# Random Variables

A ***random variable***, usually written *X*, is a variable whose possible values are numerical outcomes of a random phenomenon. There are two types of random variables, ***discrete*** and ***continuous***.

**Discrete Random Variables**

A ***discrete random variable*** is one which may take on only a countable number of distinct values such as 0,1,2,3,4,........ Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

The ***probability distribution*** of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the **probability function or the probability mass function.**

Suppose a random variable *X* may take *k* different values, with the probability that *X = xi* defined to be *P(X = xi) = pi*. The probabilities *pi* must satisfy the following:

***1:*** *0 < pi < 1 for each i*

***2:*** *p1 + p2 + ... + pk = 1.*

#### Example

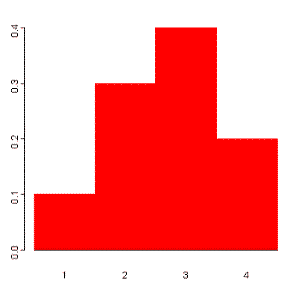
Suppose a variable X can take the values 1, 2, 3, or 4.   
The probabilities associated with each outcome are described by the following table:

Outcome 1 2 3 4

Probability 0.1 0.3 0.4 0.2

The probability that *X* is equal to 2 or 3 is the sum of the two probabilities: *P(X = 2 or X = 3) = P(X = 2) + P(X = 3)* = 0.3 + 0.4 = 0.7.

Similarly, the probability that *X* is greater than 1 is equal to 1 - *P(X = 1)* = 1 - 0.1 = 0.9, by the [complement rule](http://www.stat.yale.edu/Courses/1997-98/101/probint.htm#rule4).

This distribution may also be described by the ***probability histogram***.

**Continuous Random Variables**

A ***continuous random variable*** is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

A continuous random variable is not defined at specific values. Instead, it is defined over an *interval* of values, and is represented by the ***area under a curve*** (in advanced mathematics, this is known as an *integral*). The probability of observing any single value is equal to 0, since the number of values which may be assumed by the random variable is infinite.

Suppose a random variable *X* may take all values over an interval of real numbers. Then the probability that *X* is in the set of outcomes *A, P(A)*, is defined to be the area above *A* and under a curve. The curve, which represents a function *p(x)*, must satisfy the following:

***1:*** *The curve has no negative values (p(x) > 0 for all x)*

***2:*** *The total area under the curve is equal to 1.*

A curve meeting these requirements is known as a ***density curve***.

