"Theoretical" Reference Distributions

There are a number of theoretical (i.e. based on some kind of theory or understanding), as opposed to empirical (i.e. based on direct experience) distributions (or descriptions of the frequency at which different values of a variable are likely to occur) that can be used as reference distributions for specific statistics obtained from particular data sets. These distributions can be described using their *probability density functions* (or *pdfs*), which are analogous to the histogram, and *cumulative density functions* (or *cdfs*), which are analogous to the cumulative frequency curve.

Uniform distribution

The uniform distribution is defined for a particular interval on the number line, say $[a,b]$ (read as "the interval from a to b"), such that the probability of observing and value within this interval is p, while the probability of observing outside of this interval is zero. Let X represent a random variable (and let x_i be the *i*-th observation of athe random variable X, $(i = 1, \ldots n)$, or a variable whose observed values are determined by chance. If X is continuous (i.e., it can take on any value, not just integer values), then the pdf of the uniform distribution is

$$pdf_{uniform} = P(X)$$
$$= 1/(b-a), \text{ if } a \leq x \leq b,$$
$$= 0, \text{ otherwise.}$$

Binomial distribution

The binomial distribution is a discrete distribution that applies to situations where only two outcomes are possible (e.g. rain, no rain). The pdf of the binomial distribution is

$$pdf_{binomial} = P(X) = \frac{n!(p^X)(q^{n-X})}{X!(n-X)!}$$

where:

$n =$ number of events or "trials,"

$p =$ probability of the "given" outcome on a single trial,

$q = (1-p)$, or the probability of the "other" outcome in a single trial,

$X =$ the number of times the "given" outcome occurs within the n trials, and

$n! = n$ factorial, if $n > 0$, $n! = [n(n-1)(n-2), \ldots, (2)(1)]$, or if $n = 0$, $n! = 1$.

Poisson distribution

The Poisson distribution is a discrete distribution that illustrates the probability of observing a particular number of events in an area or time interval, when the mean number of events per area or time are known. The pdf of the Poisson distribution is

$$pdf_{Poisson} = P(X) = \frac{(e^{-z})(z^X)}{X!} = \frac{z^X}{(e^z)(X!)}, \quad (X = 0,1,2,\ldots)$$

where:

$X =$ frequency of occurrence of the event

$z =$ mean frequency of occurrence

$e = 2.71823\ldots,$ and

$X! = X$ factorial.

Normal distribution

The normal distribution arises frequency in practice as a consequence of the Central Limit Theorem, and the fact that many phenomena that are observed in practice represent integration of processes over time or space. The normal distribution is a continuous distribution, and its pdf is given by

$$pdf_{normal} = P(X) = \frac{1}{(2\pi\sigma^2)^{0.5}} e^{-(X-\mu)^2/2\sigma^2}$$

where

$X =$ the value of a statistic that is known to be normally distributed,

$\mu =$ the mean (or location) of the distribution,

$\sigma =$ the standard deviation (or the scale or dispersion) of the distribution

μ and σ are referred to as the parameters of the normal distribution.

The standard normal distribution applies to the simple case when $\mu = 0,$ and $\sigma = 1.$ The pdf of the standard normal distribution is

$$pdf_{std.\,normal} = P(Z) = \frac{1}{(2\pi)^{0.5}} e^{-Z^2/2}$$

where:

$Z =$ the value of a statistic that is known to a standard normal distribution.