	3	4 or more
Observed	7	10	11	9	13
Expected	5.54	12.19	13.41	9.83	9.03

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

$= \frac{(7-5.54)^2}{5.54} + \frac{(10-12.19)^2}{12.19} + \frac{(11-13.41)^2}{13.41} + \frac{(9-9.83)^2}{9.83} + \frac{(13-9.03)^2}{9.03}$

$= 0.3848 + 0.3934 + 0.4331 + 0.0701 + 1.7454$

$= 3.027$

Degrees of freedom = 5 − 1 = 4

10% critical value = 7.779

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

Since 3.027 < 7.779, there is insufficient evidence to reject H_0.

There is no evidence to suggest that the number of bags sold per day cannot be modelled by a Poisson distribution with mean 2.2.

(b) A chi-squared goodness of fit test for the Poisson distribution with mean 2·5 is conducted. This uses the same number of degrees of freedom as part (a) and gives a test statistic of 1·53. State, with a reason, which of these two Poisson models is a better fit for the data. [1]

Po(2.5) is the better fit, since its χ^2 test statistic (1.53) is smaller than 3.027 (with the same degrees of freedom).

© WJEC CBAC Ltd.

stemup.co.uk

4. An author poses the following question:

Does using cash for transactions affect people's financial behaviour?

She collects data on 'Cash transactions as a % of all transactions' and 'Household debt as a % of net disposable income' from a random sample of 25 countries. The table below shows the data she collected. There are missing values, p and q, for Malta and Denmark respectively.

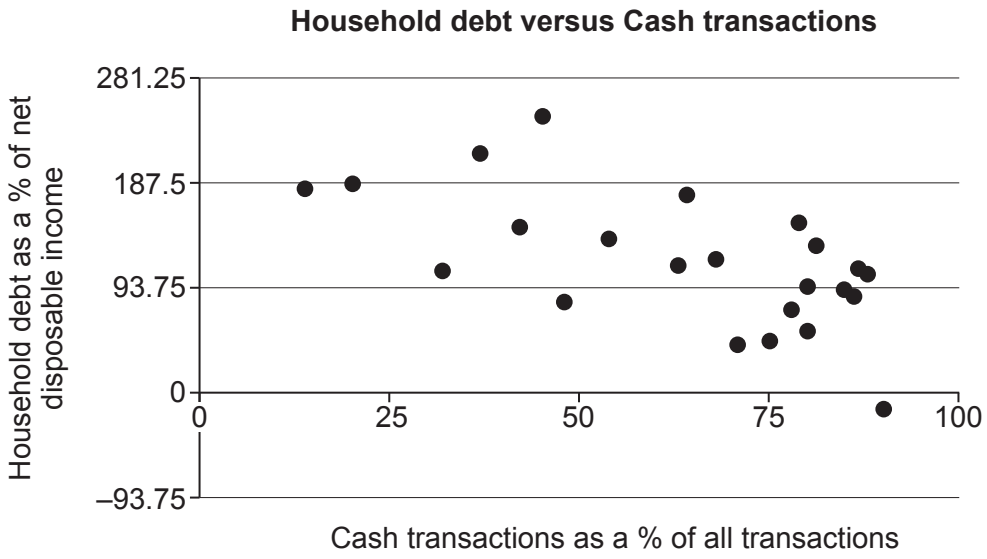
Country	Cash transactions as a % of all transactions x	Household debt as a % of net disposable income y	Country	Cash transactions as a % of all transactions x	Household debt as a % of net disposable income y
Malta	92	p	**France**	68	120
Mexico	90	–14	**Luxembourg**	64	177
Greece	88	107	**Belgium**	63	113
Spain	87	110	**Finland**	54	137
Italy	86	87	**Estonia**	48	82
Austria	85	91	**The Netherlands**	45	247
Portugal	81	131	**UK**	42	147
Slovenia	80	56	**Australia**	37	214
Germany	80	95	**USA**	32	109
Ireland	79	154	**Sweden**	20	187
Slovakia	78	74	**South Korea**	14	182
Lithuania	75	46	**Denmark**	q	261
Latvia	71	43			

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

The summary statistics and scatter diagram below are for the other 23 countries.



$\sum x = 1467$ $\sum y = 2695$ $\sum x^2 = 105073$ $S_{xx} = 11503 \cdot 91304$ $S_{yy} = 78669 \cdot 30435$

$\sum y^2 = 394453$ $\sum xy = 152999$ $S_{xy} = -18895 \cdot 13043$

(a) Using the summary statistics for the 23 countries, calculate and interpret Pearson's product moment correlation coefficient. [3]

$r = \frac{S_{xy}}{\sqrt{(S_{xx} S_{yy})}} = \frac{-18895.13043}{\sqrt{(11503.91304 \times 78669.30435)}}$

$r = -0.628093... = -0.628$

There is moderate (fairly strong) negative correlation: as the percentage of cash transactions increases, household debt as a % of net disposable income tends to decrease, i.e. countries that rely more on cash tend to have less debt.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) Calculate the equation of the least squares regression line of Household debt as a % of net disposable income (y) on Cash transactions as a % of all transactions (x). [5]

$b = \frac{S_{xy}}{S_{xx}} = \frac{-18895.13043}{11503.91304}$

$= -1.6424959... = -1.64$

$\bar{x} = \frac{1467}{23} = 63.7826...,\quad \bar{y} = \frac{2695}{23} = 117.1739...$

$a = \bar{y} - b\bar{x} = \frac{2695}{23} + 1.6424959... \times \frac{1467}{23}$

$= 117.1739... + 104.7627...$

$= 221.936... = 222$

$y = 222 - 1.64x$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

The regression line x on y is given below.

$$x = -0 \cdot 24y + 91 \cdot 92$$

(c) By selecting the appropriate regression line in each case, estimate the values of p and q in the table. [2]

Malta: x is known, so use the y on x line:

$p = 222 - 1.64 \times 92 = 71.12 \quad (p \approx 71)$

Denmark: y is known, so use the x on y line:

$q = -0.24 \times 261 + 91.92 = 29.28 \quad (q \approx 29)$

(d) Comment on the reliability of your answers in part (c). [1]

Not very reliable, since both are extrapolation: x = 92 is above the largest x value (90) and y = 261 is above the largest y value (247) in the data.

(e) Interpret the negative value of y for Mexico. [1]

In Mexico, net disposable income may be negative, i.e. household outgoings are greater than income, so debt as a % of it is negative (or it may be an erroneous value).

© WJEC CBAC Ltd.

stemup.co.uk

5. Lily is interested in the relationship between the way in which students learned Welsh and their attitude towards the Welsh language.

Students were categorised as having learned Welsh in one of three ways:

- from one Welsh-speaking parent/carer at home,
- from two Welsh-speaking parents/carers at home,
- at school only, for those with no Welsh-speaking parents/carers at home.

The students were asked to rate their attitude towards the Welsh language from 'Very negative' to 'Very positive'.

The following data for a random sample of 253 students were collected as part of a project.

	Learned Welsh			
Attitude	**From two parents/carers**	**From one parent/carer**	**At school only**	**Total**
Very negative	2	14	30	**46**
Slightly negative	4	20	21	**45**
Neutral	12	17	8	**37**
Slightly positive	21	19	11	**51**
Very positive	25	21	28	**74**
Total	**64**	**91**	**98**	**253**

Lily intends to carry out a chi-squared test for independence at the 5% level. She produces the following tables which are incomplete.

Expected Frequencies	**Learned Welsh**		
Attitude	**From two parents/carers**	**From one parent/carer**	**At school only**
Very negative	11·64	16·55	17·82
Slightly negative	11·38	16·19	17·43
Neutral	9·36	13·31	14·33
Slightly positive	12·90	18·34	19·75
Very positive	F	26·62	28·66

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

Chi-Squared Contributions	Learned Welsh		
Attitude	**From two parents/carers**	**From one parent/carer**	**At school only**
Very negative	7·98	0·39	8·33
Slightly negative	4·79	0·90	0·73
Neutral	0·74	1·02	G
Slightly positive	5·08	0·02	3·88
Very positive	2·11	1·19	0·02
Total	20·70	3·52	H

(a) Calculate the values of F, G and H. [4]

$F = \frac{\text{row total} \times \text{column total}}{\text{grand total}} = \frac{74 \times 64}{253} = 18.72$

$G = \frac{(O-E)^2}{E} = \frac{(8-14.33)^2}{14.33} = \frac{40.0689}{14.33} = 2.796\ldots = 2.80$

$H = 8.33 + 0.73 + G + 3.88 + 0.02$

$= 8.33 + 0.73 + 2.80 + 3.88 + 0.02$

$= 15.76$

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(b) Carry out Lily's chi-squared test for independence at the 5% level. [6]

H_0: Attitude towards Welsh is independent of how Welsh was learned.

H_1: Attitude towards Welsh is not independent of how Welsh was learned.

$\chi^2 = 20.70 + 3.52 + H = 20.70 + 3.52 + 15.76$

$\chi^2 = 39.98$

Degrees of freedom $= (5 - 1)(3 - 1) = 8$

5% critical value = 15.507

Since $39.98 > 15.507$, there is sufficient evidence to reject H_0.

There is evidence to suggest that how a student learned Welsh affects their attitude towards the Welsh language.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

(c) By referring to the figures in the tables on pages 16 and 17, give two comments on the relationship between the way students learned Welsh and their attitude towards the Welsh language. [2]

1. Students who learned Welsh from two parents/carers are more likely to have a positive attitude: far fewer were very negative than expected (2 observed vs 11.64 expected, contribution 7.98).

2. Students who learned Welsh at school only show a real mix of attitudes: many more than expected were very negative (30 vs 17.82, contribution 8.33), yet 28 of them were very positive.

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

6. Penelope makes 8 cakes per week. Each cake costs £20 to make and sells for £60. She always sells at least 5 cakes per week. Any cakes left at the end of the week are donated to a food bank. The probability that 5 cakes are sold in a week is 0·3. She is twice as likely to sell 6 cakes in a week as she is to sell 7 cakes in a week. The expected profit per week is £206.

Construct a probability distribution for the weekly profit. [7]

Cost = 8 × £20 = £160 per week. Profit = 60n − 160 when n cakes are sold.

n = 5: £140, n = 6: £200, n = 7: £260, n = 8: £320

Let P(profit = 260) = p, then P(profit = 200) = 2p,

P(profit = 320) = 1 − 0.3 − 2p − p = 0.7 − 3p

Profit £y	140	200	260	320
P(Y = y)	0.3	2p	p	0.7 − 3p

E(Y) = Σ y P(Y = y)

= 0.3 × 140 + 2p × 200 + p × 260 + (0.7 − 3p) × 320

206 = 42 + 400p + 260p + 224 − 960p

206 = 266 − 300p

300p = 60 ⇒ p = 0.2

Probability distribution of the weekly profit:

Profit £y	140	200	260	320
P(Y = y)	0.3	0.4	0.2	0.1

© WJEC CBAC Ltd.

stemup.co.uk

Examiner only

END OF PAPER

© WJEC CBAC Ltd.

stemup.co.uk

Question number	Additional page, if required. Write the question number(s) in the left-hand margin.

Examiner only

© WJEC CBAC Ltd.

stemup.co.uk

Question number	Additional page, if required. Write the question number(s) in the left-hand margin.

Examiner only

© WJEC CBAC Ltd.

stemup.co.uk

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

© WJEC CBAC Ltd.

stemup.co.uk