Compendium of Common Probability Distributions

Michael P. McLaughlin

July, 2014

Compendium of Common Probability Distributions   Second Edition, v2.7

Second printing, with corrections.

This document created using LaTeX and TeXShop. Figures produced using Mathematica.

Contents

Preface

This *Compendium* is part of the documentation for the software package, *Regress+*. The latter is a utility intended for univariate mathematical modeling and addresses both deteministic models (equations) as well as stochastic models (distributions). Its focus is on the modeling of empirical data so the models it contains are fully-parametrized variants of commonly used formulas.

This document describes the distributions available in *Regress+* (v2.7). Most of these are well known but some are not described explicitly in the literature. This *Compendium* supplies the formulas and parametrization as utilized in the software plus additional formulas, notes, etc. plus one or more plots of the density (mass) function.

There are 59 distributions in *Regress+* partitioned into four categories:

- Continuous: Symmetric (9)
- Continuous: Skewed (27)
- Continuous: Mixtures (11)
- Discrete: Standard (6)
- Discrete: Mixtures (6)

Formulas, where appropriate, include the following:

- Probability Density (Mass) Function: PDF
- Cumulative Distribution: CDF
- Characteristic Function: CF
- Central Moments (dimensioned): Mean, Variance, Skewness, Kurtosis
- Mode
- Quartiles: First, Second (Median), Third

Additional items include

- Notes relevant to use of the distribution as a model
- Possible aliases and special cases
- Characterization(s) of the distribution

In some cases, where noted, a particular formula may not available in any simple closed form. Also, the parametrization in *Regress+* may be just one of several used in the literature. In some cases, the distribution itself is specific to *Regress+*., e.g., **Normal&Normal1**.

Finally, there are many more distributions described in the literature. It is hoped that, for modeling purposes, those included here will prove sufficient in most studies.

Comments and suggestions are, of course, welcome. The appropriate email address is given below.

MICHAEL P. MCLAUGHLIN
MCLEAN, VA
APRIL 2013
MPMCL 'AT' CAUSASCIENTIA.ORG

Legend

$$\sim = \text{(is) distributed as}$$

$$\text{i.i.d.} = \text{independent and identically distributed}$$

$$\text{iff} = \text{if and only if}$$

$$u = \text{a Uniform[0, 1] random variate}$$

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} \quad \text{; binomial coefficient}$$

$$\lfloor z \rfloor = \mathit{floor}(z)$$

$$\mathit{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp\left(-t^2\right) dt \quad \text{; error function}$$

$$\Phi(z) = \frac{1}{2}\left[1 + \mathit{erf}\left(\frac{z}{\sqrt{2}}\right)\right] \quad \text{; standard normal CDF}$$

$$\begin{aligned} \Gamma(z) &= \int_0^\infty t^{z-1} \exp(-t)\, dt \quad \text{; Gamma function} \\ &= (z-1)! \quad \text{; for integer z} > 0 \end{aligned}$$

$$\Gamma(c; z) = \int_z^\infty t^{c-1} \exp(-t)\, dt \quad \text{; incomplete Gamma function}$$

$$\mathit{Beta}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad \text{; complete Beta function}$$

$$\mathit{Beta}(z; a, b) = \int_0^z t^{a-1}(1-t)^{b-1}\, dt \quad \text{; incomplete beta function}$$

$$\gamma = 0.57721566\ldots \quad \text{; EulerGamma}$$

$$\psi(z) = \Gamma'(z)/\Gamma(z) \quad \text{; digamma function}$$

$$I_n(z), K_n(z) = \text{the two solutions (modified Bessel functions) of } z^2\,y'' + z\,y' - \left(z^2 + n^2\right) y$$

$$\zeta(z) = \sum_{k=1}^{\infty} k^{-s} \quad \text{; Riemann zeta function, } s > 1$$

$${}_1F_1(z; a, b) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 \exp(z\,t)\,t^{a-1}(1-t)^{b-a-1}\,dt \quad \text{; confluent hypergeometric function}$$

$${}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}\,(1-t)^{c-b-1}(1-t\,z)^{-a}\,dt \quad \text{; Gauss hypergeometric function}$$

$$U(a, b, z) = \frac{1}{\Gamma(a)} \int_0^{\infty} \exp(-t)\,t^{b-1}\,dt \quad \text{; confluent hypergeometric function (second kind)}$$

$$Ei(z) = -\int_{-z}^{\infty} t^{-1}\exp(-t)\,dt \quad \text{; one of several exponential integrals}$$

$$L_n(x) = \text{one solution of } x\,y'' + (1-x)\,y' + n\,y \quad \text{; Laguerre polynomial}$$

$$Q_m(a, b) = \int_b^{\infty} x \left(\frac{x}{a}\right)^{m-1} I_{m-1}(ax)\exp\left(-\frac{x^2 + a^2}{2}\right) dx \quad \text{; Marcum Q function}$$

$$OwenT(z, \alpha) = \frac{1}{2\,\pi} \int_0^{\alpha} \frac{1}{1 + x^2} \exp\left[-\frac{z^2}{2}\left(1 + x^2\right)\right] dx \quad \text{; Owen's T function}$$

$$H_n(z) = \sum_{k=1}^{n} x^{-z} \quad \text{; Harmonic number}$$

$$Li_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^n} \quad \text{; polylogarithm function}$$

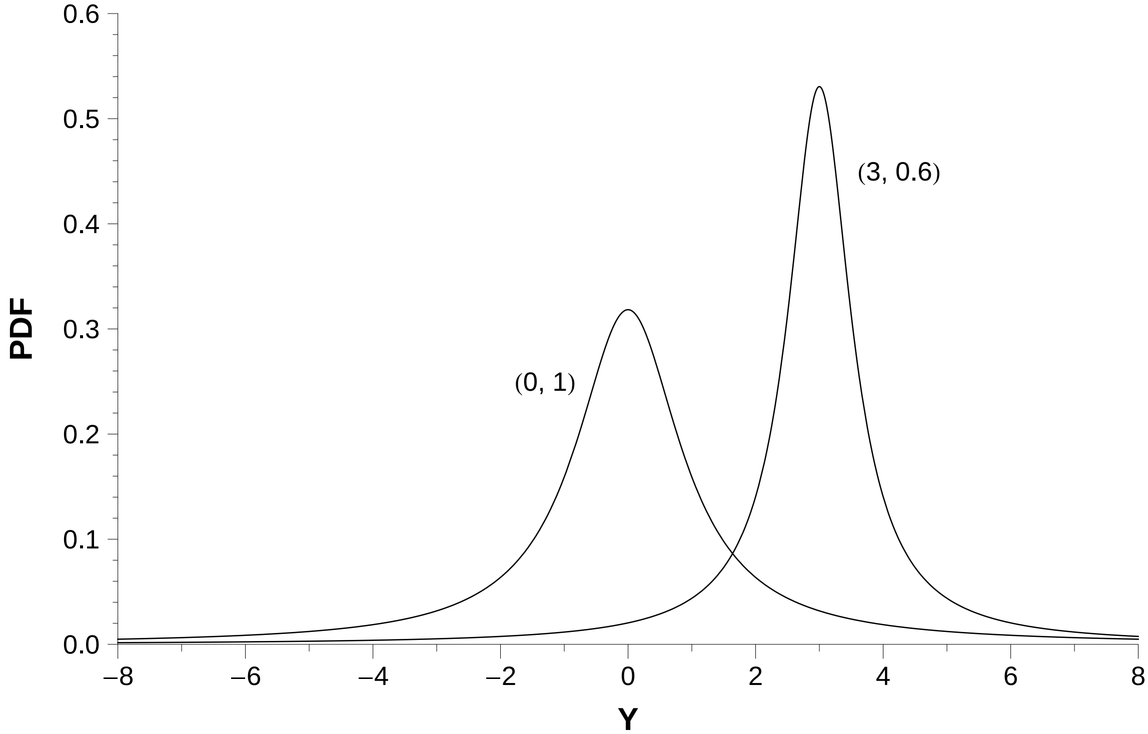
Part I

Continuous: Symmetric

These distributions are symmetric about their mean, usually denoted by parameter A.

Cauchy(A,B)

$\mathrm{B} > 0$



$$\mathrm{PDF} = \frac{1}{\pi B}\left[1 + \left(\frac{y-A}{B}\right)^2\right]^{-1}$$

$$\mathrm{CDF} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{y-A}{B}\right)$$

$$\mathrm{CF} = \exp(iAt - B|t|)$$

Parameters – A (θ): Location, B (λ): Scale

Moments, etc.

$$\text{Moments do not exist.}$$

$$\mathrm{Mode} = \mathrm{Median} = \mathrm{A}$$

$$\mathrm{Q1} = A - B \quad \mathrm{Q3} = A + B$$

$$\mathrm{RandVar} = A + B\tan\left[\pi\left(u - \frac{1}{2}\right)\right]$$

Notes

1. Since there are no finite moments, the location parameter (ostensibly the mean) does not have its usual interpretation for a symmetric distribution.

Aliases and Special Cases

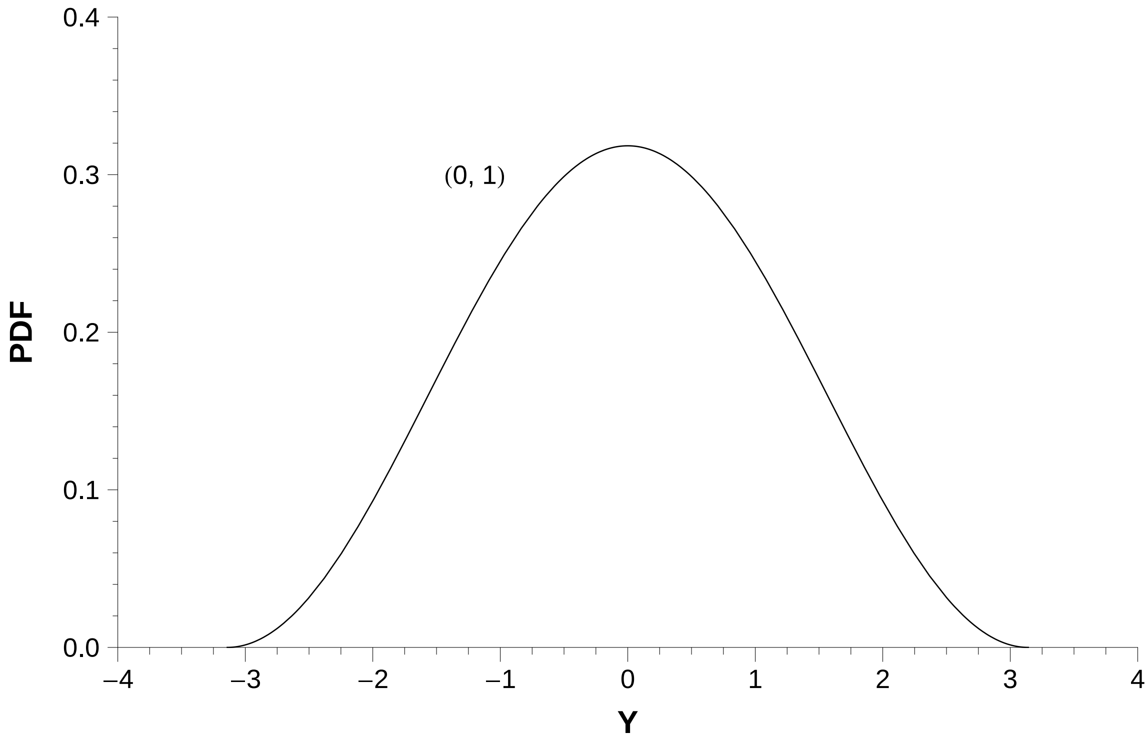
1. The Cauchy distribution is sometimes called the *Lorentz* distribution.

Characterizations

1. If U and V are $\sim$ **Normal**(0, 1), the ratio $U/V \sim Cauchy(0, 1)$.
2. If $Z \sim$ Cauchy, then $W = (a + b\,Z)^{-1} \sim Cauchy$.
3. If particles emanate from a fixed point, their points of impact on a straight line $\sim$ Cauchy.

Cosine(A,B)

$$\mathrm{A} - \pi\mathrm{B} \leq \mathrm{y} \leq \mathrm{A} + \pi\mathrm{B}, \quad B > 0$$



$$\mathrm{PDF} = \frac{1}{2\pi B}\left[1 + \cos\left(\frac{y-A}{B}\right)\right]$$

$$\mathrm{CDF} = \frac{1}{2\pi}\left[\pi + \frac{y-A}{B} + \sin\left(\frac{y-A}{B}\right)\right]$$

$$\mathrm{CF} = \frac{\exp(iAt)\sin(B\pi t)}{B\pi t\,(1-Bt)(1+Bt)}$$

Parameters – A: Location, B: Scale

Moments, etc.

$$\mathrm{Mean} = \mathrm{Median} = \mathrm{Mode} = A$$

$$\mathrm{Variance} = B^2\,(\pi^2 - 6)/3$$

$$\mathrm{Skewness} = 0$$

$$\mathrm{Kurtosis} = B^4\,(\pi^4 - 20\pi^2 + 120)/5$$

$$\mathrm{Q1} \approx A - 0.83171\,B \quad \mathrm{Q3} \approx A + 0.83171\,B$$

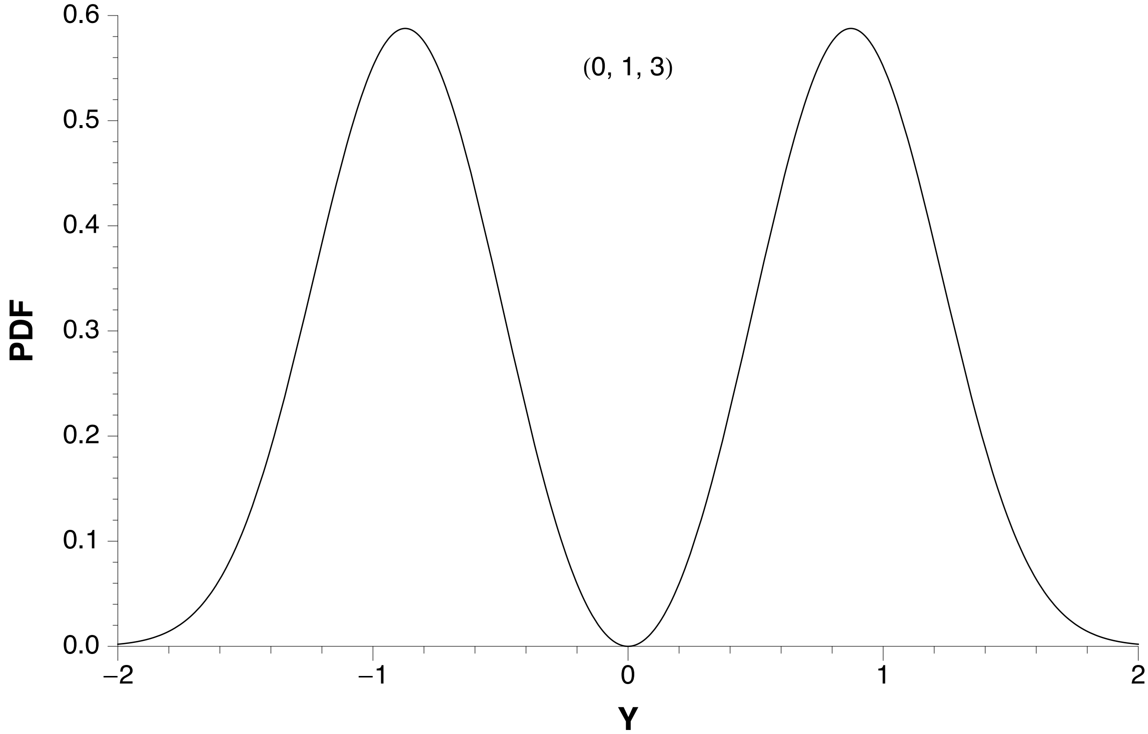
Notes

Aliases and Special Cases

Characterizations

1. The Cosine distribution is sometimes used as a simple, and more computationally tractable, alternative to the **Normal** distribution.

DoubleWeibull(A,B,C) $\qquad \mathrm{B, C} > 0$



$$\text{PDF} = \frac{C}{2B}\left|\frac{y-A}{B}\right|^{C-1}\exp\left(-\left|\frac{y-A}{B}\right|^{C}\right)$$

$$\text{CDF} = \begin{cases} \dfrac{1}{2}\exp\left[-\left(\dfrac{y-A}{B}\right)^{C}\right], & y \le A \\[2ex] 1-\dfrac{1}{2}\exp\left[-\left(\dfrac{y-A}{B}\right)^{C}\right], & y > A \end{cases}$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A: Location, B: Scale, C: Shape

Moments, etc.

$$\text{Mean} = \text{Median} = A$$

$$\text{Variance} = \Gamma\left(\frac{C+2}{C}\right)B^2$$

$$\text{Skewness} = 0$$

$$\text{Kurtosis} = \textit{undefined}$$

$$\text{Mode} = \textit{undefined} \quad \textbf{(bimodal when C > 1)}$$

$$\text{Q1, Q3} : \textit{no simple closed form}$$

Notes

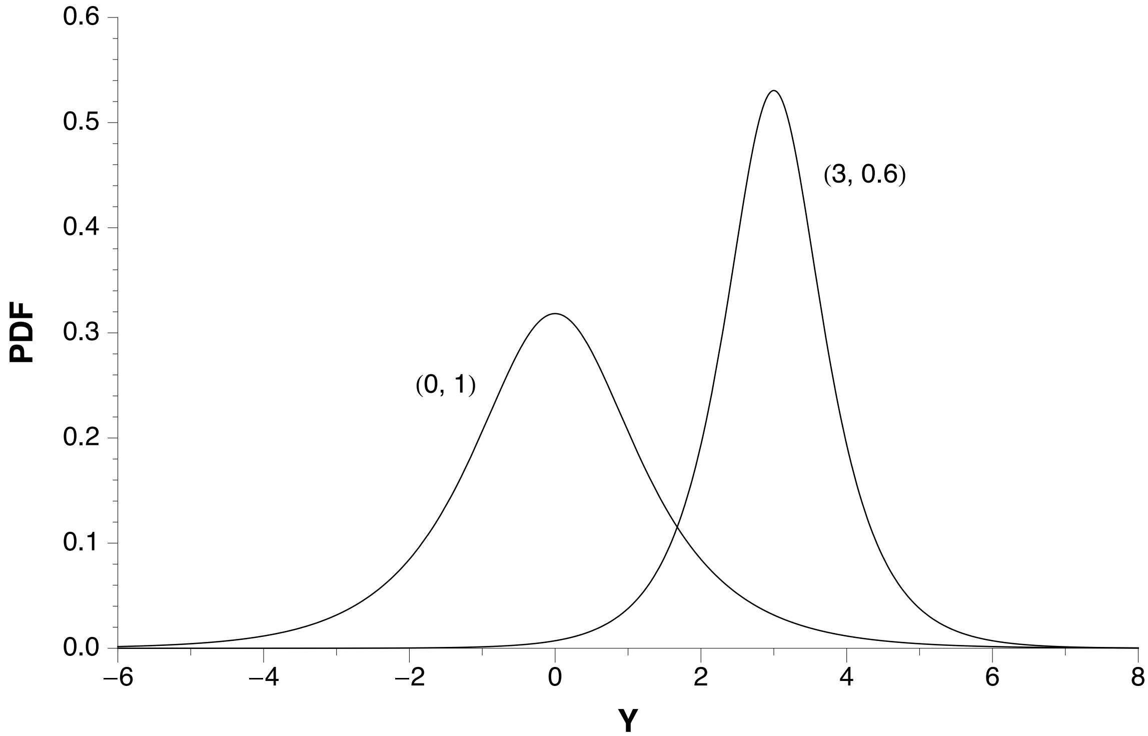
Aliases and Special Cases

1. The Double Weibull distribution becomes the **Laplace** distribution when C = 1.

Characterizations

1. The Double Weibull distribution is the signed analogue of the **Weibull** distribution.

HyperbolicSecant(A,B) $\quad B > 0$



$$\text{PDF} = \frac{1}{\pi\, B}\,\text{sech}\left(\frac{y-A}{B}\right)$$

$$\text{CDF} = \frac{2}{\pi}\tan^{-1}\left[\exp\left(\frac{y-A}{B}\right)\right]$$

$$\text{CF} = \exp(iAt)\,\text{sech}\left(\frac{\pi Bt}{2}\right)$$

Parameters – A (μ): Location, B (θ): Scale

Moments, etc.

$$\text{Mean} = \text{Median} = \text{Mode} = A$$

$$\text{Variance} = \frac{\pi^2}{4}B^2$$

$$\text{Skewness} = 0$$

$$\text{Kurtosis} = \frac{5\pi^4}{16}B^4$$

$$\text{Q1} = A - B\log\left(1+\sqrt{2}\right) \quad \text{Q3} = A + B\log\left(1+\sqrt{2}\right)$$

$$\text{RandVar} = A + B\log\left[\tan\left(\frac{\pi \text{u}}{2}\right)\right]$$

Notes

1. The HyperbolicSecant distribution is related to the **Logistic** distribution.

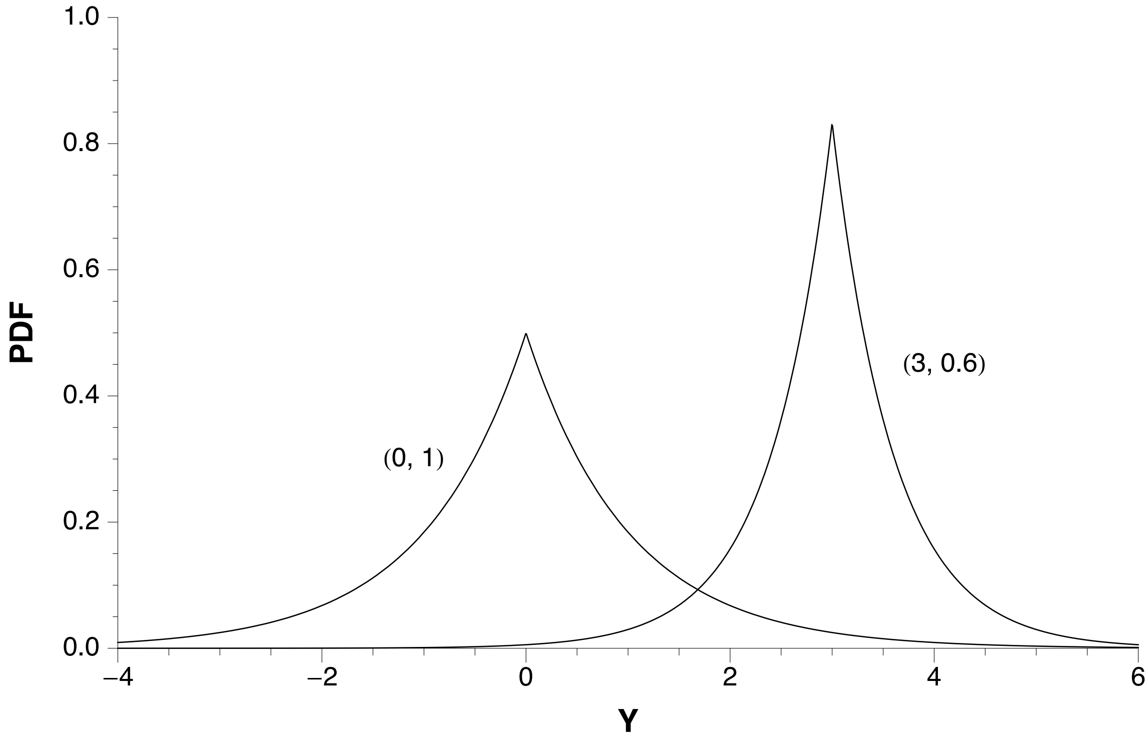
Aliases and Special Cases

Characterizations

1. The HyperbolicSecant distribution is often used in place of the **Normal** distribution when the tails are less than the latter would produce.

2. The HyperbolicSecant distribution is frequently used in lifetime analyses.

Laplace(A,B) $\qquad$ $\mathrm{B} > 0$



$$\mathrm{PDF} = \frac{1}{2\,B}\exp\left(-\frac{|y-A|}{B}\right)$$

$$\mathrm{CDF} = \begin{cases} \dfrac{1}{2}\exp\left(\dfrac{y-A}{B}\right), & y \le A \\ 1-\dfrac{1}{2}\exp\left(\dfrac{A-y}{B}\right), & y > A \end{cases}$$

$$\mathrm{CF} = \frac{\exp(iAt)}{1+B^2\,t^2}$$

Parameters – A (μ): Location, B (λ): Scale

Moments, etc.

$$\mathrm{Mean} = \mathrm{Median} = \mathrm{Mode} = A$$

$$\mathrm{Variance} = 2B^2$$

$$\mathrm{Skewness} = 0$$

$$\mathrm{Kurtosis} = 24B^4$$

$$\mathrm{Q1} = A - B\log(2) \quad \mathrm{Q3} = A + B\log(2)$$

$$\mathrm{RandVar} = A - B\log(u),\ \text{with a random sign}$$

Notes

Aliases and Special Cases

1. The Laplace distribution is often called the *double-exponential* distribution.

2. It is also known as the *bilateral exponential* distribution.

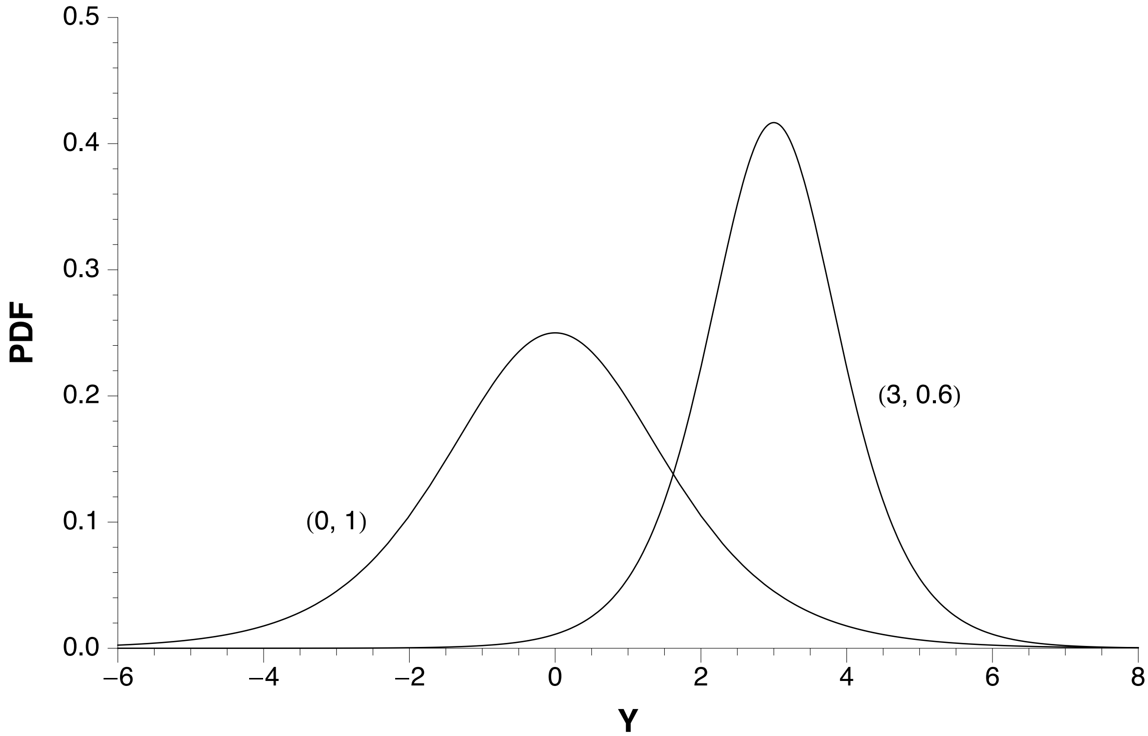
Characterizations

1. The Laplace distribution is the signed analogue of the **Exponential** distribution.

2. Errors of real-valued observations are often $\sim$Laplace or $\sim$**Normal**.

Logistic(A,B)

$\mathrm{B} > 0$



$$\mathrm{PDF} = \frac{1}{B}\exp\left(-\frac{y-A}{B}\right)\left[1+\exp\left(-\frac{y-A}{B}\right)\right]^{-2}$$

$$\mathrm{CDF} = \left[1+\exp\left(-\frac{y-A}{B}\right)\right]^{-1}$$

$$\mathrm{CF} = \frac{\pi B t}{\sinh(\pi B t)}\exp(iAt)$$

Parameters – A: Location, B: Scale

Moments, etc.

$$\mathrm{Mean} = \mathrm{Median} = \mathrm{Mode} = A$$

$$\mathrm{Variance} = \frac{\pi^2 B^2}{3}$$

$$\mathrm{Skewness} = 0$$

$$\mathrm{Kurtosis} = \frac{7\,\pi^4 B^4}{15}$$

$$\mathrm{Q1} = A - B\log(3) \quad \mathrm{Q3} = A + B\log(3)$$

$$\mathrm{RandVar} = A + B\log\left(\frac{u}{1-u}\right)$$

Notes

1. The Logistic distribution is often used as an approximation to other symmetric distributions due to the mathematical tractability of its CDF.

Aliases and Special Cases

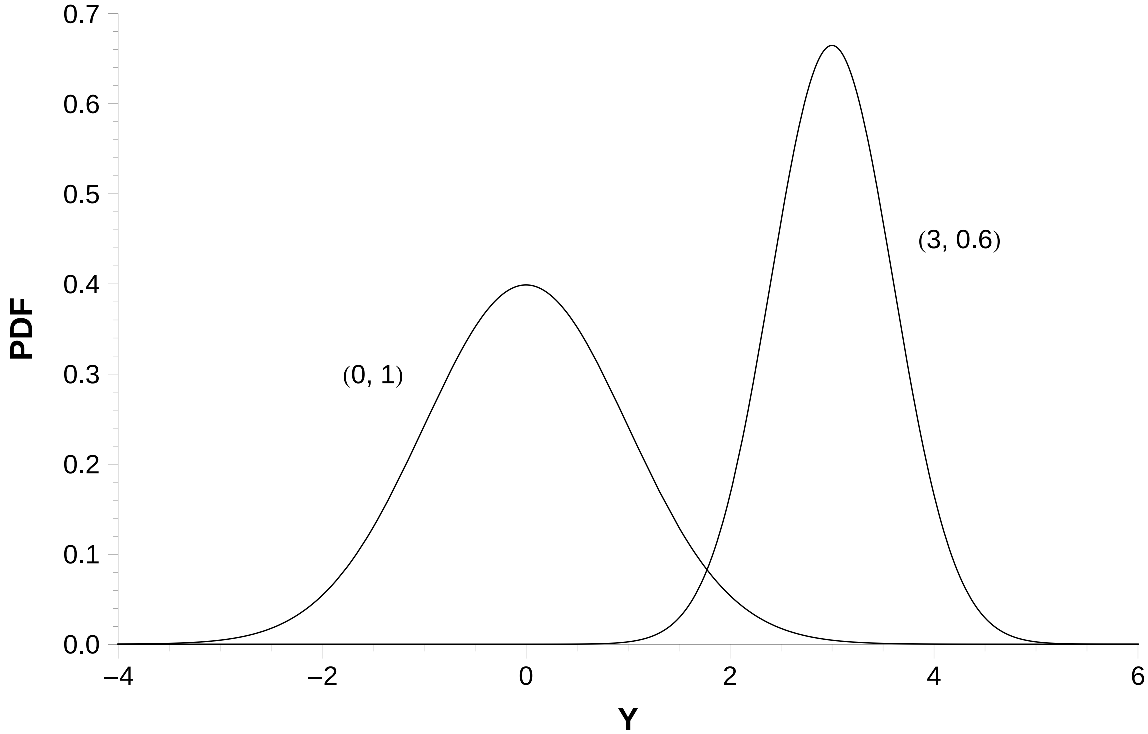
1. The Logistic distribution is sometimes called the *Sech-squared* distribution.

Characterizations

1. The Logistic distribution is used to describe many phenomena that follow the logistic law of growth.

2. If lo and hi are the minimum and maximum of a random sample (size = N), then, as $N \to \infty$, the asymptotic distribution of the midrange $\equiv (hi - lo)/2$ is $\sim$Logistic.

Normal(A,B) $\quad$ B > 0



$$\text{PDF} = \frac{1}{B\sqrt{2\pi}}\exp\left[-\frac{(y-A)^2}{2\,B^2}\right]$$

$$\text{CDF} = \frac{1}{2}\left[1-\textbf{erf}\left(\frac{A-y}{B\sqrt{2}}\right)\right] \equiv \Phi\left(\frac{y-A}{B}\right)$$

$$\text{CF} = \exp\left(iAt - \frac{B^2t^2}{2}\right)$$

Parameters – A (μ): Location, B (σ): Scale

Moments, etc.

$$\text{Mean} = \text{Median} = \text{Mode} = A$$

$$\text{Variance} = B^2$$

$$\text{Skewness} = 0$$

$$\text{Kurtosis} = 3B^4$$

$$\text{Q1} \approx A - 0.67449\,B \quad \text{Q3} \approx A + 0.67449\,B$$

Notes

1. The distribution is generally expressed in terms of the standard variable, z:

$$z = \frac{y - A}{B}$$

2. The sample standard deviation, s, is the maximum-likelihood estimator of B but is biased with respect to the population value. The latter may be estimated as follows:

$$B = \sqrt{\frac{N}{N-1}}\, s$$

   where N is the sample size.

Aliases and Special Cases

1. The Normal distribution is also called the *Gaussian* distribution and very often, in non-technical literature, the *bell curve*.
2. Its CDF is closely related to the error function, $erf(z)$.
3. The FoldedNormal and HalfNormal are special cases.

Characterizations

1. Let $Z_1, Z_2, \ldots, Z_N$ be i.i.d. be random variables with finite values for their mean, μ, and variance, σ^2. Then, for any real number, z,

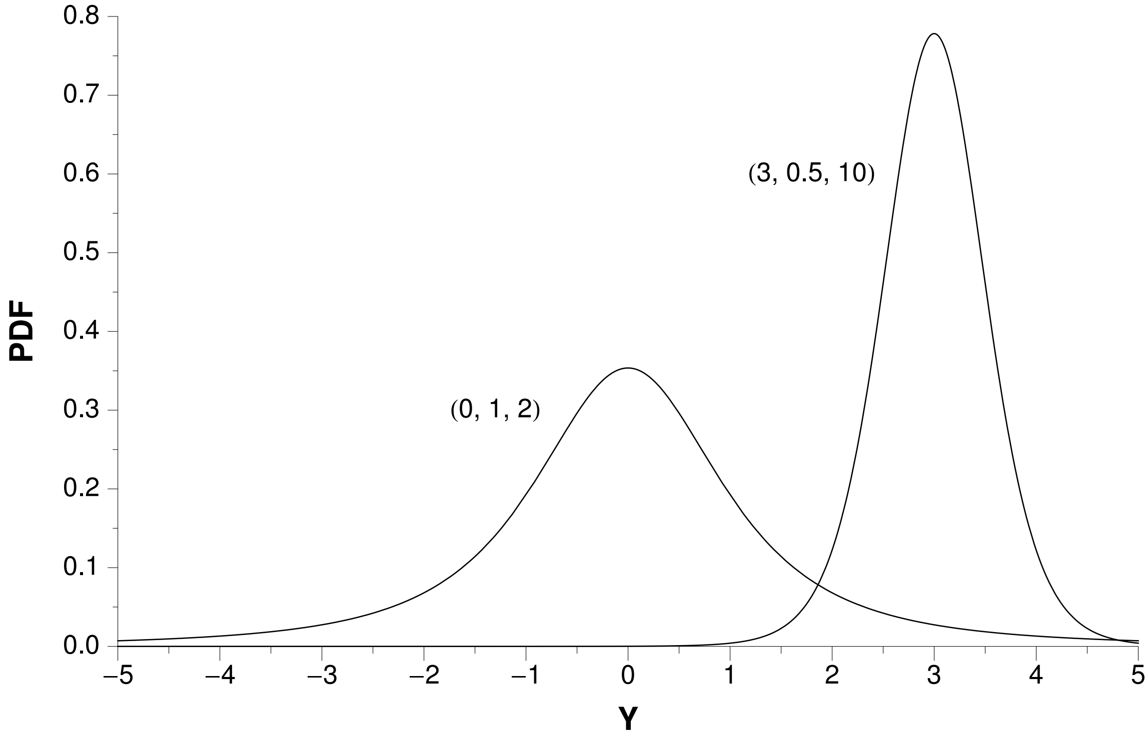
$$\lim_{N\to\infty} Prob\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{Z_i-\mu}{\sigma}\right) \le z\right] = \Phi(z)$$

   known as the Central Limit Theorem.

2. If $X \sim Normal(A, B)$, then $Y = aX + b \sim Normal(a\,A + b, a\,B)$
3. If $X \sim Normal(A, B)$ and $Y \sim Normal(C, D)$, then $S = X + Y$ (i.e., the *convolution* of X and Y) $\sim Normal(A + C, \sqrt{B^2 + D^2})$.
4. Errors of real-valued observations are often ∼Normal or ∼**Laplace**.

StudentsT(A,B,C)

$\mathrm{B}, \mathrm{C} > 0$



$$\mathrm{PDF} = \frac{1}{B\sqrt{C}\,Beta\left(\frac{C}{2}, \frac{1}{2}\right)}\left[1 + \frac{(y-A)^2}{B^2\,C}\right]^{-(C+1)/2}$$

$$\mathrm{CDF} = \begin{cases} \dfrac{\Gamma\left((C+1)/2\right)}{2\sqrt{\pi}\,\Gamma\left(C/2\right)} Beta\left(\dfrac{C}{C+z^2}; \dfrac{C}{2}, \dfrac{1}{2}\right), & z \leq 0 \\ \dfrac{1}{2} + \dfrac{\Gamma\left((C+1)/2\right)}{2\sqrt{\pi}\,\Gamma\left(C/2\right)} Beta\left(\dfrac{z^2}{C+z^2}; \dfrac{1}{2}, \dfrac{C}{2}\right), & z > 0 \end{cases} \quad \text{where } z = \frac{y-A}{B}$$

$$\mathrm{CF} = \frac{2^{1-C/2}\,C^{C/4}\,(B\,|t|)^{C/2}\exp(iAt)}{\Gamma(C/2)}\,K_{C/2}\left(B\,|t|\,\sqrt{C}\right)$$

Parameters – A (μ): Location, B: Scale, C: Shape (degrees of freedom)

Moments, etc.

$$\text{Mean} = \text{Median} = \text{Mode} = A$$

$$\text{Variance} = \frac{C}{C-2}B^2$$

$$\text{Skewness} = 0$$

$$\text{Kurtosis} = \frac{3\,C^2}{8 - 6\,C + C^2}B^4$$

$$\mathrm{Q1},\ \mathrm{Q3} : \textit{no simple closed form}$$

Notes

1. Moment r exists iff $C > r$.

Aliases and Special Cases

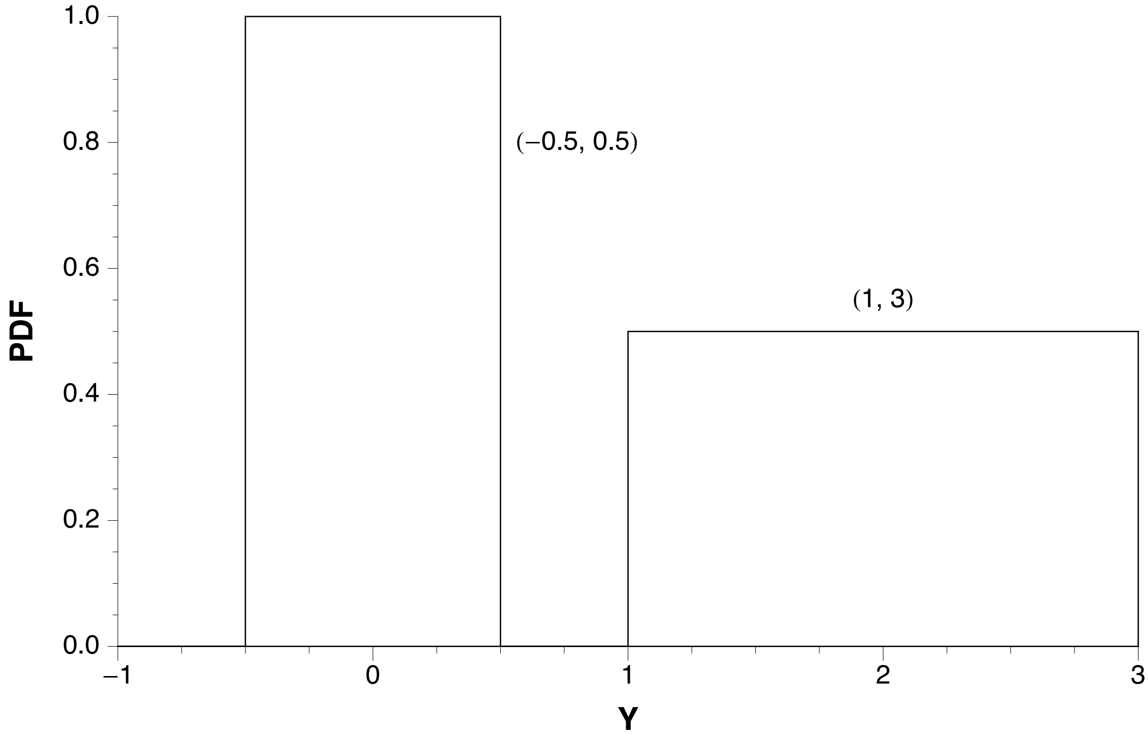
1. The StudentsT distribution if often called simply the t distribution.
2. The StudentsT distribution approaches a **Normal** distribution in the limit $C \to \infty$.

Characterizations

1. The StudentsT distribution is used to characterize small samples (typically, N $<$ 30) from a **Normal** population.
2. The StudentsT distribution is equivalent to a *parameter-mix* distribution in which a **Normal** distribution has a variance modeled as $\sim$**InverseGamma**. This is symbolized

$$\mathbf{Normal}(\mu,\sigma) \bigwedge_{\sigma^2} \mathbf{InverseGamma}(A,B)$$

Uniform(A,B) $\quad\quad A \leq y \leq B$



$$\text{PDF} = \frac{1}{B - A}$$

$$\text{CDF} = \frac{y - A}{B - A}$$

$$\text{CF} = \frac{i\,(\exp(iBt) - \exp(iAt))}{(A - B)\,t}$$

Parameters – A: Location, B: Scale (upper bound)

Moments, etc.

$$\text{Mean} = \text{Median} = \frac{A + B}{2}$$

$$\text{Variance} = \frac{(B - A)^2}{12}$$

$$\text{Skewness} = 0$$

$$\text{Kurtosis} = \frac{(B - A)^4}{80}$$

$$\text{Mode} = \textit{undefined}$$

$$\mathrm{Q1} = \frac{3\,A + B}{4} \quad \mathrm{Q3} = \frac{A + 3\,B}{4}$$

$$\mathrm{RandVar} = A + u\,(B - A)$$

Notes

1. The parameters of the Uniform distribution are sometimes given as the mean and half-width of the domain.

Aliases and Special Cases

1. The Uniform distribution is often called the *Rectamgular* distribution.

Characterizations

1. The Uniform distribution is often used to describe total ignorance within a bounded interval.

2. Pseudo-random number generators typically return X $\sim$Uniform(0, 1) as their output.

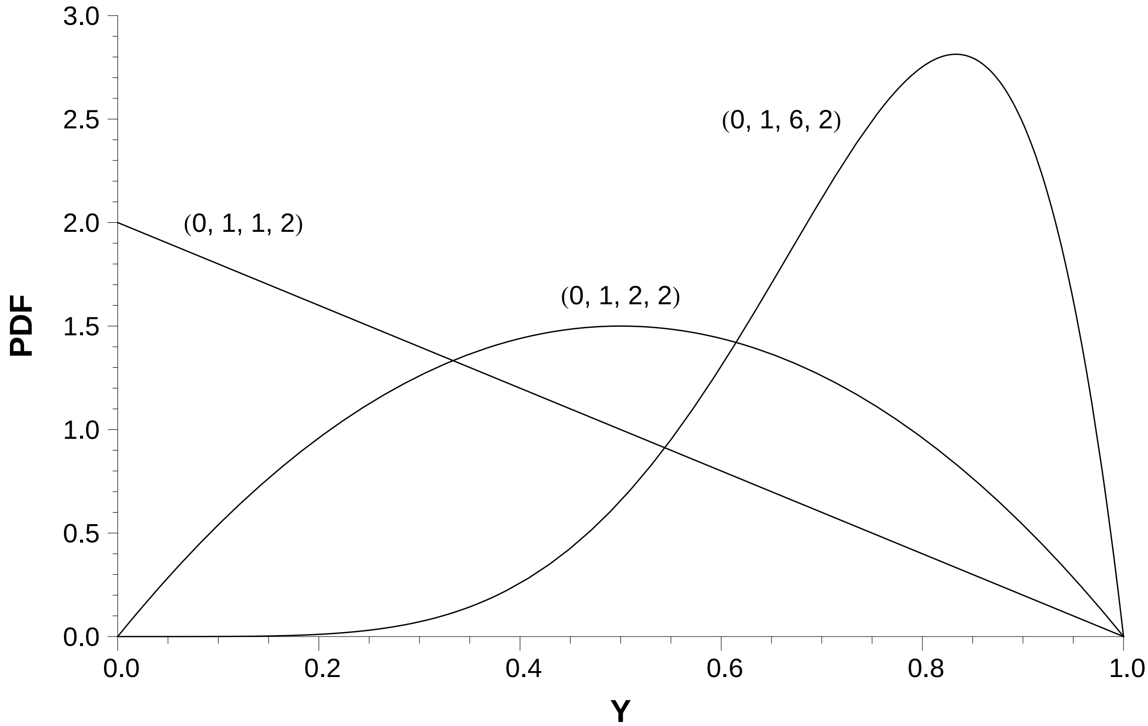
Part II

Continuous: Skewed

These distributions are skewed, most often to the right. In some cases, the direction of skewness is controlled by the value of a parameter. To get a model that is skewed to the left, it is often sufficient to model the negative of the random variate.

Beta(A,B,C,D)

$\mathrm{A} < \mathrm{y} < \mathrm{B}, \quad \mathrm{C}, \mathrm{D} > 0$



$$\mathrm{PDF} = \frac{(B-A)^{1-C-D}}{Beta(C,D)}(y-A)^{C-1}(B-y)^{D-1}$$

$$\mathrm{CDF} = Beta\left(\frac{A-y}{A-B}; C, D\right) / Beta(C,D)$$

$$\mathrm{CF} = \exp(iAt)\,{}_1F_1(C; C+D, (B-A)it)$$

Parameters – A: Location, B: Scale (upper bound), C, D: Shape

Moments, etc.

$$\mathrm{Mean} = \frac{AD+BC}{C+D}$$

$$\mathrm{Variance} = \frac{CD(B-A)^2}{(C+D+1)(C+D)^2}$$

$$\mathrm{Skewness} = \frac{2CD(C-D)(B-A)^3}{(C+D+2)(C+D+1)(C+D)^3}$$

$$\mathrm{Kurtosis} = \frac{3CD[CD(D-2)+2D^2+(D+2)C^2](B-A)^4}{(C+D+3)(C+D+2)(C+D+1)(C+D)^4}$$

$$\mathrm{Mode} = \frac{A(D-1)+B(C-1)}{C+D-2}$$

$\mathrm{Median},\ \mathrm{Q1},\ \mathrm{Q3}$: *no simple closed form*

Notes

1. The Beta distribution is so flexible that it is often used to mimic other distributions provided suitable bounds can be found.

2. If both C and D are large, this distribution is roughly symmetric and some other model is indicated.

Aliases and Special Cases

1. The standard Beta distribution is Beta(0, 1, C, D).

2. Beta(0, 1, C, 1) is often called the *Power-function* distribution.

3. Beta(0, B, C, D) is often called the *Scaled Beta* distribution (with scale = B).

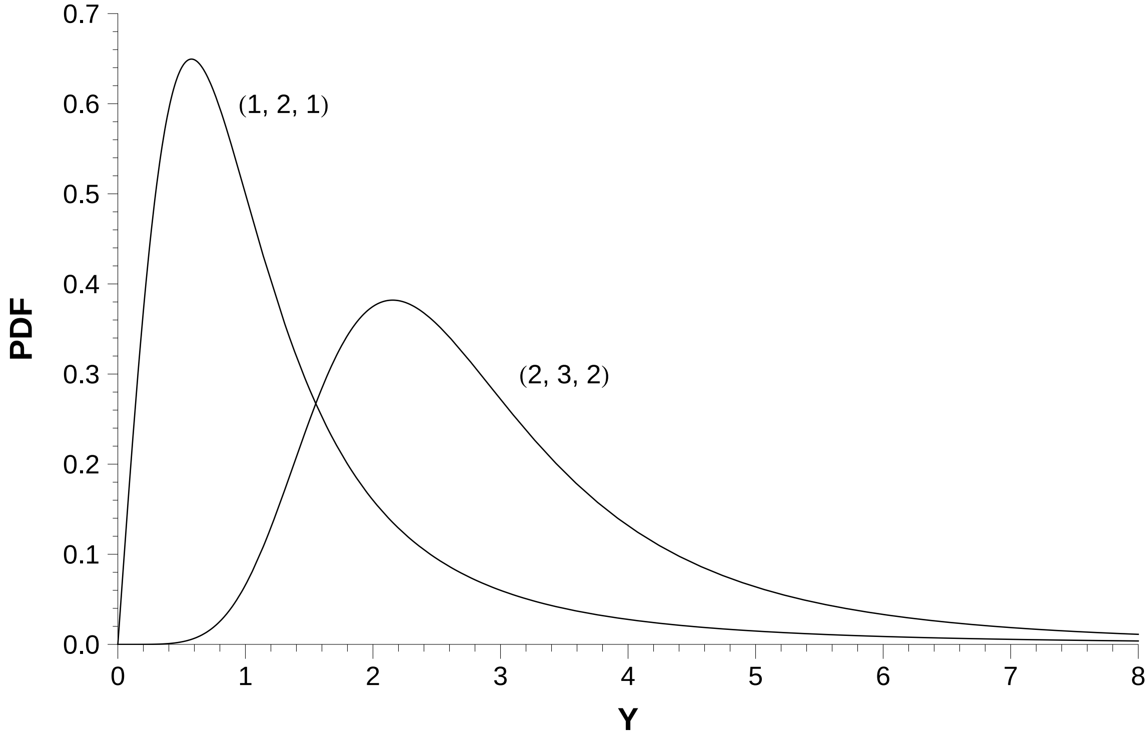
Characterizations

1. If X_j^2, j = 1, 2 $\sim$ standard *Chi-square* with ν_j degrees of freedom, respectively, then $Z = (X_1^2)/(X_1^2 + X_2^2) \sim \mathrm{Beta}(0, 1, \nu_1/2, \nu_2/2)$.

2. More generally, $Z = W_1/(W_1 + W_2) \sim \mathrm{Beta}(0, 1, p_1, p_2)$ if $W_j \sim$ **Gamma**(0, σ, p_j), for any scale, σ.

3. If $Z_1, Z_2, \ldots, Z_n \sim$ **Uniform**(0, 1) are sorted to give the corresponding order statistics $Z_{(1)}, Z_{(2)}, \ldots, Z_{(n)}$, then the r^{th} order statistic, $Z_{(r)} \sim \mathrm{Beta}(0, 1, r, n - r + 1)$.

Burr3(A,B,C) $\quad\quad\quad\quad\quad\quad\quad\quad \mathrm{y}>0, \quad \mathrm{A}, \mathrm{B}, \mathrm{C}>0$



$$\mathrm{PDF}=\frac{BC}{A}\left(\frac{y}{A}\right)^{-B-1}\left[1+\left(\frac{y}{A}\right)^{-B}\right]^{-C-1}$$

$$\mathrm{CDF}=\left[1+\left(\frac{y}{A}\right)^{-B}\right]^{-C}$$

$$\mathrm{CF}=\textit{no simple closed form}$$

Parameters – A: Scale, B, C: Shape

Moments, etc. $[\mu_r \equiv A^r C\,Beta((BC+r)/B,(B-r)/B)]$

$$\mathrm{Mean}=\mu_1$$

$$\mathrm{Variance}=-\mu_1^2+\mu_2$$

$$\mathrm{Skewness}=2\mu_1^3-3\mu_1\mu 2+\mu_3$$

$$\mathrm{Kurtosis}=-3\mu_1^4+6\mu_1^2\mu_2-4\mu_1\mu_3+\mu_4$$

$$\mathrm{Mode}=A\left(\frac{BC-1}{B+1}\right)^{1/B}$$

$$\mathrm{Median}=A\left(2^{1/C}-1\right)^{-1/B}$$

$$\mathrm{Q1}=A\left(4^{1/C}-1\right)^{-1/B} \quad \mathrm{Q3}=A\left((4/3)^{1/C}-1\right)^{-1/B}$$

$$\mathrm{RandVar}=A\left(u^{-1/C}-1\right)^{-1/B}$$

Notes

1. Moment r exists iff $B > r$.

Aliases and Special Cases

1. The Burr III distribution is also called the *Dagum* distribution or the *inverse Burr* distribution.

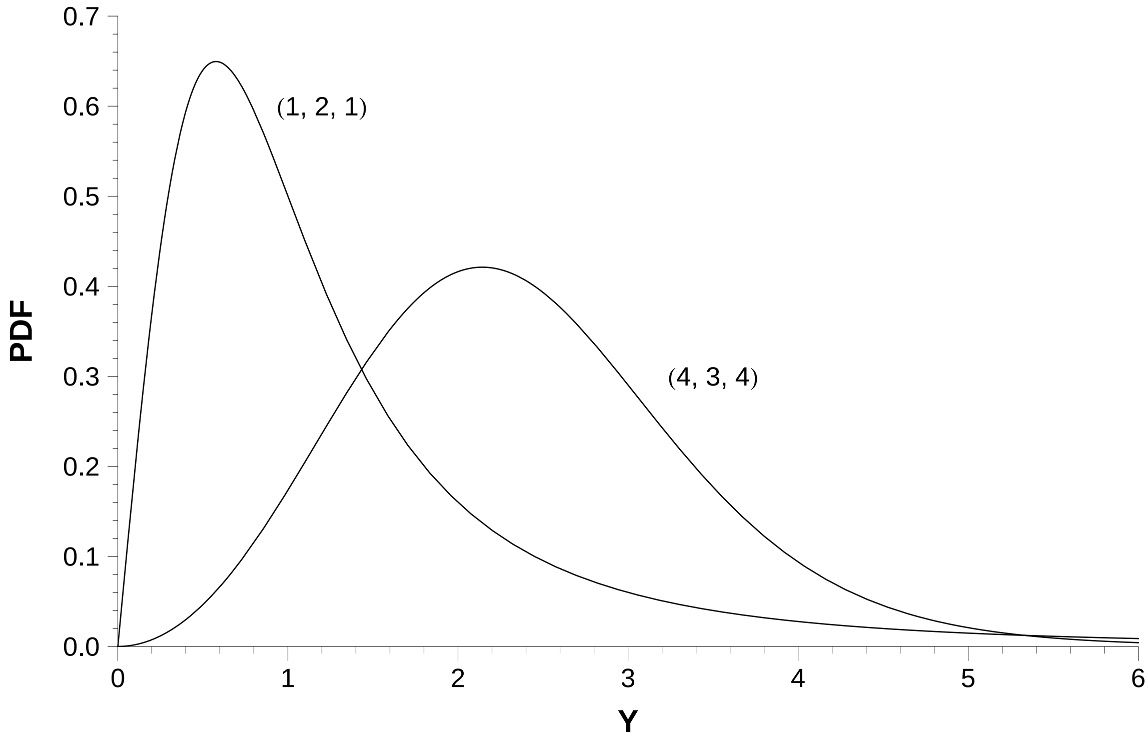
Characterizations

1. If $Y \sim$ **Burr XII**, then $1/Y \sim$ Burr III.

2. Burr distributions are often used when there is a need for a flexible, left-bounded, continuous model.

Burr12(A,B,C)

$y > 0, \quad A, B, C > 0$



$$\text{PDF} = \frac{BC}{A}\left(\frac{y}{A}\right)^{B-1}\left[1+\left(\frac{y}{A}\right)^{B}\right]^{-C-1}$$

$$\text{CDF} = 1-\left[1+\left(\frac{y}{A}\right)^{B}\right]^{-C}$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A: Scale, B, C: Shape

Moments, etc. $[\mu_r \equiv A^r C Beta((BC-r)/B, (B+r)/B)]$

$$\text{Mean} = \mu_1$$
$$\text{Variance} = -\mu_1^2 + \mu_2$$
$$\text{Skewness} = 2\mu_1^3 - 3\mu_1\mu 2 + \mu_3$$
$$\text{Kurtosis} = -3\mu_1^4 + 6\mu_1^2\mu_2 - 4\mu_1\mu_3 + \mu_4$$

$$\text{Mode} = A\left(\frac{B-1}{BC+1}\right)^{1/B}$$

$$\text{Median} = A\left(2^{1/C}-1\right)^{1/B}$$

$$\text{Q1} = A\left((4/3)^{1/C}-1\right)^{1/B} \quad \text{Q3} = A\left(4^{1/C}-1\right)^{1/B}$$

$$\text{RandVar} = A\left(u^{-1/C}-1\right)^{1/B}$$

Notes

1. Moment r exists iff $B\,C > r$.

Aliases and Special Cases

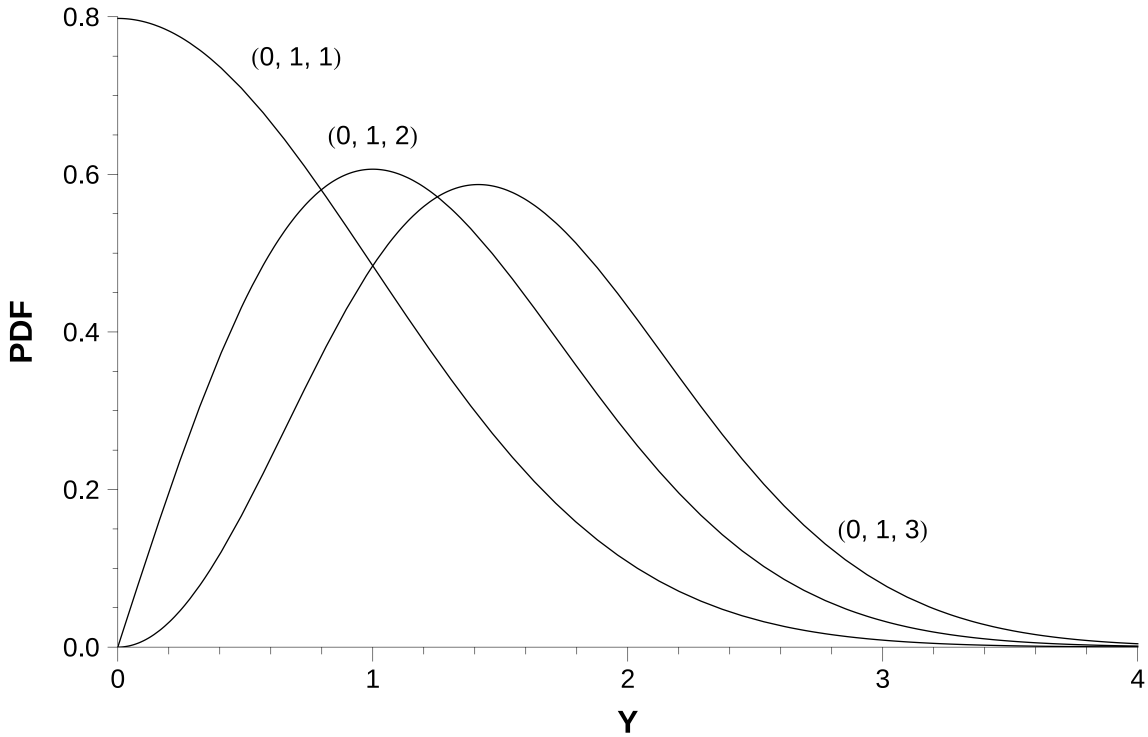
1. The Burr XII distribution is also called the *Singh-Maddala* distribution or the *generalized log-logistic* distribution.
2. When $B = 1$, the Burr XII distribution becomes the **Pareto II** distribution.
3. When $C = 1$, the Burr XII distribution becomes a special case of the **Fisk** distribution.

Characterizations

1. This distribution is often used to model incomes and other quantities in econometrics.
2. Burr distributions are often used when there is a need for a flexible, left-bounded, continuous model.

Chi(A,B,C)

$\mathrm{y} > \mathrm{A}, \quad \mathrm{B}, \mathrm{C} > 0$



$$\mathrm{PDF} = \frac{2^{1-C/2}}{\Gamma(C/2)} \left(\frac{y-A}{B}\right)^{C-1} \exp\left(-\frac{(y-A)^2}{2\,B^2}\right)$$

$$\mathrm{CDF} = 1 - \frac{\Gamma\left(\frac{C}{2}, \frac{(y-A)^2}{2\,B^2}\right)}{\Gamma(C/2)}$$

$$\mathrm{CF} = \exp(iAt)\Big[\,{}_1F_1\left(C/2, 1/2, -(B^2t^2)/2\right) + \frac{i\sqrt{2}B\,t\,\Gamma((C+1)/2)}{\Gamma(C/2)}\,{}_1F_1\left((C+1)/2, 3/2, -(B^2t^2)/2\right)\Big]$$

Parameters – A: Location, B: Scale, C (ν): Shape (also, degrees of freedom)

Moments, etc. $\left[\gamma \equiv \Gamma\left(\frac{c+1}{2}\right), \delta \equiv \Gamma\left(\frac{c}{2}\right)\right]$

$$\mathrm{Mean} = A + \sqrt{2}B\frac{\gamma}{\delta}$$

$$\mathrm{Variance} = B^2\left[C - 2\frac{\gamma^2}{\delta^2}\right]$$

$$\mathrm{Skewness} = \frac{\sqrt{2}B^3\gamma\left[(1-2C)\delta^2 + 4\gamma^2\right]}{\delta^3}$$

$$\mathrm{Kurtosis} = \frac{B^4\left[C(C+2)\delta^4 + 4(C-2)\gamma^2\delta^2 - 12\gamma^4\right]}{\delta^4}$$

$$\text{Mode} = A + B\sqrt{C-1}$$

$$\text{Median, Q1, Q3}: \textit{no simple closed form}$$

Notes

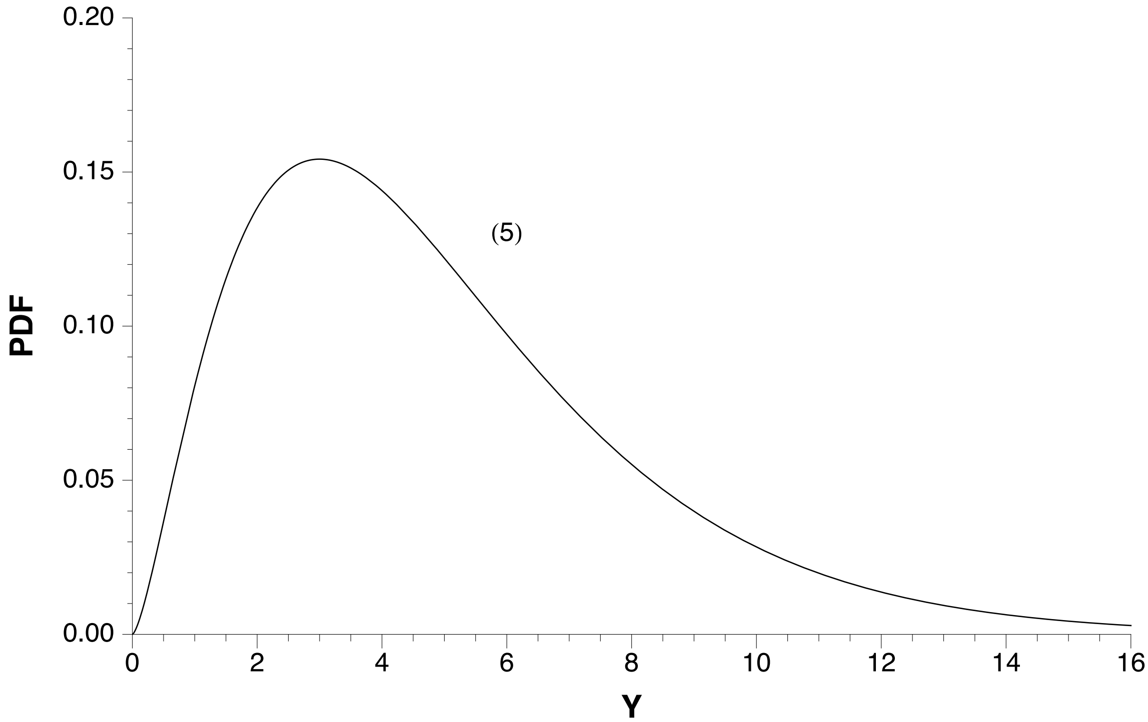
Aliases and Special Cases

1. Chi(A, B, 1) is the **HalfNormal** distribution.
2. Chi(0, B, 2) is the *Rayleigh* distribution.
3. Chi(A, B, 3) is the *Maxwell* distribution.

Characterizations

1. If Y ∼*Chi-square*, its positive square root is ∼Chi.
2. If X, Y ∼**Normal**(0, B), the distance from the origin to the point (X, Y) is ∼*Rayleigh*.
3. In a spatial pattern generated by a Poisson process, the distance between any pattern element and its nearest neighbor is ∼*Rayleigh*.
4. The speed of a random molecule, at any temperature, is ∼*Maxwell*.

ChiSquare(A) $y > 0, \quad A > 0$



$$\text{PDF} = \frac{y^{A/2-1}}{2^{A/2}\exp(y/2)\,\Gamma(A/2)}$$

$$\text{CDF} = 1 - \frac{\Gamma(A/2, y/2)}{\Gamma(A/2)}$$

$$\text{CF} = (1 - 2it)^{-A/2}$$

Parameters – A (ν): Shape (degrees of freedom)

Moments, etc.

$$\text{Mean} = A$$

$$\text{Variance} = 2\,A$$

$$\text{Skewness} = 8\,A$$

$$\text{Kurtosis} = 12\,A(A+4)$$

$$\text{Mode} = A - 2 \quad ; \text{A} \geq 2, \text{else } 0$$

$$\text{Median, Q1, Q3} : \textit{no simple closed form}$$

Notes

1. In general, parameter A need not be an integer.

Aliases and Special Cases

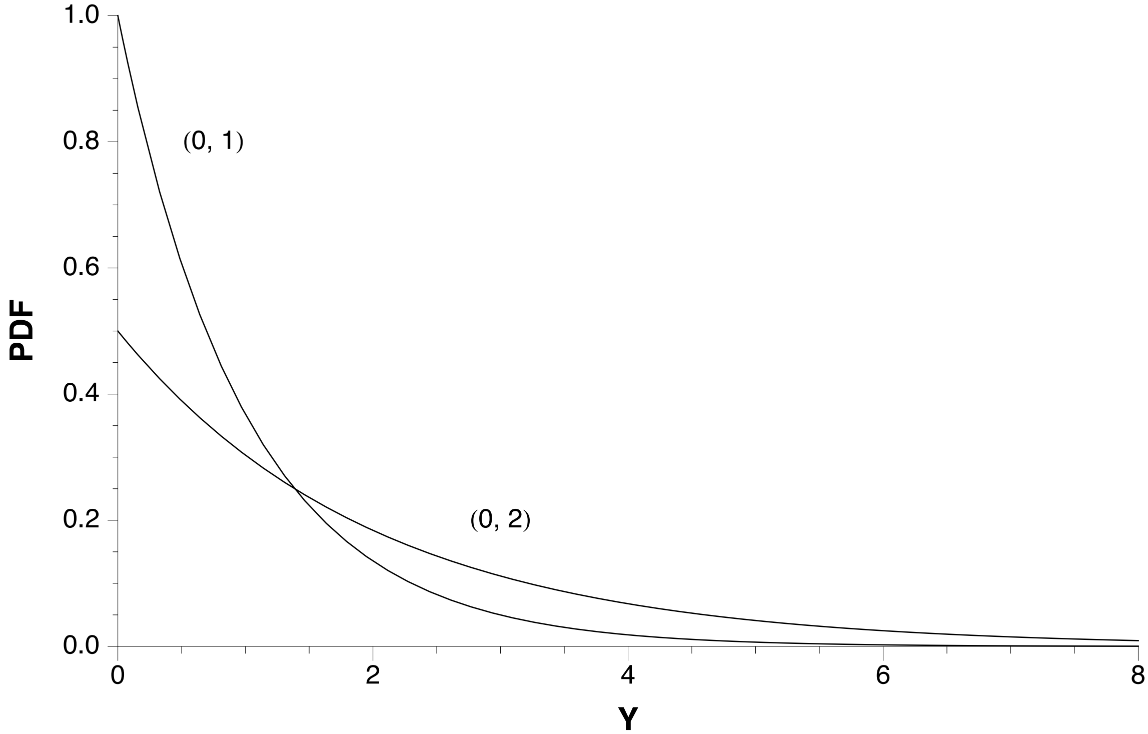
1. The Chi Square distribution is just a special case of the **Gamma** distribution and is included here purely for convenience.

Characterizations

1. If $Z_1, Z_2, \ldots, Z_A$ ~**Normal**(0, 1), then $W = \sum_{k=1}^{A} Z^2$ ~ChiSquare(A).

Exponential(A,B)

$y \geq A, \quad B > 0$



$$\text{PDF} = \frac{1}{B}\exp\left(\frac{A-y}{B}\right)$$

$$\text{CDF} = 1 - \exp\left(\frac{A-y}{B}\right)$$

$$\text{CF} = \frac{\exp(iAt)}{1-iBt}$$

Parameters – A: Location, B: Scale

Moments, etc.

$$\text{Mean} = A + B$$

$$\text{Variance} = B^2$$

$$\text{Skewness} = 2B^3$$

$$\text{Kurtosis} = 9B^4$$

$$\text{Mode} = A$$

$$\text{Median} = A + B\log(2)$$

$$\text{Q1} = A + B\log\left(\frac{4}{3}\right) \qquad \text{Q3} = A + B\log(4)$$

$$\text{RandVar} = A - B\log(u)$$

Notes

1. The one-parameter version of this distribution, Exponential(0, B), is far more common than the general parametrization shown here.

Aliases and Special Cases

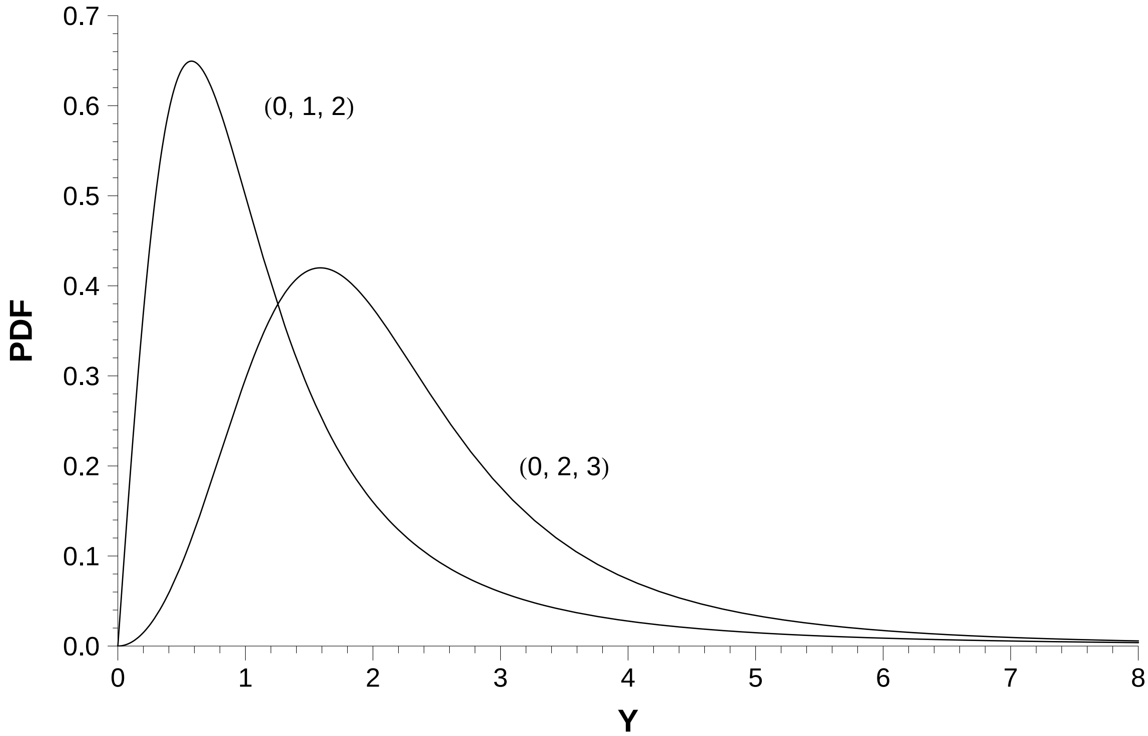
1. The Exponential distribution is sometimes called the *Negative exponential* distribution.

2. The discrete analogue of the Exponential distribution is the **Geometric** distribution.

Characterizations

1. If the future lifetime of a system at any time, t, has the same distribution for all t, then this distribution is the Exponential distribution. This behavior is known as the memoryless property.

Fisk(A,B,C) $\qquad\qquad y > A, \quad B, C > 0$



$$\text{PDF} = \frac{C}{B}\left(\frac{y-A}{B}\right)^{C-1}\left[1+\left(\frac{y-A}{B}\right)^{C}\right]^{-2}$$

$$\text{CDF} = \frac{1}{1+\left(\frac{y-A}{B}\right)^{-C}}$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A: Location, B: Scale, C: Shape

Moments, etc. (see Note #4)

$$\text{Mean} = A + \frac{\pi B}{C}\csc\left(\frac{\pi}{C}\right)$$

$$\text{Variance} = \frac{2\pi B^2}{C}\csc\left(\frac{2\pi}{C}\right) - \frac{\pi^2 B^2}{C^2}\csc\left(\frac{\pi}{C}\right)^2$$

$$\text{Skewness} = \frac{2\pi^3 B^3}{C^3}\csc\left(\frac{\pi}{C}\right)^3 - \frac{6\pi^2 B^3}{C^2}\csc\left(\frac{\pi}{C}\right)\csc\left(\frac{2\pi}{C}\right) + \frac{3\pi B^3}{C}\csc\left(\frac{3\pi}{C}\right)$$

$$\text{Kurtosis} = -\frac{3\pi^4 B^4}{C^4}\csc\left(\frac{\pi}{C}\right)^4 + \frac{12\pi^3 B^4}{C^3}\csc\left(\frac{\pi}{C}\right)^2\csc\left(\frac{2\pi}{C}\right) - \frac{12\pi^2 B^4}{C^2}\csc\left(\frac{\pi}{C}\right)\csc\left(\frac{3\pi}{C}\right) + \frac{4\pi B^4}{C}\csc\left(\frac{4\pi}{C}\right)$$

$$\text{Mode} = A + B\,\sqrt[C]{\frac{C-1}{C+1}}$$

$$\text{Median} = A + B$$

$$\text{Q1} = A + \frac{B}{\sqrt[C]{3}} \quad \text{Q3} = A + B\,\sqrt[C]{3}$$

$$\text{RandVar} = A + B\,\sqrt[C]{\frac{u}{1-u}}$$

Notes

1. The Fisk distribution is right-skewed.
2. To model a left-skewed distribution, try modeling $w = -y$.
3. The Fisk distribution is related to the **Logistic** distribution in the same way that a **LogNormal** is related to the **Normal** distribution.
4. Moment r exists iff $C > r$.

Aliases and Special Cases

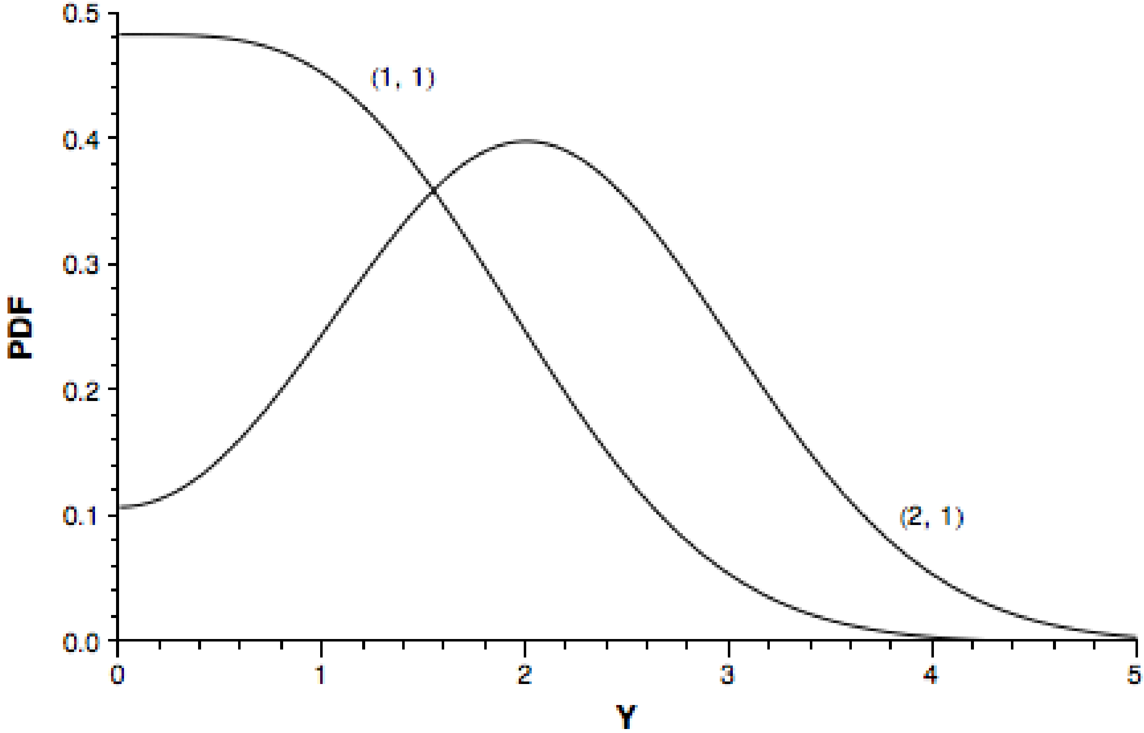
1. The Fisk distribution is also known as the *LogLogistic* distribution.

Characterizations

1. The Fisk distribution is often used in income and lifetime analysis.

FoldedNormal(A,B)

$y \geq A \geq 0, \quad B > 0$



$$\text{PDF} = \frac{1}{B\sqrt{2\pi}}\left[\exp\left(-\frac{(y-A)^2}{2B^2}\right) + \exp\left(-\frac{(y+A)^2}{2B^2}\right)\right]$$

$$\text{CDF} = \Phi\left(\frac{A+y}{B}\right) - \Phi\left(\frac{A-y}{B}\right)$$

$$\text{CF} = \exp\left(\frac{-2iAt - B^2t^2}{2}\right)\left[1 - \Phi\left(\frac{A - iB^2t}{B}\right) + \exp(2iAt)\,\Phi\left(\frac{A+iB^2t}{B}\right)\right]$$

Parameters – A (μ): Location, B (σ): Scale, both for the corresponding unfolded **Normal**

Moments, etc.

$$\text{Mean} = B\sqrt{\frac{2}{\pi}}\exp\left(-\frac{A^2}{2B^2}\right) - A\left[2\,\Phi\left(\frac{A}{B}\right) - 1\right]$$

$$\text{Variance} = A^2 + B^2 - \text{Mean}^2$$

$$\text{Skewness} = 2\,\text{Mean}^3 - 3\left(A^2+B^2\right)\,\text{Mean} + B\left(A^2+2B^2\right)\sqrt{\frac{2}{\pi}}\exp\left(-\frac{A^2}{2B^2}\right) + \left(A^3 + 3AB^2\right)\left[2\,\Phi\left(\frac{A}{B}\right) - 1\right]$$

$$\text{Kurtosis} = A^4 + 6A^2B^2 + 3B^4 + 6\left(A^2+B^2\right)\,\text{Mean}^2 - 3\,\text{Mean}^4 - 4\,\text{Mean}\left[B\left(A^2+2B^2\right)\sqrt{\frac{2}{\pi}}\exp\left(-\frac{A^2}{2B^2}\right) + \left(A^3+3AB^2\right)\left[2\,\Phi\left(\frac{A}{B}\right) - 1\right]\right]$$

$$\text{Mode, Median, Q1, Q3}: \textit{no simple closed form}$$

Notes

1. This distribution is indifferent to the sign of A so, to avoid ambiguity, *Regress+* restricts A to be positive.

2. Mode > 0 when $A > B$.

Aliases and Special Cases

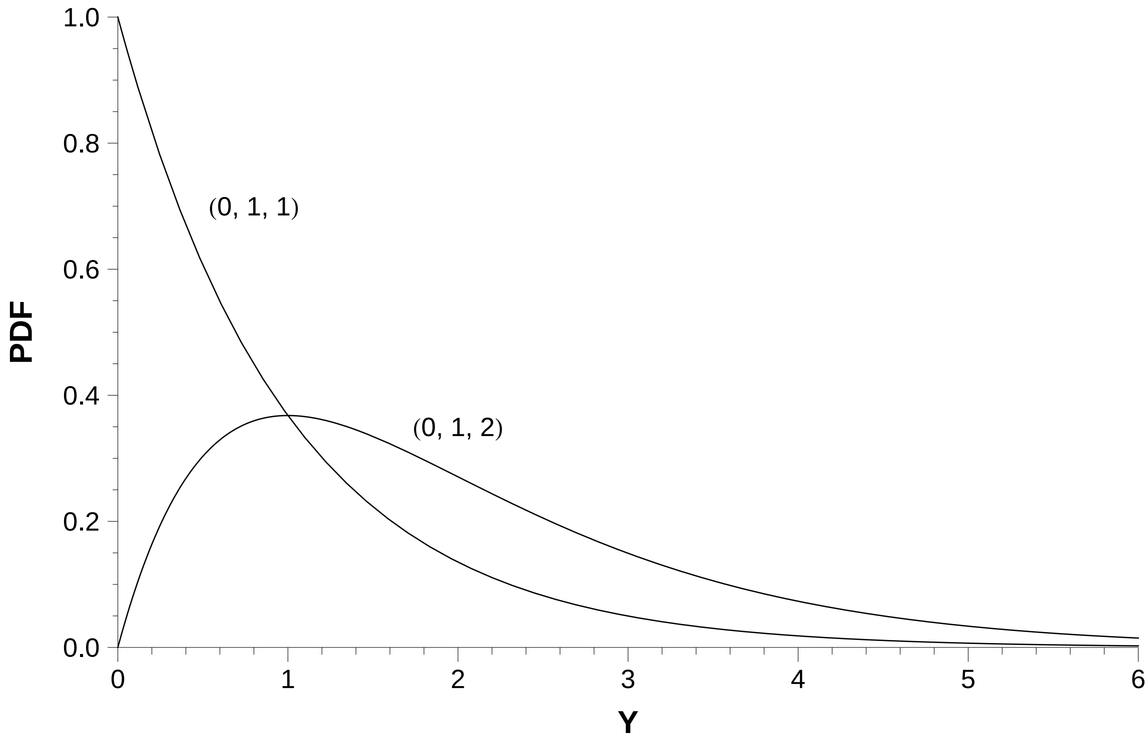
1. If $A = 0$, the FoldedNormal distribution reduces to the **HalfNormal** distribution.

2. The FoldedNormal distribution is identical to the distribution of χ' (*Non-central Chi*) with one degree of freedom and non-centrality parameter $(A/B)^2$.

Characterizations

1. If Z ~**Normal(A, B)**, $|Z|$ ~FoldedNormal(A, B).

Gamma(A,B,C) $\quad\quad y > A, \quad B, C > 0$



$$\text{PDF} = \frac{1}{B\,\Gamma(C)}\left(\frac{y-A}{B}\right)^{C-1}\exp\left(-\frac{y-A}{B}\right)$$

$$\text{CDF} = 1 - \frac{\Gamma\left(C, \frac{y-A}{B}\right)}{\Gamma(C)}$$

$$\text{CF} = \exp(iAt)(1 - iBt)^{-C}$$

Parameters – A: Location, B (β): Scale, C (α): Shape

Moments, etc.

$$\text{Mean} = A + BC$$

$$\text{Variance} = B^2 C$$

$$\text{Skewness} = 2B^3 C$$

$$\text{Kurtosis} = 3B^4 C(C+2)$$

$$\text{Mode} = A + B(C-1) \quad ; \text{C} \geq 1, \text{else } 0$$

Median, Q1, Q3 : *no simple closed form*

Notes

1. The Gamma distribution is right-skewed.

2. To model a left-skewed distribution, try modeling $w = -y$.

Aliases and Special Cases

1. Gamma(0, B, C), where C is an integer > 0, is the *Erlang*distribution.

2. Gamma(A, B, 1) is the **Exponental** distribution.

3. Gamma(0, $2, \nu/2$) is the *Chi-square* distribution with ν degrees of freedom.

4. The Gamma distribution approaches a **Normal** distribution in the limit $C \to \infty$.

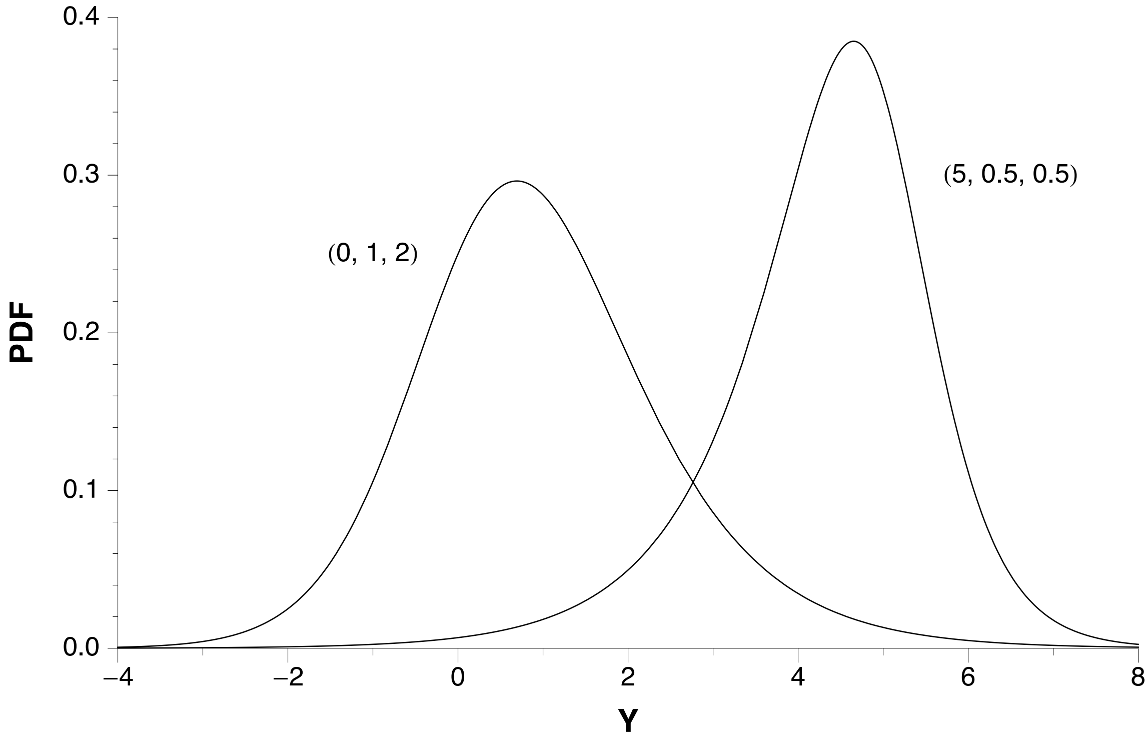
Characterizations

1. If $Z_1, Z_2, \ldots, Z_\nu$ ~**Normal**(0, 1), then $W = \sum_{k=1}^{\nu} Z_k^2$ ~Gamma($0, 2, \nu/2$).

2. If $Z_1, Z_2, \ldots, Z_n$ ~**Exponential**(A, B), then $W = \sum_{k=1}^{n} Z_k$ ~Erlang(A, B, n).

3. If Z_1 ~Gamma(A, B, C_1) and Z_2 ~Gamma(A, B, C_2), then $(Z_1 + Z_2)$ ~Gamma(A, B, $\mathrm{C}_1 + \mathrm{C}_2$).

GenLogistic(A,B,C) $\quad B, C > 0$



$$\text{PDF} = \frac{C}{B}\exp\left(-\frac{y-A}{B}\right)\left[1+\exp\left(-\frac{y-A}{B}\right)\right]^{-C-1}$$

$$\text{CDF} = \left[1+\exp\left(-\frac{y-A}{B}\right)\right]^{-C}$$

$$\text{CF} = \frac{\exp(iAt)}{\Gamma(C)}\,\Gamma(1-iBt)\,\Gamma(C+iBt)$$

Parameters – A: Location, B: Scale, C: Shape

Moments, etc.

$$\text{Mean} = A + [\gamma + \psi(C)]\,B$$

$$\text{Variance} = \left[\frac{\pi^2}{6} + \psi'(C)\right] B^2$$

$$\text{Skewness} = [\psi''(C) - \psi''(1)]\,B^3$$

$$\text{Kurtosis} = \left[\psi'''(C) + \psi'(C)\left[\pi^2 + 3\psi'(C)\right] + \frac{3\pi^4}{20}\right] B^4$$

$$\text{Mode} = A + B\log(C)$$

$$\text{Median} = A - B\log\left(\sqrt[C]{2} - 1\right)$$

$$\text{Q1} = A - B\log\left(\sqrt[C]{4} - 1\right) \qquad \text{Q3} = A - B\log\left(\sqrt[C]{4/3} - 1\right)$$

$$\text{RandVar} = A - B\log\left(\frac{1}{\sqrt[C]{u}} - 1\right)$$

Notes

1. This distribution is Type I of several generalizations of the **Logistic** distribution.
2. It is left-skewed when $C < 1$ and right-skewed when $C > 1$.

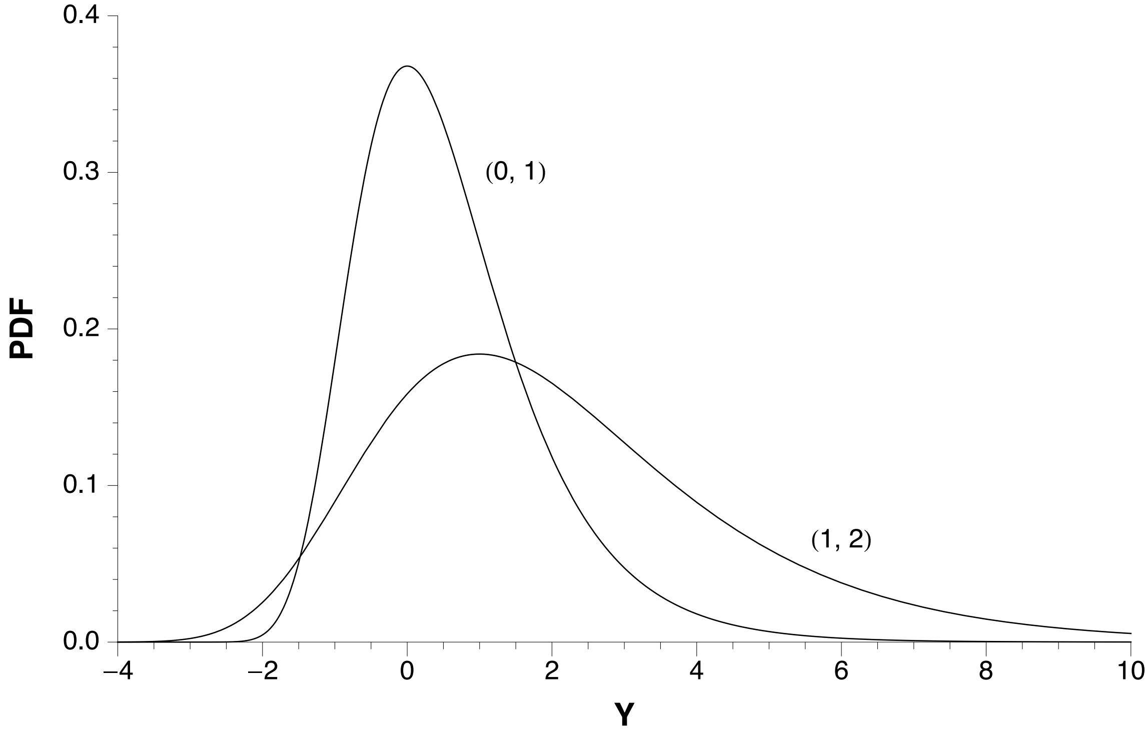
Aliases and Special Cases

1. This distribution is also known as the *skew-logistic* distribution.

Characterizations

1. This distribution has been used in the analysis of extreme values.

Gumbel(A,B) $\qquad B > 0$



$$\text{PDF} = \frac{1}{B} \exp\left(\frac{A-y}{B}\right) \exp\left[-\exp\left(\frac{A-y}{B}\right)\right]$$

$$\text{CDF} = \exp\left[-\exp\left(\frac{A-y}{B}\right)\right]$$

$$\text{CF} = \exp(iAt)\,\Gamma(1-iBt)$$

Parameters – A: Location, B: Scale

Moments, etc.

$$\text{Mean} = A + B\gamma$$

$$\text{Variance} = \frac{(\pi B)^2}{6}$$

$$\text{Skewness} = 2\,\zeta(3)B^3$$

$$\text{Kurtosis} = \frac{3(\pi B)^4}{20}$$

$$\text{Mode} = A$$

$$\text{Median} = A - B\log(\log(2))$$

$$\text{Q1} = A - B\log(\log(4)) \quad \text{Q3} = A - B\log(\log(4/3))$$

$$\text{RandVar} = A - B\log(-\log(u))$$

Notes

1. The Gumbel distribution is one of the class of extreme-value distributions.
2. It is right-skewed.
3. To model the analogous left-skewed distribution, try modeling $w = -y$.

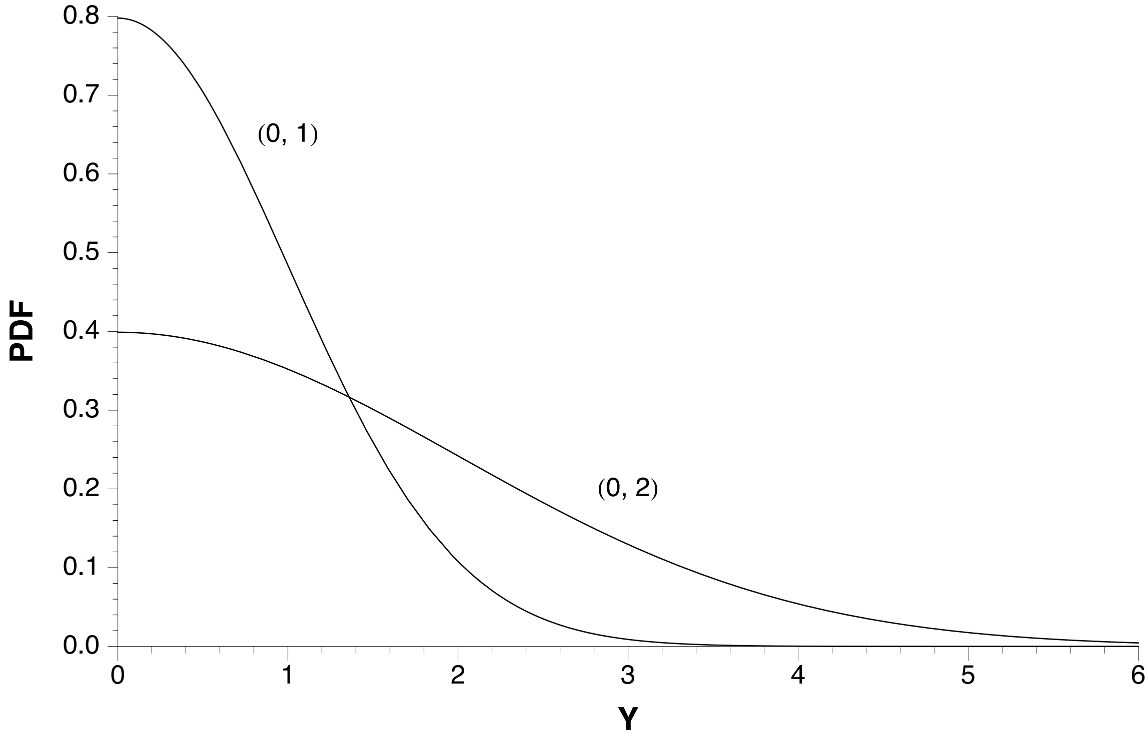
Aliases and Special Cases

1. The Gumbel distribution is sometimes called the *LogWeibull* distribution.
2. It is also known as the *Gompertz* distribution.
3. It is also known as the *Fisher-Tippett* distribution.

Characterizations

1. Extreme-value distributions are the limiting distributions, as N $\rightarrow \infty$, of the greatest value among N i.i.d. variates selected from a continuous distribution. By replacing y with $-y$, the smallest values may be modeled.
2. The Gumbel distribution is often used to model maxima from samples (of the same size) in which the variate is unbounded.

HalfNormal(A,B) $y \geq A, \quad B > 0$



$$\text{PDF} = \frac{1}{B}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{(y-A)^2}{2\,B^2}\right)$$

$$\text{CDF} = 2\,\Phi\left(\frac{y-A}{B}\right) - 1$$

$$\text{CF} = \exp\left(iAt - \frac{B^2t^2}{2}\right)\left[1 + erf\left(\frac{iBt}{\sqrt{2}}\right)\right]$$

Parameters – A: Location, B: Scale

Moments, etc.

$$\text{Mean} = A + B\sqrt{\frac{2}{\pi}}$$

$$\text{Variance} = \left(1 - \frac{2}{\pi}\right)B^2$$

$$\text{Skewness} = \sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right)B^3$$

$$\text{Kurtosis} = ((\pi(3\pi - 4) - 12)\,\frac{B^4}{\pi^2}$$

$$\text{Mode} = A$$

$$\text{Median} \approx A + 0.67449\,B$$

$$\text{Q1} \approx A + 0.31864\,B \quad \text{Q3} \approx A + 1.15035\,B$$

Notes

1. The HalfNormal distribution is used in place of the **Normal** distribution when only the magnitudes (absolute values) of observations are available.

Aliases and Special Cases

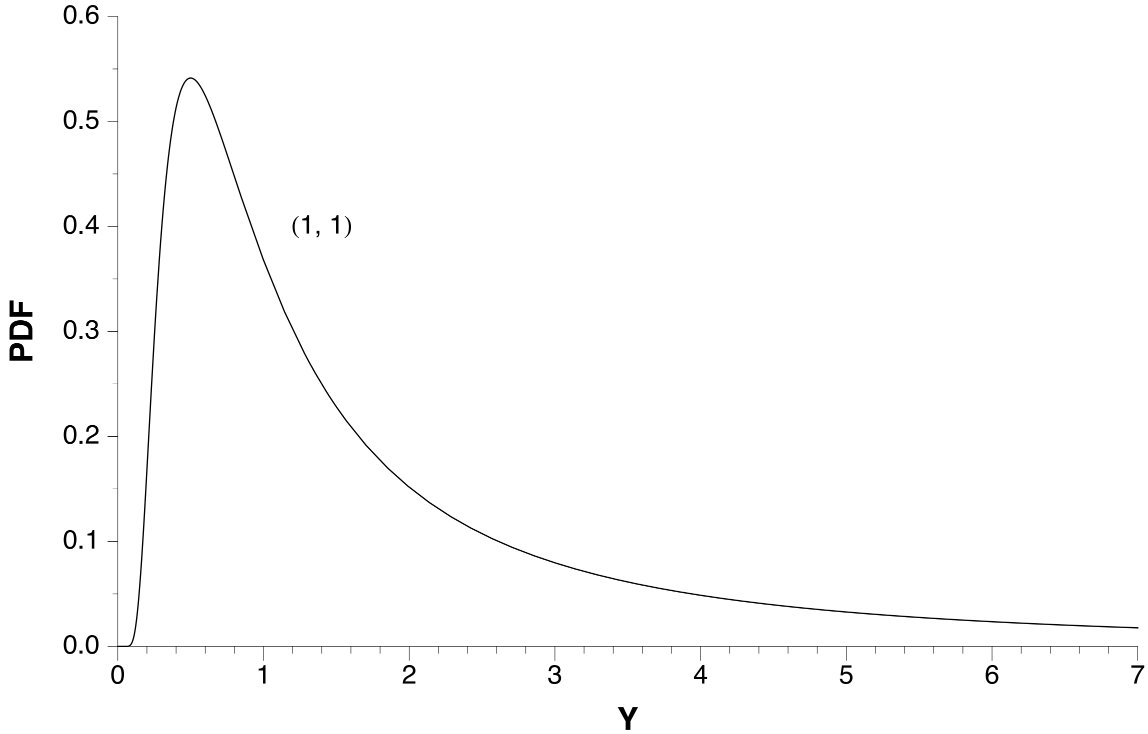
1. The HalfNormal distribution is a special case of both the **Chi** and the **FoldedNormal** distributions.

Characterizations

1. If X $\sim$**Normal**(A, B) is folded (to the right) about its mean, A, the resulting distribution is HalfNormal(A, B).

InverseGamma(A,B)

$y > 0, \quad A, B > 0$



$$\text{PDF} = \frac{A^B}{\Gamma(B)} y^{-(B+1)} \exp\left(-\frac{A}{y}\right)$$

$$\text{CDF} = \frac{1}{\Gamma(B)}\,\Gamma\left(B, \frac{A}{y}\right)$$

$$\text{CF} = \frac{2\,(-iat)^{B/2}}{\Gamma(B)} K_B\left(2\sqrt{-iat}\right)$$

Parameters – A: Scale, B: Shape

Moments, etc.

$$\text{Mean} = \frac{A}{B-1}$$

$$\text{Variance} = \frac{A^2}{(B-2)(B-1)^2}$$

$$\text{Skewness} = \frac{4\,A^3}{(B-3)(B-2)(B-1)^3}$$

$$\text{Kurtosis} = \frac{3\,(B+5)A^4}{(B-4)(B-3)(B-2)(B-1)^4}$$

$$\text{Mode} = \frac{A}{B+1}$$

$$\text{Median, Q1, Q3}: \textit{no simple closed form}$$

Notes

1. Moment r exists iff $B > r$.

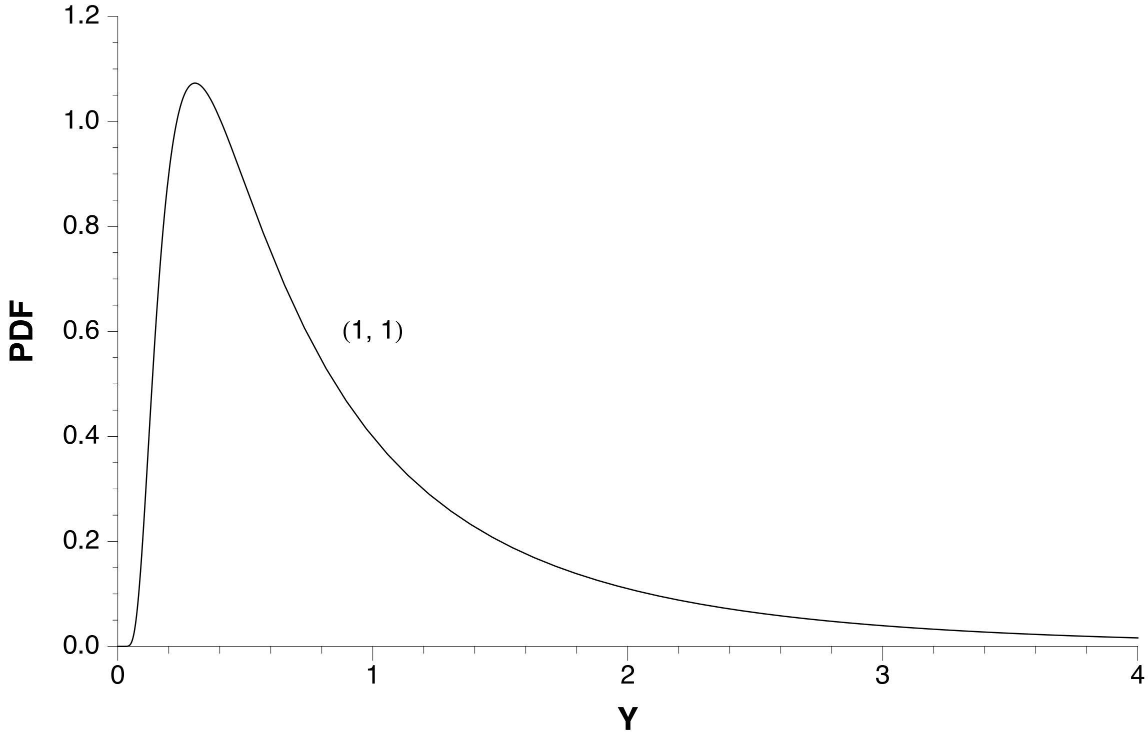
Aliases and Special Cases

1. The InverseGamma distribution is a special case of the *Pearson (Type 5)* distribution.

Characterizations

1. If $X \sim$**Gamma**$(0, \beta, \gamma)$, then $1/X \sim$InverseGamma$(1/\beta, \gamma)$.

InverseNormal(A,B) $\quad\quad y > 0, \quad A, B > 0$



$$\text{PDF} = \sqrt{\frac{B}{2\pi y^3}}\,\exp\left[-\frac{B}{2\,y}\left(\frac{y-A}{A}\right)^2\right]$$

$$\text{CDF} = \Phi\left[\sqrt{\frac{B}{y}}\left(\frac{y-A}{A}\right)\right] + \exp\left(\frac{2\,B}{A}\right)\Phi\left[-\sqrt{\frac{B}{y}}\left(\frac{y+A}{A}\right)\right]$$

$$\text{CF} = \exp\left[\frac{B}{A}\left(1-\sqrt{1-\frac{2itA^2}{B}}\right)\right]$$

Parameters – A: Location and scale, B: Scale

Moments, etc.

$$\text{Mean} = A$$

$$\text{Variance} = \frac{A^3}{B}$$

$$\text{Skewness} = \frac{3A^5}{B^2}$$

$$\text{Kurtosis} = \frac{15A^7}{B^3} + \frac{3A^6}{B^2}$$

$$\text{Mode} = \frac{A}{2B}\left(\sqrt{9A^2 + 4B^2} - 3A\right)$$

Median, Q1, Q3 : *no simple closed form*

Notes

Aliases and Special Cases

1. The InverseNormal distribution is also called the *Wald* distribution.

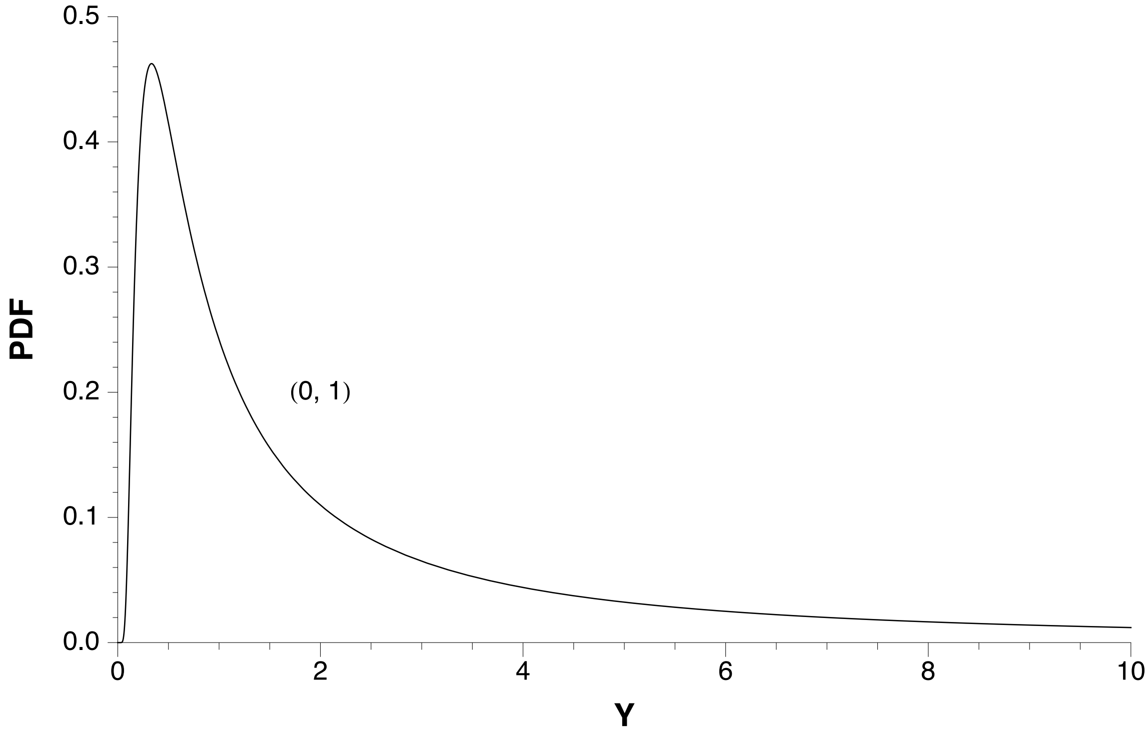
Characterizations

1. If X_i ~**InverseNormal**(A, B), then $\sum_{i=1}^{n} X_i$ ~**InverseNormal**(nA, n^2B).

2. If X ~**InverseNormal**(A, B), then kX ~**InverseNormal**(kA, kB).

Lévy(A,B)

$y \geq A, \quad B > 0$



$$\text{PDF} = \sqrt{\frac{B}{2\,\pi}}\,(y-A)^{-3/2}\exp\left(-\frac{B}{2\,(y-A)}\right)$$

$$\text{CDF} = 1 - \mathit{erf}\left(\sqrt{\frac{B}{2\,(y-A)}}\right)$$

$$\text{CF} = \exp\left(iAt - \sqrt{-2iBt}\right)$$

Parameters – A (μ): Location, B (σ): Scale

Moments, etc.

Moments do not exist.

$$\text{Mode} = A + \frac{B}{3}$$

$$\text{Median} \approx A + 2.19811\,B$$

$$\text{Q1} \approx A + 0.75568\,B \quad \text{Q3} \approx A + 9.84920\,B$$

Notes

1. The Lévy distribution is one of the class of stable distributions.

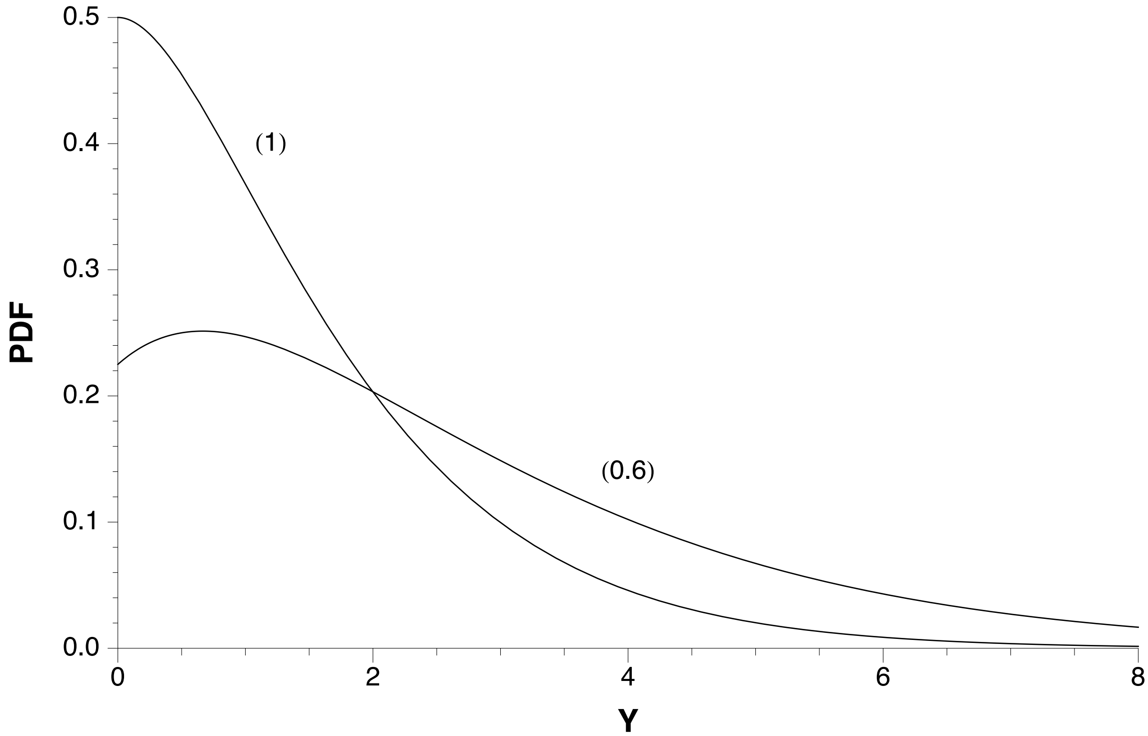
Aliases and Special Cases

1. The Lévy distribution is sometimes called the *Lévy alpha-stable* distribution.

Characterizations

1. The Lévy distribution is sometimes used in financial modeling due to its long tail.

Lindley(A) $y > 0$



$$\text{PDF} = \frac{A^2}{A+1}(y+1)\exp(-Ay)$$

$$\text{CDF} = 1 - \frac{A(y+1)+1}{A+1}\exp(-Ay)$$

$$\text{CF} = \frac{A^2(A+1-it)}{(a+1)(a-it)^2}$$

Parameters – A: Scale

Moments, etc.

$$\text{Mean} = \frac{A+2}{A(A+1)}$$

$$\text{Variance} = \frac{2}{A^2} - \frac{1}{(A+1)^2}$$

$$\text{Skewness} = \frac{4}{A^3} - \frac{2}{(A+1)^3}$$

$$\text{Kurtosis} = \frac{3\,(8 + A(32 + A(44 + 3A(A+8))))}{A^4(A=1)^4}$$

$$\text{Mode} = \begin{cases} \dfrac{1-A}{A}, & A \le 1 \\ 0, & A > 1 \end{cases}$$

$\text{Median},\ \text{Q1},\ \text{Q3}$: *no simple closed form*

Notes

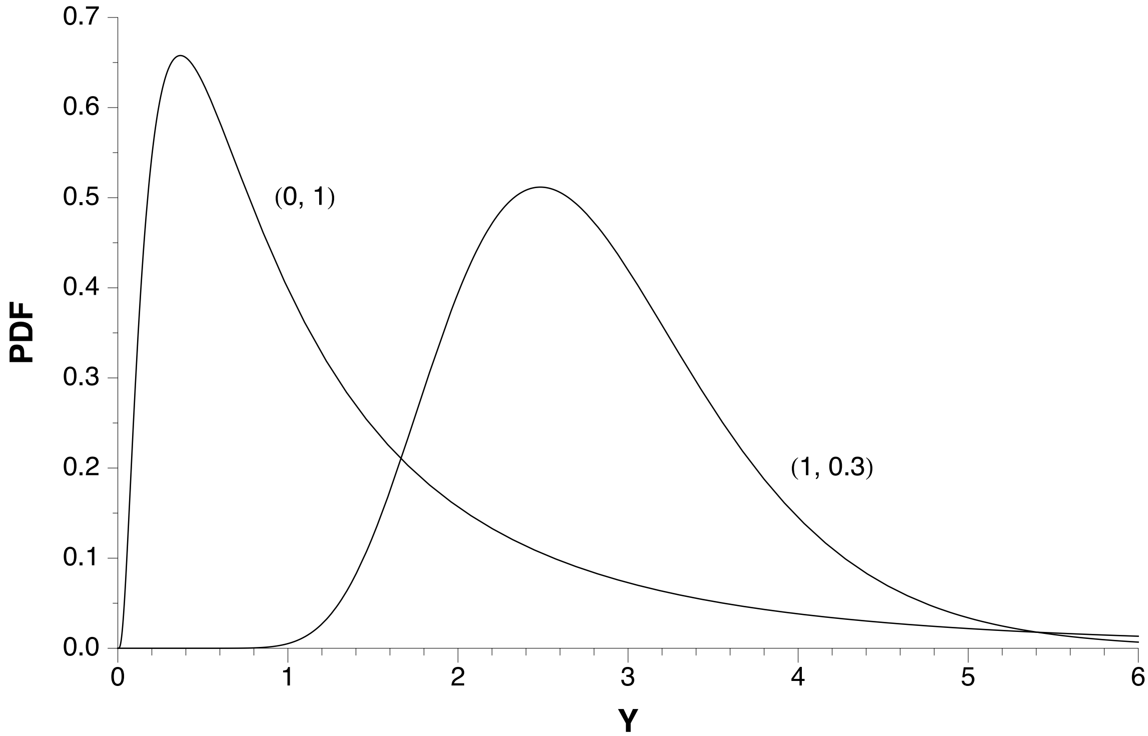
Aliases and Special Cases

Characterizations

1. The Lindley distribution is sometimes used in applications of queueing theory.

LogNormal(A,B)

$y > 0, \quad B > 0$



$$\text{PDF} = \frac{1}{B\,y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(y) - A}{B}\right)^2\right]$$

$$\text{CDF} = \Phi\left(\frac{\log(y) - A}{B}\right)$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A (μ): Location, B (σ): Scale, both measured in log space

Moments, etc.

$$\text{Mean} = \exp\left(A + \frac{B^2}{2}\right)$$

$$\text{Variance} = \left(\exp(B^2) - 1\right)\exp\left(2A + B^2\right)$$

$$\text{Skewness} = \left(\exp(B^2) - 1\right)^2\left(\exp(B^2) + 2\right)\exp\left(3A + \frac{3B^2}{2}\right)$$

$$\text{Kurtosis} = \left(\exp(B^2) - 1\right)^2\left[-3 + \exp\left(2B^2\right)\left(3 + \exp\left(B^2\right)\left(\exp\left(B^2\right) + 2\right)\right)\right]$$

$$\text{Mode} = \exp\left(A - B^2\right)$$

$$\text{Median} = \exp(A)$$

$$\mathrm{Q1} \approx \exp(A - 0.67449\,B) \quad \mathrm{Q3} \approx \exp(A + 0.67449\,B)$$

Notes

1. There are several alternate forms for the PDF, some of which have more than two parameters.

2. Parameters A and B are the mean and standard deviation of y in (natural) log space. Therefore, their units are similarly transformed.

Aliases and Special Cases

1. The LogNormal distribution is sometimes called the *Cobb-Douglas* distribution, especially when applied to econometric data.

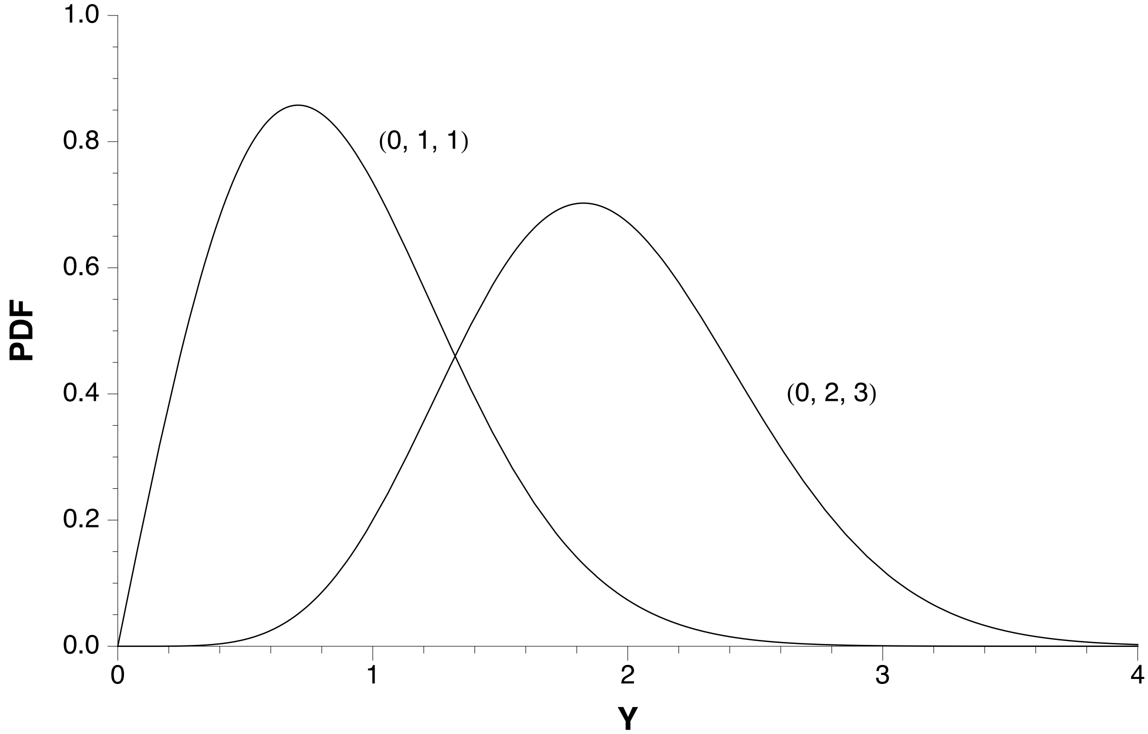
2. It is also known as the *antilognormal* distribution.

Characterizations

1. As the PDF suggests, the LogNormal distribution is the distribution of a random variable which, in log space, is $\sim$**Normal**.

Nakagami(A,B,C)

$y > \mathrm{A}, \quad \mathrm{B}, \mathrm{C} > 0$



$$\text{PDF} = \frac{2\,C^C}{B\,\Gamma(C)} \left(\frac{y-A}{B}\right)^{2C-1} \exp\left[-C\left(\frac{y-A}{B}\right)^2\right]$$

$$\text{CDF} = 1 - \frac{\Gamma\left[C, C\left(\frac{y-A}{B}\right)^2\right]}{\Gamma(C)}$$

$$\text{CF} = \exp(iAt)\frac{\Gamma(C+1/2)}{\sqrt{\pi}}\,U\left(C, \frac{1}{2}, -\frac{B^2t^2}{4\,C}\right)$$

Parameters – A: Location, B: Scale, C: Shape (also, degrees of freedom)

Moments, etc.

$$\text{Mean} = A + \frac{\Gamma\left(C+\frac{1}{2}\right)}{\sqrt{C}\,\Gamma(C)}B$$

$$\text{Variance} = \left[1 - \frac{C\,\Gamma\left(C+\frac{1}{2}\right)^2}{\Gamma(C+1)^2}\right]B^2$$

$$\text{Skewness} = \left[\frac{2\,\Gamma\left(C+\frac{1}{2}\right)^3 + \frac{1}{2}(1-4C)\Gamma(C)^2\Gamma\left(C+\frac{1}{2}\right)}{C^{3/2}\,\Gamma(C)^3}\right]B^3$$

$$\text{Kurtosis} = \left[-\frac{3\,\Gamma\left(C+\frac{1}{2}\right)^4}{C^2\,\Gamma(C)^4} + \frac{1}{C} + \frac{2\,(C-1)\Gamma\left(C+\frac{1}{2}\right)^2}{\Gamma(C+1)^2} + 1\right]B^4$$

$$\text{Mode} = A + \frac{B}{\sqrt{2}}\sqrt{\frac{2\,C-1}{C}}$$

$$\text{Median, Q1, Q3} : \textit{no simple closed form}$$

Notes

1. The Nakagami distribution is usually defined only for $y > 0$.
2. Its scale parameter is also defined differently in many references.

Aliases and Special Cases

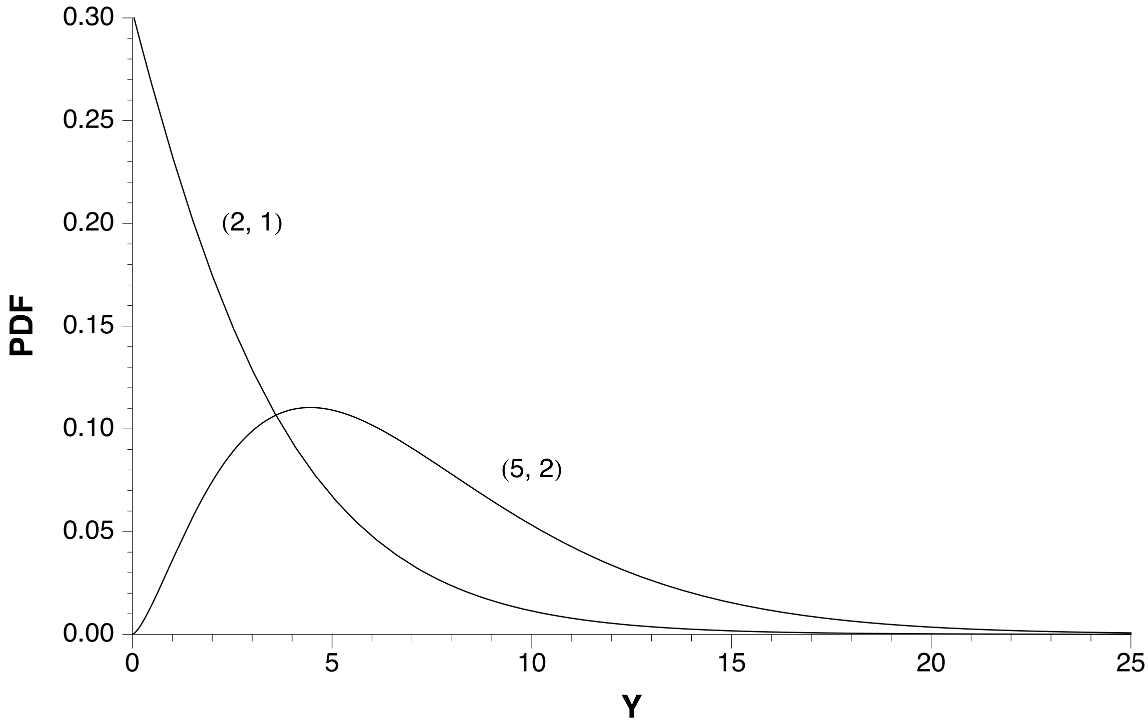
1. cf. **Chi** distribution.

Characterizations

1. The Nakagami distribution is a generalization of the **Chi** distribution.
2. It is sometimes used as an approximation to the **Rice** distribution.

NoncentralChiSquare(A,B)

$$\mathrm{y} \geq 0, \quad \mathrm{A} > 0, \quad \mathrm{B} > 0$$



$$\mathrm{PDF} = \frac{1}{2} \exp\left(-\frac{y+B}{2}\right) \left(\frac{y}{B}\right)^{A/4-1/2} I_{A/2-1}\left(\sqrt{B\,y}\right)$$

$$\mathrm{CDF} = 1 - Q_{A/2}\left(\sqrt{B}, \sqrt{y}\right)$$

$$\mathrm{CF} = \exp\left(-\frac{B\,t}{i+2\,t}\right) (1-2it)^{-A/2}$$

Parameters – A: Shape (degrees of freedom), B: Scale (noncentrality)

Moments, etc.

$$\mathrm{Mean} = A + B$$

$$\mathrm{Variance} = 2(A + 2\,B)$$

$$\mathrm{Skewness} = 8(A + 3\,B)$$

$$\mathrm{Kurtosis} = 12\left(A^2 + 4\,A(1+B) + 4\,B(4+B)\right)$$

$\mathrm{Mode},\ \mathrm{Median},\ \mathrm{Q1},\ \mathrm{Q3}$: *no simple closed form*

Notes

1. *Noncentrality* is a parameter common to many distributions.

Aliases and Special Cases

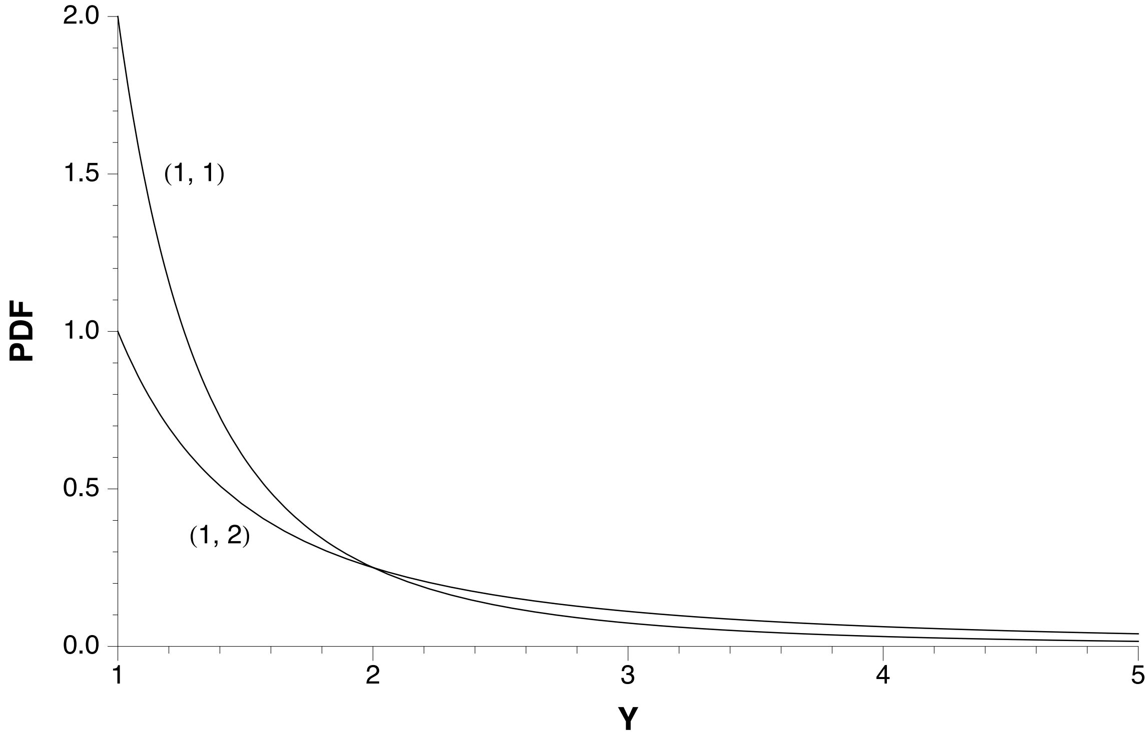
1. As $B \to 0$, the NoncentralChiSquare distribution becomes the **ChiSquare** distribution.

Characterizations

1. The NoncentralChiSquare distribution describes the sum of squares of $X_i \sim \mathbf{Normal}(\mu_\mathbf{i}, \mathbf{1})$ where $i = 1 \ldots A$.

Pareto1(A,B)

$0 < \mathrm{A} \leq \mathrm{y}, \quad \mathrm{B} > 0$



$$\mathrm{PDF} = \frac{B A^B}{y^{B+1}}$$

$$\mathrm{CDF} = 1 - \left(\frac{A}{y}\right)^B$$

$$\mathrm{CF} = B\,(-iAy)^B\,\Gamma(-b, -iAy)$$

Parameters – A: Location and scale, B: Shape

Moments, etc.

$$\mathrm{Mean} = \frac{A\,B}{B-1}$$

$$\mathrm{Variance} = \frac{A^2 B}{(B-2)(B-1)^2}$$

$$\mathrm{Skewness} = \frac{2A^3 B(B+1)}{(B-3)(B-2)(B-1)^3}$$

$$\mathrm{Kurtosis} = \frac{3A^4 B(3B^2+B+2)}{(B-4)(B-3)(B-2)(B-1)^4}$$

$$\mathrm{Mode} = A$$

$$\text{Median} = A\sqrt[B]{2}$$

$$\text{Q1} = A\sqrt[B]{\frac{4}{3}} \quad \text{Q3} = A\sqrt[B]{4}$$

$$\text{RandVar} = \frac{A}{\sqrt[B]{u}}$$

Notes

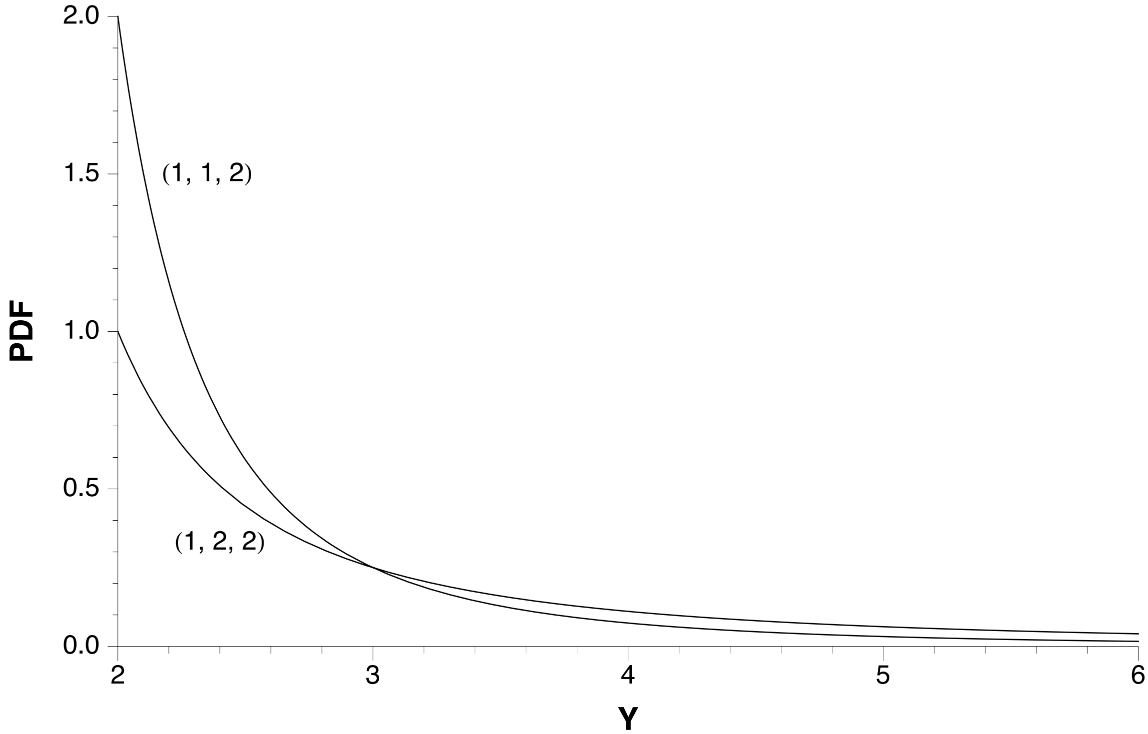
1. Moment r exists iff $B > r$.
2. The name Pareto is applied to a class of distributions.

Aliases and Special Cases

Characterizations

1. The Pareto distribution is often used as an income distribution. That is, the probability that a random income in some defined population exceeds a minimum, A, is $\sim$Pareto.

Pareto2(A,B,C) $\mathrm{y} \geq \mathrm{C}, \quad \mathrm{A}, \mathrm{B} > 0$



$$\mathrm{PDF} = \frac{B}{A}\left(\frac{A}{y+A-C}\right)^{B+1}$$

$$\mathrm{CDF} = 1 - \left(\frac{A}{y+A-C}\right)^{B}$$

$$\mathrm{CF} = B\,(-iAy)^{B} \exp\left(iy(C-A)\right)\Gamma(-b,-iAy)$$

Parameters – A: Scale, B: Shape, C: Location

Moments, etc.

$$\mathrm{Mean} = \frac{A}{B-1} + C$$

$$\mathrm{Variance} = \frac{A^2 B}{(B-2)(B-1)^2}$$

$$\mathrm{Skewness} = \frac{2A^3 B(B+1)}{(B-3)(B-2)(B-1)^3}$$

$$\mathrm{Kurtosis} = \frac{3A^4 B(3B^2+B+2)}{(B-4)(B-3)(B-2)(B-1)^4}$$

$$\mathrm{Mode} = A$$

$$\text{Median} = A\left(\sqrt[B]{2} - 1\right) + C$$

$$\text{Q1} = A\left(\sqrt[B]{\frac{4}{3}} - 1\right) + C \quad \text{Q3} = A\left(\sqrt[B]{4} - 1\right) + C$$

$$\text{RandVar} = A\left(\frac{1}{\sqrt[B]{u}} - 1\right) + C$$

Notes

1. Moment r exists iff $B > r$.

2. The name Pareto is applied to a class of distributions.

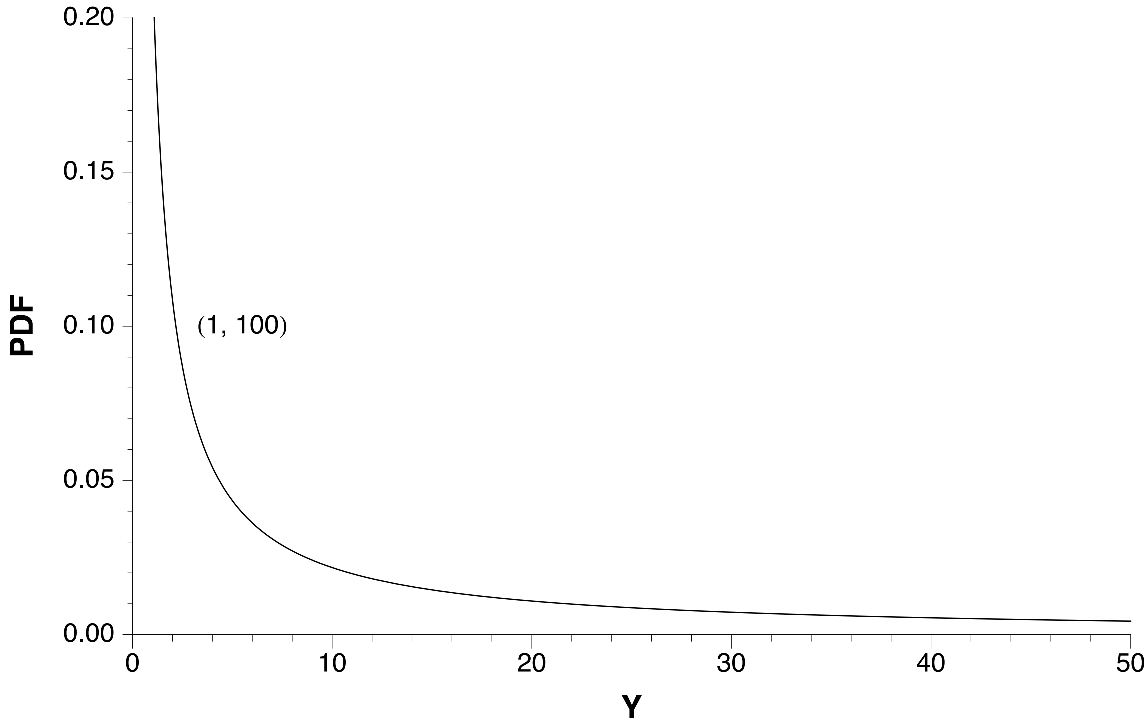
Aliases and Special Cases

Characterizations

1. The Pareto distribution is often used as an income distribution. That is, the probability that a random income in some defined population exceeds a minimum, A, is $\sim$Pareto.

Reciprocal(A,B)

$0 < \mathrm{A} \leq \mathrm{y} \leq \mathrm{B}$



$$\mathrm{PDF} = \frac{1}{y \log\left(\frac{B}{A}\right)}$$

$$\mathrm{CDF} = \frac{\log(y)}{\log\left(\frac{B}{A}\right)}$$

$$\mathrm{CF} = \frac{Ei(iBt) - Ei(iAt)}{\log\left(\frac{B}{A}\right)}$$

Parameters – A, B: Shape

Moments, etc. $d = \log\left(\frac{B}{A}\right)$

$$\mathrm{Mean} = \frac{B - A}{d}$$

$$\mathrm{Variance} = \frac{(B - A)(d(A + B) + 2(A - B))}{2\, d^2}$$

$$\mathrm{Skewness} = \frac{2d^2\left(B^3 - A^3\right) - 9d(A + B)(A - B)^2 - 12(A - B)^3}{6\, d^3}$$

$$\mathrm{Kurtosis} = \frac{3d^3\left(B^4 - A^4\right) - 16d^2\left(A^2 + AB + B^2\right)(A - B)^2 - 36d(A + B)(A - B)^3 - 36(A - B)^4}{12d^4}$$

$$\text{Mode} = A$$

$$\text{Median} = \sqrt{\frac{B}{A}}$$

$$\text{Q1} = \left(\frac{B}{A}\right)^{1/4} \qquad \text{Q3} = \left(\frac{B}{A}\right)^{3/4}$$

$$\text{RandVar} = \left(\frac{B}{A}\right)^{u}$$

Notes

1. The Reciprocal distribution is unusual in that is has no conventional location or shape parameter.

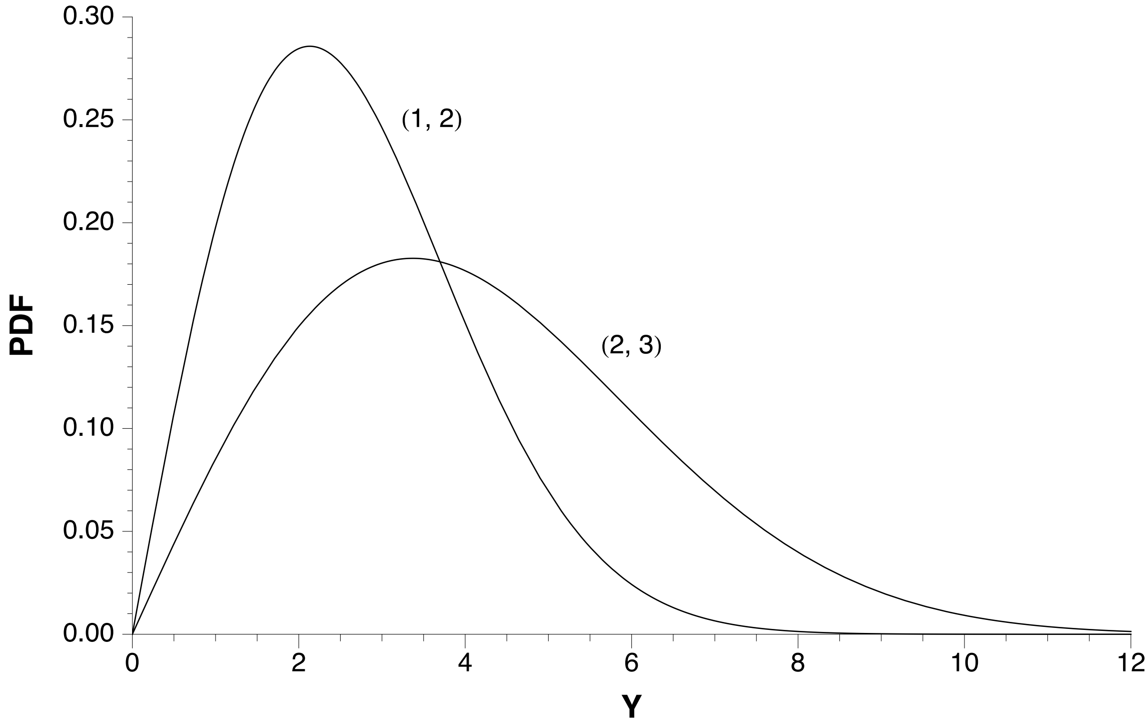
Aliases and Special Cases

Characterizations

1. The Reciprocal distribution is often used to descrobe *1/f noise*.

Rice(A,B)

y > 0, A ≥ 0, B > 0



$$\text{PDF} = \frac{y}{B^2}\exp\left(-\frac{y^2+A^2}{2B^2}\right) I_0\left(\frac{A\,y}{B^2}\right)$$

$$\text{CDF} = 1 - Q_1\left(\frac{A}{B}, \frac{y}{B}\right)$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A: Shape (noncentrality), B: Scale

Moments, etc. $f = L_{1/2}\left(-\frac{A^2}{2B^2}\right)$

$$\text{Mean} = B\sqrt{\frac{\pi}{2}}\, f$$

$$\text{Variance} = A^2 + 2B^2 - \frac{\pi B^2}{2}\, f^2$$

$$\text{Skewness} = B^3\sqrt{\frac{\pi}{2}}\left[\left(-6 - \frac{3A^2}{2B^2}\right) f + \pi\, f^3 + 3\, L_{3/2}\left(-\frac{A^2}{2B^2}\right)\right]$$

$$\text{Kurtosis} = A^4 + 8A^2B^2 + 8B^4 + \frac{3\pi B^2}{4}\, f\left[4\left(A^2+2B^2\right) f - \pi B^2 f^3 - 8B^2\, L_{3/2}\left(-\frac{A^2}{2B^2}\right)\right]$$

$$\text{Mode, Median, Q1, Q3}: \textit{no simple closed form}$$

Notes

1. *Noncentrality* is a parameter common to many distributions.

Aliases and Special Cases

1. The Rice distribution is often called the *Rician* distribution.

2. When $A = 0$, the Rice distribution becomes the *Rayleigh* distribution.

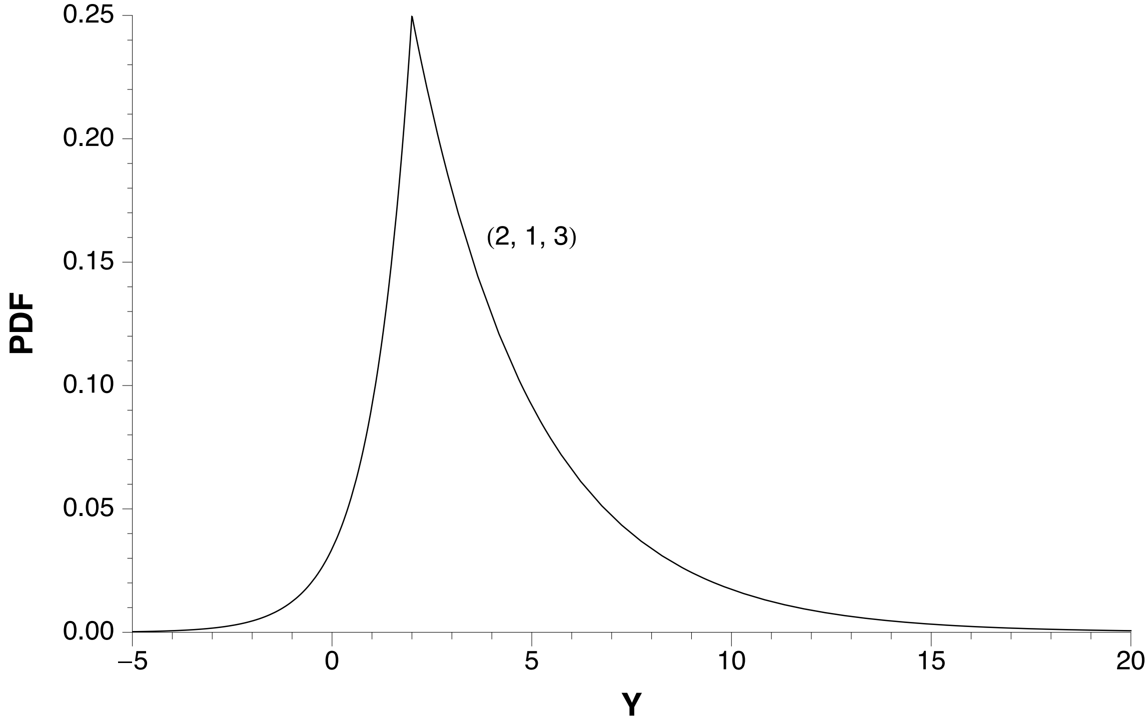
Characterizations

1. The Rice distribution may be considered a noncentral Chi disribution with two degrees of freedom.

2. If X ∼**Normal**(m_1, B) and Y ∼**Normal**(m_2, B), the distance from the origin to (X, Y) ∼Rice$(\sqrt{m_1^2 + m_2^2}, B)$.

3. In communications theory, the Rice distribution is often used to describe the combined power of signal plus noise with the noncentrality parameter corresponding to the signal. The noise would then be ∼*Rayleigh*.

SkewLaplace(A,B,C) $\qquad \mathrm{B, C} > 0$



$$\mathrm{PDF} = \frac{1}{B+C}\begin{cases} \exp\left(\dfrac{y-A}{B}\right), & y \le A \\ \exp\left(\dfrac{A-y}{C}\right), & y > A \end{cases}$$

$$\mathrm{CDF} = \begin{cases} \dfrac{B}{B+C}\exp\left(\dfrac{y-A}{B}\right), & y \le A \\ 1 - \dfrac{C}{B+C}\exp\left(\dfrac{A-y}{C}\right), & y > A \end{cases}$$

$$\mathrm{CF} = \frac{\exp(iAt)}{(Bt-i)(Ct+i)}$$

Parameters – A: Location, B, C: Scale

Moments, etc.

$$\mathrm{Mean} = A - B + C$$

$$\mathrm{Variance} = B^2 + C^2$$

$$\mathrm{Skewness} = 2\left(C^3 - B^3\right)$$

$$\mathrm{Kurtosis} = 9\,B^4 + 6\,B^2C^2 + 9\,C^4$$

$$\mathrm{Mode} = A$$

$$\text{Median, Q1, Q3} : \textit{vary with skewness}$$

$$\text{RandVar} = \begin{cases} A + B\,\log\left((B+C)\,u/B\right), & u \leq \dfrac{B}{B+C} \\ A - C\,\log\left((B+C)(1-u)/C\right), & u > \dfrac{B}{B+C} \end{cases}$$

Notes

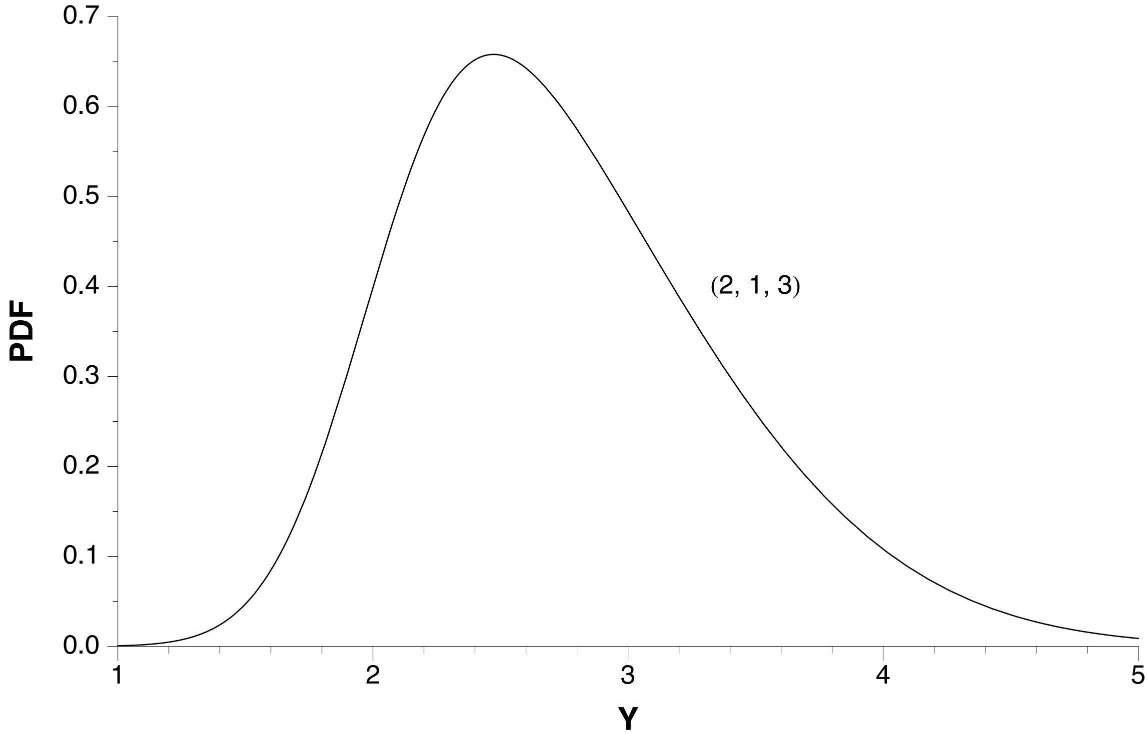
1. Skewed variants of standard distributions are common. This form is just one possibility.
2. In this form of the SkewLaplace distribution, the skewness is controlled by the relative size of the scale parameters.

Aliases and Special Cases

Characterizations

1. The SkewLaplace distribution is used to introduce asymmetry into the **Laplace** distribution.

SkewNormal(A,B,C) $\quad\quad$ B > 0



$$\text{PDF} = \frac{1}{B}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{(y-A)^2}{2B^2}\right)\Phi\left(\frac{C\,(y-A)}{B}\right)$$

$$\text{CDF} = \Phi\left(\frac{y-A}{B}\right) - 2\,OwenT\left(\frac{y-A}{B}, C\right)$$

$$\text{CF} = \exp\left(iAt - \frac{B^2t^2}{2}\right)\left[1 + erf\left(\frac{iBCt}{\sqrt{2}\sqrt{1+C^2}}\right)\right]$$

Parameters – A: Location, B, C: Scale

Moments, etc.

$$\text{Mean} = A + \frac{C}{\sqrt{1+C^2}}\sqrt{\frac{2}{\pi}}\,B$$

$$\text{Variance} = \left[1 - \frac{2\,C^2}{\pi\,(1+C^2)}\right]B^2$$

$$\text{Skewness} = \frac{\sqrt{2}\,(\pi-4)\,C^3}{\pi^{3/2}\,(1+C^2)^{3/2}}\,B^3$$

$$\text{Kurtosis} = \frac{6\,\pi\,C^2(\pi-2) + 3\,\pi^2 + (\pi\,(3\,\pi-4)-12)\,C^4}{\pi^2\,(1+C^2)^2}\,B^4$$

$$\text{Mode, Median, Q1, Q3} : \textit{no simple closed form}$$

Notes

1. Skewed variants of standard distributions are common. This form is just one possibility.
2. Here, skewness has the same sign as parameter C.

Aliases and Special Cases

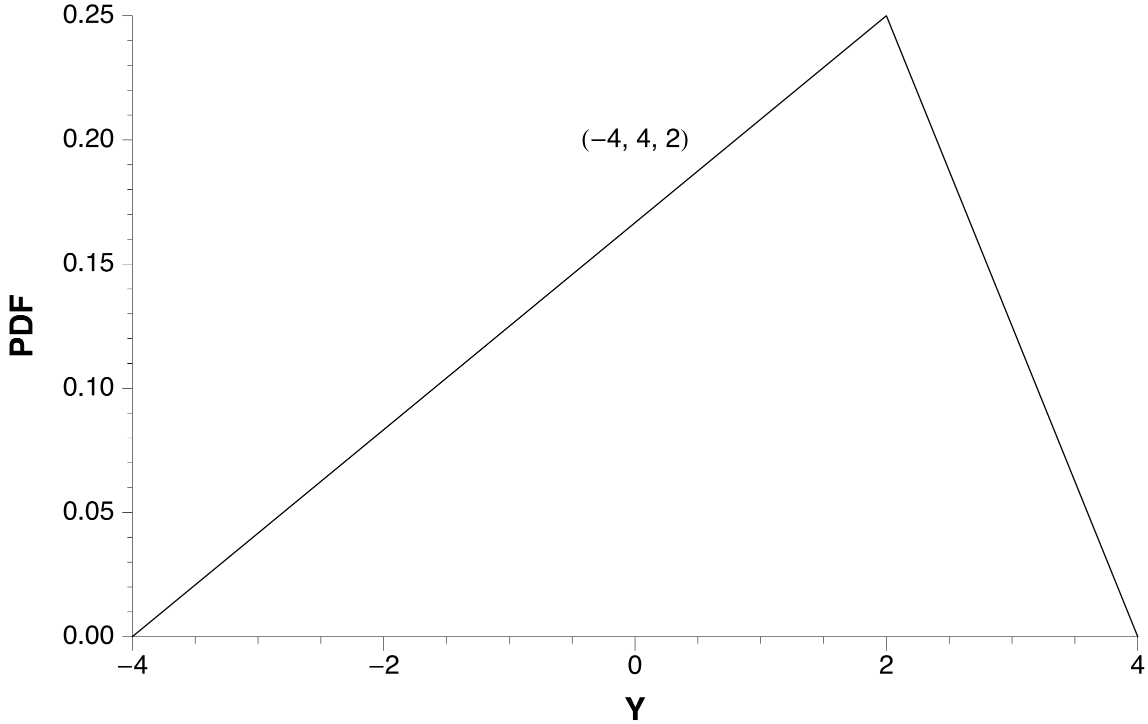
1. This variant of the SkewNormal distribution is sometimes called the *Azzalini* skew normal form.

Characterizations

1. The SkewNormal distribution is used to introduce asymmetry into the **Normal** distribution.
2. When C = 0, the SkewNormal distribution becomes the **Normal** distribution.

Triangular(A,B,C)

$\mathrm{A} \leq \mathrm{y}, \mathrm{C} \leq \mathrm{B}$



$$\text{PDF} = \begin{cases} \dfrac{2\,(y-A)}{(B-A)(C-A)}, & y \leq C \\ \dfrac{2\,(B-y)}{(B-A)(B-C)}, & y > C \end{cases}$$

$$\text{CDF} = \begin{cases} \dfrac{(y-A)^2}{(B-A)(C-A)}, & y \leq C \\ 1 - \dfrac{(B-y)^2}{(B-A)(B-C)}, & y > C \end{cases}$$

$$\text{CF} = \frac{2\ ((a-b)\exp(ict) + \exp(iat)(b-c) + (c-a)\exp(ibt))}{(b-a)(a-c)(b-c)\,t^2}$$

Parameters – A: Location, B: Scale (upper bound), C: Shape (mode)

Moments, etc.

$$\text{Mean} = \frac{1}{3}(A+B+C)$$

$$\text{Variance} = \frac{1}{18}\left(A^2 - A(B+C) + B^2 - BC + C^2\right)$$

$$\text{Skewness} = \frac{1}{270}(A+B-2C)(2A-B-C)(A-2B+C)$$

$$\text{Kurtosis} = \frac{1}{135}\left(A^2 - A(B+C) + B^2 - BC + C^2\right)^2$$

$$\text{Mode} = C$$

$$\text{Median} = \begin{cases} A + \dfrac{1}{\sqrt{2}}\sqrt{(B-A)(C-A)}, & \dfrac{A+B}{2} \le C \\ B - \dfrac{1}{\sqrt{2}}\sqrt{(B-A)(B-C)}, & \dfrac{A+B}{2} > C \end{cases}$$

$$\text{Q1} = \begin{cases} A + \dfrac{1}{2}\sqrt{(B-A)(C-A)}, & \dfrac{1}{4} \le \dfrac{C-A}{B-A} \\ B - \dfrac{\sqrt{3}}{2}\sqrt{(B-A)(B-C)}, & \dfrac{1}{4} > \dfrac{C-A}{B-A} \end{cases}$$

$$\text{Q3} = \begin{cases} A + \dfrac{\sqrt{3}}{2}\sqrt{(B-A)(C-A)}, & \dfrac{3}{4} \le \dfrac{C-A}{B-A} \\ B - \dfrac{1}{2}\sqrt{(B-A)(B-C)}, & \dfrac{3}{4} > \dfrac{C-A}{B-A} \end{cases}$$

$$\text{RandVar} = \begin{cases} A + \sqrt{u\,(B-A)(C-A)}, & u \le \dfrac{C-A}{B-A} \\ B - \sqrt{(1-u)(B-A)(B-C)}, & u > \dfrac{C-A}{B-A} \end{cases}$$

Notes

Aliases and Special Cases

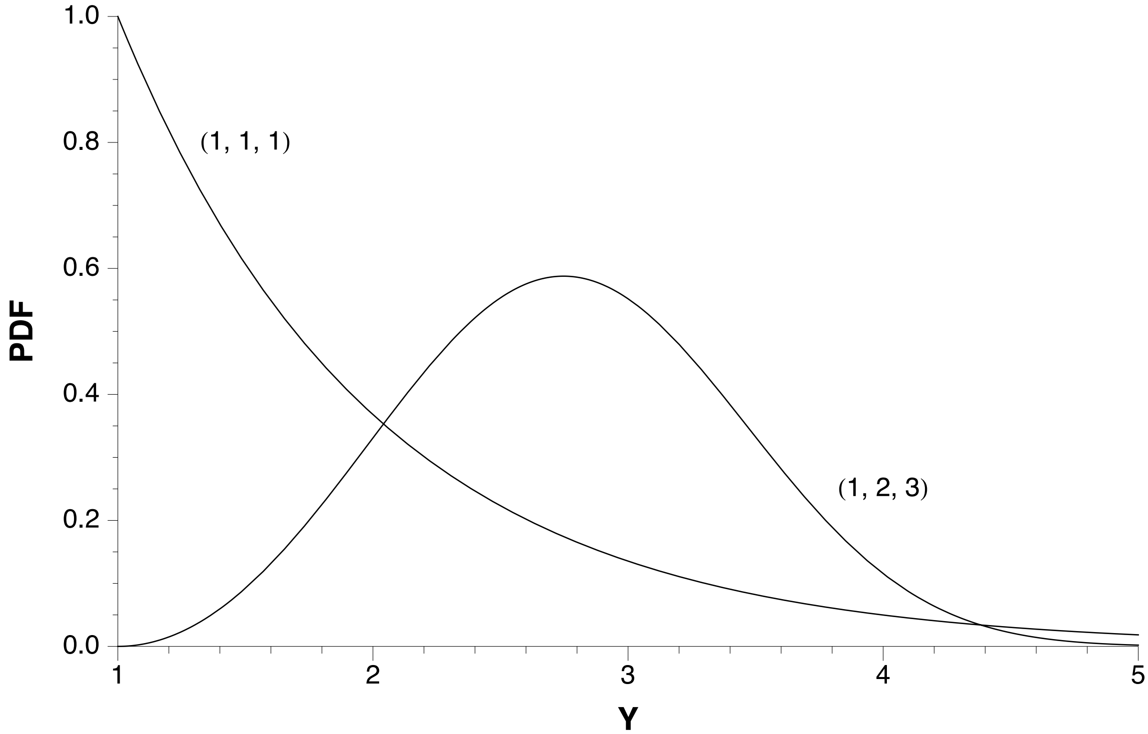
Characterizations

1. If X $\sim$**Uniform**(a, b) and Z $\sim$**Uniform**(c, d) and (b - a) = (d - c), then (X + Z) $\sim$Triangular(a + c, b + d, (a + b + c + d)/2).

2. The Triangular distribution is often used in simulations of bounded data with very little prior information.

Weibull(A,B,C)

$y > A \quad B, C > 0$



$$\text{PDF} = \frac{C}{B}\left(\frac{y-A}{B}\right)^{C-1}\exp\left[-\left(\frac{y-A}{B}\right)^{C}\right]$$

$$\text{CDF} = 1 - \exp\left[-\left(\frac{y-A}{B}\right)^{C}\right]$$

$$\text{CF} = \textit{no simple closed form}$$

Parameters – A: Location, B: Scale, C:Shape

Moments, etc. $G(n) = \Gamma\left(\frac{C+n}{C}\right)$

$$\text{Mean} = A + B\,G(1)$$

$$\text{Variance} = \left[-G^2(1) + G(2)\right]B^2$$

$$\text{Skewness} = \left[2\,G^3(1) - 3\,G(1)G(2) + G(3)\right]B^3$$

$$\text{Kurtosis} = \left[-3\,G^4(1) + 6\,G^2(1)G(2) - 4\,G(1)G(3) + G(4)\right]B^4$$

$$\text{Mode} = \begin{cases} A, & C \le 1 \\ A + B\sqrt[C]{\dfrac{C-1}{C}}, & C > 1 \end{cases}$$

$$\text{Median} = A + B\sqrt[C]{\log(2)}$$

$$\text{Q1} = A + B\sqrt[C]{\log(4/3)} \quad \text{Q3} = A + B\sqrt[C]{\log(4)}$$

$$\text{RandVar} = A + B\sqrt[C]{-\log(u)}$$

Notes

1. The Weibull distribution is roughly symmetric for C near 3.6. When C is smaller/larger, the distribution is left/right-skewed, respectively.

Aliases and Special Cases

1. The Weibull distribution is sometimes known as the *Fréchet* distribution.

2. Weibull(A,B,1) is the **Exponential**(A,B) distribution.

3. Weibull(0,1,2) is the *Rayleigh* distribution.

Characterizations

1. If $X = \left(\frac{y-A}{B}\right)^C$ is ∼**Exponential**(0,1), then y ∼Weibull(A,B,C). Thus, the Weibull distribution is a generalization of the **Exponential** distribution.

2. The Weibull distribution is often used to model extreme events.

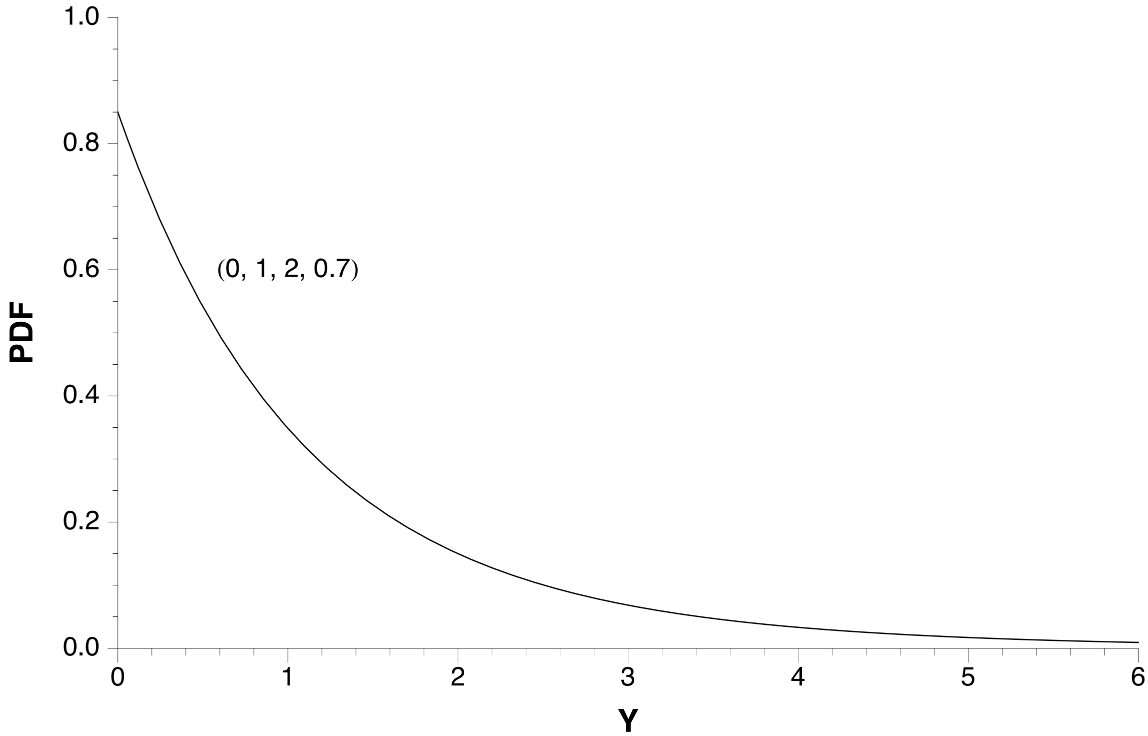
Part III

Continuous: Mixtures

Distributions in this section are mixtures of two distributions. Except for the **InvGammaLaplace** distribution, they are all *component mixtures*.

Expo(A,B)&Expo(A,C)

$\mathrm{B, C} > 0, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = \frac{p}{B}\exp\left(\frac{A-y}{B}\right) + \frac{1-p}{C}\exp\left(\frac{A-y}{C}\right)$$

$$\mathrm{CDF} = p\left[1-\exp\left(\frac{A-y}{B}\right)\right] + (1-p)\left[1-\exp\left(\frac{A-y}{C}\right)\right]$$

$$\mathrm{CF} = \exp(iAt)\left(\frac{p}{1-iBt} + \frac{1-p}{1-iCt}\right)$$

Parameters – A: Location, B, C (λ_1, λ_2): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = A + p\,B + (1-p)\,C$$

$$\mathrm{Variance} = C^2 + 2\,p\,B(B-C) - p^2(B-C)^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

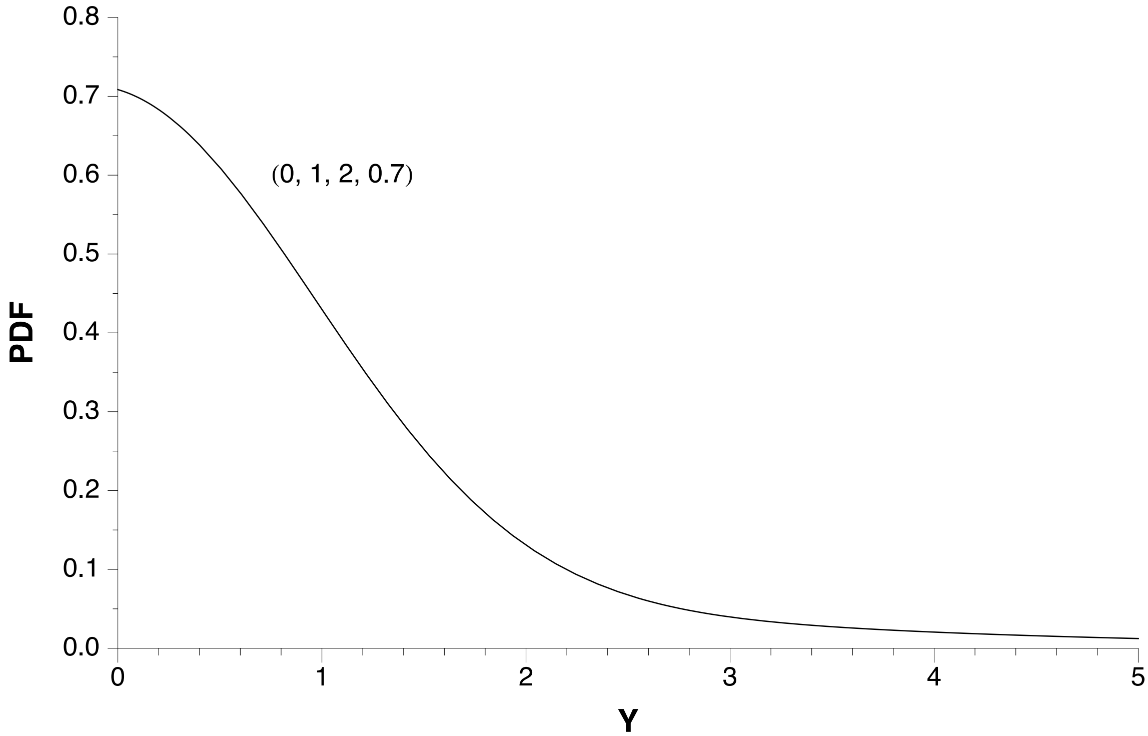
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

HNormal(A,B)&Expo(A,C) $\quad \mathrm{B, C > 0, \quad 0 \leq p \leq 1}$



$$\mathrm{PDF} = \frac{p}{B}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{(y-A)^2}{2\,B^2}\right) + \frac{1-p}{C}\exp\left(\frac{A-y}{C}\right)$$

$$\mathrm{CDF} = 1 - 2\,p\,\Phi\left(\frac{A-y}{B}\right) - (1-p)\left(\frac{A-y}{C}\right)$$

$$\mathrm{CF} = \exp(iAt)\left[p\,\exp\left(-\frac{B^2t^2}{2}\right)(1 + \mathit{erf}\left(\frac{iBt}{\sqrt{2}}\right) + \frac{1-p}{1-iCt}\right]$$

Parameters – A: Location, B, C: Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = A + \sqrt{\frac{2}{\pi}}\left[p\,B + (1-p)\,C\right]$$

$$\mathrm{Variance} = p\,B^2 + (1-p)\,C^2 - \frac{2}{\pi}\left[p\,B + (1-p)\,C\right]^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

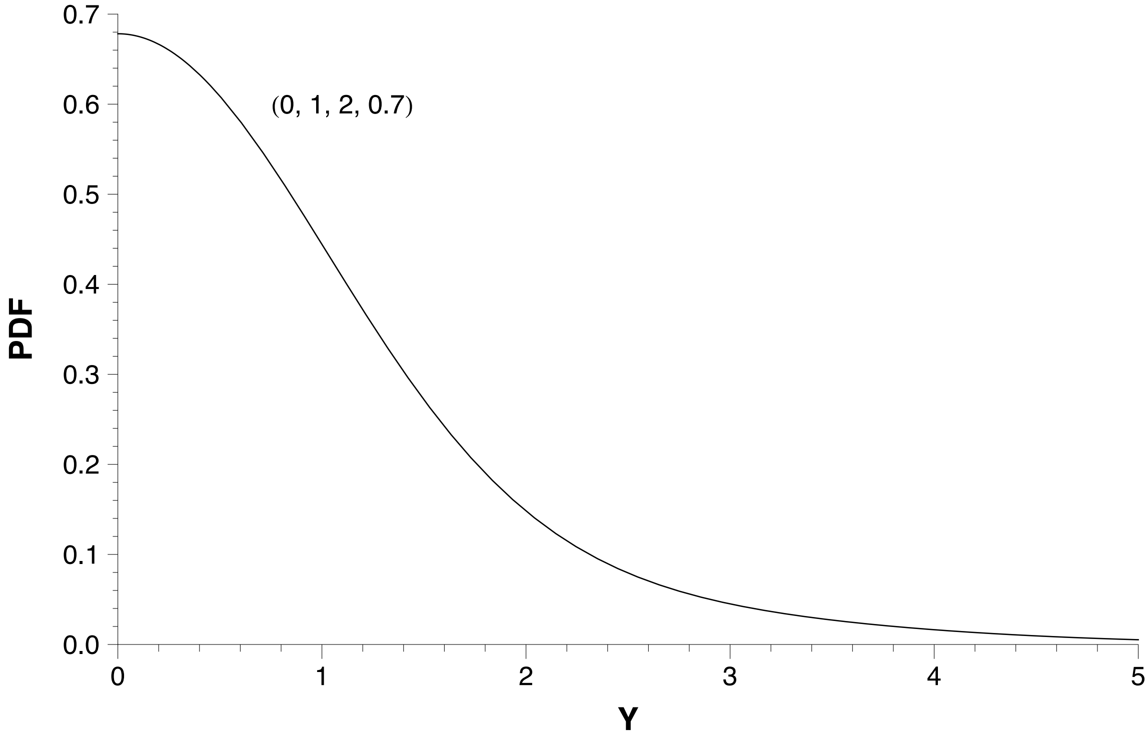
Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

HNormal(A,B)&HNormal(A,C)

$B, C > 0, \quad 0 \leq p \leq 1$



$$\text{PDF} = \sqrt{\frac{2}{\pi}} \left[\frac{p}{B} \exp\left(-\frac{(y-A)^2}{2\,B^2} \right) + \frac{1-p}{C} \exp\left(-\frac{(y-A)^2}{2\,C^2} \right) \right]$$

$$\text{CDF} = 1 - 2\,p\,\Phi\left(\frac{A-y}{B} \right) - 2\,(1-p)\Phi\left(\frac{A-y}{C} \right)$$

$$\text{CF} = \exp(iAt) \left[p\ \exp\left(-\frac{B^2 t^2}{2} \right) (1 + erf\left(\frac{iBt}{\sqrt{2}} \right) + (1-p)\exp\left(-\frac{C^2 t^2}{2} \right) (1 + erf\left(\frac{iCt}{\sqrt{2}} \right) \right]$$

Parameters – A: Location, B, C: Scale, p: Weight of component #1

Moments, etc.

$$\text{Mean} = A + \sqrt{\frac{2}{\pi}} \left[p\,B + (1-p)\,C \right]$$

$$\text{Variance} = p\,B^2 + (1-p)\,C^2 - \frac{2}{\pi} \left[p\,B + (1-p)\,C \right]^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

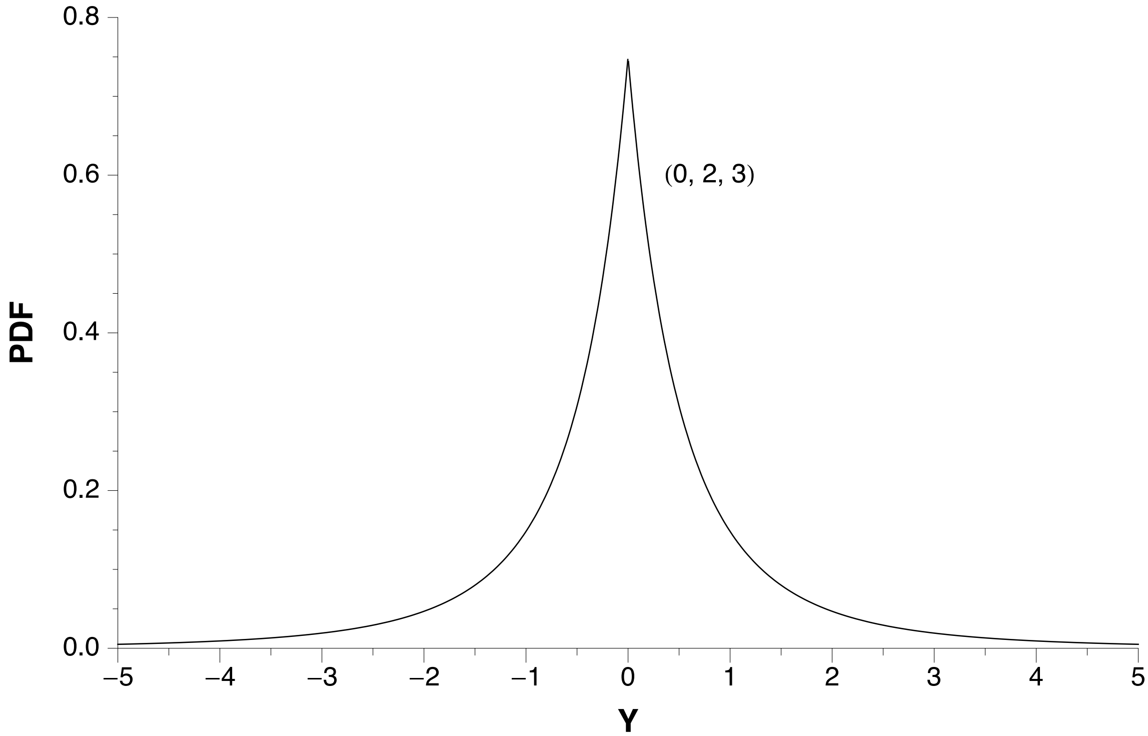
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

InverseGammaLaplace(A,B,C) $\quad\quad \mathrm{B, C} > 0$



$$\mathrm{PDF} = \frac{C}{2\,B}\left(1+\frac{|y-A|}{B}\right)^{-(C+1)}$$

$$\mathrm{CDF} = \begin{cases} \dfrac{1}{2}\left(1+\dfrac{|y-A|}{B}\right)^{-C}, & y \le A \\ 1-\dfrac{1}{2}\left(1+\dfrac{|y-A|}{B}\right)^{-C}, & y > A \end{cases}$$

$$\mathrm{CF} = \textit{no simple closed form}$$

Parameters – A: Location, B: Scale, C: Shape

Moments, etc.

$$\mathrm{Mean} = A$$

$$\mathrm{Variance} = \frac{2\,\Gamma(C-2)}{\Gamma(C)}B^2$$

Notes

1. Moment r exists iff $C > r$.

Aliases and Special Cases

Characterizations

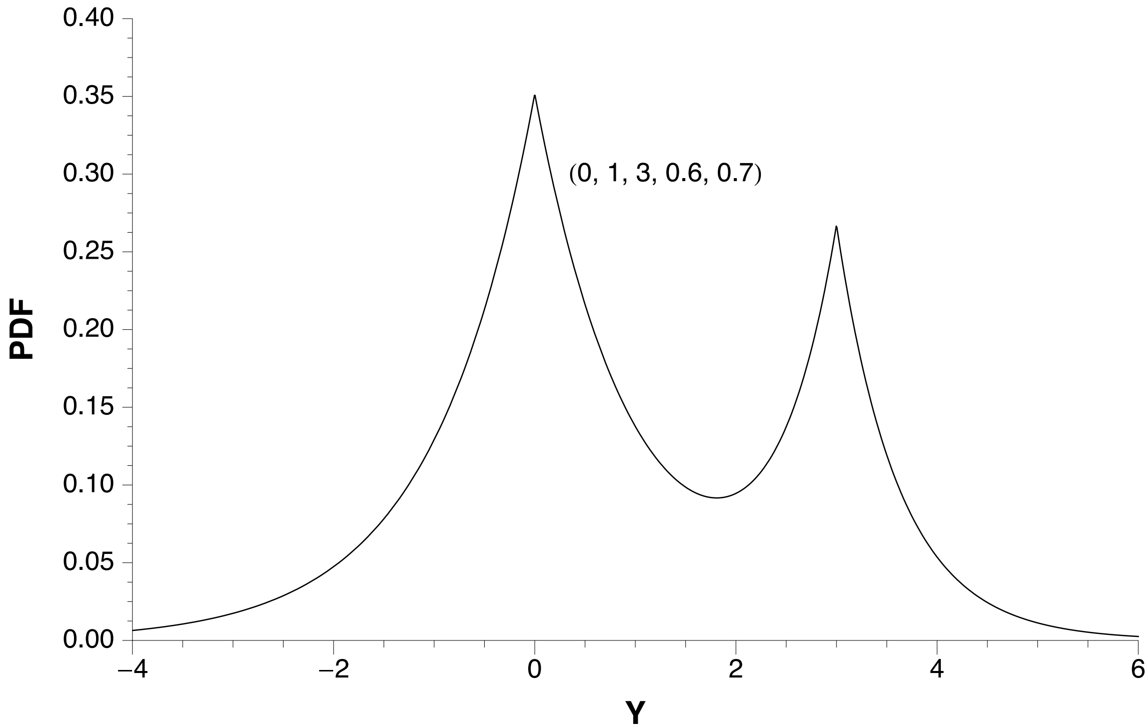
1. The InverseGammaLaplace distribution is a *parameter-mix* distribution in which a **Laplace** distribution has scale parameter, B (λ), modeled as $\sim$**InverseGamma**. This is symbolized

$$\mathbf{Laplace}(\mu, \lambda) \bigwedge_{\lambda} \mathbf{InverseGamma}(A, B)$$

2. The analogous InverseGammaNormal distribution is **StudentsT**.

3. The InverseGammaLaplace distribution is essentially a reflected **Pareto2** distribution.

Laplace(A,B)&Laplace(C,D) $\quad \mathrm{B, D > 0,} \quad 0 \le \mathrm{p} \le 1$



$$\mathrm{PDF} = \frac{p}{2\,B}\exp\left(-\frac{|y-A|}{B}\right) + \frac{1-p}{2\,D}\exp\left(-\frac{|y-C|}{D}\right)$$

$$\mathrm{CDF} = p\begin{cases} \dfrac{1}{2}\exp\left(\dfrac{y-A}{B}\right), & y \le A \\ 1-\dfrac{1}{2}\exp\left(\dfrac{A-y}{B}\right), & y > A \end{cases} + (1-p)\begin{cases} \dfrac{1}{2}\exp\left(\dfrac{y-C}{D}\right), & y \le C \\ 1-\dfrac{1}{2}\exp\left(\dfrac{C-y}{D}\right), & y > C \end{cases}$$

$$\mathrm{CF} = \frac{p\exp(iAt)}{1+B^2\,t^2} + \frac{(1-p)\exp(iCt)}{1+D^2\,t^2}$$

Parameters – A, C (μ_1, μ_2): Location, B, D (λ_1, λ_2): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = p\,A + (1-p)\,C$$

$$\mathrm{Variance} = p\left[2\,B^2 + (1-p)(A-C)^2\right] + 2\,(1-p)D^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.
2. **Warning!** Mixtures usually have several local optima.

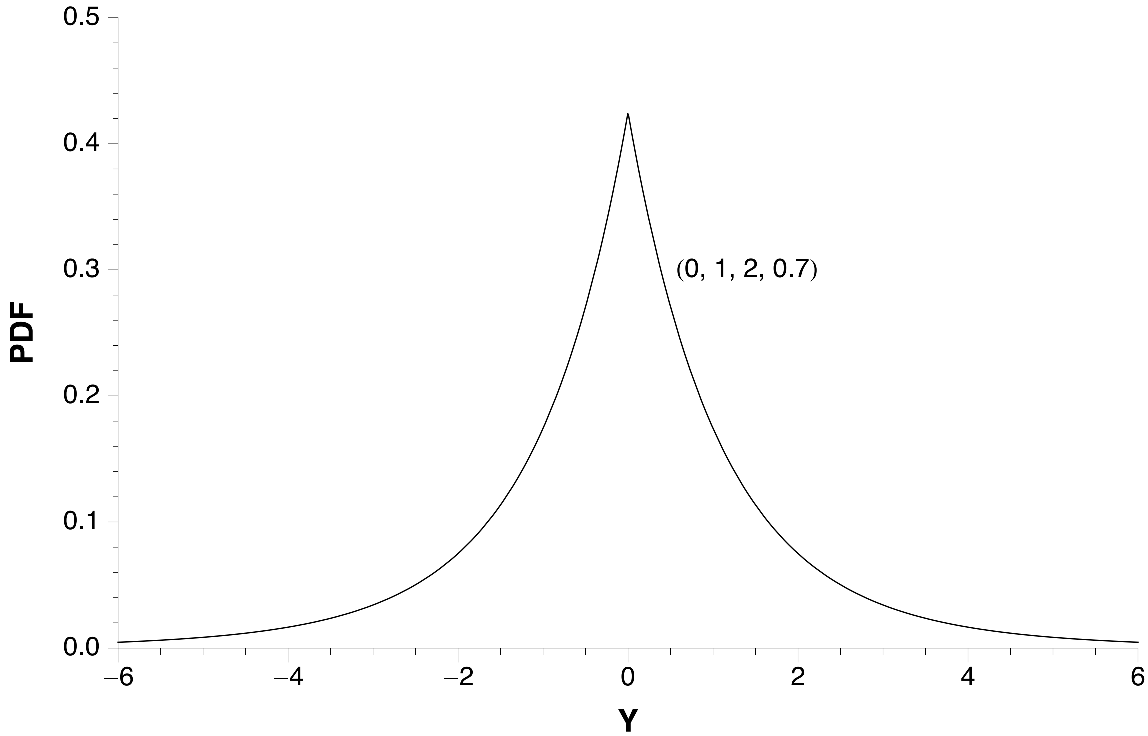
Aliases and Special Cases

1. This binary mixture is often referred to as the *double double-exponential* distribution.

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Laplace(A,B)&Laplace(A,C) $\quad\quad \mathrm{B, C} > 0, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = \frac{p}{2\,B} \exp\left(-\frac{|y-A|}{B}\right) + \frac{1-p}{2\,C} \exp\left(-\frac{|y-A|}{C}\right)$$

$$\mathrm{CDF} = \begin{cases} \frac{1}{2}\left[p \exp\left(\frac{y-A}{B}\right) + (1-p) \exp\left(\frac{y-A}{C}\right)\right], & y \leq A \\ \frac{1}{2}\left[2 - p \exp\left(\frac{A-y}{B}\right) - (1-p) \exp\left(\frac{A-y}{C}\right)\right], & y > A \end{cases}$$

$$\mathrm{CF} = \exp(iAt)\left(\frac{p}{1+B^2\,t^2} + \frac{(1-p)}{1+C^2\,t^2}\right)$$

Parameters – A, C (μ): Location, B, C (λ_1, λ_2): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = A$$

$$\mathrm{Variance} = 2\left[p\,B^2 + (1-p)\,C^2\right]$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

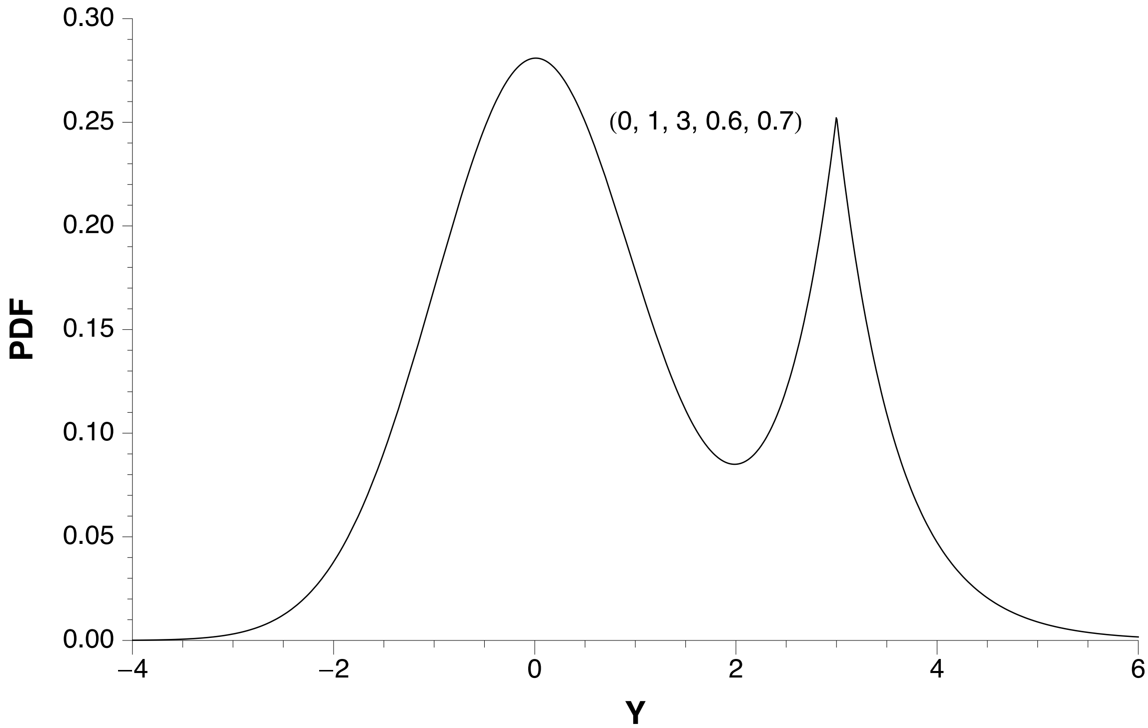
1. This is a special case of the **Laplace(A,B)&Laplace(C,D)** distribution.

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Normal(A,B)&Laplace(C,D)

$\mathrm{B, D} > 0, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = \frac{p}{B\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,B^2}\right] + \frac{1-p}{2\,D} \exp\left[-\frac{|y-C|}{D}\right]$$

$$\mathrm{CDF} = p\,\Phi\left(\frac{y-A}{B}\right) + (1-p)\begin{cases} \dfrac{1}{2}\exp\left(\dfrac{y-C}{D}\right), & y \leq C \\ 1 - \dfrac{1}{2}\exp\left(\dfrac{C-y}{D}\right), & y > C \end{cases}$$

$$\mathrm{CF} = p\,\exp\left(iAt - \frac{B^2t^2}{2}\right) + \frac{(1-p)\exp(iCt)}{1 + D^2\,t^2}$$

Parameters – A, C (μ_1, μ_2): Location, B, D (σ, λ): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = p\,A + (1-p)\,C$$

$$\mathrm{Variance} = p\left[B^2 + (1-p)(A-C)^2\right] + 2\,(1-p)D^2$$

Notes

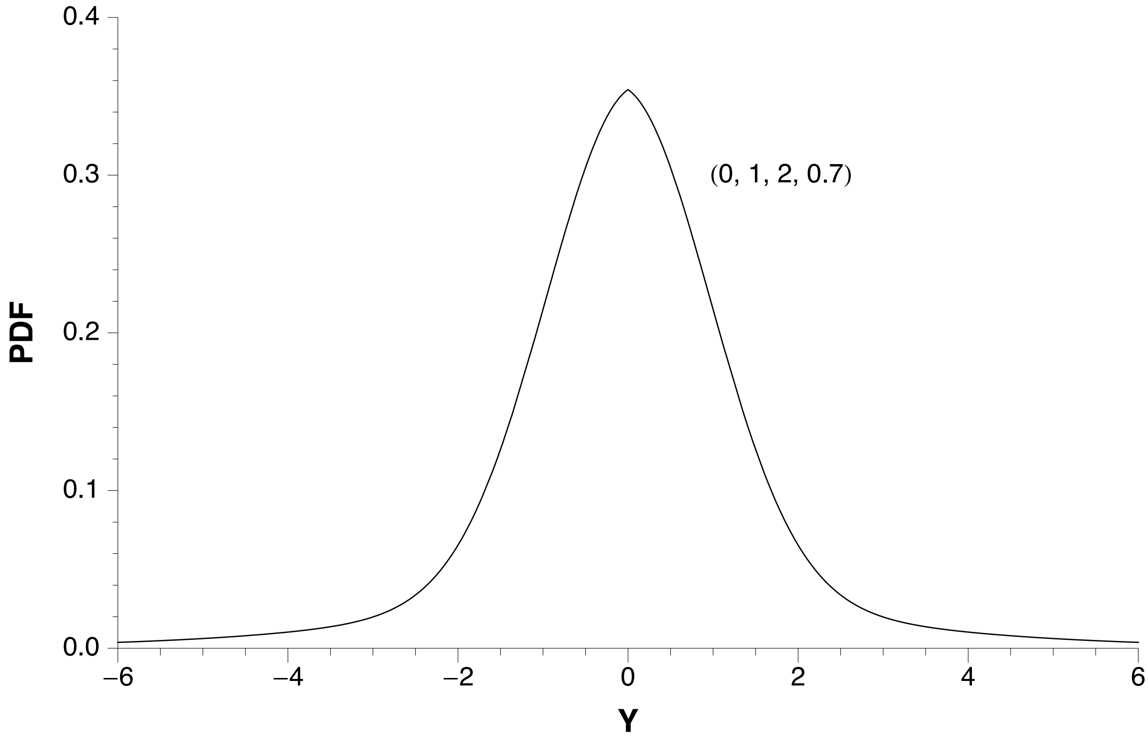
1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Normal(A,B)&Laplace(A,C) $\quad \mathrm{B,C} > 0, \quad 0 \le \mathrm{p} \le 1$



$$\mathrm{PDF} = \frac{p}{B\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,B^2}\right] + \frac{1-p}{2\,C} \exp\left[-\frac{|y-A|}{C}\right]$$

$$\mathrm{CDF} = p\,\Phi\left(\frac{y-A}{B}\right) + (1-p)\begin{cases} \dfrac{1}{2}\exp\left(\dfrac{y-A}{C}\right), & y \le A \\ 1-\dfrac{1}{2}\exp\left(\dfrac{A-y}{C}\right), & y > A \end{cases}$$

$$\mathrm{CF} = \exp(iAt)\left[p\,\exp\left(-\frac{B^2 t^2}{2}\right) + \frac{(1-p)}{1+C^2\,t^2}\right]$$

Parameters – A (μ): Location, B, C (σ, λ): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = p\,A + (1-p)\,C$$

$$\mathrm{Variance} = p\,B^2 + 2\,(1-p)\,C^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

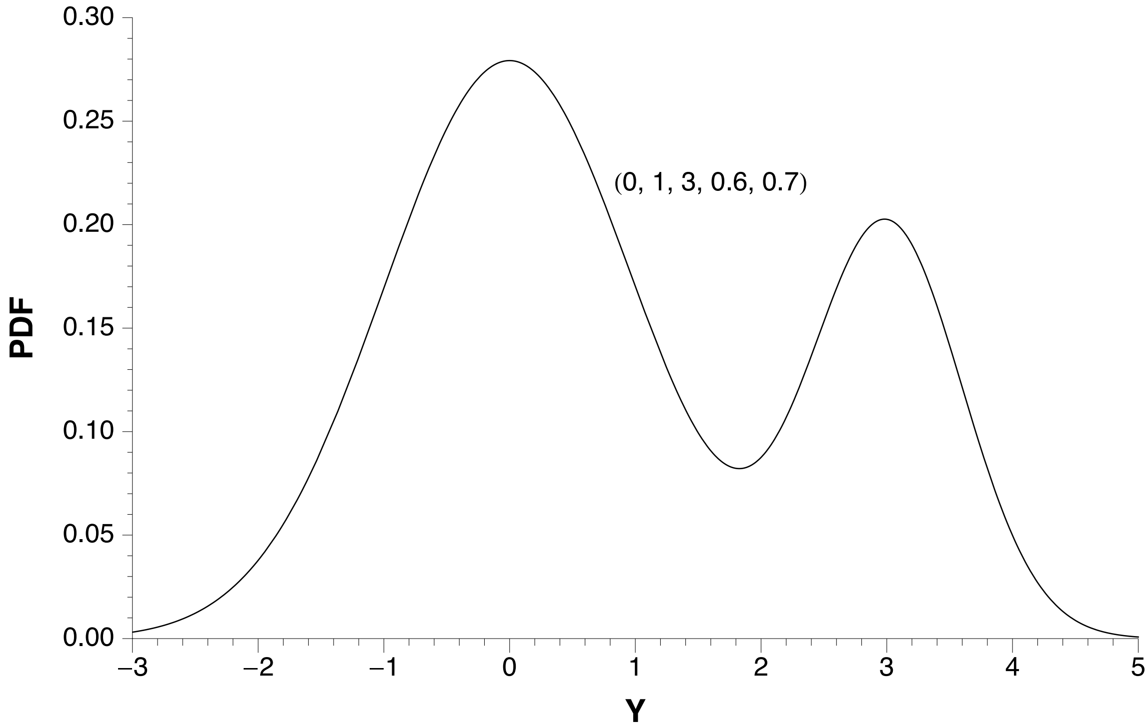
1. This is a special case of the **Normal(A,B)&Laplace(C,D)** distribution.

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Normal(A,B)&Normal(C,D)

$\mathrm{B, D} > 0, \quad 0 \le \mathrm{p} \le 1$



$$\mathrm{PDF} = \frac{p}{B\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,B^2}\right] + \frac{1-p}{D\sqrt{2\pi}} \exp\left[-\frac{(y-C)^2}{2\,D^2}\right]$$

$$\mathrm{CDF} = p\,\Phi\left(\frac{y-A}{B}\right) + (1-p)\,\Phi\left(\frac{y-C}{D}\right)$$

$$\mathrm{CF} = p\,\exp\left(iAt - \frac{B^2t^2}{2}\right) + (1-p)\,\exp\left(iCt - \frac{D^2t^2}{2}\right)$$

Parameters – A, C (μ_1, μ_2): Location, B, D (σ_1, σ_2): Scale, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = p\,A + (1-p)\,C$$

$$\mathrm{Variance} = p\left[B^2 + (1-p)(A-C)^2\right] + (1-p)D^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

3. Whether or not this mixture is bimodal depends partly on parameter p. Obviously, if p is small enough, it will be unimodal regardless of the remaining parameters. If

$$(A - C)^2 > \frac{8\, B^2 D^2}{B^2 + D^2}$$

then there will be some values of p for which this mixture is bimodal. However, if

$$(A - C)^2 < \frac{27\, B^2 D^2}{4\,(B^2 + D^2)}$$

then it will be unimodal.

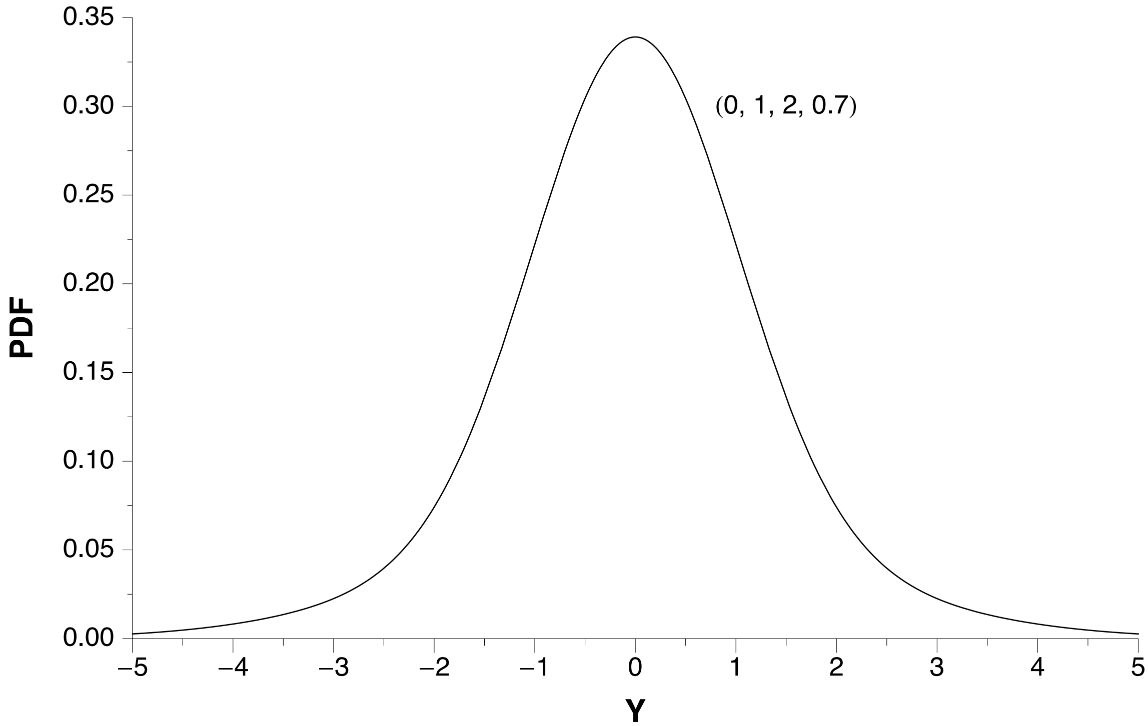
Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Normal(A,B)&Normal(A,C)

$\mathrm{B, C > 0, \quad 0 \leq p \leq 1}$



$$\text{PDF} = \frac{p}{B\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,B^2}\right] + \frac{1-p}{C\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,C^2}\right]$$

$$\text{CDF} = p\,\Phi\left(\frac{y-A}{B}\right) + (1-p)\,\Phi\left(\frac{y-A}{C}\right)$$

$$\text{CF} = \exp(iAt)\left[p\,\exp\left(-\frac{B^2t^2}{2}\right) + (1-p)\,\exp\left(-\frac{C^2t^2}{2}\right)\right]$$

Parameters – A (μ): Location, B, C (σ_1, σ_2): Scale, p: Weight of component #1

Moments, etc.

$$\text{Mean} = A$$

$$\text{Variance} = p\,B^2 + (1-p)\,C^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

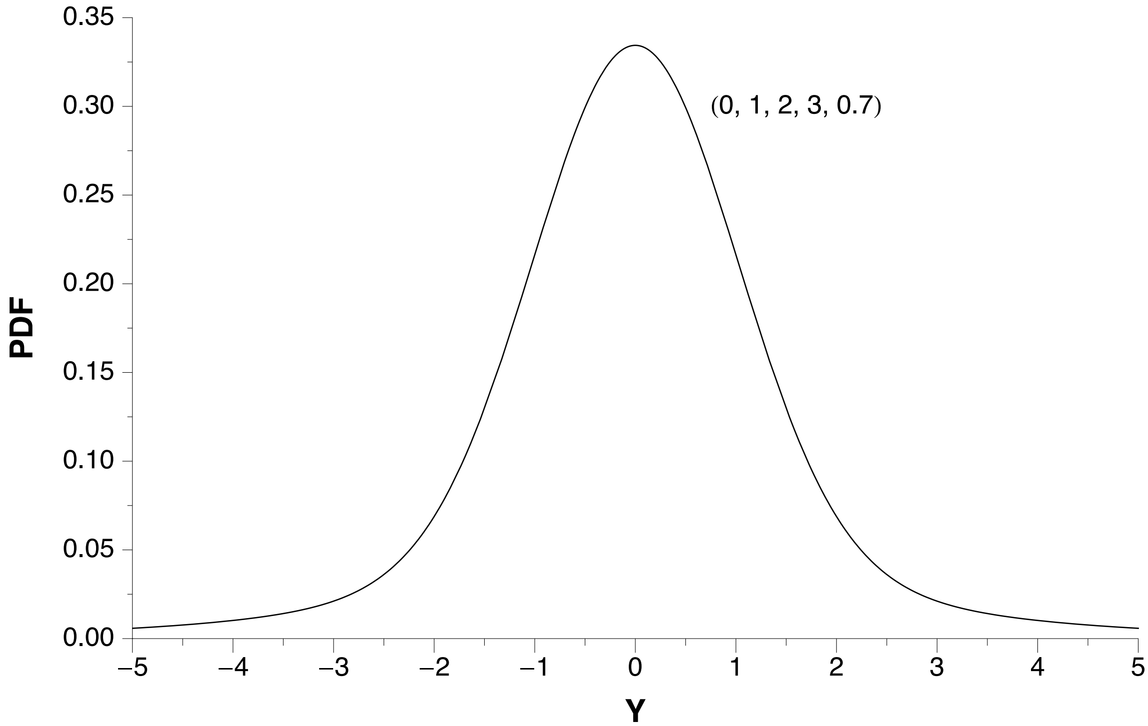
Aliases and Special Cases

1. This is a special case of the **Normal(A,B)&Normal(C,D)** distribution.

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Normal(A,B)&StudentsT(A,C,D) $\quad \mathrm{B, C, D} > 0, \quad 0 \le \mathrm{p} \le 1$



$$\text{PDF} = \frac{p}{B\sqrt{2\pi}} \exp\left[-\frac{(y-A)^2}{2\,B^2}\right] + \frac{1-p}{C\,\sqrt{D} Beta\left(\frac{D}{2}, \frac{1}{2}\right)} \left[1 + \frac{(y-A)^2}{C^2\,D}\right]^{-(D+1)/2}$$

$$\text{CDF} = p\,\Phi\left(\frac{y-A}{B}\right) + \ (1-p) \begin{cases} \dfrac{\Gamma\left((D+1)/2\right)}{2\,\sqrt{\pi}\,\Gamma\left(D/2\right)} Beta\left(\dfrac{z^2}{D+z^2}; \dfrac{D}{2}, \dfrac{1}{2}\right), & z \le 0 \\ \dfrac{1}{2} + \dfrac{\Gamma\left((D+1)/2\right)}{2\,\sqrt{\pi}\,\Gamma\left(D/2\right)} Beta\left(\dfrac{z^2}{D+z^2}; \dfrac{1}{2}, \dfrac{D}{2}\right), & z > 0 \end{cases}$$

$$\text{where } z = \frac{y-A}{C}$$

$$\text{CF} = \exp(iAt) \left[p \, \exp\left(-\frac{B^2t^2}{2}\right) + (1-p)\, \frac{(C|t|)^{D/2}\,D^{D/4}}{2^{(D-2)/2}\,\Gamma(D/2)} K_{D/2}(C\,|t|\sqrt{D})\right]$$

Parameters – A (μ): Location, B, C: Scale, D: Shape p: Weight of component #1

Moments, etc.

$$\text{Mean} = A$$

$$\text{Variance} = p\,B^2 + \frac{(1-p)\,D}{D-2}\,C^2$$

Notes

1. Moment r exists iff $D > r$.

2. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

3. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

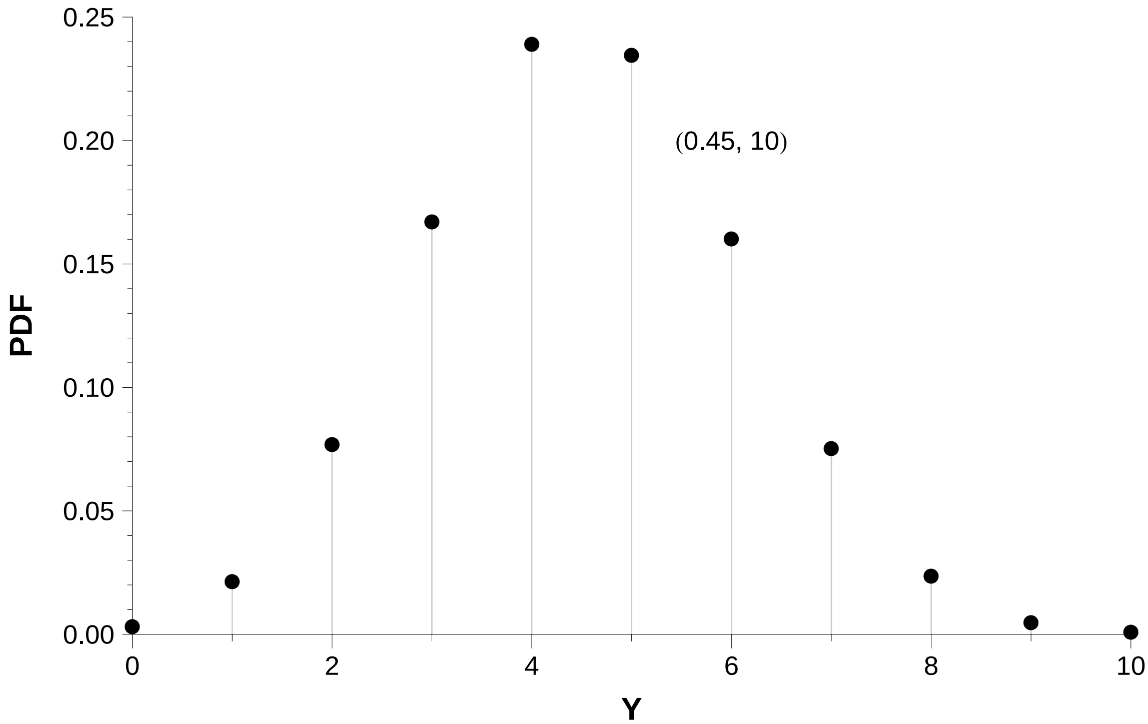
Part IV

Discrete: Standard

This section contains six discrete distributions, "discrete" meaning that their support includes only the positive integers.

Binomial(A,B)

$y = 0, 1, 2, \ldots, B > 0, \quad 0 < A < 1$



$$\text{PDF} = \binom{B}{y} A^y (1-A)^{B-y}$$

$$\text{CDF} = \binom{B}{y} (B-y) Beta(1-A; B-y, y+1)$$

$$\text{CF} = [1 - A(1 - \exp(it))]^B$$

Parameters – A (p): Prob(success), B (n): Number of Bernoulli trials

Moments, etc.

$$\text{Mean} = AB$$

$$\text{Variance} = A(1-A)B$$

$$\text{Mode} = \lfloor A(B+1) \rfloor$$

Notes

1. If $A(B+1)$ is an integer, then Mode also equals $A(B+1)-1$.

Aliases and Special Cases

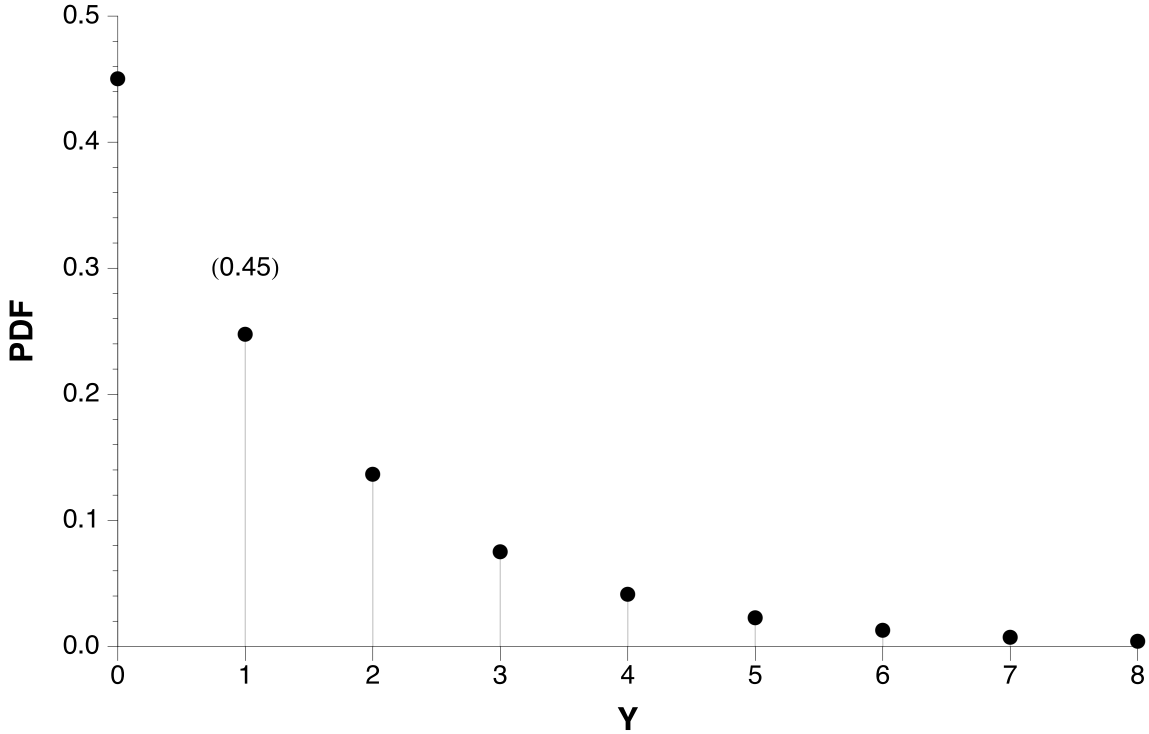
1. Binomial(A,1) is called the *Bernoulli* distribution.

Characterizations

1. In a sequence of B independent ('Bernoulli') trials with Prob(success) = A, the probability of exactly y successes is $\sim$Binomial(A,B).

Geometric(A)

$y = 0, 1, 2, \ldots, \quad 0 < A < 1$



$$\text{PDF} = A(1-A)^y$$

$$\text{CDF} = 1 - (1-A)^{y+1}$$

$$\text{CF} = \frac{A}{1-(1-A)\exp(it)}$$

Parameters – A (p): Prob(success)

Moments, etc.

$$\text{Mean} = \frac{1-A}{A}$$

$$\text{Variance} = \frac{1-A}{A^2}$$

$$\text{Mode} = 0$$

Notes

1. The Geometric distribution is the discrete analogue of the **Exponential** distribution.

2. There is an alternate form of this distribution which takes values starting at 1. See below.

Aliases and Special Cases

1. The Geometric distribution is sometimes called the *Furry* distribution.

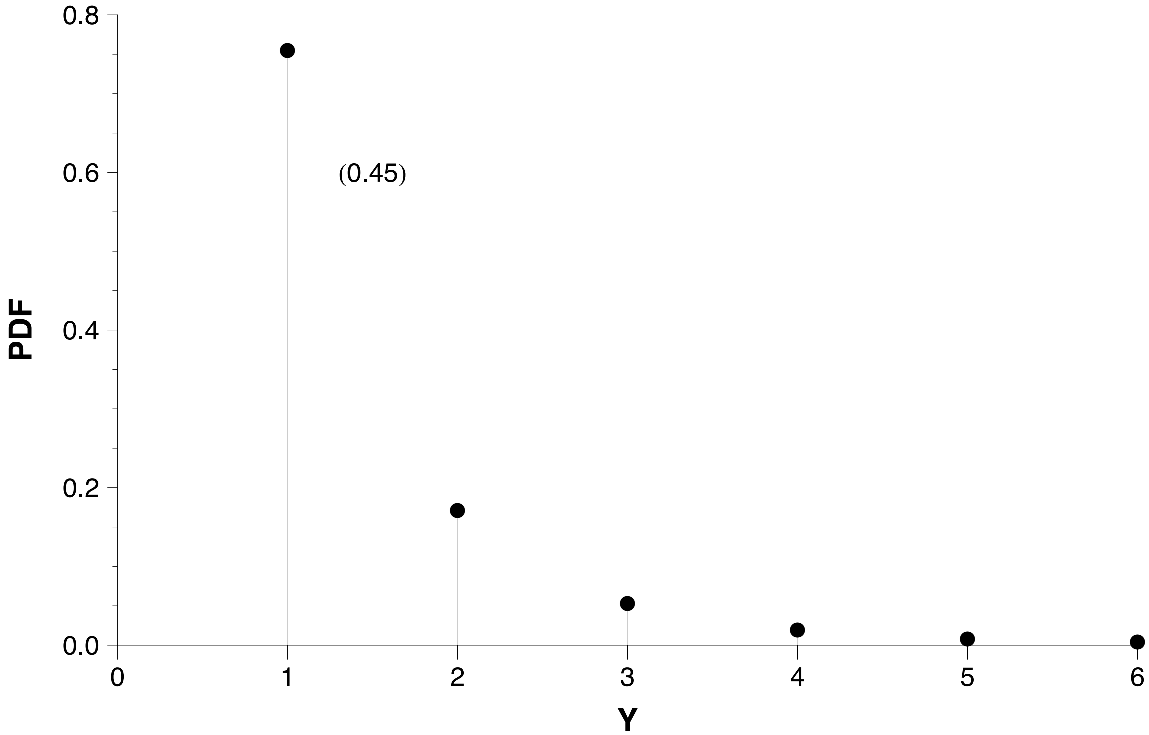
Characterizations

1. In a series of Bernoulli trials with Prob(success) = A, the number of failures preceding the first success is $\sim$Geometric(A).

2. In the alternate form cited above, it is the number of trials needed to realize the first success.

Logarithmic(A)

$y = 1, 2, 3, \ldots, \quad 0 < A < 1$



$$\text{PDF} = \frac{-A^y}{y \log(1-A)}$$

$$\text{CDF} = 1 + \frac{Beta(A, y+1, 0)}{\log(1-A)}$$

$$\text{CF} = \frac{1 - A\exp(it)}{\log(1-A)}$$

Parameters – A (p): Shape

Moments, etc.

$$\text{Mean} = \frac{A}{(A-1)\log(1-A)}$$

$$\text{Variance} = \frac{A\,(A + \log(1-A))}{(A-1)^2 \log^2(1-A)}$$

$$\text{Mode} = 1$$

Notes

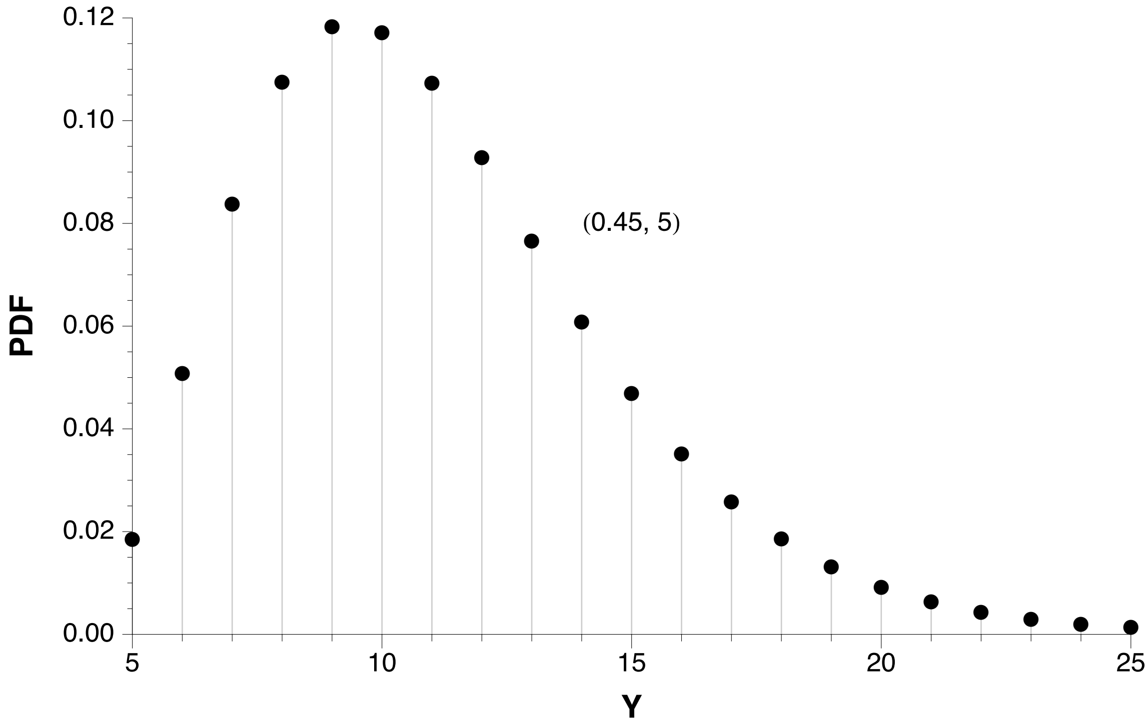
Aliases and Special Cases

1. The Logarithmic distribution is sometimes called the *Logarithmic Series* distribution or the *Log-series* distribution.

Characterizations

1. The Logarithmic distribution has been used to model the number of items of a product purchased by a buyer in a given period of time.

NegativeBinomial(A,B) $\quad 0 < \mathrm{y} = \mathrm{B}, \mathrm{B}+1, \mathrm{B}+2, \ldots, \quad 0 < \mathrm{A} < 1$



$$\mathrm{PDF} = \binom{y-1}{B-1} A^B (1-A)^{y-B}$$

$$\mathrm{CDF} = 1 - \binom{y}{B-1} \, {}_2F_1(1, y+1; y-B+2; 1-A) A^B (1-A)^{y-B+1}$$

$$\mathrm{CF} = \exp(iBt) A^B \left[1 + (A-1)\exp(it)\right]^{-B}$$

Parameters – A (p): Prob(success), B (k): a constant, target number of successes

Moments, etc.

$$\text{Mean} = \frac{B}{A}$$

$$\text{Variance} = \frac{(1-A)B}{A^2}$$

$$\text{Mode} = \left\lfloor \frac{A+B-1}{A} \right\rfloor$$

Notes

1. If $(B-1)/A$ is an integer, then Mode also equals $(B-1)/A$.

Aliases and Special Cases

1. The NegativeBinomial distribution is also known as the *Pascal* distribution.

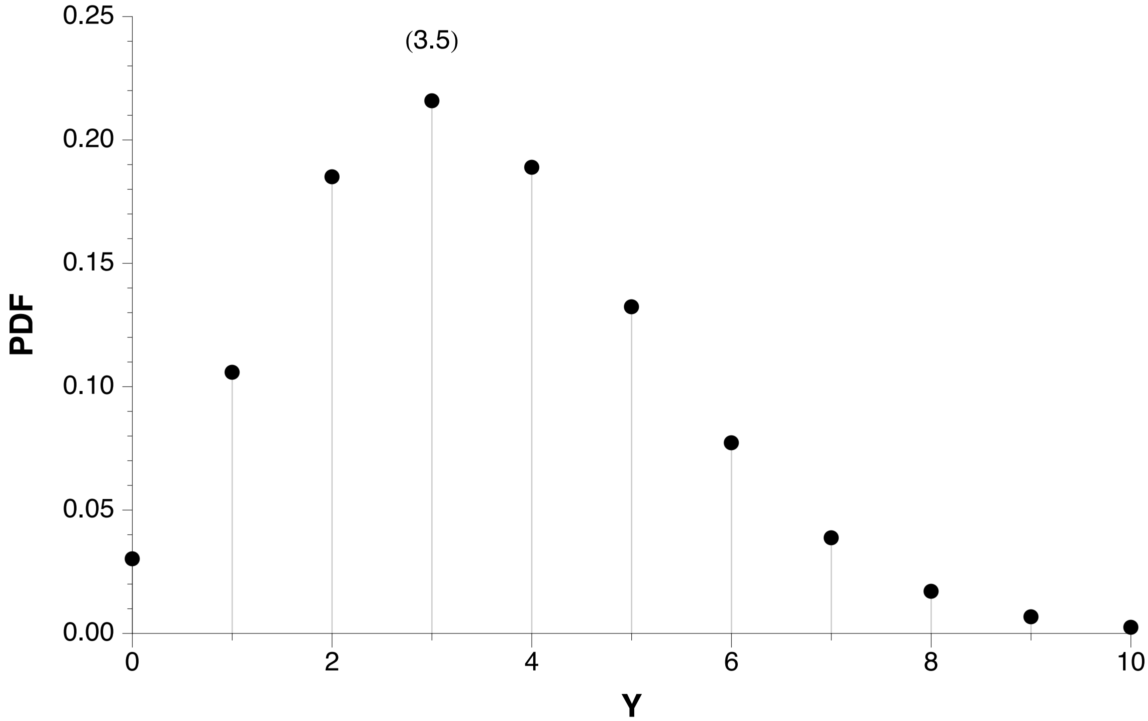
2. It is also called the *Polya* distribution.

3. If $B = 1$, the NegativeBinomial distribution becomes the **Geometric** distribution.

Characterizations

1. If Prob(success) = A, the number of Bernoulli trials required to realize the B^{th} success is $\sim$NegativeBinomial(A,B).

Poisson(A) $\qquad\qquad y = 0, 1, 2, \ldots, \quad A > 0$



$$\text{PDF} = \frac{A^y}{y\,!} \exp(-A)$$

$$\text{CDF} = \frac{\Gamma(1 + \lfloor y \rfloor, A)}{\Gamma(1 + \lfloor y \rfloor)}$$

$$\text{CF} = \exp\left[A\left(\exp(it) - 1\right)\right]$$

Parameters – A (λ): Expectation

Moments, etc.

$$\text{Mean} = \text{Variance} = A$$

$$\text{Mode} = \lfloor A \rfloor$$

Notes

1. If A is an integer, then Mode also equals $A - 1$.

Aliases and Special Cases

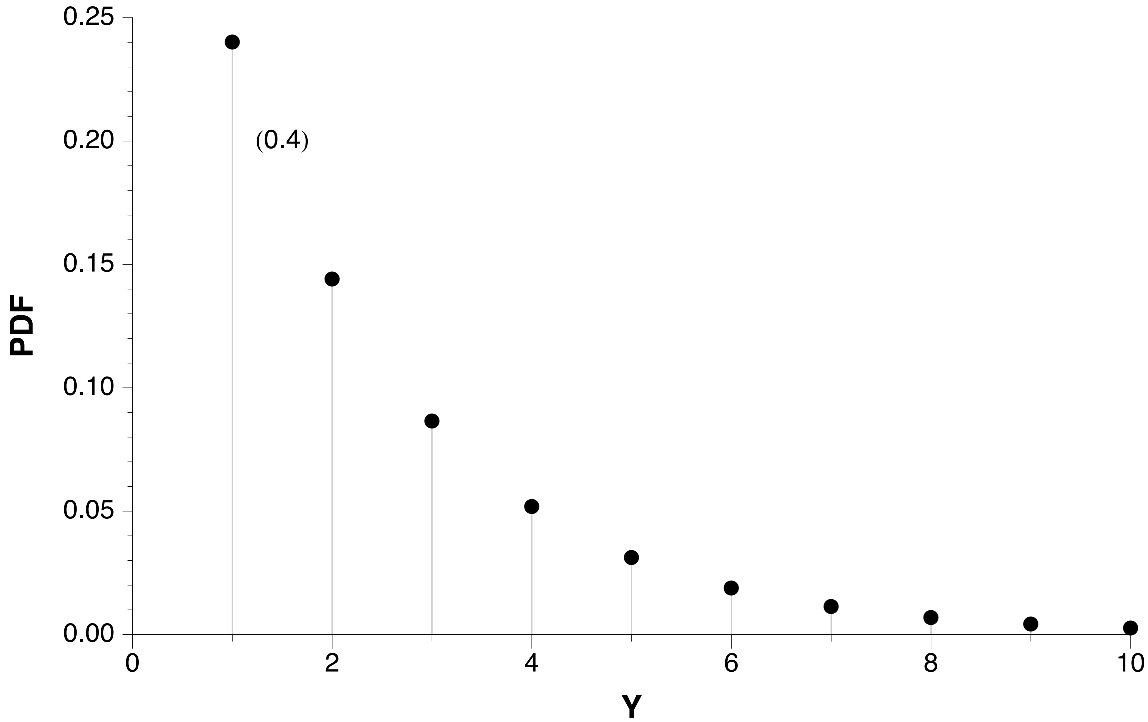
Characterizations

1. The Poisson distribution is often used to model rare events. As such, it is a good approximation to **Binomial**(A, B) when, by convention, $A\,B < 5$.

2. In queueing theory, when interarrival times are $\sim$**Exponential**, the number of arrivals in a fixed interval is $\sim$Poisson.

3. Errors in observations with integer values (i.e., miscounting) are $\sim$Poisson.

Zipf(A) $\qquad y = 1, 2, 3, \ldots, \quad A > 0$



$$\text{PDF} = \frac{y^{-(A+1)}}{\zeta(A+1)}$$

$$\text{CDF} = \frac{H_y(A+1)}{\zeta(A+1)}$$

$$\text{CF} = \frac{Li_{A+1}\left(\exp(it)\right)}{\zeta(A+1)}$$

Parameters – A: Shape

Moments, etc.

$$\text{Mean} = \frac{\zeta(A)}{\zeta(A+1)}$$

$$\text{Variance} = \frac{\zeta(A+1)\zeta(A-1) - \zeta^2(A)}{\zeta^2(A+1)}$$

$$\text{Mode} = 1$$

Notes

1. Moment r exists iff $A > r$.

Aliases and Special Cases

1. The Zipf distribution is sometimes called the *discrete* ***Pareto*** distribution.

Characterizations

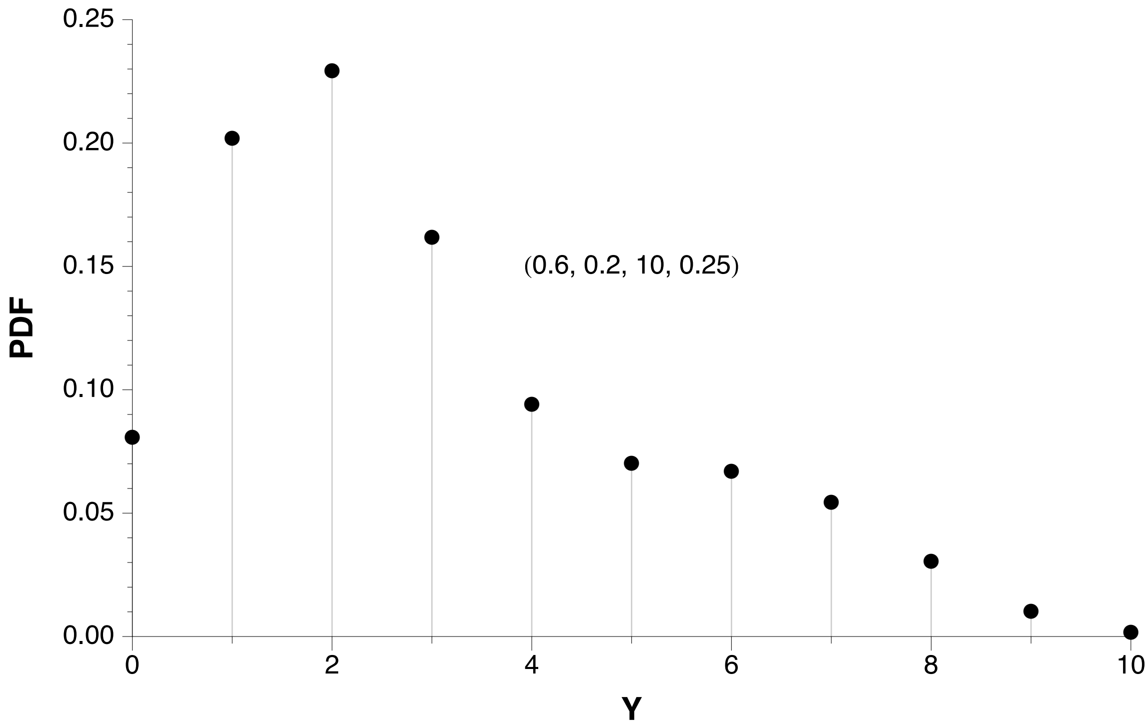
1. The Zipf distribution is used to describe the rank, y, of an ordered item as a function of the item's frequency.

Part V

Discrete: Mixtures

This section includes six binary component mixtures of discrete distributions.

Binomial(A,C)&Binomial(B,C) $\mathrm{y} = 0, 1, 2, \ldots, \mathrm{C} > 0, \quad 0 < \mathrm{A}, \mathrm{B} < 1, \quad 0 \leq \mathrm{p} \leq 1$



$$\text{PDF} = \binom{C}{y} \left[p\, A^y (1-A)^{C-y} + (1-p) B^y (1-B)^{C-y}\right]$$

$$\text{CDF} = \binom{C}{y} (C-y)[p\, Beta(1-A; C-y, y+1) + (1-p) Beta(1-B; C-y, y+1)]$$

$$\text{CF} = p\, [1 - A(1-\exp(it))]^C + (1-p)\, [1 - B(1-\exp(it))]^C$$

Parameters – A, B (p_1, p_2): Prob(success), C (n): Number of trials, p: Weight of component #1

Moments, etc.

$$\text{Mean} = (p\, A + (1-p) B)\, C$$

$$\text{Variance} = \left[p\, A(1-A) + (1-p) \left(B\,(1-B) + p\, C\,(A-B)^2\right)\right]\, C$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

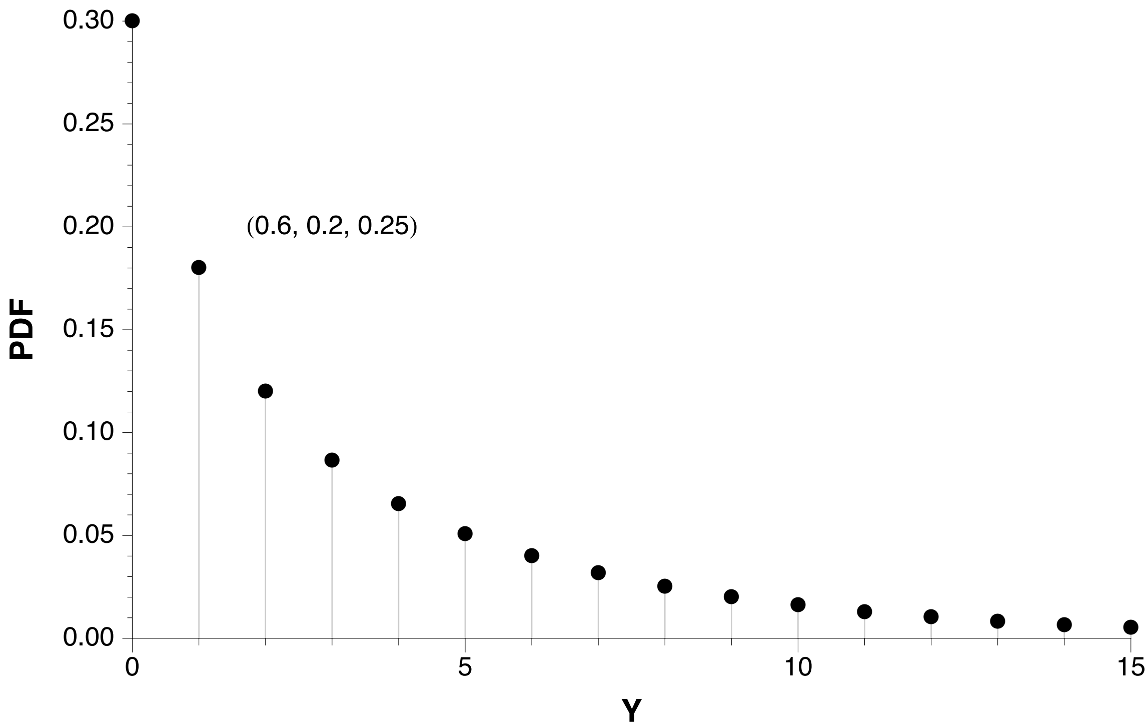
Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Geometric(A)&Geometric(B)

$y = 0, 1, 2, \ldots, \quad 0 \leq p \leq 1$



$$\text{PDF} = p\,A(1-A)^y + (1-p)B(1-B)^y$$

$$\text{CDF} = p\left[1-(1-A)^{y+1}\right] + (1-p)\left[1-(1-B)^{y+1}\right]$$

$$\text{CF} = \frac{p\,A}{1-(1-A)\exp(it)} + \frac{(1-p)B}{1-(1-B)\exp(it)}$$

Parameters – A, B (p_1, p_2): Prob(success), p: Weight of component #1

Moments, etc.

$$\text{Mean} = \frac{p}{A} + \frac{1-p}{B} - 1$$

$$\text{Variance} = \frac{A^2(1-B) - (A-2)(A-B)p\,B - (A-B)^2\,p^2}{A^2B^2}$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

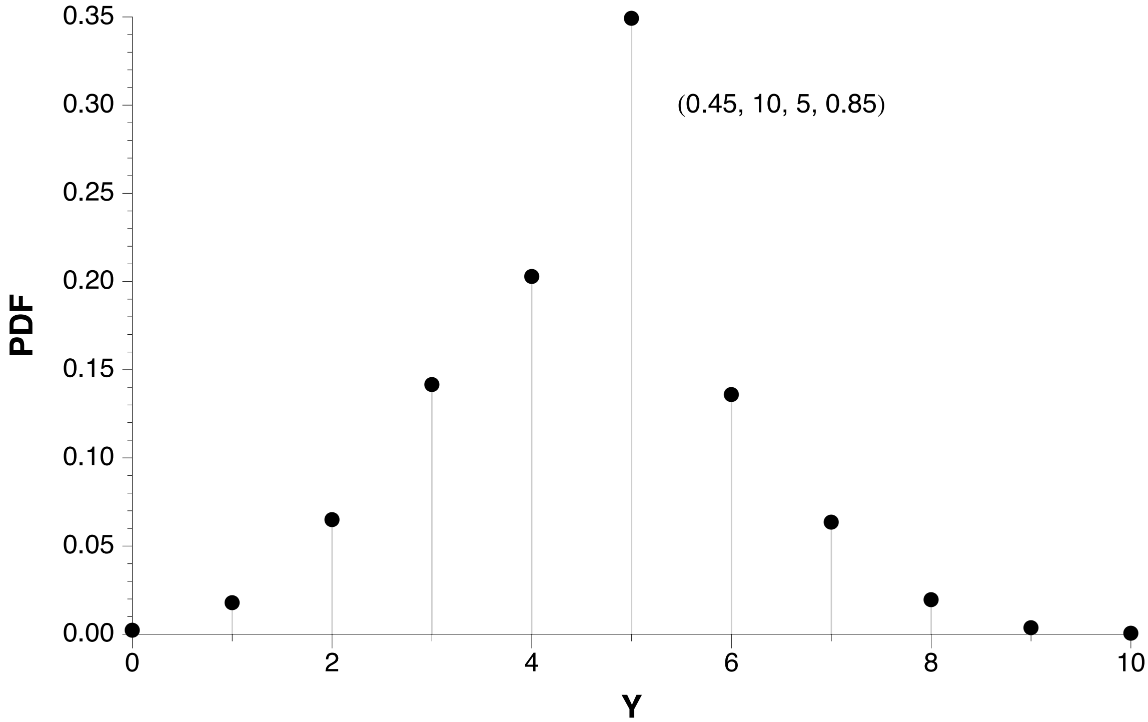
Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

InflatedBinomial(A,B,C)

$$\mathrm{y, C = 0, 1, 2, \ldots, B > 0, \quad 0 \leq p \leq 1}$$



$$\mathrm{PDF} = \begin{cases} p\dbinom{B}{y} A^y (1-A)^{B-y}, & y \neq C \\ p\dbinom{B}{y} A^y (1-A)^{B-y} + 1 - p, & y = C \end{cases}$$

$$\mathrm{CDF} = \begin{cases} p\dbinom{B}{y} (B-y) Beta(1-A; B-y, y+1), & y < C \\ p\dbinom{B}{y} (B-y) Beta(1-A; B-y, y+1) + 1 - p, & y \geq C \end{cases}$$

$$\mathrm{CF} = p\,[1 + A(\exp(it) - 1)]^B + (1-p)\exp(iCt)$$

Parameters – A: Prob(success), B: Number of trials, C: Inflated y, p: Weight of non-inflated component

Moments, etc.

$$\mathrm{Mean} = p\,AB + (1-p)C$$

$$\mathrm{Variance} = p\left[AB(1 - 2\,C(1-p)) + A^2 B\,(B\,(1-p) - 1) + C^2\,(1-p)\right]$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

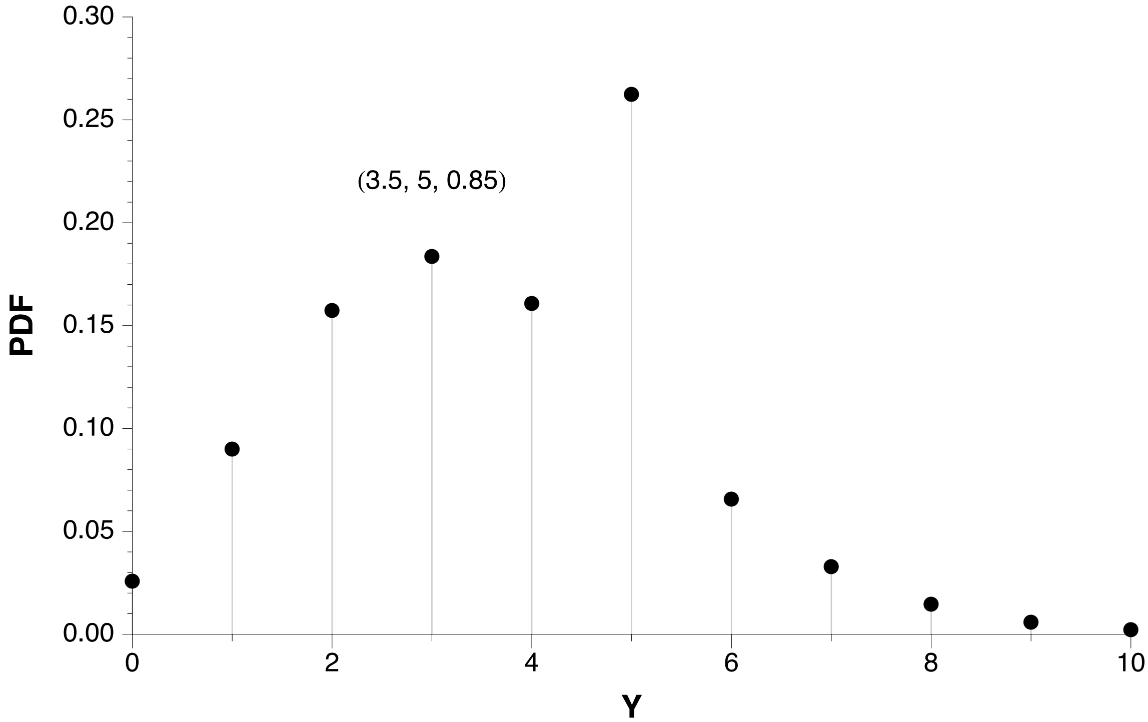
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

InflatedPoisson(A,B) $\quad \mathrm{y, B} = 0, 1, 2, \ldots, \quad \mathrm{A} > 0, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = \begin{cases} p\,\dfrac{A^y}{y\,!}\exp(-A), & y \neq B \\ p\,\dfrac{A^y}{y\,!}\exp(-A) + 1 - p, & y = B \end{cases}$$

$$\mathrm{CDF} = \begin{cases} p\,\dfrac{\Gamma(1 + \lfloor y \rfloor, A)}{\Gamma(1 + \lfloor y \rfloor)}, & y < B \\ p\,\dfrac{\Gamma(1 + \lfloor y \rfloor, A)}{\Gamma(1 + \lfloor y \rfloor)}) + 1 - p, & y \geq B \end{cases}$$

$$\mathrm{CF} = p\,\exp\left[A\left(\exp(it) - 1\right)\right] + (1 - p)\exp(iBt)$$

Parameters – A (λ): Poisson expectation, B: Inflated y, p: Weight of non-inflated component

Moments, etc.

$$\mathrm{Mean} = p\,A + (1 - p)B$$

$$\mathrm{Variance} = p\left[A + (1 - p)(A - B)^2\right]$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

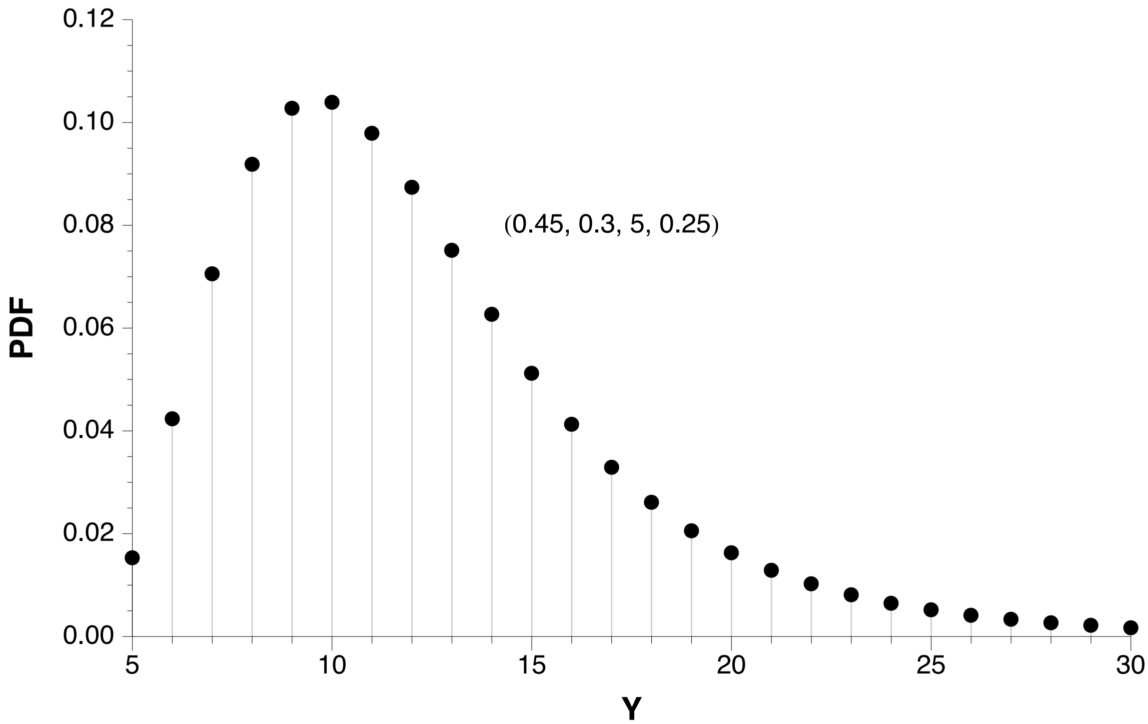
2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

NegBinomial(A,C)&NegBinomial(B,C) $0 < \mathrm{y} = \mathrm{C}, \mathrm{C}+1, \ldots, \quad 0 < \mathrm{A}, \mathrm{B} < 1, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = p \binom{y-1}{C-1} A^C (1-A)^{y-C} + (1-p) \binom{y-1}{C-1} B^C (1-B)^{y-C}$$

$$\begin{aligned}\mathrm{CDF} = & \, p \left[1 - \binom{y}{C-1} {}_2F_1(1, y+1; y-C+2; 1-A) A^C (1-A)^{y-C+1} \right] \\ & + (1-p) \left[1 - \binom{y}{C-1} {}_2F_1(1, y+1; y-C+2; 1-B) B^C (1-B)^{y-C+1} \right]\end{aligned}$$

$$\mathrm{CF} = \exp(iCt) \left[p\, A^C \left(1 + (A-1)\exp(it)\right)^{-C} + (1-p) B^C \left(1 + (B-1)\exp(it)\right)^{-C} \right]$$

Parameters – A, B: Prob(success), C: a constant, target number of successes, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = \frac{p\, C}{A} + \frac{(1-p) C}{B}$$

$$\mathrm{Variance} = \frac{C}{A^2 B^2} \left[p\, B^2 (1 + C(1-p)) - pAB(B + 2\, C(1-p)) - A^2 (1-p)(B - p\, C - 1) \right]$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

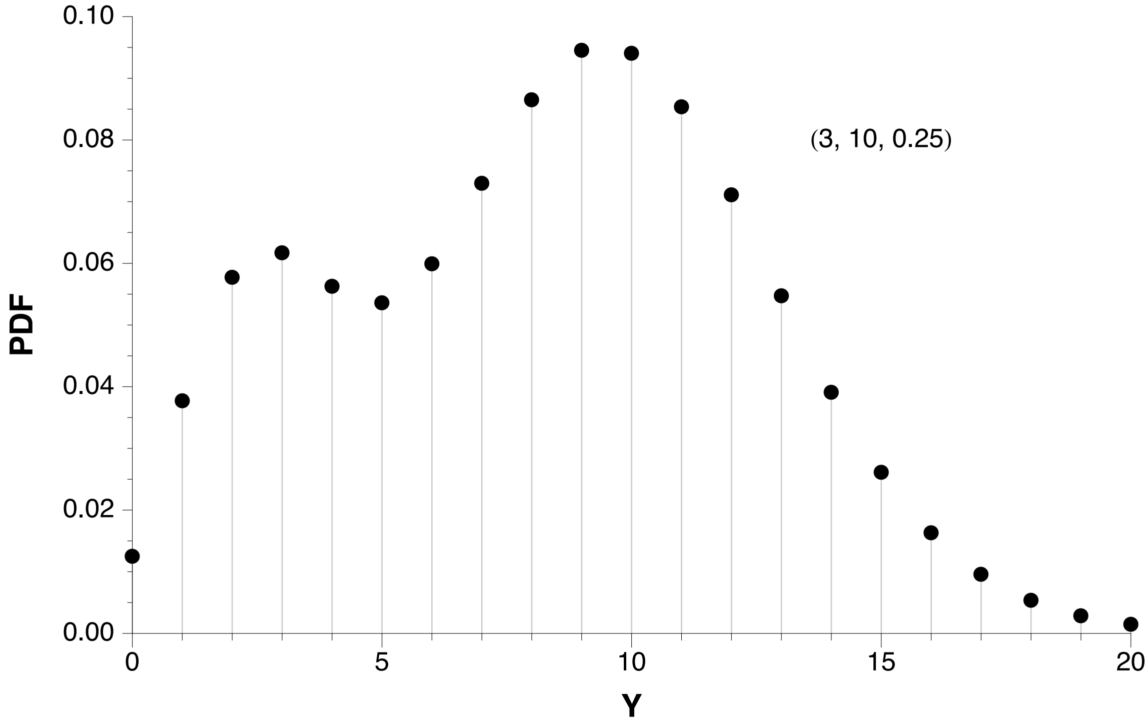
Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.

Poisson(A)&Poisson(B)

$\mathrm{y} = 0, 1, 2, \ldots, \quad 0 \leq \mathrm{p} \leq 1$



$$\mathrm{PDF} = p\,\frac{A^y}{y\,!}\exp(-A) + (1-p)\frac{B^y}{y\,!}\exp(-B)$$

$$\mathrm{CDF} = p\,\exp\left[A\left(\exp(it) - 1\right)\right] + (1-p)\exp\left[B\left(\exp(it) - 1\right)\right]$$

$$\mathrm{CF} = p\,\exp\left[A\left(\exp(it) - 1\right)\right] + (1-p)\exp\left[B\left(\exp(it) - 1\right)\right]$$

Parameters – A, B (λ_1, λ_2): Expectation, p: Weight of component #1

Moments, etc.

$$\mathrm{Mean} = p\,A + (1-p)B$$

$$\mathrm{Variance} = p\,(A-B)(1+A-B) + B - p^2(A-B)^2$$

Notes

1. Binary mixtures may require hundreds of datapoints for adequate optimization and, even then, often have unusually wide confidence intervals, especially for parameter p.

2. **Warning!** Mixtures usually have several local optima.

Aliases and Special Cases

Characterizations

1. The usual interpretation of a binary mixture is that it represents an undifferentiated composite of two populations having parameters and weights as described above.