



The Inverse Gamma Distribution and Some of Its Applications

Ilhan M. Izmirli

Abstract: The inverse gamma distribution is of considerable theoretical and applied importance because of its distinctive positive support, right-skewed form, and capacity to model quantities for which large values are possible but increasingly unlikely. It arises naturally in a variety of contexts, including Bayesian statistics, reliability theory, stochastic modeling, finance, and the analysis of variance and scale parameters. Despite these applications and its historical significance within the Pearson system, the inverse gamma distribution receives comparatively limited attention in most - if not all - introductory probability and statistics curricula. The purpose of this paper is therefore twofold: first, to examine the historical development, mathematical properties, and applications of the inverse gamma distribution within the Pearson system; and second, to argue for its inclusion as a useful and pedagogically significant distribution in an undergraduate curriculum.

PEARSON SYSTEM OF PROBABILITY DISTRIBUTIONS

The *Pearson system of probability distributions* was introduced by Karl Pearson in 1895 as a flexible family of continuous probability distributions designed, in large part, to provide mathematical models for empirical data exhibiting varying degrees of skewness. Pearson subsequently extended and refined the system in a series of influential papers, particularly those published in 1901 and 1916, establishing one of the earliest systematic frameworks for the classification of continuous probability distributions according to their algebraic and geometric properties.

In his original 1895 paper, Pearson identified several distributional forms, designated Types I through IV, in addition to the normal distribution, which was subsequently associated with Type V in the development of the system. His classification was based principally on the nature of the support of the distribution—whether bounded, semi-infinite, or unbounded—and on the possible presence or absence of skewness. In 1901, Pearson substantially revised the classification, redefining Type V and introducing Type VI. These developments established the principal members of the Pearson system, including Types I, III, IV, V, and VI. In a subsequent paper published in 1916, Pearson introduced additional forms and subtypes, extending the system to Types VII through XII.

Several important probability distributions such as the *beta distribution*, *Cauchy distribution*, *chi-squared distribution*, *continuous uniform distribution*, *F-distribution*, and the *normal distribution* are encompassed by the Pearson system. The *gamma distribution*, to give another example, emerged from Pearson's earlier investigations and became identified with the Pearson Type III distribution. The Pearson Type IV distribution includes *Student's t-distribution* as a special case, illustrating the capacity of the Pearson system to unify distributions that later became fundamental in statistical theory. Of particular interest in the present paper is the *inverse gamma distribution*, which Pearson

incorporated into the system as the Type V distribution in his 1901 work. See Pearson (1893, 1895a, 1895b, 1902, 1905, and 1916) for more detailed information.

The Pearson system of distributions can be expressed as a family of unimodal continuous probability distribution functions that satisfy the differential equation:

$$f'(x) = -\frac{x - a_1}{b_2x^2 + b_1x + b_0}f(x)$$

The domain of the distribution is all real numbers except for those x for which $B(x) = b_2x^2 + b_1x + b_0 = 0$. It is worth noting that, since $f'(a_1) = 0$, the parameter a_1 is the stationary point of the function.

Following Pearson, we distinguish between two cases to find a closed form of the general solution

$$f(x) = C \exp\left(-\int \frac{x - a_1}{b_2x^2 + b_1x + b_0} dx\right)$$

Here, of course, C is an arbitrary constant.

Case 1. The discriminant D of $B(x) = b_2x^2 + b_1x + b_0$ is negative

In this case, we put

$$t = x + \frac{b_1}{2b_2}$$

and the quadratic expression in the denominator is transformed into

$$b_2(t^2 + \alpha^2)$$

where

$$\alpha^2 = \frac{-D}{2b_2}$$

We thus obtain

$$\begin{aligned} f(t) &= \exp\left(-\frac{1}{2}\ln(t^2 + \alpha^2) + \frac{2b_2a_1 - b_1}{2b_2\alpha}\text{ArcTan}\left(\frac{t}{\alpha}\right) + C\right) \\ &= K\left(1 + \frac{t^2}{\alpha^2}\right)^{-1/2b_2} \exp\left(\frac{2b_2\alpha - b_1}{2b_2^2\alpha}\text{ArcTan}\left(\frac{t}{\alpha}\right)\right) \end{aligned}$$

Support of the function is the entire real line. With a certain choice of normalizing constants this is the density function of the Pearson type IV and VII distributions, of which the normal distribution is a special case. The Pearson type VII distribution is equivalent to the non-standardized Student's t -distribution. For more details see Elderton (1969).

Case 2: The discriminant D of $B(x) = b_2x^2 + b_1x + b_0$ is non-negative. This has been, for obvious reasons, called the *logarithmic case*.

This case gives rise to several well-known probability distributions. For example, the beta distribution is a special instance of this case.

If $b_1^2 - 4b_2b_0 \geq 0$ the equation $B(x) = b_2x^2 + b_1x + b_0 = 0$ has two solutions a_1 and a_2

$$a_1 = \frac{-b_1 - \sqrt{b_1^2 - 4b_2b_0}}{2b_2}$$

$$a_2 = \frac{-b_1 + \sqrt{b_1^2 - 4b_2b_0}}{2b_2}$$

Let us assume $a_1 \neq a_2$. Consequently,

$$B(x) = b_2(x - a_1)(x - a_2)$$

Thus,

$$f(x) = K \exp \left(-\frac{1}{b_2} \int \frac{(x - a)}{(x - a_1)(x - a_2)} dx \right)$$

Clearly,

$$\begin{aligned} \int \frac{(x - a)}{(x - a_1)(x - a_2)} dx &= \frac{(a_1 - a)}{(a_1 - a_2)} \ln(x - a_1) - \frac{(a_2 - a)}{(a_1 - a_2)} \ln(x - a_2) + C' \\ &= \ln(x - a_1)^{\frac{a_1 - a}{a_1 - a_2}} - \ln(x - a_2)^{\frac{a_2 - a}{a_1 - a_2}} + C' \\ &= \ln \left(\frac{(x - a_1)^{\frac{a_1 - a}{a_1 - a_2}}}{(x - a_2)^{\frac{a_2 - a}{a_1 - a_2}}} \right) + C' \end{aligned}$$

Putting

$$\gamma = \frac{1}{b_2(a_1 - a_2)}$$

we obtain

$$f(x) = C \left(1 - \frac{x}{a_1} \right)^{-\gamma(a_1 - a)} \left(1 - \frac{x}{a_2} \right)^{\gamma(a_2 - a)}$$

As a special case, suppose $a_1 < 0 < a_2$. Clearly, $f(x)$ is supported on the interval $a_1 < x < a_2$. This implies,

$$0 < \frac{x - a_1}{a_2 - a_1} < 1$$

Let us put

$$t = \frac{x - a_1}{a_2 - a_1}$$

Then,

$$f(t) = C \left(\frac{a_1 - a_2}{a_1} t \right)^{\gamma(a_1 - a)} \left(\frac{a_2 - a_1}{a_2} (1 - t) \right)^{\gamma(a_2 - a)}$$

is supported on the interval $(0, 1)$. Let

$$\alpha = \gamma(a_1 - a)$$

and

$$\beta = \gamma(a_2 - a)$$

Then, with a suitable choice of the constant C ,

$$f(t) = C \frac{(a_2 - a_1)^{\beta - \alpha}}{a_1^\alpha a_2^\beta} t^\alpha (1 - t)^\beta$$

follows a beta distribution with parameters α and β .

THE INVERSE GAMMA DISTRIBUTION

We now show that, from a mathematical point of view, there is no more difficulty in presenting the inverse gamma distribution at an introductory level, then say, the beta distribution, which is a staple of probability and statistics classes of similar status: they are both simply different special cases of the Pearson system.

In the differential equation

$$f'(x) = -\frac{x - a_1}{b_2 x^2 + b_1 x + b_0} f(x)$$

Let us let $b_1 = b_0 = 0$. Then, for all $x > 0$, we have

$$\frac{f'(x)}{f(x)} = \left(-\frac{1}{b_2 x} - \frac{a_1}{b_2 x^2} \right)$$

Putting

$$\alpha = \frac{1}{b_2} - 1$$

and

$$\frac{1}{\beta} = -\frac{a_1}{b_2}$$

we get

$$\frac{f'(x)}{f(x)} = \left(-\frac{\alpha + 1}{x} + \frac{1}{\beta x^2} \right)$$

Consequently,

$$\ln(f(x)) = \ln\left(\frac{1}{x^{\alpha+1}}\right) - \frac{1}{\beta x} + C$$

that is,

$$f(x) = K x^{-(\alpha+1)} e^{-1/\beta x}$$

which for

$$K = \frac{1}{\Gamma(\alpha)\beta^\alpha}$$

gives us the inverse gamma distribution

$$f(x) = \frac{\beta^\alpha x^{-(\alpha+1)} e^{-\beta/x}}{\Gamma(\alpha)}$$

for $x > 0$.

Here, the **shape parameter** α controls the height: as α gets larger, the probability density function gets taller. The **scale parameter** β controls the spread.

This probability density function indeed defines the inverse gamma distribution, for if

$$Y \sim \text{Gamma}(\alpha, \beta),$$

then the reciprocal transformation

$$X = \frac{1}{Y}$$

would have the above density function.

One can easily compute the n th moment of this distribution.

$$\begin{aligned} E(X^n) &= \int_0^\infty x^n \frac{\beta^\alpha x^{-(\alpha+1)} e^{-\beta/x}}{\Gamma(\alpha)} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{n-\alpha-1} e^{-\beta/x} dx \end{aligned}$$

Substituting

$$u = \frac{\beta}{x}$$

we get

$$\begin{aligned} E(X^n) &= \frac{\beta^n}{\Gamma(\alpha)} \int_0^\infty u^{(\alpha-n)-1} e^{-u} du \\ &= \frac{\beta^n \Gamma(\alpha - n)}{\Gamma(\alpha)} \end{aligned}$$

Now we can determine the center and dispersion parameters of this distribution. By putting $n = 1$, we get for $\alpha > 1$

$$E(X) = \mu = \frac{\beta}{\alpha - 1}$$

and putting $n = 2$, we get for $\alpha > 2$

$$\text{Var}(X) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)}$$

Moreover, setting

$$f'(x) = 0$$

we get

$$f(x) = \frac{\beta^\alpha e^{-\beta/x}}{\Gamma(\alpha)x^{2(\alpha+1)}} x^{\alpha-1} (\beta - (\alpha + 1)x) = 0$$

implying the distribution has a mode at

$$x = \frac{\beta}{\alpha + 1}$$

Working with the inverse gamma distribution in R is also straightforward. The first step is to install the `invgamma` package, which provides functions for working with this distribution.

```
install.packages("invgamma")
```

```
library(invgamma)
```

Then,

1. To generate n random values, we use the write

```
rinvgamma(n, shape, rate)
```

2. To compute values of the probability density function, we write

```
dinvgamma(x, shape, rate)
```

3. To compute probabilities using the cumulative distribution function, we write

```
pinvgamma(q, shape, rate)
```

and

4. For the quantile function we write

```
qinvgamma(p, shape, rate)
```

For example, the following elementary R code gives us the graph of the probability density function of the inverse gamma distribution with parameters $\alpha = 4$ and $\beta = 2$:

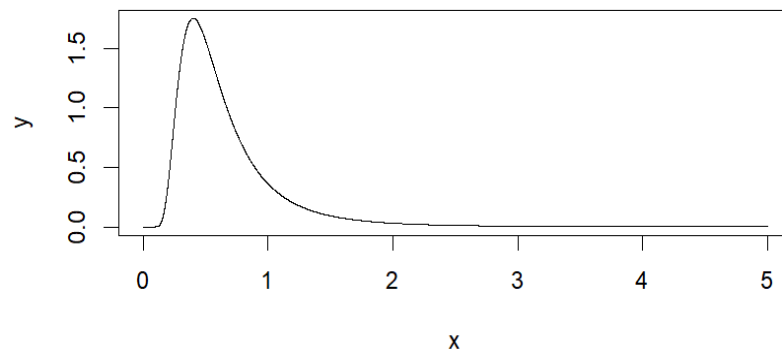
```
install.packages ("invgamma")
```

```
library (invgamma)
```

```
x<- seq(0,5, by = 0.005)
```

```
y<- dinvgamma(x, 4, 2)
```

```
plot(x, y, main = "Graph of The Inverse Gamma Distribution with Parameters 4 and 2" , type= "l")
```

Graph of The Inverse Gamma Distribution with Parameters 4 and 1

MAJOR APPLICATIONS OF THE INVERSE GAMMA DISTRIBUTION

The inverse gamma distribution is particularly suited to modeling positive, highly right-skewed quantities for which unusually large values are possible. It arises naturally as the reciprocal transformation of a gamma-distributed random variable and has important applications in *Bayesian variance modeling*, *heavy-tailed phenomena*, *reliability theory*, *hydrology*, and *stochastic volatility*, demonstrating both its theoretical importance and its practical statistical utility. In recent years, its applications in machine learning and reliability theory have also become increasingly significant. For more information see Hoff (2009) or Gelman et al (2013).

Bayesian Statistics

The inverse gamma distribution is widely used as a prior or posterior distribution for *variance parameters*. In Bayesian normal models, for example, an inverse gamma prior for the variance often leads to analytically tractable posterior distributions. It therefore plays an important role in Bayesian inference for normal and hierarchical models.

The main function of the inverse gamma distribution in Bayesian probability is its utilization as a *marginal posterior* (a way to summarize uncertain quantities) or as a *conjugate prior* (a *prior* is a probability distribution that represents our beliefs about a quantity, without taking any evidence into account). In other words, it is used to model uncertain quantities.

Financial Risk and Asset-return Modeling

The inverse gamma distribution can be used to model *stochastic volatility and variance*, particularly when the variance exhibits substantial right skewness. Since financial volatility can occasionally become exceptionally large, the heavy upper tail of the inverse gamma distribution can provide a useful representation of extreme variance.

Reliability and Survival Analysis

The inverse gamma distribution is also used in survival analysis. The distribution can model *lifetimes*, *failure rates*, and *waiting times* when the underlying phenomenon exhibits substantial positive skewness. Its heavy tail makes it especially useful when a small number of observations may have exceptionally long lifetimes.

Hydrology and Environmental Sciences

Inverse gamma models have been applied to variables associated with rainfall, precipitation, streamflow, drought severity, and extreme environmental events. These quantities are positive and frequently exhibit pronounced skewness, characteristics that are compatible with the distribution.

Rainfall data are characterized by variability across time and geographical locations. Daily precipitation measurements, for example, frequently contain many zero observations corresponding to dry days, together with positive observations that may range from light precipitation to exceptionally heavy rainfall. Consequently, rainfall data generally cannot be modeled adequately by a single continuous distribution without considering the distinction between dry and wet periods.

For positive rainfall amounts, the inverse gamma distribution may provide a useful model when the observed data exhibit pronounced right skewness and a sufficiently heavy upper tail. The shape parameter controls the rate at which the upper tail decreases, while the scale parameter determines the characteristic magnitude of rainfall amounts. Estimating these parameters from observed precipitation records allows researchers to evaluate whether the distribution provides an adequate representation of the empirical data.

Of course, several other distributions, depending on the region, observation interval, and scientific objective, can also be used to model this phenomenon: the gamma, lognormal, Weibull, and generalized gamma distributions may, at times, provide even better fits.

However, in most cases, the goal is the estimation of the probability of observing rainfall amounts above a specified threshold. Such estimates can contribute to the assessment of unusually heavy precipitation and its implications for agricultural production, drainage systems, and flood-risk management. The inverse gamma distribution's tail behavior makes it particularly indispensable in such case

A second application concerns the comparison of rainfall patterns across different locations or periods. By estimating the distributional parameters for separate datasets, researchers can investigate changes in the magnitude and variability of positive precipitation amounts. Such comparisons must account for differences in seasonality, observation practices, and climatic conditions.

The model can also be investigated computationally using R. After estimating the parameters, one can compare the fitted density with a histogram of observed positive rainfall amounts and examine goodness-of-fit measures, quantile-quantile plots, and tail probabilities. Model comparisons with competing distributions are essential, particularly when the objective is to estimate the probability of extreme precipitation.

The inverse gamma distribution provides a mathematically tractable model for positive random variables exhibiting right skewness and potentially heavy upper tails. These characteristics make it a candidate for investigating positive rainfall amounts, especially when extreme observations are of particular interest. Nevertheless, the presence of dry days, seasonal variation, and spatial dependence requires additional consideration in a comprehensive rainfall model. In practice, the inverse gamma distribution should be evaluated alongside alternative probability models using observed precipitation data and appropriate statistical diagnostics. Its principal value in rainfall research lies not in universal applicability, but in providing a flexible candidate for the statistical description of certain precipitation regimes and the analysis of their associated risks.

Signal Processing and Communications

The inverse gamma distribution can arise in models involving *random noise power*, *signal-to-noise ratios*, and *fading phenomena*. In particular, reciprocal relationships between gamma-distributed quantities naturally produce inverse gamma distributions.

Random-effects and Hierarchical Models

In hierarchical statistical models, inverse gamma distributions are frequently employed as distributions for *between-group variance components*. Their mathematical tractability makes them particularly convenient in Bayesian hierarchical formulations.

Physics and Statistical Mechanics

Inverse gamma distributions occur in certain models of fluctuating energies, diffusion processes, and anomalous stochastic phenomena. More generally, their heavy-tailed behavior makes them relevant when physical systems display substantial variability or rare extreme observations.

Population and Biological Modeling

The distribution can be useful for modeling highly skewed positive biological quantities, including certain measures of population variability, organismal lifetimes, and biological rates, particularly when extreme observations are substantially more probable than under a normal or exponential model.

REFERENCES

- Elderton, Sir W. P and N. L. Johnson (1969). *Systems of Frequency Curves*. Cambridge: Cambridge University Press.
- Gelman, Andrew; John B. Carlin; Hal S. Stern; David B. Dunson; Aki Vehtari; and Donald B. Rubin. (2013). *Bayesian Data Analysis* (Third Edition). London: Chapman & Hall/CRC.
- Hoff, Peter. (2009). *A First Course in Bayesian Statistical Methods*. New York: Springer.

Pearson, Karl (1893). Contributions to the mathematical theory of evolution. *Proceedings of the Royal Society*. 54 (326-330): 329-333.

Pearson, Karl (1895a). Contributions to the mathematical theory of evolution, II: Skew Variation in homogenous material. *Philosophical Transactions of the Royal Society*. 186: 343-414.

Pearson, Karl (1895b). Contributions to the mathematical theory of evolution, X: Supplement to a memoire on skew variation. *Philosophical Transactions of the Royal Society A*. 197 (287-299): 443-459.

Pearson, Karl (1902). On the mathematical theory of errors of judgement, with special reference to the personal equation. *Philosophical Transactions of the Royal Society A*. 198 (300-311): 235-299.

Pearson, Karl (1905). *Mathematical contributions to the theory of evolution, XIV: On the general theory of skew correlation and non-linear regression*. *Drapers' Company Research Memoirs - Biometric series*. Vol. II.

Pearson, Karl (1916). *Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoire on skew variation*. *Philosophical Transactions of the Royal Society A*. 216 (538-548): 429-457.