Gamma distribution

From Wikipedia, the free encyclopedia

In probability theory and statistics, the **gamma distribution** is a two-parameter family of continuous probability distributions. The common exponential distribution and chi-squared distribution are special cases of the gamma distribution. There are three different parametrizations 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.
3. With a shape parameter k and a mean parameter $\mu = k/\beta$.

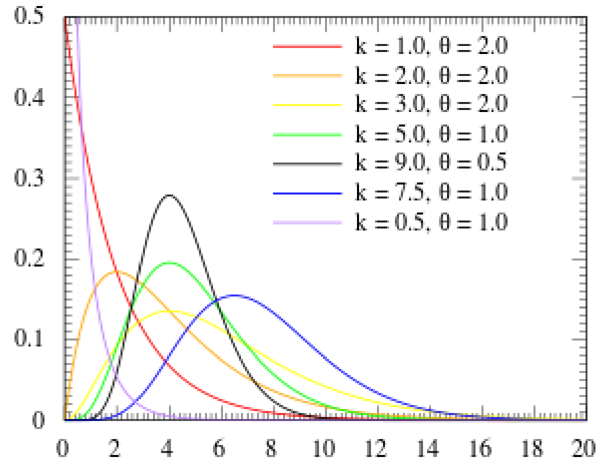
In each of these three forms, both parameters are positive real numbers.

The gamma distribution is the maximum entropy probability distribution 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).[2]

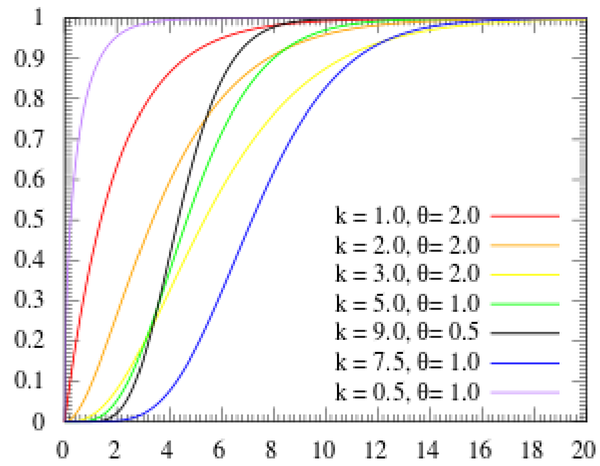
Contents

Gamma

Probability density function



Cumulative distribution function



Parameters	■ $k > 0$ shape ■ $\theta > 0$ scale	■ $\alpha > 0$ shape ■ $\beta > 0$ rate
Support	$x \in (0, \infty)$	$x \in (0, \infty)$

Parameterizations

The parameterization with k and θ appears to be more common in econometrics and certain other applied fields, where e.g. 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.[3]

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 (aka rate) parameters, such as the λ of an exponential distribution or a Poisson distribution[4] – 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

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

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

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

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

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:[5]

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

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

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

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

The **cumulative distribution function** is the regularized gamma function:

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

PDF	$\frac{1}{\Gamma(k)\theta^k}x^{k-1}e^{-\frac{x}{\theta}}$	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ [1]
CDF	$1-\frac{1}{\Gamma(k)}\Gamma\left(k,\frac{x}{\theta}\right)$	$1-\frac{1}{\Gamma(\alpha)}\Gamma(\alpha,\beta x)$
Mean	$\mathbf{E}[X]=k\theta$ $\mathbf{E}[\ln X]=\psi(k)+\ln(\theta)$ (see digamma function)	$\mathbf{E}[X]=\frac{\alpha}{\beta}$ $\mathbf{E}[\ln X]=\psi(\alpha)-\ln(\beta)$ (see digamma function)
Median	No simple closed form	No simple closed form
Mode	$(k-1)\theta$ for $k\geq 1$	$\frac{\alpha-1}{\beta}$ for $\alpha\geq 1$
Variance	$\mathrm{Var}[X]=k\theta^2$ $\mathrm{Var}[\ln X]=\psi_1(k)$ (see trigamma function)	$\mathrm{Var}[X]=\frac{\alpha}{\beta^2}$ $\mathrm{Var}[\ln X]=\psi_1(\alpha)$ (see trigamma function)
Skewness	$\frac{2}{\sqrt{k}}$	$\frac{2}{\sqrt{\alpha}}$
Excess kurtosis	$\frac{6}{k}$	$\frac{6}{\alpha}$
Entropy	$k+\ln\theta+\ln[\Gamma(k)]$ $+(1-k)\psi(k)$	$\alpha-\ln\beta+\ln[\Gamma(\alpha)]$ $+(1-\alpha)\psi(\alpha)$
MGF	$(1-\theta t)^{-k}$ for $t<\frac{1}{\theta}$	$\left(1-\frac{t}{\beta}\right)^{-\alpha}$ for $t<\beta$
CF	$(1-\theta i\,t)^{-k}$	$\left(1-\frac{i\,t}{\beta}\right)^{-\alpha}$

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):[5]

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

Properties

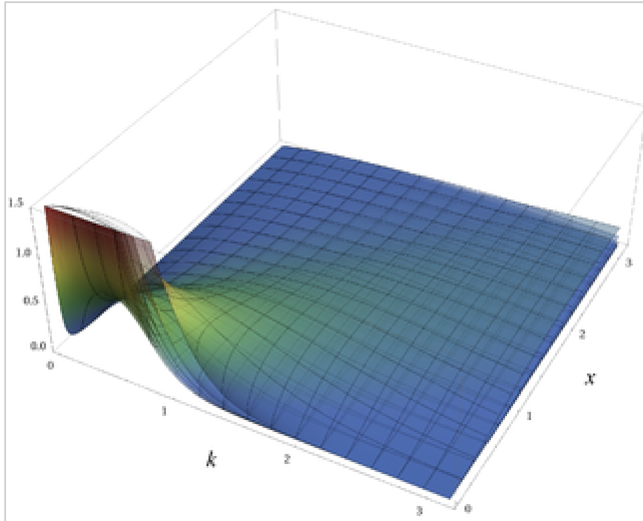


Illustration of the gamma PDF for parameter values over k and x with θ set to 1, 2, 3, 4, 5 and 6. One can see each θ layer by itself here [2] (http://commons.wikimedia.org/wiki/File:Gamma-PDF-3D-by-k.png) as well as by k [3] (http://commons.wikimedia.org/wiki/File:Gamma-PDF-3D-by-Theta.png) and x. [4]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

Skewness

The skewness is equal to $2/\sqrt{k}$, it depends only on the shape parameter (k) and approaches a normal distribution when k is large (approximately when $k > 10$).

Median calculation

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

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

A formula for approximating the median for any gamma distribution, when the mean is known, has been derived based on the fact that the ratio $\mu/(\mu - \nu)$ is approximately a linear function of k when $k \geq 1$.[6] The approximation formula is

$$\nu \approx \mu \frac{3k - 0.8}{3k + 0.2},$$

where $\mu(= k\theta)$ is the mean.

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

$$m - \frac{1}{3} < \lambda(m) < m,$$

where $\lambda(m)$ denotes the median of the $\mathrm{Gamma}(m, 1)$ distribution.[7]

K. P. Choi later showed that the first five terms in the asymptotic expansion of the median are

$$\lambda(m) = m - \frac{1}{3} + \frac{8}{405m} + \frac{184}{25515m^2} + \frac{2248}{3444525m^3} - \cdots$$

by comparing the median to Ramanujan's θ function.[8]

Later, it was shown that $\lambda(m)$ is a convex function of m.[9]

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 (1982) and Moschopoulos (1984).

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

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

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

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.

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[10]

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

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

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

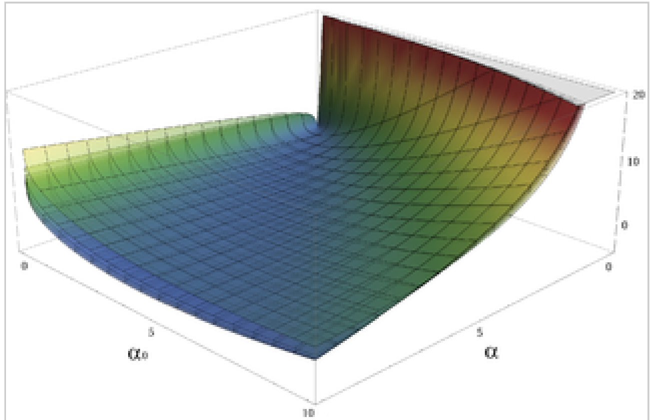


Illustration of the Kullback–Leibler (KL) divergence for two gamma PDFs. Here $\beta = \beta_0 + 1$ which are set to 1, 2, 3, 4, 5 and 6. The typical asymmetry for the KL divergence is clearly visible.

Laplace transform

The Laplace transform of the gamma PDF is

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

Differential equation

$$\left\{\beta x f'(x)+f(x)(-\alpha\beta+\beta+x)=0; f(1)=\frac{e^{-1/\beta}\beta^{-\alpha}}{\Gamma(\alpha)}\right\}$$

$$\left\{x f'(x)+f(x)(-k+\theta x+1)=0; f(1)=\frac{e^{-\theta}\left(\frac{1}{\theta}\right)^{-k}}{\Gamma(k)}\right\}$$

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

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

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.[11] An explicit form for the Newton–Raphson update of this initial guess is:[12]

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

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

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

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). 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,[13]: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.[13]:406 For arbitrary values of the shape parameter, one can apply the Ahrens and Dieter[14] modified acceptance–rejection method Algorithm GD (shape $k \geq 1$), or transformation method[15] when $0 < k < 1$. Also see Cheng and Feast Algorithm GKM 3[16] or

Marsaglia's squeeze method.[17]

The following is a version of the Ahrens-Dieter acceptance–rejection method:[14]

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 integral 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.[13]:401–428

For example, Marsaglia's simple transformation-rejection method relying on a one normal and one uniform random number:[18]

1. Setup: d = a - 1/3, c = 1/sqrt(9d).
2. Generate: v=(1+c*x)^3, with x standard normal.
3. if v > 0 and log(UNI) < 0.5 · x^2 + d − dv + d log(v) return dv.
4. go back to step 2.

With $1 \leq a = \alpha = k$ 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.

Applications

The gamma distribution has been used to model the size of insurance claims[19] and rainfalls.[20] This means that aggregate insurance claims and the amount of rainfall accumulated in a reservoir are modelled by a gamma process. The gamma distribution is also used to model errors in multi-level Poisson regression models, because the combination of the Poisson distribution and a gamma distribution is a negative binomial distribution.

In wireless communication, the gamma distribution is used to model the multi-path fading of signal power.

In neuroscience, the gamma distribution is often used to describe the distribution of inter-spike intervals.[21][22]

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.[23]

In genomics, the gamma distribution was applied in peak calling step (i.e. in recognition of signal) in ChIP-chip[24] and ChIP-seq[25] 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.

Related distributions

Special cases

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:[26]

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

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.

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.[27]

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.

Related distributions and properties

- If $X \sim$ Gamma(1, 1/λ) (shape -scale parametrization), then X has an exponential distribution with rate parameter λ.
- If $X \sim$ Gamma(ν/2, 2)(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 $cQ \sim$ 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 $1/\theta$. 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

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

- If $X \sim$ Gamma(k, θ), then $\sqrt{X}$ follows a generalized gamma distribution with parameters p = 2, d = 2k, and $a = \sqrt{\theta}$.
- More generally, if $X \sim$ 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 $X \sim$ Gamma(k, θ), then 1/$X \sim$ 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, $\frac{X_1}{X_2} \sim \beta'(\alpha_1, \alpha_2, 1, \frac{\theta_1}{\theta_2})$
- 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, $\frac{X_1}{X_2} \sim \beta'(\alpha_1, \alpha_2, 1, \frac{\beta_2}{\beta_1})$
- If $X \sim$ Gamma(α, θ) and $Y \sim$ Gamma(β, θ) are independently distributed, then X/(X + Y) has a beta distribution with parameters α and β.

- If $X_i \sim \mathrm{Gamma}(\alpha_i, 1)$ are independently distributed, then the vector $(X_1/S, ..., X_n/S)$, where $S = X_1 + ... + X_n$, follows a Dirichlet distribution with parameters $\alpha_1, ..., \alpha_n$.
- For large k the gamma distribution converges to Gaussian 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.

Notes

1. http://ocw.mit.edu/courses/mathematics/18-443-statistics-for-applications-fall-2006/lecture-notes/lecture6.pdf
2. Park, Sung Y.; Bera, Anil K. (2009). "Maximum entropy autoregressive conditional heteroskedasticity model" (PDF). *Journal of Econometrics*. Elsevier: 219–230. doi:10.1016/j.jeconom.2008.12.014. Retrieved 2011-06-02.
3. See Hogg and Craig (2005, Remark 3.3.1) for an explicit motivation
4. *Scalable Recommendation with Poisson Factorization* (http://arxiv.org/pdf/1311.1704v3.pdf), Prem Gopalan, Jake M. Hofman, David Blei, arXiv.org 2014
5. Papoulis, Pillai, *Probability, Random Variables, and Stochastic Processes*, Fourth Edition
6. Banneheka BMSG, Ekanayake GEMUPD (2009) "A new point estimator for the median of gamma distribution". *Viyodaya J Science*, 14:95–103
7. 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, [1] (http://dx.doi.org/10.1016/0167-7152(86)90044-1).
8. Choi, K.P. "On the Medians of the Gamma Distributions and an Equation of Ramanujan" (http://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.
9. Berg,Christian and Pedersen, Henrik L. "Convexity of the median in the gamma distribution" (http://arxiv.org/pdf/math/0609442.pdf).
10. W.D. Penny, [www.fil.ion.ucl.ac.uk/~wpenny/publications/densities.ps KL-Divergences of Normal, Gamma, Dirichlet, and Wishart densities]
11. Minka, Thomas P. (2002) "Estimating a Gamma distribution". http://research.microsoft.com/en-us/um/people/minka/papers/minka-gamma.pdf
12. Choi, S.C.; Wette, R. (1969) "Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias", *Technometrics*, 11(4) 683–690
13. Devroye, Luc (1986). *Non-Uniform Random Variate Generation*. New York: Springer-Verlag. ISBN 0-387-96305-7. See Chapter 9, Section 3.
14. 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.. See Algorithm GD, p. 53.
15. Ahrens, J. H.; Dieter, U. (1974). "Computer methods for sampling from gamma, beta, Poisson and binomial distributions". *Computing*. **12**: 223–246. CiteSeerX 10.1.1.93.3828. doi:10.1007/BF02293108.
16. Cheng, R.C.H., and Feast, G.M. Some simple gamma variate generators. Appl. Stat. 28 (1979), 290–295.
17. Marsaglia, G. The squeeze method for generating gamma variates. Comput, Math. Appl. 3 (1977), 321–325.
18. 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.
19. p. 43, Philip J. Boland, Statistical and Probabilistic Methods in Actuarial Science, Chapman & Hall CRC 2007
20. Aksoy, H. (2000) "Use of Gamma Distribution in Hydrological Analysis" (http://journals.tubitak.gov.tr/engineering/issues/muh-00-24-6/muh-24-6-7-9909-13.pdf), *Turk J. Engin Environ Sci*, 24, 419 – 428.
21. 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)

22. 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)
23. 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.
24. DJ Reiss, MT Facciotti and NS Baliga (2008) "Model-based deconvolution of genome-wide DNA binding" (http://bioinformatics.oxfordjournals.org/content/24/3/396.full.pdf+html), *Bioinformatics*, 24, 396–403
25. 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
26. 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).
27. Dubey, Satya D. (December 1970). "Compound gamma, beta and F distributions". *Metrika*. **16**: 27–31. doi:10.1007/BF02613934.

References

- R. V. Hogg and A. T. Craig (1978) *Introduction to Mathematical Statistics*, 4th edition. New York: Macmillan. (See Section 3.3.)'
- P. G. Moschopoulos (1985) *The distribution of the sum of independent gamma random variables*, **Annals of the Institute of Statistical Mathematics**, 37, 541–544
- A. M. Mathai (1982) *Storage capacity of a dam with gamma type inputs*, **Annals of the Institute of Statistical Mathematics**, 34, 591–597

External links

- Hazewinkel, Michiel, ed. (2001), "Gamma-distribution", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Weisstein, Eric W. "Gamma distribution". *MathWorld*.

The Wikibook *Statistics* has a page on the topic of: ***Gamma distribution***

- Engineering Statistics Handbook (http://www.itl.nist.gov/div898/handbook/eda/section3/eda366b.htm)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Gamma_distribution&oldid=753900436"

Categories: Continuous distributions | Factorial and binomial topics | Conjugate prior distributions | Exponential family distributions | Infinitely divisible probability distributions | Survival analysis

---