Probability Generating Function (pgf)

The probability generating function is an example of a generating function of a sequence.[1]

In mathematics, a **generating function** is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers a_n that is indexed by the natural numbers.[2]

The *ordinary generating function* of a sequence a_n is

$$G(a_n;t) = \sum_{n=0}^{\infty} a_n t^n$$

When the term *generating function* is used without qualification, it is usually taken to mean an ordinary generating function.[2]

Examples of Generating Functions

A key generating function is the constant sequence 1, 1, 1, 1, 1, 1, 1, 1, 1, ..., whose ordinary generating function is

$$\sum_{n=0}^{\infty} t^n = \frac{1}{1-t}$$

The left-hand side is the Maclaurin series expansion of the right-hand side. Alternatively, the right-hand side expression can be justified by multiplying the power series on the left by $1 - t$, and checking that the result is the constant power series 1, in other words that all coefficients except the one of t^0 vanish. Moreover there can be no other power series with this property. The left-hand side therefore designates the multiplicative inverse of $1 - t$ in the ring of power series. Expressions for the ordinary generating function of other sequences are easily derived from this one. For instance, the substitution $t \to at$ gives the generating function for the geometric sequence $1, a, a^2, a^3, \ldots$ for any constant a:

$$\sum_{n=0}^{\infty} (at)^n = \frac{1}{1-at}.$$

(The equality also follows directly from the fact that the left-hand side is the Maclaurin series expansion of the right-hand side.) In particular,

$$\sum_{n=0}^{\infty} (-1)^n t^n = \frac{1}{1+t}.$$

One can also introduce regular "gaps" in the sequence by replacing x by some power of x, so for instance for the sequence 1, 0, 1, 0, 1, 0, 1, 0, ... one gets the generating function

$$\sum_{n=0}^{\infty} t^{2n} = \frac{1}{1-t^2}.$$

By squaring the initial generating function, or by finding the derivative of both sides with respect to x and making a change of running variable $n \to n-1$, one sees that the coefficients form the sequence 1, 2, 3, 4, 5, ..., so one has

$$\sum_{n=0}^{\infty} (n+1) t^n = \frac{1}{(1-t)^2},$$

and the third power has as coefficients the triangular numbers 1, 3, 6, 10, 15, 21, ... whose term n is the binomial coefficient $\binom{n+2}{2}$, so that

$$\sum_{n=0}^{\infty} \binom{n+2}{2} t^n = \frac{1}{(1-t)^3}$$

More generally, for any positive integer k, it is true that

$$\sum_{n=0}^{\infty} \binom{n+k}{k} t^n = \frac{1}{(1-t)^{k+1}}$$

Note that, since

$$2\binom{n+2}{2} - 3\binom{n+1}{1} + \binom{n}{0} = 2\frac{(n+1)(n+2)}{2} - 3(n+1) + 1 = n^2,$$

one can find the ordinary generating function for the sequence 0, 1, 4, 9, 16, ... of square numbers by linear combination of binomial-coefficient generating sequences;

$$G(n^2;t) = \sum_{n=0}^{\infty} n^2 t^n = \frac{2}{(1-t)^3} - \frac{3}{(1-t)^2} + \frac{1}{1-t} = \frac{t(t+1)}{(1-t)^3}$$

If a_n is the probability mass function of a discrete random variable, then its ordinary generating function is called a probability-generating function.[2]

Hence for a probability generating function if we define $P(T=n) = a_n$, we will have the following:

$$\sum_{n=0}^{\infty} a_n = 1 \text{ and } G(a_n;t) = \sum_{n=0}^{\infty} a_n t^n = \sum_{n=0}^{\infty} P(T=n) t^n = E(t^n) \quad *$$

*Recall that for a discrete random variable *X*, $E\big(f(x)\big) = \sum_{x=0}^{\infty} f(x) P(X=x)$.

If *X* is a discrete random variable taking values in the non-negative integers {0,1, ...}, then the *probability generating function* of *X* can be defined as

$$G(z) = \sum_{n=0}^{\infty} z^x P(X=x) = E(z^x). \quad ^3$$

References:

[1] http://en.wikipedia.org/wiki/Probability-generating_function#Definition
[2] http://en.wikipedia.org/wiki/Generating_function
[3] http://en.wikipedia.org/wiki/Probability-generating_function