Probability Generating Functions

Andreas Klappenecker

Texas A&M University

Definition

Let X be a discrete random variable defined on a probability space with probability measure Pr. Assume that X has non-negative integer values. The **probability generating function** of X is defined by

$$G_X(z) = \mathrm{E}[z^X] = \sum_{k=0}^{\infty} \Pr[X = k] z^k.$$

This series converges for all z with $|z| \leqslant 1$.

Expected Value

Expectation

The expectation value can be expressed by

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} k \Pr[X = k] = G_X'(1), \tag{1}$$

where $G_X'(z)$ denotes the derivative of $G_X(z)$.

Indeed, $G_X'(z) = \sum_{k=0}^{\infty} k \Pr[X = k] z^{k-1} = \sum_{k=1}^{\infty} k \Pr[X = k] z^{k-1}$.

Second Moment

$$\mathrm{E}[X^2] = G''_X(1) + G'_X(1)$$

Indeed,

$$G'_X(z) = \sum_{k=1}^{\infty} k \Pr[X = k] z^{k-1}$$

and

$$G''_X(z) = \sum_{k=2}^{\infty} k(k-1) \Pr[X = k] z^{k-2} = \sum_{k=2}^{\infty} (k^2 - k) \Pr[X = k] z^{k-1}.$$

Variance

$$\begin{aligned} \mathrm{Var}[X] &= E[X^2] - E[X]^2 \\ &= G_X''(1) + G_X'(1) - G_X'(1)^2. \end{aligned}$$

Example

Let X be a random variable that has Bernoulli distribution with parameter p. The probability generating function is given by

$$G_X(z) = (1-p) + pz.$$

Hence $G_X'(z) = p$, and $G''(z) = 0$. We obtain $\mathrm{E}[X] = G_X'(1) = p$ and

$$\mathrm{Var}[X] = G_X''(1) + G_X'(1) - G_X'(1)^2 = 0 + p - p^2 = p(1-p).$$

Example

The probability generating function of a geometrically distributed random variable X is

$$G(z) = \sum_{k=1}^{\infty} p(1-p)^{k-1} z^k = pz \sum_{k=0}^{\infty} (1-p)^k z^k = \frac{pz}{1-(1-p)z}.$$

Some calculus shows that

$$G'(z) = \frac{p}{(1-(1-p)z)^2}, \qquad G''(z) = \frac{2p(1-p)}{(1-(1-p)z)^3}.$$

Therefore, the expectation value is $\mathrm{E}[X] = G_X'(1) = 1/p$. The variance is given by

$$\mathrm{Var}[X] = G''(1) + G'(1) - G'(1)^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Proposition

Let $X_1, \ldots, X_n$ be independent $\mathbf{Z}_{\geqslant}$-valued random variables with probability generating functions $G_{X_1}(z), \ldots, G_{X_n}(z)$. The probability generating function of $X = X_1 + \cdots + X_n$ is given by the product

$$G_X(z) = \prod_{k=1}^{n} G_{X_k}(z).$$

Proof.

It suffices to show this for two random variables X and Y. The general case can be established by a straightforward induction proof.

$$
\begin{aligned}
G_X(z)G_Y(z) &= \left(\sum_{k=0}^{\infty} \Pr[X = k]z^k\right)\left(\sum_{k=0}^{\infty} \Pr[Y = k]z^k\right) \\
&= \sum_{k=0}^{\infty} z^k \left(\sum_{\ell=0}^{k} \Pr[X = \ell]\Pr[Y = k - \ell]\right) \\
&= \sum_{k=0}^{\infty} z^k \left(\sum_{\ell=0}^{k} \Pr[X = \ell, Y = k - \ell]\right) \\
&= \sum_{k=0}^{\infty} \sum_{\ell=0}^{k} \Pr[X + Y = k]z^k = G_{X+Y}(z) \qquad \square
\end{aligned}
$$

Example

Recall that the Bernoulli distribution with parameter p has generating function $(1-p)+pz$. If $X_1, \ldots, X_n$ are independent random variables that are Bernoulli distributed with parameter p, then $X = X_1 + \cdots + X_n$ is, by definition, binomially distributed with parameters n and p. The generating function of X is

$$G_X(z) = ((1-p)+pz)^n = \sum_{k=0}^{n} \binom{n}{k} (1-p)^{n-k} p^k z^k.$$

Example (Continued.)

We have

$$G_X'(z) = np((1-p) + pz)^{n-1}$$

The expected value is given by

$$\mathrm{E}[X] = G_X'(1) = np.$$

Example (Continued.)

We have

$$G_X'(z) = np((1-p)+pz)^{n-1}$$

and

$$G_X''(z) = n(n-1)p^2((1-p)+pz)^{n-2}.$$

The expected value is given by

$$\begin{aligned}\mathrm{Var}[X] &= G_X''(1) + G_X'(1) - G_X'(1)^2 \\ &= (n^2-n)p^2 + np - n^2p^2 \\ &= -np^2 + np = np(1-p)\end{aligned}$$

Uniqueness Theorem

Proposition

Let X and Y be discrete random variables with probability generating functions $G_X(z)$ and $G_Y(z)$, respectively. Then the probability generating function

$$G_X(z) = G_Y(z)$$

if and only if the probability distributions

$$\Pr[X = k] = \Pr[Y = k]$$

for all integers $k \geqslant 0$.

Proof.

If the probability distributions are the same, then evidently $G_X(z) = G_Y(z)$.

Conversely, suppose that the generating functions $G_X(z)$ and $G_Y(z)$ are the same. Since the radius of convergence is at least 1, we can expand the two generating funcions into power series

$$G_X(z) = \sum_{k=0}^{\infty} \Pr[X = k] z^k$$
$$G_Y(z) = \sum_{k=0}^{\infty} \Pr[Y = k] z^k$$

These two power series must have identical coefficients, since the generating functions are the same. Therefore, $\Pr[X = k] = \Pr[Y = k]$ for all $k \geqslant 0$, as claimed.

Number of Inversions

Inversion of a Permutation

Definition

Let $(a_1, a_2, \ldots, a_n)$ be a permutation of the set $\{1, 2, \ldots, n\}$. The pair (a_i, a_j) is called an **inversion** if and only if $i < j$ and $a_i > a_j$.

Example

The permutation $(3, 4, 1, 2)$ has the inversions

$$\{(3, 1), (3, 2), (4, 1), (4, 2)\}.$$

Definition

Let $I_n(k)$ denote the **number of permutations** on n points with k inversions.

Example

We have $I_n(0) = 1$, since only the identity has no inversions.

Example

We have $I_n(1) = n - 1$. Indeed, a permutation π can have a single inversion if and only if π is equal to a transposition of neighboring elements $(k+1, k)$ for some k in the range $1 \leqslant k \leqslant n-1$.

Example

Since no permutation can have more than $\binom{n}{2}$ inversions, we have

$$I_n(k) = 0 \quad \text{for all} \quad k > \binom{n}{2}.$$

Example

By reversal of the permutations, we have the symmetry

$$I_n\left(\binom{n}{2} - k\right) = I_n(k)$$

Probability Generating Function

Suppose that we choose permutations π uniformly at random from the symmetric group S_n.

Let X_n denote the random variable on S_n that assigns a permutation π its number of inversions. Then the probability generating function

$$G_{X_n}(z) = \sum_{k=0}^{\binom{n}{2}} \Pr[X_n = k] z^k$$

is given by

$$G_{X_n}(z) = \sum_{k=0}^{\binom{n}{2}} \frac{I_n(k)}{n!} z^k.$$

Question

Can we relate the generating functions $G_{X_n}(z)$ and $G_{X_{n-1}}(z)$?

Probability Generating Function

Question

Can we relate the generating functions $G_{X_n}(z)$ and $G_{X_{n-1}}(z)$?

Observation

Suppose that we have a permutation π_{n-1} on $\{1, 2, \dots, n-1\}$. If we insert the element n at position j with $1 \leqslant j \leqslant n$, then we get an additional $n - j$ inversions.

Example

1	2	8	3	4	5	6	7

Probability Generating Function

Example

1	2	3	4	5	6	7	8	additional inversions: 0
1	2	3	4	5	6	8	7	additional inversions: 1
				⋮				
8	1	2	3	4	5	6	7	additional inversions: 7

Observation

We need to insert n uniformly at random to obtain a uniformly distributed permutation on n elements from uniformly distributed permutations on $n-1$ elements.

Proposition

$$G_{X_n}(z) = \begin{cases} \dfrac{(1 + z + z^2 + \cdots + z^{n-1})}{n} G_{X_{n-1}}(z) & \textit{if } n > 1, \\ 1 & \textit{if } n = 1. \end{cases}$$

Corollary

$$G_{X_n}(z) = \prod_{k=1}^{n} \frac{1 + z + z^2 + \cdots + z^{k-1}}{k}$$

$$= \prod_{k=1}^{n} \frac{1 - z^k}{k(1 - z)} = \frac{1}{n!} \prod_{k=1}^{n} \frac{1 - z^k}{1 - z}$$

Factorization: Expected Value

In other words, the generating function $G_{X_n}(z)$ is the product of generating functions of discrete uniform random variables U_k on $\{0, 1, \ldots, k-1\}$,

$$G_{X_n}(z) = \prod_{k=1}^{n} G_{U_k}(z), \quad \text{where} \quad G_{U_k}(z) = \frac{1 + z + \cdots + z^{k-1}}{k}.$$

By the product rule for the derivative of products of n functions, we have

$$G'_{U_n}(z) = \prod_{k=1}^{n} G_{U_k}(z) \sum_{\ell=1}^{n} \frac{\frac{1}{n}(1 + 2z + 3z^2 + \cdots + \cdots + (\ell-1)z^{\ell-2})}{G_{U_\ell}(z)}.$$

Then

$$\mathrm{E}[X_n] = G'_{U_n}(1) = \sum_{k=1}^{n} \frac{k-1}{2} = \frac{n(n-1)}{4}.$$

Expected Value (Alternative Way)

Example (Creating a Permutation by Inserting One Element at a Time)

1							
2	1						
2	1	3					
$\vdots$		$\ddots$					
8	2	1	3	5	7	6	8

Observation

$$X_n = U_1 + U_2 + \cdots + U_n$$

$$\mathrm{E}[X_n] = \mathrm{E}[U_1] + \mathrm{E}[U_2] + \cdots + \mathrm{E}[U_n]$$

Expected Value

Proposition

$$\mathrm{E}[X_n] = \sum_{k=1}^{n} \mathrm{E}[U_k] = \sum_{k=1}^{n} \frac{k-1}{2} = \frac{n(n-1)}{4}.$$

Variance

Proposition

$$\mathrm{Var}[X_n] = \frac{2n^3 + 3n^2 - 5n}{72}.$$

Proof.

Since $X_n = U_1 + U_2 + \cdots + U_n$ and the U_k are mutually independent, we get

$$\begin{aligned}
\mathrm{Var}[X_n] &= \sum_{k=1}^{n} \mathrm{Var}[U_n] = \sum_{k=1}^{n} \frac{k^2-1}{12} = \frac{1}{12}\left(\sum_{k=1}^{n} k^2 - n\right) \\
&= \frac{1}{12}\left(\frac{2n^3+3n^2+n}{6} - n\right) \\
&= \frac{2n^3+3n^2-5n}{72}.
\end{aligned}$$

$\square$