Diploma Programme
Programme du diplôme
Programa del Diploma

Mathematics
Higher level
Paper 3 – statistics and probability

Thursday 15 November 2018 (afternoon)

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 **[50 marks]**.

8818 – 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: 9]

Two independent random variables X and Y follow Poisson distributions.
Given that $\mathrm{E}(X) = 3$ and $\mathrm{E}(Y) = 4$, calculate

(a) $\mathrm{E}(2X + 7Y)$; [2]

(b) $\mathrm{Var}(4X - 3Y)$; [3]

(c) $\mathrm{E}(X^2 - Y^2)$. [4]

2. [Maximum mark: 9]

The times t, in minutes, taken by a random sample of 75 workers of a company to travel to work can be summarized as follows

$$\sum t = 2165\,,\ \sum t^2 = 76475\,.$$

Let T be the random variable that represents the time taken to travel to work by a worker of this company.

(a) Find unbiased estimates of

(i) the mean of T;

(ii) the variance of T. [3]

(b) Assuming that T is normally distributed, find

(i) the 90% confidence interval for the mean time taken to travel to work by the workers of this company;

(ii) the 95% confidence interval for the mean time taken to travel to work by the workers of this company. [3]

Before seeing these results the managing director believed that the mean time was 26 minutes.

(c) Explain whether your answers to part (b) support her belief. [3]

3. [Maximum mark: 21]

(a) Mr Sailor owns a fish farm and he claims that the weights of the fish in one of his lakes have a mean of 550 grams and standard deviation of 8 grams.

Assume that the weights of the fish are normally distributed and that Mr Sailor's claim is true.

(i) Find the probability that a fish from this lake will have a weight of more than 560 grams.

(ii) The maximum weight a hand net can hold is 6 kg. Find the probability that a catch of 11 fish can be carried in the hand net. [6]

(b) Kathy is suspicious of Mr Sailor's claim about the mean and standard deviation of the weights of the fish. She collects a random sample of fish from this lake whose weights are shown in the following table.

Fish	A	B	C	D	E	F	G	H
Weight (g)	545	554	548	551	558	541	543	549

Using these data, test at the 5% significance level the null hypothesis $H_0: \mu = 550$ against the alternative hypothesis $H_1: \mu < 550$, where μ grams is the population mean weight.

(i) State the distribution of your test statistic, including the parameter.

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

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

(c) Kathy decides to use the same fish sample to test at the 5% significance level whether or not there is a positive association between the weights and the lengths of the fish in the lake. The following table shows the lengths of the fish in the sample. The lengths of the fish can be assumed to be normally distributed.

Fish	A	B	C	D	E	F	G	H
Length (mm)	351	365	355	353	357	349	348	354

(i) State suitable hypotheses for the test.

(ii) Find the product-moment correlation coefficient r.

(iii) State the p-value and interpret it in this context. [6]

(d) Use an appropriate regression line to estimate the weight of a fish with length 360 mm. [3]

4. [Maximum mark: 11]

Let $X \sim \mathrm{Geo}(p)$.

(a) Show that the probability generating function of X is given by

$$G_X(t) = \frac{pt}{1-qt} \text{ where } q = 1 - p. \quad [3]$$

(b) Hence prove that $\mathrm{E}(X) = \frac{1}{p}$. [4]

(c) Find the probability generating function of the random variable $Y = 2X + 3$. [4]