AIMS Mathematics

AIMS Mathematics, 9(6): 16436–16452.
DOI: 10.3934/math.2024796
Received: 26 February 2024
Revised: 17 April 2024
Accepted: 23 April 2024
Published: 10 May 2024

http://www.aimspress.com/journal/Math

Research article

Neutrosophic geometric distribution: Data generation under uncertainty and practical applications

Muhammad Aslam* and Mohammed Albassam

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 215511, Saudi Arabia

* **Correspondence:** aslam_ravian@hotmail.com.

Abstract: This paper introduces the geometric distribution in the context of neutrosophic statistics. The research outlines the essential properties of this new distribution and introduces novel algorithms for generating imprecise geometric data. The study explores the practical applications of this distribution in the industry, highlighting differences in data generated under deterministic and indeterminate conditions using detailed tables, simulation studies, and real-world applications. The results indicate that the level of uncertainty has a substantial impact on data generation from the geometric distribution. These findings suggest updating classical statistical algorithms to better handle the generation of imprecise data. Therefore, decision-makers should exercise caution when using data from the geometric distribution in uncertain environments.

Keywords: simulation; algorithm; neutrosophic data; classical statistics; industry
Mathematics Subject Classification: 62A86

1. Introduction

The geometric distribution is a discrete probability distribution employed to model the occurrence of the first success in a series of Bernoulli trials. Widely recognized in statistics, this distribution finds applications across various industries and fields. Control chart design involves assessing the distribution of the average run length, a task often accomplished using the geometric distribution. This distribution and its extended variants have found utility in diverse areas. For instance, Beria [1] delved into the confidence interval of the geometric distribution, while Chen [2] investigated goodness-of-fit tests for the geometric distribution. Bertoli-Barsotti and Lando [3] applied the geometric distribution to citation analysis, while Slim et al. [4] utilized it for graduation rate analysis. Gao [5] proposed an extension of the geometric distribution for a replacement policy, and Altun [6] presented a generalization with discussions on its applications. Additionally, Almazah et al. [7] introduced an extended geometric distribution, applying it to various real datasets. Zhang et al. [8] explored the application of geometric distribution in magnetic fields. Ghosh et al. [9] introduced the bivariate geometric distribution, discussing its potential applications. Abbas [10] proposed a Bayesian approach to the bivariate geometric distribution, applied in reliability analysis. Further instances of the geometric distribution and its extended forms can be found in the works of Akdoğan et al. [11], Nadarajah and Bakar [12], Hassan and Abdelghafar [13], and Ramadan et al. [14]. The broad range of applications underscores the versatility and significance of the geometric distribution in statistical analyses and modeling.

The credit for introducing neutrosophic statistics goes to Smarandache [15], who developed neutrosophic descriptive statistics. Neutrosophic statistics, applicable under uncertainty and characterized by the degree of indeterminacy, provide additional information. This approach is flexible and well-suited for analyzing imprecise data. It is important to note that neutrosophic statistical results revert to classical statistics when the data contains no uncertainty. Classical statistics is traditionally used to analyze crisp data, where values are precise and unambiguous. In contrast, neutrosophic statistics is designed to handle incomplete, ambiguous, and indeterminate data. Neutrosophic statistics offers a more robust approach to analyzing uncertain data compared with classical statistics. Neutrosophic statistics finds applications across a wide range of fields, including social sciences, political science, machine learning, artificial intelligence, medical science, and environmental studies. For instance, in environmental studies, climate pattern data often contains gaps and uncertainties. Neutrosophic statistics can be employed to analyze such data, including climate patterns and pollution levels, effectively capturing its inherent uncertainties and complexities. In a significant contribution, Florentin Smarandache [16] demonstrated the efficiency of neutrosophic statistical analysis for neutrosophic data compared to interval statistics. Chen et al. [17] as well as Chen et al. [18] explored neutrosophic statistical methods for analyzing neutrosophic data. Duan et al. [19]. Granados [20] introduced the neutrosophic geometric distribution using the neutrosophic random variable, and Granados et al. [21] presented various neutrosophic statistical distributions. Aslam [22] introduced sine-cosine and convolution methods for generating neutrosophic data. Aslam [23] introduced a neutrosophic simulation procedure for the DUS-Weibull distribution. Aslam [24] presented an algorithm for generating data from the Weibull distribution. Aslam and Alamri [25] proposed an algorithm for generating neutrosophic data using the accept-reject method. Jdid et al. [26] conducted a study on the neutrosophic simulation process for the exponential distribution. These contributions collectively highlight the growing importance and diverse applications of neutrosophic statistics in handling uncertain and imprecise data.

While there is a considerable amount of literature on geometric distribution, the traditional geometric distribution in classical statistics is not suitable for handling imprecise data. A review of existing literature did not yield any studies on the geometric distribution using neutrosophic statistics. In this paper, we introduce the geometric distribution using the innovative concept of the neutrosophic random variable. Our investigation highlights a research gap in the development of algorithms for generating imprecise geometric data. To fill this gap, we introduce the neutrosophic random variable and propose the neutrosophic geometric distribution. We characterize this new geometric distribution by discussing its fundamental properties. Additionally, we present algorithms tailored for this distribution to generate imprecise data. We demonstrate the efficiency of the proposed algorithms through simulation studies and showcase the application of this new distribution in the industry. We expect that uncertainty will play a significant role in the generation of geometric data within this novel framework.

2. Neutrosophic random variable

Suppose that $X_L$be a random variable having mean μ and variance σ^2 based on this, we define a neutrosophic random variable $X_N = X_L + X_L I_N; I_N \epsilon [I_L, I_U]$ the neutrosophic random variable, where $X_L I_N$ be the indeterminate part and $I_N \epsilon [I_L, I_U]$ be the indeterminacy. It is worth noting that the introduced neutrosophic random variable is an extension of the classical random variable. When $I_L = 0$, the proposed neutrosophic random variable reverts to the classical random variable. As mentioned in Granados [20], neutrosophic logic is the generalization of fuzzy logic where an indeterminacy component is added additionally. Note that $I_N^2 = I_N, \ldots, I_N^n = I_N, 0.I_N = 0; n\epsilon\mathrm{N}$. Based on this information, the expected properties of the neutrosophic random variable $X_N = X_L + X_L I_N$ are given as

1) $E(X_N) = E(X_L + X_L I_N) = (1 + I_N)\mu$.
2) $E(X_N + c) = (X_L + X_L I_N) + c = (1 + I_N)\mu + c$, where c is a constant.
3) $E(aX_N + c) = a(X_L + X_L I_N) + c = a(1 + I_N)\mu + c$, where a and c is are constant.
4) For two random variables X_N and Y_N; $E(X_N + Y_N) = (1 + I_N)\mu_y + (1 + I_N)\mu_x$.

Based on this information, the variance properties of the neutrosophic random variable $X_N = X_L + X_L I_N$ are given as

1) $Var(X_N) = Var(X_L + X_L I_N) = (1 + I_N)^2\sigma^2$.
2) $Var(aX_N) = a^2 Var(X_L + X_L I_N) = a^2(1 + I_N)^2\sigma^2$.
3) $Var(aX_N + bY_N) = a^2(1 + I_N)^2\sigma_x^2 + b^2(1 + I_N)^2\sigma_y^2 + 2abI_N Cov(X_N, Y_N)$.
4) For two random variables X_N and Y_N; $Var(X_N + Y_N) = (1 + I_N)^2\sigma_x^2 + (1 + I_N)^2\sigma_y^2 + 2I_N Cov(X_N, Y_N)$.
5) For two independent random variables X_N and Y_N; $Var(X_N + Y_N) = (1 + I_N)^2\sigma_x^2 + (1 + I_N)^2\sigma_y^2$.

3. Neutrosophic geometric distribution

Consider the random variable X_L representing a deterministic value, utilized to determine the number of successes with a success probability of p. Let $X_N = X_L + X_L I_N$, where $X_L I_N$ represents the indeterminate part and $I_N \epsilon [I_L, I_U]$ signifies the indeterminacy. The neutrosophic random variable is an extension of the classical random variable. It converges to a classical random variable when I_L=0.

Theorem 3.1. The probability function of X_N is expressed as follows:

$$f\left(\frac{X_N}{(1+I_N)}\right) = p(1-p)^{\frac{X_N - 1 - I_N}{(1+I_N)}}; X_N = (1 + I_N), 2(1 + I_N), \ldots.. \quad (1)$$

Proof: As the Bernoulli random variables exhibit independence, the probability of the kth trial being the initial success is expressed as

$$P\left(B_{(1+I_N)}=0\right).,,,.P\left(B_{\left(\frac{X_N}{(1+I_N)}\right)}=0\right).P\left(B_{\left(\frac{X_N}{(1+I_N)}+1\right)}=1\right)$$

$$=p(1-p)^{\frac{X_N-1-I_N}{(1+I_N)}}.$$

Note that p represents the probability of success, while $(1-p=q)$ denotes the probability of failure.

Theorem 3.2. Show that the proposed distribution constitutes a complete probability mass function.

Proof: We know

$$\sum_{X_N=(1+I_N)}^{\infty} f(X_N)=p\sum_{X_N=(1+I_N)}^{\infty}(q)^{\frac{X_N-1-I_N}{(1+I_N)}}$$

$$=p\left[q^0+q^{\frac{2+I_N}{(1+I_N)}}+q^{\frac{3+I_N}{(1+I_N)}+\cdots}\right]$$

$$=p\left[\frac{1}{1-q}\right]=1.$$

The cumulative distribution function of X_N is given by

$$F\left(\frac{X_N}{(1+I_N)}\right)=1-(1-p)^{\frac{X_N}{(1+I_N)}};\ X_N=(1+I_N),2(1+I_N),\ldots \tag{2}$$

The expected value of X_N is given by

$$E(X_N)=E(X_L+X_LI_N)=(1+I_N)/p. \tag{3}$$

The higher order moment is expressed by

$$E(X_N^r)=(1+I_N)^r/p. \tag{4}$$

The variance of X_N is given by

$$Var(X_N)=Var(X_L+X_LI_N)=(1+I_N)^2(1-p)/p^2. \tag{5}$$

Theorem 3.3. Show that the moment generating function (mgf) of the neutrosophic geometric function is given as

$$M(t_N)=(1+I_N)pe^{t_N}\frac{1}{1-e^{t_N}(1-p)}.$$

Proof: We define

$$M(t_N)=(1+I_N)p\sum_{X_N=(1+I_N)}^{\infty}e^{\frac{t_NX_N}{(1+I_N)}}(1-p)^{\frac{X_N-1-I_N}{(1+I_N)}}$$

$$=(1+I_N)pe^{t_N}\sum_{X_N=(1+I_N)}^{\infty} e^{t_N\left(\frac{X_N}{(1+I_N)}-1\right)}(1-p)^{\frac{X_N-1-I_N}{(1+I_N)}}$$

$$=(1+I_N)pe^{t_N}\sum_{X_N=(1+I_N)}^{\infty} e^{t_N\left(\frac{X_N-1-I_N}{(1+I_N)}\right)}(1-p)^{\frac{X_N-1-I_N}{(1+I_N)}}.$$

$$M(t_N)=\frac{pe^{t_N(1+I_N)}}{1-e^{t_N}(1-p)}. \quad (6)$$

Theorem 3.4. Let $X_N = X_L + X_L I_N$ and $Y_N = Y_L + Y_L I_N$ be two independent neutrosophic random variable. Let $Z_N = X_N + Y_N$, the mgf is given by

$$M_{Z_N}(t_N) = M_{X_N}(t_N).M_{Y_N}(t_N).$$

Proof:

$$M_{Z_N}(t_N) = E(e^{t_N Z_N}) = E\left(e^{t_N(X_N+Y_N)}\right) = E(e^{t_N X_N}).E(e^{t_N Y_N}).$$

$$M_{Z_N}(t_N) = M_{X_N}(t_N).M_{Y_N}(t_N). \quad (7)$$

4. The proposed algorithm-I

The algorithm presented in Thomopoulos [27], categorized under classical statistics, is effective for generating deterministic data. Nevertheless, its utility is constrained in uncertain environments. To evaluate the effects of uncertainty, it is crucial to augment the current algorithm by integrating features that account for uncertainty. To accomplish this, we will implement the following process to generate data utilizing the geometric distribution.

Step-1: Specify the degree of uncertainty I_N.

Step-2: Generate random samples from continuous uniform distribution $u \sim U(0,1)$.

Step-3: Compute $X_N = inetger\ [\{([\ln(1-u)/\ln(1-p)]+1)\}(1+I_N)]$.

Step-4: Return X_N.

The algorithm is explained with the aid of Figure 1.

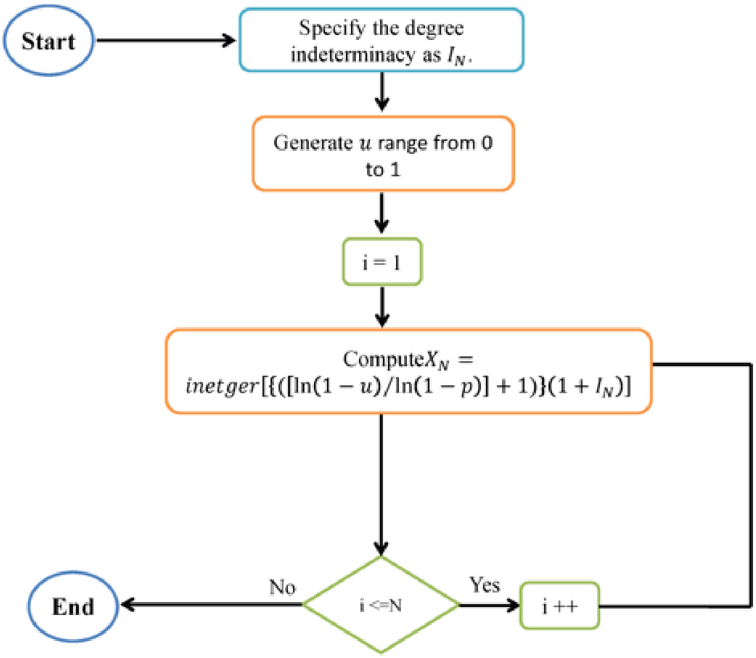

Figure 1. The proposed algorithm-I.

It is worth noting that the suggested algorithm-I simplifies to the algorithm outlined in Thomopoulos [27] when the parameter I_L is set to zero.

5. The proposed algorithm-II

Within this section, we introduce Algorithm-II, designed to generate data from the geometric distribution using $u \sim U(0,1)$ and the probability of the first success. This algorithm is an extension of the form presented in classical statistics. When I_L=0, the proposed Algorithm-II simplifies the algorithm used for generating data from the negative binomial distribution. The subsequent routine will be employed for generating data from the geometric distribution.

Step-1: Specify the degree of uncertainty I_N.
Step-2: Generate random samples from continuous uniform distribution $u \sim U(0,1)$.
Step-3: Compute $X_N = inetger\ [\{([\ln(u)/\ln(q)])\}(1+I_N)]$.
Step-4: Return X_N.

The algorithm is elucidated with the assistance of Figure 2.

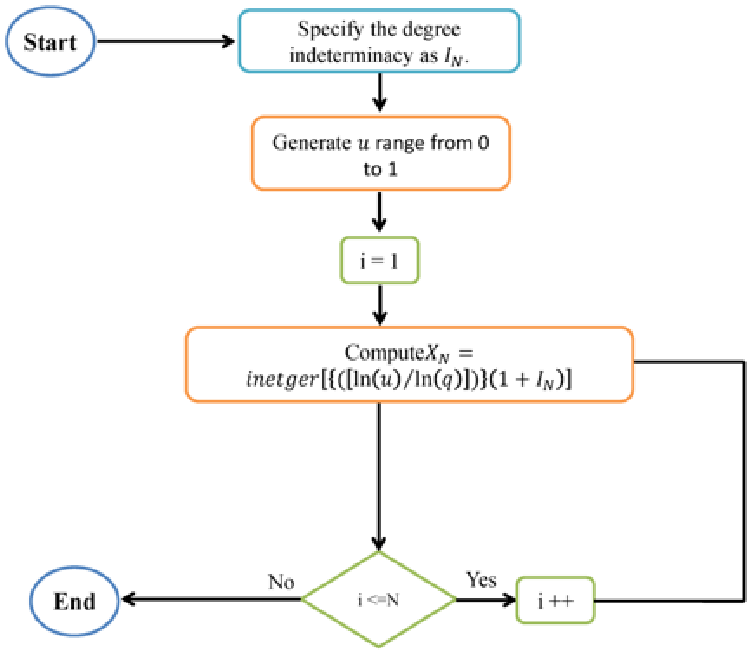

Figure 2. The proposed algorithm-II.

Note that the proposed algorithm-I and algorithm-II are slightly more complex than the classical statistics-based algorithms. This complexity arises because our algorithm-I and algorithm-II incorporate the degree of uncertainty in data generation, whereas the existing algorithms do not account for this uncertainty.

6. Simulation

Within this segment, we shall showcase simulation studies employing the proposed algorithms. Initially, we will delve into the data generation process using algorithm-I for the geometric distribution, followed by the presentation of data generated using algorithm-II for the same distribution.

6.1. *Simulation using algorithm-I*

Utilizing algorithm-I, the simulation procedure is executed, and data derived from the geometric distribution for different combinations of p and I_N are outlined in Tables 1–4. Specifically, Table 1 corresponds to p=0.10, Table 2 to p=0.20, Table 3 to p =0.30, and Table 4 to p =0.40. A discernible ascending pattern is evident in the geometric distribution data across Tables 1–4 as I_N values escalate from 0.1 to 0.9. For instance, with I_N=0.1 and p=0.10, X_N yields a value of 4, whereas with I_N=0.9 and p=0.10, X_N attains a value of 7. This progression is visually depicted in Figure 3, where it is observable that the data curve for I_N=0.1 is positioned below that of I_N=0.5 and I_N=0.9. Presently, we illustrate the data trend of X_N while holding I_N constant at 0.50, with varying values of p set at 0.10, 0.20, 0.30, and 0.40 in Figure 4. Upon examination of Figure 4, it is evident that X_N values exhibit a decline as the parameter p increases from 0.10 to 0.40. To illustrate, when p=0.10 and

I_N=0.50, X_N attains a value of 5, whereas with p=0.40 and I_N=0.50, X_N diminishes to a value of 2.

Table 1. Random variates using algorithm-I when $p = 0.10$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
3	4	4	5	5	5	6	6	7	7
6	7	7	8	9	9	10	11	11	12
2	3	3	3	3	4	4	4	5	5
32	35	39	42	45	49	52	55	58	62
18	20	21	23	25	27	29	30	32	34
14	15	17	18	20	21	22	24	25	27
15	17	18	20	22	23	25	26	28	30
8	9	9	10	11	12	13	13	14	15
2	2	2	3	3	3	3	4	4	4
8	8	9	10	11	12	12	13	14	15
1	1	1	1	1	1	1	1	2	2
17	18	20	22	24	25	27	29	31	32
20	22	24	26	28	30	32	34	36	38
31	34	37	40	43	46	49	52	56	59
16	18	19	21	23	24	26	28	29	31

Table 2. Random variates using algorithm-I when $p = 0.20$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
2	2	2	3	3	3	3	4	4	4
3	3	4	4	5	5	5	6	6	6
1	2	2	2	2	2	2	3	3	3
15	17	19	20	22	23	25	27	28	30
9	10	10	11	12	13	14	15	16	17
7	8	8	9	10	10	11	12	13	13
7	8	9	10	11	11	12	13	14	15
4	4	5	5	6	6	7	7	7	8
1	1	1	2	2	2	2	2	2	3
4	4	5	5	6	6	6	7	7	8
1	1	1	1	1	1	1	1	1	2
8	9	10	11	12	12	13	14	15	16
10	11	12	13	14	15	16	17	18	19
15	16	18	19	21	22	24	25	27	28
8	9	10	10	11	12	13	14	15	15

Table 3. Random variates using algorithm-I when $p = 0.30$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
1	2	2	2	2	2	3	3	3	3
2	2	3	3	3	3	4	4	4	5
1	1	1	1	2	2	2	2	2	2
10	11	12	13	14	15	16	17	18	19
6	6	7	7	8	9	9	10	10	11
4	5	5	6	6	7	7	8	8	9
5	5	6	6	7	8	8	9	9	10
3	3	3	4	4	4	5	5	5	5
1	1	1	1	1	2	2	2	2	2
3	3	3	4	4	4	4	5	5	5
1	1	1	1	1	1	1	1	1	1
5	6	6	7	8	8	9	9	10	11
6	7	8	8	9	10	10	11	12	12
9	10	11	12	13	14	15	16	17	18
5	6	6	7	7	8	8	9	10	10

Table 4. Random variates using algorithm-I when $p = 0.40$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
1	1	1	2	2	2	2	2	2	3
2	2	2	2	3	3	3	3	3	4
1	1	1	1	1	2	2	2	2	2
7	8	9	9	10	11	12	12	13	14
4	5	5	5	6	6	7	7	8	8
3	4	4	4	5	5	5	6	6	7
4	4	4	5	5	6	6	6	7	7
2	2	2	3	3	3	3	4	4	4
1	1	1	1	1	1	2	2	2	2
2	2	2	3	3	3	3	4	4	4
1	1	1	1	1	1	1	1	1	1
4	4	5	5	6	6	6	7	7	8
4	5	5	6	6	7	7	8	8	9
7	7	8	9	10	10	11	12	12	13
4	4	5	5	5	6	6	7	7	8

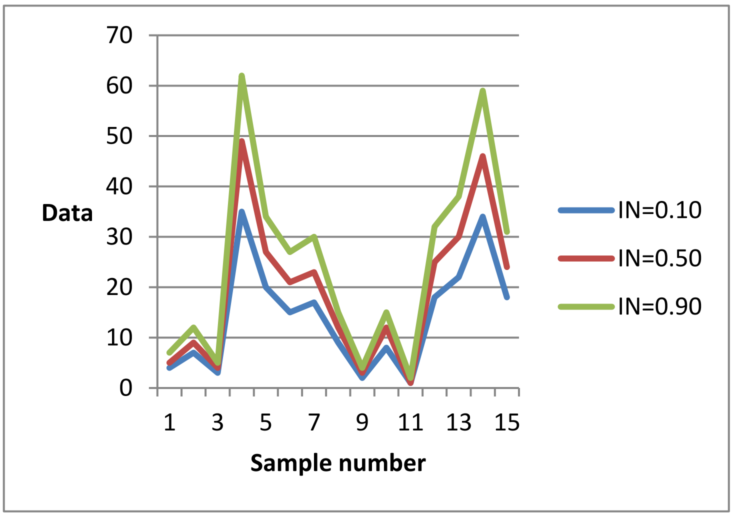

Figure 3. Data curves for various I_N and p=0.10.

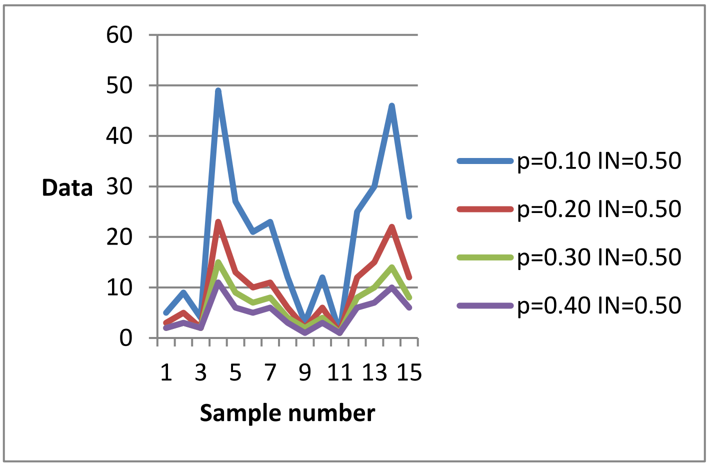

Figure 4. Data curves for $I_N = 0.50$ and various p.

6.2. Simulation using algorithm-II

Applying algorithm-II, the simulation process is conducted, and data originating from the geometric distribution for various combinations of p and I_N are delineated in Tables 5–8. Specifically, Table 5 corresponds to p=0.10, Table 6 to p=0.20, Table 7 to p=0.30, and Table 8 to p=0.40. An observable ascending pattern is discernible in the geometric distribution data across Tables 5–8 as I_N values escalate from 0.1 to 0.9. For instance, with I_N=0.1 and p=0.10, X_N yields a value of 13, while with I_N=0.9 and p=0.10, X_N attains a value of 23. This progression is visually represented in Figure 5, where the data curve for I_N=0.1 is positioned below that of I_N=0.5 and I_N=0.9. Now, we depict the trend in X_N data, maintaining I_N at 0.50, and varying p values at 0.10, 0.20, 0.30, and 0.40 in Figure 6. Upon scrutiny of Figure 6, it is apparent that X_N values experience a decline as the parameter p increases from 0.10 to 0.40. To exemplify, when p=0.10 and I_N=0.50, X_N reaches a value of 18, while with p=0.40 and I_N=0.50, X_N diminishes to a value of 3. We used Excel for data simulation, which is available from the authors upon reasonable request

Table 5. Random variates using algorithm-II when $p = 0.10$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
12	13	14	16	17	18	19	21	22	23
7	8	9	10	10	11	12	13	13	14
16	18	20	21	23	25	26	28	30	31
0	0	0	0	0	0	0	0	0	0
1	1	2	2	2	2	2	2	3	3
2	2	3	3	3	4	4	4	4	5
2	2	2	2	3	3	3	3	4	4
5	6	7	7	8	8	9	10	10	11
18	20	22	24	26	28	30	32	34	35
6	6	7	7	8	9	9	10	10	11
41	45	49	53	57	62	66	70	74	78
1	2	2	2	2	2	3	3	3	3
1	1	1	1	1	1	2	2	2	2
0	0	0	0	0	0	0	0	0	0
2	2	2	2	2	3	3	3	3	3

Table 6. Random variates using algorithm-II when $p = 0.20$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
5	6	7	7	8	8	9	9	10	11
3	4	4	4	5	5	5	6	6	6
7	8	9	10	11	11	12	13	14	15
0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	1
1	1	1	1	1	1	2	2	2	2
1	1	1	1	1	1	1	1	1	2
2	3	3	3	3	4	4	4	5	5
8	9	10	11	12	13	14	15	16	16
2	3	3	3	4	4	4	4	5	5
19	21	23	25	27	29	31	33	35	37
0	0	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	1	1

Table 7. Random variates using algorithm-II when $p = 0.30$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
3	4	4	4	5	5	5	6	6	6
2	2	2	2	3	3	3	3	4	4
4	5	5	6	6	7	7	8	8	9
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1
1	1	2	2	2	2	2	3	3	3
5	6	6	7	7	8	8	9	10	10
1	1	2	2	2	2	2	3	3	3
12	13	14	15	17	18	19	20	21	23
0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1

Table 8. Random variates using algorithm-II when $p = 0.40$.

$I_N = 0$	$I_N = 0.1$	$I_N = 0.2$	$I_N = 0.3$	$I_N = 0.4$	$I_N = 0.5$	$I_N = 0.6$	$I_N = 0.7$	$I_N = 0.8$	$I_N = 0.9$
2	2	3	3	3	3	4	4	4	4
1	1	1	2	2	2	2	2	2	3
3	3	4	4	4	5	5	5	6	6
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	2	2	2
3	4	4	5	5	5	6	6	7	7
1	1	1	1	1	1	2	2	2	2
8	9	10	11	11	12	13	14	15	16
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

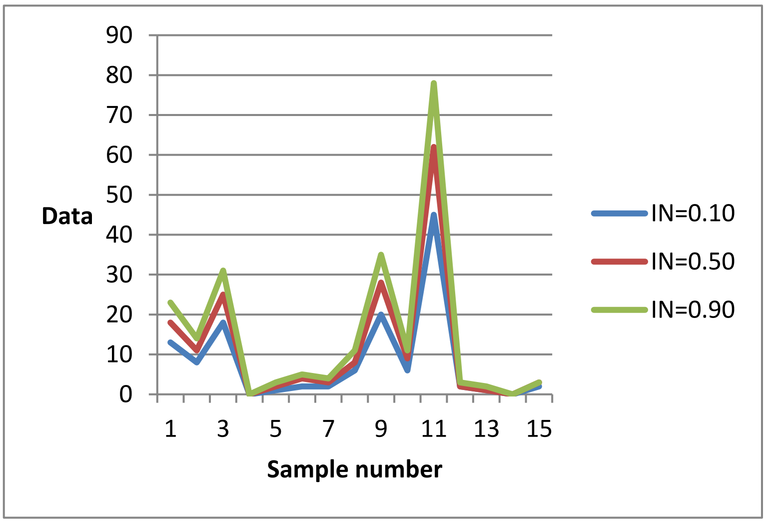

Figure 5. Data curves for various I_N and p=0.10.

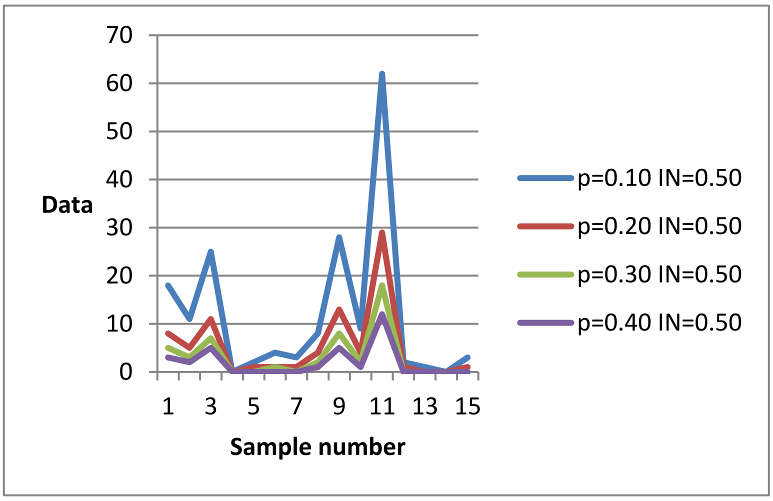

Figure 6. Data curves for $I_N = 0.50$ and various p.

7. Comparative studies

In this section, we delve into a comparative analysis of the data generated by the two proposed algorithms. It is worth noting that these algorithms extend from pre-existing ones in classical statistics, as outlined in Thomopoulos [27]. Notably, the proposed algorithms revert to the existing ones under classical statistics when I_L is set to 0. Our initial focus will be on contrasting the performance of the proposed algorithm-I with the existing algorithm in classical statistics. Subsequently, we will proceed to compare the performance of the proposed algorithm-II with the existing algorithm under classical statistics.

7.1. Comparison using algorithm-I

The X_N values in Tables 1–4, corresponding to I_L=0, represent the geometric distribution values generated by the existing algorithm in classical statistics. The data for the classical geometric distribution is detailed in Tables 1–4. Observing these tables reveals that the X_N values when I_L=0 are smaller compared to the values for other I_L. For instance, in Table 1, when I_L=0, the X_N value is 3, whereas it is 5 when I_N=0.30. The behavioral patterns in the geometric data generated by the proposed algorithm-I and the existing algorithm under classical statistics are illustrated in Figure 7. Analysis of Figure 7 indicates an increasing trend in the data as I_N rises from 0 to 0.3.

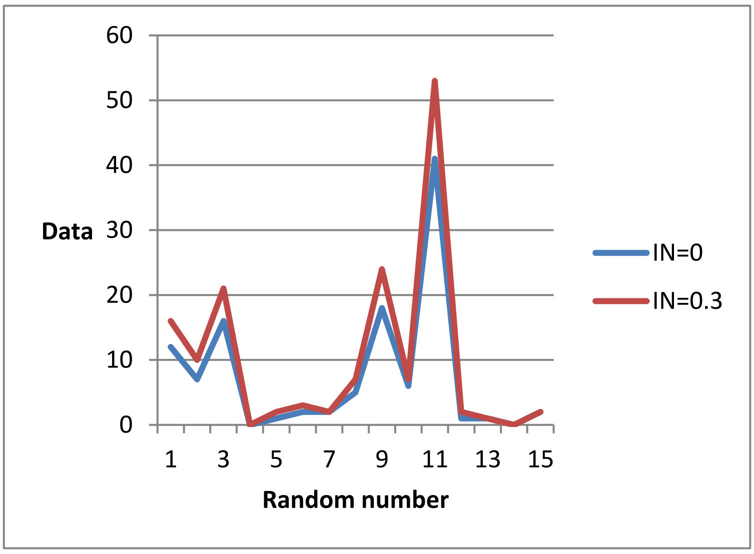

Figure 7. Data curves for p=0.10.

7.2. Comparison using algorithm-II

The X_N values in Tables 5–8, corresponding to I_L=0, represent the geometric distribution values generated by the existing algorithm in classical statistics. The data for the classical geometric distribution is presented in Tables 5–8. Examining these tables reveals that the X_N values when I_L=0 are smaller compared to the values for other I_N. For instance, in Table 5, when I_L=0, the X_N value is 12, whereas it is 16 when I_N=0.30. The behavioral patterns in the geometric data generated by proposed algorithm-II and the existing algorithm under classical statistics are depicted in Figure 8. Analysis of Figure 8 indicates an increasing trend in the data as I_N rises from 0 to 0.3. Additionally, the curve using the data generated from the existing algorithm is positioned below the curve of data when I_N=0.30.

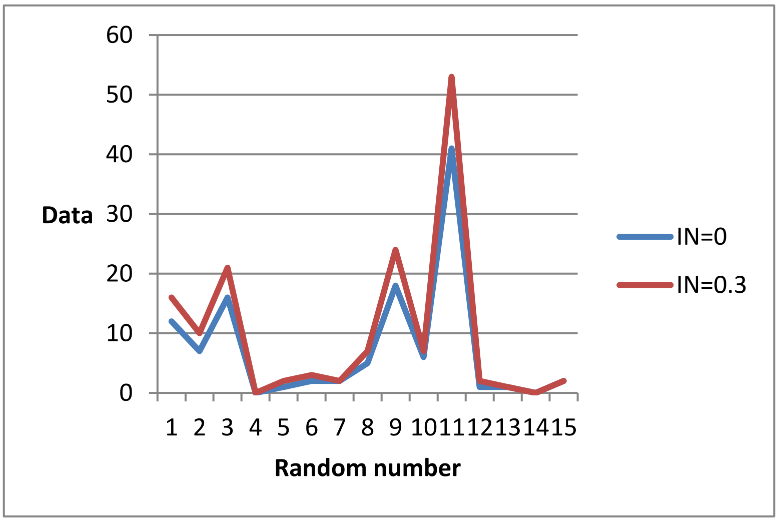

Figure 8. Data curves for p=0.10.

8. Application

Consider a scenario where a quality engineer aims to inspect products until encountering the first defective one. The probability of encountering the first defective item is 0.10. Throughout the inspection process, there exists a degree of indeterminacy set at 0.1. The inquiry revolves around determining the average number of products the engineer needs to inspect before identifying the first defective item. This average number of products required for the initial detection is calculated as follows:

$$E(X_N) = (1 + I_N)/p = 1 + 0.10/0.1 = 11.$$

The industrial engineer, on average, must examine 11 items to identify the first defective product. The average number of products inspected to detect the initial defective item, employing classical statistics, is expressed as follows:

$$E(X_N) = 1/p = 1/0.1 = 10.$$

The industrial engineer, on average, has to examine 10 items before finding the first defective product.

9. Utilization, updating software, and limitations

The simulation study and application section highlights a notable distinction in outputs when applying statistical distribution in both determinate and indeterminate environments. It is crucial to recognize that the level of uncertainty significantly influences the generation of geometric data. Currently, there is a lack of computer software dedicated to producing imprecise geometric data. Consequently, existing algorithms are ineffective in generating imprecise data under uncertain conditions. Given these findings, it is recommended to enhance current algorithms for generating geometric data in uncertain environments. The proposed algorithms offer a viable solution for generating imprecise data applicable in various fields such as industry, artificial intelligence, data analytics, and machine learning. However, it is important to acknowledge certain limitations of the proposed algorithms—they are specifically designed for generating imprecise geometric data. Furthermore, these algorithms are most effective in scenarios where data uncertainty or complexity in data recording is prevalent. Their application is optimized when quantifying the degree of uncertainty in intricate processes.

10. Concluding remarks

The primary contribution of this paper is the introduction of the geometric distribution within the framework of neutrosophic statistics, along with a detailed exploration of its key properties. We also present new algorithms designed for generating imprecise geometric data. The paper discusses the practical applications of this distribution in industry, highlighting differences in data generated under deterministic and indeterminate conditions through tables, simulation studies, and application analyses. Our study indicates the need to update existing algorithms based on classical statistics to accommodate the generation of imprecise data. We conclude that the degree of uncertainty significantly impacts data generation, and using traditional algorithms in uncertain environments can misguide decision-makers. We provide an overview of the fundamental properties of the proposed geometric distribution and suggest that future research should further investigate its statistical characteristics. We also propose the accept-reject method as an extension for future studies. Lastly, we suggest that other statistical distributions could be explored to expand upon the findings of this study in future research.

Author contributions

Muhammad Aslam and Mohammed Albassam wrote the paper. All authors of this article have been contributed equally. All authors have read and approved the final version of the manuscript for publication.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors are deeply thankful to the editor and reviewers for their valuable suggestions to improve the quality and presentation of the paper.

Conflict of interest

No conflict of interest regarding the paper.

References

1. M. Beria, Confidence interval estimation for a geometric distribution, *UNLV Retrospective Theses & Dissertations*, **2015** (2005), 1924. http://doi.org/10.25669/too9-2lgp
2. F. Y. Chen, The goodness-of-fit tests for geometric models, *Dissertations*, 2013 (2013), 350. https://digitalcommons.njit.edu/dissertations/350
3. L. Bertoli-Barsotti, T. Lando, A geometric model for the analysis of citation distributions, *International Journal of Mathematical Models and Methods in Applied Sciences*, **9** (2015), 315–319.
4. A. Slim, G. L. Heileman, M. Hickman and C. T. Abdallah, A geometric distributed probabilistic model to predict graduation rates, *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, San Francisco, CA, USA, 2017, 1–8. http://doi.org10.1109/UIC-ATC.2017.8397646
5. W. K. Gao, An extended geometric process and its application in replacement policy, *P. I. Mech. Eng. O-J. Ris.*, **234** (2020), 88–103. http://doi.org10.1177/1748006X19868891

6. E. Altun, A new generalization of geometric distribution with properties and applications, *Commun. Stat.-Simul. C.*, **49** (2020), 793–807. https://doi.org/10.1080/03610918.2019.1639739
7. M. M. A. Almazah, T. Erbayram, Y. Akdoğan, M. M. A. Sobhi, A. Z. Afify, A new extended geometric distribution: properties, regression model, and actuarial applications, *Mathematics*, **9** (2021), 1336. http://doi.org10.3390/math9121336
8. Z. Y. Zhang, X. T. Tang, Q. Huang, W. J. Lee, Preemptive medium-low voltage arc flash detection with geometric distribution analysis on magnetic field, *IEEE T. Ind. Appl.*, **57** (2021), 2129–2137. http://doi.org10.1109/TIA.2021.3057314
9. I. Ghosh, F. Marques, S. Chakraborty, A bivariate geometric distribution via conditional specification: properties and applications, *Commun. Stat.-Simul. C.*, **52** (2023), 5925–5945. http://doi.org10.1080/03610918.2021.2004419
10. N. Abbas, On classical and Bayesian reliability of systems using bivariate generalized geometric distribution, *J. Stat. Theory Appl.*, **22** (2023), 151–169. https://doi.org/10.1007/s44199-023-00058-4
11. Y. Akdoğan, C. Kuş, A. Asgharzadeh, İ. Kınacı, F. Sharafi, Uniform-geometric distribution, *J. Stat. Comput. Sim.*, **86** (2016), 1754–1770. http://doi.org10.1080/00949655.2015.1081907
12. S. Nadarajah, S. A. A. Bakar, An exponentiated geometric distribution, *Appl. Math. Model.*, **40** (2016), 6775–6784. https://doi.org/10.1016/j.apm.2015.11.010
13. A. S. Hassan, M. A. Abdelghafar, Exponentiated Lomax geometric distribution: properties and applications, *Pak. J. Stat. Oper. Res.*, **13** (2017), 545–566. https://doi.org/10.18187/pjsor.v13i3.1437
14. A. T. Ramadan, A. H. Tolba, B. S. El-Desouky, A unit half-logistic geometric distribution and its application in insurance, *Axioms*, **11** (2022), 676. https://doi.org/10.3390/axioms11120676
15. F. Smarandache, *Introduction to neutrosophic statistics: infinite study*, Columbus: Romania-Educational Publisher, 2014. http://doi.org10.13140/2.1.2780.1289
16. F. Smarandache, Neutrosophic statistics is an extension of interval statistics, while Plithogenic statistics is the most general form of statistics (second version), *International Journal of Neutrosophic Science*, **19** (2022), 148–165. http://doi.org10.54216/IJNS.190111
17. J. Q. Chen, J. Ye, S. G. Du, Scale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics, *Symmetry*, **9** (2017), 208. https://doi.org/10.3390/sym9100208
18. J. Q. Chen, J. Ye, S. G. Du, R. Yong, Expressions of rock joint roughness coefficient using neutrosophic interval statistical numbers, *Symmetry*, **9** (2017), 123. https://doi.org/10.3390/sym9070123
19. W. Q. Duan, Z. Khan, M. Gulistan, A. Khurshid, Neutrosophic exponential distribution: modeling and applications for complex data analysis, *Complexity*, **2021** (2021), 5970613. https://doi.org/10.1155/2021/5970613
20. C. Granados, Some discrete neutrosophic distributions with neutrosophic parameters based on neutrosophic random variables, *Hacet. J. Math. Stat.*, **51** (2022), 1442–1457. http://doi.org10.15672/hujms.1099081
21. C. Granados, A. K. Das, D. A. S. Birojit, Some continuous neutrosophic distributions with neutrosophic parameters based on neutrosophic random variables, *Advances in the Theory of Nonlinear Analysis and its Application*, **6** (2023), 380–389. https://doi.org/10.31197/atnaa.1056480

22. M. Aslam, Simulating imprecise data: sine-cosine and convolution methods with neutrosophic normal distribution, *J. Big Data*, **10** (2023), 143. https://doi.org/10.1186/s40537-023-00822-4
23. M. Aslam, Truncated variable algorithm using DUS-neutrosophic Weibull distribution, *Complex Intell. Syst.*, **9** (2023), 3107–3114. https://doi.org/10.1007/s40747-022-00912-5
24. M. Aslam, Uncertainty-driven generation of neutrosophic random variates from the Weibull distribution, *J. Big Data*, **10** (2023), 177. https://doi.org/10.1186/s40537-023-00860-y
25. M. Aslam, F. S. Alamri, Algorithm for generating neutrosophic data using accept-reject method, *J. Big Data*, **10** (2023), 175. https://doi.org/10.1186/s40537-023-00855-9
26. M. Jdid, R. Alhabib, A. A. Salama, The basics of neutrosophic simulation for converting random numbers associated with a uniform probability distribution into random variables follow an exponential distribution, *Neutrosophic Sets Sy.*, **53** (2023), 358–366. http://doi.org10.5281/zenodo.7536049
27. N. T. Thomopoulos, *Essentials of Monte Carlo simulation: Statistical methods for building simulation models*, New York: Springer, 2013. http://doi.org10.1007/978-1-4614-6022-0

AIMS Press