# DISCRETE PROBABILITY DISTRIBUTIONS

https://phitter.io 2024-03-02

### Abstract

This document provides an overview of the continuous probability distributions utilized in Phitter. It includes a detailed description for each distribution, covering aspects such as the definition, domain, parameter definitions and domains, probability density function, cumulative distribution function, percentile point function, raw moments, mean, variance, skewness, kurtosis, median, and mode in a concise and clear manner.

# Contents

1	Berr	noulli Distribution	5
	1.1	Distribution definition	5
	1.2	Distribution domain	5
	1.3	Parameters domain and parameters constraints	5
	1.4	Cumulative distribution function	5
	1.5	Probability density function	5
	1.6	Percent point function/Sample	
	1.7	Parametric centered moments	
	1.8	Parametric mean	
	1.9	Parametric variance	
	1.10	Parametric skewness	
		Parametric kurtosis	
		Parametric median	
		Parametric mode	
		Additional information and definitions	
		Spreadsheet documents	
	1.10	Spreadsheet documents	U
2	Bino	omial Distribution	7
	2.1	Distribution definition	7
	2.2	Distribution domain	
	2.3	Parameters domain and parameters constraints	
	2.4	Cumulative distribution function	
	2.5	Probability density function	
	$\frac{2.6}{2.6}$	Percent point function/Sample	
	$\frac{2.0}{2.7}$	Parametric centered moments	
	2.8	Parametric mean	
	2.9	Parametric variance	
		Parametric skewness	
		Parametric kurtosis	
		Parametric median	
		Parametric mode	
		Additional information and definitions	
	2.15	Spreadsheet documents	8
3	Geo	ometric Distribution	9
•	3.1	Distribution definition	_
	3.2	Distribution domain	
	3.3	Parameters domain and parameters constraints	
	3.4	Cumulative distribution function	9
	3.5	Probability density function	
	3.6	Percent point function/Sample	
	3.7	Parametric centered moments	
	3.8	Parametric mean	
	3.9	Parametric variance	
		Parametric skewness	
		Parametric kurtosis	
		Parametric median	
		Parametric mode	
		Additional information and definitions	
	3.15	Spreadsheet documents	10

4	Hyr	pergeometric Distribution	11
-	4.1	Distribution definition	
	4.1		
		Distribution domain	
	4.3	Parameters domain and parameters constraints	
	4.4	Cumulative distribution function	
	4.5	Probability density function	
	4.6	Percent point function/Sample	11
	4.7	Parametric centered moments	
	4.8	Parametric mean	11
	4.9	Parametric variance	11
	4.10	Parametric skewness	
		Parametric kurtosis	
		Parametric median	
		Parametric mode	
		Additional information and definitions	
	4.15	Spreadsheet documents	12
5	Lor	arithmic Distribution	13
J	5.1	Distribution definition	
	5.2	Distribution domain	
	5.3	Parameters domain and parameters constraints	
	5.4	Cumulative distribution function	
	5.5	Probability density function	
	5.6	Percent point function/Sample	
	5.7	Parametric centered moments	
	5.8	Parametric mean	
	5.9	Parametric variance	
	5.10	Parametric skewness	13
	5.11	Parametric kurtosis	13
	5.12	Parametric median	13
	5.13	Parametric mode	13
	5.14	Additional information and definitions	14
		Spreadsheet documents	
6	_	gative Binomial Distribution	<b>15</b>
	6.1	Distribution definition	
	6.2	Distribution domain	
	6.3	Parameters domain and parameters constraints	15
	6.4	Cumulative distribution function	15
	6.5	Probability density function	15
	6.6	Percent point function/Sample	
	6.7	Parametric centered moments	
	6.8	Parametric mean	
	6.9	Parametric variance	
	0.0	Parametric skewness	
		Parametric kurtosis	
		Parametric median	
		Parametric mode	15
		Additional information and definitions	
	6.15	Spreadsheet documents	16
7	Pois	sson Distribution	17
•	7.1	Distribution definition	17
	$7.1 \\ 7.2$	Distribution domain	
	7.3	Parameters domain and parameters constraints	
	7.4	Cumulative distribution function	
	7.5	FIODADIII V QCIISITV TUIICTIOII	J (

	7.6	Percent point function/Sample	7
	7.7	Parametric centered moments	7
	7.8	Parametric mean	7
	7.9	Parametric variance	7
	7.10	Parametric skewness	7
	7.11	Parametric kurtosis	7
	7.12	Parametric median	7
	7.13	Parametric mode	7
	7.14	Additional information and definitions	3
	7.15	Spreadsheet documents	3
8		form Distribution 19	_
	8.1	Distribution definition	
	8.2	Distribution domain	
	8.3	Parameters domain and parameters constraints	
	8.4	Cumulative distribution function	
	8.5	Probability density function	
	8.6	Percent point function/Sample	
	8.7	Parametric centered moments	
	8.8	Parametric mean	
	8.9	Parametric variance	)
		Parametric skewness	
		Parametric kurtosis	
	8.12	Parametric median	)
	8.13	Parametric mode	9
	8.14	Additional information and definitions	)
	8.15	Spreadsheet documents	)

## 1 Bernoulli Distribution

### 1.1 Distribution definition

$$X \sim \text{Bernoulli}(p)$$

### 1.2 Distribution domain

$$x \in \{0, 1\}$$

### 1.3 Parameters domain and parameters constraints

$$p \in (0,1) \subseteq \mathbb{R}$$

### 1.4 Cumulative distribution function

$$F_X(x) = \begin{cases} 1 - p & \text{if } x = 0\\ 1 & \text{if } x = 1 \end{cases}$$

### 1.5 Probability density function

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

### 1.6 Percent point function/Sample

$$F_X^{-1}(u) = \begin{cases} 1 & \text{if } u \le p \\ 0 & \text{if } u > p \end{cases}$$

#### 1.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{1} x^k f_X(x) = p$$

### 1.8 Parametric mean

$$Mean(X) = \mu_1' = p$$

### 1.9 Parametric variance

Variance(X) = 
$$(\mu'_2 - \mu'_1^2) = p(1-p)$$

### 1.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{1.5}} = \frac{1 - 2p}{\sqrt{p(1 - p)}}$$

### 1.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 + \frac{1 - 6p(1 - p)}{p(1 - p)}$$

### 1.12 Parametric median

Median(X) = 
$$\begin{cases} 0 & \text{if } p < 1/2\\ [0,1] & \text{if } p = 1/2\\ 1 & \text{if } p > 1/2 \end{cases}$$

### 1.13 Parametric mode

$$Mode(X) = \begin{cases} 0 & \text{if } p < 1/2\\ 0, 1 & \text{if } p = 1/2\\ 1 & \text{if } p > 1/2 \end{cases}$$

5

- $\bullet\,$  Excel file from GitHub repository
- $\bullet$  Google spreadsheet document

## 2 Binomial Distribution

### 2.1 Distribution definition

$$X \sim \text{Binomial}(n, p)$$

### 2.2 Distribution domain

$$x \in \mathbb{N} \equiv \{0, 1, 2, \dots\}$$

### 2.3 Parameters domain and parameters constraints

$$n \in \mathbb{N}, p \in (0,1) \subseteq \mathbb{R}$$

### 2.4 Cumulative distribution function

$$F_X(x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} = I(1-p, n-x, 1+x)$$

## 2.5 Probability density function

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

## 2.6 Percent point function/Sample

$$F_X^{-1}(u) = \arg\min_{x} |F_X(x) - u|$$

### 2.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{\infty} x^k f_X(x) = \sum_{i=0}^k \frac{n!}{(n-i)!} S(k,i) p^i$$

### 2.8 Parametric mean

$$Mean(X) = \mu'_1 = np$$

### 2.9 Parametric variance

Variance
$$(X) = (\mu'_2 - \mu'_1{}^2) = np(1-p)$$

### 2.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{1.5}} = \frac{1 - 2p}{\sqrt{np(1 - p)}}$$

### 2.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 + \frac{1 - 6p(1 - p)}{np(1 - p)}$$

### 2.12 Parametric median

$$Median(X) = \lfloor np \rfloor \vee \lceil np \rceil$$

### 2.13 Parametric mode

$$Mode(X) = \lfloor (n+1)p \rfloor \vee \lceil (n+1)p \rceil - 1$$

7

• Computing an analytic expression for the inverse of the cumulative distribution function is not feasible. However, it is possible to calculate the Percentile Point Function by approximating it to the nearest integer.

• u : Uniform[0,1] random varible

•  $\lfloor x \rfloor$ : Floor function

•  $\lceil x \rceil$ : Ceiling Function

•  $I\left(x,a,b\right)$  : Regularized incomplete beta function

• S(a,b) : Stirling numbers of the second kind =  $\frac{1}{b!}\sum_{j=0}^{b}(-1)^{b-j}\binom{b}{j}j^a$ 

- Excel file from GitHub repository
- Google spreadsheet document

## 3 Geometric Distribution

### 3.1 Distribution definition

$$X \sim \text{Geometric}(p)$$

### 3.2 Distribution domain

$$x \in \mathbb{N} \equiv \{0, 1, 2, \dots\}$$

### 3.3 Parameters domain and parameters constraints

$$p \in (0,1) \subseteq \mathbb{R}$$

### 3.4 Cumulative distribution function

$$F_X(x) = 1 - (1 - p)^{\lfloor x \rfloor}$$

### 3.5 Probability density function

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

## 3.6 Percent point function/Sample

$$F_X^{-1}(u) = \left[\frac{\ln(1-u)}{\ln(1-p)}\right]$$

### 3.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{\infty} x^k f_X(x) = \sum_{x=0}^{\infty} (1-p)^x p \cdot x^k$$

### 3.8 Parametric mean

$$Mean(X) = \mu_1' = \frac{1}{p}$$

### 3.9 Parametric variance

Variance
$$(X) = (\mu'_2 - \mu'^2_1) = \frac{1-p}{p^2}$$

### 3.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{1.5}} = \frac{2 - p}{\sqrt{1 - p}}$$

### 3.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 9 + \frac{p^2}{1 - p}$$

## 3.12 Parametric median

$$Median(X) = \left\lceil \frac{-1}{\log_2(1-p)} \right\rceil$$

$$Mode(X) = 1$$

•  $\lfloor x \rfloor$  : Floor function

•  $\lceil x \rceil$ : Ceiling Function

- $\bullet$  Excel file from GitHub repository
- Google spreadsheet document

## 4 Hypergeometric Distribution

4.1 Distribution definition

$$X \sim \text{Hypergeometric}(N, K, n)$$

4.2 Distribution domain

$$x \in \{\max(0, n + K - N), \min(n, K)\}\$$

4.3 Parameters domain and parameters constraints

$$N \in \mathbb{N}, K \in \{0 \dots N\}, n \in \{0 \dots N\}$$

4.4 Cumulative distribution function

$$F_X(x) = \sum_{i=0}^{x} \binom{K}{i} \binom{N-K}{n-i} / \binom{N}{n}$$

4.5 Probability density function

$$f_X(x) = \binom{K}{x} \binom{N-K}{n-x} / \binom{N}{n}$$

4.6 Percent point function/Sample

$$F_X^{-1}(u) = \arg\min_{x} |F_X(x) - u|$$

4.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=\max(0,n+K-N)}^{\min(n,K)} x^k f_X(x)$$

4.8 Parametric mean

$$Mean(X) = \mu_1' = \frac{nK}{N}$$

4.9 Parametric variance

Variance
$$(X) = (\mu'_2 - \mu'^2_1) = n \frac{K}{N} \frac{N - K}{N} \frac{N - n}{N - 1}$$

4.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{1.5}} = \frac{(N - 2K)(N - 1)^{\frac{1}{2}}(N - 2n)}{[nK(N - K)(N - n)]^{\frac{1}{2}}(N - 2)}$$

4.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 + \frac{1}{nK(N-K)(N-n)(N-2)(N-3)}$$

4.12 Parametric median

$$Median(X) = F_X^{-1}(0.5)$$

4.13 Parametric mode

$$Mode(X) = \left\lfloor \frac{(n+1)(K+1)}{N+2} \right\rfloor$$

11

• Computing an analytic expression for the inverse of the cumulative distribution function is not feasible. However, it is possible to calculate the Percentile Point Function by approximating it to the nearest integer.

• u : Uniform[0,1] random varible

•  $\lfloor x \rfloor$ : Floor function

• [x]: Ceiling Function

- Excel file from GitHub repository
- Google spreadsheet document

## 5 Logarithmic Distribution

### 5.1 Distribution definition

$$X \sim \text{Logarithmic}(p)$$

### 5.2 Distribution domain

$$x \in \mathbb{N}_{\geqslant 1} \equiv \{1, 2, \dots\}$$

## 5.3 Parameters domain and parameters constraints

$$p \in (0,1) \subseteq \mathbb{R}$$

### 5.4 Cumulative distribution function

$$F_X(x) = \sum_{i=0}^{x} \frac{1}{-\ln(1-p)} \frac{p^i}{i}$$

### 5.5 Probability density function

$$f_X(x) = \frac{1}{-\ln(1-p)} \frac{p^x}{x}$$

### 5.6 Percent point function/Sample

$$F_X^{-1}(u) = \arg\min_{x} |F_X(x) - u|$$

### 5.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{\infty} x^k f_X(x) = \frac{(k-1)!}{-\ln(1-p)} \left(\frac{p}{1-p}\right)^k$$

#### 5.8 Parametric mean

Mean(X) = 
$$\mu'_1 = \frac{1}{-\ln(1-p)} \frac{p}{1-p}$$

### 5.9 Parametric variance

Variance(X) = 
$$(\mu'_2 - \mu'_1)^2 = -\frac{p^2 + p \ln(1-p)}{(1-p)^2 (\ln(1-p))^2}$$

### 5.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3}{(\mu'_2 - \mu'^2_1)^{1.5}}$$

### 5.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2}$$

## 5.12 Parametric median

$$Median(X) = F_X^{-1}(0.5)$$

$$Mode(X) = 1$$

- Computing an analytic expression for the inverse of the cumulative distribution function is not feasible. However, it is possible to calculate the Percentile Point Function by approximating it to the nearest integer.
- u : Uniform[0,1] random varible

- Excel file from GitHub repository
- Google spreadsheet document

## 6 Negative Binomial Distribution

6.1 Distribution definition

 $X \sim \text{NegativeBinomial}(r, p)$ 

6.2 Distribution domain

$$x \in \mathbb{N} \equiv \{0, 1, 2, \dots\}$$

6.3 Parameters domain and parameters constraints

$$r\in\mathbb{N}_{\geqslant 1}, p\in(0,1)\subseteq\mathbb{R}$$

6.4 Cumulative distribution function

$$F_X(x) = I(p, r, x+1)$$

6.5 Probability density function

$$f_X(x) = \binom{r+x-1}{x} p^r (1-p)^x$$

6.6 Percent point function/Sample

$$F_X^{-1}(u) = \arg\min_{x} |F_X(x) - u|$$

6.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{\infty} x^k f_X(x)$$

6.8 Parametric mean

$$Mean(X) = \mu_1' = \frac{r(1-p)}{p}$$

6.9 Parametric variance

Variance(X) = 
$$(\mu'_2 - \mu'^2_1) = \frac{r(1-p)}{p^2}$$

6.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1}{(\mu'_2 - \mu'^2_1)^{1.5}} = \frac{2 - p}{\sqrt{r(1 - p)}}$$

6.11 Parametric kurtosis

$$Kurtosis(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 + \frac{6}{r} + \frac{p^2}{r(1-p)}$$

6.12 Parametric median

$$Median(X) = F_X^{-1}(0.5)$$

$$Mode(X) = \lfloor (r-1)(1-p)/p \rfloor$$

- Computing an analytic expression for the inverse of the cumulative distribution function is not feasible. However, it is possible to calculate the Percentile Point Function by approximating it to the nearest integer.
- u : Uniform[0,1] random varible
- $\lfloor x \rfloor$  : Floor function

- Excel file from GitHub repository
- Google spreadsheet document

## 7 Poisson Distribution

7.1 Distribution definition

$$X \sim \text{Poisson}(\lambda)$$

7.2 Distribution domain

$$x \in \mathbb{N} \equiv \{0, 1, 2, \dots\}$$

7.3 Parameters domain and parameters constraints

$$\lambda \in \mathbb{R}^+$$

7.4 Cumulative distribution function

$$F_X(x) = e^{-\lambda} \sum_{i=0}^{x} \frac{\lambda^i}{i!} = 1 - \frac{\gamma(x+1,\lambda)}{x!} = 1 - P(x-1,\lambda)$$

7.5 Probability density function

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

7.6 Percent point function/Sample

$$F_X^{-1}(u) = \arg\min_{x} |F_X(x) - u|$$

7.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=0}^{\infty} x^k f_X(x)$$

7.8 Parametric mean

$$Mean(X) = \mu'_1 = \lambda$$

7.9 Parametric variance

$$Variance(X) = (\mu_2' - \mu_1'^2) = \lambda$$

7.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{1.5}} = \lambda^{-1/2}$$

7.11 Parametric kurtosis

$$\operatorname{Kurtosis}(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 + \lambda^{-1}$$

7.12 Parametric median

$$Median(X) = \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$$

$$Mode(X) = \lfloor \lambda \rfloor$$

- Computing an analytic expression for the inverse of the cumulative distribution function is not feasible. However, it is possible to calculate the Percentile Point Function by approximating it to the nearest integer.
- u : Uniform[0,1] random varible
- $\lfloor x \rfloor$ : Floor function
- $P\left(a,x\right)=\frac{\gamma(a,x)}{\Gamma(a)}$  : Regularized lower incomplete gamma function
- $\gamma(a, x)$ : Lower incomplete Gamma function

- Excel file from GitHub repository
- Google spreadsheet document

## 8 Uniform Distribution

## 8.1 Distribution definition

$$X \sim \text{Uniform}(a, b)$$

#### 8.2 Distribution domain

$$x \in \{a, a+1, \dots, b-1, b\}$$

### 8.3 Parameters domain and parameters constraints

$$a \in \mathbb{N}, b \in \mathbb{N}, a < b$$

### 8.4 Cumulative distribution function

$$F_X(x) = \frac{x - a + 1}{b - a + 1}$$

## 8.5 Probability density function

$$f_X(x) = \frac{1}{b - a + 1}$$

### 8.6 Percent point function/Sample

$$F_X^{-1}(u) = [u(b-a+1) + a - 1]$$

### 8.7 Parametric centered moments

$$E[X^k] = \mu'_k = \sum_{x=a}^b x^k f_X(x) = \frac{1}{b-a+1} \sum_{x=a}^b x^k$$

### 8.8 Parametric mean

$$Mean(X) = \mu_1' = \frac{a+b}{2}$$

### 8.9 Parametric variance

Variance(X) = 
$$(\mu'_2 - \mu'_1)^2 = \frac{(b-a+1)^2 - 1}{12}$$

### 8.10 Parametric skewness

Skewness(X) = 
$$\frac{\mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1}{(\mu'_2 - \mu'^2_1)^{1.5}} = 0$$

### 8.11 Parametric kurtosis

$$\operatorname{Kurtosis}(X) = \frac{\mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2} = 3 - \frac{6((b-a+1)^2 + 1)}{5((b-a+1)^2 - 1)}$$

### 8.12 Parametric median

$$Median(X) = \frac{a+b}{2}$$

$$Mode(X) \in [a, b]$$

•  $\lceil x \rceil$  : Ceiling Function

- Excel file from GitHub repository
- ullet Google spreadsheet document

## References

- [1] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.
- [2] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17:261–272, 2020.