Discrete Random Variables – 2022 GCE Statistics Further Math A Y542

1. June/2022/Paper_ Y542/01/No.1

A researcher wishes to find people who say that they support a specific plan. Each day the researcher interviews people at random, one after the other, until they find one person who says that they support this plan. The researcher does not then interview any more people that day. The total number of people interviewed on any one day is denoted by R.

(a) Assume that in fact 1% of the population would say that they support the plan.

(i) State an appropriate distribution with which to model R, giving the value(s) of any parameter(s). **[1]**

(ii) Find $\mathrm{P}(50 < R \leqslant 150)$. **[2]**

The researcher incorrectly believes that the variance of a random variable X with **any** discrete probability distribution is given by the formula $[\mathrm{E}(X)]^2 - \mathrm{E}(X)$.

(b) Show that, for the type of distribution stated in part **(a)**, they will obtain the correct value of the variance, regardless of the value(s) of the parameter(s). **[2]**

2. June/2022/Paper_ Y542/01/No.3

In this question you must show detailed reasoning.

A discrete random variable V has the following probability distribution, where p and q are constants.

v	0	1	2	3
$\mathrm{P}(V = v)$	p	q	0.12	0.2

It is given that $\mathrm{E}(V) = \mathrm{Var}(V)$.

Determine the value of p and the value of q. **[8]**

3. June/2022/Paper_ Y542/01/No.4

The manager of a car breakdown service uses the distribution Po(2.7) to model the number of punctures, R, in a 24-hour period in a given rural area. The manager knows that, for this model to be valid, punctures must occur randomly and independently of one another.

(a) State a further assumption needed for the Poisson model to be valid. **[1]**

(b) State the value of the standard deviation of R. **[1]**

(c) Use the model to calculate the probability that, in a randomly chosen period of 168 hours, at least 22 punctures occur. **[3]**

The manager uses the distribution Po(0.8) to model the number of flat batteries in a 24-hour period in the same rural area, and he assumes that instances of flat batteries are independent of punctures. A day begins and ends at midnight, and a "bad" day is a day on which there are more than 6 instances, in total, of punctures and flat batteries.

(d) Assume first that both the manager's models are correct.

Calculate the probability that a randomly chosen day is a "bad" day. **[2]**

(e) It is found that 12 of the next 100 days are "bad" days.

Comment on whether this casts doubt on the validity of the manager's models. **[2]**