1 Basic definitions

Definition 1.1. Support The set of values a random variable can take (set of all the numerical realizations of outcomes) e.g. for a binary random variable X, it's support is $R_X = \{0, 1\}$

Definition 1.2. Sample Point In an experiment, a sample point is one of the possible outcomes of experiment denoted by ω

Definition 1.3. Independent Events Two events A and B are independent if and only if

$$P(A \cap B) = P(A) * P(B) \tag{1}$$

Definition 1.4. Jointly/Mutually Independent Events Let $E_1, ..., E_n$ be n events.
$E_1, ..., E_n$ are jointly independent(or mutually independent) if and only if for any sub-collection of k events $(k \leq n) E_{i1}, ..., E_{ik}$:

$$P(\bigcap_{j=1}^{k} E_{ij}) = \prod_{j=1}^{k} P(E_{ij}) \tag{2}$$

Definition 1.5. Random Variable A random Variable X is a function from sample space Ω to set of real numbers $\mathbb{R}$, i.e $X : \Omega -> \mathbb{R}$.

The real number $X(\omega)$ associated with sample point $\omega \in \Omega$ is called a realization of the random variable. The set of all possible realizations is called the support and denoted by R_X

Definition 1.6. Probability Mass Function The PMF of a discrete random variable X is a function $p_X : \mathbb{R} -> [0, 1]$ such that

$$p_X = P(X = x) \quad \forall x \in \mathbb{R} \tag{3}$$

where $P(X = x)$ is the probability of realization of random variable X will be equal to x. Basically PMF is numerical realizations $->$ respective Probabilities.

Definition 1.7. Distribution Function/Cumulative Distribution Function If X is a random variable, its distribution/cdf is a function $F_X : \mathbb{R} -> [0, 1]$ such that

$$F_X(x) = P(X \leq x) \quad \forall x \in \mathbb{R} \tag{4}$$

Definition 1.8. Expected Value of a Random Variable The expected value of random variable X is the weighted average of the values that X can take on where each possible realization value is weighted by its respective probability i.e.

$$\mathbf{E}[X] = \sum_{x \in R_X} x p_X(x) \tag{5}$$

Definition 1.9. Deviation of a Random Variable The Deviation of a Random Variable is its difference from its mean value/Expected value. It is denoted by

$$\overline{X} = X - \mathbf{E}[X] \tag{6}$$

Deviation of Random Variable is also a random variable in it's own respect.

Definition 1.10. Variance It is a measure of dispersion of a random variable. Let X be a random variable. The variance of X, denoted by $Var[X]$ is defined as follows:

$$Var[X] = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[\overline{X}^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2 \tag{7}$$

You can also think of it as mean of the square of the deviations of original random variable.

Definition 1.11. Standard Deviation It is a measure of dispersion of a random variable. Let X be a random variable. The standard deviation of X, denoted by $stdev[X]$ or $std[X]$ is defined as follows:

$$stdev[X] = \sqrt{Var[X]} \tag{8}$$

You can also think of it as rms deviation.

Definition 1.12. Covariance It is a measure of association between two random variables. Covariance of two random variables X and Y is , provided Expected values are well defined,

$$Cov[X,Y] = \mathbf{E}[(X - \mathbf{E}[X]) * (Y - \mathbf{E}[Y])] = \mathbf{E}[\overline{X} * \overline{Y}] \tag{9}$$

where we say deviation of X is $\overline{X} = X - \mathbf{E}[X]$ and
deviation of Y is $\overline{Y} = Y - \mathbf{E}[Y]$
or in other words, covariance is expectation of product of deviations.

Definition 1.13. Random Vector It is a multidimensional generalisation of the concept of Random Variable. Associated probability functions have the word "joint" in front of them e.g. Joint PMF, Joint cdf etc correspond to a Random Vector.

2 Logistic

A logistic function or logistic curve is a common 'S' shaped curve with equation:

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \tag{10}$$

where

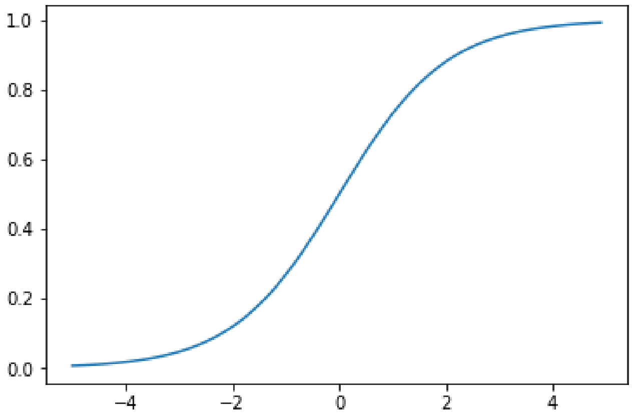
- e = natural log
- x_0 = x-value of sigmoid's midpoint
- L = the curve's maximum value
- k = the steepness of the curve

2.1 Standard Logistic

The standard logistic function is the logistic function with parameters given $(k = 1, x_0 = 0, L = 1)$ i.e.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{11}$$

which when plotted looks like



Why is logistic function so important? Because it can take any real input $x, (x \in R)$, whereas the output always takes values between 0 and 1, and hence is interpretable as probability.