WIKIPEDIA

Gamma distribution

In probability theory and statistics, the **gamma distribution** is a two-parameter family of continuous probability distributions. The exponential distribution, Erlang distribution, and chi-square distribution are special cases of the gamma distribution. There are two different parameterizations in common use:

1. With a shape parameter k and a scale parameter θ.
2. With a shape parameter $\alpha = k$ and an inverse scale parameter $\beta = 1/\theta$, called a rate parameter.

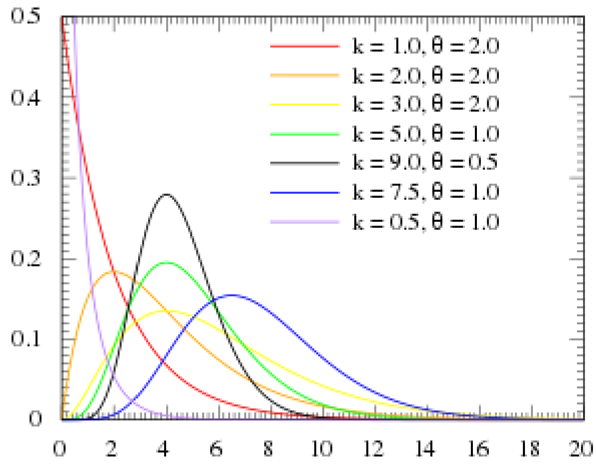
In each of these forms, both parameters are positive real numbers.

The gamma distribution is the maximum entropy probability distribution (both with respect to a uniform base measure and with respect to a $1/x$ base measure) for a random variable X for which $\mathbf{E}[X] = k\theta = \alpha/\beta$ is fixed and greater than zero, and $\mathbf{E}[\ln(X)] = \psi(k) + \ln(\theta) = \psi(\alpha) - \ln(\beta)$ is fixed (ψ is the digamma function).[1]

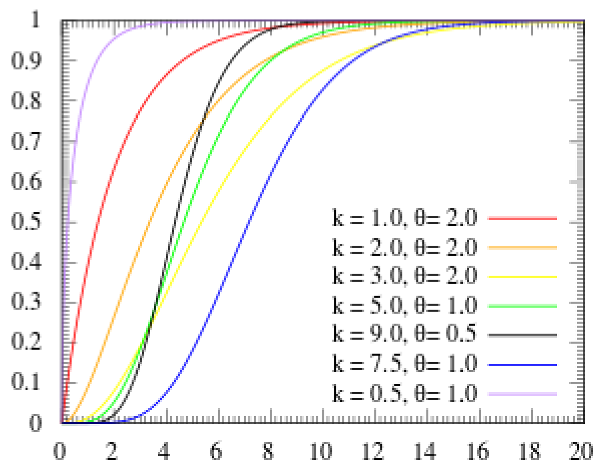
Contents

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	$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$
CDF	$F(x) = \frac{1}{\Gamma(k)} \gamma\left(k, \frac{x}{\theta}\right)$	$F(x) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \beta x)$
Mean	$k\theta$	$\frac{\alpha}{\beta}$
Median	No simple closed form	No simple closed form
Mode	$(k-1)\theta$ for $k \geq 1$, 0 for $k < 1$	$\frac{\alpha - 1}{\beta}$ for $\alpha \geq 1$, 0 for $\alpha < 1$
Variance	$k\theta^2$	$\frac{\alpha}{\beta^2}$

Skewness	$\frac{2}{\sqrt{k}}$	$\frac{2}{\sqrt{\alpha}}$
Ex. kurtosis	$\frac{6}{k}$	$\frac{6}{\alpha}$
Entropy	$k + \ln\theta + \ln\Gamma(k) + (1-k)\psi(k)$	$\alpha - \ln\beta + \ln\Gamma(\alpha) + (1-\alpha)\psi(\alpha)$
MGF	$(1-\theta t)^{-k}$ for $t < \frac{1}{\theta}$	$\left(1-\frac{t}{\beta}\right)^{-\alpha}$ for $t<\beta$
CF	$(1-\theta it)^{-k}$	$\left(1-\frac{it}{\beta}\right)^{-\alpha}$
Method of Moments	$k = \frac{E[X]^2}{V[X]}$ $\theta = \frac{V[X]}{E[X]}$	$\alpha = \frac{E[X]^2}{V[X]}$ $\beta = \frac{E[X]}{V[X]}$

Definitions

The parameterization with k and θ appears to be more common in econometrics and certain other applied fields, where for example the gamma distribution is frequently used to model waiting times. For instance, in life testing, the waiting time until death is a random variable that is frequently modeled with a gamma distribution. See Hogg and Craig[2] for an explicit motivation.

The parameterization with α and β is more common in Bayesian statistics, where the gamma distribution is used as a conjugate prior distribution for various types of inverse scale (rate) parameters, such as the λ of an exponential distribution or a Poisson distribution[3] – or for that matter, the β of the gamma distribution itself. The closely related inverse-gamma distribution is used as a conjugate prior for scale parameters, such as the variance of a normal distribution.

If k is a positive integer, then the distribution represents an Erlang distribution; i.e., the sum of k independent exponentially distributed random variables, each of which has a mean of θ.

Characterization using shape α and rate β

The gamma distribution can be parameterized in terms of a shape parameter $\alpha = k$ and an inverse scale parameter $\beta = 1/\theta$, called a rate parameter. A random variable X that is gamma-distributed with shape α and rate β is denoted

$$X \sim \Gamma(\alpha, \beta) \equiv \mathrm{Gamma}(\alpha, \beta)$$

The corresponding probability density function in the shape-rate parametrization is

$$f(x;\alpha,\beta) = \frac{x^{\alpha-1}e^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)} \quad \text{for } x > 0 \quad \alpha, \beta > 0,$$

where $\Gamma(\alpha)$ is the gamma function. For all positive integers, $\Gamma(\alpha) = (\alpha - 1)!$.

The cumulative distribution function is the regularized gamma function:

$$F(x;\alpha,\beta) = \int_0^x f(u;\alpha,\beta)\,du = \frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)},$$

where $\gamma(\alpha, \beta x)$ is the lower incomplete gamma function.

If α is a positive integer (i.e., the distribution is an Erlang distribution), the cumulative distribution function has the following series expansion:[4]

$$F(x;\alpha,\beta) = 1 - \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!} e^{-\beta x} = e^{-\beta x} \sum_{i=\alpha}^{\infty} \frac{(\beta x)^i}{i!}.$$

Characterization using shape *k* and scale *θ*

A random variable X that is gamma-distributed with shape k and scale θ is denoted by

$$X \sim \Gamma(k,\theta) \equiv \mathrm{Gamma}(k,\theta)$$

The probability density function using the shape-scale parametrization is

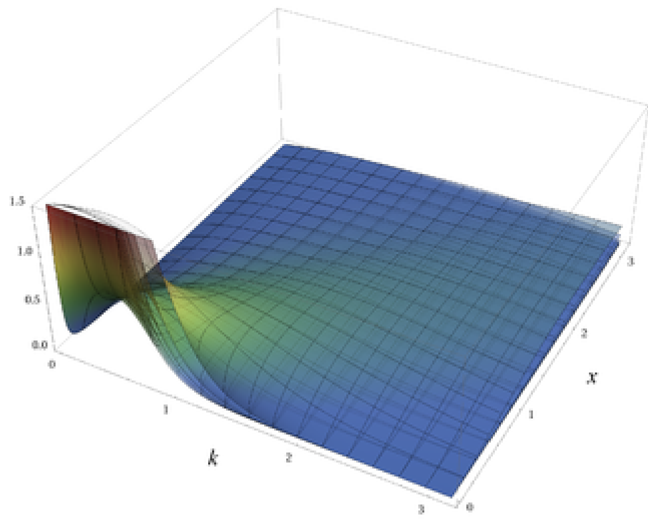


Illustration of the gamma PDF for parameter values over k and x with θ set to 1, 2, 3, 4, 5 and 6. One can see each θ layer by itself here [2] (https://commons.wikimedia.org/wiki/File:Gamma-PDF-3D-by-k.png) as well as by k [3] (https://commons.wikimedia.org/wiki/File:Gamma-PDF-3D-by-Theta.png) and x. [4] (https://commons.wikimedia.org/wiki/File:Gamma-PDF-3D-by-x.png).

$$f(x;k,\theta) = \frac{x^{k-1}e^{-x/\theta}}{\theta^k \Gamma(k)} \quad \text{for } x > 0 \text{ and } k, \theta > 0.$$

Here $\Gamma(k)$ is the gamma function evaluated at k.

The cumulative distribution function is the regularized gamma function:

$$F(x;k,\theta) = \int_0^x f(u;k,\theta)\,du = \frac{\gamma\left(k, \frac{x}{\theta}\right)}{\Gamma(k)},$$

where $\gamma\left(k, \frac{x}{\theta}\right)$ is the lower incomplete gamma function.

It can also be expressed as follows, if k is a positive integer (i.e., the distribution is an Erlang distribution):[4]

$$F(x;k,\theta) = 1 - \sum_{i=0}^{k-1} \frac{1}{i!}\left(\frac{x}{\theta}\right)^i e^{-x/\theta} = e^{-x/\theta} \sum_{i=k}^{\infty} \frac{1}{i!}\left(\frac{x}{\theta}\right)^i.$$

Both parametrizations are common because either can be more convenient depending on the situation.

Properties

Skewness

The skewness of the gamma distribution only depends on its shape parameter, k, and it is equal to $2/\sqrt{k}$.

Higher moments

The nth raw moment is given by:

$$\mathrm{E}[X^n] = \theta^n \frac{\Gamma(n+k)}{\Gamma(k)}.$$

Median approximations and bounds

Unlike the mode and the mean, which have readily calculable formulas based on the parameters, the median does not have a closed-form equation. The median for this distribution is the value ν such that

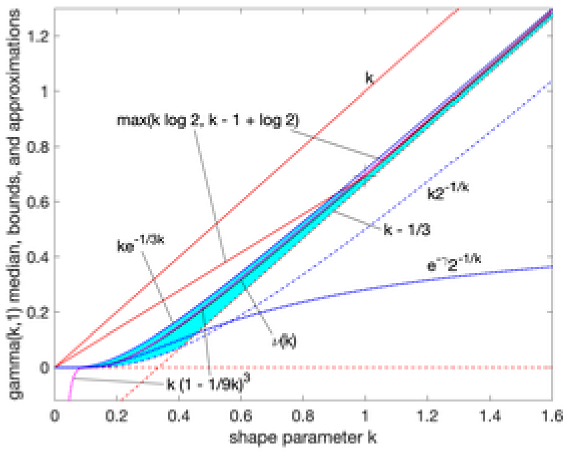
$$\frac{1}{\Gamma(k)\theta^k} \int_0^{\nu} x^{k-1} e^{-x/\theta} dx = \frac{1}{2}.$$

A rigorous treatment of the problem of determining an asymptotic expansion and bounds for the median of the gamma distribution was handled first by Chen and Rubin, who proved that (for $\theta = 1$)

$$k - \frac{1}{3} < \nu(k) < k,$$

where $\mu(k) = k$ is the mean and $\nu(k)$ is the median of the $\mathrm{Gamma}(k, 1)$ distribution.[5] For other values of the scale parameter, the mean scales to $\mu = k\theta$, and the median bounds and approximations would be similarly scaled by θ.

K. P. Choi found the first five terms in a Laurent series asymptotic approximation of the median by comparing the median to Ramanujan's θ function.[6] Berg and Pedersen found more terms:[7]

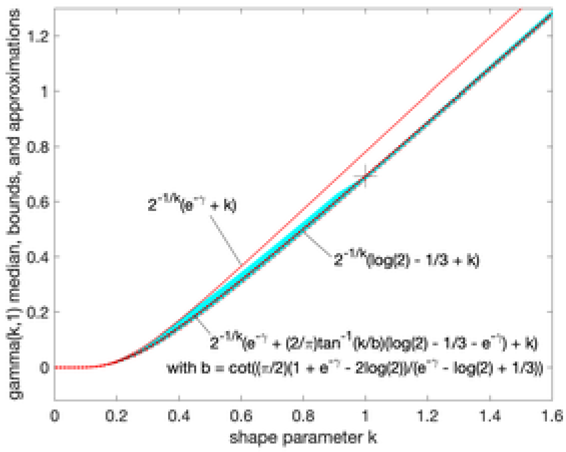


Bounds and asymptotic approximations to the median of the gamma distribution. The cyan colored region indicates the large gap between published lower and upper bounds.

$$\nu(k) = k - \frac{1}{3} + \frac{8}{405k} + \frac{184}{25515k^2} + \frac{2248}{3444525k^3} - \frac{19006408}{15345358875k^4} - O\left(\frac{1}{k^5}\right) + \cdots$$

Partial sums of these series are good approximations for high enough k; they are not plotted in the figure, which is focused on the low-k region that is less well approximated.

Berg and Pedersen also proved many properties of the median, showed that it is a convex function of k,[8] and that the asymptotic behavior near $k = 0$ is $\nu(k) \approx e^{-\gamma} 2^{-1/k}$ (where γ is the Euler–Mascheroni constant), and that for all $k > 0$ the median is bounded by $k2^{-1/k} < \nu(k) < ke^{-1/3k}$.[7]



Two gamma distribution median asymptotes which are conjectured to be bounds (upper solid red and lower dashed red), of the from $\nu(k) \approx 2^{-1/k}(A + k)$, and an interpolation between them that makes an approximation (dotted red) that is exact at k = 1 and has maximum relative error of about 0.6%. The cyan shaded region is the remaining gap between upper and lower bounds (or conjectured bounds) including these new (as of 2021) conjectured bounds and the proven bounds in the previous figure.

A closer linear upper bound, for $k \geq 1$ only, was provided in 2021 by Gaunt and Merkle,[9] relying on the Berg and Pedersen result that the slope of $\nu(k)$ is everywhere less than 1:

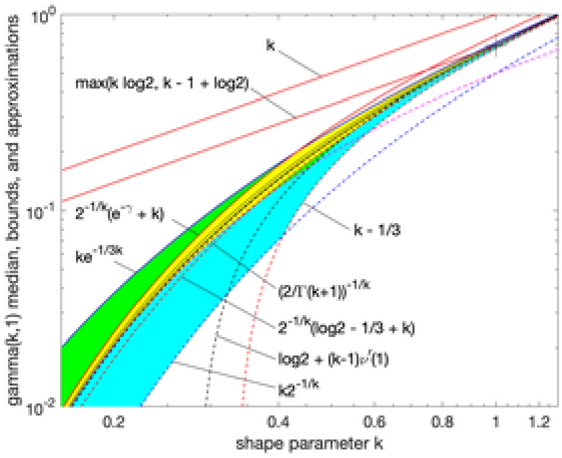
> $\nu(k) \leq k - 1 + \log 2$ for $k \geq 1$ (with equality at $k = 1$)

which can be extended to a bound for all $k > 0$ by taking the max with the chord shown in the figure, since the median was proved convex.[8]

An approximation to the median that is asymptotically accurate at high k and reasonable down to $k = 0.5$ or a bit lower follows from the Wilson–Hilferty transformation:

$$\nu(k) = k\left(1 - \frac{1}{9k}\right)^3$$

which goes negative for $k < 1/9$.



Log–log plot of upper (solid) and lower (dashed) bounds to the median of a gamma distribution, and the gaps between them. The green, yellow, and cyan regions represent the gap before the Lyon 2021 paper. The green and yellow narrow that gap with the lower bounds that Lyon proved. The yellow is further narrowed by Lyon's conjectured bounds. Mostly within the yellow, closed-form rational-function-interpolated bounds are plotted, along with the numerically calculated value of the median (dotted). Tighter interpolated bounds exist but are not plotted, as they would not be resolved at this scale.

In 2021, Lyon proposed several closed-form approximations of the form $\nu(k) \approx 2^{-1/k}(A + Bk)$. He conjectured closed-form values of A and B for which this approximation is an asymptotically tight upper or lower bound for all $k > 0$. In particular:[10]

$\nu_{L\infty}(k) = 2^{-1/k}(\log 2 - \frac{1}{3} + k)$ is a lower bound, asymptotically tight as $k \to \infty$

$\nu_U(k) = 2^{-1/k}(e^{-\gamma} + k)$ is an upper bound, asymptotically tight as $k \to 0$

Lyon also derived two other lower bounds that are not closed-form expressions, including this one based on solving the integral expression substituting 1 for e^{-x}:

$\nu(k) > \left(\frac{2}{\Gamma(k+1)}\right)^{-1/k}$ (approaching equality as $k \to 0$)

and the tangent line at $k = 1$ where the derivative was found to be $\nu'(1) \approx 0.9680448$:

$\nu(k) \geq \nu(1) + (k-1)\nu'(1)$ (with equality at $k = 1$)
$\nu(k) \geq \log(2) + (k-1)(\gamma - 2\,\mathrm{Ei}(-\log 2) - \log\log 2)$

where Ei is the exponential integral.[10]

Additionally, he showed that interpolations between bounds can provide excellent approximations or tighter bounds to the median, including an approximation that is exact at $k = 1$ (where $\nu(1) = \log 2$) and has a maximum relative error less than 0.6%. Interpolated approximations and bounds are all of the form

$$\nu(k) \approx \tilde{g}(k)\nu_{L\infty}(k) + (1 - \tilde{g}(k))\nu_U(k)$$

where $\tilde{g}$ is an interpolating function running monotonically from 0 at low k to 1 at high k, approximating an ideal, or exact, interpolator $g(k)$:

$$g(k) = \frac{\nu_U(k) - \nu(k)}{\nu_U(k) - \nu_{L\infty}(k)}$$

For the simplest interpolating function considered, a first-order rational function

$$\tilde{g}_1(k) = \frac{k}{b_0 + k}$$

the tightest lower bound has

$$b_0 = \frac{\frac{8}{405} + e^{-\gamma}\log 2 - \frac{\log^2 2}{2}}{e^{-\gamma} - \log 2 + \frac{1}{3}} - \log 2 \approx 0.143472$$

and the tightest upper bound has

$$b_0 = \frac{e^{-\gamma} - \log 2 + \frac{1}{3}}{1 - \frac{e^{-\gamma}\pi^2}{12}} \approx 0.374654$$

The interpolated bounds are plotted (mostly inside the yellow region) in the log–log plot shown. Even tighter bounds are available using different interpolating functions, but not usually with closed-form parameters like these.[10]

Summation

If X_i has a Gamma(k_i, θ) distribution for i = 1, 2, ..., N (i.e., all distributions have the same scale parameter θ), then

$$\sum_{i=1}^{N} X_i \sim \mathrm{Gamma}\left(\sum_{i=1}^{N} k_i, \theta\right)$$

provided all X_i are independent.

For the cases where the X_i are independent but have different scale parameters see Mathai [11] or Moschopoulos.[12]

The gamma distribution exhibits infinite divisibility.

Scaling

If

$$X \sim \mathrm{Gamma}(k, \theta),$$

then, for any $c > 0$,

$$cX \sim \mathrm{Gamma}(k, c\,\theta),$$ by moment generating functions,

or equivalently, if

$$X \sim \mathrm{Gamma}\left(\alpha, \beta\right)$$ (shape-rate parameterization)

$$cX \sim \mathrm{Gamma}\left(\alpha, \frac{\beta}{c}\right),$$

Indeed, we know that if X is an exponential r.v. with rate λ then cX is an exponential r.v. with rate λ/c; the same thing is valid with Gamma variates (and this can be checked using the moment-generating function, see, e.g.,these notes (http://www.stat.washington.edu/thompson/S341_10/Notes/week4.pdf), 10.4-(ii)): multiplication by a positive constant c divides the rate (or, equivalently, multiplies the scale).

Exponential family

The gamma distribution is a two-parameter exponential family with natural parameters $k - 1$ and $-1/\theta$ (equivalently, $\alpha - 1$ and $-\beta$), and natural statistics X and $\ln(X)$.

If the shape parameter k is held fixed, the resulting one-parameter family of distributions is a natural exponential family.

Logarithmic expectation and variance

One can show that

$$\mathrm{E}[\ln(X)] = \psi(\alpha) - \ln(\beta)$$

or equivalently,

$$\mathrm{E}[\ln(X)] = \psi(k) + \ln(\theta)$$

where ψ is the digamma function. Likewise,

$$\operatorname{var}[\ln(X)] = \psi^{(1)}(\alpha) = \psi^{(1)}(k)$$

where $\psi^{(1)}$ is the trigamma function.

This can be derived using the exponential family formula for the moment generating function of the sufficient statistic, because one of the sufficient statistics of the gamma distribution is $\ln(x)$.

Information entropy

The information entropy is

$$\begin{aligned}\mathrm{H}(X) &= \mathrm{E}[-\ln(p(X))] \\ &= \mathrm{E}[-\alpha\ln(\beta) + \ln(\Gamma(\alpha)) - (\alpha - 1)\ln(X) + \beta X] \\ &= \alpha - \ln(\beta) + \ln(\Gamma(\alpha)) + (1 - \alpha)\psi(\alpha).\end{aligned}$$

In the k, θ parameterization, the information entropy is given by

$$\mathrm{H}(X) = k + \ln(\theta) + \ln(\Gamma(k)) + (1 - k)\psi(k).$$

Kullback–Leibler divergence

The Kullback–Leibler divergence (KL-divergence), of Gamma(α_p, β_p) ("true" distribution) from Gamma(α_q, β_q) ("approximating" distribution) is given by[13]

$$\begin{aligned}D_{\mathrm{KL}}(\alpha_p, \beta_p; \alpha_q, \beta_q) = (\alpha_p - \alpha_q)\psi(\alpha_p) - \log\Gamma(\alpha_p) + \log\Gamma(\alpha_q) \\ + \alpha_q(\log\beta_p - \log\beta_q) + \alpha_p\frac{\beta_q - \beta_p}{\beta_p}.\end{aligned}$$

Written using the k, θ parameterization, the KL-divergence of Gamma(k_p, θ_p) from Gamma(k_q, θ_q) is given by

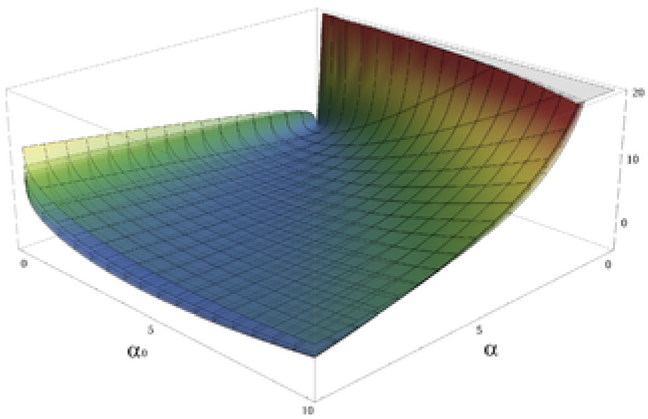


Illustration of the Kullback–Leibler (KL) divergence for two gamma PDFs. Here $\beta = \beta_0 + 1$ which are set to 1, 2, 3, 4, 5 and 6. The typical asymmetry for the KL divergence is clearly visible.

$$\begin{aligned} D_{\mathrm{KL}}(k_p, \theta_p; k_q, \theta_q) = (k_p - k_q)\psi(k_p) - \log\Gamma(k_p) + \log\Gamma(k_q) \\ + k_q(\log\theta_q - \log\theta_p) + k_p\frac{\theta_p - \theta_q}{\theta_q}. \end{aligned}$$

Laplace transform

The Laplace transform of the gamma PDF is

$$F(s) = (1 + \theta s)^{-k} = \frac{\beta^\alpha}{(s+\beta)^\alpha}.$$

Related distributions

General

- Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed random variables following an exponential distribution with rate parameter λ, then $\sum_i X_i$ ~ Gamma(n, λ) where n is the shape parameter and λ is the rate, and $\bar{X} = \frac{1}{n}\sum_i X_i \sim \mathbf{Gamma}(n, n\lambda)$ where the rate changes $n\lambda$.
- If X ~ Gamma(1, 1/λ) (in the shape–scale parametrization), then X has an exponential distribution with rate parameter λ.
- If X ~ Gamma(ν/2, 2) (in the shape–scale parametrization), then X is identical to $\chi^2(\nu)$, the chi-squared distribution with ν degrees of freedom. Conversely, if $Q \sim \chi^2(\nu)$ and c is a positive constant, then cQ ~ Gamma(ν/2, 2c).
- If k is an integer, the gamma distribution is an Erlang distribution and is the probability distribution of the waiting time until the kth "arrival" in a one-dimensional Poisson process with intensity 1/θ. If

$$X \sim \Gamma(k \in \mathbf{Z}, \theta), \qquad Y \sim \mathrm{Pois}\left(\frac{x}{\theta}\right),$$

then

$$P(X > x) = P(Y < k).$$

- If X has a Maxwell–Boltzmann distribution with parameter a, then

$$X^2 \sim \Gamma\left(\frac{3}{2}, 2a^2\right).$$

- If X ~ Gamma(k, θ), then $\log X$ follows an exponential-gamma (abbreviated exp-gamma) distribution.[14] It is sometimes referred to as the log-gamma distribution.[15] Formulas for its mean and variance are in the section #Logarithmic expectation and variance.
- If X ~ Gamma(k, θ), then $\sqrt{X}$ follows a generalized gamma distribution with parameters p = 2, d = 2k, and $a = \sqrt{\theta}$.
- More generally, if X ~ Gamma(k,θ), then X^q for $q > 0$ follows a generalized gamma distribution with parameters p = 1/q, d = k/q, and $a = \theta^q$.
- If X ~ Gamma(k, θ) with shape k and scale θ, then 1/X ~ Inv-Gamma(k, θ^{-1}) (see Inverse-gamma distribution for derivation).
- Parametrization 1: If $X_k \sim \Gamma(\alpha_k, \theta_k)$ are independent, then $\frac{\alpha_2 \theta_2 X_1}{\alpha_1 \theta_1 X_2} \sim \mathrm{F}(2\alpha_1, 2\alpha_2)$, or equivalently, $\frac{X_1}{X_2} \sim \beta'\left(\alpha_1, \alpha_2, 1, \frac{\theta_1}{\theta_2}\right)$
- Parametrization 2: If $X_k \sim \Gamma(\alpha_k, \beta_k)$ are independent, then $\frac{\alpha_2 \beta_1 X_1}{\alpha_1 \beta_2 X_2} \sim \mathrm{F}(2\alpha_1, 2\alpha_2)$, or equivalently, $\frac{X_1}{X_2} \sim \beta'\left(\alpha_1, \alpha_2, 1, \frac{\beta_2}{\beta_1}\right)$
- If X ~ Gamma(α, θ) and Y ~ Gamma(β, θ) are independently distributed, then X/(X + Y) has a beta distribution with parameters α and β, and X/(X + Y) is independent of X + Y, which is Gamma(α + β, θ)-distributed.
- If X_i ~ Gamma(α_i, 1) are independently distributed, then the vector (X_1/S, ..., X_n/S), where S = X_1 + ... + X_n, follows a Dirichlet distribution with parameters α_1, ..., α_n.
- For large k the gamma distribution converges to normal distribution with mean $\mu = k\theta$ and variance $\sigma^2 = k\theta^2$.
- The gamma distribution is the conjugate prior for the precision of the normal distribution with known mean.
- The Wishart distribution is a multivariate generalization of the gamma distribution (samples are positive-definite matrices rather than positive real numbers).
- The gamma distribution is a special case of the generalized gamma distribution, the generalized integer gamma distribution, and the generalized inverse Gaussian distribution.
- Among the discrete distributions, the negative binomial distribution is sometimes considered the discrete analogue of the gamma distribution.
- Tweedie distributions – the gamma distribution is a member of the family of Tweedie exponential dispersion models.

Compound gamma

If the shape parameter of the gamma distribution is known, but the inverse-scale parameter is unknown, then a gamma distribution for the inverse scale forms a conjugate prior. The compound distribution, which results from integrating out the inverse scale, has a closed-form solution, known as the compound gamma distribution.[16]

If instead the shape parameter is known but the mean is unknown, with the prior of the mean being given by another gamma distribution, then it results in K-distribution.

Weibull and Stable count

The gamma distribution $f(x;k)$ $(k > 1)$ can be expressed as the product distribution of a Weibull distribution and a variant form of the stable count distribution. Its shape parameter k can be regarded as inverse of Lévy's stability parameter in the stable count distribution:

$$f(x;k) = \int_0^\infty \frac{1}{u} W_k\left(\frac{x}{u}\right) \left[k u^{k-1} \mathfrak{N}_{\frac{1}{k}}\left(u^k\right) \right] du,$$

where $\mathfrak{N}_\alpha(\nu)$ is a standard stable count distribution of shape $\alpha = 1/k$, and $W_k(x)$ is a standard Weibull distribution of shape k.

Statistical inference

Parameter estimation

Maximum likelihood estimation

The likelihood function for N iid observations $(x_1, ..., x_N)$ is

$$L(k,\theta) = \prod_{i=1}^N f(x_i;k,\theta)$$

from which we calculate the log-likelihood function

$$\ell(k,\theta) = (k-1)\sum_{i=1}^N \ln(x_i) - \sum_{i=1}^N \frac{x_i}{\theta} - Nk\ln(\theta) - N\ln(\Gamma(k))$$

Finding the maximum with respect to θ by taking the derivative and setting it equal to zero yields the maximum likelihood estimator of the θ parameter:

$$\hat{\theta} = \frac{1}{kN}\sum_{i=1}^N x_i$$

Substituting this into the log-likelihood function gives

$$\ell = (k-1)\sum_{i=1}^N \ln(x_i) - Nk - Nk\ln\left(\frac{\sum x_i}{kN}\right) - N\ln(\Gamma(k))$$

Finding the maximum with respect to k by taking the derivative and setting it equal to zero yields

$$\ln(k) - \psi(k) = \ln\left(\frac{1}{N}\sum_{i=1}^N x_i\right) - \frac{1}{N}\sum_{i=1}^N \ln(x_i)$$

where ψ is the digamma function. There is no closed-form solution for k. The function is numerically very well behaved, so if a numerical solution is desired, it can be found using, for example, Newton's method. An initial value of k can be found either using the method of moments, or using the approximation

$$\ln(k) - \psi(k) \approx \frac{1}{2k}\left(1 + \frac{1}{6k+1}\right)$$

If we let

$$s = \ln\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) - \frac{1}{N}\sum_{i=1}^{N}\ln(x_i)$$

then k is approximately

$$k \approx \frac{3 - s + \sqrt{(s-3)^2 + 24s}}{12s}$$

which is within 1.5% of the correct value.[17] An explicit form for the Newton–Raphson update of this initial guess is:[18]

$$k \leftarrow k - \frac{\ln(k) - \psi(k) - s}{\frac{1}{k} - \psi'(k)}.$$

Closed-form estimators

Consistent closed-form estimators of k and θ exists that are derived from the likelihood of the generalized gamma distribution.[19]

The estimate for the shape k is

$$\hat{k} = \frac{N\sum_{i=1}^{N} x_i}{N\sum_{i=1}^{N} x_i \ln(x_i) - \sum_{i=1}^{N}\ln(x_i)\sum_{i=1}^{N} x_i}$$

and the estimate for the scale θ is

$$\hat{\theta} = \frac{1}{N^2}\left(N\sum_{i=1}^{N} x_i \ln(x_i) - \sum_{i=1}^{N}\ln(x_i)\sum_{i=1}^{N} x_i\right)$$

If the rate parameterization is used, the estimate of $\hat{\beta} = \frac{1}{\hat{\theta}}$.

These estimators are not strictly maximum likelihood estimators, but are instead referred to as mixed type log-moment estimators. They have however similar efficiency as the maximum likelihood estimators.

Although these estimators are consistent, they have a small bias. A bias-corrected variant of the estimator for the scale θ is

$$\tilde{\theta} = \frac{N}{N-1}\hat{\theta}$$

A bias correction for the shape parameter k is given as[20]

$$\tilde{k} = \hat{k} - \frac{1}{N}\left(3\hat{k} - \frac{2}{3}\left(\frac{\hat{k}}{1+\hat{k}}\right) - \frac{4}{5}\frac{\hat{k}}{(1+\hat{k})^2}\right)$$

Bayesian minimum mean squared error

With known k and unknown θ, the posterior density function for theta (using the standard scale-invariant prior for θ) is

$$P(\theta \mid k, x_1, \dots, x_N) \propto \frac{1}{\theta}\prod_{i=1}^{N} f(x_i; k, \theta)$$

Denoting

$$y \equiv \sum_{i=1}^{N} x_i, \qquad P(\theta \mid k, x_1, \dots, x_N) = C(x_i)\theta^{-Nk-1}e^{-y/\theta}$$

Integration with respect to θ can be carried out using a change of variables, revealing that $1/\theta$ is gamma-distributed with parameters $\alpha = Nk$, $\beta = y$.

$$\int_0^\infty \theta^{-Nk-1+m}e^{-y/\theta}\,d\theta = \int_0^\infty x^{Nk-1-m}e^{-xy}\,dx = y^{-(Nk-m)}\Gamma(Nk-m)$$

The moments can be computed by taking the ratio (m by $m = 0$)

$$\mathrm{E}[x^m] = \frac{\Gamma(Nk-m)}{\Gamma(Nk)}y^m$$

which shows that the mean ± standard deviation estimate of the posterior distribution for θ is

$$\frac{y}{Nk-1} \pm \sqrt{\frac{y^2}{(Nk-1)^2(Nk-2)}}.$$

Bayesian inference

Conjugate prior

In Bayesian inference, the **gamma distribution** is the conjugate prior to many likelihood distributions: the Poisson, exponential, normal (with known mean), Pareto, gamma with known shape σ, inverse gamma with known shape parameter, and Gompertz with known scale parameter.

The gamma distribution's conjugate prior is:[21]

$$p(k,\theta \mid p,q,r,s) = \frac{1}{Z} \frac{p^{k-1} e^{-\theta^{-1} q}}{\Gamma(k)^r \theta^{ks}},$$

where Z is the normalizing constant, which has no closed-form solution. The posterior distribution can be found by updating the parameters as follows:

$$\begin{aligned} p' &= p \prod\nolimits_i x_i, \\ q' &= q + \sum\nolimits_i x_i, \\ r' &= r + n, \\ s' &= s + n, \end{aligned}$$

where n is the number of observations, and x_i is the ith observation.

Occurrence and applications

Consider a sequence of events, with the waiting time for each event being an exponential distribution with rate β. Then the waiting time for the n-th event to occur is the gamma distribution with integer shape $\alpha = n$. This construction of the gamma distribution allows it to model a wide variety of phenomena where several sub-events, each taking time with exponential distribution, must happen in sequence for a major event to occur.[22] Examples include the waiting time of cell-division events,[23] number of compensatory mutations for a given mutation,[24] waiting time until a repair is necessary for a hydraulic system,[25] and so on.

The gamma distribution has been used to model the size of insurance claims[26] and rainfalls.[27] This means that aggregate insurance claims and the amount of rainfall accumulated in a reservoir are modelled by a gamma process – much like the exponential distribution generates a Poisson process.

The gamma distribution is also used to model errors in multi-level Poisson regression models, because a mixture of Poisson distributions with gamma distributed rates has a known closed form distribution, called negative binomial.

In wireless communication, the gamma distribution is used to model the multi-path fading of signal power; see also Rayleigh distribution and Rician distribution.

In oncology, the age distribution of cancer incidence often follows the gamma distribution, whereas the shape and scale parameters predict, respectively, the number of driver events and the time interval between them.[28][29]

In neuroscience, the gamma distribution is often used to describe the distribution of inter-spike intervals.[30][31]

In bacterial gene expression, the copy number of a constitutively expressed protein often follows the gamma distribution, where the scale and shape parameter are, respectively, the mean number of bursts per cell cycle and the mean number of protein molecules produced by a single mRNA during its lifetime.[32]

In genomics, the gamma distribution was applied in peak calling step (i.e. in recognition of signal) in ChIP-chip[33] and ChIP-seq[34] data analysis.

The gamma distribution is widely used as a conjugate prior in Bayesian statistics. It is the conjugate prior for the precision (i.e. inverse of the variance) of a normal distribution. It is also the conjugate prior for the exponential distribution.

Generating gamma-distributed random variables

Given the scaling property above, it is enough to generate gamma variables with $\theta = 1$ as we can later convert to any value of β with simple division.

Suppose we wish to generate random variables from Gamma($n + \delta$, 1), where n is a non-negative integer and $0 < \delta < 1$. Using the fact that a Gamma(1, 1) distribution is the same as an Exp(1) distribution, and noting the method of generating exponential variables, we conclude that if U is uniformly distributed on (0, 1], then $-\ln(U)$ is distributed Gamma(1, 1) (i.e. inverse transform sampling). Now, using the "α-addition" property of gamma distribution, we expand this result:

$$-\sum_{k=1}^{n} \ln U_k \sim \Gamma(n, 1)$$

where U_k are all uniformly distributed on (0, 1] and independent. All that is left now is to generate a variable distributed as Gamma(δ, 1) for $0 < \delta < 1$ and apply the "α-addition" property once more. This is the most difficult part.

Random generation of gamma variates is discussed in detail by Devroye,[35]:401–428 noting that none are uniformly fast for all shape parameters. For small values of the shape parameter, the algorithms are often not valid.[35]:406 For arbitrary values of the shape parameter, one can apply the Ahrens and Dieter[36] modified acceptance–rejection method Algorithm GD (shape $k \geq 1$), or transformation method[37] when $0 < k < 1$. Also see Cheng and Feast Algorithm GKM 3[38] or Marsaglia's squeeze method.[39]

The following is a version of the Ahrens-Dieter acceptance–rejection method:[36]

1. Generate U, V and W as iid uniform (0, 1] variates.
2. If $U \leq \frac{e}{e+\delta}$ then $\xi = V^{1/\delta}$ and $\eta = W\xi^{\delta-1}$. Otherwise, $\xi = 1 - \ln V$ and $\eta = We^{-\xi}$.
3. If $\eta > \xi^{\delta-1}e^{-\xi}$ then go to step 1.
4. ξ is distributed as $\Gamma(\delta, 1)$.

A summary of this is

$$\theta\left(\xi - \sum_{i=1}^{\lfloor k \rfloor} \ln(U_i)\right) \sim \Gamma(k, \theta)$$

where $\lfloor k \rfloor$ is the integer part of k, ξ is generated via the algorithm above with $\delta = \{k\}$ (the fractional part of k) and the U_k are all independent.

While the above approach is technically correct, Devroye notes that it is linear in the value of k and in general is not a good choice. Instead he recommends using either rejection-based or table-based methods, depending on context.[35]:401–428

For example, Marsaglia's simple transformation-rejection method relying on one normal variate X and one uniform variate U:[40]

1. Set $d = a - \frac{1}{3}$ and $c = \frac{1}{\sqrt{9d}}$.
2. Set $v = (1 + cX)^3$.
3. If $v > 0$ and $\ln U < \frac{X^2}{2} + d - dv + d \ln v$ return dv, else go back to step 2.

With $1 \le a = \alpha = k$ generates a gamma distributed random number in time that is approximately constant with k. The acceptance rate does depend on k, with an acceptance rate of 0.95, 0.98, and 0.99 for k=1, 2, and 4. For $k < 1$, one can use $\gamma_\alpha = \gamma_{1+\alpha} U^{1/\alpha}$ to boost k to be usable with this method.

References

1. Park, Sung Y.; Bera, Anil K. (2009). "Maximum entropy autoregressive conditional heteroskedasticity model" (https://web.archive.org/web/20160307144515/http://wise.xmu.edu.cn/uploadfiles/paper-masterdownload/2009519932327055475115776.pdf) (PDF). *Journal of Econometrics*. **150** (2): 219–230. CiteSeerX 10.1.1.511.9750 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.511.9750). doi:10.1016/j.jeconom.2008.12.014 (https://doi.org/10.1016%2Fj.jeconom.2008.12.014). Archived from the original (http://www.wise.xmu.edu.cn/Master/Download/..%5C..%5CUploadFiles%5Cpaper-masterdownload%5C2009519932327055475115776.pdf) (PDF) on 2016-03-07. Retrieved 2011-06-02.
2. Hogg, R. V.; Craig, A. T. (1978). *Introduction to Mathematical Statistics* (4th ed.). New York: Macmillan. pp. Remark 3.3.1. ISBN 0023557109.
3. *Scalable Recommendation with Poisson Factorization* (https://arxiv.org/abs/1311.1704), Prem Gopalan, Jake M. Hofman, David Blei, arXiv.org 2014
4. Papoulis, Pillai, *Probability, Random Variables, and Stochastic Processes*, Fourth Edition
5. Jeesen Chen, Herman Rubin, Bounds for the difference between median and mean of gamma and poisson distributions, Statistics & Probability Letters, Volume 4, Issue 6, October 1986, Pages 281–283, ISSN 0167-7152 (https://www.worldcat.org/search?fq=x0:jrnl&q=n2:0167-7152), [1] (https://dx.doi.org/10.1016/0167-7152(86)90044-1).
6. Choi, K. P. "On the Medians of the Gamma Distributions and an Equation of Ramanujan" (https://www.ams.org/journals/proc/1994-121-01/S0002-9939-1994-1195477-8/S0002-9939-1994-1195477-8.pdf), Proceedings of the American Mathematical Society, Vol. 121, No. 1 (May, 1994), pp. 245–251.
7. Berg, Christian & Pedersen, Henrik L. (March 2006). "The Chen–Rubin conjecture in a continuous setting" (https://www.intlpress.com/site/pub/files/_fulltext/journals/maa/2006/0013/0001/MAA-2006-0013-0001-a004.pdf) (PDF). *Methods and Applications of Analysis*. **13** (1): 63–88. doi:10.4310/MAA.2006.v13.n1.a4 (https://doi.org/10.4310%2FMAA.2006.v13.n1.a4). S2CID 6704865 (https://api.semanticscholar.org/CorpusID:6704865). Retrieved 1 April 2020.
8. Berg, Christian and Pedersen, Henrik L. "Convexity of the median in the gamma distribution" (https://arxiv.org/abs/math/0609442).
9. Gaunt, Robert E., and Milan Merkle (2021). "On bounds for the mode and median of the generalized hyperbolic and related distributions". *Journal of Mathematical Analysis and Applications*. **493** (1): 124508. arXiv:2002.01884 (https://arxiv.org/abs/2002.01884). doi:10.1016/j.jmaa.2020.124508 (https://doi.org/10.1016%2Fj.jmaa.2020.124508). S2CID 221103640 (https://api.semanticscholar.org/CorpusID:221103640).

10. Lyon, Richard F. (13 May 2021). "On closed-form tight bounds and approximations for the median of a gamma distribution" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8118309). *PLOS One*. **16** (5): e0251626. arXiv:2011.04060 (https://arxiv.org/abs/2011.04060). Bibcode:2021PLoSO..1651626L (https://ui.adsabs.harvard.edu/abs/2021PLoSO..1651626L). doi:10.1371/journal.pone.0251626 (https://doi.org/10.1371%2Fjournal.pone.0251626). PMC 8118309 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8118309). PMID 33984053 (https://pubmed.ncbi.nlm.nih.gov/33984053).
11. Mathai, A. M. (1982). "Storage capacity of a dam with gamma type inputs". *Annals of the Institute of Statistical Mathematics*. **34** (3): 591–597. doi:10.1007/BF02481056 (https://doi.org/10.1007%2FBF02481056). ISSN 0020-3157 (https://www.worldcat.org/issn/0020-3157). S2CID 122537756 (https://api.semanticscholar.org/CorpusID:122537756).
12. Moschopoulos, P. G. (1985). "The distribution of the sum of independent gamma random variables". *Annals of the Institute of Statistical Mathematics*. **37** (3): 541–544. doi:10.1007/BF02481123 (https://doi.org/10.1007%2FBF02481123). S2CID 120066454 (https://api.semanticscholar.org/CorpusID:120066454).
13. W.D. Penny, [www.fil.ion.ucl.ac.uk/~wpenny/publications/densities.ps KL-Divergences of Normal, Gamma, Dirichlet, and Wishart densities]
14. "ExpGammaDistribution—Wolfram Language Documentation" (https://reference.wolfram.com/language/ref/ExpGammaDistribution.html).
15. "scipy.stats.loggamma — SciPy v1.8.0 Manual" (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.loggamma.html#scipy.stats.loggamma). *docs.scipy.org*.
16. Dubey, Satya D. (December 1970). "Compound gamma, beta and F distributions". *Metrika*. **16**: 27–31. doi:10.1007/BF02613934 (https://doi.org/10.1007%2FBF02613934). S2CID 123366328 (https://api.semanticscholar.org/CorpusID:123366328).
17. Minka, Thomas P. (2002). "Estimating a Gamma distribution" (https://tminka.github.io/papers/minka-gamma.pdf) (PDF).
18. Choi, S. C.; Wette, R. (1969). "Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias". *Technometrics*. **11** (4): 683–690. doi:10.1080/00401706.1969.10490731 (https://doi.org/10.1080%2F00401706.1969.10490731).
19. Ye, Zhi-Sheng; Chen, Nan (2017). "Closed-Form Estimators for the Gamma Distribution Derived from Likelihood Equations" (https://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1209129). *The American Statistician*. **71** (2): 177–181. doi:10.1080/00031305.2016.1209129 (https://doi.org/10.1080%2F00031305.2016.1209129). S2CID 124682698 (https://api.semanticscholar.org/CorpusID:124682698).
20. Louzada, Francisco; Ramos, Pedro L.; Ramos, Eduardo (2019). "A Note on Bias of Closed-Form Estimators for the Gamma Distribution Derived from Likelihood Equations" (https://www.tandfonline.com/doi/abs/10.1080/00031305.2018.1513376). *The American Statistician*. **73** (2): 195–199. doi:10.1080/00031305.2018.1513376 (https://doi.org/10.1080%2F00031305.2018.1513376). S2CID 126086375 (https://api.semanticscholar.org/CorpusID:126086375).
21. Fink, D. 1995 A Compendium of Conjugate Priors (http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.157.5540&rep=rep1&type=pdf). In progress report: Extension and enhancement of methods for setting data quality objectives. (DOE contract 95-831).
22. Jessica., Scheiner, Samuel M., 1956- Gurevitch (2001). "13. Failure-time analysis" (https://www.google.com/books/edition/Design_and_Analysis_of_Ecological_Experi/AgsTDAAAQBAJ?hl=en&gbpv=1&dq=gamma+distribution+failure+waiting+time&pg=PA235&printsec=frontcover). *Design and analysis of ecological experiments* (http://worldcat.org/oclc/43694448). Oxford University Press. ISBN 0-19-513187-8. OCLC 43694448 (https://www.worldcat.org/oclc/43694448).

23. Golubev, A. (March 2016). "Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression" (https://dx.doi.org/10.1016/j.jtbi.2015.12.027). *Journal of Theoretical Biology*. **393**: 203–217. Bibcode:2016JThBi.393..203G (https://ui.adsabs.harvard.edu/abs/2016JThBi.393..203G). doi:10.1016/j.jtbi.2015.12.027 (https://doi.org/10.1016%2Fj.jtbi.2015.12.027). ISSN 0022-5193 (https://www.worldcat.org/issn/0022-5193). PMID 26780652 (https://pubmed.ncbi.nlm.nih.gov/26780652).
24. Poon, Art; Davis, Bradley H; Chao, Lin (2005-07-01). "The Coupon Collector and the Suppressor Mutation" (https://dx.doi.org/10.1534/genetics.104.037259). *Genetics*. **170** (3): 1323–1332. doi:10.1534/genetics.104.037259 (https://doi.org/10.1534%2Fgenetics.104.037259). ISSN 1943-2631 (https://www.worldcat.org/issn/1943-2631). PMC 1451182 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1451182). PMID 15879511 (https://pubmed.ncbi.nlm.nih.gov/15879511).
25. Vineyard, Michael; Amoako-Gyampah, Kwasi; Meredith, Jack R (July 1999). "Failure rate distributions for flexible manufacturing systems: An empirical study" (https://dx.doi.org/10.1016/s0377-2217(98)00096-4). *European Journal of Operational Research*. **116** (1): 139–155. doi:10.1016/s0377-2217(98)00096-4 (https://doi.org/10.1016%2Fs0377-2217%2898%2900096-4). ISSN 0377-2217 (https://www.worldcat.org/issn/0377-2217).
26. p. 43, Philip J. Boland, Statistical and Probabilistic Methods in Actuarial Science, Chapman & Hall CRC 2007
27. Wilks, Daniel S. (1990). "Maximum Likelihood Estimation for the Gamma Distribution Using Data Containing Zeros" (https://www.jstor.org/stable/26196366). *Journal of Climate*. **3** (12): 1495–1501. Bibcode:1990JCli....3.1495W (https://ui.adsabs.harvard.edu/abs/1990JCli....3.1495W). doi:10.1175/1520-0442(1990)003<1495:MLEFTG>2.0.CO;2 (https://doi.org/10.1175%2F1520-0442%281990%29003%3C1495%3AMLEFTG%3E2.0.CO%3B2). ISSN 0894-8755 (https://www.worldcat.org/issn/0894-8755). JSTOR 26196366 (https://www.jstor.org/stable/26196366).
28. Belikov, Aleksey V. (22 September 2017). "The number of key carcinogenic events can be predicted from cancer incidence" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5610194). *Scientific Reports*. **7** (1): 12170. Bibcode:2017NatSR...712170B (https://ui.adsabs.harvard.edu/abs/2017NatSR...712170B). doi:10.1038/s41598-017-12448-7 (https://doi.org/10.1038%2Fs41598-017-12448-7). PMC 5610194 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5610194). PMID 28939880 (https://pubmed.ncbi.nlm.nih.gov/28939880).
29. Belikov, Aleksey V.; Vyatkin, Alexey; Leonov, Sergey V. (2021-08-06). "The Erlang distribution approximates the age distribution of incidence of childhood and young adulthood cancers" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8351573). *PeerJ*. **9**: e11976. doi:10.7717/peerj.11976 (https://doi.org/10.7717%2Fpeerj.11976). ISSN 2167-8359 (https://www.worldcat.org/issn/2167-8359). PMC 8351573 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8351573). PMID 34434669 (https://pubmed.ncbi.nlm.nih.gov/34434669).
30. J. G. Robson and J. B. Troy, "Nature of the maintained discharge of Q, X, and Y retinal ganglion cells of the cat", J. Opt. Soc. Am. A 4, 2301–2307 (1987)
31. M.C.M. Wright, I.M. Winter, J.J. Forster, S. Bleeck "Response to best-frequency tone bursts in the ventral cochlear nucleus is governed by ordered inter-spike interval statistics", Hearing Research 317 (2014)
32. N. Friedman, L. Cai and X. S. Xie (2006) "Linking stochastic dynamics to population distribution: An analytical framework of gene expression", *Phys. Rev. Lett.* 97, 168302.
33. DJ Reiss, MT Facciotti and NS Baliga (2008) "Model-based deconvolution of genome-wide DNA binding" (https://web.archive.org/web/20121117144623/http://bioinformatics.oxfordjournals.org/content/24/3/396.full.pdf+html), *Bioinformatics*, 24, 396–403
34. MA Mendoza-Parra, M Nowicka, W Van Gool, H Gronemeyer (2013) "Characterising ChIP-seq binding patterns by model-based peak shape deconvolution" (http://www.biomedcentral.com/1471-2164/14/834), *BMC Genomics*, 14:834

35. Devroye, Luc (1986). *Non-Uniform Random Variate Generation* (http://luc.devroye.org/rnbookindex.html). New York: Springer-Verlag. ISBN 978-0-387-96305-1. See Chapter 9, Section 3.
36. Ahrens, J. H.; Dieter, U (January 1982). "Generating gamma variates by a modified rejection technique". *Communications of the ACM*. **25** (1): 47–54. doi:10.1145/358315.358390 (https://doi.org/10.1145%2F358315.358390). S2CID 15128188 (https://api.semanticscholar.org/CorpusID:15128188).. See Algorithm GD, p. 53.
37. Ahrens, J. H.; Dieter, U. (1974). "Computer methods for sampling from gamma, beta, Poisson and binomial distributions". *Computing*. **12** (3): 223–246. CiteSeerX 10.1.1.93.3828 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.93.3828). doi:10.1007/BF02293108 (https://doi.org/10.1007%2FBF02293108). S2CID 37484126 (https://api.semanticscholar.org/CorpusID:37484126).
38. Cheng, R.C.H., and Feast, G.M. Some simple gamma variate generators. Appl. Stat. 28 (1979), 290–295.
39. Marsaglia, G. The squeeze method for generating gamma variates. Comput, Math. Appl. 3 (1977), 321–325.
40. Marsaglia, G.; Tsang, W. W. (2000). "A simple method for generating gamma variables". *ACM Transactions on Mathematical Software*. **26** (3): 363–372. doi:10.1145/358407.358414 (https://doi.org/10.1145%2F358407.358414). S2CID 2634158 (https://api.semanticscholar.org/CorpusID:2634158).

External links

- "Gamma-distribution" (https://www.encyclopediaofmath.org/index.php?title=Gamma-distribution), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Weisstein, Eric W. "Gamma distribution" (https://mathworld.wolfram.com/GammaDistribution.html). *MathWorld*.
- ModelAssist (2017) Uses of the gamma distribution in risk modeling, including applied examples in Excel (http://www.epixanalytics.com/modelassist/AtRisk/Model_Assist.htm#Distributions/Continuous_distributions/Gamma.htm).
- Engineering Statistics Handbook (http://www.itl.nist.gov/div898/handbook/eda/section3/eda366b.htm)