Probability I - Random Variables

Arkaraj Mukherjee

August 29, 2025

For a sample space Ω, event space $\mathcal{F}$ and probability P, a *random variable* is a function $X : \Omega \to \mathbb{R}$. the set of values the random variable can take, i.e. it's range is called the *support* of the RV (short form of Random Variable) often denoted by

$$\operatorname{supp} X : = \{X(\omega) | \omega \in \Omega\}$$

RVs X for which $\operatorname{supp} X$ is atmost countable i.e. countably infinite or finite are called *discrete RVs.* The probabilities of the RV taking on some values in its support are governed by the *probability mass function* (*pmf.* in short) often denoted by p, i.e.

$$\begin{aligned} p : \operatorname{supp} X &\longrightarrow [0, 1] \\ x &\mapsto P(X = x) := P(X^{-1}\{x\}) \end{aligned}$$

In practice we extend this definition of p on $\operatorname{supp} X$ all the way to $\mathbb{R}$ by fixing $p(x) = 0$ for all $x \in \mathbb{R} \setminus \operatorname{supp} X$. For discrete RVs, if we ennumerate $\operatorname{supp} X = \{x_i | i = 1, 2, \ldots\}$ then from the axioms of probability we get that,

$$\sum_{i=1}^{\infty} p(x_i) = \underbrace{\sum_{i=1}^{\infty} P(X^{-1}(\{x_i\})) = P\left(\bigcup_{i=1}^{\infty} X^{-1}(\{x_i\})\right) = P(\Omega) = 1}_{\text{Why are these three equalities true?}}$$

The *cumulative distribution function* (cdf. for short) of a random variable is a function on the real line defined as,

$$\begin{aligned} F : \mathbb{R} &\to [0, 1] \\ x &\mapsto P(X \leq x) := P(X^{-1}((-\infty, x])) \end{aligned}$$

This is a non-decreasing step-function with jump discontinuitites and from the definition we can see that (try proving it)

$$p(x) = F(x) - \lim_{\substack{h \to x \\ h < x}} F(h)$$

The limit on the right is sometimes known as the *left limit of F at x*. From the axioms of probability again we can see that

$$F(x) = P(X^{-1}((-\infty, x])) = P\left(\bigcup_{x_i < x} X^{-1}(\{x_i\})\right) = \sum_{x_i < x} P(X^{-1}(\{x_i\})) = \sum_{x_i < x} p(x_i)$$

You should notice that we can also deduce the equation with the left limit from this as well. The *expectation* of a RV is a weighted average of the values it takes, formally this is written as

$$\mathbb{E}[X] := \sum_{i=1}^{\infty} x_i p(x_i)$$

It may not always exist(i.e. converge). Suppose we have a RV X, then for some function f on $\operatorname{supp} X$, $Y = f(X)$ is also a RV i.e. a function $\Omega \to \mathbb{R}$ and $\operatorname{supp} Y = f(\operatorname{supp} X)$. Also,

$$\mathbb{E}\left[Y\right] = \sum_{y \in \operatorname{supp} Y} yP(f(X) = y) = \sum_{y \in \operatorname{supp} Y} \sum_{x \in f^{-1}(y)} yP(X = x)$$

$$= \sum_{y \in \operatorname{supp} Y} \sum_{x \in f^{-1}(y)} f(x)P(X = x) = \sum_{x \in \bigcup_{y \in f(\operatorname{supp} X)} f^{-1}(y)} f(x)P(X = x)$$

$$= \sum_{x \in \operatorname{supp} X} f(x)P(X = x)$$

This is a major result. The *k-th moment* of a RV is defined as $\mathbb{E}\left[X^k\right] := \mathbb{E}\left[g(X)\right]$ where $g : x \mapsto x^k$.