5.1.4 Functions of Two Random Variables

Analysis of a function of two random variables is pretty much the same as for a function of a single random variable. Suppose that you have two discrete random variables X and Y, and suppose that $Z = g(X,Y)$, where $g : \mathbb{R}^2 \mapsto \mathbb{R}$. Then, if we are interested in the PMF of Z, we can write

$$\begin{aligned} P_Z(z) &= P(g(X,Y) = z) \\ &= \sum_{(x_i,y_j)\in A_z} P_{XY}(x_i,y_j), \quad \text{ where } A_z = \{(x_i,y_j) \in R_{XY} : g(x_i,y_j) = z\}. \end{aligned}$$

Note that if we are only interested in $E[g(X,Y)]$, we can directly use LOTUS, without finding $P_Z(z)$:

Law of the unconscious statistician (LOTUS) for two discrete random variables:

$$E[g(X,Y)] = \sum_{(x_i,y_j)\in R_{XY}} g(x_i,y_j)P_{XY}(x_i,y_j) \qquad (5.5)$$

Example 5.8

Linearity of Expectation: For two discrete random variables X and Y, show that $E[X+Y] = EX + EY$.

Solution

Let $g(X,Y) = X+Y$. Using LOTUS, we have

$$
\begin{aligned}
E[X+Y] &= \sum_{(x_i,y_j)\in R_{XY}} (x_i+y_j)P_{XY}(x_i,y_j) \\
&= \sum_{(x_i,y_j)\in R_{XY}} x_i P_{XY}(x_i,y_j) + \sum_{(x_i,y_j)\in R_{XY}} y_j P_{XY}(x_i,y_j) \\
&= \sum_{x_i\in R_X}\sum_{y_j\in R_Y} x_i P_{XY}(x_i,y_j) + \sum_{x_i\in R_X}\sum_{y_j\in R_Y} y_j P_{XY}(x_i,y_j) \\
&= \sum_{x_i\in R_X} x_i \sum_{y_j\in R_Y} P_{XY}(x_i,y_j) + \sum_{y_j\in R_Y} y_j \sum_{x_i\in R_X} P_{XY}(x_i,y_j) \\
&= \sum_{x_i\in R_X} x_i P_X(x_i) + \sum_{y_j\in R_Y} y_j P_Y(y_j) \qquad \text{(marginal PMF (Equation 5.1))} \\
&= EX + EY.
\end{aligned}
$$

Example 5.9

Let X and Y be two independent $Geometric(p)$ random variables. Also let $Z = X - Y$. Find the PMF of Z.

Solution

First note that since $R_X = R_Y = \mathbb{N} = \{1, 2, 3, \ldots\}$, we have $R_Z = \mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. Since $X, Y \sim Geometric(p)$, we have

$$P_X(k) = P_Y(k) = pq^{k-1}, \quad \text{for } k = 1, 2, 3, \ldots,$$

where $q = 1 - p$. We can write for any $k \in \mathbb{Z}$

$$
\begin{aligned}
P_Z(k) &= P(Z = k) \\
&= P(X - Y = k) \\
&= P(X = Y + k) \\
&= \sum_{j=1}^{\infty} P(X = Y + k | Y = j) P(Y = j) \qquad \text{(law of total probability)} \\
&= \sum_{j=1}^{\infty} P(X = j + k | Y = j) P(Y = j) \\
&= \sum_{j=1}^{\infty} P(X = j + k) P(Y = j) \qquad \text{(since } X, Y \text{ are independent)} \\
&= \sum_{j=1}^{\infty} P_X(j + k) P_Y(j).
\end{aligned}
$$

Now, consider two cases: $k \geq 0$ and $k < 0$. If $k \geq 0$, then

$$
\begin{aligned}
P_Z(k) &= \sum_{j=1}^{\infty} P_X(j+k) P_Y(j) \\
&= \sum_{j=1}^{\infty} pq^{j+k-1} pq^{j-1} \\
&= p^2 q^k \sum_{j=1}^{\infty} q^{2(j-1)} \\
&= p^2 q^k \frac{1}{1-q^2} && \big(\text{geometric sum (Equation 1.4)}\big) \\
&= \frac{p(1-p)^k}{2-p}.
\end{aligned}
$$

For $k < 0$, we have

$$
\begin{aligned}
P_Z(k) &= \sum_{j=1}^{\infty} P_X(j+k) P_Y(j) \\
&= \sum_{j=-k+1}^{\infty} pq^{j+k-1} pq^{j-1} && (\text{since } P_X(j+k) = 0 \text{ for } j < -k+1) \\
&= p^2 \sum_{j=-k+1}^{\infty} q^{k+2(j-1)} \\
&= p^2 \left[q^{-k} + q^{-k+2} + q^{-k+4} + \ldots \right] \\
&= p^2 q^{-k} \left[1 + q^2 + q^4 + \ldots \right] \\
&= \frac{p}{(1-p)^k (2-p)} && \big(\text{geometric sum (Equation 1.4)}\big).
\end{aligned}
$$

To summarize, we conclude

$$
P_Z(k) = \begin{cases} \frac{p(1-p)^{|k|}}{2-p} & k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}
$$