Probability Generating Functions are a useful tool with reference to *discrete* probability distributions, that is probabilities *P(x)* that are a function of *the positive integers x* = 0,1,2,3,4,5.... i.e. *Uniform, Geometric, Binomial and Poisson Distributions*, but *not* a Normal distribution of *continuous* random variables (i.e. where *x* can take any real value).

Definition: $G_x(z)$ is the Probability Generating Function (PGF) associated with discrete random variable *x*. *p(x)* is the probability of *x* occurring. *z* is a real number and $G_x(z)$ is the *expectation* of the quantity z^x.

$$\boxed{G_x(z) = E\left[z^x\right] = \sum_{x=0}^{\infty} p(x) z^x}$$

If we can construct $G_x(z)$ for a particular probability distribution, then its algebraic properties enable us to *efficiently compute useful properties of the distribution* (such as the *mean*, *variance* and the distribution of *combinations* of random variables). It works very much like a *Moment Generating Function* (MGF), but is *restricted to discrete random variables.*

$$\boxed{G_x(1) = \sum_{x=0}^{\infty} p(x) = 1}$$

by *definition* of the probability distribution *p(x)*

Mean

$$\left.\frac{\partial G_x}{\partial z}\right|_{z=1} = \sum_{x=0}^{\infty} x p(x) z^{x-1}\Big|_{z=1} = \sum_{x=0}^{\infty} x p(x) = E[x]$$

$$\therefore \boxed{\mu = E[x] = G_x^{'}(1)}$$

Variance

$$\left.\frac{\partial^2 G_x}{\partial z^2}\right|_{z=1} = \sum_{x=0}^{\infty} x(x-1)\, p(x) z^{x-2}\Big|_{z=1} = \sum_{x=0}^{\infty} x^2 p(x) - \sum_{x=0}^{\infty} x p(x) = V[x] + \mu^2 - \mu$$

$$\therefore \boxed{\sigma^2 = V[x] = G_x^{''}(1) - \mu^2 + \mu}$$

Example: Uniform distribution

$$p(x) = \tfrac{1}{N}$$

$$x = 1 \ldots N$$

$$G_x(z) = \tfrac{1}{N} z^1 + \tfrac{1}{N} z^2 + \tfrac{1}{N} z^3 \ldots + \tfrac{1}{N} z^N$$

$$G_x^{'}(z) = \frac{\partial G_x}{\partial z} = \tfrac{1}{N} z^0 + \tfrac{2}{N} z^1 + \tfrac{3}{N} z^2 + .. + \tfrac{N}{N} z^{N-1}$$

$$G_x^{'}(1) = \tfrac{1}{N} + \tfrac{2}{N} + \ldots \tfrac{N}{N} = \tfrac{1}{2} N\left(\tfrac{1}{N} + \tfrac{N}{N}\right) = \tfrac{1}{2}(N+1)$$

Sum of arithmetic progression

$$\therefore \boxed{\mu = \tfrac{1}{2}(N+1)}$$

Mean

$$G_x^{''}(z) = \frac{\partial^2 G_x}{\partial z^2} = \tfrac{2}{N} z^0 + \tfrac{6}{N} z^1 + \ldots + \tfrac{N(N-1)}{N} z^{N-2}$$

$$G_x^{''}(1) = \tfrac{2}{N} + \tfrac{6}{N} + \ldots + \tfrac{N(N-1)}{N} = \tfrac{1}{N}\sum_{n=1}^{N-1} n(n+1) = \tfrac{1}{N}\sum_{n=1}^{N-1}\left(n^2 + n\right)$$

$$\sum_{n=1}^{N} n = \tfrac{1}{2} N(N+1)$$

$$\sum_{n=1}^{N} n^2 = \tfrac{1}{6} N(2N+1)N+1)$$

Quote these summation formulae

$$\tfrac{1}{N}\sum_{n=1}^{N-1}\left(n^2 + n\right) = \tfrac{1}{N}\left(\tfrac{1}{6} N(2N+1)(N+1) + \tfrac{1}{2} N(N+1) - N^2 - N\right)$$

$$\tfrac{1}{N}\sum_{n=1}^{N-1}\left(n^2 + n\right) = \tfrac{1}{6}(2N+1)(N+1) + \tfrac{1}{2}(N+1) - (N+1)$$

$$\tfrac{1}{N}\sum_{n=1}^{N-1}\left(n^2 + n\right) = \tfrac{1}{12}(N+1)\{4N + 2 + 6 - 12\}$$

$$\tfrac{1}{N}\sum_{n=1}^{N-1}\left(n^2 + n\right) = \tfrac{1}{12}(N+1)(4N-4)$$

$$\therefore \sigma^2 = \tfrac{1}{12}(N+1)(4N-4) + \tfrac{1}{2}(N+1) - \tfrac{1}{4}(N+1)^2$$

$$\sigma^2 = \tfrac{1}{12}(N+1)(4N - 4 + 6 - 3N - 3)$$

$$\boxed{\sigma^2 = \tfrac{1}{12}(N+1)(N-1)}$$

Variance

$$\boxed{\begin{aligned} \mu &= E[x] = G_x^{'}(1) \\ \sigma^2 &= V[x] = G_x^{''}(1) - \mu^2 + \mu \end{aligned}}$$

$$\mu = E[x] = G_x^{'}(1)$$
$$\sigma^2 = V[x] = G_x^{''}(1) - \mu^2 + \mu$$

Example: non uniform distribution

$x = 1, 2, 3$

$p(1) = 0.5$

$p(2) = 0.2$

$p(3) = 0.3$

$$G_x(z) = 0.5z + 0.2z^2 + 0.3z^3$$
$$G_x^{'}(z) = 0.5 + 0.4z + 0.9z^2$$
$$G_x^{''}(z) = 0.4 + 1.8z$$

$\mu = G_x^{'}(1) = 0.5 + 0.4 + 0.9 = 1.8$

$\sigma = G_x^{''}(1) - \mu^2 + \mu = 0.4 + 1.8 - 1.8^2 + 1.8 = \boxed{0.76}$

Example: Geometric Distribution

$x = 1 \ldots \infty$

$P(x) = (1-p)^{x-1} p$

$G_x(z) = pz + (1-p)pz^2 + (1-p)^2 pz^3 + \ldots + (1-p)^{N-1} pz^N + \ldots$

Geometric progression with first term *pz* and common ratio *(1-p)z*

$$G_x(z) = \lim_{N \to \infty} \left(\frac{pz(1-(1-p)^N z^N)}{1-(1-p)z} \right)$$

$$G_x(z) = \frac{pz}{1-(1-p)z} \quad , \quad s.t. \ |(1-p)z| < 1$$

Unless *p* = 1, this will always be true when *z* = 1 since *p* is in the range [0,1]

$$G_x^{'}(z) = \frac{(1-(1-p)z)p - pz(-(1-p))}{(1-(1-p)z)^2} = \frac{p - pz + p^2 z + pz - p^2 z}{(1-(1-p)z)^2} = \frac{p}{(1-(1-p)z)^2}$$

$$G_x^{''}(z) = \frac{-2p}{(1-(1-p)z)^3}(-(1-p)) = \frac{2p(1-p)}{(1-(1-p)z)^3}$$

$$G_x^{'}(1) = \frac{p}{p^2} = \frac{1}{p} \qquad G_x^{''}(1) = \frac{2p(1-p)}{p^3} = \frac{2(1-p)}{p^2}$$

$$\therefore \mu = \frac{1}{p} \qquad \therefore \sigma^2 = \frac{2(1-p)}{p^2} - \frac{1}{p^2} + \frac{1}{p} = \frac{2-2p-1+p}{p^2} = \frac{1-p}{p^2}$$

Example: Binomial Distribution

$x = 0, 1, 2, \ldots, N \qquad p(x \mid N, p) = \frac{N!}{(N-x)!x!} p^x (1-p)^{N-x}$

$$G_x(z) = \sum_{x=0}^{N} p(x \mid N, p) z^x$$

$$G_x(z) = \frac{N!}{(N-0)!0!}(pz)^0 (1-p)^N + \frac{N!}{(N-1)!1!}(pz)^1 (1-p)^{N-1} + \ldots$$

$$\ldots + \frac{N!}{(N-N)!N!}(pz)^N (1-p)^{N-N}$$

$$G_x(z) = (pz + 1 - p)^N$$

$$G_x^{'}(z) = Np(pz + 1 - p)^{N-1}$$

$$G_x^{''}(z) = N(N-1)p^2 (pz + 1 - p)^{N-2}$$

$$\mu = G_x^{'}(1) = Np$$

$$\sigma = G_x^{''}(1) - \mu^2 + \mu = N(N-1)p^2 - N^2 p^2 + Np = -Np^2 + Np = Np(1-p)$$

Example: Poisson Distribution

$x = 0, 1, 2, ..$

$$p(x \mid \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

$$G_x(z) = \sum_{x=0}^{N} p(x \mid \lambda) z^x$$

$$G_x(z) = e^{-\lambda} \left(\frac{(\lambda z)^0}{0!} + \frac{(\lambda z)^1}{1!} + \ldots + \frac{(\lambda z)^x}{x!} + \ldots \right)$$

$$G_x(z) = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}$$

$$G_x^{'}(z) = \lambda e^{\lambda(z-1)}$$

$$G_x^{''}(z) = \lambda^2 e^{\lambda(z-1)}$$

$$\mu = G_x^{'}(1) = \lambda$$

$$\sigma = G_x^{''}(1) - \mu^2 + \mu = \lambda^2 - \lambda^2 + \lambda$$

$$\sigma = \lambda$$

Combining discrete random variables

$$\boxed{y=\sum_{i=1}^{n} a_i x_i}$$

Consider a weighted sum of random variables $x_1, x_2, \ldots x_n$

$$\therefore G_y(z)=E\left[z^y\right]=E\left[z^{\sum_{i=1}^{n} a_i x_i}\right]=E\left[z^{a_1x_1}z^{a_2x_2}\ldots z^{a_nx_n}\right]$$

If two random variables have zero covariance, this means there is no correlation between them. In this case we say each random variable is *independent*.

$$\operatorname{cov}[x,y]=E[xy]-E[x]E[y]$$

$$\operatorname{cov}[x,y]=0 \Rightarrow E[xy]=E[x]E[y]$$

Therefore if n random variables $x_1, x_2, \ldots x_n$ are all *independent*

$$G_y(z)=E\left[z^{a_1x_1}z^{a_2x_2}\ldots z^{a_nx_n}\right]=E\left[z^{a_1x_1}\right]E\left[z^{a_2x_2}\right]\ldots E\left[z^{a_nx_n}\right]$$

$$\therefore \boxed{G_y(z)=G_{x_1}(z^{a_1})\times G_{x_2}(z^{a_2})\times\ldots G_{x_n}(z^{a_n})}$$

So if random variables x,y are *independent*

$$\boxed{G_{x+y}(z)=G_x(z)\times G_y(z)}$$

Example: Sum of two Poisson distributed variables

$$x,y=0,1,2,..$$

$$G_x(z)=e^{\lambda_1(z-1)}$$

$$G_y(z)=e^{\lambda_2(z-1)}$$

$$G_{x+y}(z)=G_x(z)\times G_y(z)$$

$$G_{x+y}(z)=e^{(\lambda_1+\lambda_2)(z-1)}$$

$$\boxed{\mu=G_x'(1)=\lambda_1+\lambda_2}$$

$$\sigma=G_x''(1)-\mu^2+\mu=(\lambda_1+\lambda_2)^2-(\lambda_1+\lambda_2)^2+\lambda_1+\lambda_2$$

$$\boxed{\sigma=\lambda_1+\lambda_2}$$

Which proves the result:

$$\boxed{\begin{array}{l} X\sim \mathrm{Po}(\lambda_1),\quad Y\sim \mathrm{Po}(\lambda_2) \\ X+Y\sim \mathrm{Po}(\lambda_1+\lambda_2)\end{array}}$$

Example: Sum of two six sided dice being rolled

Outcomes for each dice are 1,2,3,4,5,6. Probabilities for each one are 1/6. Let x be the outcome of dice 1 and y be the outcome of dice 2. Assume both are independent.

$$G_x(z)=\tfrac{1}{6}z^1+\tfrac{1}{6}z^2+\tfrac{1}{6}z^3+\tfrac{1}{6}z^4+\tfrac{1}{6}z^5+\tfrac{1}{6}z^6$$

$$G_y(z)=\tfrac{1}{6}z^1+\tfrac{1}{6}z^2+\tfrac{1}{6}z^3+\tfrac{1}{6}z^4+\tfrac{1}{6}z^5+\tfrac{1}{6}z^6$$

The PGF for the sum of x and y is therefore:

$$G_{x+y}(z)=G_x(z)\times G_y(z)$$

$$G_{x+y}(z)=\tfrac{1}{36}\left(z^1+z^2+z^3+z^4+z^5+z^6\right)^2$$

$$G_{x+y}(z)=\tfrac{1}{36}\left(x^2+2x^3+3x^4+4x^5+5x^6+6x^7+5x^8+4x^9+3x^{10}+2x^{11}+x^{12}\right)$$

Now from the definition of the PGF $w=x+y;\quad G_w(z)=\sum_{w=0}^{\infty}p(w)z^w$

Hence:

w	0	1	2	3	4	5	6	7	8	9	10	11	12
$p(w)$	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$\mu=\sum_{w=0}^{12}wp(w)=\tfrac{1}{36}\left(2+6+12+20+30+42+40+36+30+22+12\right)$$

$$\mu=\tfrac{252}{36}=\boxed{7}$$

$$\sigma^2=\sum_{w=0}^{12}w^2p(w)-\mu^2$$

$$\sigma^2=\tfrac{1}{36}\left(4+18+48+100+180+294+320+324+300+242+144\right)$$

$$\sigma^2=\tfrac{1974}{36}-7^2=\boxed{5\tfrac{5}{6}}$$

This result is consistent with what could be obtained using an *addition table*.

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

However, the PGF approach is *powerful* since it could be extended to *more than two dice*, or *dice with different numbers etc.*

Outcome	Freq	Probability
2	1	1/36 = 0.028
3	2	2/36 = 0.056
4	3	3/36 = 0.083
5	4	4/36 = 0.111
6	5	5/36 = 0.139
7	6	6/36 = 0.167
8	5	5/36 = 0.139
9	4	4/36 = 0.111
10	3	3/36 = 0.083
11	2	2/36 = 0.056
12	1	1/36 = 0.028