DISCRETE DISTRIBUTIONS : WORKED EXAMPLES

1. Suppose that two systems are set in operation on day 1. The probability that system 1 fails for the first time on day n is denoted $p_1(n)$ (that is, the conditional probability, defined as a function of n, that the system fails on day n *given* that it has not failed on any preceding day); similarly, the probability that system 2 fails for the first time on day n is denoted $p_2(n)$ Let X_1 and X_2 be the discrete random variables corresponding to the days on which system 1 and system 2 fail for the first time, respectively.

Show that the probability mass functions of X_1 and X_2 are given by

$$f_{X_i}(n) = p_i(n)\prod_{k=1}^{n-1}(1-p_i(k)) \qquad n = 1,2,... \quad \text{for } i = 1,2$$

and zero elsewhere, respectively.

Suppose that, $p_1(.)$ and $p_2(.)$ are specified by

$$\begin{aligned} p_1(n) &= \frac{1}{n+1} && n = 1,2,... \\ p_2(n) &= 1 - e^{-\lambda n} && n = 1,2,... \end{aligned}$$

where λ is a positive real constant, so that p_1 is decreasing and p_2 is increasing with n.

Show that p_1 and p_2 lead to valid probability models for X_1 and X_2, and find the mass and distribution functions of X_1 and X_2.

SOLUTION :
Let D_k denote the event that system 1 fails on day k, for $k = 1,2,....$ Then, if S_n denotes the event that the first failure occurs on day n then

$$S_n = D_1' \cap D_2' \cap D_3' \cap ... \cap D_{n-1}' \cap D_n$$

so that, by the chain rule

$$\begin{aligned} P(S_n) &= P(D_1')\, P(D_2'|D_1')P(D_3'|D_1' \cap D_2')...P(D_{n-1}'|D_1' \cap D_2' \cap D_3' \cap ... \\ &\quad ... \cap D_{n-2}')P(D_n|D_1' \cap D_2' \cap D_3' \cap ... \cap D_{n-1}') \\ &= (1-p_1(1))(1-p_1(2))(1-p_1(3))...(1-p_1(n-1))p_1(n) \end{aligned}$$

as for each k, we have

$$P(D_k|D_1' \cap D_2' \cap D_3' \cap ... \cap D_{k-1}') = p_1(k) \qquad P(D_k'|D_1' \cap D_2' \cap D_3' \cap ... \cap D_{k-1}') = 1 - p_1(k)$$

as these conditional probabilities are given in the question. Hence, for $i = 1,2$,

$$f_{X_i}(n) = P[X_i = n] = P(S_n) = p_i(n)\prod_{k=1}^{n-1}(1-p_i(k)) \qquad n = 1,2,...$$

Now, we also have, using the chain rule, that

$$P[X_i > n] = P(D_1' \cap D_2' \cap D_3' \cap ... \cap D_{n-1}' \cap D_n') = \prod_{k=1}^{n}(1-p_i(k))$$

so for system 1 we have

$$P[X_1 > n] = \prod_{k=1}^{n}(1 - p_1(k)) = \prod_{k=1}^{n}\left(1 - \frac{1}{k+1}\right) = \prod_{k=1}^{n}\left(\frac{k}{k+1}\right) = \frac{1}{n+1}$$

and hence the cdf of X_1 is given by

$$F_{X_1}(n) = P[X_1 \le n] = 1 - P[X_1 > n] = \frac{n}{n+1} \qquad n = 1, 2, ...$$

which is a valid cdf, as it takes values on $[0,1]$, is non-decreasing in n, and is a right-continuous (step) function. By differencing, we have

$$\begin{aligned} f_{X_1}(n) &= P[X_1 \le n] - P[X_1 \le n-1] \\ &= (1 - P[X_1 > n]) - (1 - P[X_1 > n-1]) \\ &= \left(1 - \frac{1}{n+1}\right) - \left(1 - \frac{1}{n}\right) \\ &= \frac{1}{n(n+1)} \qquad\qquad n = 1, 2, 3, ... \end{aligned}$$

Similarly, for system 2,

$$P[X_2 > n] = \prod_{k=1}^{n}(1 - p_2(k)) = \prod_{k=1}^{n} e^{-\lambda k} = \exp\left\{-\lambda \sum_{k=1}^{n} k\right\} = \exp\left\{-\frac{\lambda}{2}n(n+1)\right\}$$

and thus

$$F_{X_2}(n) = P[X_2 \le n] = 1 - P[X_2 > n] = 1 - \exp\left\{-\frac{\lambda}{2}n(n+1)\right\} \qquad n = 1, 2, ..$$

which is again a valid cdf. Finally, we have

$$\begin{aligned} f_{X_2}(n) &= P[X_2 \le n] - P[X_2 \le n-1] \\ &= (1 - P[X_2 > n]) - (1 - P[X_2 > n-1]) \\ &= \left(1 - \exp\left\{-\frac{\lambda}{2}n(n+1)\right\}\right) - \left(1 - \exp\left\{-\frac{\lambda}{2}n(n-1)\right\}\right) \\ &= \left(1 - e^{-\lambda n}\right)\exp\left\{-\frac{\lambda}{2}n(n-1)\right\} \qquad\qquad n = 1, 2, 3, ... \end{aligned}$$

2. In a cricket ball throwing contest, a competitor is permitted as many throws as they like, and the longest of their throws is recorded. The number of throws that they take is a discrete random variable, T, taking values on range $\mathbb{T} = \{1, 2, ...\}$ with probability mass function given by

$$f_T(t) = \left(\frac{9}{10}\right)\left(\frac{1}{10}\right)^{t-1} \qquad t = 1, 2, ..$$

Let random variables $X_1, X_2, ..., X_T$ (where, note, T is not known before the contest starts) denote the measured throws of a given competitor. Suppose that the throws are mutually independent, and have identical probability distributions specified via a "reliability" function r_X that specifies the probability that a single throw exceeds a given value, that is, for $x > 0$,

$$r_X(x) = P[X_i > x] = \frac{1}{1+x} \qquad x > 0$$

for each i.

Verify that r_X specifies a valid probability model. If the number of throws taken is **known** to be t, let the longest throw recorded for a given competitor be a random variable M, so that

$$M = \max\{X_1, X_2, ..., X_t\},$$

find the reliability function for M, r_M say, given by

$$r_M(x) = P[M > x]$$

for $x > 0$.

Now, suppose that a competitor records a longest throw that is longer than a distance m. Find an expression for the conditional probability that the number of throws taken in total is equal to t, for $t = 1, 2, ...$

SOLUTION :

First, r_X specifies a valid probability model, as, because of the probability axioms, we require that

$$0 \leq r_X(x) \leq 1 \quad \text{for all } x > 0$$

and also that, here,

$$\lim_{x \to 0} r_M(x) = P[X > 0] = 1 \qquad \text{and} \qquad \lim_{x \to \infty} r_M(x) = \lim_{x \to \infty} P[X > x] = 0.$$

Now, given that $T = t$, that is that $t \geq 1$ throws were taken, we consider the event $[M \leq m]$ (the event that the largest of the t throws is less than m). We have that

$$\begin{aligned}
P[M \leq m \,|\, T = t] &= P[(X_1 \leq m) \cap (X_2 \leq m) \cap ... \cap (X_t \leq m) \,|\, T = t] && \text{(no throw longer than } m\text{)} \\
&= P[X_1 \leq m]P[X_2 \leq m]...P[X_t \leq m] && \text{(by mutual independence)} \\
&= (1 - r_X(m))(1 - r_X(m))...(1 - r_X(m)) \\
&= \left(\frac{m}{1+m}\right)^t && \text{(from above)}
\end{aligned}$$

Hence, using the theorem of total probability, using the partition given by $[T=1], [T=2], \ldots$ we have

$$
\begin{aligned}
P[M \leq m] &= \sum_{t=1}^{\infty} P[M \leq m \,|\, T=t] P[T=t] \\
&= \sum_{t=1}^{\infty} \left(\frac{m}{1+m}\right)^{t} \left(\frac{9}{10}\right) \left(\frac{1}{10}\right)^{t-1} \\
&= \left(\frac{9}{10}\right)\left(\frac{m}{1+m}\right) \sum_{s=0}^{\infty} \left(\frac{m}{10(1+m)}\right)^{s} \quad \text{where } s = t-1 \\
&= \left(\frac{9}{10}\right)\left(\frac{m}{1+m}\right) \frac{1}{\left(1 - \dfrac{m}{10(1+m)}\right)} \quad \text{(summing a geometric progression)} \\
&= \frac{9m}{10+9m}
\end{aligned}
$$

Hence

$$
r_M(m) = 1 - P[M \leq m] = \frac{10}{10+9m}.
$$

Now, given $[M > m]$, we have by Bayes Theorem

$$
P[T=t \,|\, M>m] = \frac{P[M>m \,|\, T=t] P[T=t]}{P[M>m]} = \frac{\left\{1 - \left(\dfrac{m}{m+1}\right)^{t}\right\} \left(\dfrac{9}{10}\right)\left(\dfrac{1}{10}\right)^{t-1}}{\left(\dfrac{10}{10+9m}\right)}
$$