Diploma Programme
Programme du diplôme
Programa del Diploma

Mathematics
Higher level
Paper 3 – statistics and probability

Tuesday 10 November 2020 (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]**.

8820 – 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: 8]

Peter, Dan and Eva decide to observe the colour of each car as it enters their school. The colours of the cars are assumed to be independent of one another.

The probability that a car entering their school is red is $\frac{1}{12}$.

(a) Peter observes 20 cars. Find the probability that at least three of these cars are red. [2]

(b) Dan observes cars until a red one enters. Find the probability that he observes at least six cars. [2]

(c) Eva observes cars until three red cars enter.

(i) Find the probability that Eva observes exactly 20 cars.

(ii) Find the expected number of cars that Eva observes. [4]

2. [Maximum mark: 8]

A discrete random variable X has a probability generating function given by $G_X(t) = \frac{2+t}{3}$.

(a) Write down the value of $\mathrm{P}(X = 1)$. [1]

A discrete random variable Y has a probability generating function given by

$$G_Y(t) = \frac{kt+1}{k^2 - t^2}, \text{ where } k > 0.$$

(b) Show that $k = 2$. [3]

(c) Given that X and Y are independent, find

(i) $G_{X+Y}(t)$, the probability generating function for $X + Y$;

(ii) $\mathrm{E}(X + Y)$. [4]

3. [Maximum mark: 15]

A shop sells carrots and broccoli. The weights of carrots can be modelled by a normal distribution with variance 25 grams^2 and the weights of broccoli can be modelled by a normal distribution with variance 80 grams^2. The shopkeeper claims that the mean weight of carrots is 130 grams and the mean weight of broccoli is 400 grams.

(a) Assuming that the shopkeeper's claim is correct, find the probability that the weight of six randomly chosen carrots is more than two times the weight of one randomly chosen broccoli. [6]

Dong Wook decides to investigate the shopkeeper's claim that the mean weight of carrots is 130 grams. He plans to take a random sample of n carrots in order to calculate a 98% confidence interval for the population mean weight.

(b) Find the least value of n required to ensure that the width of the confidence interval is less than 2 grams. [3]

Anjali thinks the mean weight, μ grams, of the broccoli is less than 400 grams. She decides to perform a hypothesis test, using a random sample of size 8. Her hypotheses are

$$H_0 : \mu = 400;\ H_1 : \mu < 400.$$

She decides to reject H_0 if the sample mean is less than 395 grams.

(c) Find the significance level for this test. [3]

(d) Given that the weights of the broccoli actually follow a normal distribution with mean 392 grams and variance 80 grams^2, find the probability of Anjali making a Type II error. [3]

4. [Maximum mark: 19]

A discrete random variable X has a probability distribution given by

x	0	1	2
$\mathrm{P}(X = x)$	p	$3p$	$1 - 4p$

where $0 < p < \frac{1}{4}$.

(a) Find an expression for $\mathrm{E}(X)$, in terms of p. [2]

(b) Show that $\mathrm{Var}(X) = p(7 - 25p)$. [3]

Christine and Sarah want to estimate the value of p. They take a random sample of n observations of X.

(c) Christine calculates the sample mean, $\overline{X}$, and proposes $C = \frac{2 - \overline{X}}{5}$ as an estimator for p.

(i) Show that C is an unbiased estimator for p.

(ii) Find $\mathrm{Var}(C)$. [5]

(d) Sarah counts the number of zeros, Y, in the sample of size n. She proposes $S = \frac{Y}{n}$ as an estimator for p.

(i) Write down the distribution for Y.

(ii) Show that S is an unbiased estimator for p.

(iii) Show that $\mathrm{Var}(S) = \frac{p(1-p)}{n}$. [5]

(e) (i) Sketch a graph of the ratio $\frac{\mathrm{Var}(C)}{\mathrm{Var}(S)}$ for $0 < p < \frac{1}{4}$, indicating clearly the scale on the y-axis.

(ii) Hence determine, with a reason, which of C or S is the more efficient estimator. [4]