Probability

When we are looking into the future, an expected event will happen with a certain probability. This probability ranges from 0 (“will not happen” as in *Two suns will simultaneously rise tomorrow*) to 1 (“is guaranteed to happen” as in *The sun will set tonight*.)

Note that **probability** is firmly associated with **future**. The probability that something happened yesterday is always 0 or 1.

Random variable

A random variable is a set of homogeneous objects (numbers, colours, letters) with a probability associated with each object. Considering that *probability* is part of the future, so is a *random variable*. The set can be infinite, etc.

Note that the sum of the probabilities of all the objects in the set must add to 1.

A random variable has no specific value; the whole set represents it.

Standard notation calls for naming random variables $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$.

If the random variable has a countable ($\subseteq \aleph_0$) set of values, the variable is called **discrete**; otherwise, it is called **continuous**.

The definitions vary somewhat for the two types, the former using summation and the latter integral calculus.

Probability distribution–discrete

If the set of values of a random variable $\mathcal{X}$ is finite, a probability is associated with each member of the set. **Probability distribution** is the fancy term for the collection of probabilities.

Quite formally, the **probability distribution function** $f : \mathcal{X} \rightarrow < 0, 1 >$ or $f(\mathcal{X})$ of a discrete random variable $\mathcal{X}$ is

$$f(x) = \begin{cases} p(x) & x \in \mathcal{X} \\ 0 & x \notin \mathcal{X} \end{cases}$$

For a continuous distribution $\mathcal{X}$, the probability $P(\mathcal{X} = x) = 0$ for all values of x (artificial exceptions being ignored), so an alternate definition based on the cumulative distribution must be used. We call it the **probability density function** $f(\mathcal{X})$ (name conveniently chosen to abbreviate **pdf** as well).

For practical purposes, a **pdf** is equivalent to the random variable that it represents.

Cumulative distribution

Cumulative distribution function:

$$F(x) = P(\mathcal{X} \leq x)$$

If we know the $f(\mathcal{X})$ of a variable $\mathcal{X}$,

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

(the integral degenerates to a summation for discrete variables).

The cumulative distribution function gives the actual distribution of probabilities by interval: the probability of a random variable taking a value from interval (a, b) equals $F(b) - F(a)$.

For continuous distributions, the probability density function $f(x)$ equals:

$$f(x) = \frac{d}{dx}F(x)$$

This pair of definitions is circular.

Variate aka Random Number

When one picks a specific object from the set forming a random variable, that object becomes a **variate**, more commonly called a random number.

A variate is a constant and there is nothing *random* about it.

Intuitively, a variate is part of the past in the same way as a probability is part of the future.

Variate vs. Random Variable

Consider a deck of 52 cards spread on a table face down.

This is a random variable made of a set of 52 objects; a probability of $\frac{1}{52}$ is associated with each object.

We touch one card; this card can be any of the 52, hence it represents the whole variable. The probability that this card is a $2\clubsuit$ is $\frac{1}{52}$.

Now we turn the card face up and discover that it is a $K\spadesuit$. This $K\spadesuit$ is now **a fact** (with probability 1) and there is nothing random about it anymore. We call it a **variate**. Note that it became a variate at the moment when it was turned face up, i.e. when the card lost its random nature.

Example

1. We go to the racetrack. In the next race, 6 horses will be running. Looking at them expertly, we determine that the probabilities of victory are:

Horse	1	2	3	4	5	6
Probability	0.2	0.4	0.1	0.1	0.15	0.05

   These 6 horses form a random variable with a **pdf** given by the second row of the table.

2. The race begins. What happens during the race is inconsequential.

3. The race ends. Horse 5 won. This “5” is a variate; it does not make much sense to talk about probabilities now. If we must, we could create a new random variable which would have $p(5) = 1, p(1, 2, 3, 4, 6) = 0$.

Properties of a random variables

Every random variable $\mathcal{X}$ has two properties: a **mean** and a **variance**.

- **Mean** (or expectation) μ or E()

  The general formula for discrete distributions is:

  $$\mu = E(\mathcal{X}) = \sum_{i=1}^{n} p(x_i)x_i$$

  (The assumption is that $\mathcal{X}$ has n members.)

  For continuous variables:

  $$\mu = \int_{-\infty}^{\infty} xf(x)dx$$

It is a mistake to assume that $P(x \leq \mu) = 1/2$

The mean, used in isolation, tells us very little about the random variable. When combined with the variance, it becomes very useful.

- **Variance** $V()$ or σ^2

  The general formula is:

$$\sigma^2 = E(\mathcal{X}^2) - \mu^2$$

  For discrete distributions:

$$\sigma^2 = V(X) = \sum_{i=1}^{n} x_i^2 p(x_i) - \mu^2$$

The **standard deviation** $\sigma = \sqrt{V(X)}$ has the same unit as the mean.

The variance (and standard deviation) is independent of the mean:

$$V(\mathcal{X} + c) = V(\mathcal{X})$$

Uniform distribution

A random variable $\mathcal{X}$ is **uniformly distributed** on the interval (a, b) if its **pdf** is given by:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \geq b \\ 0 & \text{otherwise} \end{cases}$$

The expert will quickly calculate

$$\mu = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

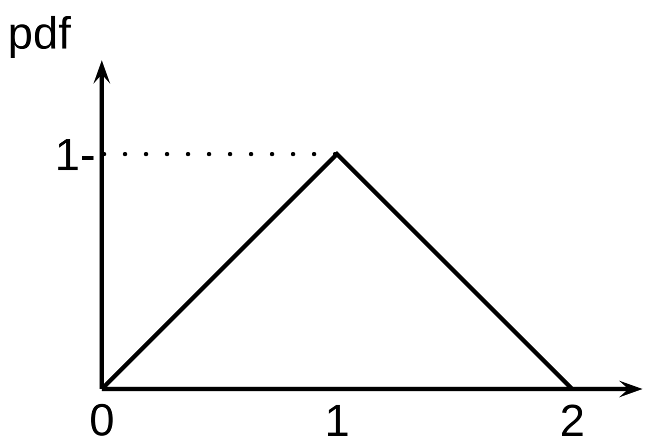
$$\sigma^2 = \frac{(b-a)^2}{12}$$

Uniform variates can be generated by calling a standard library function which returns a variate uniformly distributed on the interval (0,1). We denote a uniform variate as $U(0, 1)$.

Typical library function name: **drand48()**. when not available, the expression **(double)random() / MAXINT** will do.

Triangular distribution

Adding two uniformly distributed random variables yields a random variable that has a **triangular distribution**. Its **pdf** looks like this:



To generate a triangular variate T, one uses the formula:

$$T = U_1(0,1) + U_2(0,1)$$

where U_1 and U_2 are two independent uniform variates.