GCE Examinations

Statistics
Module S2

Advanced Subsidiary / Advanced Level

Paper E

Time: 1 hour 30 minutes

Instructions and Information

Candidates may use any calculator except those with a facility for symbolic algebra and / or calculus.

Full marks may be obtained for answers to ALL questions.

Mathematical and statistical formulae and tables are available.

This paper has 7 questions.

Advice to Candidates

You must show sufficient working to make your methods clear to an examiner. Answers without working will gain no credit.



Written by Shaun Armstrong & Chris Huffer

© Solomon Press

These sheets may be copied for use solely by the purchaser's institute.

1. *(a)* State one advantage and one disadvantage in using a census rather than a sample survey in statistical work.

(2 marks)

(b) Give an example of a situation in which you would choose to use a census rather than a sample survey and explain why.

(2 marks)

---

2. An advert for *Tatty's Crisps* claims that 1 in 10 bags contain a free scratchcard game.

Tatty's Crisps can be bought in a Family Pack containing 10 bags. Find the probability that the bags in one of these Family Packs contain

(a) no scratchcards, **(2 marks)**

(b) more than 2 scratchcards. **(2 marks)**

Tatty's Crisps can also be bought wholesale in boxes containing 50 bags. A pub Landlord notices that her customers only found 2 scratchcards in the crisps from one of these boxes.

(c) Stating your hypotheses clearly, test at the 10% level of significance whether or not this gives evidence of there being fewer free scratchcards than is claimed by the advert.

(4 marks)

---

3. A class of children are each asked to draw a line that they think is 10 cm long without using a ruler. The teacher models how many centimetres each child's line is longer than 10 cm by the random variable X and believes that X has the following probability density function:

$$f(x) = \begin{cases} \frac{1}{8}, & {}^{-}4 \le x \le 4, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Write down the name of this distribution. **(1 mark)**

(b) Define fully the cumulative distribution function $F(x)$ of X. **(4 marks)**

(c) Calculate the proportion of children making an error of less than 15% according to this model.

(3 marks)

(d) Give two reasons why this may not be a very suitable model. **(2 marks)**

© *Solomon Press*

4. A bag contains 40 beads of the same shape and size. The ratio of red to green to blue beads is 1 : 3 : 4 and there are no beads of any other colour.

In an experiment, a bead is picked at random, its colour noted and the bead replaced in the bag. This is done ten times.

(a) Suggest a suitable distribution for modelling the number of times a blue bead is picked out and give the value of any parameters needed.

(2 marks)

(b) Explain why this distribution would not be suitable if the beads were not replaced in the bag.

(1 mark)

(c) Find the probability that of the ten beads picked out

(i) five are blue,

(ii) at least one is red. **(6 marks)**

The experiment is repeated, but this time a bead is picked out and replaced n times.

(d) Find in the form $a^n < b$, where a and b are exact fractions, the condition which n must satisfy in order to have at least a 99% chance of picking out at least one red bead.

(3 marks)

---

5. A charity receives donations of more than £10 000 at an average rate of 25 per year.

Find the probability that the charity receives

(a) exactly 30 such donations in one year, **(3 marks)**

(b) less than 3 such donations in one month. **(5 marks)**

(c) Using a suitable approximation, find the probability that the charity receives more than 45 donations of more than £10 000 in the next two years.

(5 marks)

Turn over

© *Solomon Press*

6. The length of time, in tens of minutes, that patients spend waiting at a doctor's surgery is modelled by the continuous random variable T, with the following cumulative distribution function:

$$\mathrm{F}(t) = \begin{cases} 0, & t < 0, \\ \frac{1}{135}(54t + 9t^2 - 4t^3), & 0 \le t \le 3, \\ 1, & t > 3. \end{cases}$$

(a) Find the probability that a patient waits for more than 20 minutes. **(3 marks)**

(b) Show that the median waiting time is between 11 and 12 minutes. **(3 marks)**

(c) Define fully the probability density function f(t) of T. **(3 marks)**

(d) Find the modal waiting time in minutes. **(4 marks)**

(e) Give one reason why this model may need to be refined. **(1 mark)**

---

7. A student collects data on the number of bicycles passing outside his house in 5-minute intervals during one morning.

(a) Suggest, with reasons, a suitable distribution for modelling this situation. **(3 marks)**

The student's data is shown in the table below.

Number of bicycles	0	1	2	3	4	5	6 or more
Frequency	7	14	10	2	1	2	0

(b) Show that the mean and variance of these data are 1.5 and 1.58 (to 3 significant figures) respectively and explain how these values support your answer to part *(a)*.

(6 marks)

An environmental organisation declares a "car free day" encouraging the public to leave their cars at home. The student wishes to test whether or not there are more bicycles passing along his road on this day and records 16 bicycles in a half-hour period during the morning.

(c) Stating your hypotheses clearly, test at the 5% level of significance whether or not there are more than 1.5 bicycles passing along his road per 5-minute interval that morning.

(5 marks)

END

© *Solomon Press*