3.2.3 Functions of Random Variables

If X is a random variable and $Y = g(X)$, then Y itself is a random variable. Thus, we can talk about its PMF, CDF, and expected value. First, note that the range of Y can be written as

$$R_Y = \{g(x) | x \in R_X\}.$$

If we already know the PMF of X, to find the PMF of $Y = g(X)$, we can write

$$\begin{aligned} P_Y(y) &= P(Y = y) \\ &= P(g(X) = y) \\ &= \sum_{x:g(x)=y} P_X(x) \end{aligned}$$

Let's look at an example.

Example 3.16
Let X be a discrete random variable with $P_X(k) = \frac{1}{5}$ for $k = -1, 0, 1, 2, 3$. Let $Y = 2|X|$. Find the range and PMF of Y.

Solution

First, note that the range of Y is

$$\begin{aligned} R_Y &= \{2|x| \text{ where } x \in R_X\} \\ &= \{0, 2, 4, 6\}. \end{aligned}$$

To find $P_Y(y)$, we need to find $P(Y = y)$ for $y = 0, 2, 4, 6$. We have

$$\begin{aligned} P_Y(0) &= P(Y = 0) = P(2|X| = 0) \\ &= P(X = 0) = \tfrac{1}{5}; \\ P_Y(2) &= P(Y = 2) = P(2|X| = 2) \\ &= P\big((X = -1) \text{ or } (X = 1)\big) \\ &= P_X(-1) + P_X(1) = \tfrac{1}{5} + \tfrac{1}{5} = \tfrac{2}{5}; \end{aligned}$$

$$\begin{aligned} P_Y(4) &= P(Y=4) = P(2|X|=4) \\ &= P(X=2) + P(X=-2) = \tfrac{1}{5}; \\ P_Y(6) &= P(Y=6) = P(2|X|=6) \\ &= P(X=3) + P(X=-3) = \tfrac{1}{5}. \end{aligned}$$

So, to summarize,

$$P_Y(k) = \begin{cases} \frac{1}{5} & \text{for } k = 0, 4, 6 \\ \frac{2}{5} & \text{for } k = 2 \\ 0 & \text{otherwise} \end{cases}$$

Expected Value of a Function of a Random Variable (LOTUS)

Let X be a discrete random variable with PMF $P_X(x)$, and let $Y = g(X)$. Suppose that we are interested in finding EY. One way to find EY is to first find the PMF of Y and then use the expectation formula $EY = E[g(X)] = \sum_{y \in R_Y} y P_Y(y)$. But there is another way which is usually easier. It is called the law of the unconscious statistician (LOTUS).

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

$$E[g(X)] = \sum_{x_k \in R_X} g(x_k) P_X(x_k) \tag{3.2}$$

You can prove this by writing $EY = E[g(X)] = \sum_{y \in R_Y} y P_Y(y)$ in terms of $P_X(x)$. In practice it is usually easier to use LOTUS than direct definition when we need $E[g(X)]$.

Example 3.17

Let X be a discrete random variable with range $R_X = \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi\}$, such that $P_X(0) = P_X(\frac{\pi}{4}) = P_X(\frac{\pi}{2}) = P_X(\frac{3\pi}{4}) = P_X(\pi) = \frac{1}{5}$. Find $E[\sin(X)]$.

Solution

Using LOTUS, we have

$$\begin{aligned}E[g(X)] &= \sum_{x_k \in R_X} g(x_k)P_X(x_k) \\ &= \sin(0) \cdot \frac{1}{5} + \sin(\frac{\pi}{4}) \cdot \frac{1}{5} + \sin(\frac{\pi}{2}) \cdot \frac{1}{5} + \sin(\frac{3\pi}{4}) \cdot \frac{1}{5} + \sin(\pi) \cdot \frac{1}{5} \\ &= 0 \cdot \frac{1}{5} + \frac{\sqrt{2}}{2} \cdot \frac{1}{5} + 1 \cdot \frac{1}{5} + \frac{\sqrt{2}}{2} \cdot \frac{1}{5} + 0 \cdot \frac{1}{5} \\ &= \frac{\sqrt{2}+1}{5}.\end{aligned}$$

Example 3.18

Prove $E[aX+b] = aEX+b$ (linearity of expectation).

Solution

Here $g(X) = aX+b$, so using LOTUS we have

$$\begin{aligned}E[aX+b] &= \sum_{x_k \in R_X} (ax_k+b)P_X(x_k) \\ &= \sum_{x_k \in R_X} ax_kP_X(x_k) + \sum_{x_k \in R_X} bP_X(x_k) \\ &= a\sum_{x_k \in R_X} x_kP_X(x_k) + b\sum_{x_k \in R_X} P_X(x_k) \\ &= aEX+b.\end{aligned}$$
