OCR
Oxford Cambridge and RSA

AS Level Further Mathematics A

Y532/01 Statistics

Thursday 17 May 2018 – Afternoon

Time allowed: 1 hour 15 minutes

7065367946

You must have:
- Printed Answer Booklet
- Formulae AS level Further Mathematics A

You may use:
- a scientific or graphical calculator

INSTRUCTIONS
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Complete the boxes provided on the Printed Answer Booklet with your name, centre number and candidate number.
- Answer **all** the questions.
- **Write your answer to each question in the space provided in the Printed Answer Booklet.** If additional space is required, use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- Do **not** write in the barcodes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.
- The acceleration due to gravity is denoted by $g\,\text{m}\,\text{s}^{-2}$. Unless otherwise instructed, when a numerical value is needed, use $g = 9.8$.

INFORMATION
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- **You are reminded of the need for clear presentation in your answers.**
- The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **4** pages.

1 A book reviewer estimates that the probability that he receives a delivery of books to review on any one weekday is 0.1. The first weekday in September on which he receives a delivery of books to review is the Xth weekday of September.

(i) State an assumption needed for X to be well modelled by a geometric distribution. **[1]**

(ii) Find $\mathrm{P}(X = 11)$. **[2]**

(iii) Find $\mathrm{P}(X \leqslant 8)$. **[2]**

(iv) Find $\mathrm{Var}(X)$. **[2]**

(v) Give a reason why a geometric distribution might not be an appropriate model for the first weekday in a calendar year on which the reviewer receives a delivery of books to review. **[1]**

2 The probability distribution for the discrete random variable W is given in the table.

w	1	2	3	4
$\mathrm{P}(W = w)$	0.25	0.36	x	x^2

(i) Show that $\mathrm{Var}(W) = 0.8571$. **[7]**

(ii) Find $\mathrm{Var}(3W + 6)$. **[1]**

3 In the manufacture of fibre optical cable (FOC), flaws occur randomly. Whether any point on a cable is flawed is independent of whether any other point is flawed. The number of flaws in 100 m of FOC of standard diameter is denoted by X.

(i) State a further assumption needed for X to be well modelled by a Poisson distribution. **[1]**

Assume now that X can be well modelled by the distribution Po(0.7).

(ii) Find the probability that in 300 m of FOC of standard diameter there are exactly 3 flaws. **[2]**

The number of flaws in 100 m of FOC of a larger diameter has the distribution Po(1.6).

(iii) Find the probability that in 200 m of FOC of standard diameter and 100 m of FOC of the larger diameter the total number of flaws is at least 4. **[3]**

4 Judith believes that mathematical ability and chess-playing ability are related. She asks 20 randomly chosen chess players, with known British Chess Federation (BCF) ratings X, to take a mathematics aptitude test, with scores Y. The results are summarised as follows.

$$n = 20,\ \Sigma x = 3600,\ \Sigma x^2 = 660\,500,\ \Sigma y = 1440,\ \Sigma y^2 = 105\,280,\ \Sigma xy = 260\,990$$

(i) Calculate the value of Pearson's product-moment correlation coefficient r. **[2]**

(ii) State an assumption needed to be able to carry out a significance test on the value of r. **[1]**

(iii) Assume now that the assumption in part **(ii)** is valid. Test at the 5% significance level whether there is evidence that chess players with higher BCF ratings are better at mathematics. **[4]**

(iv) There are two different grading systems for chess players, the BCF system and the international ELO system. The two sets of ratings are related by

$$\text{ELO rating} = 8 \times \text{BCF rating} + 650.$$

Magnus says that the experiment should have used ELO ratings instead of BCF ratings. Comment on Magnus's suggestion. **[1]**

5 **(i)** A team of 9 is chosen at random from a class consisting of 8 boys and 12 girls.
Find the probability that the team contains no more than 3 girls. **[4]**

(ii) A group of n people, including Mr and Mrs Laplace, are arranged at random in a line. The probability that Mr and Mrs Laplace are placed next to each other is less than 0.1. Find the smallest possible value of n. **[4]**

6 **In this question you must show detailed reasoning.**

The random variable T has a binomial distribution. It is known that $\mathrm{E}(T) = 5.625$ and the standard deviation of T is 1.875. Find the values of the parameters of the distribution. **[5]**

7 An environmentalist measures the mean concentration, c milligrams per litre, of a particular chemical in a group of rivers, and the mean mass, m pounds, of fish of a certain species found in those rivers. The results are given in the table.

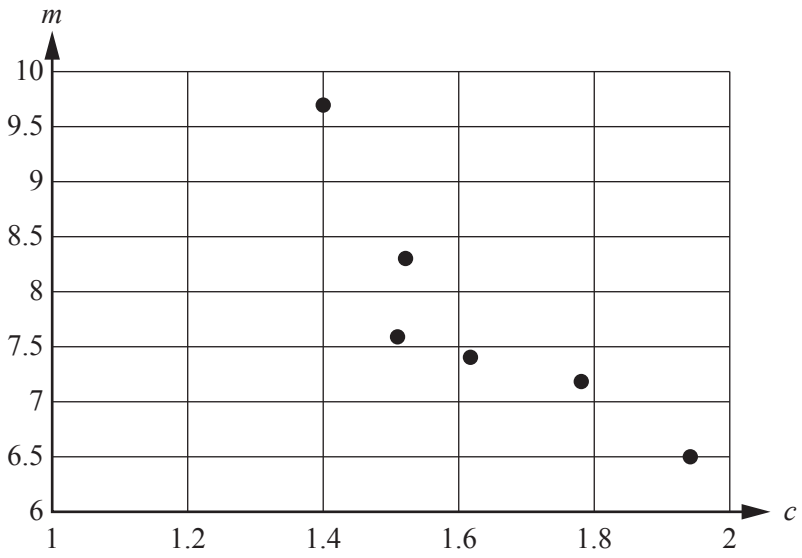
c	1.94	1.78	1.62	1.51	1.52	1.4
m	6.5	7.2	7.4	7.6	8.3	9.7

(i) State which, if either, of m and c is an independent variable. **[1]**

(ii) Calculate the equation of the least squares regression line of c on m. **[3]**

(iii) State what effect, if any, there would be on your answer to part **(ii)** if the masses of the fish had been recorded in kilograms rather than pounds. (1 kg $\approx$ 2.2 pounds.) **[1]**

(iv) The data is illustrated in the scatter diagram. Explain what is meant by 'least squares', illustrating your answer using the copy of this diagram in the Printed Answer Booklet. **[3]**



8 The table shows the results of a random sample drawn from a population which is thought to have the distribution U(20).

Range	$1 \leqslant x \leqslant 8$	$9 \leqslant x \leqslant 12$	$13 \leqslant x \leqslant 20$
Observed frequency	12	y	$28 - y$

Find the range of values of y for which the data are not consistent with the distribution at the 5% significance level. **[9]**

END OF QUESTION PAPER