Eur. J. Stat. 5 (2025) 6
doi: 10.28924/ada/stat.5.6

Introducing a New Unit Gamma Distribution: Properties and Applications

Christophe Chesneau

Department of Mathematics, LMNO, University of Caen-Normandie, 14032 Caen, France
christophe.chesneau@gmail.com

Abstract. This article fills a gap in distribution theory and statistics by introducing a new, simple and intuitive two-parameter unit distribution derived from the gamma distribution. It serves as a complementary option to the existing unit gamma distribution. The main features are explored through both theoretical and practical approaches. Specifically, the shapes of the corresponding probability density and hazard rate functions are studied, an understandable stochastic comparison with the existing unit gamma distribution is provided, moments and incomplete moments are expressed, moment skewness and kurtosis are computed, random numbers are generated, and a new family of distributions is proposed. A statistical application demonstrates how the two parameters can be estimated quite effectively and the fit to a real data set is tested. It is also shown that the new distribution is able to outperform four well-known two-parameter unit distributions: the beta distribution, the Kumaraswamy distribution, the unit Weibull distribution and, more importantly, the existing unit gamma distribution. An appendix lists the main codes used in the application.

1. Introduction

1.1. **Context.** In the field of distribution theory, much effort has been devoted to the development of new and flexible distributions using various techniques. See, for example, [2,3,13,18]. A particular focus has been on lifetime distributions, which are continuous distributions with support on the interval $(0, +\infty)$. More recently, however, attention has shifted to distributions with support on the interval $(0, 1)$, known as unit distributions. This shift is largely driven by the growing need to model modern data that lie within this interval, such as rates, proportions, normalized values and percentages.

Historically, the availability of unit distributions has been limited. Classic examples include the beta distribution [16], the Kumaraswamy distribution [24], the Topp-Leone distribution [39], the unit gamma (UG) distribution [15,34], the arcsine distribution [4], and the standard two-sided power distribution [40]. As each unit distribution has its own advantages and limitations, new proposals have emerged to enrich the modeling possibilities. These include the unit Burr III distribution [30,37], the

Received: 30 Jan 2025.
Key words and phrases. gamma distribution; mathematical transformation; probability; moments; maximum likelihood estimation; real data analysis.

unit Lindley distribution [29], the unit Weibull (UW) distribution [28], the unit Gompertz distribution [27], the unit Burr XII distribution [21], the unit inverse Gaussian distribution [14], the arcsecant hyperbolic normal distribution [22], the logit-slash distribution [20], the unit power-logarithmic distribution [7], the unit half-normal distribution [5], the family of variable-power distributions [8], the unit modified Burr III distribution [17], the family of composed unit distributions [10] and the family of negation unit distributions [9].

1.2. **On the UG distribution.** The motivation for this article stems from the following observation: despite the wide panel of unit distributions, the UG distribution seems to be the only unit variant of the gamma distribution in the literature. This led us to investigate another variant, which turns out to be interesting enough to be highlighted. Before going into further detail, a review of the UG distribution is necessary. First, we define the $Gam(\alpha, \beta)$ distribution, i.e., the (classical) gamma distribution with parameters $\alpha, \beta > 0$, by the following cumulative distribution function (cdf):

$$F_\circ(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)}\gamma(\alpha, \beta x), & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{1}$$

where $\gamma(a, b) = \int_0^b x^{a-1} e^{-x} dx$ is the lower incomplete gamma function at $a, b > 0$ and $\Gamma(a) = \lim_{b \to +\infty} \gamma(a, b) = \int_0^{+\infty} x^{a-1} e^{-x} dx$ is the standard gamma function at $a > 0$. The corresponding probability density function (pdf) is obtained by differentiating $F_\circ(x)$, which leads to

$$f_\circ(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{2}$$

The gamma distribution is one of the best known lifetime distributions, extending the scope of the exponential distribution by introducing the shape parameter α. It is widely used in practice to efficiently analyze waiting times, lifetimes and other positively skewed phenomena. Among other things, it is the basis for the definition of the UG distribution. More precisely, given a random variable X with this gamma distribution, the distribution of the random variable

$$Y = e^{-X} \tag{3}$$

is with support $(0, 1)$, and it is known as the $UG(\alpha, \beta)$ distribution, i.e., the UG distribution with parameters α, β. The underlying idea is therefore to apply an exponential decay to the values of X, converting them to the interval $(0, 1)$. In this way, we also capture the key properties of the gamma distribution, such as its flexibility in shape and skewness, while adapting these properties for modeling quantities that lie within this interval. Based on the cdf of the gamma distribution in Equation (1) and the transformation in Equation (3), the cdf of the $UG(\alpha, \beta)$ distribution is defined by

$$F_\star(x) = \begin{cases} 1 - \dfrac{1}{\Gamma(\alpha)}\gamma[\alpha, -\beta\log(x)], & x \in (0,1), \\ 1, & x \geq 1, \\ 0, & x \leq 0. \end{cases} \tag{4}$$

The corresponding pdf follows by differentiating $F_\star(x)$, which leads to

$$f_\star(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\beta-1}[-\log(x)]^{\alpha-1}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{5}$$

This function can have various increasing, decreasing, constant and unimodal shapes. In addition, the corresponding hazard rate function (hrf) allows for decreasing and bathtub shapes. These features make the UG distribution a flexible alternative to other two-parameter unit distributions, and, in particular, a direct competitor to the beta and Kumaraswamy distributions. From a statistical point of view, the various methods for estimating the two parameters involved are well known. As a result, the UG distribution has been used in several areas, including the estimation of bacterial or viral densities in biological studies, the construction of regression models, and the development of control charts for rates and proportions. Discussions of these properties and applications can be found in [1, 12, 15, 26, 32, 34, 35, 38].

1.3. **Contributions.** The exponential transformation of a random variable with the gamma distribution, as described in Equation (3), is thus central to the definition of the UG distribution. Surprisingly, the current literature lacks attempts to explore alternative transformations for generating different unit versions of the gamma distribution. Classical ratio-type transformations, which have been successfully applied to other lifetime distributions, are a natural consideration. See, for example, [5, 17, 29, 30].

In this context, given a random variable X with the $Gam(\alpha, \beta)$ distribution, we thought to consider the following transformed random variable:

$$Z = \frac{1}{1+X}. \tag{6}$$

The idea is to apply a simple polynomial decay to the values of X, converting them to the interval $(0,1)$. The distribution of Z thus forms a new unit gamma distribution that follows the spirit of the UG distribution in terms of "capturing the flexibility of the gamma distribution while adapting it for modeling quantities that lie within $(0,1)$", but with a different functional structure. We call it the new unit gamma (NUG) distribution, also written as $NUG(\alpha, \beta)$ to indicate the two parameters involved.

Preliminary theoretical and practical investigations on the NUG distribution show interesting results, enough to attract specialists in distribution theory and statisticians. Among these results, we focus first on the determination of the most important functions, namely the cdf, pdf, hrf and the quantile function (qf). The shapes of the cdf, pdf and hrf are discussed mainly through graphical analysis. In particular, we show how the NUG and UG distributions are complementary in this respect. In a more theoretical work, we prove that the NUG distribution stochastically dominates the UG distribution in a first order sense. Moments, incomplete moments, measures of skewness and kurtosis are studied using formulas and numerics. The process of generating random numbers based on the NUG distribution is detailed. The theory concludes with a discussion of a new family of distributions. The statistical application of the NUG distribution is then developed through a simple and accurate estimation method. A concrete application is given by using a referenced data set in [11], on the proportion of income spent on food for a sample of households. The result obtained shows an acceptable performance, better than that of its direct two-parameter competitors, namely the beta distribution, the Kumaraswamy distribution, the UW distributions and, more interestingly, the UG distribution.

1.4. **Organization.** The contributions described above are divided into several sections, as follows: The main functions of the NUG distribution are examined in detail in Section 2. Section 3 is devoted to its technical characteristics. The statistical work is contained in Section 4. The article ends with some concluding remarks in Section 5. The codes corresponding to the data fitting performance of the NUG distribution are given in the appendix.

2. The NUG Distribution

2.1. **Determination of the cdf and graphical analysis.** The expression of the cdf of the NUG distribution is given in the proposition below.

Proposition 2.1. *Let $\alpha, \beta > 0$. The cdf of the $NUG(\alpha, \beta)$ distribution is expressed as*

$$F(x) = \begin{cases} 1 - \dfrac{1}{\Gamma(\alpha)}\gamma\left[\alpha, \beta\left(\dfrac{1}{x} - 1\right)\right], & x \in (0,1), \\ 1, & x \geq 1, \\ 0, & x \leq 0. \end{cases} \tag{7}$$

Proof. As mentioned in the introduction, given a random variable X with the $Gam(\alpha, \beta)$ distribution, the NUG distribution is defined by the distribution of

$$Z = \frac{1}{1+X}.$$

Since the support of X is $(0, +\infty)$, that of Z is $(0,1)$, from which we immediately derive $F(x) = 0$ for $x \leq 0$ and $F(x) = 1$ for $x \geq 1$. For $x \in (0,1)$, using the probability operator $\mathbb{P}$ and the cdf of the $Gam(\alpha, \beta)$ distribution given in Equation (1), we get

$$F(x) = \mathbb{P}(Z \leq x) = \mathbb{P}\left(\frac{1}{1+X} \leq x\right) = 1 - \mathbb{P}\left(X \leq \frac{1}{x} - 1\right)$$
$$= 1 - F_\circ\left(\frac{1}{x} - 1\right) = 1 - \frac{1}{\Gamma(\alpha)}\gamma\left[\alpha, \beta\left(\frac{1}{x} - 1\right)\right].$$

This completes the proof. □

As for the cdf of the UG distribution given in Equation (4), we note that the incomplete gamma function plays a key role in defining this cdf. Figure 1 shows examples of curves of this cdf with different parameter configurations. Note that the free software R is used [33], with the basic function `curve`.

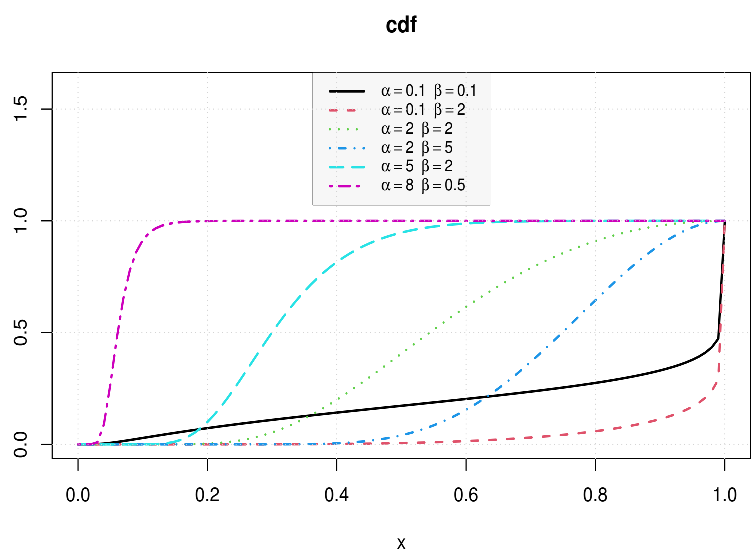


FIGURE 1. Examples of curves of the cdf of the NUG distribution

A variety of increasingly convex and concave shapes can be observed. This gives an indication of the flexibility of the NUG distribution, which will be refined by studying the shapes of the corresponding pdf and hrf.

2.2. **Determination of the pdf and graphical analysis.** The expression of the pdf of the NUG distribution is given in the proposition below.

Proposition 2.2. *Let $\alpha, \beta > 0$. The pdf of the $NUG(\alpha, \beta)$ distribution is expressed as*

$$f(x) = \begin{cases} \dfrac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x}, & x \in (0,1), \\ 0, & x \notin (0,1). \end{cases} \tag{8}$$

Proof. We proceed by differentiating the cdf of the NUG distribution determined in Equation (7). For $x \notin (0, 1)$, it is clear that $f(x) = F'(x) = 0$. For $x \in (0, 1)$, using the standard differentiation rules of the composite function, we have

$$f(x) = F'(x) = \left\{1 - \frac{1}{\Gamma(\alpha)}\gamma\left[\alpha, \beta\left(\frac{1}{x} - 1\right)\right]\right\}' = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{x^2}\left(\frac{1}{x} - 1\right)^{\alpha-1} e^{-\beta(1/x-1)}$$
$$= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x}.$$

This ends the proof. □

Compared to the pdf of the UG distribution expressed in Equation (5), the pdf of the NUG distribution deals mainly with simple power functions, which are more manageable than the power of the minus logarithmic function. We also note the presence of an exponential term, which plays a crucial role for values around $x = 0$. The pdf of the UG distribution does not have this feature.

Figure 2 shows examples of curves of this pdf with different parameter configurations.

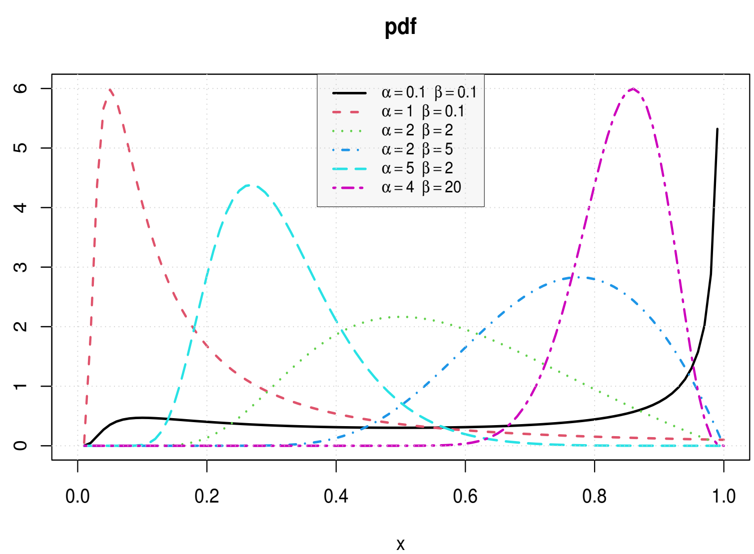


Figure 2. Examples of curves of the pdf of the NUG distribution

We observe different types of shapes, but mainly unimodal shapes. Compared to the pdf of the UG distribution, the pdf of the NUG distribution lacks decreasing shapes, but can reach a wider variety of unimodal shapes, including left-skewed, nearly symmetric and right-skewed unimodal shapes.

2.3. **Determination of the hrf and graphical analysis.** The expression of the hrf of the NUG distribution is given in the proposition below.

Proposition 2.3. *Let $\alpha, \beta > 0$. The hrf of the $NUG(\alpha, \beta)$ distribution is expressed as*

$$h(x) = \begin{cases} \beta^{\alpha} e^{\beta} \dfrac{x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x}}{\gamma\left[\alpha, \beta\left(1/x - 1\right)\right]}, & x \in (0,1), \\ 0, & x \notin (0,1). \end{cases}$$

Proof. The desired expression follows from the definition of a hrf, that is

$$h(x) = \frac{f(x)}{1 - F(x)},$$

for $x \in (0,1)$, and $h(x) = 0$ for $x \notin (0,1)$, combined with the expressions of $F(x)$ and $f(x)$ given in Equations (7) and (8), respectively. This ends the proof. □

As with the pdf, the shapes of the corresponding hrf are informative about the modeling power of a distribution. With this in mind, Figure 3 shows examples of curves of the hrf of the NUG distribution with different parameter configurations.

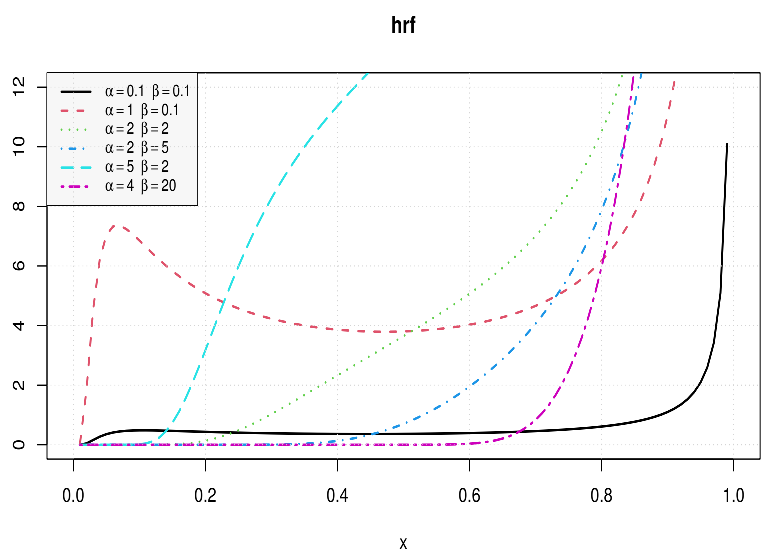


FIGURE 3. Examples of curves of the hrf of the NUG distribution

We observe that it allows increasing and bathtub shapes. This contrasts with the hrf of the UG distribution, which allows decreasing shapes in addition to bathtub shapes. Thus, the NUG and UG distributions can be seen as complementary in this respect.

2.4. **Determination of the qf.** The expression of the qf of the NUG distribution is given in the proposition below.

Proposition 2.4. *Let $\alpha, \beta > 0$. The qf of the NUG(α, β) distribution is expressed as*

$$Q(y) = \begin{cases} \dfrac{\beta}{\beta + \gamma^{-1}[\alpha, (1-y)\Gamma(\alpha)]}, & y \in (0,1), \\ 0, & y \notin (0,1), \end{cases}$$

where $\gamma^{-1}(a, b)$ is the inverse lower incomplete gamma function with respect to b, i.e., $\gamma[a, \gamma^{-1}(a, b)] = b$, at $a, b > 0$.

Proof. By the general definition of a qf, we must have $F[Q(y)] = y$ for $y \in (0, 1)$. Using the expression of $F(x)$ in Equation (7), for $y \in (0, 1)$, the following equivalences hold true:

$$\begin{aligned} F[Q(y)] = y &\Leftrightarrow 1 - \frac{1}{\Gamma(\alpha)}\gamma\left\{\alpha, \beta\left[\frac{1}{Q(y)} - 1\right]\right\} = y \\ &\Leftrightarrow \gamma\left\{\alpha, \beta\left[\frac{1}{Q(y)} - 1\right]\right\} = (1-y)\Gamma(\alpha) \\ &\Leftrightarrow \beta\left[\frac{1}{Q(y)} - 1\right] = \gamma^{-1}\left[\alpha, (1-y)\Gamma(\alpha)\right] \\ &\Leftrightarrow \frac{1}{Q(y)} = 1 + \frac{1}{\beta}\gamma^{-1}\left[\alpha, (1-y)\Gamma(\alpha)\right] \\ &\Leftrightarrow Q(y) = \frac{\beta}{\beta + \gamma^{-1}[\alpha, (1-y)\Gamma(\alpha)]}. \end{aligned}$$

This ends the proof. □

Various quantile measures can be derived from this qf. Examples include the median given by $M = Q(1/2)$, the interquartile range given by $I = Q(3/4) - Q(1/4)$, and the quartile skewness given by $K = [Q(3/4) - M]/[M - Q(1/4)]$. We can also use this qf to generate random numbers from the NUG distribution, although a more direct approach can be considered, as described in Subsection 3.5. For more details on these measures, see [19, 31].

The next section is devoted to some technical characteristics of the NUG distribution.

3. Technical Characteristics

3.1. **Stochastic comparison with the UG distribution.**

The proposition below shows a first-order stochastic dominance involving the UG and NUG distributions. The complete theory on the notion of stochastic dominance can be found in [36].

Proposition 3.1. *Let $\alpha, \beta > 0$. Then the NUG(α, β) distribution first-order stochastic dominates the UG(α, β) distribution, i.e., for any $x \in \mathbb{R}$, we have*

$$F(x) \leq F_{\star}(x),$$

where $F_{\star}(x)$ and $F(x)$ are given in Equations (4) and (7), respectively.

Proof. We consider the random variable transformations at the basis of the UG and NUG distributions. More precisely, given a random variable X with the $Gam(\alpha, \beta)$ distribution, $Y = e^{-X}$ has the $UG(\alpha, \beta)$ distribution and $Z = 1/(1+X)$ has the $NUG(\alpha, \beta)$ distribution. Using the following well-known exponential inequality: $e^x \geq 1 + x$ for $x \geq 0$ (which is in fact extended to $x \in \mathbb{R}$), we have

$$Y = e^{-X} \leq \frac{1}{1+X} = Z.$$

As a result, for any $x \in \mathbb{R}$, we have $\{Z \leq x\} \subseteq \{Y \leq x\}$, which implies that

$$F(x) = \mathbb{P}(Z \leq x) \leq \mathbb{P}(Y \leq x) = F_\star(x).$$

This ends the proof. □

This result highlights the different stochastic characteristics of the NUG and UG distributions, despite having the same support and being defined on the same gamma distribution baseline.

3.2. **Moments.** Due to the complexity of the pdf, the moments associated with the NUG distribution do not have simple analytical expressions. However, they can be expressed as specific series. This is precisely formulated in the result below.

Proposition 3.2. *Let $\alpha \in (0, +\infty) \backslash \mathbb{N}$, $\beta > 0$, $r \in \mathbb{R}$ and Z be a random variable with the $NUG(\alpha, \beta)$ distribution. Then we have*

$$\mathbb{E}(Z^r) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha - 1}{k} \beta^k \Gamma(\alpha - r - k, \beta),$$

where $\mathbb{E}$ is the expectation operator, $\Gamma(a, b) = \Gamma(a) - \gamma(a, b) = \int_b^{+\infty} x^{a-1} e^{-x} dx$ is the upper incomplete gamma function at $a, b > 0$, and $\binom{a}{k} = a(a-1)\dots(a-k+1)/k!$ is the generalized binomial coefficient at $a > 0$ and $k \in \mathbb{N}$.

For the special case where $\alpha \in \mathbb{N} \backslash \{0\}$, the following finite series expansion holds:

$$\mathbb{E}(Z^r) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{\alpha-1} (-1)^k \binom{\alpha - 1}{k} \beta^k \Gamma(\alpha - r - k, \beta).$$

Proof. Based on the pdf of the NUG distribution defined in Equation (8), the law of the unconscious statistician gives

$$\mathbb{E}(Z^r) = \int_{-\infty}^{+\infty} x^r f(x) dx = \int_0^1 x^r f(x) dx = \int_0^1 x^r \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1} (1-x)^{\alpha-1} e^{-\beta/x} dx. \quad (9)$$

For $x \in (0, 1)$ and $\alpha \in (0, +\infty) \backslash \mathbb{N}$, the generalized binomial decomposition ensures that

$$(1-x)^{\alpha-1} = \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha - 1}{k} x^k.$$

Using this and exchanging the symbols integral and sum by the Fubini theorem, we obtain

$$\begin{aligned}
&\int_0^1 x^r \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x} dx \\
&= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \int_0^1 x^{r-\alpha-1} \left[\sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} x^k \right] e^{-\beta/x} dx \\
&= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \int_0^1 x^{r-\alpha-1+k} e^{-\beta/x} dx. \qquad (10)
\end{aligned}$$

Applying the change of variables $y = \beta/x$, we get

$$\begin{aligned}
&\int_0^1 x^{r-\alpha-1+k} e^{-\beta/x} dx = \int_{+\infty}^{\beta} \left(\frac{\beta}{y}\right)^{r-\alpha-1+k} e^{-y} \left(-\frac{\beta}{y^2} dy\right) \\
&= \beta^{r-\alpha+k} \int_{\beta}^{+\infty} y^{\alpha-r-k-1} e^{-y} dy \\
&= \beta^{r-\alpha+k} \Gamma(\alpha - r - k, \beta). \qquad (11)
\end{aligned}$$

Combining Equations (9), (10) and (11), we find that

$$\begin{aligned}
\mathbb{E}(Z^r) &= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^{r-\alpha+k} \Gamma(\alpha - r - k, \beta) \\
&= \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \Gamma(\alpha - r - k, \beta).
\end{aligned}$$

For the special case where $\alpha \in \mathbb{N}\backslash\{0\}$, we can use the standard binomial decomposition instead of the generalized binomial decomposition, which gives

$$(1-x)^{\alpha-1} = \sum_{k=0}^{\alpha-1} (-1)^k \binom{\alpha-1}{k} x^k.$$

The desired finite series expression follows with the same arguments as above. This concludes the proof. □

In this proposition, we can see that $r \in \mathbb{R}$, including the inverse moments associated with the NUG distribution by taking r negative.

In the case where $\alpha \in (0, +\infty)\backslash\mathbb{N}$, we can derive an acceptable finite series approximation of the moments by replacing $+\infty$ with a large integer. More precisely, in the setting of Proposition 3.2, taking $\delta = 100$, we have

$$\mathbb{E}(Z^r) \approx \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{\delta} (-1)^k \binom{\alpha-1}{k} \beta^k \Gamma(\alpha - r - k, \beta).$$

This is of some computational interest as an alternative to integral approximation techniques.

On the other hand, the moments are also essential for defining standard measures of skewness and kurtosis. This aspect is considered in Subsection 3.4.

We end this subsection with a moment result involving the UG and NUG distributions.

Proposition 3.3. *Let $\alpha, \beta > 0$, Y be a random variable with the $UG(\alpha, \beta)$ distribution, Z be a random variable with the $NUG(\alpha, \beta)$ distribution and $g : (0, 1) \to \mathbb{R}$ be an non-decreasing function. Then we have*

$$\mathbb{E}[g(Y)] \leq \mathbb{E}[g(Z)].$$

This inequality is reversed if g is assumed to be non-increasing rather than non-decreasing.

Proof. By the definitions of the UG and NUG distributions, we can introduce a random variable X with the $Gam(\alpha, \beta)$ distribution such that $Y = e^{-X}$ and $Z = 1/(1 + X)$. Using the well-known exponential inequality: $e^x \geq 1 + x$ for $x \geq 0$, we have

$$Y = e^{-X} \leq \frac{1}{1 + X} = Z.$$

It follows from the non-decreasing property of g that $g(Y) \leq g(Z)$, and by taking the expectation on both sides, we get

$$\mathbb{E}[g(Y)] \leq \mathbb{E}[g(Z)].$$

Obviously this inequality is reversed if g is assumed to be non-increasing instead of non-decreasing. This completes the proof. □

This result shows that some moment measures or functions of the UG and NUG distributions can be compared. For example, in the framework of Proposition 3.3, we have $\mathbb{E}(Y^2) \leq \mathbb{E}(Z^2)$ and, for any $s \geq 0$, $\mathbb{E}(e^{-sZ}) \leq \mathbb{E}(e^{-sY})$, which is an inequality involving the Laplace transform of the UG and NUG distributions at s.

3.3. **Incomplete moments.** The series methodology developed for the moments associated with the NUG distribution can be extended to the incomplete moments, as shown below.

Proposition 3.4. *Let $\alpha \in (0, +\infty)\backslash\mathbb{N}$, $\beta > 0$, $r \in \mathbb{R}$, $\epsilon \in (0, 1)$ and Z be a random variable with the $NUG(\alpha, \beta)$ distribution. Then we have*

$$\mathbb{E}(Z^r 1_{\{Z \leq \epsilon\}}) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha - 1}{k} \beta^k \Gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right),$$

where $1_{\{Z \leq \epsilon\}} = 1$ if the event $\{Z \leq \epsilon\}$ is realized, otherwise $1_{\{Z \leq \epsilon\}} = 0$. For the special case where $\alpha \in \mathbb{N}\backslash\{0\}$, the following finite series expansion holds:

$$\mathbb{E}(Z^r 1_{\{Z \leq \epsilon\}}) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{\alpha-1} (-1)^k \binom{\alpha - 1}{k} \beta^k \Gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right).$$

Proof. The proof follows the line of that in Proposition 3.2. However, for the sake of completeness, we explain it in detail below. Considering that $\epsilon \in (0,1)$, the law of the unconscious statistician gives

$$\mathbb{E}(Z^r 1_{\{Z\leq\epsilon\}}) = \int_{-\infty}^{\epsilon} x^r f(x)dx = \int_{0}^{\epsilon} x^r f(x)dx = \int_{0}^{\epsilon} x^r \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x}dx. \quad (12)$$

It follows from the generalized binomial decomposition that, for $x \in (0,\epsilon) \subseteq (0,1)$ and $\alpha \in (0,+\infty)\backslash\mathbb{N}$,

$$(1-x)^{\alpha-1} = \sum_{k=0}^{+\infty}(-1)^k \binom{\alpha-1}{k} x^k.$$

Using this and exchanging the symbols integral and sum by the Fubini theorem, we obtain

$$\begin{aligned}
&\int_{0}^{\epsilon} x^r \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x^{-\alpha-1}(1-x)^{\alpha-1} e^{-\beta/x}dx \\
&= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \int_{0}^{\epsilon} x^{r-\alpha-1}\left[\sum_{k=0}^{+\infty}(-1)^k \binom{\alpha-1}{k} x^k\right] e^{-\beta/x}dx \\
&= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty}(-1)^k \binom{\alpha-1}{k} \int_{0}^{\epsilon} x^{r-\alpha-1+k} e^{-\beta/x}dx. \quad (13)
\end{aligned}$$

Applying the change of variables $y = \beta/x$, we get

$$\begin{aligned}
&\int_{0}^{\epsilon} x^{r-\alpha-1+k} e^{-\beta/x}dx = \int_{+\infty}^{\beta/\epsilon} \left(\frac{\beta}{y}\right)^{r-\alpha-1+k} e^{-y}\left(-\frac{\beta}{y^2}dy\right) \\
&= \beta^{r-\alpha+k}\int_{\beta/\epsilon}^{+\infty} y^{\alpha-r-k-1}e^{-y}dy \\
&= \beta^{r-\alpha+k}\Gamma\left(\alpha-r-k,\frac{\beta}{\epsilon}\right). \quad (14)
\end{aligned}$$

Combining Equations (12), (13) and (14), we establish that

$$\begin{aligned}
\mathbb{E}(Z^r 1_{\{Z\leq\epsilon\}}) &= \frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty}(-1)^k \binom{\alpha-1}{k} \beta^{r-\alpha+k}\Gamma\left(\alpha-r-k,\frac{\beta}{\epsilon}\right) \\
&= \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty}(-1)^k \binom{\alpha-1}{k} \beta^{k}\Gamma\left(\alpha-r-k,\frac{\beta}{\epsilon}\right).
\end{aligned}$$

For the special case where $\alpha \in \mathbb{N}\backslash\{0\}$, using the standard binomial decomposition instead of the generalized binomial decomposition gives the claimed formula. This ends the proof. □

Clearly, if we apply $\epsilon \to 1$, Proposition 3.4 becomes Proposition 3.2; it thus can be viewed as a generalization.

Like for the moments, in the case where $\alpha \in (0,+\infty)\backslash\mathbb{N}$, we can derive an acceptable finite series approximation of the incomplete moments by replacing $+\infty$ with a large integer. More precisely, in the setting of Proposition 3.4, taking $\delta = 100$, we have

$$\mathbb{E}(Z^r 1_{\{Z\leq\epsilon\}}) \approx \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{\delta} (-1)^k \binom{\alpha-1}{k} \beta^k \Gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right).$$

A result about complementary incomplete moments, which is a consequence of Propositions 3.2 and 3.4, concludes this subsection.

Proposition 3.5. *Let $\alpha \in (0,+\infty)\backslash\mathbb{N}$, $\beta > 0$, $r \in \mathbb{R}$, $\epsilon \in (0,1)$ and Z be a random variable with the $NUG(\alpha,\beta)$ distribution. Then we have*

$$\mathbb{E}(Z^r 1_{\{Z>\epsilon\}}) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right).$$

For the special case where $\alpha \in \mathbb{N}\backslash\{0\}$, the following finite series expansion holds:

$$\mathbb{E}(Z^r 1_{\{Z>\epsilon\}}) = \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{\alpha-1} (-1)^k \binom{\alpha-1}{k} \beta^k \gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right).$$

Proof. We clearly have

$$\mathbb{E}(Z^r 1_{\{Z>\epsilon\}}) = \mathbb{E}(Z^r) - \mathbb{E}(Z^r 1_{\{Z\leq\epsilon\}}).$$

It follows from this, Propositions 3.2 and 3.4, and $\gamma(a,b) = \Gamma(a) - \Gamma(a,b)$ that

$$\begin{aligned}
\mathbb{E}(Z^r 1_{\{Z>\epsilon\}}) &= \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \Gamma(\alpha - r - k) \\
&\quad - \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \Gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right) \\
&= \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \left[\Gamma(\alpha - r - k) - \Gamma\left(\alpha - r - k, \frac{\beta}{\epsilon}\right)\right] \\
&= \frac{\beta^r e^\beta}{\Gamma(\alpha)} \sum_{k=0}^{+\infty} (-1)^k \binom{\alpha-1}{k} \beta^k \gamma(\alpha - r - k).
\end{aligned}$$

The finite series formula can be derived in a similar way. This concludes the proof. □

Such incomplete moments are useful for characterizing partial properties of the NUG distribution, such as conditional means or truncated moments. We do not develop this aspect further.

3.4. **Moments skewness and kurtosis.** Even if we already have some knowledge about the skewness and kurtosis of the NUG distribution thanks to the shape analysis of its pdf, it can be interesting to have some numerical benchmarks on these aspects. We can therefore study some moment measures associated with the NUG distribution, such as the moment skewness and moment kurtosis. Given a random variable Z with the $NUG(\alpha,\beta)$ distribution, the moment skewness is defined by

$$C(Z) = \mathbb{E}\left\{\left[\frac{Z - \mathbb{E}(Z)}{\sigma(Z)}\right]^3\right\},$$

where

$$\sigma(Z) = \sqrt{\mathbb{E}\{[Z - \mathbb{E}(Z)]^2\}}.$$

It can be interpreted as follows: if $C(Z)$ is greater than 0, the NUG distribution is positively skewed, if $C(Z)$ is less than 0, the NUG is negatively skewed, and if $C(Z) \approx 0$, the NUG distribution is nearly symmetric.

In addition to the moment skewness, the moment kurtosis is defined by

$$D(Z) = \mathbb{E}\left\{\left[\frac{Z - \mathbb{E}(Z)}{\sigma(Z)}\right]^4\right\}.$$

It can be interpreted as follows: with the usual reference to the standard normal distribution as a benchmark, if $D(Z)$ is greater than 3, the NUG distribution is leptokurtic, if $D(Z)$ is less than 3, the NUG distribution is platykurtic, and if $D(Z) \approx 3$, the NUG distribution is mesokurtic.

With a view to analyzing the skewness and kurtosis of the NUG distribution, Table 1 presents the values of $\mathbb{E}(Z)$, $\mathbb{E}(Z^2)$, $\sigma(Z)$, $C(Z)$ and $D(Z)$ with different parameter configurations. The software R is used with the basic function `integrate`.

Table 1. Numerical values of $\mathbb{E}(Z)$, $\mathbb{E}(Z^2)$, $\sigma(Z)$, $C(Z)$ and $D(Z)$, where Z is a random variable with the $NUG(\alpha, \beta)$ distribution with different parameter configurations

	$E(Z)$	$E(Z^2)$	$\sigma(Z)$	$C(Z)$	$D(Z)$
$\alpha = 0.1,\ \beta = 0.3$	0.887	0.832	0.213	-2.059	6.176
$\alpha = 1,\ \beta = 0.5$	0.461	0.269	0.237	0.481	2.214
$\alpha = 0.2,\ \beta = 1$	0.891	0.822	0.169	-1.842	5.692
$\alpha = 2,\ \beta = 1$	0.404	0.193	0.173	0.715	3.005
$\alpha = 1,\ \beta = 2$	0.723	0.555	0.180	-0.417	2.298
$\alpha = 3,\ \beta = 8$	0.744	0.564	0.106	-0.285	2.654
$\alpha = 4,\ \beta = 6$	0.622	0.400	0.115	0.056	2.596
$\alpha = 8,\ \beta = 3$	0.291	0.090	0.075	0.766	3.840

From this table, we can see that $C(Z)$ can be greater than 0, less than 0 or approximately 0, meaning that the NUG distribution can have all skewness states. Furthermore, $D(Z)$ can be greater than 3, less than 3 or approximately 3, meaning that the NUG distribution can have all kurtosis states. This highlights the flexibility of the NUG distribution in these aspects.

3.5. **Random number generation.** Thanks to the stochastic structure of the NUG distribution, generating random numbers from a random variable Z with the $NUG(\alpha, \beta)$ distribution is straightforward. If n is the desired value, the process is as follows:

(1) Fix $\alpha, \beta > 0$.

(2) Generate n random numbers from a random variable X with the $Gam(\alpha, \beta)$ distribution, say

$$x_1, \ldots, x_n.$$

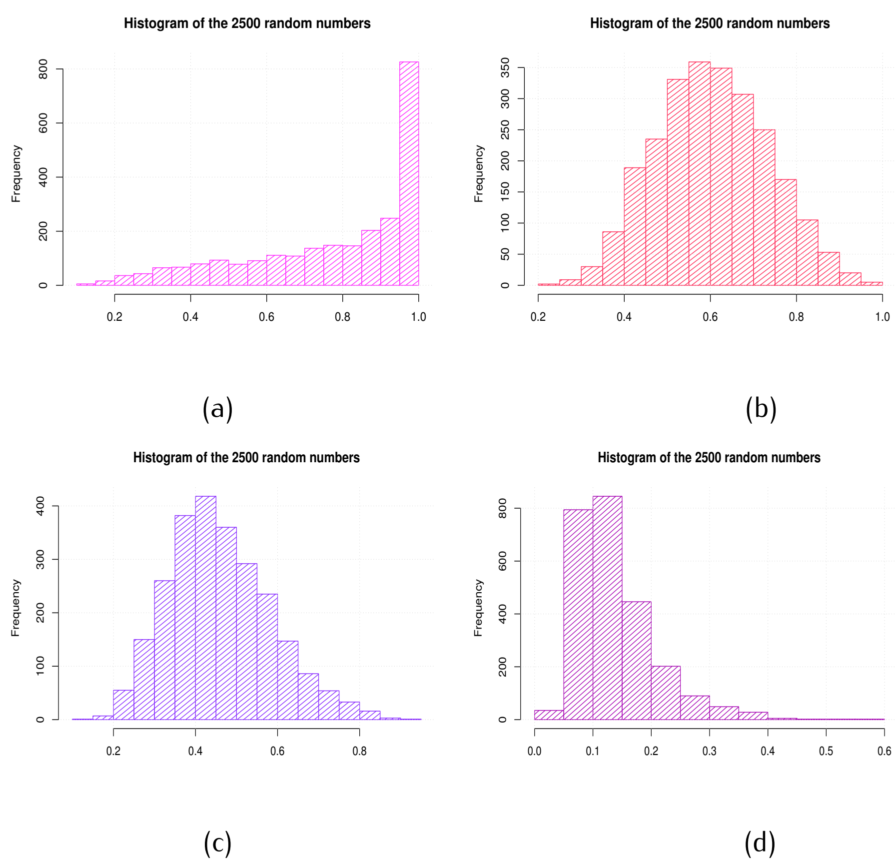
For example, with the software R, we can use the basic function `rgamma`.

(3) For any $i = 1, \ldots, n$, calculate

$$z_i = \frac{1}{1 + x_i}.$$

(4) The desired numbers are $z_1, \ldots, z_n$.

We now make a simple graphical representation of this process. We generate four samples of $n = 2500$ random numbers from Z, each taken under a particular parameter configuration, and plot the corresponding frequency histograms in Figure 4.



(a) (b)

(c) (d)

FIGURE 4. Frequency histograms of four sample of $n = 2500$ random numbers generated from a random variable Z with the $NUG(\alpha, \beta)$ distribution for the following parameter configurations: (a) $\alpha = 0.4$ and $\beta = 0.8$, (b) $\alpha = 3$ and $\beta = 4$, (c) $\alpha = 4$ and $\beta = 3$, and (d) $\alpha = 4$ and $\beta = 0.5$

As expected, the general shapes of the frequency histograms correspond to those observed for the pdf of the NUG distribution in Figure 1.

3.6. **NUG family of distributions.** With a simple composed scheme, the NUG distribution can be used to generate various kinds of distributions, with different supports. Given a baseline continuous distribution with the cdf $G(x)$, $x \in \mathbb{R}$, using the cdf of the NUG distribution determined in Equation (7), we define the NUG family of distributions by the following composite cdf:

$$\begin{aligned} F_{fam}(x) &= F[G(x)] \\ &= 1 - \frac{1}{\Gamma(\alpha)}\gamma\left[\alpha, \beta\left(\frac{1}{G(x)} - 1\right)\right], \quad x \in \mathbb{R}. \end{aligned}$$

Denoting $g(x)$ the pdf associated with $G(x)$ and using the pdf of the NUG distribution given in Equation (8), the pdf of the NUG family of distributions is expressed as

$$\begin{aligned} f_{fam}(x) &= g(x)f[G(x)] \\ &= g(x)\frac{\beta^\alpha e^\beta}{\Gamma(\alpha)}G(x)^{-\alpha-1}(1-x)^{\alpha-1}e^{-\beta/G(x)}, \quad x \in \mathbb{R}. \end{aligned}$$

We briefly specify this family by considering the standard normal distribution for the baseline, i.e., with the integral form cdf given by

$$G(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt, \quad x \in \mathbb{R}$$

and the corresponding pdf expressed as

$$g(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R}.$$

The NUG normal distribution is thus defined by the following pdf:

$$f_\dagger(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\frac{\beta^\alpha e^\beta}{\Gamma(\alpha)}\Phi(x)^{-\alpha-1}\left[1-\Phi(x)\right]^{\alpha-1}e^{-\beta/\Phi(x)}, \quad x \in \mathbb{R}.$$

By construction, it forms a new skewed version of the normal distribution. Figure 5 shows example curves of this pdf with different parameter configurations.

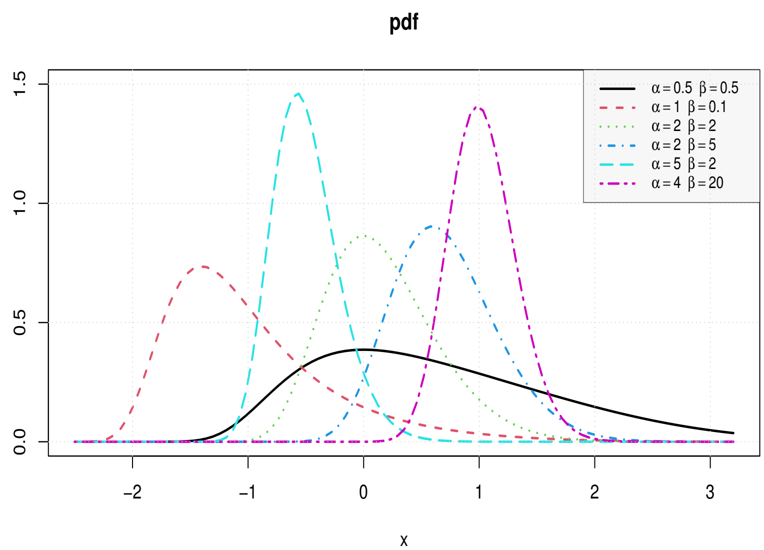


Figure 5. Sample of curves of the pdf of the NUG normal distribution

Different skewed shapes are observed, especially right-skewed shapes with different degrees of skewness.

The different types of distributions that can generate the NUG distribution deserve a full study, which we leave to future work.

The next section is devoted to the statistical application of the NUG distribution.

4. Statistical application

This section is devoted to the statistical application of the NUG distribution, with emphasis on its remarkable modeling accuracy.

4.1. **Estimation method.** The maximum likelihood (ML) is one of the best known parametric estimation methods. Its theory and practice are fully understood, which guarantees its effectiveness in most statistical scenarios [6]. In the context of the $NUG(\alpha, \beta)$ distribution, under the assumption that α and β are unknown, the ML method is described below. Let n be the number of data and $x_1, \ldots, x_n$ be data that lie in the interval $(0, 1)$ and that are assumed to be possibly in distributional adequacy with the NUG distribution. Based on the pdf given in Equation (8), the ML estimates of α and β are given by the following "argmaxima":

$$(\hat{\alpha}, \hat{\beta}) = \text{argmax}_{(\alpha,\beta)\in(0,+\infty)^2} \sum_{i=1}^{n} \log[f(x_i)]$$
$$= \text{argmax}_{(\alpha,\beta)\in(0,+\infty)^2} \sum_{i=1}^{n} \log\left[\frac{\beta^\alpha e^\beta}{\Gamma(\alpha)} x_i^{-\alpha-1}(1-x_i)^{\alpha-1} e^{-\beta/x_i}\right].$$

Note that the logarithmic term can be developed further, but is not of interest here; after this effort in a parallel work, it appears that the ML estimates $\hat{\alpha}$ and $\hat{\beta}$ have no closed form expression. However, they can be calculated using scientific software such as R and the basic function `nlminb`. Once these estimates are obtained, we can derive the estimated pdf by a classical substitution approach, that is

$$\hat{f}(x) = \frac{\hat{\beta}^{\hat{\alpha}} e^{\hat{\beta}}}{\Gamma(\hat{\alpha})} x^{-\hat{\alpha}-1}(1-x)^{\hat{\alpha}-1} e^{-\hat{\beta}/x} \tag{15}$$

for $x \in (0,1)$. This function can be thought of as the best fit that the pdf of the NUG distribution can make to the data under consideration. Ideally, it should efficiently fit the shape of the corresponding normalized histogram.

4.2. **Distribution comparison.** In this study, given a data set, two different criteria are considered to compare different distributions. These are the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), which are briefly discussed below in the context of the NUG distribution.

The AIC is given by

$$AIC = 2k - 2\sum_{i=1}^{n} \log\left[\hat{f}(x_i)\right],$$

where $\hat{f}(x)$ is given in Equation (15) and k is the number of parameters (in this case, $k = 2$). The AIC evaluates the quality of the distribution by balancing the goodness of fit with the complexity of the distribution. A lower AIC value indicates a better-fitting distribution.

The BIC is given by

$$BIC = k\log(n) - 2\sum_{i=1}^{n} \log\left[\hat{f}(x_i)\right].$$

It is similar to the AIC but imposes a stronger penalty on complex distributions. While both AIC and BIC assess the fit and complexity of the distribution, the BIC is more stringent when it comes to penalizing the inclusion of additional parameters. As with the AIC, a lower BIC value indicates a better-fitting distribution.

We also consider four famous unit distributions as competitors, also defined with two parameters $\alpha, \beta > 0$: the beta, Kumaraswamy, UW and UG distributions, which are briefly described below.

The beta distribution is defined by the following pdf:

$$f_V(x) = \begin{cases} \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & x \in (0,1), \\ 0, & x \notin (0,1), \end{cases}$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ is the standard beta function. The beta distribution is one of the most commonly used unit distributions. See also the book [16], which gives a complete overview of its theory and applications.

The Kumaraswamy distribution is defined by the following pdf:

$$f_{\triangle}(x) = \begin{cases} \alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}, & x \in (0,1), \\ 0, & x \notin (0,1). \end{cases}$$

It can be presented as a more manageable functional alternative to the pdf of the beta distribution. The details of this distribution can be found in [24].

The UW distribution is defined by the following pdf:

$$f_{\diamond}(x) = \begin{cases} \dfrac{\alpha}{\beta^{\alpha}} \dfrac{1}{x}[-\log(x)]^{\alpha-1} e^{-[-\log(x)/\beta]^{\alpha}}, & x \in (0,1), \\ 0, & x \notin (0,1). \end{cases}$$

This is a slightly modified parameter version of [28], more closely related to the pdf implemented in the function `dweibull` of R.

Finally, we recall that the UG distribution is defined by the pdf in Equation (5).

The remainder of this section is devoted to the concrete application of the NUG distribution in fitting a real data set, and its fitting comparisons with the above competitors.

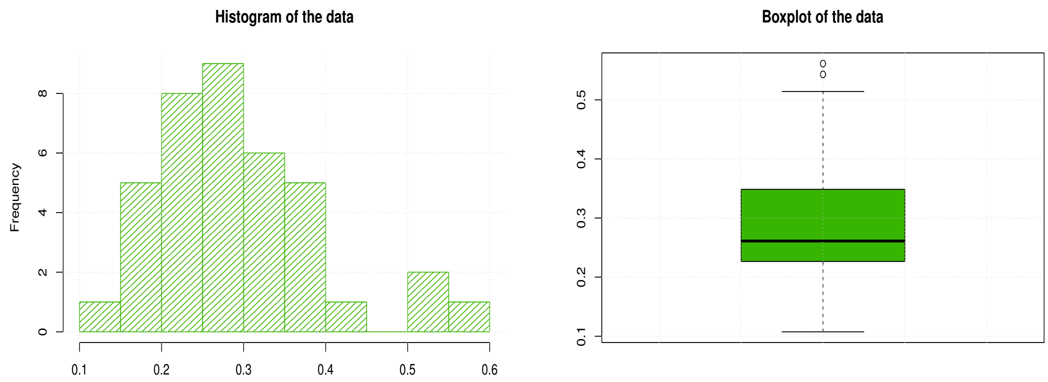
4.3. **Real data analysis.** We consider a data set derived from [11], described as the proportion of income spent on food for 38 households in a large US city. It is freely available in the R package `betareg` and can be retrieved precisely by the following commands involving a division operation: `FoodExpenditure$food / FoodExpenditure$income`. This data set has also been used in [10] for a similar purpose.

First, a summary of the data is given in Table 2.

Minimum	First quartile	Median	Mean	Third quartile	Maximum
0.1075	0.2269	0.2611	0.2897	0.3469	0.5612

Table 2. Basic summary of the consided data

We also represent the frequency histogram of the data and the corresponding boxplot in Figure 6.



(a) (b)

Figure 6. (a) Frequency histogram and (b) boxplot of the data

From this figure, we can see that the data are right skewed and have two values that can be considered as outliers. Given its shape, the NUG distribution is a candidate for a two-parameter unit distribution that can fit them.

After carrying out the ML estimation of the parameters of the unit distributions under consideration and the calculation of their respective AIC and BIC, we summarize the results obtained in Table 3.

Distribution	$\hat{\alpha}$	$\hat{\beta}$	AIC	BIC
NUG	4.543241	1.578676	-68.271292	-64.996120
UW	4.128786	1.424609	-67.868722	-64.593549
UG	13.26089	10.22977	-67.12632	-63.85115
beta	6.07164	14.82210	-66.69289	-63.41772
Kumaraswamy	2.954554	26.965413	-62.978199	-59.703027

Table 3. Parameter estimates, AIC, and BIC for five different two-parameter unit distributions, including the NUG distribution

With the lowest AIC and BIC, the NUG distribution can be considered the best for fitting. It is followed by the UW distribution, which is known to be particularly efficient for such fitting exercises [28]. Furthermore, to highlight the significance of the results, we mention that the NUG distribution can outperform certain three-parameter unit distributions. In particular, it is indicated in [10, Table 2] that, for the same data set, the I-UDa distribution has an AIC of -67.337 and the II-UDa distribution has an AIC of -67.400, both of which are greater than that of the NUG distribution, i.e., $AIC = -68.271292$ (and also greater than that of the UW distribution, i.e., $AIC = -67.868722$).

We also note that, for the NUG distribution, the ML estimates of α and β are given by

$$\hat{\alpha} = 4.543241, \quad \hat{\beta} = 1.578676.$$

As a result, the corresponding estimated pdf is

$$\hat{f}(x) = \frac{1.578676^{4.543241} e^{1.578676}}{\Gamma(4.543241)} x^{-4.543241-1}(1-x)^{4.543241-1} e^{-1.578676/x}$$

for $x \in (0,1)$, with $\Gamma(4.543241) \approx 12.35455$.

To visualize the efficiency of our approach, the curve of this estimated pdf is plotted over the normalized histogram of the data in Figure 7.

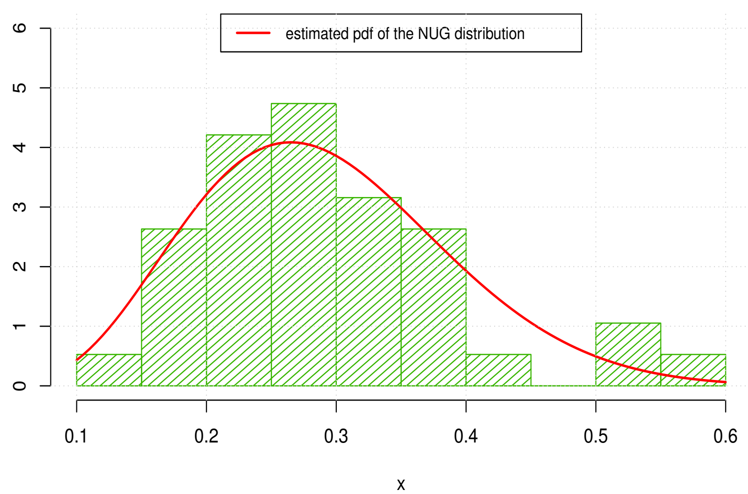


Figure 7. Curve of the estimated pdf of the NUG distribution over the normalized histogram of the data

We can see that the red curve captures well the overall shape of the histogram, including the two outliers. This simple but significant data analysis shows that the NUG distribution should be considered among the notable two-parameter unit distributions. It clearly has potential for further applications in various areas of statistics.

To ensure reproducibility, we conclude this section by noting that the main codes used for this analysis are provided in the appendix.

5. Conclusion

This article shows that much remains to be done in the theory and practice of unit distributions, using simple transformations and the well-known lifetime distributions. This claim has been illustrated by considering the NUG distribution, which is a new and simple two-parameter unit variant of the gamma distribution. We have compared it with the existing unit gamma distribution, i.e., the UG distribution, to see how they complement each other. The theoretical properties of the NUG have been examined in detail with formulas for the main functions, moments and various stochastic properties. A random number generation process is presented and validated with a numerical study, also supported graphically. The theory concludes with a brief presentation of the NUG family of distributions. Statistical analysis on a famous real data set shows that the NUG distribution fits better than four famous two-parameter unit distributions, namely the beta distribution, the Kumaraswamy distribution, the unit Weibull distribution and, more importantly, the existing unit gamma distribution.

The logical perspectives of this work include the points below.

- A more detailed study of the NUG family of distributions, together with applications of some members of this family in data analysis.
- A possible regression model using the NUG distribution as the response variable in a generalized regression framework, assessing its fit relative to traditional unit distributions based on real-world proportional data.
- The extension of the NUG distribution to allow for more flexible structures, such as the inclusion of shape parameters aiming to provide greater control over skewness and kurtosis.
- The investigation of potential copula models using the NUG distribution, exploring its applicability in modeling dependence structures in multivariate data.
- The use of the NUG distribution in machine learning, particularly in probabilistic modeling, generative adversarial networks (GANs), and neural network-based density estimation techniques.

All these aspects need to be further explored to improve the theoretical understanding and practical applicability of the NUG distribution in statistical modeling and data analysis.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] M.F. Akram, S. Ali, I. Shah, G. Marcon, Unit Interval Time and Magnitude Monitoring Using Beta and Unit Gamma Distributions, J. Math. 2022 (2022), 7951748. https://doi.org/10.1155/2022/7951748.

[2] M.A. Aljarrah, C. Lee, F. Famoye, On Generating T-X Family of Distributions Using Quantile Functions, J. Stat. Distrib. Appl. 1 (2014), 24. https://doi.org/10.1186/2195-5832-1-2

[3] A. Alzaatreh, C. Lee, F. Famoye, A New Method for Generating Families of Continuous Distributions, Metron 71 (2013), 63–79. https://doi.org/10.1007/s40300-013-0007-y.

[4] B.C. Arnold, R.A. Groeneveld, Some Properties of the Arcsine Distribution, J. Amer. Stat. Assoc. 75 (1980), 173–175. https://doi.org/10.1080/01621459.1980.10477449.

[5] H.S. Bakouch, A.S. Nik, A. Asgharzadeh, H.S. Salinas, A Flexible Probability Model for Proportion Data: Unit-Half-Normal Distribution, Communications in Statistics: Case Studies, Data Anal. Appl. 7 (2021), 271–288. https://doi.org/10.1080/23737484.2021.1882355.

[6] G. Casella, R.L. Berger, Statistical Inference, Duxbury Press, Pacific Grove, CA, (2002).

[7] C. Chesneau, Study of a Unit Power-Logarithmic Distribution, Open J. Math. Sci. 5 (2021), 218–235. https://doi.org/10.30538/oms2021.0159.

[8] C. Chesneau, a Collection of New Variable-Power Parametric Cumulative Distribution Functions for $(0,1)$-Supported Distributions, Res. Com. Math. Math. Sci. 15 (2023), 89-152.

[9] C. Chesneau, Negation-Type Unit Distributions: Concept, Theory and Examples, Math. Pannon. 30 (2024), 191–212. https://doi.org/10.1556/314.2024.00018.

[10] F. Condino, F. Domma, Unit Distributions: A General Framework, Some Special Cases, and the Regression Unit-Dagum Models, Mathematics 11 (2023), 2888. https://doi.org/10.3390/math11132888.

[11] F. Cribari-Neto, A. Zeileis, Beta Regression in R, J. Stat. Softw. 34 (2010), 1-24. https://doi.org/10.18637/jss.v034.i02

[12] S. Dey, F.B. Menezes, J. Mazucheli, Comparison of Estimation Methods for Unit-Gamma Distribution, J. Data Sci. 17 (2021), 768–801. https://doi.org/10.6339/JDS.201910_17(4).0009.

[13] N. Eugene, C. Lee, F. Famoye, Beta-Normal Distribution and Its Applications, Commun. Stat. - Theory Methods 31 (2002), 497–512. https://doi.org/10.1081/STA-120003130.

[14] M.E. Ghitany, J. Mazucheli, A.F.B. Menezes, F. Alqallaf, The Unit-Inverse Gaussian Distribution: A New Alternative to Two-Parameter Distributions on the Unit Interval, Commun. Stat. - Theory Methods 48 (2019), 3423–3438. https://doi.org/10.1080/03610926.2018.1476717.

[15] A. Grassia, On a Family of Distributions With Argument Between 0 and 1 Obtained by Transformation of the Gamma Distribution and Derived Compound Distributions, Austrian J. Stat. 19 (1977), 108-114. https://doi.org/10.1111/j.1467-842X.1977.tb01277.x.

[16] A.K. Gupta, S. Nadarajah, Handbook of Beta Distribution and Its Applications, Marcel Dekker, New York, USA, (2004).

[17] M.A. Haq, S. Hashmi, K. Aidi, P.F.L. Ramos, Unit Modified Burr-Iii Distribution: Estimation, Characterizations and Validation Test, Ann. Data Sci. 10 (2023), 415-449. https://doi.org/10.1007/s40745-020-00298-6.

[18] M. Jones, Families of Distributions Arising from the Distributions of Order Statistics, Test 13 (2004) 1-43. https://doi.org/10.1007/BF02602999.

[19] J.F. Kenney, E.S. Keeping, Mathematics of Statistics, Princeton, New Jersey, USA, (1962).

[20] M. Korkmaz, A New Heavy-Tailed Distribution Defined on the Bounded Interval: The Logit Slash Distribution and Its Applications, J. Appl. Stat. 473 (2019), 2097-2119. https://doi.org/10.1080/02664763.2019.1704701.

[21] M. Korkmaz, C. Chesneau, On the Unit Burr-XII Distribution With the Quantile Regression Modeling and Applications, Comput. Appl. Math. 40 (2021), 1-26. https://doi.org/10.1007/s40314-021-01418-5.

[22] M. Korkmaz, C. Chesneau, Z. Korkmaz, On the Arcsecant Hyperbolic Normal Distribution. Properties, Quantile Regression Modeling and Applications, Symmetry 13 (2021), 117. https://doi.org/10.3390/sym13010117.

[23] S. Kotz, J.R. Van Dorp, Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications, World Scientific Publishing, Singapore, (2004).

[24] P. Kumaraswamy, A Generalized Probability Density Function for Double-Bounded Random Processes, J. Hydrol. 46 (1980), 79-88. https://doi.org/10.1016/0022-1694(80)90036-0.

[25] A.W. Marshall, I. Olkin, Life Distributions, Springer, New York, USA, (2007).

[26] J. Mazucheli, A.F.B. Menezes, S. Dey, Improved Maximum-Likelihood Estimators for the Parameters of the Unit-Gamma Distribution, Commun. Stat. - Theory Methods 47 (2018), 3767-3778. https://doi.org/10.1080/03610926.2017.1361993.

[27] J. Mazucheli, A.F.B. Menezes, S. Dey, Unit-Gompertz Distribution with Applications, Statistica 79 (2019), 26-43. https://doi.org/10.6092/issn.1973-2201/8497.

[28] J. Mazucheli, A. Menezes, M. Ghitany, The Unit-Weibull Distribution and Associated Inference, J. Appl. Probab. Stat. 13 (2018), 1-22.

[29] J. Mazucheli, A. Menezes, S. Chakraborty, On the One Parameter Unit-Lindley Distribution and Its Associated Regression Model for Proportion Data, J. Appl. Stat. 46 (2019), 700-714. https://doi.org/10.1080/02664763.2018.1511774.

[30] K. Modi, V. Gill, Unit Burr III Distribution with Application, J. Stat. Manag. Syst. 23 (2020), 579-592. https://doi.org/10.1080/09720510.2019.1646503.

[31] J.J. Moors, A Quantile Alternative for Kurtosis, J. R. Stat. Soc. D, 37 (1998), 25-32. https://doi.org/10.2307/2348376.

[32] A.M. Mousa, A.A. El-Sheikh, M.A. Abdel-Fattah, A Gamma Regression for Bounded Continuous Variables, Adv. Appl. Stat. 49 (2016), 305–326. https://doi.org/10.17654/AS049040305.

[33] R Core Team, R: A Language and Environment for Statistical Computing [Internet], Vienna, Austria, (2016). https://www.R-project.org.

[34] M.V. Ratnaparkhl, J.E. Mosimann, On the Normality of Transformed Beta and Unit-Gamma Random Variables, Commun. Stat. - Theory Methods 19 (1990), 3833-3854. https://doi.org/10.1080/03610929008830416.

[35] E.O. Rochá, C.L.N. Azevedo, J.M.A. Mota, M.J. Batista, J.S. Nobre, Bayesian Inference for Unit Gamma Distribution, Rev. Cad. Pedagóg. 21 (2024), e7690. https://doi.org/10.54033/cadpedv21n9-101.

[36] M. Shaked, J.G. Shanthikumar, Stochastic Orders, Springer Verlag, New York, USA, (2007).

[37] D.P. Singh, M. Jha, Y. Tripathi, L. Wang, Reliability Estimation in a Multicomponent Stress-Strength Model for Unit Burr III Distribution under Progressive Censoring, Qual. Technol. Quant. Manag. 19 (2022), 605-632. https://doi.org/10.1080/16843703.2022.2049508.

[38] P.R. Tadikamalla, On a Family of Distributions Obtained by the Transformation of the Gamma Distribution, J. Stat. Comput. Simul. 13 (1981), 209-214. https://doi.org/10.1080/00949658108810497.

[39] C.W. Topp, F.C. Leone, A Family of J-Shaped Frequency Functions, J. Am. Stat. Assoc. 50 (1955) 209-219. https://doi.org/10.2307/2281107.

[40] J.R. Van Dorp, S. Kotz, The Standard Two-Sided Power Distribution and Its Properties, Am. Stat. 56 (2002), 90-99. https://doi.org/10.1198/000313002317572745.

Appendix

The main R codes for the real data analysis performed with the NUG distribution are given below.

```r
# Load the necessary library
library(betareg)

# Load the necessary data
data("FoodExpenditure")

# Define the transformed proportion data
datta <- FoodExpenditure$food / FoodExpenditure$income

# Define the NUG log-likelihood function
R <- function(theta, datta) {
  x <- datta
  alpha <- theta[1]
  beta <- theta[2]
  g <- (1/x^2) * dgamma(1/x - 1, shape = alpha, rate = beta)
  S <- -sum(log(g))
  return(S)
}

# Estimate parameters using numerical optimization
d <- nlminb(start = c(0.5, 1), R, lower = c(0, 0),
              upper = c(100, 100), datta = datta)

# Define the pdf of the NUG distribution
g <- function(x, theta) {
  alpha <- theta[1]
  beta <- theta[2]
    h <- (1/x^2) * dgamma(1/x - 1, shape = alpha, rate = beta)
  return(h)
}

# Plot histogram of the data and the estimated pdf of the NUG distribution
hist(datta, prob = TRUE, breaks = 12, main = "", xlab = "x",
     ylab = "", ylim = c(0, 6), col = "#36b612", density = 20)
curve(g(x, c(d$par[1], d$par[2])), col = "red", lty = 1,
      lwd = 2, add = TRUE)
legend("top", legend = c("Estimated pdf of the NUG distribution"),
       col = c("red"), lwd = 2, lty = 1, cex = 0.8)
grid()

# Compute the AIC and BIC
AIC <- 2 * d$objective + 2 * 2
BIC <- 2 * d$objective + 2 * log(length(datta))

# Output estimated parameters and model selection criteria
c(d$par[1], d$par[2], AIC, BIC)
```