CDF

Remark 1: Note that $Y = F_X(X) \sim \text{Unif}(0,1)$ regardless of the distribution of X as long as F_X is invertible. In the case when F_X is not invertible, modifications can be made to obtain similar result.
Remark 2: Inverting the result gives $X \sim F_X^{-1}(Y)$ where $Y \sim \text{Unif}(0,1)$. This is useful for simulating data from a given distribution with cdf F_X. Start by sampling from $\text{Unif}(0,1)$, and apply F_X^{-1} to the sample. The resulting sample will be from a distribution with cdf F_X.

Distributions

Bernoulli – Discrete

Likelihood:

Note that this form of the Bernoulli distribution pmf makes it especially easy to multiply; indeed, we could write

$$
\begin{aligned}
L_n(X_1, \ldots, X_n | \lambda) &= L(X_1|\lambda) \ldots L(X_n|\lambda) \\
&\propto (\lambda^{X_1}(1-\lambda)^{1-X_1}) \ldots \lambda^{X_n}(1-\lambda)^{1-X_n} \\
&\propto (\lambda^{X_1} \ldots \lambda^{X_n})((1-\lambda)^{1-X_1} \ldots (1-\lambda)^{1-X_n}) \\
&\propto \lambda^{\sum X_i}(1-\lambda)^{n-\sum X_i}.
\end{aligned}
$$

$$L_1\left(X_1, p\right) = p^{X_1}(1-p)^{1-X_1}.$$

Bernoulli

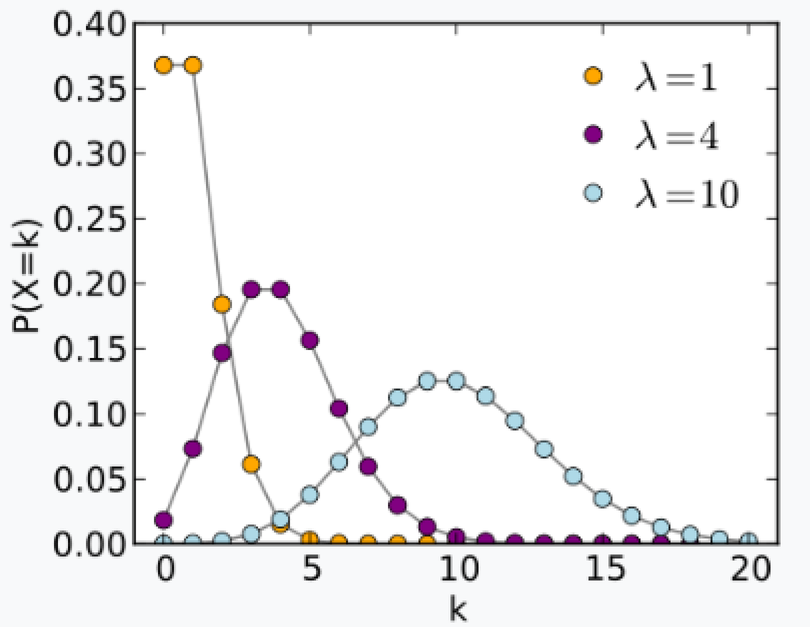
Parameters	$0 \leq p \leq 1$ $q = 1 - p$
Support	$k \in \{0, 1\}$
pmf	$\begin{cases} q = 1-p & \text{if } k = 0 \\ p & \text{if } k = 1 \end{cases}$
CDF	$\begin{cases} 0 & \text{if } k < 0 \\ 1-p & \text{if } 0 \leq k < 1 \\ 1 & \text{if } k \geq 1 \end{cases}$
Mean	p
Median	$\begin{cases} 0 & \text{if } p < 1/2 \\ [0, 1] & \text{if } p = 1/2 \\ 1 & \text{if } p > 1/2 \end{cases}$
Mode	$\begin{cases} 0 & \text{if } p < 1/2 \\ 0, 1 & \text{if } p = 1/2 \\ 1 & \text{if } p > 1/2 \end{cases}$
Variance	$p(1-p) = pq$
Skewness	$\frac{1-2p}{\sqrt{pq}}$
Ex. kurtosis	$\frac{1-6pq}{pq}$
Entropy	$-q \ln q - p \ln p$
MGF	$q + pe^t$
CF	$q + pe^{it}$
PGF	$q + pz$
Fisher information	$\frac{1}{pq}$

Poisson distribution

- Discrete
- Counts nb of events with a fixed arrival rate lamda in a fixed time/ space
- Fisher 1/lambda

Poisson

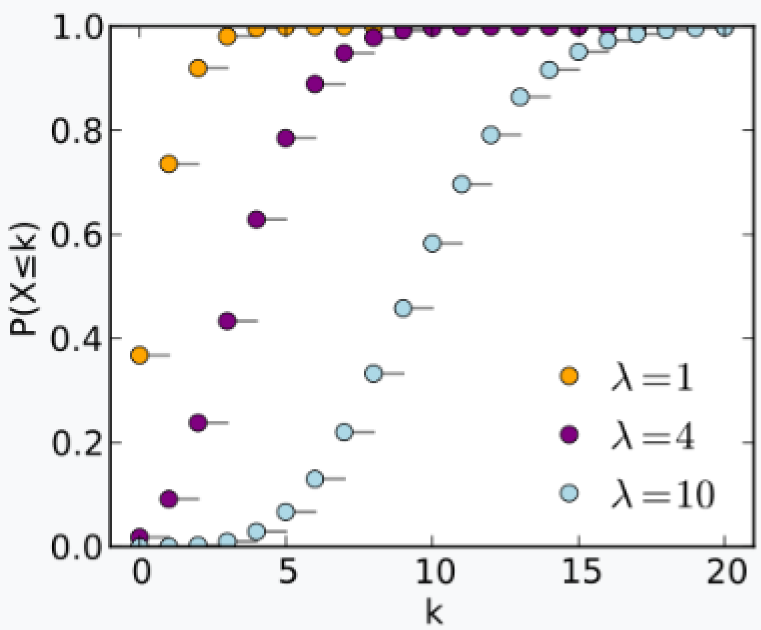
Probability mass function



The horizontal axis is the index *k*, the number of occurrences. *λ* is the expected number of occurrences, which need not be an integer. The vertical axis is the probability of *k* occurrences given *λ*. The function is defined only at integer values of *k*. The connecting lines are only guides for the eye.

-

Cumulative distribution function



The horizontal axis is the index *k*, the number of occurrences. The CDF is discontinuous at the integers of *k* and flat everywhere else because a variable that is Poisson distributed takes on only integer values.

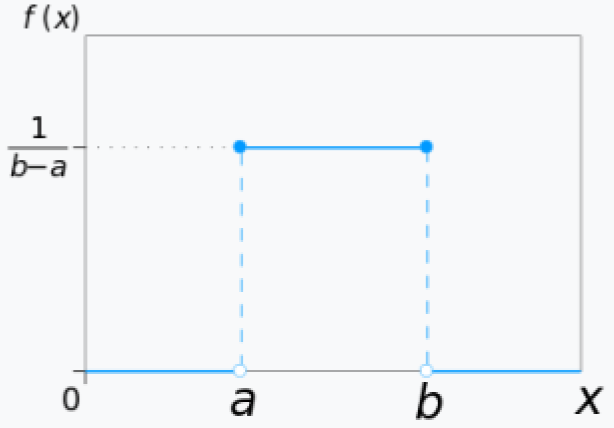
Parameters	$\lambda > 0$ (real) — rate
Support	$k \in \mathbb{N} \cup \{0\}$
pmf	$\frac{\lambda^k e^{-\lambda}}{k!}$
CDF	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor!}$, or $e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$, or $Q(\lfloor k+1 \rfloor, \lambda)$ (for $k \geq 0$, where $\Gamma(x, y)$ is the upper incomplete gamma function, $\lfloor k \rfloor$ is the floor function, and Q is the regularized gamma function)
Mean	λ
Median	$\approx \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$
Mode	$\lceil \lambda \rceil - 1, \lfloor \lambda \rfloor$
Variance	λ

Uniform

- Bounded support
- Between mix Xi and max Xi
- No Fisher information

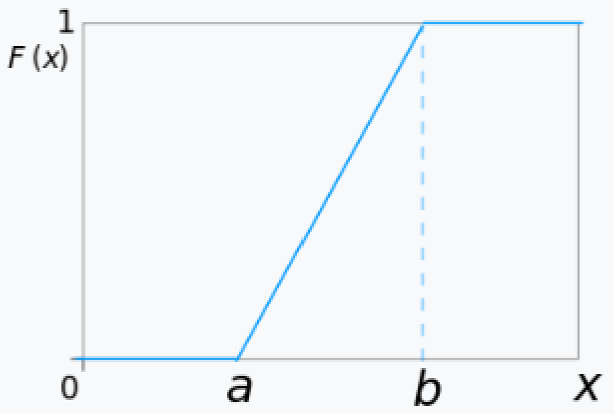
Uniform

Probability density function



Using maximum convention

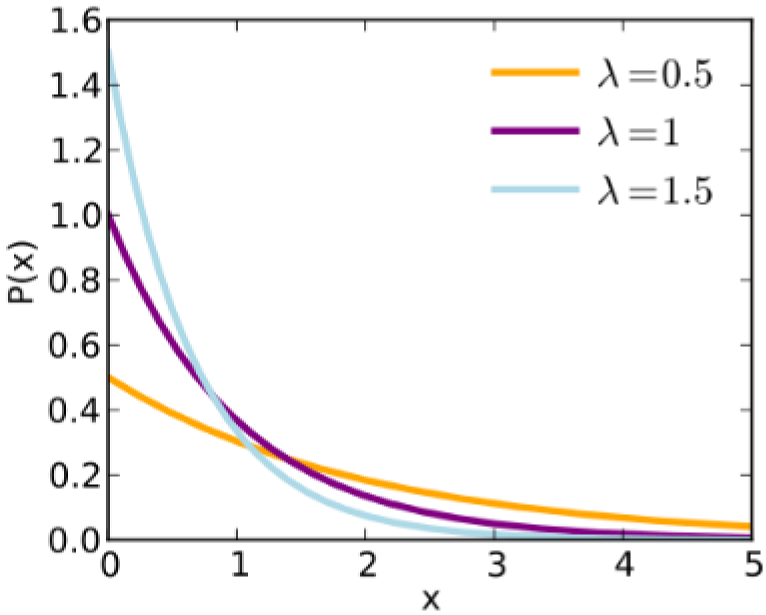
Cumulative distribution function



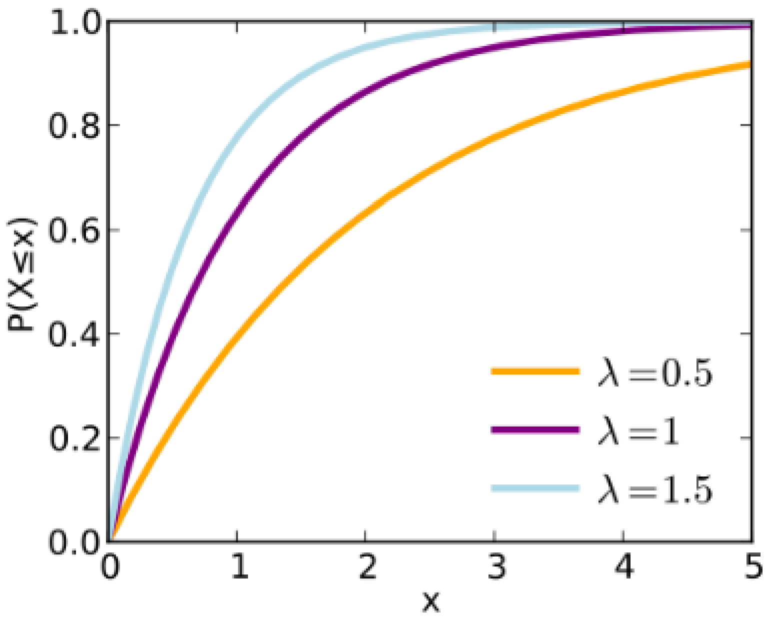
Notation	$\mathcal{U}(a,b)$ or $\mathrm{unif}(a,b)$
Parameters	$-\infty < a < b < \infty$
Support	$x \in [a,b]$
PDF	$\begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$
CDF	$\begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } x \in [a,b) \\ 1 & \text{for } x \geq b \end{cases}$
Mean	$\frac{1}{2}(a+b)$
Median	$\frac{1}{2}(a+b)$
Mode	any value in (a,b)
Variance	$\frac{1}{12}(b-a)^2$

Exponential distribution

- Continuous
- Memoryless
- Fisher = lambda^(-2)

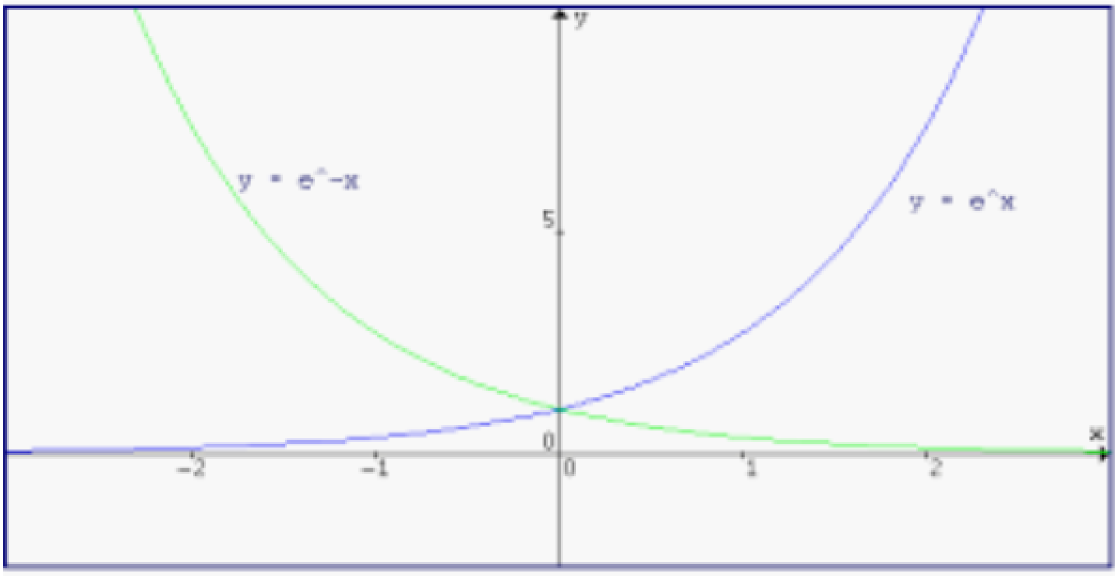


Cumulative distribution function



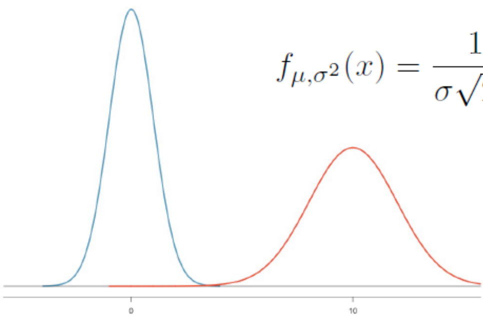
Parameters	$\lambda > 0$ rate, or inverse scale
Support	$x \in [0, \infty)$
PDF	$\lambda e^{-\lambda x}$
CDF	$1 - e^{-\lambda x}$
Quantile	$-\ln(1 - F) / \lambda$
Mean	λ^{-1} (= β)
Median	$\lambda^{-1}\ln(2)$
Mode	0
Variance	λ^{-2} (= β^2)

-



Gaussian

Gaussian density (pdf)



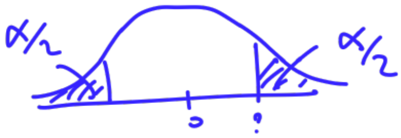
$$f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Figure 1: Two pdfs: $\mathcal{N}(0,1)$ and $\mathcal{N}(10,4)$

- Tails decay very fast (like $e^{-\frac{x^2}{2\sigma^2}}$): almost in finite interval.
- There is no closed form for their cumulative distribution function (CDF). We use tables (or computers):

$$\mathbb{P}(X \leq x) = F_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

Quantiles



Definition

Let α in $(0, 1)$. The quantile of order $1 - \alpha$ of a random variable X is the number q_α such that

$$\mathbb{P}(X \leq q_\alpha) = 1 - \alpha$$

$\alpha = .1 \Rightarrow q_\alpha$ is the 90th percentile

Let F denote the CDF of X:

- $F(q_\alpha) = 1-\alpha$
- If F is invertible, then $q_\alpha = F^{-1}(1-\alpha)$
- $\mathbb{P}(X > q_\alpha) = \alpha$
- If $X = Z \sim \mathcal{N}(0, 1)$: $\mathbb{P}(|X| > q_{\alpha/2}) = \alpha$

Some important quantiles of the $Z \sim \mathcal{N}(0, 1)$ are:

α	2.5%	5%	10%
q_α	1.96	1.65	1.28

We get that $\mathbb{P}(|Z| > 1.96) = 5\%$

Distribution of Sample Variance of Gaussian: The Chi-Squared Distribution

It is also the length of vector, distance from the center (norm of a vector)

The χ^2 distribution with d degrees of freedom is by definition the distribution of

$$Z_1^2 + Z_2^2 \ldots + Z_d^2 \qquad \text{where } Z_i \overset{iid}{\sim} \mathcal{N}(0, 1)$$

or equivalently the distribution of

$$\|\mathbf{Z}\|^2 \qquad \text{where } \mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{1}),$$

whose components are independent because the off-diagonal elements of the covariance matrix $\mathbf{1}$ are all 0.

<u>Compute quantiles in R</u>

q <- qchisq(alpha, degreeoffreedom, lower.tail = FALSE)

Student's t-distribution :

The t-test is to test the mean of a Gaussian when variance is unknown.

Works when sample size n is small

T statistic = (beta_hat – true_beta)/ std dev of beta_hat

In R

dt()

pt (q,df=degrees of freedom, lower.tail=FALSE)

qt (mass, df=degrees of freedom, lower.tail=FALSE)

rt()

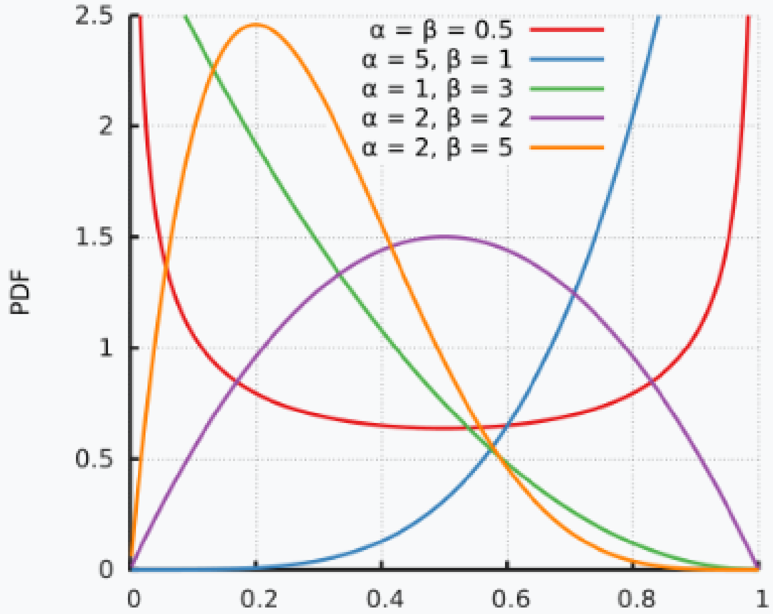
F-distribution, Fisher–Snedecor distribution, is a continuous probability distribution that arises frequently as the null distribution of a test statistic, most notably in the analysis of variance (ANOVA), e.g., F-test.

Beta distribution

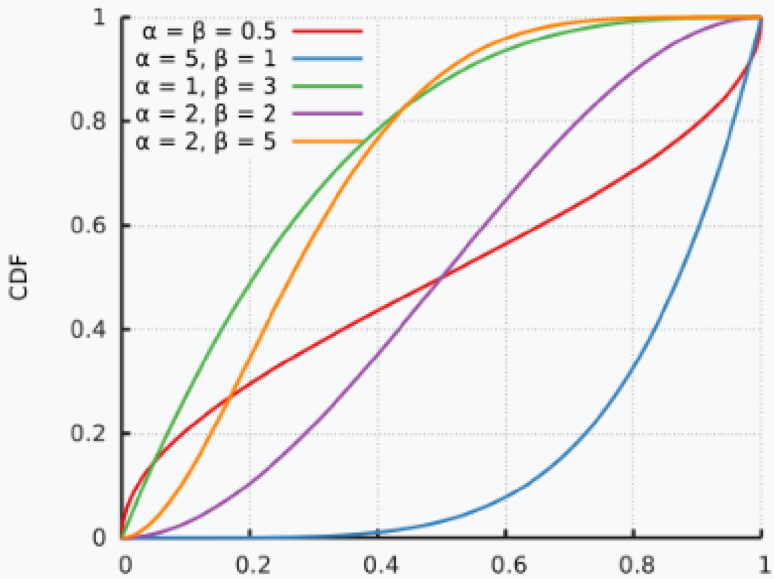
- Defined on the interval [0, 1]
- Parametrized by two positive shape parameters, denoted by α and β, that appear as exponents of the random variable and control the shape of the distribution.
- Special case of the Dirichlet distribution.

Beta

Probability density function



Cumulative distribution function



Notation	Beta(α, β)
Parameters	$\alpha > 0$ shape (real) $\beta > 0$ shape (real)
Support	$x \in [0, 1]$ or $x \in (0, 1)$
PDF	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$ where $\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

Mean	$\operatorname{E}[X] = \frac{\alpha}{\alpha+\beta}$ $\operatorname{E}[\ln X] = \psi(\alpha) - \psi(\alpha+\beta)$ $\operatorname{E}[X \ln X] = \frac{\alpha}{\alpha+\beta}\,[\psi(\alpha+1) - \psi(\alpha+\beta+1)]$ (see digamma function and see section: Geometric mean)
Median	$I_{\frac{1}{2}}^{[-1]}(\alpha,\beta) \text{ (in general)}$ $\approx \frac{\alpha-\frac{1}{3}}{\alpha+\beta-\frac{2}{3}} \text{ for } \alpha,\beta>1$
Mode	$\frac{\alpha-1}{\alpha+\beta-2}$ for α, $\beta > 1$ any value in $(0,1)$ for α, $\beta = 1$ {0, 1} (bimodal) for α, $\beta < 1$ 0 for $\alpha \le 1$, $\beta > 1$ 1 for $\alpha > 1$, $\beta \le 1$
Variance	$\operatorname{var}[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ $\operatorname{var}[\ln X] = \psi_1(\alpha) - \psi_1(\alpha+\beta)$ (see trigamma function and see section: Geometric variance)

Skewness	$\frac{2\,(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$
Ex. kurtosis	$\frac{6[(\alpha-\beta)^2(\alpha+\beta+1)-\alpha\beta(\alpha+\beta+2)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}$
Entropy	$\ln \mathrm{B}(\alpha,\beta)-(\alpha-1)\psi(\alpha)-(\beta-1)\psi(\beta) +(\alpha+\beta-2)\psi(\alpha+\beta)$
MGF	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
CF	${}_1F_1(\alpha;\alpha+\beta;i\,t)$ (see Confluent hypergeometric function)
Fisher information	$\begin{bmatrix}\operatorname{var}[\ln X] & \operatorname{cov}[\ln X,\ln(1-X)]\\ \operatorname{cov}[\ln X,\ln(1-X)] & \operatorname{var}[\ln(1-X)]\end{bmatrix}$ see section: Fisher information matrix

Gamma distribution

Continuous probability distribution

Exponential distribution, Erlang distribution, Chi-squared are special cases of gamma.

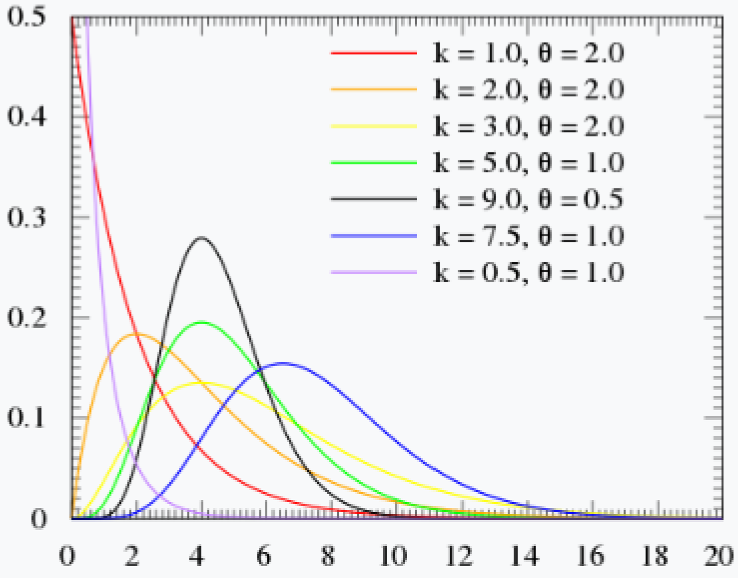
Parametrization in common use:

- a shape parameter $\alpha = k > 0$
- an inverse scale parameter $\beta = 1/\theta$, called a rate parameter, > 0.

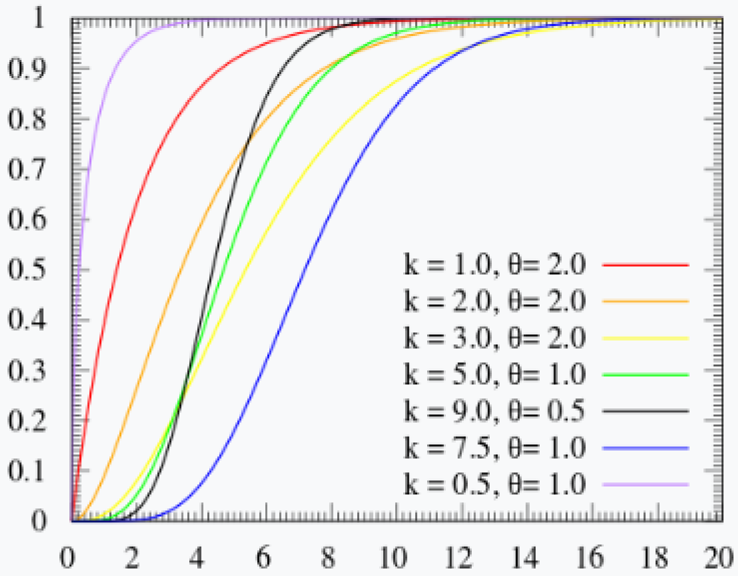
No Fisher information?

Gamma

Probability density function



Cumulative distribution function



Parameters	• $k > 0$ shape • $\theta > 0$ scale	• $\alpha > 0$ shape • $\beta > 0$ rate
Support	$x \in (0, \infty)$	$x \in (0, \infty)$
PDF	$\frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$
CDF	$\frac{1}{\Gamma(k)} \gamma\left(k, \frac{x}{\theta}\right)$	$\frac{1}{\Gamma(\alpha)} \gamma(\alpha, \beta x)$

Mean	$\operatorname{E}[X] = k\theta$	$\operatorname{E}[X] = \frac{\alpha}{\beta}$
Median	No simple closed form	No simple closed form
Mode	$(k-1)\theta \text{ for } k \geq 1$	$\frac{\alpha-1}{\beta} \text{ for } \alpha \geq 1$
Variance	$\operatorname{Var}(X) = k\theta^2$	$\operatorname{Var}(X) = \frac{\alpha}{\beta^2}$
Skewness	$\frac{2}{\sqrt{k}}$	$\frac{2}{\sqrt{\alpha}}$
Excess kurtosis	$\frac{6}{k}$	$\frac{6}{\alpha}$
Entropy	$k + \ln\theta + \ln\Gamma(k) + (1-k)\psi(k)$	$\alpha - \ln\beta + \ln\Gamma(\alpha) + (1-\alpha)\psi(\alpha)$