Diploma Programme
Programme du diplôme
Programa del Diploma

Mathematics
Higher level
Paper 3 – statistics and probability

Thursday 16 November 2017 (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]**.

8817 – 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: 7]

A continuous random variable T has a probability density function defined by

$$f(t)=\begin{cases}\dfrac{t\left(4-t^2\right)}{4}, & 0\le t\le 2\\ 0, & \text{otherwise}\end{cases}.$$

(a) Find the cumulative distribution function $F(t)$, for $0\le t\le 2$. [3]

(b) (i) Sketch the graph of $F(t)$ for $0\le t\le 2$, clearly indicating the coordinates of the endpoints.

(ii) Given that $\mathrm{P}(T<a)=0.75$, find the value of a. [4]

2. [Maximum mark: 8]

Anne is a farmer who grows and sells pumpkins. Interested in the weights of pumpkins produced, she records the weights of eight pumpkins and obtains the following results in kilograms.

$$7.7 \quad 7.5 \quad 8.4 \quad 8.8 \quad 7.3 \quad 9.0 \quad 7.8 \quad 7.6$$

Assume that these weights form a random sample from a $\mathrm{N}\left(\mu,\, \sigma^2\right)$ distribution.

(a) Determine unbiased estimates for μ and σ^2. [3]

(b) Anne claims that the mean pumpkin weight is 7.5 kilograms. In order to test this claim, she sets up the null hypothesis $\mathrm{H}_0 : \mu=7.5$.

(i) Use a two-tailed test to determine the p-value for the above results.

(ii) Interpret your p-value at the 5% level of significance, justifying your conclusion. [5]

3. [Maximum mark: 12]

A random variable X is distributed with mean μ and variance σ^2. Two independent random samples of sizes n_1 and n_2 are taken from the distribution of X. The sample means are $\overline{X_1}$ and $\overline{X_2}$ respectively.

(a) Show that $U = a\overline{X_1} + (1-a)\overline{X_2}$, $a \in \mathbb{R}$, is an unbiased estimator of μ. [3]

(b) (i) Show that $\mathrm{Var}(U) = a^2\dfrac{\sigma^2}{n_1} + (1-a)^2\dfrac{\sigma^2}{n_2}$.

(ii) Find, in terms of n_1 and n_2, an expression for a which gives the most efficient estimator of this form.

(iii) Hence find an expression for the most efficient estimator and interpret the result. [9]

4. [Maximum mark: 8]

The random variables U, V follow a bivariate normal distribution with product moment correlation coefficient ρ.

(a) State suitable hypotheses to investigate whether or not U, V are independent. [2]

A random sample of 12 observations on U, V is obtained to determine whether there is a correlation between U and V. The sample product moment correlation coefficient is denoted by r. A test to determine whether or not U, V are independent is carried out at the 1% level of significance.

(b) Find the least value of $|r|$ for which the test concludes that $\rho \neq 0$. [6]

5. [Maximum mark: 15]

The random variable X follows a Poisson distribution with mean λ. The probability generating function of X is given by $G_X(t) = \mathrm{e}^{\lambda(t-1)}$.

(a) (i) Find expressions for $G'_X(t)$ and $G''_X(t)$.

(ii) Hence show that $\mathrm{Var}(X) = \lambda$. [5]

The random variable Y, independent of X, follows a Poisson distribution with mean μ.

(b) By considering the probability generating function, $G_{X+Y}(t)$, of $X+Y$, show that $X+Y$ follows a Poisson distribution with mean $\lambda + \mu$. [3]

(c) (i) Show that $\mathrm{P}(X = x \mid X + Y = n) = \binom{n}{x}\left(\frac{\lambda}{\lambda+\mu}\right)^x\left(1-\frac{\lambda}{\lambda+\mu}\right)^{n-x}$, where n, x are non-negative integers and $n \geq x$.

(ii) Identify the probability distribution given in part (c)(i) and state its parameters. [7]