Neutrosophic Sets and Systems, Vol. 61, 2023
University of New Mexico

Neutrosophic Structure of the Geometric Model with Applications

Ahmedia Musa M. Ibrahim[1], Fuad S. Alduais[2,3], Zahid Khan [4,*] and Adnan Amin [5] and Katrina Lane-Krebs[6]

1 Finance Department, College of Business Administration in Hawtat Bin Tamim, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, am.ibrahim@psau.edu.sa
2 Mathematics Department, College of Humanities and Science in Al-Kharj , Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, f.alduais@psau.edu.sa
3 Business Administration Department, Administrative Science College, Thamar University, Thamar, Yemen
4 Department of Quantitative Methods, University of Pannonia, Veszprem, Hungary
5 Department of Mathematics and Statistics, Hazara University Mansehra, Pakistan, aminadnank@gmail.com
6 Higher Education Division, Central Queens Lands (CQ) University, Rockhampton, Australia k.lane-krebs@cqu.edu.au

* Correspondence: zahidkhan@hu.edu.pk

Abstract: In practical scenarios, it is common to encounter fuzzy data that contains numerous imprecise observations. The uncertainty associated with this type of data often leads to the use of interval statistical measures and the proposal of neutrosophic versions of probability distributions to better handle such data. We present a unique methodology that is based on the maximum likelihood approach and neutrosophic approach for estimating parameter of the proposed neutrosophic geometric distribution (NGD). The proposed methodology is supported by key likelihood inference results. The proposed distribution is specifically designed to handle variables with imprecise observation, hence effectively addressing a wide range of situations often encountered in the analysis of uncertain data. To evaluate the efficacy of the proposed neutrosophic model, we have carried out a comprehensive simulation experiment that rigorously examined the performance of the proposed model. The practical utility of NGD in the analysis of incomplete data is further exemplified through real-world applications.

Keywords: Neutrosophic logic, uncertain analysis, probability model, estimation, simulation

1. Introduction

Statistical distributions are a powerful tool for describing and predicting real-world events. The geometric distribution is possibly the most common distribution in statistical applications [1] . The geometric distribution is widely employed in various domains such as finance, investment, scientific research, and engineering, making it the most frequently utilized distribution [2]. The geometric distribution is a discrete probability distribution that is commonly employed to model the probability of attaining success in a sequence of independent trials with two possible outcomes [3]. Through the use of geometric distribution, it becomes possible to ascertain the likelihood of attaining success subsequent to a designated quantity of attempts [4]. The geometric distribution exhibits a multitude of uses in practical, real-world situations. As an illustration, it can be employed to simulate the quantity of endeavors required to achieve win in a game of probability or the quantity of unsuccessful tries prior to attaining success in a manufacturing procedure [5].

The geometric distribution is also used in banking to figure out how likely it is that a loan will not be paid back or how many trades are needed to make a profit [6], [7]. In the field of epidemiology, geometric distribution can also be used to model how many contacts a person with a disease has before they spread it to other people [8]. Additionally, it can be used in telecommunications to determine how many tries are needed to make a call in a busy network [9].

The geometric distribution is an important part of probability theory and has been studied a lot for its uses in many different areas [10]. Figuring out the chance of getting the first victory after a certain number of tries is what the geometric probability mass function is based on [11]. Well-known scientists like Feller [2] and Ross [3] have spent a great deal of time studying and exploring this idea. They have come up with detailed explanations and studies of its properties. In queuing theory, the geometric distribution is a key tool for finding out how long people will have to wait. Kleinrock's efforts [4] have shown that this can be used.

Barlow and Proschan [2] employ this probability distribution within the domain of reliability engineering to examine the duration required for the initial failure occurrence in systems. Furthermore, researchers in the field of epidemiology, such as Thelwell et al. [12], employ this tool as a means to get valuable understanding regarding the intricacies of disease transmission. The research conducted by Mandelbrot emphasises the importance of the Geometric distribution in the assessment of financial risk [13]. Furthermore, Preston's research delves into the use of this concept in the field of environmental science, namely in the modelling of species abundance [14]. The geometric distribution is widely employed in many disciplines, including information theory[15] , machine learning for pattern identification [10], game theory for strategic interactions [16], and educational research for comprehending learning patterns [17]–[20].

Fuzzy sets serve as the fundamental construct underlying the notion of fuzzy set theory. The notion of fuzzy sets is a crucial aspect within the framework of fuzzy set theory [21]. Fuzzy sets are mathematical constructs that enable the incorporation of partial membership or degrees of truth inside their representations [22]. The aforementioned frameworks offer a versatile structure for addressing ambiguity and imprecision across many domains, including but not limited to artificial intelligence, decision-making, and pattern recognition [23]. The integration of fuzzy sets within the framework of fuzzy set theory enables a more sophisticated and authentic methodology for modelling intricate systems and representing imprecise data [23]–[26]. The use of fuzzy set theory enables a more detailed modelling of complex systems, allowing for effective capture of imprecise information. Fuzzy control has been effectively employed in the automobile sector to regulate diverse systems, including automatic gearbox, suspension, engine, temperature control, and antilock brakes [27]. Furthermore, washing machines employ fuzzy control algorithms to adapt their washing approach according on several criteria, including the detected degree of filth, kind of cloth, size of the load, and water level [28]. The neutrosophy idea, initially proposed by Smarandache, is increasingly being recognised and used due to its capacity to offer a more adaptable and all-encompassing approach in addressing uncertainty and imprecision within the context of data analysis [29]. Neutrosophic statistics provide an expanded range of options for the representation and analysis of data, hence enabling to achievement of enhanced precision and dependability in the obtained outcomes [30], [31]. This strategy demonstrates significant use in scenarios when conventional statistical methods prove inadequate, consequently gaining greater popularity within the discipline of uncertain data analysis [32]–[35]. The proposal of NGD in this work is driven by the recognition of the significant role geometric distribution plays in statistical applications. Its wide applicability and the prevalence of uncertainty in real data make NGD an important consideration.

The proposed distribution and its key characteristics are described in Section 2. The estimation procedure for unknown parameters under the neutrosophic logic is presented in Section 3. In Section 4, the quantile function of the proposed model is formulated and the procedure for simulating data is explained. The significance of theoretical findings is concisely explained by analyzing a real-world examples in Section 5. Finally, Section 6 provides the final remarks of the study.

2. Proposed Model

This section presents a summary statistic of the proposed model and describes some of its important functions. The summary statistics of the proposed model provide a concise overview of its key characteristics. Additionally, the description of important functions commonly used in applied probability distribution theory helps to understand how the model can be utilized in practical applications. The geometric distribution holds significant importance in the field of statistics, being one of the fundamental distributions.

The formula provided below represents the neutrosophic probability density function (DF_n).

$$g_n(\mathcal{X}) = \mathcal{P}_n(1-\mathcal{P}_n)^{\mathcal{X}} \; ; \; \mathcal{X} \geq 0 \tag{1}$$

where $0 < \mathcal{P}_n = [\mathcal{P}_l, \mathcal{P}_u] < 1$ is the neutrosophic parameter of the NGD. To calculate the probability of waiting exactly r trials before the first successful event, we need to know the probability of success in a single trial ($\mathcal{P}_n$). The probability of failure (q_n) can be calculated as 1 minus $\mathcal{P}_n$. This scenario is known as a special case of the negative binomial distribution. It should be noted that the suggested model differs from the existing framework of the geometric model, where the parameter is precisely determined. The suggested model becomes equal to the classical model, when the indeterminate portion of the suggested model is zero, i.e., $\mathcal{P}_l = \mathcal{P}_u = \mathcal{P}$. The neutrosophic probability density function, often denoted as DF_n, is a mathematical function that describes the likelihood of a neutrosophic random variable taking on a particular interval value due to imprecision in $\mathcal{P}_n$. It provides valuable information about the distribution of the neutrosophic variable and can be used to calculate probabilities of different outcomes. Based on (1), the NGD is depicted in Figure1.

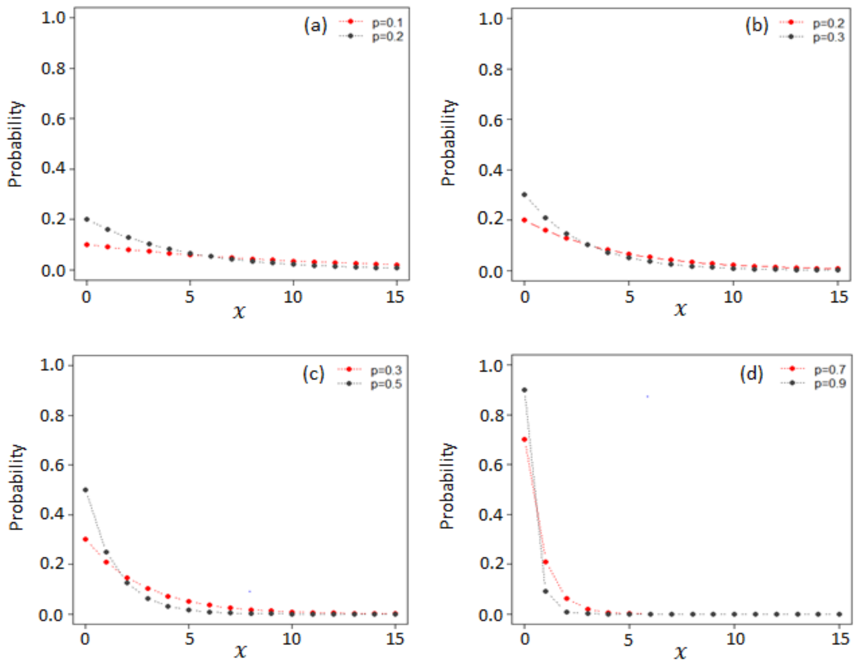

Figure 1: Density plots of the proposed NGD with different vague values of parameter

Figure 1 illustrates that there is a distinct interval probability for every value of the random variable $\mathcal{X}$.As illustrated in Figure 1(a), for instance, $\mathcal{P}_n = [0.1, 0.2]$ approximation for $\mathcal{X} = 1$, and the same is true for other values. The graph of DF_n shows that the likelihood of different outcomes occurring within a given range. It provides a visual representation of the probability of each possible outcome. By examining the shape and characteristics of the DF_n, one can gain insights into the likelihood and spread of values within the distribution. The neutrosophic probability mass function (PMF_n) of any density is another fascinating feature of probability theory applications. To describe the distribution of a discrete random variable, we can use the PMF_n. This function assigns probabilities to each possible value that the random variable can take. The $PMF_n$is a cooperatively linked variant of the DF_n and may be calculated as:

$$\mathcal{G}_n(\mathcal{X}) = 1 - (1 - \mathcal{P}_n)^{\mathcal{X}} \quad (2)$$

It should be noted that the PMF_n can be applied to any real number in the set R. However, if an argument does not belong to the possible values that the variable can take (i.e., the support of the sample space), then the PMF_n will have a value of zero. Conversely, if an argument does belong to the support of the sample space, then the PMF_n will have a positive value. This means that the PMF_n assigns probabilities to specific values within the sample space. It is important to note that the sum of all the probabilities assigned by the PMF_n must equal 1. The graph of PMF_n with imprecise values of NGD with different interval values of $\mathcal{P}_n$ is shown in Figure 2.

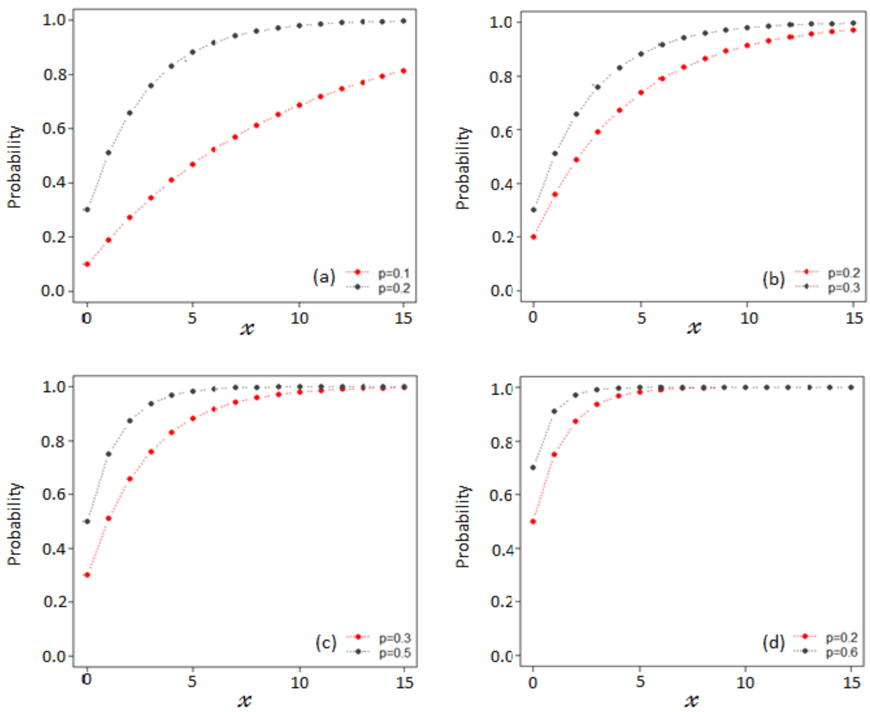

Figure 2: The graph of PMF_n of the proposed model

The PMF_n graph provides a visual representation of the probabilities linked to neutrosophic random variable. This graph illustrates discrete outcomes on the horizontal axis and their corresponding neutrosophic probability on the vertical axis. Each data point on the graph represents the probability of a certain result, with taller hight indicating more likely events. Importantly, the total of all probabilities shown on the graph equals one. The peaks spots on the graph depict the most probable occurrences, providing a distinct comparative examination of the likelihood of various events. The graph's discrete form, characterized by distinct double points, sets it apart from the classical plot of the geometric distribution. The PMF_n graph is a useful tool for comprehending and forecasting the unpredictability linked to discrete events in statistical research.

The suggested model's survival function can be described as follows in the neutrosophic framework: In the given statistical approach, the survival function plays a significant role in determining the probability of an individual's life surviving for a specific duration. Referred to as the survival rate, this function can be defined within the neutrosophic framework according to suggested model as:

$$\mathcal{S}_n(\mathcal{X}) = (1-\mathcal{P}_n)^{\mathcal{X}} \tag{3}$$

The graph of the survival function which is also known as reliability function is depicted in Figure 3.

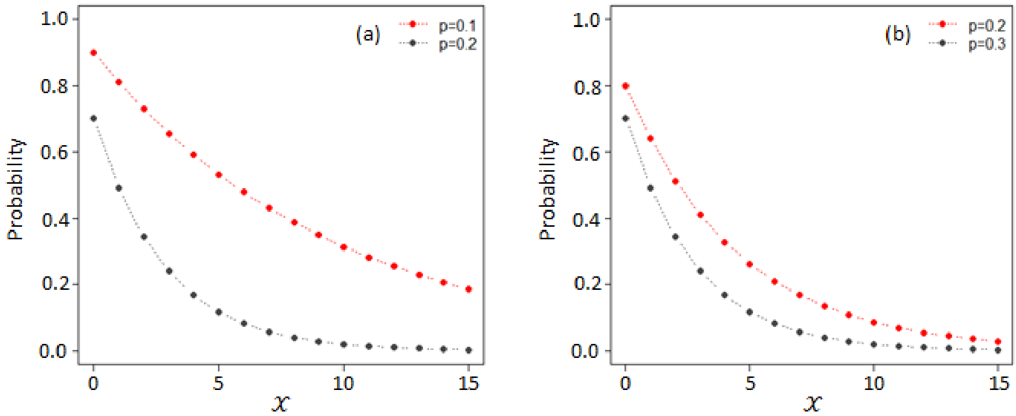

Figure 3: The survival function of the suggested NGD

The neutrosophic hazard function (HF_n), often known as the impending failure rate, is another important function in reliability analysis. For the given model, it is the ratio of the survival and density functions, which may be computed as follows:

$$\hbar_n(\mathcal{X}) = \frac{g_n(x)}{\mathcal{S}_n(x)} = \mathcal{P}_n \tag{4}$$

The function $\hbar_n(x)$ calculates an individual or item failure probability over a short period of time. The HF_n may increase, decrease, stay constant, or reflect a more complex process. In this way the suggested model is memoryless in the family of discrete probability distribution like the exponential distribution in the class of continuous distributions.

Several theorems can be used to establish statistical properties of the proposed distribution. Some of these theorems include the derivations of important statistical measures in neutrosophic framework that can help to understand the behavior of the distribution for analyzing the vague dataset. These theorems provide a solid foundation for making reliable inferences and drawing meaningful conclusions.

Theorem 1 If x follows the NGD then $E(\mathcal{X}) = \frac{1-\mathcal{P}_n}{\mathcal{P}_n}$

Proof: By definition, the mean of the NGD is given by:

$$E(\mathcal{X}) = \sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_n)^{\mathcal{X}}\mathcal{P}_n\mathcal{X}$$
$$= (1-\mathcal{P}_n)\mathcal{P}_n\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_n)^{\mathcal{X}-1}\mathcal{X}$$
$$= [(1-\mathcal{P}_l)\mathcal{P}_l\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_l)^{\mathcal{X}-1}\mathcal{X},\ (1-\mathcal{P}_{u.})\mathcal{P}_u\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_u)^{\mathcal{X}-1}\mathcal{X}] \tag{5}$$

Equation (5) further yielded:

$$(1-\mathcal{P}_l)\mathcal{P}_l\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_l)^{\mathcal{X}-1}\mathcal{X} = \frac{1-\mathcal{P}_l}{\mathcal{P}_l}$$

and

$$(1-\mathcal{P}_{u.})\mathcal{P}_u\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_u)^{\mathcal{X}-1}\mathcal{X} = \frac{1-\mathcal{P}_u}{\mathcal{P}_u}$$

So,

$\left[\frac{1-\mathcal{P}_l}{\mathcal{P}_l}, \frac{1-\mathcal{P}_u}{\mathcal{P}_u}\right] = \frac{1-\mathcal{P}_n}{\mathcal{P}_n}$, hence proved.

Theorem 2 If x follows the NGD, then $\tilde{V}_n(\mathcal{X}) = \frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}$ is the variance of the proposed model.

Proof: The variance of the NGD is given by:

$$\tilde{V}_n(x) = E(\mathcal{X}^2) - [E(\mathcal{X})]^2 \quad (6)$$

Now

$$E(\mathcal{X}^2) = \sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_n)^{\mathcal{X}}\mathcal{P}_n\mathcal{X}^2$$
$$= [(1-\mathcal{P}_{u.})^2\mathcal{P}_l\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_l)^{\mathcal{X}-1}\mathcal{X}^2,\ (1-\mathcal{P}_{u.})^2\mathcal{P}_u\sum_{\mathcal{X}=0}^{\infty}(1-\mathcal{P}_u)^{\mathcal{X}-1}\mathcal{X}^2] \quad (7)$$

Simplification of (7) provided:

$$\left[\frac{2-3\mathcal{P}_l+\mathcal{P}_l^2}{\mathcal{P}_l^2},\ \frac{2-3\mathcal{P}_u+\mathcal{P}_u^2}{\mathcal{P}_u^2}\right] = \frac{2-3\mathcal{P}_n+\mathcal{P}_n^2}{\mathcal{P}_n^2}$$

Thus (6) becomes:

$$\tilde{V}_n(x) = \left[\frac{1-\mathcal{P}_l}{\mathcal{P}_l^2},\ \frac{1-\mathcal{P}_u}{\mathcal{P}_u^2}\right] = \frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}$$

Theorem 3 Show that k^{th} moment of the NGD is $\frac{\mathcal{P}_n}{1-(1-\mathcal{P}_n)e^k}$

Proof: By definition the k^{th} moment of the NGD is given by:

$$\mu_{kn} = \sum_{\mathcal{X}=0}^{\infty} e^{k\mathcal{X}}(1-\mathcal{P}_n)^{\mathcal{X}}\mathcal{P}_n$$
$$= \mathcal{P}_n\sum_{\mathcal{X}=0}^{\infty}[e^k(1-\mathcal{P}_n)]^{\mathcal{X}}$$
$$= [\mathcal{P}_l\sum_{\mathcal{X}=0}^{\infty}[e^k(1-\mathcal{P}_l)]^{\mathcal{X}},\ \mathcal{P}_u\sum_{\mathcal{X}=0}^{\infty}[e^k(1-\mathcal{P}_u)]^{\mathcal{X}}] \quad (8)$$

From (8), we can write;

$$\mathcal{P}_l\sum_{\mathcal{X}=0}^{\infty}[e^k(1-\mathcal{P}_l)]^{\mathcal{X}} = \frac{\mathcal{P}_l}{1-(1-\mathcal{P}_l)e^k}$$

and

$$\mathcal{P}_u\sum_{\mathcal{X}=0}^{\infty}[e^k(1-\mathcal{P}_u)]^{\mathcal{X}} = \frac{\mathcal{P}_u}{1-(1-\mathcal{P}_u)e^k}$$

Hence

$\mu_{kn} = \left[\frac{\mathcal{P}_l}{1-(1-\mathcal{P}_l)e^k}, \frac{\mathcal{P}_u}{1-(1-\mathcal{P}_u)e^k}\right] = \frac{\mathcal{P}_n}{1-(1-\mathcal{P}_n)e^k}$ is required result.

$where\ k = 1,2,3,\dots$ is a general expression for the kth row moment about the origin of the NGD. By using the following relations, moments about the mean for NGD can be derived as:

$$\mu'_{1n} = \mu_{1n} = \frac{1-\mathcal{P}_n}{\mathcal{P}_n}$$
$$\mu'_{2n} = \mu_{2n} - (\mu_{1n})^2 = \frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}$$
$$\mu'_{3n} = \mu_{3n} - 3\mu_{2n}\mu_{1n} + 2(\mu_{1n})^3 = (1-\mathcal{P}_n)(1+(1-\mathcal{P}_n))\mathcal{P}_n$$
$$\mu'_{4n} = \mu_{4N} - 4\mu_{3n}\mu_{1n} + 6\mu_{2n}{\mu_{1n}}^2 - 3{\mu_{1n}}^4 = \left(\frac{9(1-\mathcal{P}_n^2)}{\mathcal{P}_n^4}\right) + \left(\frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}\right)$$

Theorem 4 The coefficient of skewness of the NGD is $\frac{(1+(1-\mathcal{P}_n))}{(1-\mathcal{P}_n)^{1/2}}$

Proof: By definition, the coefficient of skewness for NGD is given by:

$$\alpha_3 = \frac{\mu'_{3n}}{(\mu'_{2n})^{3/2}} \quad (9)$$

Where $\mu'_{3n} = (1-\mathcal{P}_n)(1+(1-\mathcal{P}_n))\mathcal{P}_n$ and $\mu'_{2n} = \frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}$

Substituting in (9) yielded;

$$\alpha_3 = \frac{(1+(1-\mathcal{P}_n))}{(1-\mathcal{P}_n)^{1/2}}$$

where $\alpha_3 \in [\alpha_l, \alpha_u]$.

Theorem 5 Show that the coefficient of kurtosis for NGD is $\left(9 + \mathcal{P}_n^2/1-\mathcal{P}_n\right)$

Proof: By definition, the coefficient of kurtosis is given by:

$$\alpha_4 = \frac{\mu'_{4n}}{{\mu'_{2n}}^2} \quad (10)$$

Where $\mu'_{4n} = \left(\frac{9(1-\mathcal{P}_n^2)}{\mathcal{P}_n^4}\right) + \left(\frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}\right)$ and $\mu'_{2n} = \frac{1-\mathcal{P}_n}{\mathcal{P}_n^2}$

Substituting in (10) yielded:

$\alpha_4 = \left(9 + {\mathcal{P}_n^2}/{1-\mathcal{P}_n}\right)$

where $\alpha_4 = [\alpha_l, \alpha_u]$.

In the same way, other important distributional properties can also be explored through the neutrosophic framework. These properties offer a comprehensive approach to analyzing uncertainties and vagueness.

3. Estimation Procedure

The maximum likelihood estimate (MLE) is a widely used method in many real-world applications. It aims to determine the parameter value(s) that provide the highest probability of the observed data occurring. In uncertain environments, MLE differs from the classical approach as it provides interval estimates of neutrosophic parameters instead of a single point estimate. This allows for a more comprehensive representation of uncertainty and variability in the data. By providing interval estimates, MLE under the neutrosophic structure accounts for the inherent ambiguity and imprecision present in uncertain environments, making it a valuable tool in decision-making processes. In this part, a well-known MLE technique is used to determine the neutrosophic parameter of the proposed NGD. The ML technique is defined by considering the parameters unknown and calculating the joint density of all observations is a dataset that are assumed to be identical and dispersed independently. Once the likelihood of the NGD is established, maxima of the function are determined. These ML estimators are essential in the statistical viewpoint because of minimal variance and asymptotic unbiasedness properties. Let $y_1, y_2, \dots, y_k$ are identical and independently observations from the k subjects which follow the parametric model given in (1) then the joint density is given by:

$$\mathcal{L}(\mathcal{P}_n|\,\mathcal{X}) = \prod_{i=1}^{k} g_n(\mathcal{X}|\mathcal{P}_n)$$
$$= \prod_{i=1}^{k} \mathcal{P}_n(1-\mathcal{P}_n)^{\mathcal{X}_i}$$
$$= \mathcal{P}_n \prod_{i=1}^{k} (1-\mathcal{P}_n)^{\mathcal{X}_i} \quad (11)$$

Taking the logarithm of (11) and symbolizing it by $\omega_n(\mathcal{T}_i|\mathcal{P}_n)$,

$$\omega_n(\mathcal{T}_i|\mathcal{P}_n) = log\left[\mathcal{P}_n \prod_{i=1}^{k}(1-\mathcal{P}_n)^{\mathcal{X}_i}\right] \quad (12)$$

Simplification of (12) yielded;

$$\omega_n(\mathcal{T}_i|\mathcal{P}_n) = klog(\mathcal{P}_n) + \left(\sum_1^k \mathcal{X}_i - k\right)\log(1-\mathcal{P}_n) \quad (13)$$

Partially differentiating (13) by unknown values and equating to zero implies:

$$\left[\frac{\delta\omega_n(\mathcal{X}_i,\mathcal{P}_n)}{\delta\mathcal{P}_n}\right] = 0 \quad (14)$$

Further solution of (14) provides the following estimates for unknown parameter of the NGD

$$\hat{\mathcal{P}}_n = \frac{k}{\left(\sum_1^k \mathcal{X}_i\right)} \quad (15)$$

Note that $\hat{\mathcal{P}}_n$ will be interval forms because of imprecise sample data.

This aligns with intuition because when observing a geometric random variable across k trials, the total number of successes observed is represented by the sum of individual trial outcomes $\sum_1^k \mathcal{X}_i$. By calculating the ratio of the number of successes to the total number of trials, we can estimate the probability $\mathcal{P}_n$. It is crucial to note that the maximum likelihood estimator (MLE) can be considered as a random variable since it is based on random data. Consequently, the MLE inherits the randomness of the underlying dataset from which it is derived. Let us take an example where we see that how the MLE estimation can be performed. We consider a situation where we assume that a manufacturing process that produce some specific items. We want to model the number of attempts needed to produce a defect produced by a manufacturing machine. In a sample of 10 attempts, we can record the number of attempts it took to produce a defective item for each attempt. For example, in this case, the recorded attempts are:

2, 5, [1,2], 3, [4,5] 2, 1, [6,7] 2, [5, 6]

Here some values such as [1,2], [4,5], [6,7] and [5,6] are imprecise. Here the value for instance [4,5] means that position of the defective item is not clearly defined. The same holds for other imprecise items. Now this data the unknown neutrosophic parameter can be estimated as:

The above data can further be written as:

[2, 2], [5, 5], [1, 2], [3, 3], [4, 5], [2, 2], [1, 1], [6, 7], [2, 2], [5, 6]

By using (15) the $\mathcal{P}_n$ can be estimated as:

$$\hat{\mathcal{P}}_n = \frac{10}{[\sum_1^k \mathcal{X}_{il}, \sum_1^k \mathcal{X}_{iu}]}$$

where $\sum_1^k \mathcal{X}_{il}$ and $\sum_1^k \mathcal{X}_{iu}$ are lower and upper values of the neutrosophic data.

Thus,

$$\hat{\mathcal{P}}_n = \frac{10}{[31,35]} \cong [0.28, 0.32]$$

Hence the estimated imprecise value lies between 0.28 and 0.32.

4. Random Data Generation

We may require information on the number of trials needed to achieve a 25%, 50% or 75% probability of success occurrence. For example in a production line where there is a 5% defective rate, we aim to determine the minimum number of inspections, denoted as a , required to ensure that the probability of observing at least one defective item reaches or exceeds 50%.

To find a such that

$$p(\mathcal{X} \leq a) \geq 0.50$$

where a is known as the 50% quantile of geometric distribution.

Generally, k percentile provides the minimum interval value of a such that

$$p(\mathcal{X} \leq a) \geq k/100 \quad (16)$$

Equation (16) can be expressed as:

$$1 - (1 - \mathcal{P}_n)^a \geq \frac{k}{100} \quad (17)$$

Further simplification of (17) yields:

$$a \leq \frac{\ln(1 - \frac{k}{100})}{\ln((1 - \mathcal{P}_n))} \quad (18)$$

Solution of (8) provides the minimum interval value of a. For example the 50% quantile for defective rate $\mathcal{P}_n = [0.1, 0.15]$ can be found utilizing (8) as:

$$a \leq \frac{\ln(1 - 0.5)}{\ln(1 - [0.1, 0.15])} \cong [4, 6]$$

This means that there is at least 50% chance to get the first success in the trial interval [4, 6]. In general the inverse distribution can be used to produce random neutrosophic variable from the model as:

$$\mathcal{G}_n(\mathcal{X})^{-1} = \frac{ln(1-u)}{ln(1-\mathcal{P}_n)}; \quad 0 < u < 1. \quad (19)$$

The (19) based on inverse transformation method and can used to generate random data from the proposed NGD.

By taking the value $\mathcal{P}_n = [0.2, 0.4]$, exact mean and variance from Theorem 1 and Theorem 2 can be calculated as follows:

$$E(\mathcal{X}) = \frac{1 - \mathcal{P}_n}{\mathcal{P}_n}$$

$$= \frac{1 - [0.2,0.4]}{[0.2,0.4]}$$

$$E(\mathcal{X}) = [1.5, 4]$$

$$V(\mathcal{X}) = \frac{1 - \mathcal{P}_n}{\mathcal{P}_n^{\ 2}}$$

$= \frac{1-[0.2,0.4]}{[0.2,0.4]^2}$

$V(\mathcal{X}) = [3.75, 20]$

Thus exact mean and variance of the proposed distribution by considering $\mathcal{P}_n = [0.2,0.4]$ are $[1.5, 4]$ and $[3.75, 20]$ respectively. Now we will see that our simulation results are also in close approximation to exact values.

The study uses a larger sequence of random numbers generated from 10,000 Monte Carlo simulations to estimate the parameter of a proposed model. The parameter range considered in the study is between 0.2 and 0.4. To obtain simulated results, a program written in R is utilized. Additionally, the program utilizes the "moments" package to analyze moment-based characteristics of the proposed distribution. The larger sequence of random numbers generated from 10,000 Monte Carlo simulations allows for a more accurate estimation of the parameter in the proposed model. By considering a parameter range between 0.2 and 0.4, the study ensures a comprehensive analysis of the distribution's characteristics. Table 1 displays the estimation results of the NGD parameter using the generated simulated data.

Table 1: Summary statistics of the NGD based on simulated data.

Properties	Estimated values
MLE Estimate	[0.20, 0.40]
Mean	[2.49, 4.98]
Variance	[3.73, 19.93]
Skewness	[2.05, 2.00]
Quartile 1	[1, 2]
Quartile 2	[2, 4]
Quartile 3	[3, 7]

The results in Table 1 show that due to uncertainty in the parameter of NGD, the characteristics of the distribution are interval based and imprecise. Furthermore, the simulated results closely approximate the true characteristics of the distribution.

5. Real Data Applications

In this section, some numerical examples have been considered to illustrate the application of the concepts discussed in this work. These examples serve to provide a practical understanding of how the concepts can be applied in real-life scenarios. By showcasing numerical calculations and their corresponding interpretations, readers can better grasp the significance and implications of the discussed concepts.

Example1: Assume that a production machine has a faulty rate ranging from 5% to 8%. Considering the unknown defective rate of the machine's products, which ranges from$\mathcal{P}_n = [0.05, 0.08]$, what is the probability range for the occurrence of the first defective item in the third inspection?

Given the defective rate $\mathcal{P}_n = [0.05, 0.08]$

Let $\mathcal{X}$ be the neutrosophic random variable which denotes the number of defective items produced by the machine.

Neutrosophically, the defective rate is in the range $\mathcal{P}_n = [0.05, 0.08]$, signifying the uncertainty or imprecision in the defective rate of the machine's products.

The probability of the first defective occurring in the third item can be calculated under this interval probability as:

$p(\mathcal{X} = 3) \qquad = [0.05, 0.08]\big[1 - [0.05, 0.08]\big]^2$
$= [0.05, 0.08][0.8464, 0.9025]$
$= [0.042, 0.072]$

By evaluating these expressions we found the probability range of [0.042,0.072] for the occurrence of the first defective item in the third position. This range takes into consideration the imprecision or uncertainty in the defective rate, which falls between 5% and 8%.

Example 2: There is an estimated possibility of [0.4, 0.6] in a specific Malaysian city that a randomly selected individual owns a motorcycle. What is the likelihood that the first motorbike owner to be encountered among the first four people interviewed in this city will be the fourth person interviewed?

We must take into account this interval in the neutrosophic context, where the estimated likelihood of owning a motorcycle is between 0.4 and 0.6 (i.e., $\mathcal{P}_n = [0.4, 0.6]$), to determine the range of probabilities for the occurrence.

Let $\mathcal{X}$ be the random variable that denotes the number of people having motorbike.

$p(\mathcal{X} = 4) = [0.4, 0.6]\big[1 - [0.4, 0.6]\big]^3$
$= [0.4, 0.6][0.064, 0.216]$
$= [0.0256, 0.1296]$

Thus, assessing these expressions according to neutrosophic arithmetic rules will yield a range of probability $[0.0256, 0.1296]$, taking into consideration the imprecision or uncertainty in the estimated likelihood of motorcycle ownership between 0.4 and 0.6, in the event that the fourth interviewee is the first to have a motorcycle.

Example 3: Calculate the probability of a student pilot passing the written test for a private pilot's license on their third attempt, assuming that the probability of passing the test is between 0.2 and 0.3.

Let $\mathcal{X}$ be the neutrosophic random variable that denotes the number of attempts a student makes to pass this test.

Now the required probability can be obtained as:

$p(\mathcal{X} = 4) = [0.2, 0.3]\big[1 - [0.2, 0.3]\big]^2$
$= [0.2, 0.3][0.49, 0.64]$
$= [0.098, 0.192]$

Example 4: What is the neutrosophic probability of encountering the first defective product within the initial six inspections, given a defective rate ranging from 0.03 to 0.05?

To solve this problem, we need to find involve the neutrosophic distribution function as described in (3).

$p(\mathcal{X} \leq 6) = 1 - [1 - \mathcal{P}_n]^6$
where $\mathcal{P}_n = [0.03, 0.05]$
Now

$p(\mathcal{X} \leq 6) = 1 - \big[1 - [0.03, 0.05]\big]^6$
$= 1 - [0.95, 0.97]^6$
$= [0.167, 0.265]$

Based on the provided imprecise defective rate (0.03 to 0.05), the neutrosophic probability of first defective item out of six inspected items fall between 0.0167 and 0.265. This range indicates that there is a relatively low probability of encountering the first defective item, but it is not entirely unlikely.

6. Concluding Remarks

The neutrosophic geometric distribution (NGD) is a revolutionary framework that has been introduced in this research. It is derived from the classical geometric distribution and aims to handle

imprecise data analysis. By doing so, it offers a reliable and generalized method for conducting modern statistical investigations for another class of data. We have extensively examined the basic characteristics of the NGD in a neutrosophic setting and clarified its essential reliability functions. To make it more useful in real-life situations, we have devised most method of the ML estimation. The effectiveness of this technique in determining the NGD parameters has been demonstrated through several numerical instances, proving its applicability in real-world situations. Furthermore, we have focused on developing the NGD's quantile function by the inverse cumulative function method. This function enabled us to generate simulation data, serving as a valuable tool for estimating parameter and providing insightful summary statistics on the behavior of the proposed model. We have considered real-life situations to demonstrate the application of NGD and enhance the comprehension of its theoretical concepts.
Furthermore, our research acts as a connection between classical structures and the innovative neutrosophic framework, enabling future developments in extending geometric distribution to neutrosophic domain and exploring its diverse applications.

Acknowledgement
This study is supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/2023/R/1445).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abbas, N. (2023). On classical and bayesian reliability of systems using bivariate generalized geometric distribution. Journal of Statistical Theory and Applications, 22(3), pp. 1-9. https://doi.org/10.1007/s44199-023-00058-4
2. Akram, M., Habib, S., & Javed, I. (2014). Intuitionistic fuzzy logic control for washing machines. Indian Journal of Science and Technology, 7(5), p.654. https://doi.org/10.17485/ijst/2014/v7i5.20
3. Almazah, M. M. A., Erbayram, T., Akdoğan, Y., Al Sobhi, M. M., & Afify, A. Z. (2021). A new extended geometric distribution: Properties, regression model, and actuarial applications. Mathematics, 9(12)., p.1336. https://doi.org/10.3390/math9121336
4. Bhati, D., & Joshi, S. (2018). Weighted geometric distribution with new characterizations of geometric distribution. Communications in Statistics - Theory and Methods, 47(6), pp. 1510-1527. https://doi.org/10.1080/03610926.2017.1321765
5. Smarandache, F. Neutrosophy: Neutrosophic probability, set, and logic: analytic synthesis and synthetic analysis; American Research Press: Santa Fe, NM, USA,1998 ; pp. 25–38.
6. Céspedes-Mota, A., Castañón, G., Martínez-Herrera, A. F., Cárdenas-Barrón, L. E., & Sarmiento, A. M. (2018). Differential evolution algorithm applied to wireless sensor distribution on different geometric shapes with area and energy optimization. Journal of Network and Computer Applications, 119. https://doi.org/10.1016/j.jnca.2018.06.006
7. Chakraborty, S., Ong, S. H., & Biswas, A. (2023). An extension of the geometric distribution with properties and applications. Austrian Journal of Statistics, 52(3), pp. 124-142. https://doi.org/10.17713/ajs.v52i3.1487
8. Chesneau, C., Bakouch, H. S., Hussain, T., & Para, B. A. (2021). The cosine geometric distribution with count data modeling. Journal of Applied Statistics, 48(1), pp. 124-137. https://doi.org/10.1080/02664763.2019.1711364

9. Ghosh, I., Marques, F., & Chakraborty, S. (2021). A bivariate geometric distribution via conditional specification: properties and applications. Communications in Statistics: Simulation and Computation., pp. 01-21. https://doi.org/10.1080/03610918.2021.2004419
10. Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. Test, 19(2) pp. 399-415. https://doi.org/10.1007/s11749-009-0169-3
11. Jayakumar, K., & Babu, M. G. (2018). Discrete Weibull geometric distribution and its properties. Communications in Statistics - Theory and Methods, 47(7), pp. 1767-1783. https://doi.org/10.1080/03610926.2017.1327074
12. Kabir, S., & Papadopoulos, Y. (2018). A review of applications of fuzzy sets to safety and reliability engineering. International Journal of Approximate Reasoning, 100, pp. 29-55. https://doi.org/10.1016/j.ijar.2018.05.005
13. Khan, Z., Al-Bossly, A., Almazah, M. M. A., & Alduais, F. S. (2021). On Statistical Development of Neutrosophic Gamma Distribution with Applications to Complex Data Analysis. Complexity, 2021, pp. 1-8. https://doi.org/10.1155/2021/3701236
14. Khan, Z., Gulistan, M., Kausar, N., & Park, C. (2021a). Neutrosophic Rayleigh model with some basic characteristics and engineering applications. IEEE Access, 9, pp. 71277–71283. https://doi.org/10.1109/ACCESS.2021.3078150
15. Liu, Q., & Xia, A. (2022). Geometric sums, size biasing and zero biasing. Electronic Communications in Probability, 27, pp. 1-3. https://doi.org/10.1214/22-ECP462
16. Louzada, F., Ramos, P. L., & Perdoná, G. S. C. (2016). Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data. Computational and Mathematical Methods in Medicine, 2016. https://doi.org/10.1155/2016/8727951
17. Mohamed Abdel-Basset, Mai Mohamed, Florentin Smarandache (2020), Comment on "a novel method for solving the fully neutrosophic linear programming problems: suggested modifications , Vol. 31, Neutrosophic Sets and Systems
18. Okorie, I. E., Akpanta, A. C., Ohakwe, J., Chikezie, D. C., & Onyemachi, C. U. (2019). On the Rayleigh-geometric distribution with applications. Heliyon, 5(8). https://doi.org/10.1016/j.heliyon.2019.e02200
19. Pankaj, Bhatti, J., & Kakkar, M. K. (2022). Reliability analysis of active redundant system using geometric distribution. Malaysian Journal of Science, 41(3) pp. 49-57. https://doi.org/10.22452/mjs.vol41no3.7
20. Pankaj, Bhatti, J., & Kakkar, M. K. (2023). Profit analysis to an industrial system possessing active redundancy form using geometric distribution. Contemporary Mathematics (Singapore), 4(3), pp. 423-434. https://doi.org/10.37256/cm.4320232515
21. Park, C., Gou, K., & Wang, M. (2022). A study on estimating the parameter of the truncated geometric distribution. In American Statistician, 76(3), pp. 257-261. https://doi.org/10.1080/00031305.2022.2034666
22. Peng, B., Xu, Z., & Wang, M. (2019). The exponentiated lindley geometric distribution with applications. Entropy, 21(5), p. 510. https://doi.org/10.3390/e21050510
23. Perrier, E., Tao, D., & Ferrie, C. (2020). Quantum geometric machine learning for quantum circuits and control. New Journal of Physics, 22(10), p.103056. https://doi.org/10.1088/1367-2630/abbf6b
24. Puri, m. L., Ralescu, d. A., & Zadeh, l. (1993). Fuzzy random variables. readings in fuzzy sets for intelligent systems, pp. 265–271. https://doi.org/10.1016/B978-1-4832-1450-4.50029-8
25. 25. Ramadan, A. T., Tolba, A. H., & El-Desouky, B. S. (2022). A Unit Half-Logistic Geometric Distribution and Its Application in Insurance. Axioms, 11(12), p. 676. https://doi.org/10.3390/axioms11120676
26. 26. Shah, F., Khan, Z., Aslam, M., & Kadry, S. (2021). Statistical development of the VSQ -control chart for extreme data with an application to the carbon fiber industry, pp. 1-11. Mathematical Problems in Engineering, 2021. https://doi.org/10.1155/2021/9766986
27. 27. F. Smarandache, Introduction to Neutrosophic Statistics. Craiova: Sitech & Education Publishing, 2014.

28. Smarandache, F., Broumi, S., Singh, P. K., Liu, C., Venkateswara Rao, V., Yang, H.-L., Patrascu, I., & Elhassouny, A. (2019). Introduction to neutrosophy and neutrosophic environment. Neutrosophic Set in Medical Image Analysis, 3–29. https://doi.org/10.1016/B978-0-12-818148-5.00001-1
29. 29. Smarandache, F., & Said, B. (2020). Neutrosophic Theories in Communication, Management and Information Technology. https://philpapers.org/rec/SMANTI
30. 30. Susilo, T., Widiarti, Kurniasari, D., & Aziz, D. (2021). Characteristics of Bayes estimator in the geometric distribution with prior beta. Journal of Physics: Conference Series, 1751(1), p. 012020. https://doi.org/10.1088/1742-6596/1751/1/012020
31. 31. Thelwell, M., Bullas, A., Kühnapfel, A., Hart, J., Ahnert, P., Wheat, J., Loeffler, M., Scholz, M., & Choppin, S. (2022), Modelling of human torso shape variation inferred by geometric morphometrics. PLoS ONE, 17, p.e0265255. https://doi.org/10.1371/journal.pone.0265255
32. 32. R. Jansi, K. Mohana, Florentin Smarandache (2019), Correlation Measure for Pythagorean Neutrosophic Sets with T and F as Dependent Neutrosophic Components , Neutrosophic Sets and Systems: Vol. 30, Neutrosophic Sets and Systems
33. 33. Florentin Smarandache (2016), Classical Logic and Neutrosophic Logic. Answers to K. Georgiev , Neutrosophic Sets and Systems: Vol. 13, Neutrosophic Sets and Systems
34. 34. Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1(1), pp. 3–28. https://doi.org/10.1016/0165-0114(78)90029-5
35. 35. Florentin Smarandache (2018), Plithogenic Set, an Extension of Crisp, Fuzzy, Intuitionistic Fuzzy, and Neutrosophic Sets - Revisited , Vol. 21, Neutrosophic Sets and Systems

Received: Aug 19, 2023. Accepted: Dec. 21, 2023