

## 5.5: Power Series Distributions

*Power Series Distributions* are discrete distributions on (a subset of)  $\mathbb{N}$  constructed from power series. This class of distributions is important because most of the special, discrete distributions are power series distributions.

### Basic Theory

#### Power Series

Suppose that  $\mathbf{a} = (a_0, a_1, a_2, \dots)$  is a sequence of nonnegative real numbers. We are interested in the power series with  $\mathbf{a}$  as the sequence of coefficients. Recall first that the *partial sum* of order  $n \in \mathbb{N}$  is

$$g_n(\theta) = \sum_{k=0}^n a_k \theta^k, \quad \theta \in \mathbb{R} \quad (5.5.1)$$

The *power series*  $g$  is then defined by  $g(\theta) = \lim_{n \rightarrow \infty} g_n(\theta)$  for  $\theta \in \mathbb{R}$  for which the limit exists, and is denoted

$$g(\theta) = \sum_{n=0}^{\infty} a_n \theta^n \quad (5.5.2)$$

Note that the series converges when  $\theta = 0$ , and  $g(0) = a_0$ . Beyond this trivial case, recall that there exists  $r \in [0, \infty]$  such that the series converges absolutely for  $|\theta| < r$  and diverges for  $|\theta| > r$ . The number  $r$  is the *radius of convergence*. From now on, we assume that  $r > 0$ . If  $r < \infty$ , the series may converge (absolutely) or may diverge to  $\infty$  at the endpoint  $r$ . At  $-r$ , the series may converge absolutely, may converge conditionally, or may diverge.

#### Distributions

From now on, we restrict  $\theta$  to the interval  $[0, r)$ ; this interval is our *parameter space*. Some of the results below may hold when  $r < \infty$  and  $\theta = r$ , but dealing with this case explicitly makes the exposition unnecessarily cumbersome.

Suppose that  $N$  is a random variable with values in  $\mathbb{N}$ . Then  $N$  has the *power series distribution* associated with the function  $g$  (or equivalently with the sequence  $\mathbf{a}$ ) and with parameter  $\theta \in [0, r)$  if  $N$  has probability density function  $f_\theta$  given by

$$f_\theta(n) = \frac{a_n \theta^n}{g(\theta)}, \quad n \in \mathbb{N} \quad (5.5.3)$$

**Proof**

To show that  $f_\theta$  is a valid discrete probability density function, note that  $a_n \theta^n$  is nonnegative for each  $n \in \mathbb{N}$  and  $g(\theta)$ , by definition, is the normalizing constant for the sequence  $(a_n \theta^n : n \in \mathbb{N})$ .

Note that when  $\theta = 0$ , the distribution is simply the point mass distribution at 0; that is,  $f_0(0) = 1$ .

The distribution function  $F_\theta$  is given by

$$F_\theta(n) = \frac{g_n(\theta)}{g(\theta)}, \quad n \in \mathbb{N} \quad (5.5.4)$$

**Proof**

This follows immediately from the definitions since  $F_\theta(n) = \sum_{k=0}^n f_\theta(k)$  for  $n \in \mathbb{N}$

Of course, the probability density function  $f_\theta$  is most useful when the power series  $g(\theta)$  can be given in closed form, and similarly the distribution function  $F_\theta$  is most useful when the power series and the partial sums can be given in closed form

#### Moments

The moments of  $N$  can be expressed in terms of the underlying power series function  $g$ , and the nicest expression is for the factorial moments. Recall that the permutation formula is  $t^{(k)} = t(t-1) \cdots (t-k+1)$  for  $t \in \mathbb{R}$  and  $k \in \mathbb{N}$ , and the *factorial moment* of  $N$  of order  $k \in \mathbb{N}$  is  $\mathbb{E}(N^{(k)})$ .

For  $\theta \in [0, r)$ , the factorial moments of  $N$  are as follows, where  $g^{(k)}$  is the  $k$ th derivative of  $g$ .

$$\mathbb{E}(N^{(k)}) = \frac{\theta^k g^{(k)}(\theta)}{g(\theta)}, \quad k \in \mathbb{N} \quad (5.5.5)$$

Proof

Recall that a power series is infinitely differentiable in the open interval of convergence, and that the derivatives can be taken term by term. Thus

$$\mathbb{E}(N^{(k)}) = \sum_{n=0}^{\infty} n^{(k)} \frac{a_n \theta^n}{g(\theta)} = \frac{\theta^k}{g(\theta)} \sum_{n=k}^{\infty} a_n n^{(k)} \theta^{n-k} = \frac{\theta^k}{g(\theta)} g^{(k)}(\theta) \quad (5.5.6)$$

The mean and variance of  $N$  are

1.  $\mathbb{E}(N) = \theta g'(\theta) / g(\theta)$
2.  $\text{var}(N) = \theta^2 \left( g''(\theta) / g(\theta) - [g'(\theta) / g(\theta)]^2 \right)$

Proof

1. This follows from the previous result on factorial moments with  $k = 1$ .
2. This also follows from the previous result since  $\text{var}(N) = \mathbb{E}(N^{(2)}) + \mathbb{E}(N)[1 - \mathbb{E}(N)]$ .

The probability generating function of  $N$  also has a simple expression in terms of  $g$ .

For  $\theta \in (0, r)$ , the probability generating function  $P$  of  $N$  is given by

$$P(t) = \mathbb{E}(t^N) = \frac{g(\theta t)}{g(\theta)}, \quad t < \frac{r}{\theta} \quad (5.5.7)$$

Proof

For  $t \in (0, r/\theta)$ ,

$$P(t) = \sum_{n=0}^{\infty} t^n f_{\theta}(n) = \frac{1}{g(\theta)} \sum_{n=0}^{\infty} a_n (t\theta)^n = \frac{g(t\theta)}{g(\theta)} \quad (5.5.8)$$

## Relations

Power series distributions are closed with respect to sums of independent variables.

Suppose that  $N_1$  and  $N_2$  are independent, and have power series distributions relative to the functions  $g_1$  and  $g_2$ , respectively, each with parameter value  $\theta < \min\{r_1, r_2\}$ . Then  $N_1 + N_2$  has the power series distribution relative to the function  $g_1 g_2$ , with parameter value  $\theta$ .

Proof

A direct proof is possible, but there is an easy proof using probability generating functions. Recall that the PGF of the sum of independent variables is the product of the PGFs. Hence the PGF of  $N_1 + N_2$  is

$$P(t) = P_1(t)P_2(t) = \frac{g_1(\theta t)}{g_1(\theta)} \frac{g_2(\theta t)}{g_2(\theta)} = \frac{g_1(\theta t)g_2(\theta t)}{g_1(\theta)g_2(\theta)}, \quad t < \min\left\{\frac{r_1}{\theta}, \frac{r_2}{\theta}\right\} \quad (5.5.9)$$

The last expression is the PGF of the power series distribution relative to the function  $g_1 g_2$ , at  $\theta$ .

Here is a simple corollary.

Suppose that  $(N_1, N_2, \dots, N_k)$  is a sequence of independent variables, each with the same power series distribution, relative to the function  $g$  and with parameter value  $\theta < r$ . Then  $N_1 + N_2 + \dots + N_k$  has the power series distribution relative to the function  $g^k$  and with parameter  $\theta$ .

In the context of this result, recall that  $(N_1, N_2, \dots, N_k)$  is a *random sample* of size  $k$  from the common distribution.

## Examples and Special Cases

### Special Distributions

The Poisson distribution with rate parameter  $\lambda \in [0, \infty)$  is a power series distribution relative to the function  $g(\lambda) = e^\lambda$  for  $\lambda \in [0, \infty)$ .

**Proof**

This follows directly from the definition, since the PDF of the Poisson distribution with parameter  $\lambda$  is  $f(n) = e^{-\lambda} \lambda^n / n!$  for  $n \in \mathbb{N}$ .

The geometric distribution on  $\mathbb{N}$  with success parameter  $p \in (0, 1]$  is a power series distribution relative to the function  $g(\theta) = 1/(1 - \theta)$  for  $\theta \in [0, 1)$ , where  $\theta = 1 - p$ .

**Proof**

This follows directly from the definition, since the PDF of the geometric distribution on  $\mathbb{N}$  is  $f(n) = (1 - p)^n p = (1 - \theta)^n \theta$  for  $n \in \mathbb{N}$ .

For fixed  $k \in (0, \infty)$ , the negative binomial distribution on  $\mathbb{N}$  with stopping parameter  $k$  and success parameter  $p \in (0, 1]$  is a power series distribution relative to the function  $g(\theta) = 1/(1 - \theta)^k$  for  $\theta \in [0, 1)$ , where  $\theta = 1 - p$ .

**Proof**

This follows from the result above on [sums of IID variables](#), but can also be seen directly, since the PDF is

$$f(n) = \binom{n+k-1}{k-1} p^k (1-p)^n = (1-\theta)^k \binom{n+k-1}{k-1} \theta^n, \quad n \in \mathbb{N} \quad (5.5.10)$$

For fixed  $n \in \mathbb{N}_+$ , the binomial distribution with trial parameter  $n$  and success parameter  $p \in [0, 1]$  is a power series distribution relative to the function  $g(\theta) = (1 + \theta)^n$  for  $\theta \in [0, \infty)$ , where  $\theta = p/(1 - p)$ .

**Proof**

Note that the PDF of the binomial distribution is

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} = (1-p)^n \binom{n}{k} \left( \frac{p}{1-p} \right)^k = \frac{1}{(1+\theta)^n} \binom{n}{k} \theta^k, \quad k \in \{0, 1, \dots, n\} \quad (5.5.11)$$

where  $\theta = p/(1 - p)$ . This shows that the distribution is a power series distribution corresponding to the function  $g(\theta) = (1 + \theta)^n$ .

The logarithmic distribution with parameter  $p \in (0, 1)$  is a power series distribution relative to the function  $g(p) = -\ln(1 - p)$  for  $p \in [0, 1)$ .

**Proof**

This follows directly from the definition, since the PDF is

$$f(n) = \frac{1}{-\ln(1 - p)} \frac{1}{n} p^n, \quad n \in \mathbb{N} \quad (5.5.12)$$

This page titled [5.5: Power Series Distributions](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.