- **Discrete Random Variables**

A *discrete random variable* is used to model a "random experiment" with a finite or countable number of possible outcomes. For example, the toss of a coin, the roll of a die, or the count of the number of telephone calls during a given time, etc.

The *sample space* $\mathcal{S}$, of the experiment is the set of all possible outcomes and contains a finite or countable number of elements. Let $S = \zeta_1, \zeta_2, \cdots$.

An *event* is a subset of S. Events consisting a single outcome are called *elementary events*.

- **Discrete Random Variables**

概率质量函数

Let X be a random variable with sample space $\mathcal{S}_X$. A *probability mass function (pmf)* for X is a mapping $p_X : \mathcal{S}_X \to [0, 1]$ from $\mathcal{S}_X$ to the closed unit interval [0, 1] satisfying

$$\sum_{x \in \mathcal{S}_X} p_X(x) = 1, \tag{4}$$

where the number $p_X(x)$ is the *probability* that the outcome of the given random experiment is x, i.e., $p_X(x) = \Pr[X = x]$.

Every event $A \in S$ has a probability $p(A) \in [0, 1]$ satisfying the following:

1. $p(A) \geq 0$
2. $p(\mathcal{S}) = 1$
3. for $A, B \in \mathcal{S}, p(A \cup B) = p(A) + p(B)$ if $A \cap B = \emptyset$

- **Discrete Random Variables**

Example: A fair coin is tossed N times, and A is the event that an even number of heads occurs. What is $\Pr[A]$?

$$\begin{aligned}\Pr[A] &= \sum_{k=0,\ k \text{ even}}^{N} \Pr[\text{exactly } k \text{ heads occur}] \\ &= \sum_{k=0,\ k \text{ even}}^{N} \binom{N}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{N-k} \\ &= \frac{1}{2^N} \sum_{k=0, k \text{ even}}^{N} \binom{N}{k} \\ &= \frac{1}{2}.\end{aligned}$$

- **Vector Random Variables**

If the elements of $\mathcal{S}_X$ are vectors of real numbers, then X is a *(real) vector random variable*.

Suppose Z is a vector random variable with a sample space in which each elements has two components (X, Y), i.e., $\mathcal{Z} = \{z_1, z_2, \cdots\} = \{(x_1, y_1), (x_2, y_2), \cdots\}$.

The *projection* of $\mathcal{S}_Z$ on its first coordinate is

$$\mathcal{S}_X = \{x : \text{for some} y, (x, y) \in \mathcal{S}_Z\}.$$

Example: If $Z = (X, Y)$ and $\mathcal{S}_Z = \{(0, 0), (1, 0), (1, 1)\}$, then $\mathcal{S}_X = \mathcal{S}_Y = \{0, 1\}$.

- **Vector Random Variables**

  The *pmf* of a vector random variable $Z = (X, Y)$ is also called the *joint pmf* of X and Y, and is denoted by

  $$p_Z(x, y) = p_{X,Y}(x, y) = \Pr(X = x, Y = y),$$

  where the comma in the last equation denotes a logical 'AND' operation.

  From $p_{X,Y}(x, y)$, we can find $p_X(x)$ as

  $$p_X(x) \equiv p(x) = \sum_{y \in \mathcal{S}_Y} p_{X,Y}(x, y);$$

  and similarly,

  $$p_Y(y) \equiv p(y) = \sum_{x \in \mathcal{S}_X} p_{X,Y}(x, y); \tag{5}$$

- Let A and B be events, with $\Pr[A] > 0$. The *conditional probability* of B given that A *occured* is

  $$\Pr[B|A] = \frac{\Pr[A \cap B]}{\Pr[A]}$$

  Thus, $\Pr[A|A] = 1$, and $\Pr[B|A] = 0$ if $A \cap B = \emptyset$.

  If $Z = (X, Y)$ and $p_X(x_k) > 0$, then

  $$\begin{aligned} p_{Y|X}(y_j|x_k) &= \Pr[Y = y_j | X = x_k] \\ &= \frac{\Pr[X = x_k, Y = y_j]}{\Pr[X = x_k]} \\ &= \frac{p_{X,Y}(x_k, y_j)}{p_X(x_k)}. \end{aligned}$$

- If $Z = (X, Y)$ and $p_X(x_k) > 0$, then

  $$p_{Y|X}(y_j|x_k) = \frac{p_{X,Y}(x_k, y_j)}{p_X(x_k)}.$$

  Then random variables X and Y are *independent* if

  $$\forall (x, y) \in \mathcal{S}_{X,Y} (p_{X,Y}(x, y) = p_X(x) p_Y(y)).$$

  If X and Y are *independent*, then

  $$p_{X|Y}(x|y) = \frac{p_{X,Y}(x, y)}{p_Y(y)} = \frac{p_X(x) p_Y(y)}{p_Y(y)} = p_X(x),$$

  and

  $$p_{Y|X}(y|x) = \frac{p_{X,Y}(x, y)}{p_X(x)} = \frac{p_X(x) p_Y(y)}{p_X(x)} = p_Y(y),$$

- If X is a random variable, the *expected value* (or mean) of X, denoted by $E[X]$, is

  $$E[X] = \sum_{x \in \mathcal{S}_X} x p_X(x).$$

  Then *expected value* of the random variable $f(X)$ is

  $$E[f(X)] = \sum_{x \in \mathcal{S}_X} f(x) p_X(x).$$

  In particular, $E[X^n]$ is the *n-th moment* of X. The *variance* of X is the second moment of $X - E[X]$, which can be computed as

  $$VAR[X] = E[X^2] - E[X]^2.$$