GCE AS/A LEVEL – NEW

2305U20-1

S18-2305U20-1

FURTHER MATHEMATICS – AS unit 2
FURTHER STATISTICS A

THURSDAY, 17 MAY 2018 – AFTERNOON

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.

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

INFORMATION FOR CANDIDATES

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.

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

1. The random variable X has the binomial distribution B(12, 0·3). The independent random variable Y has the Poisson distribution Po(4). Find

(a) $E(XY)$, [2]

(b) Var(XY). [6]

2. The length of time a battery works, in tens of hours, is modelled by a random variable X with cumulative distribution function

$$F(x)=\begin{cases}0 & \text{for } x<0,\\ \dfrac{x^3}{432}(8-x) & \text{for } 0 \leqslant x \leqslant 6,\\ 1 & \text{for } x>6.\end{cases}$$

(a) Find $P(X>5)$. [2]

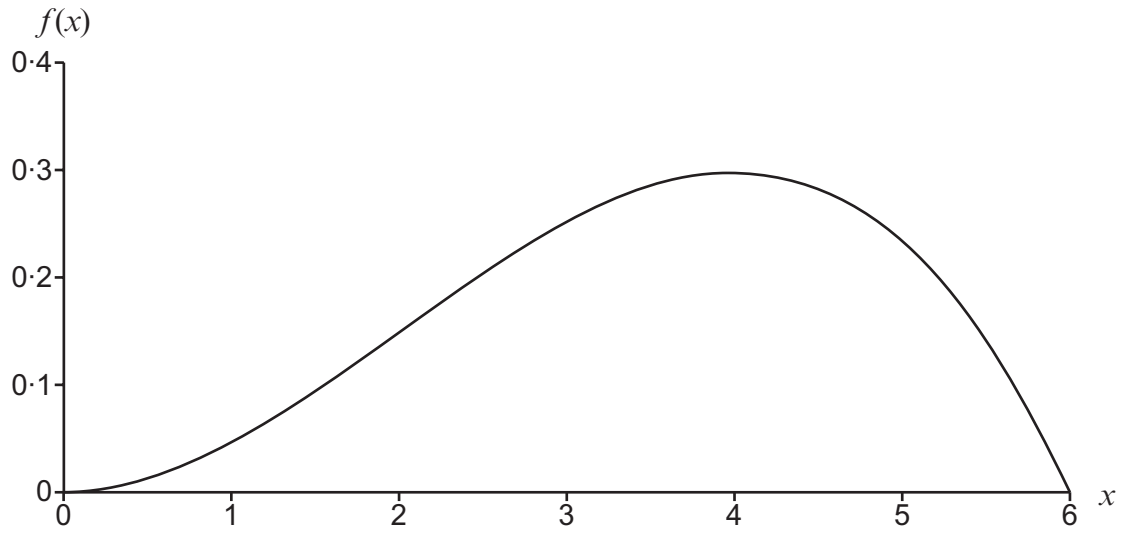
(b) A head torch uses three of these batteries. All three batteries must work for the torch to operate. Find the probability that the head torch will operate for more than 50 hours. [2]

(c) Show that the upper quartile of the distribution lies between 4·5 and 4·6. [3]

(d) Find $f(x)$, the probability density function for X. [3]

(e) Find the mean lifetime of the batteries in hours. [4]

(f) The graph of $f(x)$ is given below.



Give a reason why the model may not be appropriate. [1]

3. A game at a school fete is played with a fair coin and a random number generator which generates random integers between 1 and 52 inclusive. It costs 50 pence to play the game. First, the player tosses the coin. If it lands on tails, the player loses. If it lands on heads, the player is allowed to generate a random number. If the number is 1, the player wins £5. If the number is between 2 and 13 inclusive, the player wins £1. If the number is greater than 13, the player loses.

(a) Find the probability distribution of the player's profit. [5]

(b) Find the mean and standard deviation of the player's profit. [4]

(c) Given that 200 people play the game, calculate

(i) the expected number of players who win some money,

(ii) the expected profit for the fete. [2]

4. On a Welsh television game show, contestants are asked to guess the weights of a random sample of seven cows. The game show judges want to investigate whether there is positive correlation between the actual weights and the estimated weights. The results are shown below for one contestant.

Cow	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Actual weight, kg	614	1105	718	1001	889	770	682
Estimated weight, kg	700	1500	850	1400	750	900	800

(a) Calculate Spearman's rank correlation coefficient for this data set. [5]

(b) Stating your hypotheses clearly, determine whether or not there is evidence at the 5% significance level of a positive association between the actual weights and the weights as estimated by this contestant. [3]

(c) One of the game show judges says, "This contestant was good at guessing the weights of the cows." Comment on this statement. [1]

TURN OVER

5. A life insurance saleswoman investigates the number of policies she sells per day. The results for a random sample of 50 days are shown in the table below.

Number of policies sold	0	1	2	3	4	5	6
Number of days	2	2	9	12	15	9	1

She sees the same fixed number of clients each day. She would like to know whether the binomial distribution with parameters 6 and 0·6 is a suitable model for the number of policies she sells per day.

(a) State suitable hypotheses for 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 policies sold	0	1	2	3	4	5	6
Observed	2	2	9	12	15	9	1
Expected	0·205	1·843	6·912	d	e	9·331	2·333

(i) Calculate the values of d and e.

(ii) Carry out the test using a 10% level of significance and draw a conclusion in context. [10]

(c) What do the parameters 6 and 0·6 mean in this context? [1]

6. A student, considering options for the future, collects data on education and salary. The table below shows the highest level of education attained and the salary bracket of a random sample of 664 people.

	Fewer than 5 GCSE	**5 or more GCSE**	**3 A Levels**	**University degree**	**Post graduate qualification**	**Total**
Less than £20 000	18	32	20	28	10	108
£20 000 to £60 000	50	95	112	155	50	462
More than £60 000	3	22	29	35	5	94
Total	71	149	161	218	65	664

By conducting a chi-squared test for independence, the student investigates the relationship between the highest level of education attained and the salary earned.

(a) State the null and alternative hypotheses. [1]

(b) The table below shows the expected values. Calculate the value of k. [2]

Expected values	Fewer than 5 GCSE	5 or more GCSE	3 A Levels	University degree	Post graduate qualification
Less than £20 000	k	24·23	26·19	35·46	10·57
£20 000 to £60 000	49·40	103·67	112·02	151·68	45·23
More than £60 000	10·05	21·09	22·79	30·86	9·20

(c) The following computer output is obtained. Calculate the values of m and n. [2]

Chi Squared Contributions	Fewer than 5 GCSE	5 or more GCSE	3 A Levels	University degree	Post graduate qualification
Less than £20 000	3·604530799	m	1·46165	1·5686	0·03098
£20 000 to £60 000	0·007272735	0·72535	4E-06	0·07264	0·50396
More than £60 000	4·946619863	0·03897	1·69081	0·55498	n

```
X-squared = 19·61301, df = 8, p-value = 0·0119
```

(d) (i) Without carrying out any further calculations, explain how `X-squared = 19·61301` (the χ^2 test statistic) was calculated.

(ii) Comment on the values in the "Fewer than 5 GCSE" column of the table in part *(c)*. [2]

(e) The student says that the highest levels of education lead to the highest paying jobs. Comment on the accuracy of the student's statement. [1]

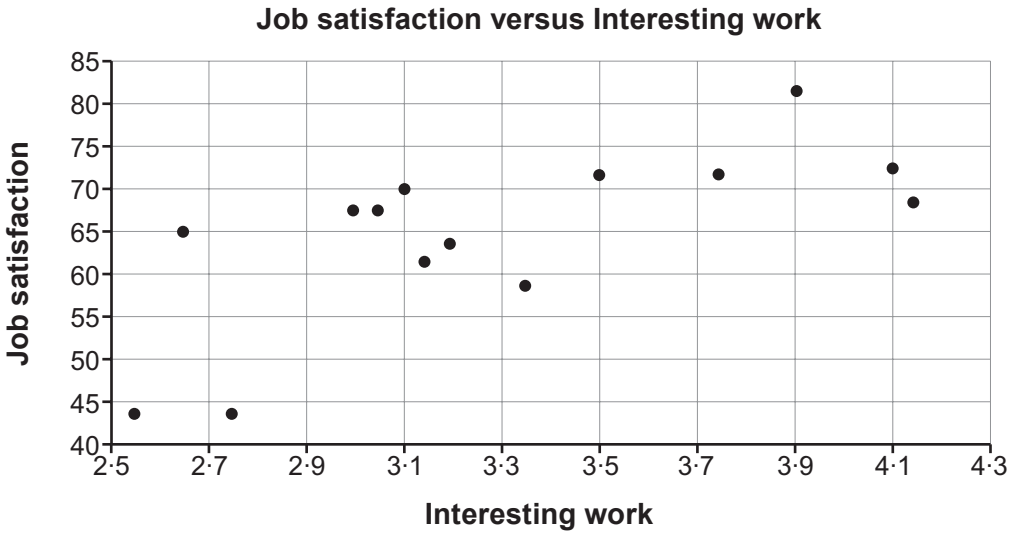
TURN OVER

7. A university professor conducted some research into factors that affect job satisfaction. The four factors considered were *Interesting work*, *Good wages*, *Job security* and *Appreciation of work done*.

The professor interviewed workers at 14 different companies and asked them to rate their companies on each of the factors. The workers' ratings were averaged to give each company a score out of 5 on each factor.

Each company was also given a score out of 100 for *Job satisfaction*.

The following graph shows the part of the research concerning *Job Satisfaction* versus *Interesting work*.



(a) Calculate the equation of the least squares regression line of *Job satisfaction* (y) on *Interesting work* (x), given the following summary statistics. [5]

$$\sum x = 46{\cdot}2, \qquad \sum y = 898, \qquad S_{xx} = 3{\cdot}48$$

$$S_{xy} = 49{\cdot}45, \qquad S_{yy} = 1437{\cdot}714, \qquad n = 14$$

(b) Give two reasons why it would be inappropriate for the professor to use this equation to calculate the score for *Interesting work* from a *Job satisfaction* score of 90. [2]

END OF PAPER

BLANK PAGE