OCR
Oxford Cambridge and RSA

A Level Further Mathematics A
Y542/01 Statistics

Practice Paper – Set 2
Time allowed: 1 hour 30 minutes

You must have:
- Printed Answer Booklet
- Formulae A 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, you should use the lined page(s) at the end of the Printed Answer 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 **75**.
- 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 **8** pages.

Answer **all** the questions.

1 An experiment involves releasing a coin on a sloping plane so that it slides down the slope and then slides along a horizontal plane at the bottom of the slope before coming to rest. The angle $\theta°$ of the sloping plane is varied, and for each value of θ, the distance d cm the coin slides on the horizontal plane is recorded. A scatter diagram to illustrate the results of the experiment is shown below, together with the least squares regression line of d on θ.
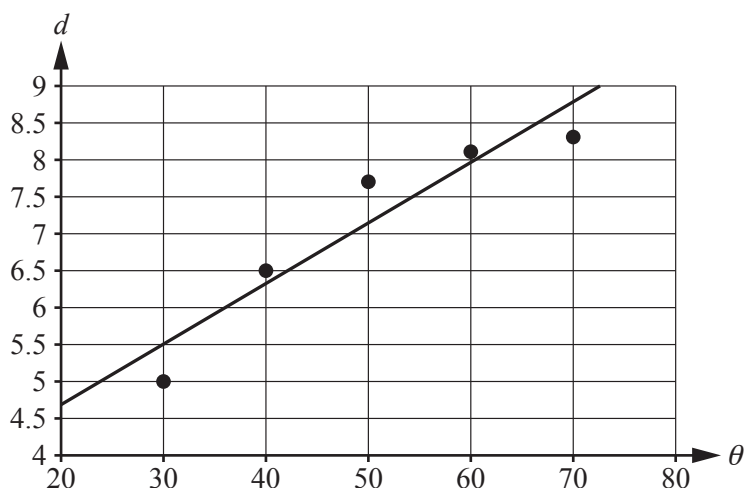
   (i) State which two of the following correctly describe the variable θ.

   Controlled variable   Correlation coefficient

   Dependent variable   Independent variable

   Response variable   Regression coefficient   **[1]**

The least squares regression line of d on θ has equation $d = 1.96 + 0.11\theta$.

   (ii) Use the diagram in the Printed Answer Booklet to explain the term "least squares".   **[2]**

   (iii) State what difference, if any, it would make to the equation of the regression line if d were measured in inches rather than centimetres. (1 inch ≈ 2.54 cm).   **[1]**

2 Shooting stars occur randomly, independently of one another and at a constant average rate of 12.0 per hour.
   On each of a series of randomly chosen clear nights I look for shooting stars for 20 minutes at a time.
   A successful night is a night on which I see at least 8 shooting stars in a 20-minute period.
   From tomorrow, I will count the number, X, of nights on which I look for shooting stars, up to and including the first successful night.

   Find $E(X)$.   **[7]**

3 A discrete random variable X has the distribution U(11).
   The mean of 50 observations of X is denoted by $\overline{X}$.

   Use an approximate method, which should be justified, to find $P(\overline{X} \leqslant 6.10)$.   **[7]**

4 A survey is carried out into the length of time for which customers wait for a response on a telephone helpline. A statistician who is analysing the results of the survey starts by modelling the waiting time, x minutes, by an exponential distribution with probability density function (PDF)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0, \\ 0 & x < 0. \end{cases}$$

(i) In this question you must show detailed reasoning.

The mean waiting time is found to be 5.0 minutes. Show that $\lambda = 0.2$. **[4]**

(ii) Use the model to calculate the probability that a customer has to wait longer than 20 minutes for a response. **[2]**

In practice it is found that no customer waits for more than 15 minutes for a response. The statistician constructs an improved model to incorporate this fact.

(iii) On the diagram in the Printed Answer Booklet, sketch the following, labelling the curves clearly:

(a) the PDF of the model using the exponential distribution, **[1]**

(b) a possible PDF for the improved model. **[2]**

5 Hal designs a 4-edged spinner with edges labelled 1, 2, 3 and 4. He intends that the probability that the spinner will land on any edge should be proportional to the number on that edge. He spins the spinner 20 times and on each spin he records the number of the edge on which it lands. The results are shown in the table.

Edge number	1	2	3	4
Frequency	3	7	4	6

Test at the 10% significance level whether the results are consistent with the intended probabilities. **[8]**

6 A bag contains 7 red counters and 5 blue counters.

  (i) Fred chooses 4 counters at random, without replacement. Show that the probability that Fred chooses exactly 2 red counters is $\frac{14}{33}$. **[3]**

  (ii) Lina chooses 4 counters at random from the bag, records whether or not exactly 2 red counters are chosen, and returns the counters to the bag. She carries out this experiment 99 times.

    (a) Find the mean of the number of experiments that result in choosing exactly 2 red counters. **[1]**

    (b) Find the variance of the number of experiments that result in choosing exactly 2 red counters. **[1]**

  (iii) Alex arranges all 12 counters in a random order in a straight line.

    A is the event: no two blue counters are next to one another.

    B is the event: all the blue counters are next to one another.

    Find $P(A \cup B)$. **[5]**

7 The table shows the values of 5 observations of bivariate data (x, y).

x	4.6	5.9	6.5	7.8	8.3
y	15.6	10.8	10.4	10.1	9.7

$$n = 5, \ \Sigma x = 33.1, \ \Sigma y = 56.6, \ \Sigma x^2 = 227.95, \ \Sigma y^2 = 664.26, \ \Sigma xy = 362.37$$

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

  (ii) State what this value of r tells you about a scatter diagram illustrating the data. **[1]**

  (iii) Test at the 5% significance level whether there is association between x and y. **[5]**

  (iv) State the value of Spearman's rank correlation coefficient r_s for the data. **[1]**

  (v) State whether r, r_s, or both or neither is changed when the values of x are replaced by

    (a) $3x - 2$, **[1]**

    (b) $\sqrt{x}$. **[1]**

8 In an experiment to investigate the effect of background music in carrying out work, ten students were each given a task. Five of the students did the task in silence and the other five did the task with background music. The scores on the tasks were as follows.

Silence	43	46	55	58	61
Background music	19	31	38	52	70

(i) Use a Wilcoxon rank-sum test to test at the 10% level whether the presence of background music affects scores. **[7]**

(ii) A statistician suggests that the experiment is redesigned so that each student takes one task in silence and another task with background music. The differences in the test scores would then be analysed using a paired-sample method. State an advantage in redesigning the experiment in this way. **[1]**

9 The continuous random variable C has the distribution $N(\mu, \sigma^2)$. The sum of a random sample of 16 observations of C is 224.0.

(i) Find an unbiased estimate of μ. **[1]**

(ii) It is given that an unbiased estimate of σ^2 is 0.24. Find the value of Σc^2. **[3]**

D is the sum of 10 independent observations of C.

(iii) Explain whether D has a normal distribution. **[1]**

The continuous random variable F is normally distributed with mean 15.0, and it is known that $P(F < 13.2) = 0.115$.

(iv) Use the unbiased estimates of μ and σ^2 to find $P(D + F > 157.0)$. **[6]**

END OF QUESTION PAPER

6

BLANK PAGE

BLANK PAGE