# M1S : EXERCISES 7

## PROPERTIES OF DISCRETE PROBABILITY DISTRIBUTIONS

**ASSESSED COURSEWORK**
(Deadline : Tuesday 7th March)

The Poisson distribution is a probability model that is commonly used to specify probabilities for the outcomes of counting experiments. If $X$ is a discrete random variable taking values on range $\mathbb{X} = \{0, 1, 2, ...\}$ with probability mass function $f_X$ given by

$$f_X(x) = \mathrm{P}\,[\ X = x\ ] = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, ...$$

for parameter $\lambda > 0$, then $X$ has a Poisson distribution with parameter $\lambda$, or $X \sim Poisson(\lambda)$.

(a)(i) Verify that the Poisson distribution is a valid discrete probability distribution.

*[2 MARKS]*

(ii) Suppose that $Z$ is a discrete random variable with $\mathrm{P}\,[\ Z = 0\ ] = 0$, but otherwise with $\mathrm{P}\,[\ Z = z\ ] \propto \mathrm{P}\,[\ X = z\ ]$ for $z \in \mathbb{R}$. Find the probability mass function of $Z$.

*[3 MARKS]*

(iii) Suppose now that $X_1 \sim Poisson(\lambda_1)$ and $X_2 \sim Poisson(\lambda_2)$ are **independent** random variables, so that

$$\mathrm{P}\,[\ (X_1 = x_1) \cap (X_2 = x_2)\ ] = \mathrm{P}\,[\ (X_1 = x_1)\ ]\,\mathrm{P}\,[\ (X_2 = x_2)\ ] \qquad \text{for all pairs } (x_1, x_2)$$

If random variable $Y$ is defined by $Y = X_1 + X_2$, prove that $Y \sim Poisson(\lambda_1 + \lambda_2)$.

Hint: consider the partition

$$(Y = y) = \bigcup_{x_1=0}^{\infty} ((X_1 = x_1) \cap (X_2 = y - x_1)) \qquad \text{for } y = 0, 1, 2, ...$$

*[5 MARKS]*

(b) For a discrete random variable $X$ with range $\mathbb{X}$ and probability mass function $f_X$, define the **probability generating function** (or pgf) of $X$, $G_X$, as a power series in $t$ by

$$G_X(t) = \sum_{x \in \mathbb{X}} f_X(x)\ t^x$$

for values of $t$ where the sum is convergent. The pgf is a function of $t$ that is specific to each mass function $f_X$; each $f_X$ has a unique corresponding $G_X$.

(i) Find the pgf of $X$ if $X \sim Poisson(\lambda)$.

*[4 MARKS]*

(ii) Prove that if $X_1$ and $X_2$ are independent random variables, and $Y = X_1 + X_2$, then

$$G_Y(t) = G_{X_1}(t)\ G_{X_2}(t)$$

*[6 MARKS]*

**EXERCISES** (not to be handed in)

1. Show that the function

$$\frac{1}{1+\lambda}\left(\frac{\lambda}{1+\lambda}\right)^x$$

for parameter $\lambda > 0$ is a valid probability mass function for a discrete random variable $X$ taking values on $\{0, 1, 2, ...\}$, and find the corresponding cumulative distribution function.

2. Suppose that discrete random variable $X$ has a *negative binomial* distribution, that is, the probability mass function of $X$ is given by

$$f_X(x) = \binom{x-1}{n-1}\theta^n(1-\theta)^{x-n} \qquad x = n, n+1, ...$$

and zero otherwise, for $n \geq 1$ and $0 < \theta \leq 1$. Recall that $X$ corresponds to the number of binary trials required to obtain $n$ successes.

Find the probability mass function of discrete random variable $Y$ defined by $Y = X - n$.

3. Recall that, if $X \sim$ *Hypergeometric*$(N, R, n)$, then if $N$ and $R$ are large, so that in particular $N, R >> n$, then

$$f_X(x) \approx \binom{n}{x}\theta^x(1-\theta)^{n-x} \qquad x = 0, 1, ..., n$$

and zero otherwise, then $X$ has an approximate Binomial distribution, with parameters $n$ and $\theta = \frac{R}{N}$.

Explain this result by considering sampling without replacement from a finite but large population.

4. A fair coin is tossed $n$ times. Let $X$ be the discrete random variable corresponding to the difference between the number of heads and the number of tails observed. Find the range and probability mass function of $X$.

5. If $X$ has a Geometric distribution with parameter $\theta$, show that, for $n, k \geq 1$,

$$\mathrm{P}[\, X = n + k \mid X > n \,] = \mathrm{P}[\, X = k \,]$$

This result is known as the *Lack of Memory* property.

6. For which values of $k$ and $\alpha$ are the following functions valid probability mass functions on the ranges given;

$$\text{(i) } f(x) = \frac{k}{x(x+1)}, \quad x = 1, 2, 3, ... \qquad \text{(ii) } f(x) = kx^\alpha, \quad x = 1, 2, 3, ....$$

7. Consider a sequence of Bernoulli random variables $X_1, ..., X_n$ each with parameter $\theta$ resulting from independent binary trials. Find the probability distributions of the random variables

$$\text{(i) } Y = \mathrm{Min}\,\{X_1, ..., X_n\} \qquad \text{(ii) } Z = \mathrm{Max}\,\{X_1, ..., X_n\}$$

[*Hint: find the ranges of* $Y$ *and* $Z$*, and consider* $\mathrm{P}[\, Y = 1 \,]$, $\mathrm{P}[\, Z = 0 \,]$]

8. Consider an event $E$ in sample space $\Omega$, and consider the *indicator function*, $I_E$, defined for $\omega \in \Omega$ by

$$I_E(\omega) = \begin{cases} 1 & \omega \in E \\ 0 & \omega \notin E \end{cases}$$

Show that $I_E$ defines a Bernoulli random variable.

Show furthermore that **any** discrete random variable can be expressed as a linear combination of indicator random variables.