Probability Generating Functions – Ideas & Exercises

(16 pages; 27/2/25)

Contents

(1) Definition

(Applies to Discrete random variables that take non-negative integer values.)

$G_X(s) = E(s^X) = \sum_{k=0}^{\infty} p_k s^k$, where $p_k = Prob(X = k)$

When $|s| \leq 1, \sum_{k=0}^{\infty} p_k s^k \leq \sum_{k=0}^{\infty} p_k = 1$, so that the series converges for $|s| \leq 1$

As well as generating probabilities, the pgf can be used to establish the mean and variance.

(2) Bernouilli & Binomial distributions

For a Bernouilli (single trial Binomial) random variable,

$G_X(s) = q + ps$

For $Y \sim B(n, p)$:

$G_Y(s) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} s^k = \sum_{k=0}^{n} \binom{n}{k} (ps)^k q^{n-k} = (q + ps)^n$

Notes

(i) $n = 1$ gives the Bernouilli distribution

(ii) $G_Y(s) = [G_X(s)]^n$, where X has the Bernouilli distribution

(generally true when $Y = X_1 + \cdots + X_n$, where the X_i have the same distribution – see later on)

(3) Poisson distribution

Find the PGF for $X \sim Po(\lambda)$ $[P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}]$

$$G_X(s) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda}\left(e^{\lambda s}\right) = e^{\lambda(s-1)}$$

(4) Mean and Variance

$G_X(s) = \sum_{k=0}^{\infty} p_k s^k$

So $G_X'(s) = \sum_{k=0}^{\infty} k p_k s^{k-1}$

and $G_X'(1) = E[X]$

Similarly, $G_X''(1) = E[X(X-1)]$

Find an expression involving the derivatives of $G_X(s)$ for $Var(X)$

$Var(X) = E(X^2) - [E(X)]^2$

$= E[X(X-1)] + E[X] - [E(X)]^2$

$= G_X''(1) + G_X'(1) - [G_X'(1)]^2$

If $X \sim P_o(\lambda)$, prove that $Var(X) = \lambda$, given that $G_X(s) = e^{\lambda(s-1)}$

$G_X(s) = e^{\lambda(s-1)}$

$G_X'(s) = \lambda e^{\lambda(s-1)}$ & $G''_X(s) = \lambda^2 e^{\lambda(s-1)}$

$Var(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2$

$= \lambda^2 + \lambda - \lambda^2 = \lambda$

(5) Establishing the distribution from the pgf

Given that $G_X(s) = \frac{ps}{1-qs}$, what is the distribution of X?

$[G_X(s) = \frac{ps}{1-qs}]$

$G_X(s) = ps(1 + qs + (qs)^2 + \cdots)$

The coefficient of s^k is $p_k = pq^{k-1}$

So X has a Geometric distribution.

(6) Sum of independent random variables

If X & Y are independent random variables, then

$G_{X+Y}(s) = G_X(s)G_Y(s)$

Proof

$G_{X+Y}(s) = E(s^{X+Y}) = E(s^X s^Y)$

$= E(s^X)E(s^Y)$ (by independence)

$= G_X(s)G_Y(s)$

Example

If $X_1 \sim P_o(\lambda_1)$ & $X_2 \sim P_o(\lambda_2)$ and X_1 & X_2 are independent,

then $G_{X_1+X_2}(s) = G_{X_1}(s)G_{X_2}(s) = e^{\lambda_1(s-1)}e^{\lambda_2(s-1)} = e^{(\lambda_1+\lambda_2)(s-1)}$

$\Rightarrow X_1 + X_2 \sim P_o(\lambda_1 + \lambda_2)$

(See STEP 2015, P3, Q12 for a variation on this.)

Let $Y = a + bX$. Then $G_Y(s) = E(s^Y) = s^a E(s^{bx}) = s^a G_X(s^b)$

PGF of Negative Binomial distribution?

$X = X_1 + \cdots + X_n$, where $X_i \sim Geo(p)$

$G_{X_i}(s) = \frac{ps}{1-qs}$, so $G_X(s) = \left(\frac{ps}{1-qs}\right)^n$

[See Appendix for derivation from 1st principles.]

$E(X) = ?$

$G_X(s) = \left(\frac{ps}{1-qs}\right)^n$

So $G_X'(s) = n\left(\frac{ps}{1-qs}\right)^{n-1}.\frac{(1-qs)p-ps(-q)}{(1-qs)^2}$

and $E(X) = G_X'(1) = n\left(\frac{p}{1-q}\right)^{n-1}.\frac{(1-q)p-p(-q)}{(1-q)^2}$

$= n\frac{p}{p^2} = \frac{n}{p}$

(7) Obtaining probabilities from the pgf

Given $G_X(s)$, the p_k can be obtained by either of the following methods:

(a) expanding $G_X(s)$, to find the coefficient of s^k

(b) $p_k = \frac{1}{k!} G_X^{(k)}(0)$ (for $k > 0$)

(8) Sum of variable number of independent random variables

If $X_1, X_2, \ldots$ & N are independent random variables, where the X_i have pgf $G_X(s)$, then $S_N = X_1 + X_2 + \cdots + X_N$ has pgf

$G_{S_N}(s) = G_N\big(G_X(s)\big)$

Proof

$G_{S_N}(s) = E(s^{S_N}) = \sum_{n=0}^{\infty} E(\,s^{S_n})P(N = n)$

$= \sum_{n=0}^{\infty} E(s^{X_1} s^{X_2} \ldots s^{X_n})\, P(N = n)$

$= \sum_{n=0}^{\infty} E(s^{X_1})E(s^{X_2}) \ldots E(s^{X_n})\, P(N = n)$

(as the X_i are independent)

$= \sum_{n=0}^{\infty} \big(G_X(s)\big)^n\, P(N = n) = G_N\big(G_X(s)\big)$

(With the same notation as above) $E(S_N) = E(N)E(X)$

Proof

$E(S_N) = G'_{S_N}(1)$

and $G'_{S_N}(s) = \frac{d}{ds}[G_N\big(G_X(s)\big)]$ (from earlier)

$= G'_N\big(G_X(s)\big)G'_X(s)$

[noting that $G'_N\big(G_X(s)\big)$ means the derivative wrt $G_X(s)$]

So $E(S_N) = G'_{S_N}(1) = G'_N\big(G_X(1)\big)G'_X(1)$

$= G'_N(1)G'_X(1) \quad$, as $G_X(1) = \sum_{k=0}^{\infty} p_k = 1$

$= E(N)E(X)$

(With the same notation as above)

$Var(S_N) = E(N)Var(X) + Var(N)[E(X)]^2$

[See PGF Exercises for proof.]

['Poisson hen'] A hen lays N eggs, where $N \sim P_o(\lambda)$, and each egg has probability p of hatching. It can be shown that the total number of eggs that hatch $\sim P_o(\lambda p)$.

[See PGF Exercises for proof.]

Appendix: pgf of Negative Binomial random variable

Approach A (From 1st Principles)

Negative Binomial: $p_k = \binom{k-1}{n-1} p^{n-1} q^{(k-1)-(n-1)} p$

(probability of nth success on kth attempt

$= P(n-1\ successes\ in\ k-1\ trials) \times P(success\ on\ kth\ trial)$)

$= \binom{k-1}{n-1} p^n q^{k-n}\ \ (k \geq n)$

$G_Y(s) = \sum_{k=n}^{\infty} \binom{k-1}{n-1} p^n q^{k-n} s^k$

$= (ps)^n \sum_{k-n=0}^{\infty} \binom{k-1}{(k-1)-(n-1)} q^{k-n} s^{k-n}$

$= (ps)^n \sum_{k-n=0}^{\infty} \binom{k-1}{k-n} (qs)^{k-n}$

$= (ps)^n \sum_{r=0}^{\infty} \binom{n+r-1}{r} (qs)^r$

$= (ps)^n \{1 + nqs + \binom{n+1}{2} (qs)^2 + \binom{n+2}{3} (qs)^3 + \cdots\}$

$= (ps)^n \{1 + nqs + \frac{(n+1)n}{2!} (qs)^2 + \frac{(n+2)(n+1)n}{3!} (qs)^3 + \cdots\}$

$= (ps)^n \{1 + (-n)(-qs) + \frac{(-n)(-n-1)}{2!} (-qs)^2$

$+ \frac{-n(-n-1)(-n-2)n}{3!} (-qs)^3 + \cdots\}$

$= (ps)^n (1-qs)^{-n} = \left(\frac{ps}{1-qs}\right)^n$

Approach B: Using Geometric pgf (see Exercises)

Notes

(a) $n = 1$ gives the Geometric distribution

(b) $G_Y(s) = [G_X(s)]^n$, where X has the Geometric distribution