# Random Variables and Distributions

**Definition: Discrete Random Variable (RV)**

Let $(\Omega, \mathcal{F}, P)$ be a discrete probability space. A random variable $X$ is a function $X : \Omega \to \mathcal{X}$ where $\mathcal{X}$ is a set, and we may assume it to be discrete.

A *real* random variable is one whose image is contained in $\mathbb{R}$. A (The *image* and the *range* of a random variable $X$ are given by the image and the range of $X$ in the function-theoretic sense.) The image of a *binary* random variable is a set $x_0, x_1$ with only two elements.

**Definition: Probability distribution**

Let $X$ be a random variable. The probability distribution of $X$ is the function $P_X : \mathcal{X} \to [0, 1]$ defined as

$$P_X(x) := P[X = x],$$

where $X = x$ denotes the event $\{\omega \in \Omega \mid X(\omega) = x\}$.

Alternatively, one can write $P_X(x) = P[X^{-1}(x)]$ to express that the probability of $x$ is precisely the $P$-measure of the pre-image of $x$ under the random variable $X$.

We say that $P_X$ is a **uniform** distribution if the associated probability measure is uniform, i.e. $P_X(x) = \frac{1}{|\mathcal{X}|}$. The **support** of a random variable or a probability distribution is defined as $\mathrm{supp}(P_X) := \{x \in \mathcal{X} \mid P_X(x) > 0\}$, the points of the range which have strictly positive probability. We often slightly abuse notation and write $\mathrm{supp}(X)$ instead. When given two or more random variables defined on the same probability space, we can consider the probability that each of the variables take on a certain value:

**Definition: Joint probability distribution**

Let $X$ and $Y$ be two random variables defined on the same probability space, with respective ranges $\mathcal{X}$ and $\mathcal{Y}$. The pair $XY$ is a random variable with probability distribution $P_{XY} : \mathcal{X} \times \mathcal{Y} \to [0, 1]$ given by

$$P_{XY}(x, y) := P[X = x, Y = y].$$

This definition naturally extends to three and more random variables. Unless otherwise stated, a collection of random variables is assumed to be defined on the same (implicit) probability space, so that their joint distribution is always well-defined. If $P_{XY} = P_X \cdot P_Y$, in the sense that $P_{XY}(x, y) = P_X(x)P_Y(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, then the random variables $X$ and $Y$ are said to be **independent**. If a set of variables $X_1, \ldots, X_n$ are all mutually independent and all have the same distribution (i.e., $P_{X_i} = P_{X_j}$ for all $i, j$), then they are **independent and identically distributed**, or **i.i.d.** From a joint distribution, we can always find out the "original'' (or **marginal**) distribution of one of the random variables (for example, $X$) by **marginalizing** out the variable that we want to discard (for example, $Y$):

$$P_X(x) = \sum_{y \in \mathcal{Y}} P_{XY}(x, y).$$

This marginalization process also works with more than two random variables. Like events, probability distributions can also be conditioned on probabilistic events:

**Definition: Conditional probability distribution**

If $\mathcal{A}$ is an event with $P[\mathcal{A}] > 0$, then the conditional probability distribution of $X$ given $\mathcal{A}$ is given by

$$P_{X|\mathcal{A}}(x) = \frac{P[X = x, \mathcal{A}]}{P[\mathcal{A}]}.$$

If $Y$ is another random variable and $P_Y(y) > 0$, then we write

$$P_{X|Y}(x|y) := P_{X|Y=y}(x) = \frac{P_{XY}(x, y)}{P_Y(y)}$$

for the conditional distribution of $X$, given $Y = y$.

Note that again, both $(\mathcal{X}, P_{X|\mathcal{A}})$ and $(\mathcal{X}, P_{X|Y=y})$ themselves form probability spaces. Note also that if $X$ and $Y$ are independent, then

$$P_{X|Y}(x|y) = \frac{P_{XY}(x, y)}{P_Y(y)} = \frac{P_X(x) \cdot P_Y(y)}{P_Y(y)} = P_X(x),$$

which aligns well with our intuition of independent variables: the distribution of $X$ remains unchanged when $Y$ is fixed to a specific value.

**Example: Fair die (continued)**

Consider again the throw of a six-sided fair die. Let the random variable $X$ describe the number of (distinct) integer divisors for the outcome, that is

$$X(1) = 1 \quad X(2) = 2 \quad X(3) = 2 \quad X(4) = 3 \quad X(5) = 2 \quad X(6) = 4$$

$X$ is a real random variable, with range $\mathcal{X} = 1, 2, 3, 4$. The associated probability distribution is

$$P_X(1) = P[1] = \frac{1}{6}, \quad P_X(2) = P[2, 3, 5] = \frac{1}{2}, \quad P_X(3) = P[4] = \frac{1}{6}, \quad P_X(4) = P[6] = \frac{1}{6}.$$

If we now condition on the event $\mathcal{A} = 2, 4, 6$ (the outcome of the die being even), we get that

$$P_{X|\mathcal{A}}(1) = 0, \quad P_{X|\mathcal{A}}(2) = \frac{1}{3}, \quad P_{X|\mathcal{A}}(3) = \frac{1}{3}, \quad P_{X|\mathcal{A}}(4) = \frac{1}{3}$$

If $X$ is a random variable and $f : \mathcal{X} \to \mathcal{Y}$ is a surjective function, then $f(X)$ is a random variable, defined by composing the map $f$ with the map $X$. Its image is $\mathcal{Y}$. Clearly,

$$P_{f(X)}(y) = \sum_{x \in \mathcal{X}: f(x) = y} P_X(x).$$

For example, $1/P_X(X)$ denotes the real random variable obtained from another random variable $X$ by composing with the map $1/P_X$ that assigns $1/P_X(x) \in \mathbb{R}$ to $x \in \mathcal{X}$.