DISCRETE AND CONTINUOUS PROBABILITY DISTRIBUTIONS

Probability mass functions

If $x \in \{x_1, x_2, x_3, \ldots\}$ is discrete, then a function $f(x_i)$ giving the probability that $x = x_i$ is called a *probability mass function.* Such a function must have the properties that

$$f(x_i) \geq 0, \quad \text{for all} \quad i, \qquad \text{and} \quad \sum_i f(x_i) = 1.$$

Example. Consider $x \in \{0, 1, 2, 3, \ldots\}$ with

$$f(x) = (1/2)^{x+1}.$$

It is certainly true that $f(x_i) \geq 0$ for all i. Also,

$$\sum_i f(x_i) = \left\{ \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots \right\} = 1.$$

To see this, we may recall that

$$\frac{1}{1-\theta} = \{1 + \theta + \theta^2 + \theta^3 + \cdots\},$$

whence

$$\frac{\theta}{1-\theta} = \{\theta + \theta^2 + \theta^3 + +\theta^4 + \cdots\}.$$

Setting $\theta = 1/2$ in the expression above gives $\theta/(1-\theta) = \frac{1}{2}/(1-\frac{1}{2}) = 1$, which is the result that we are seeking.

Probability density functions

If x is continuous, then a *probability density function* (p.d.f.) $f(x)$ may be defined such that the probability of the event $a < x \leq b$ is given by

$$P(a < x \leq b) = \int_a^b f(x)dx.$$

Notice that, when $b = a$, there is

$$P(x = a) = \int_a^a f(x)dx = 0.$$

That is to say, the integral of the continuous function $f(x)$ at a point is zero. The value of $f(x)$ at a point is described as a probability measure as opposed to a probability.

Example. Consider the exponential function

$$f(x) = \frac{1}{a}e^{-x/a}$$

defined over the interval $[0, \infty)$ and with $\alpha > 0$. There is $f(x) > 0$ and

$$\int_0^\infty \frac{1}{a}e^{-x/a}dx = \left[-e^{-x/a}\right]_0^\infty = 1.$$

Therefore, this function constitutes a valid p.d.f.

The exponential distribution provides a model for the lifespan of an electronic component, such as fuse, for which the probability of failing is liable to be independent of how long it has already survived.

Cumulative distribution functions

Corresponding to any p.d.f $f(x)$, there is a cumulative distribution function, denoted by $F(x)$, which, for any value x^*, gives the probability of the event $x \leq x^*$.

Thus , if $f(x)$ is the p.d.f. of x, which we denote by writing $x \sim f(x)$, then

$$F(x^*) = \int_{-\infty}^{x^*} f(x)dx = P(-\infty < x \leq x^*).$$

In the case where x is a discrete random variable with a probability mass function $f(x)$, also denoted by $x \sim f(x)$, there is

$$F(x^*) = \sum_{x \leq x^*} f(x) = P(-\infty < x \leq x^*).$$

Pascal's Triangle and the Binomial Expansion.

Consider the following binomial expansions:

$$\begin{aligned}(p+q)^0 &= 1,\\ (p+q)^1 &= p+q,\\ (p+q)^2 &= p^2+2pq+q^2,\\ (p+q)^3 &= p^3+3p^2q+3pq^2+q^3,\\ (p+q)^4 &= p^4+4p^3q+6p^2q^2+4pq^3+q^4,\\ (p+q)^5 &= p^5+5p^4q+10p^3q^2+10p^2q^3+5pq^4+q^5.\end{aligned}$$

The generic expansion is in the form of

$$(p+q)^n = \sum_{x=0}^{n} \frac{n!}{(n-x)!x!} p^x q^{n-x}.$$

We can find the coefficient of the binomial expansions of successive degrees via Pascal's triangle, where the numbers in each row but the first are obtained by adding two adjacent numbers in the row above:

$$\begin{array}{ccccccccccc} &&&&& 1 &&&&& \\ &&&& 1 && 1 &&&& \\ &&& 1 && 2 && 1 &&& \\ && 1 && 3 && 3 && 1 && \\ & 1 && 4 && 6 && 4 && 1 & \\ 1 && 5 && 10 && 10 && 5 && 1 \end{array}$$

Permutations. Consider the various ways of ordering the letters $\{a, b, c\}$. The first place can be filled in three ways, the second place in two ways and third place must be filled with the remaining letter. Thus, there there are altogether $3 \times 2 \times 1 = 6$ different arrangements or permutations:

$$abc, \quad cab, \quad bca, \quad cba, \quad bac \quad acb.$$

Now consider an ordering a set of n objects

$$\{x_i; i = 1, \ldots, n\},$$

Q_1: How may ways to fill the 1st place? A_1: n
Q_2: How may ways to fill the 2nd place? A_2: $n - 1$ ways,
Q_3: How may ways to fill the 3rd place? A_3: $n - 2$ ways,
$\vdots$
Q_r: How may ways to fill the rth place? A_r: $n - r + 1$ ways,
$\vdots$
Q_n: How may ways to fill the nth place? A_n: 1 way.

We can recognise that there are

$$n(n-1)(n-2)\cdots 3.2.1 = n!$$

different permutations of the objects. This is n-factorial, written as $n!$. It is the number of permutations of n objects taken n at a time. We also denote it by

$$^nP_n = n!.$$

Permutations, continued.

Q1: How many ordered sets can we recognise if r of the n objects are so alike as to be indistinguishable?

A1: Within any permutation there are r objects that are indistinguishable. The r objects can be permuted in $r!$ ways. Therefore, there are only

$$\frac{n!}{r!} \quad \text{recognizably distinct permutations.}$$

Q2: How many distinct permutations can we recognise if the n objects are divided into two sets of r and $n - r$ objects, where two objects in the same set are indistinguishable?

A2: By extending the previous argument, we should find that the answer is

$$\frac{n!}{(n-r)!r!}.$$

Q3: How many ways can we construct a permutation of r objects selected from a set of n objects?

A3: The number of ways of filling the r places is

$$\begin{aligned} {}^nP_r &= n(n-1)(n-2)\cdots(n-r+1) \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

This is also the number of distinct permutation of n objects when $n-r$ of them are indistinguishable. We cannot distinguish amongst the objects that have been omitted from the selection.

Combinations. Combinations are selections of objects in which no attention is payed to the ordering.

Q: How many ways can we construct an unordered set of r objects selected from amongst n objects?

A: The total number of permutations of r objects selected from amongst n is ${}^nP_r$.

There are $r!$ different orderings or permutations of the r objects which are to be ignored. Therefore, we divide ${}^nP_r$ by $r!$ to give the total number of combinations:

$$\begin{aligned} {}^nC_r &= \frac{1}{r!}\,{}^nP_r = \frac{n!}{(n-r)!r!} \\ &= {}^nC_{n-r}. \end{aligned}$$

This is also the number of distinct permutations of a set of n objects of which r are in one category and $n-r$ in another.

Example: the power set. The power set is the set of all sets cotaining selections from n objects. There are exactly 2^n objects in the power set.

This is demonstrated by setting $p = q = 1$ in the binomial expansion to give

$$2^n = \sum_x \binom{n}{x} = \sum_x \frac{n!}{x!(n-x)!}.$$

Each element of this sum is the number of ways of selecting x objects from amongst n, and the sum is for all values of $x = 0, 1, \ldots, n$.

The Binomial Theorem. The object is to determine the coefficient associated with the generic term $p^x q^{n-x}$ in the expansion of

$$(p+q)^n = (p+q)(p+q)\cdots(p+q),$$

where the RHS displays the n factors multiplied together. The elements of the expansion have the following coefficients:

- p^n The coefficient is unity, since there is only one way of choosing n of the p's from amongst the n factors.
- $p^{n-1}q$ This term is the product of q selected from one of the factors and $n-1$ p's provided by the remaining factors. There are $n = {}^nC_1$ ways of selecting the single q.
- $p^{n-2}q^2$ The coefficent associated with this term is the number of ways of selecting two q's from n factors which is ${}^nC_2 = n(n-1)/2$.

$\vdots$

- $p^{n-r}q^r$ The coefficent associated with this terms is the number of ways of selecting r q's from n factors which is ${}^nC_r = n!/\{(n-r)!r!\}$.

From such reasoning, it follows that

$$\begin{aligned}(p+q)^n =& p^n + {}^nC_1 p^{n-1}q + {}^nC_2 p^{n-2}q^2 + \\ & \cdots + {}^nC_r p^{n-r}q^r + \cdots \\ & + {}^nC_{n-2}p^2q^{n-2} + {}^nC_{n-1}pq^{n-1} + q^n \\ =& \sum_{x=0}^{n} \frac{n!}{(n-x)!x!} p^x q^{n-x}.\end{aligned}$$

The Binomial Probability Distribution. We wish to find, for example, the number of ways of getting a total of x heads in n tosses of a coin.

Consider the ith toss, and let $x_i = 1$ denote heads and $x_i = 0$ denote tails. Let $P(x_i = 1) = p$ and $P(x_i = 0) = 1 - p$.

The probabilty function for the outcome of the ith trial is

$$f(x_i) = p^{x_i}(1-p)^{1-x_i} \quad \text{with} \quad x_i \in \{0, 1\}.$$

If the coin is tossed n times, then the probability of any particular sequence $(x_1, x_2, \ldots, x_n)$ of heads and tails is given by

$$\begin{aligned} P(x_1, x_2, \ldots, x_n) &= f(x_1)f(x_2)\cdots f(x_n) \\ &= p^{\sum x_i}(1-p)^{n-\sum x_i} \\ &= p^x(1-p)^{n-x}, \end{aligned}$$

where we have defined $x = \sum x_i$. This follows from the independence of the trials whereby $P(x_i, x_j) = P(x_i)P(x_j)$.

Altogether there are

$$\binom{n}{x} = \frac{n!}{(n-x)!x!} = nC_x$$

different sequences $(x_1, x_2, \ldots, x_n)$ containing x heads; so the probability of getting x heads in n tosses is given by

$$b(x; n, p) = \frac{n!}{(n-x)!x!} p^x(1-p)^{n-x}.$$

This is the binomial probability mass function.