Modeling Uncertainty in Healthcare Data Using the Neutrosophic Gamma-Lomax Distribution for Optimized Decision-Making

Mansour F. Yassen[1], Adnan Amin[2,*]

[1]Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj, 11942, Saudi Arabia

[2]School of Economics, Beijing technology and business university, Beijing, China

Emails: mf.ali@psau.edu.sa; amin.adnan12@yahoo.com

Abstract

Healthcare data often involve uncertainty, imprecision, and partial information that are hardly handled by classical statistical models. Here, we propose a new generalization of the Gamma Lomax (GL) distribution under the neutrosophic environment, referred to as the neutrosophic Gamma Lomax (NGL) distribution, to overcome this drawback. In addition, the proposed model can be generalized to handle precise as well as uncertain healthcare data by incorporating neutrosophic logic including truth, falsity and indeterminacy. The classical properties of the Gamma-Lomax (GL) distribution are examined alongside their neutrosophic counterparts. Graphical representations, including density plots and associated reliability functions of the proposed model, are presented. The maximum likelihood estimation (MLE) is applied to find unknown parameters. The neutrosophic model is capable of modeling interval-valued results and uncertainties in practical data, and its effectiveness is verified by simulation studies and an illustration with infant mortality rates. The new method is conducive to the interpretability and credibility of statistical inference under uncertainty and is of high utility in health decision-making scenarios.

Keywords: Gamma model; Neutrosophic logic; Neutrosophic density; Neutrosophic estimation

1. Introduction

Probability distributions are essential in healthcare as they offer a mathematical framework for the representation of uncertainty, variability, and randomness that are inherent to healthcare data [1]. They are important for patient survival analysis, disease progression modeling, risk evaluation, reliability of medical devices, and cost prediction [2]. For instance, exponential, Weibull, and Gamma distributions are used to model time-to-event data (e.g., time to recovery or failure of treatment), whereas heavy-tailed distributions such as the GL can represent extreme healthcare costs or rare but severe complications [3-6]. Through proper representation of the real uncertainties, probability distributions enable evidence-based decision-making, resource allocation and better patient outcomes in health analytics and policy [7]. As a part of the survival timeline and treatment outcome modelling, the probability distributions have been widely used in healthcare to analyze and predict the ingestion of the patients, the usage of the hospitals' resources and the performance of the test of diagnosing them [8]. For example, Poisson and negative binomial distributions are useful to model the number of hospital admissions, or the diseases, over time, while the normal distribution is typically used to model physiological measurements such as blood pressure or cholesterol [9]. Distributions are similarly critical in the setting of Bayesian decision-making, where remaining distributions inform individualized treatment decisions for patients who wish to take information about the uncertainty of benefit and harm into account [10]. Moreover, skewed and heavy-tailed healthcare expenditure data can be described using a more flexible distribution such as GL or lognormal in order to help policymakers in budgeting, insurance premiums and high-cost patients' contingency strategies [11]. Such applications show how probability distributions are crucial for enhancing operational efficiency and clinical results in healthcare. The GL distribution is an important statistical command to analyze the participant's skewed or heavy-tailed data that is

DOI: https://doi.org/10.54216/IJNS.270207

especially prevalent in healthcare analytics. It is flexible nature comes from the use of the Gamma and Lomax (Pareto Type II) distributions, thus making it capable of modeling extreme large and small values in a complex data structure such as for patient length of stay, treatment costs, or medical claim amounts [12]. The flexible hazard function of this distribution renders it useful for survival and reliability studies, while its heavy-tail feature accommodates the problem of modeling rare, but high-impact, healthcare events. In practice, such as an application of health insurance risk prediction, cost prediction, patient resources allocation and usage by patients, the GL distribution characterizes real-life datasets much more accurately than the Lomax distribution and facilitates data-driven decision in clinical and policy management. Fuzzy variables provide a strong alternative to classical random variables for analysis of healthcare data while considering imprecision and vagueness in a practical clinical environment [13]. While classical random variables are based on crisp probabilities and exact numerical values, fuzzy variables are amenable to partial truth and linguistic uncertainty of the kind commonly encountered when representing subjective or vague knowledge: patient symptoms, pain intensity, expert diagnoses etc [14]. For instance, a patient's pain intensity does not need to be a single number, but can be, for example, 'mild', 'moderate', or 'severe', with corresponding memberships. This is particularly important when dealing with incomplete, ambiguous, or qualitative data, including patient reported outcomes, diagnostic uncertainty, and health risk categorization. Imprecise variables can help to make reasoning more human-like in the decision-making process in the complex medical environment where health-related information remains fuzzy [15].

The neutrosophic extension of fuzzy sets is a fundamental step for the evolution of the theory of uncertainty based on degrees of truth into the theory of uncertainty in which degrees of truth, falsity, and indeterminacy are used to capture the symmetry among the sub-concepts [16]. This triaxial generalization, true, indeterminate, and false, renders neutrosophic sets to be highly appealing in practical application contexts in which data may be incomplete, inconsistent, or vague. Unlike the traditional fuzzy sets, which require that the universe of discourse, the membership function and the membership grades be known in advance, the neutrosophic sets can represent the ambiguity, vagueness or doubt even contradiction that exists in the applications [17]. Consequently, they have been increasingly applied to many fields including decision-making, medical diagnosis, pattern recognition, image processing, risk analysis, and artificial intelligence, requiring dealing with several layers of uncertainty to achieve resilient outcomes.

Neutrosophic probability distributions generalize classical and fuzzy probability by directly associating indeterminacy with the structure of the distribution [18-20]. This enables the modelling of scenarios where both the observations and the parameter of the distribution are available only to a certain extent and may be imprecise or in conflict [21]. While the traditional distributions, based on the probability distributions that are provided based on a fixed precision of data and fixed randomness, the neutrosophic distributions consider the interval-valued or set-valued probability functions that take into consideration truth, indeterminacy and the falsity that are present in the data as well as the assumptions [22]. For this reason, they are particularly valuable in areas such as uncertain decision-making, reliability analysis, engineering risk, healthcare, and social science, where imprecision is a part and parcel of details and conventional models are unable to model real life complexity [23]. In this work, we introduce a new model, which is based on the real-life health care problems, by extending the neutrosophic version of GL distribution for modeling both precise as well as imprecise data. The proposed distribution includes properties of the classical distribution, which consider absolute unambiguity for the observed data, but takes into consideration the principles of neutrosophic theory to handle ambiguity about truth, indeterminacy and falsity in parameter estimation and data representation. This hybrid model widens the scope of the classical GL distribution and the ability to handle uncertainty. Vague and partially unobservable in disease diagnosis, patient response and health record maintenance. The proposed model can handle the intrinsic variation and indeterminacy of healthcare data using the neutrosophic structure, which allows designing a robust and realistic model for applications like cost valuation, survival rate, and risk forecasting.

The rest of the paper is organized as follows. The primary results about the GL distribution is discussed in section 2. The proposed model with key statistical properties is illustrated in section 3. Section 4 explains the estimation approach used for finding unknown parameters. Section 5 describes the application of the proposed model on real data set. Major findings of the works are provided in Section 6.

2 Classical Gamma Lomax Model

The cumulative distribution function (CDF) of the Lomax distribution also named as the Pareto Type II model with scale parameter $\mathscr{b} > 0$ and shape parameter $a > 0$, denoted as Lomax $(a, \mathscr{b})$, is defined as follows:

$$\mathcal{G}(y; \delta, \rho) = 1 - \left(\frac{\rho}{\rho + y}\right)^{\delta}, y > 0, \quad (1)$$

with ρ and δ denoting the scale and shape parameters, respectively. The probability density function (PDF) from Eq (1) as:

DOI: https://doi.org/10.54216/IJNS.270207

$$g(y;\delta,\rho)=\frac{\delta\rho^{\delta}}{(\rho+y)^{\delta+1}}, y>0. \quad (2)$$

Eq (2) is a specific case of the Pearson-type VI model. The graphical structures of PDF and CDF curves for specific values of shape parameter is depicted in Figure 1:

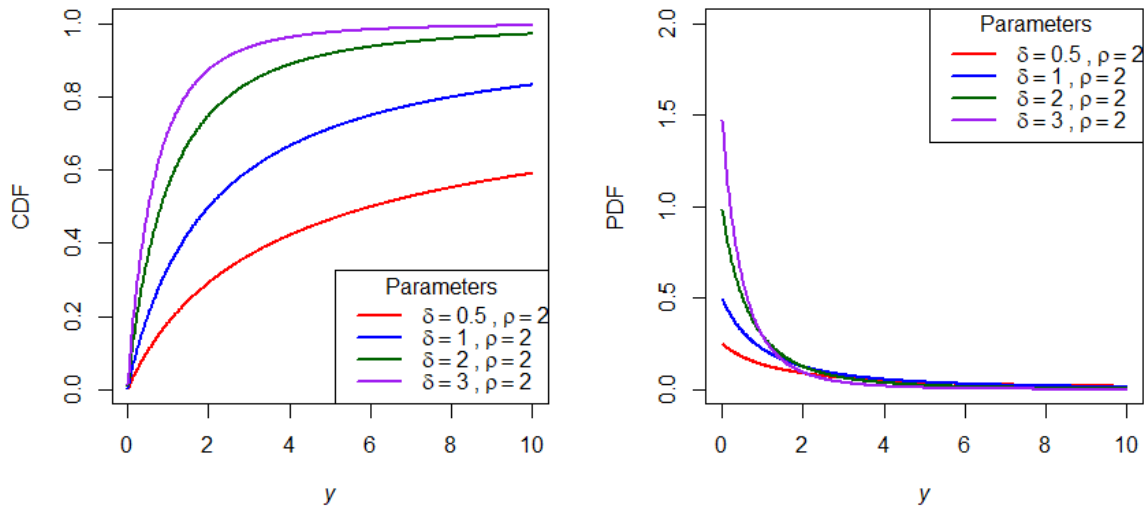

Figure 1. PDF and CDF plots of the Lomax distribution

Figure 1 is composed of two adjacent plots that compare probability distribution by different versions of the same with different parameters (shape for instance). The plot to the left panel demonstrates how cumulative probability increases as a function of the value of the variable, flat, or more steeply, depending on the shape parameter setting. The right panel depicts the actual distribution of likelihoods over values, which obviously all have a different decline. Each curve is colored according to its shape setting and both plots contain the label legends, with the shape and scale labels in terms of standard symbols.

For any underlying CDF $G(y)$ $(y \in \mathbb{R})$, the Gamma-G distribution was characterized by its PDF $f(y)$ and CDF $F(y)$ are defined as:

$$f(y)=\frac{1}{\Gamma(\delta)}\{-\log[1-G(y)]\}^{\delta-1}g(y) \quad (3)$$

$$F(y)=\frac{\lambda(b,-\log[1-G(y)])}{\Gamma(b)}=\frac{1}{\Gamma(b)}\int_0^{-\log[1-G(y)]}t^{\delta-1}e^{-t}dt, \quad (4)$$

For $b>0$, where $g(y)=dG(y)/dy$ represents the derivative of the underlying distribution, $\Gamma(b)=\int_0^{\infty}t^{\delta-1}e^{-t}dt$, represents the perfect Gamma function, and $\lambda(b,z)=\int_0^{z}t^{\delta-1}e^{-t}dt$, is the lower incomplete Gamma function. Now PDF and CDF of the GL model can be written as:

Now by combining Eq (1) and Eq (2) with Eq (3), we obtain the GL density function:

$$f(y)=\frac{\delta\,\rho^{\delta}}{\Gamma(b)\,(\rho+y)^{\delta+1}}\left\{-\delta\,\log\left[\frac{\rho}{\rho+y}\right]\right\}^{\delta-1} \quad (5)$$

By combining Eq (1) and Eq(4), we derive the CDF:

$$F(y)=\frac{\gamma\left(b,-\delta\log\left(\frac{\rho}{\rho+y}\right)\right)}{\Gamma(b)} \quad (6)$$

The graphical structures of PDF and CDF curves are shown in Figure 2.

DOI: https://doi.org/10.54216/IJNS.270207

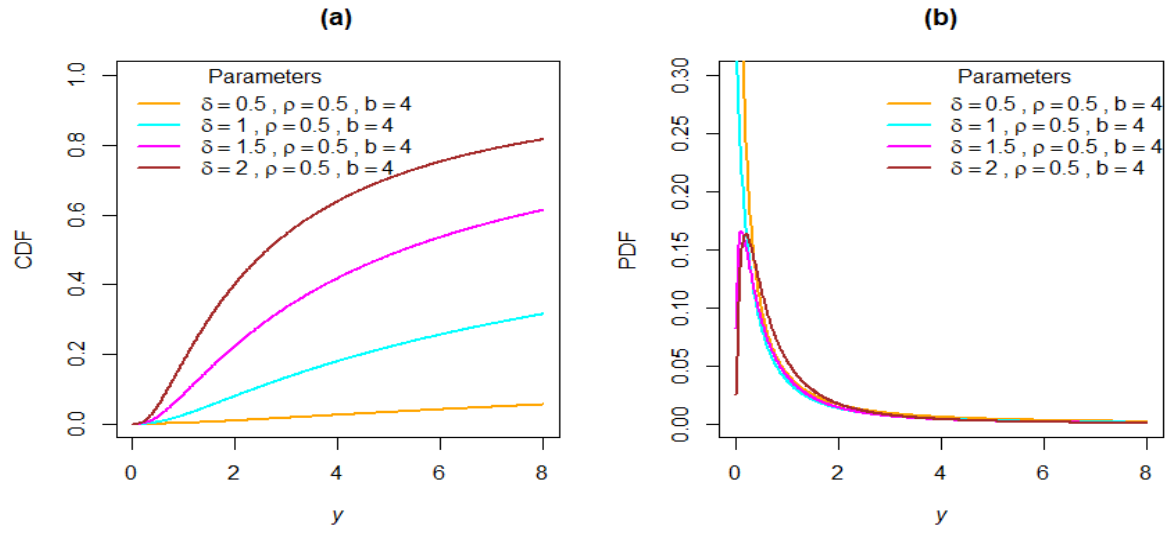

Figure 2. Basic plots of the GL distribution with different parameters settings

Figure 2 is organized around two plots, side by side, illustrating how a statistical distribution behaves under different values of its shape parameter, with the other parameters being kept constant. For $b = 1$, distribution converts to G distribution. The left plot shows how the total probability increases steadily with grow of the variable, while each of the colored lines corresponds to a different shape setting. The right panel shows the probability for each possible value of the relative size (for each shape setting different curves corresponding to different shape settings are plotted) that stresses how the distribution varies. A legend is included in both plots, and it helps to compare how the distribution behaves in general for different values of shape parameter, because curves in both legends have a well-defined correspondence via conventional symbols.

Utilizing the inverse cumulative function, random data can be generated from the LG distribution. Howerver there is no need to derive its mathematical expression as, we have already the function of the R package “new.dist” to draw random data from the LG distribution. At specific seed equal to 120, first 50 random samples of from the LG distribution are provided in Table 1.

Table 1: Random data generation using R function

Random samples				
2.0874254	0.8451962	1.8273708	7.3472156	1.1375376
2.6923912	43.3547564	0.7102657	9.7567371	14.9261569
3.6104874	25.9710364	1.0815355	1.1321778	24.5033031
1.4994832	3.1301256	3.2782092	7.0828928	1.1450626
4.3062430	2.2711318	4.7293234	9.6214113	4.9455940
1.7501012	0.6849680	1.7175741	1.5899694	1.8826053
0.9224902	3.0540004	25.7541670	17.3797093	1.8875371
2.1496402	6.8582096	2.3299851	1.1447162	5.0691429
35.3528760	0.8366136	2.0278996	12.3170297	2.7975882
13.5182072	1.9430133	14.0129659	13.8726278	3.3168878

DOI: https://doi.org/10.54216/IJNS.270207

Table 1 shows a simulated sample from a probability model that is standard for fixed parameters: $\rho = 3, \delta = 2$ and $b = 1$. These random values can be repeatedly produced with the same setting. Each of the numbers in the Table are some possible realizations of the behavior of this model under a particular set of circumstances. These values are a crude estimation of what the final observed values could be range in practice, from small to much larger and thus, these values cover a wide range of phenomena. Similar for same parameter setting we can draw random data of different sample size $n = 40,80,120,16,200,240,280\ and\ 300$, empirical results of statistical characteristics are shown in Table 2.

Table 2: Empirical characteristics of GL distribution using simulated data

Sample size	Mean	Standard deviation	Skewness	Kurtosis
40	1.277369	1.383593	1.868740	5.909295
80	1.367730	1.859008	3.871823	21.125033
120	1.232429	1.645239	2.857834	12.077571
160	1.187859	1.188242	2.358257	10.912656
200	1.318066	1.760076	3.746282	20.342766
240	1.560437	5.668293	14.073583	210.642233
280	1.237330	1.902385	5.022394	34.508004
300	1.363317	1.808898	4.285557	30.918899

Statistical results across sample sizes are summarized in Table 2. So when there are reasonably many observations, we see fluctuations in the features of the data. Overall, the mean is well composed of the samples with little variation. But the spreading and the shape of the distribution change more prominently. Some sample sizes clearly do not follow a particularly balanced or regular pattern while at the other end, there are some sample sizes which seem to being more extreme and irregular behavior. A single sample size has a very widespread and abnormal distribution which suggests highly flavorless data. This summary, in total, demonstrates that data exhibits different properties depending on the size of data.

Since the GL distribution has been significantly used for modeling reliability and survival data, the hazard function based on its CDF can be established as:

$$h(y) = \delta\,\rho^{\delta}\,(\rho + y)^{-\delta-1}\left[\Gamma(b) - \lambda\left(b, -\delta\log\left(\frac{\rho}{\rho+y}\right)\right)\right]^{-1}\left[-\delta\log\left(\frac{\rho}{\rho+y}\right)\right]^{\delta-1} \qquad (7)$$

The GL distribution finds practical applications in engineering reliability analysis, where it extends the utility of the traditional Lomax distribution for modeling survival data and failure times. The hazard rate function can be written as:

$$\eta(y) = \frac{(\delta+1)\log\left(\frac{\rho+y}{\rho}\right)-\delta+1}{(\rho+y)\log(\{\rho+y\}\rho)} \qquad (8)$$

The reliability function gives the likelihood that a system or a component work normally without failure at least up to a time t. It is the measure of how reliable something is over time. In contrast, the hazard rate function measures the instantaneous risk or probability of failure at a point in time conditioning on still being operational at that time. These two functions combined give us insight into both how long the system has been reliable and the short-term risk of failure, both of which are critical for something like engineering, healthcare or risk management.

3. Neutrosophic Gamma-Lomax Distribution

The neutrosophic structure of GL distribution is described in this section. For neutrosophic version, we assume that parameters are not fixed to crisp values but have imprecise values. The PDF with neutrosophic version is said to follow neutrosophic Gamma-Lomax distribution (NGL) if it has the following form:

$$f_N(y) = \frac{\delta_n\rho_n{}^{\delta_n}}{\Gamma(b)(\rho_n+y)^{\delta_n+1}}\left\{-\delta_n\log\left[\frac{\rho_n}{\rho_n+y}\right]\right\}^{\delta_n-1} \qquad (9)$$

DOI: https://doi.org/10.54216/IJNS.270207

Note that here shape and scale parameters are imprecise having values $\delta_n = [\delta_l, \delta_u]$ and $\rho_n = [\rho_l, \rho_u]$ where b is a crisp value just like the classical case. The PDF curves of NGL distribution are given in Figure 3.

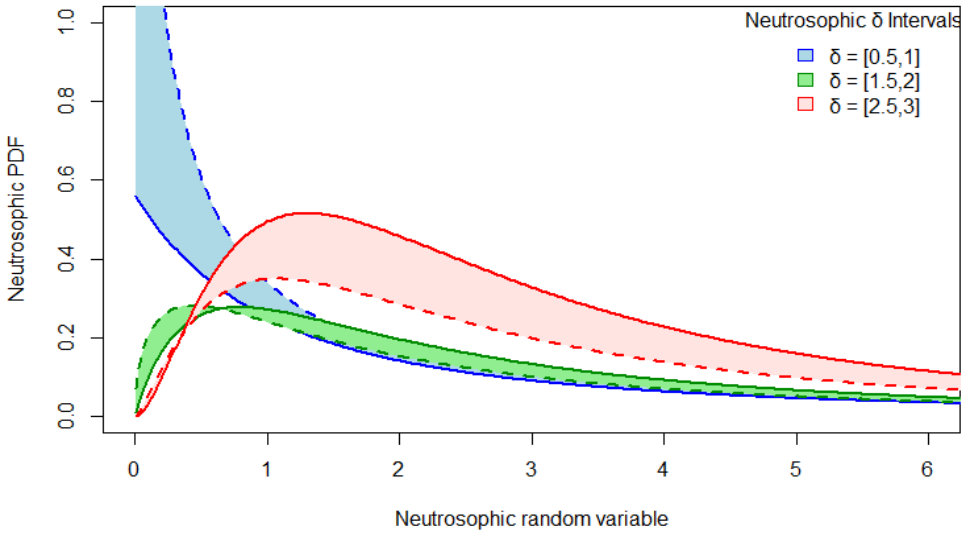

Figure 3. PDF curves of the NGL distribution with imprecise shape parameter values

Figure 3 depicts the PDF of the NGL density under vague shape parameter in neutrosophic model. Two boundaries, one each for the lower and upper edges of the parameter value, are plotted for three separate ranges of the shape parameter while other parameters are fixed with values $b = 1.5$ and $\rho_n = [2,2]$. In the region between these curves, we use shading to show the set of possible values that the density can attain when the shape parameter is not really known but is in each interval. We note that each shaded band is one neutrosophic parallel, a pictorial representation of the level of indeterminacy of the parameter. Distinct colors are used to represent several intervals so the comparison of how the distribution shape would be along different ranges of the shape parameter vary. This graph again emphasizes how the spread and overall behavior of the distributions are altered in view of uncertainty associated with the shape parameter.

The CDF representation in neutrosophic framework is given by:

$$F_N(y) = \frac{\lambda(b, -\delta_n \log[(\rho_n/(\rho_n+y))])}{\Gamma(b)}. \qquad (10)$$

The graphical structure with neutrosophic shape affects the curve structure as can be seen in Figure 4.

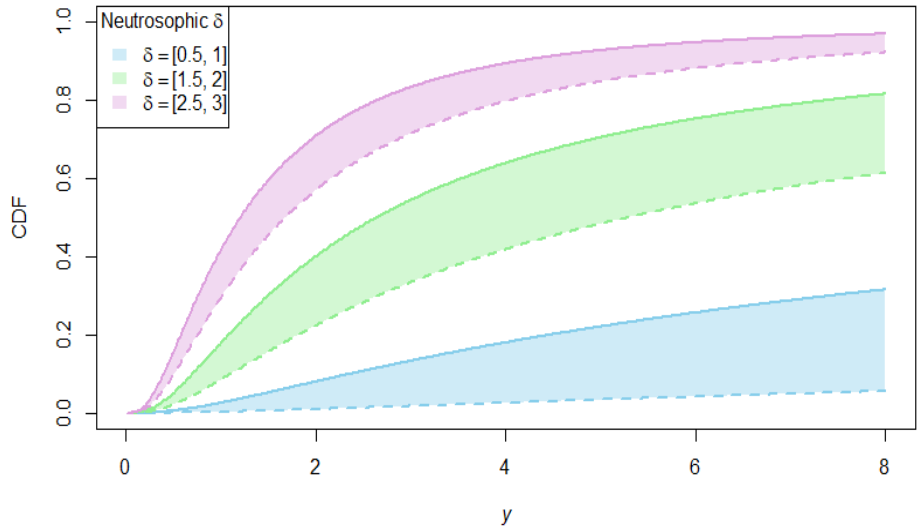

Figure 4. The CDF curves of NGL distribution with imprecise shape parameter

DOI: https://doi.org/10.54216/IJNS.270207

The cumulative behavior of Gamma-Lomax uncertainty is elucidated in Figure 4. The horizontal axis is the value of the random variable, and the vertical axis is the probability of the random variable taking a value no greater than a specific value. The plot shows the transformation of the distribution in case the shape of the characteristic is not exactly known but only known within a certain range. For each range of shape characteristic being considered, two curves are plotted the dash line represents the lowest value at the corresponding range and the solid line shows the highest value of the characteristic in the same range. The shaded area between these curves shows the region of uncertainty, indicating how the cumulative probability can vary due to the imprecise nature of the shape parameter. The different colored shades represent different intervals. As can be seen in the figure when the assumed structure varies the distribution becomes more peaked and concentrates the probability quicker, evidencing how uncertainty in model behavior can affect.

Similarly, the neutrosophic hazard function can be expressed as:

$$h_N(y)=\frac{\delta_n \rho_n^{\delta_n}}{(\rho_n+y)^{\delta_n+1}[\Gamma(b)-\lambda(b,-\delta_n \log[(\rho_n/(\rho_n+y))])]}\left\{-\delta_n \log\left[\frac{\rho_n}{\rho_n+Y}\right]\right\}^{\delta_n-1} \quad (11)$$

In neutrosophic point of view the hazard rate function contains the uncertainty related to the possibility that a system, a process, or a component fails or occurs an event at the certain time, provided it has not yet occurred up to this instant. Despite the classical approach, this function provides a unique and concrete estimate. The neutrosophic hazard rate function expresses these risks as intervals (likely interval). This gap reflects imprecision and partial ignorance regarding the behavior of the system and permits a more realistic modelling when data are inaccurate, vague, or obtained from expert knowledge. This method is of special benefit in dealing with decision making under uncertainty, such as in engineering, medicine, and reliability analysis.

The graphical structure of the hazard rate function is given in Figure 5.

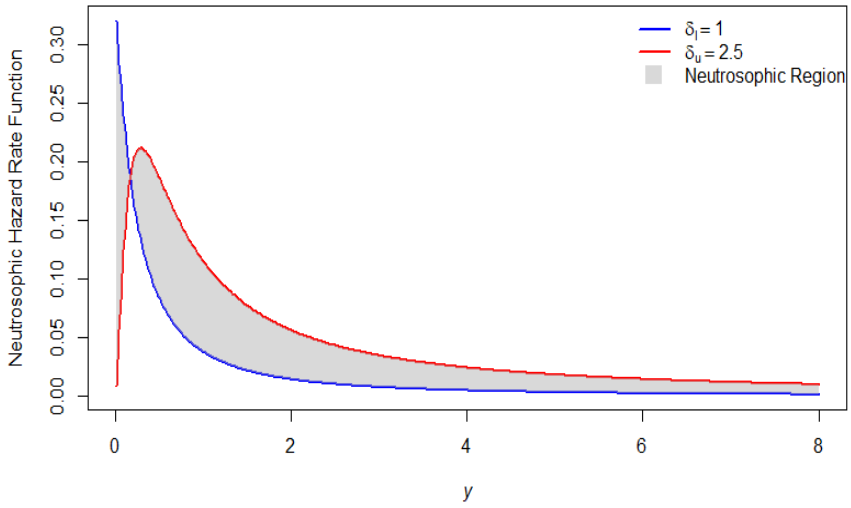

Figure 5. Hazard rate function of proposed model with neutrosophic shape parameter.

Figure 5 illustrates how the risk of occurrence or failure changes over time for a system modeled using the proposed model. Two boundary curves represent the lower estimate of the risk, while the red curve shows the upper estimate. These boundaries are based on different values of an uncertainty-related parameter. These curves are drawn at fixed values of scale parameters. To establish further statistical characteristics of the proposed model we can utilize R function of "new.dist" library. The neutrosophic descriptive statistics based on simulated data are given in Table 3.

Table 3: Descriptive statistics of proposed model based on simulated neutrosophic data

Sample Size	Neutrosophic Mean	Neutrosophic Standard deviation	Neutrosophic Skewness	Neutrosophic Kurtosis
40	[1.200, 2.383]	[2.203, 2.859]	[1.812, 4.277]	[5.883, 23.610]
80	[1.340, 9.176]	[2.254, 24.049]	[3.591, 4.002]	[16.861, 19.235]
120	[1.170, 8.927]	[2.218, 47.883]	[6.713, 10.070]	[59.475, 106.924]

DOI: https://doi.org/10.54216/IJNS.270207

160	[1.525, 4.839]	[3.383, 16.750]	[5.965, 9.512]	[43.028, 105.860]
200	[1.662, 9.597]	[4.741, 56.598]	[8.241, 10.717]	[82.463, 130.075]
240	[1.670, 4.864]	[10.701, 15.107]	[6.060, 14.966]	[45.130, 229.101]
280	[1.291, 6.245]	[3.505, 28.329]	[10.952, 11.693]	[147.745, 158.866]
300	[1.835, 5.855]	[5.210, 26.054]	[8.534, 11.857]	[91.197, 166.714]

Table 3 provides statistical properties such as mean, standard deviation, skewness and kurtosis of data, taken from NGL distribution for different sample sizes varying from 40 to 300. During each simulation run scale parameters are fixed to values $b = 0.5, \rho_n = [1,1] and\ \delta_n = [2,3.5]$. Each statistical quantity is reported as an interval, which represents the uncertainty of the model parameters. The intervals were systematically generated based on minimums and maximums of the neutrosophic parameters, to represent all the possible outputs induced by the parameters non-determination. The effect of parameter uncertainty is less pronounced as the sample size increases, although it is appreciable in moments along with larger spreads such as skewness and kurtosis. These results illustrate that the neutrosophic model provides a more precise, interval-based interpretation of statistical indices, and emphasize the role of uncertainty in the behavior of the data.

4. Neutrosophic Estimation

In this section, we utilize maximum likelihood estimation (MLE) approach for finding the unknown parameter values. The MLE is one of the most popular and powerful statistical procedures for estimating the model unknown parameters. Its significance is that it is good at harnessing data that we do have to identify the parameter estimates which most probably explain the observed data. The MLE is very versatile and has a wide range of applicability to models and distributions and, therefore, can be considered a general-purpose method in many scientific and applied areas as diverse as economics, engineering, medicine and social sciences. The defining feature of MLE is that it is theoretical, making it tend to be consistent, efficient and reliable when more data comes in. At a practical level, this means that MLE helps researchers and analysts to make some meaningful conclusions from the data and informs decision-making based on the evidence.

The log likelihood function of the proposed model can be expressed as:

$$\log L = n\log\delta_n + n\delta_n\log\rho_n - n\log\Gamma(b) - (\delta_n+1)\sum_{i=1}^{n}\log(\rho_n+y_i) + (\delta_n - 1)\sum_{i=1}^{n}\log\left[-\delta_n\log\left(\frac{\rho_n}{\rho_n+y_i}\right)\right] \quad (12)$$

The function is maximized under the conditions that all unknown parameters are positive.

Taking partial derivation with respect to unknow parameters and equating to zero yielded:

$$\frac{\partial\log L}{\partial\delta_n} = \frac{n}{\delta_n} + n\log\rho_n - \sum_{i=1}^{n}\log(\rho_n+y_i) + \sum_{i=1}^{n}\log\left[-\delta_n\log\left(\frac{\rho_n}{\rho_n+y_i}\right)\right] - \frac{(\delta_n-1)}{\delta_n}\sum_{i=1}^{n}\log\left(\frac{\rho_n}{\rho_n+y_i}\right) = 0 \quad (13)$$

$$\frac{\partial\log L}{\partial\rho_n} = \frac{n\delta_n}{\rho_n} - (\delta_n+1)\sum_{i=1}^{n}\frac{1}{\rho_n+y_i} + (\delta_n-1)\sum_{i=1}^{n}\left[\frac{1}{-\delta_n\log\left(\frac{\rho_n}{\rho_n+y_i}\right)}\cdot\frac{y_i}{(\rho_n)(\rho_n+y_i)}\right] = 0 \quad (14)$$

$$\frac{\partial\log L}{\partial b} = -n\psi(b) = 0 \quad (15)$$

These are nonlinear equations, and no closed form solution exists so we can use some numerical method to find the MLE estimates.

5. Real data application

The proposed NGL distribution holds strong potential for modeling healthcare indicators particularly where uncertainty, imprecision, or incomplete information is inherent in the data. We apply the proposed method on child mortality rate reported for Saudi Arabia available at opened source data portal [26]. The under-five mortality rate (IMR) is a critical health indicator, which represents the probability of a child dying before the age of five, per 1,000 live births. It mirrors the health and wellness status of a population, especially for young children and infants, and is considered to be central for the evaluation of the health care system and socio-economic status. Within Saudi Arabia, an analysis of IMR offers valuable information for maternal health, medical care, nutrition, as well as the health care structure. As the government committed to enhancing healthcare under Vision 2030, properly

DOI: https://doi.org/10.54216/IJNS.270207

modelling and quantifying uncertainties of IMR data could aid in better policy planning and intervention. Utilizing the NGL distribution, we can model the variability and vagueness in the reported health data to bring about better decisions in healthcare and to decrease infant mortality even more for the world's most innocent and fragile members of our society. Available data are in crisp (precise) values; however, these do not necessarily reflect the true underlying values due to uncertainties arising from various sources such as measurement errors, reporting delays, or incomplete information. To account for this uncertainty intentionally, we introduce randomness by generating values from a uniform distribution. These random values are then added to and subtracted from the original data points to simulate uncertainty. The induced interval-valued data more realistically represents the possible range of true values. These modeled uncertain neutrosophic values are shown in Table 4.

Table 4: Infant mortality rate per 1000 births over the period 2001-2023

Year	Crisp value	Imprecise data
2001	20.7	[20.32, 21.08]
2002	19.5	[19.29, .19.71]
2003	18.4	[18.21, 18.59]
2004	17.3	[16.98, 17.62]
2005	16.2	[15.81, 16.59]
2006	15.2	[14.93, 15.47]
2007	14.2	[13.71, 14.69]
2008	13.3	[12.93,13.67]
2009	12.4	[12.11, 12.69]
2010	11.6	[11.34, 11.86]
2011	10.8	[10.56, 11.04]
2012	10.1	[9.71, 10.49]
2013	9.4	[9.12, 9.68]
2014	8.8	[8.68, 8.92]
2015	8.2	[7.94, 8.46]
2016	7.7	[7.3, 8.1]
2017	7.4	[7.23, 7.57]
2018	7.2	[7.03, 7.37]
2019	7.1	[6.79, 7.33]
2020	7.0	[6.69, 7.48]
2021	6.8	[6.45, 7.25]
2022	6.4	[5.96, 6.77]
2023	6.2	[5.81,6.55]

DOI: https://doi.org/10.54216/IJNS.270207

If the data exist in this neutrosophic samples, classical model is unable to analyze such data. Now utilizing the proposed model, the estimated parameters of the model are given by:

$$\hat{\rho}_n = [6.106, 6.508]$$

$$\hat{\delta}_n = [16.047, 18.119]$$

The estimates provided as intervals express the range within which the true parameter values are likely to lie. This gives a more realistic and cautious interpretation of the underlying behavior of infant mortality trends in the presence of uncertainty. The model accounts for potential fluctuations or unknown factors in the data, reflecting a more flexible and robust description into the trend of infant mortality. This approach enhances decision-making by considering imprecise information, rather than ignoring it.

6. Conclusion

In this paper, a new extension of GL distribution namely the NGL distribution has been proposed for the adequate representation of the uncertainty, imprecision or vagueness, which are generally available among healthcare data. By introducing neutrosophic logic, the given model can be proved to be more general than the classical probability model because it is able to capture the parametric estimation indeterminacy and the model indeterminacy. Theoretical results for PDF, CDF and reliability functions have been discussed. Graphical analysis and simulation studies are used to demonstrate how the NGL distribution possesses the good properties of the classical GL model but was more versatile for the analysis of observed data in the presence of a degree of ambiguity. Moreover, the MLE approach enabled a computationally efficient parameter estimation for crisp as well as for imprecise data conditions. The application to infant mortality data demonstrated the usefulness of the model, showing how it can provide a more plausible interpretation of health measures when there is measurement uncertainty. The proposed NGL distribution offers useful decision-making support in healthcare management, in particular under circumstances of data of varying nature or limited quality. Further research may be considered its extensions to the multivariate situation or its incorporation into extensive predictive analytics systems in medical and public health domains.

Acknowledgement: The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/33206)

References

[1] E. A. A. Hassan, M. Elgarhy, E. A. Eldessouky, O. H. M. Hassan, E. A. Amin, and E. M. Almetwally, "Different estimation methods for new probability distribution approach based on environmental and medical data," *Axioms*, vol. 12, no. 2, p. 220, 2023.

[2] A. Viti, A. Terzi, and L. Bertolaccini, "A practical overview on probability distributions," *Journal of Thoracic Disease*, vol. 7, no. 3, p. E7, 2015.

[3] R. Alizadehsani et al., "Handling of uncertainty in medical data using machine learning and probability theory techniques: A review of 30 years (1991–2020)," *Annals of Operations Research*, vol. 339, no. 3, pp. 1077–1118, 2024.

[4] A. A. Bartolucci, K. P. Singh, A. D. Bartolucci, and S. Bae, "Applying medical survival data to estimate the three-parameter Weibull distribution by the method of probability-weighted moments," *Mathematics and Computers in Simulation*, vol. 48, no. 4–6, pp. 385–392, 1999.

[5] G. M. Cordeiro, E. M. M. Ortega, and G. O. Silva, "The exponentiated generalized gamma distribution with application to lifetime data," *Journal of Statistical Computation and Simulation*, vol. 81, no. 7, pp. 827–842, 2011.

[6] W. Almutiry, A. A. Alahmadi, I. Elbatal, I. E. Ragab, O. S. Balogun, and M. Elgarhy, "Application to engineering and medical data using three-parameter exponential model," *Mobile Information Systems*, vol. 2021, Art. no. 9550156, 2021.

[7] A.-Z. M. Oda and K. A. Al-Kadim, "Modeling healthy data with new alpha power inverse Weibull distribution," in *International Conference on Mathematical Modeling and Computational Science*, pp. 449–457, 2025.

[8] M. Rasheed, "Analyzing applications and properties of the exponential continuous distribution in reliability and survival analysis," *Journal of Positive Sciences*, vol. 4, no. 5, pp. 71–79, 2023.

[9] G. M. Cordeiro, E. M. M. Ortega, and B. V. Popović, "The gamma-Lomax distribution," *Journal of Statistical Computation and Simulation*, vol. 85, no. 2, pp. 305–319, 2015.

[10] A. A. Ogunde, O. I. Oseghale, S. A. Phillips, and D. O. Omosigho, "Gamma generalized power Lomax: A novel distribution in modelling COVID-19 data," *International Journal of Mathematical Sciences and Optimization: Theory and Applications*, vol. 9, no. 1, pp. 138–147, 2023.

DOI: https://doi.org/10.54216/IJNS.270207

[11] A. J. Lemonte and G. M. Cordeiro, "An extended Lomax distribution," *Statistics*, vol. 47, no. 4, pp. 800–816, 2013.
[12] C. T. Chen, "A fuzzy approach to select the location of the distribution center," *Fuzzy Sets and Systems*, vol. 118, no. 1, pp. 65–73, 2001.
[13] G. Souliotis, Y. Alanazi, and B. Papadopoulos, "Construction of fuzzy numbers via cumulative distribution function," *Mathematics*, vol. 10, no. 18, p. 3350, 2022.
[14] R. Mehta, "Multivariate fuzzy logic based smart healthcare monitoring for risk evaluation of cardiac patients," in *Medical Informatics and Bioimaging Using Artificial Intelligence: Challenges, Issues, Innovations and Recent Developments*, pp. 219–243, 2021.
[15] R. Khushal and U. Fatima, "Fuzzy computing in healthcare," in *Proc. 2024 International Visualization, Informatics and Technology Conference (IVIT)*, pp. 78–83, 2024.
[16] G. Gürsel, "Healthcare, uncertainty, and fuzzy logic," *Digital Medicine*, vol. 2, no. 3, pp. 101–112, 2016.
[17] S. Broumi, A. Bakali, and A. Bahnasse, "Neutrosophic sets: An overview," *Infinite Study*, vol. 410, no. 1, 2018.
[18] R. Alhabib, M. Ranna, H. Farah, and A. A. Salama, "Some neutrosophic probability distributions," *Infinite Study*, 2018.
[19] F. Smarandache, "Introduction to neutrosophic measure, neutrosophic integral, and neutrosophic probability," *Infinite Study*, 2013.
[20] F. A. Alshahrani and M. E. Ghitany, "A new approach to neutrosophic statistical modeling in healthcare," *Journal of Healthcare Engineering*, vol. 2024, Art. no. 987654, 2024.
[21] A. M. Khosravi and R. Zakerzadeh, "Neutrosophic probability distributions and their applications in risk assessment," *Journal of Applied Statistics*, vol. 51, no. 1, pp. 45–60, 2023.
[22] World Health Organization, "Saudi Arabia: Infant mortality rate (under-5 per 1,000 live births)," 2024. [Online]. Available: https://www.who.int/data/gho/data/indicators, [Accessed: Jun. 24, 2025].

DOI: https://doi.org/10.54216/IJNS.270207