Diploma Programme
Programme du diplôme
Programa del Diploma

Mathematics
Higher level
Paper 3 – statistics and probability

Friday 18 November 2016 (morning)

1 hour

Instructions to candidates

- Do not open this examination paper until instructed to do so.
- Answer all the questions.
- Unless otherwise stated in the question, all numerical answers should be given exactly or correct to three significant figures.
- A graphic display calculator is required for this paper.
- A clean copy of the **mathematics HL and further mathematics HL formula booklet** is required for this paper.
- The maximum mark for this examination paper is **[60 marks]**.

8816 – 7210

Please start each question on a new page. Full marks are not necessarily awarded for a correct answer with no working. Answers must be supported by working and/or explanations. In particular, solutions found from a graphic display calculator should be supported by suitable working. For example, if graphs are used to find a solution, you should sketch these as part of your answer. Where an answer is incorrect, some marks may be given for a correct method, provided this is shown by written working. You are therefore advised to show all working.

1. [Maximum mark: 17]

In this question you may assume that these data are a random sample from a bivariate normal distribution, with population product moment correlation coefficient ρ.
Richard wishes to do some research on two types of exams which are taken by a large number of students. He takes a random sample of the results of 10 students, which are shown in the following table.

Student	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Exam 1	51	70	10	22	99	33	45	8	65	82
Exam 2	52	64	8	25	90	43	50	50	70	50

(a) For these data find the product moment correlation coefficient, r. [2]

Using these data, it is decided to test, at the 1% level, the null hypothesis $H_0 : \rho = 0$ against the alternative hypothesis $H_1 : \rho > 0$.

(b) (i) State the distribution of the test statistic (including any parameters).

(ii) Find the p-value for the test.

(iii) State the conclusion, in the context of the question, with the word “correlation” in your answer. Justify your answer. [6]

Richard decides to take the exams himself. He scored 11 on Exam 1 but his result on Exam 2 was lost.

(c) Using a suitable regression line, find an estimate for his score on Exam 2, giving your answer to the nearest integer. [3]

Caroline believes that the population mean mark on Exam 2 is 6 marks higher than the population mean mark on Exam 1. Using the original data from the 10 students, it is decided to test, at the 5% level, this hypothesis against the alternative hypothesis that the mean of the differences, $d =$ exam 2 mark – exam 1 mark, is less than 6 marks.

(d) (i) State the distribution of your test statistic (including any parameters).

(ii) Find the p-value.

(iii) State the conclusion, justifying the answer. [6]

2. [Maximum mark: 17]

John rings a church bell 120 times. The time interval, T_i , between two successive rings is a random variable with mean of 2 seconds and variance of $\frac{1}{9}$ seconds^2.

Each time interval, T_i, is independent of the other time intervals. Let $X = \sum_{i=1}^{119} T_i$ be the total time between the first ring and the last ring.

(a) Find

(i) $\mathrm{E}(X)$;

(ii) $\mathrm{Var}(X)$. [3]

(b) Explain why a normal distribution can be used to give an approximate model for X. [2]

(c) Use this model to find the values of A and B such that $\mathrm{P}(A < X < B) = 0.9$, where A and B are symmetrical about the mean of X. [7]

The church vicar subsequently becomes suspicious that John has stopped coming to ring the bell and that he is letting his friend Ray do it. When Ray rings the bell the time interval, T_i, has a mean of 2 seconds and variance of $\frac{1}{25}$ seconds^2.

The church vicar makes the following hypotheses:

H_0: Ray is ringing the bell; H_1: John is ringing the bell.

He records four values of X. He decides on the following decision rule:

If $236 \le X \le 240$ for all four values of X he accepts H_0, otherwise he accepts H_1.

(d) Calculate the probability that he makes a Type II error. [5]

3. [Maximum mark: 15]

Alun answers mathematics questions and checks his answer after doing each one. The probability that he answers any question correctly is always $\frac{6}{7}$, independently of all other questions. He will stop for coffee immediately following a second incorrect answer. Let X be the number of questions Alun answers before he stops for coffee.

(a) (i) State the distribution of X, including its parameters.

(ii) Calculate $\mathrm{E}(X)$.

(iii) Calculate $\mathrm{P}(X = 5)$. [6]

Nic answers mathematics questions and checks his answer after doing each one. The probability that he answers any question correctly is initially $\frac{6}{7}$. After his first incorrect answer, Nic loses confidence in his own ability and from this point onwards, the probability that he answers any question correctly is now only $\frac{4}{7}$.

Both before and after his first incorrect answer, the result of each question is independent of the result of any other question. Nic will also stop for coffee immediately following a second incorrect answer. Let Y be the number of questions Nic answers before he stops for coffee.

(b) (i) Calculate $\mathrm{E}(Y)$.

(ii) Calculate $\mathrm{P}(Y = 5)$. [9]

4. [Maximum mark: 11]

Two independent discrete random variables X and Y have probability generating functions $G(t)$ and $H(t)$ respectively. Let $Z = X + Y$ have probability generating function $J(t)$.

(a) Write down an expression for $J(t)$ in terms of $G(t)$ and $H(t)$. [1]

(b) By differentiating $J(t)$, prove that

(i) $\mathrm{E}(Z) = \mathrm{E}(X) + \mathrm{E}(Y)$;

(ii) $\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. [10]