Random variables

X is a random variable if, informally, it is a function $X : S \to \mathbb{R}$ and we can quantify the "likelihood" of its values.

Discrete random variables

A discrete random variable is a kind of random variable. The first property is that instead of mapping to $\mathbb{R}$ it must map to a **countable** set, such as $\mathbb{Z}^+$. For example, let S be the sample space associated with flipping two coins, $S = \{(H,H),(H,T),(T,H),(T,T)\}$, and let X be the number of heads. In this case, $X : S \to \{0,1,2\}$, and we have $\mathbb{P}(X = 0) = \frac{1}{4}$, $\mathbb{P}(X = 1) = \frac{1}{2}$ and $\mathbb{P}(X = 2) = \frac{1}{4}$.

In general, for a discrete variable X with possible values $\{x_1, x_2 \cdots x_n\}$ we can calculate the values $\mathbb{P}(X = x_n) = p_n = f(x_n)$, where p_i is a probability, that is, $\Sigma_{i=1}^{n} p_i = 1$; note that n does not have to be finite. f is called the **probability mass function**, and is in this case the **discrete distribution function** of X.

Uniform variable on integers

A uniform variable on integers is any discrete random variable X such that $\mathbb{P}(X = x_1) = \mathbb{P}(X = x_2) = \cdots = \mathbb{P}(X = x_n) = \frac{1}{n}$. An example is the **Bernoulli distribution**, which is defined $X = \mathrm{Bn}(p), \mathbb{P}(X = x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \end{cases}$.

Binomial random variable

The context for this kind of discrete random variable is an experiment that consists of n independent **Bernoulli trials**. A Bernoulli trial is a trial that can be either a success, with probability p, or a failure, with probability $1-p$. That is, each trial follows the Bernoulli distribution.

Let X be the random variable that counts the number of successes in such an experiment, out of n. then, X is called a **Binomial random variable**, with parameters n and p. We write $X \sim \mathrm{B}(n,p)$, such that $\mathbb{P}(X = k) = p_k = \binom{n}{k} p^k (1-p)^{m-k}$.

As an example, suppose that it is known that the memory sticks produced by some company will be defective with probability 0.01, independently of each other. The company sells the sticks in packages of ten. It offers a money back guarantee that at most 1 in 10 is defective. You buy three packages; that is the probability that you return exactly one of them?

Answer: Let Y be the number of returned packages. Then, $Y \sim \mathrm{B}(3, p)$, and we need to find $\mathbb{P}(Y = 1)$. p is the probability of a single package being returned. Let X be the number of defective memory sticks in a package: $X \sim \mathrm{B}(10, 0.01)$. $\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X = 0, 1) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1)) = 1 - (\binom{10}{0} \cdot 0.01^0 \cdot 0.99^10 + \binom{10}{1} \cdot 0.01^1 \cdot 0.99^9) = 1 - 0.9957 = 0.004266$. Therefore, $\mathbb{P}(Y = 1) = \binom{3}{1} 0.004266^1 0.9957^2 = 0.0127$.

Continuous random variables

A continuous random variable is another kind of random variable. Let $X : S \to \mathbb{R}$, and define $\mathbb{P}(X \in B) = \mathbb{P}(s \in S | X(s) \in B)$. Then, X is a continuous random variable if there exists a function $f(x) \geq 0 : \mathbb{R} \to [0, \infty)$ such that $\mathbb{P}(X \in B) = \int_B f(x)dx$, for any $B \subseteq \mathbb{R}$. In this case, f is called the **probability density function** of X.

As a consequence of this, we have $\int_{-\infty}^{\infty} f(x)dx = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(S) = 1$. The rule for computing probabilities of events that can be described via a continuous random variable is as follows: the probability $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$. Note that the result is equivalent regardless of whether the inequalities are strict or not.

Cumulative distribution function

Let X be a random variable. Then $F : \mathbb{R} \to [0, 1]$ defined by $F(x) \equiv \mathbb{P}(X \leq x)$ is called the **cumulative distribution function** of X.