Modeling bladder cancer survival function based on neutrosophic inverse Gompertz distribution

Oday Esam Al-Saqal [1], Zeina Ameer Hadied [2], Zakariya Yahya Algamal [3,*]

[1] Department of Sharia, University of Mosul, Mosul, Iraq.
[2] University of Mosul, Mosul, Iraq.
[3] Department of Statistics and Informatics, University of Mosul, Mosul, Iraq.

Emails: Oday.esam@uomosul.edu.iq; zena_saeed@uomosul.edu.iq; zakariya.algamal@uomosul.edu.iq

Abstract

In the field of survival analysis, the inverse Gompertz distribution is used to mimic human lifetime data patterns. The goal of the neutrosophic inverse Gompertz distribution (NIGD) is to describe a range of indeterminate survival data. The defined distribution is very helpful for modeling somewhat positively skewed unknown data. The main statistical characteristics of the created NIGD, such as the neutrosophic moments, hazard rate, and survival function, are covered in this paper. Additionally, the well-known maximum likelihood estimation method is used to estimate the neutrosophic parameters. A simulation study is conducted to see whether the projected neutrosophic parameters were reached. Not to mention that possible real-world uses of NIGD have been discussed using actual data. To show how well the suggested model performed in comparison to the present distributions, real data were used.

Keywords: Survival analysis; Neutrosophic statistics; bladder cancer; inverse Gompertz distribution; hazard function.

1. Introduction

Neutrosophic statistics is an application of classical statististics by using neutrosophic logic for uncertain, imprecise, incomplete and inconsistent data to classify the data. Similarly, the elements of neutrosophic statistical distribution can contain parameters, which are mean, variance and various other statistical attributes and they themselves may be neutrosophic in nature. The use of fuzzy logic was expanded by [1] to create neutrosophy, which enables the depiction of uncertainty, ambiguity, and contradiction. This is because traditional statistics tends to assume that the data is deterministic or crisp, wherein each member in the given set is classified as having a definite value. Nevertheless, information in real-world circumstances sometimes lacks general clarity of some data or it is not enough for complete identification. Neutosophic statistics provide an example of how to work with such limitations, such as with ambiguous, scarce, or inconsistent information [2,3,4].

Neutosophic statistics considers three factors: Unlike a sub-set of the mathematical set theory, truth membership, indeterminacy membership and falsity membership are the aspects of the fuzzy set theory. Each component is a sign of the degree of truth or untruth, in the degree of certainty in an observation or a hypothesis. These degrees are presented in the same way as by means of membership functions described in the case of the fuzzy sets [2, 3]. Numerous industries, such as decision-making, pattern identification, data mining, and image processing, use neutrosophic statistics [4,5,6,7].
Neutrosophic statistical distributions have been attempted to fit data which possess different level of uncertainties, which make it a more expandable tool when facing with imprecise information. These distributions are frequently used in situations where randomness is a natural characteristic of the given problem, for instance,

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024

in decision making under ambiguity or when data provide only an approximate indication of the true state of affairs [8,9,10,11,12].
It is necessary that these distributions capture the level of truth, falsity, and uncertainty of the survival probability at the various time points and also the currently available survival data. This can be achieved through mathematical formalizations and tools that are affiliated with neutrosophic logic. A lot of articles on neutrosophic probability distribution exist [11,2, 13,14,15,16,17,18,19,20,21,22,23,24].
Neutrosophic information is particularly important thanks to the analysis of survival statistics, which is its vital function. One of the statistical methods which concentrate on the time to some event occurs as known as survival analysis [11]. The whole survival analysis is based on the probability distributions of the temporal data. It enables the depiction of having a limited or fragmented comprehension of the events.
Applications for the inverse Gompertz distribution can be found in many domains, including survival analysis. In the form of neutrosophy, we extended the applications of the inverse Gompertz distribution in this study when the data is in interval form and has a certain amount of indeterminacy. Many qualities are investigated under the newly proposed distribution and their applications are described with the help of simulated and real data application based on bladder cancer.

2. Neutrosophic inverse Gompertz distribution

The inverse Gompertz distribution, proposed by Eliwa, et al. [25], considers the inverse of a random variable with a Gompertz distribution. More specifically, if a random variable Y has a Gompertz distribution, then the random variable X=1 / Y follows an inverse Gompertz distribution with density and cumulative distribution functions defined, respectively, by:

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

$$F(x) = e^{\frac{-\alpha}{\beta}\left(e^{\frac{\beta}{x-1}}\right)} \quad (2)$$

where $x > 0$ and $\beta, \alpha > 0$.

"The concept of neutrosophic probability as a function $NP :\to [0,\ 1]^3$ was originally presented by [2], where U is a neutrosophic sample space and defined the probability mapping to take the form $NP(S) = (ch(S),\ ch(neut\ S),\ ch(anti\ S)) = (\eta,\ \beta,\ \tau)$ $where\ 0 \le \eta, \beta, \tau \le 1\ and\ 0 \le \eta + \beta + \tau \le 3$. The term Ψ represents the set of sample space, R represents the set of real numbers, and Υ denotes a sample space event, X_N and Y_N denote neutrosophic r.v.

Suppose the neutrosophic variable could be expressed as: $x_N = x_L + x_U I_N$ where $I_N \in \{I_L, I_U\}$ and x_L and $x_U I_N$ denote the determined and indeterminate parts, respectively. Assume that the neutrosophic random variable $x_N \in \{x_L, x_U\}$ follows the inverse Gompertz having neutrosophic scale parameter $\beta_N \in \{\beta_L, \beta_U\}$ and neutrosophic shape parameter $\alpha_N \in \{\alpha_L, \alpha_U\}$ where the letters L and U are the lower values and the upper values, respectively". Then, the neutrosophic probability density function (NPDF) of neutrosophic inverse Gompertz (NIGD) is given by

$$f(x_N) = \frac{\alpha_N}{x_N^2} e^{\frac{-\alpha_N}{\beta_N}\left(e^{\frac{-\alpha_N}{x_N-1}}\right)+\frac{\beta_N}{x_N}}, \quad x_N > 0, \alpha_N > 0, \beta_N > 0 \quad (3)$$

The neutrosophic cumulative density function (NCDF), the neutrosophic survival, and neutrosophic hazard functions are given below, respectively:

$$F(x_N) = e^{\frac{-\alpha_N}{\beta_N}\left(e^{\frac{\beta_N}{x_N}}-1\right)}, \quad (4)$$

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024

$$S(x_N) = 1 - e^{\frac{-\alpha_N}{\beta_N}\left(e^{\frac{\beta_N}{x_N}}-1\right)} \quad (5)$$

$$h(x_N) = \frac{\alpha_N}{x_N^{\,2}} e^{\frac{\beta_N}{x_N}} \left(e^{\frac{\alpha_N}{\beta_N}\left(e^{\frac{\beta_N}{x_N}}-1\right)} - 1 \right)^{-1} \quad (6)$$

3. Statistical Properties of NIGD

In this section, statistical properties of the NIGD are covered.

Moments: The r^{th} moment about origin is given by

$$\mu'_{r_N} = E(X_N^{\,r}) = \sum_{i,m=0}^{\infty} \sum_{j=0}^{i} \frac{(-1)^{i+1} \alpha_N^{\,i+1} (1+i)^m \beta_N^{\,m-i}}{j!\,m!\,(i-j)!} \int_0^{\infty} x_N^{\,r-m-2} e^{\frac{-j\beta_N}{x_N}} \, dx_N \quad (7)$$

It is worth of note that, for the r th raw moment to exist, the constraint $\alpha_N > r$ must be satisfied.

4- Parameter Estimation of NIGD

Maximum likelihood estimation (MLE) method is mostly used in estimating NIGD parameters. Let $x_{N1}, x_{N2}, \ldots, x_{Nn}$ be a random sample of size n from the NIGD. The log-likelihood function is then given by

$$L(\alpha_N, \beta_N) = n \ln \alpha_N + \beta_N \sum_{i=1}^{n} \frac{1}{x_{Ni}} - \frac{\alpha_N}{\beta_N} \sum_{i=1}^{n} \left(e^{\frac{\beta_N}{x_{Ni}}} - 1 \right) - 2 \sum_{i=1}^{n} \ln(x_{Ni}) \quad (8)$$

Thus, the maximum likelihood estimates $\hat{\alpha}_N, \hat{\beta}_N$ for α_N, β_N are the solutions to the nonlinear equations:

$$\hat{\alpha}_N \sum_{i=1}^{n} \left(e^{\frac{\hat{\beta}_N}{x_{Ni}}} - 1 \right) - n \hat{\beta}_N = 0 \quad (9)$$

$$\sum_{i=1}^{n} \frac{1}{x_{Ni}} + \frac{\hat{\alpha}_N}{\left(\hat{\beta}_N\right)^2} \sum_{i=1}^{n} \left(e^{\frac{\hat{\beta}_N}{x_{Ni}}} - 1 \right) - \frac{\hat{\alpha}_N}{\hat{\beta}_N} \sum_{i=1}^{n} \left(\frac{1}{x_{Ni}} e^{\frac{\hat{\beta}_N}{x_{Ni}}} \right) = 0 \quad (10)$$

5. Simulation results

A Monte Carlo simulation is run in R software with several sample sizes, $n = 20, 50, 200, 350$ and neutrosophic parameters in three cases: (1) $\beta_N \in [1, 1.8]$ and $\alpha_N \in [1.3, 2]$, (2) $\beta_N \in [1.5, 2.5]$ and $\alpha_N \in [2, 3]$, and (3) $\beta_N \in [2.5, 4]$ and $\alpha_N \in [3, 4.5]$. The simulation is replicated for 5000 times. For every number of n, performance metrics such as the neutrosophic mean square error (NMSE), the neutrosophic average bias (NAB), and the neutrosophic average of the estimators are obtained. The results are given in Tables 1 -3. These Tables show that, for both neutrosophic parameters, the NAB and NMSE decrease with increasing sample sizes, as would be predicted. Moreover, the results of the study indicate that the neutrosophic MLE for the NIGD provides precise estimation with a larger sample size.

Table 1: Average performance for case 1

n	NAB	NMSE

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024

	β_N	α_N	β_N	α_N
20	[0.0175, 0.0183]	[0.0224, 0.0237]	[0.0324, 0.0376]	[0.0422, 0.0430]
50	[0.0134, 0.0151]	[0.0153, 0.0164]	[0.0337, 0.0324]	[0.0364, 0.0383]
200	[0.0121, 0.0134]	[0.0136, 0.0153]	[0.0318, 0.0333]	[0.0349, 0.0367]
350	[0.0038, 0.0061]	[0.0067, 0.0091]	[0.0241, 0.0271]	[0.0273, 0.0291]

Table 2: Average performance for case 2

n	NAB		NMSE	
	β_N	α_N	β_N	α_N
20	[0.0235, 0.0261]	[0.0283, 0.0292]	[0.0467, 0.0471]	[0.0471, 0.0512]
50	[0.0208, 0.0222]	[0.0237, 0.0255]	[0.0431, 0.0431]	[0.0453, 0.0484]
200	[0.0191, 0.0207]	[0.0218, 0.0231]	[0.0414, 0.0418]	[0.0331, 0.0362]
350	[0.0147, 0.0168]	[0.0161, 0.0169]	[0.0343, 0.0362]	[0.0274, 0.0288]

Table 3: Average performance for case 3

n	NAB		NMSE	
	β_N	α_N	β_N	α_N
20	[0.0244, 0.0257]	[0.0281, 0.0292]	[0.0463, 0.0481]	[0.0474, 0.0512]
50	[0.0201, 0.0210]	[0.0237, 0.0256]	[0.0424, 0.0435]	[0.0453, 0.0495]
200	[0.0182, 0.0198]	[0.0223, 0.0231]	[0.0411, 0.0426]	[0.0337, 0.0367]
350	[0.0136, 0.0148]	[0.0151, 0.0164]	[0.0361, 0.0378]	[0.0281, 0.0294]

6. Applications

This section's practical application used a real-world dataset to measure interest in the NIGD distribution. The data under consideration is a compilation of 128 cancer patients' months-long remission periods. The remission times shown here are based on a subset of bladder cancer data and are mainly intended for descriptive purposes

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024

[26]. Based on the findings of the goodness of fit test based on the Kolmogorov–Smirnov (KS) test, the IGD is one of the plausible distributions for the remission times. According to [16], the information presented here demonstrates that the remission times for a number of cancer patients, including those listed in [7.26, 8.2], [12, 14.77], [15, 17.2], [5.3, 7.1], [75.02, 81], and [1.5, 3.2], are not exact but rather given at irregular intervals.

We evaluate the model suitability of the proposed NIGD with the neutrosophic exponential distribution (NED) applications for complicated data processing studied by [16]. The log-likelihood value (LogL), Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Kolmogorov-Smirnov (KS) test are the techniques used to determine which model best fits the data. The best match is indicated by the model with the highest LogL values and the lowest AIC, BIC, and KS statistic values. Moreover, a larger p-value indicates that the model most closely matches the neutrosophic data. Table 4 lists the model sufficiency indicators and the neutrosophic maximum likelihood estimators. The outcomes show that, in terms of data, the NIGD is better than the NED. The table's bold values indicate the effectiveness of the proposed model.

Table 4: The cancer patients' data neutrosophic distributions

	NED	NIGD
Parameter	$\alpha_N = [0.1081, 0.10822]$	$\beta_N = [1.198, 1.541]$ $\alpha_N = [0.2614,\ 0.6542]$
LogL	[10.352, 13.241]	**[78.5491, 80.2257]**
AIC	[63.508, 65.334]	**[151.224, 153.3614]**
BIC	[60.218, 61.229]	**[149.6621, 152.0581]**
KS-value	[0.752, 0.774]	**[0.163, 0.177]**
KS- p-value	$[1.135\times10^{-6}, 1.188\times10^{-6}]$	**[0.961, 0.978]**

7. Conclusions

This article proposes a neutrosophic inverse Gompertz distribution (NIGD). For survival and dependability uncertainty, a variety of application data can be used with this well-established distribution. The neutrosophic moments, neutrosophic hazard rate, and neutrosophic survival function have all been studied as the main statistical components of the developed NIGD. After construction, the neutrosophic MLEs have shown MSEs and neutrosophic average bias for a range of sample sizes. A simulation research was carried out to ascertain whether the computed neutrosophic parameters were met. Simulation data indicates that two important criteria in accurately estimating an unknown parameter are the sample size and the neutrosophic parametric value. The collecting of remission times from 128 cancer patients is another piece of evidence in favor of using the NIGD in neutrosophic conditions.

Funding: "This research received no external funding"

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

References

[1] Smarandache, F., A unifying field in Logics: Neutrosophic Logic, in *Philosophy*: American Research Press, pp. 1-141, 1999.
[2] Smarandache, F., *Introduction to neutrosophic statistics*. Infinite Study, 2014.
[3] Smarandache, F., 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,* vol. 19, no. 1, pp. 148-165, 2022.
[4] Guan, H., Dai, Z., Guan, S., and Zhao, A., A Neutrosophic Forecasting Model for Time Series Based on First-Order State and Information Entropy of High-Order Fluctuation, *Entropy (Basel),* vol. 21, no. 5, May 1 2019.

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024

[5] Mao, X., Guoxi, Z., M. Fallah, and S. A. Edalatpanah, "A Neutrosophic-Based Approach in Data Envelopment Analysis with Undesirable Outputs," *Mathematical Problems in Engineering,* vol. 2020, pp. 1-8, 2020.
[6] Aslam, M., A New Failure-Censored Reliability Test Using Neutrosophic Statistical Interval Method *International Journal of Fuzzy Systems,* vol. 21, no. 4, pp. 1214-1220, 2019.
[7] Taş, F., Topal,S., and Smarandache,F., Clustering Neutrosophic Data Sets and Neutrosophic Valued Metric Spaces, *Symmetry*, vol. 10, no. 10, 2018.
[8] Alanaz, M. M.,and Algamal, Z. Y., Neutrosophic exponentiated inverse Rayleigh distribution: Properties and Applications, *International Journal of Neutrosophic Science,* vol. 21, no. 4, pp. 36-42, 2023.
[9] Alanaz, M. M. Mustafa, , M. Y., and Z. Y. Algamal, Neutrosophic Lindley distribution with application for Alloying Metal Melting Point, *International Journal of Neutrosophic Science,* vol. 21, no. 4, pp. 65-71, 2023.
[10] Algamal, Z. Y., Alobaidi, N. N., Hamad, A. A., Alanaz, M. M., and M. Y. Mustafa, Neutrosophic Beta-Lindley distribution: Mathematical properties and modeling bladder cancer data, *International Journal of Neutrosophic Science,* vol. 23, no. 2, pp. 186-194, 2024.
[11] Bibani, A. A., and Algamal, Z. Y., Survival Function Estimation for Fuzzy Gompertz Distribution with neutrosophic data, *International Journal of Neutrosophic Science,* vol. 21, no. 3, pp. 137-142, 2023.
[12] Mustafa, M. Y., and Algamal, Z. Y., Neutrosophic inverse power Lindley distribution: A modeling and application for bladder cancer patients, *International Journal of Neutrosophic Science,* vol. 21, no. 2, pp. 216-223, 2023.
[13] Ahsan-ul-Haq, M. J. N. S. S., Neutrosophic Kumaraswamy distribution with engineering application, vol. 49, pp. 269-276, 2022.
[14] Albassam, M., Ahsan-ul-Haq, M., and Aslam, M., Weibull distribution under indeterminacy with applications, *AIMS Mathematics*, vol. 8, no. 5, pp. 10745-10757, 2023.
[15] Alsoboh, A., Amourah, A., Darus, M., and Sharefeen, R. I. A. , Applications of Neutrosophic q-Poisson distribution Series for Subclass of Analytic Functions and Bi-Univalent Functions, *Mathematics,* vol. 11, no. 4, 2023.
[16] Duan, W.-Q., Khan, Z., Gulistan, M., Khurshid, A. and Stevic, Z., Neutrosophic Exponential Distribution: Modeling and Applications for Complex Data Analysis, *Complexity,* vol. 2021, pp. 1-8, 2021.
[17] Granados, C. J. H. J. o. M., and Statistics, Some discrete neutrosophic distributions with neutrosophic parameters based on neutrosophic random variables, vol. 51, no. 5, pp. 1442-1457, 2022.
[18] Hassan, M. K. H. and Aslam, M., DUS-neutrosophic multivariate inverse Weibull distribution: properties and applications, *Complex & Intelligent Systems,* 2023.
[19] Khan, Z., Almazah, M. M. A., Hamood Odhah, O., Alshanbari, H. M., and Mehmood, T., Generalized Pareto Model: Properties and Applications in Neutrosophic Data Modeling, *Mathematical Problems in Engineering,* vol. 2022, pp. 1-11, 2022.
[20] Khan, Z., Amin, A., Khan, S. A., Gulistan, M. J. N. S. and Systems, Statistical development of the neutrosophic Lognormal model with application to environmental data, vol. 47, no. 1, p. 1, 2021.
[21] Rao, G. S., Neutrosophic Log-Logistic Distribution Model in Complex Alloy Metal Melting Point Applications, *International Journal of Computational Intelligence Systems,* vol. 16, no. 1, 2023.
[22] Shah, F., Aslam, M., Khan, Z., Almazah, M. M. A., Alduais, F. S., and Gulzar, M.,"On Neutrosophic Extension of the Maxwell Model: Properties and Applications, *Journal of Function Spaces,* vol. 2022, pp. 1-9, 2022.
[23] Khan, Z., Gulistan, M., Kausar, N., and Park, C., Neutrosophic Rayleigh Model With Some Basic Characteristics and Engineering Applications, *IEEE Access,* vol. 9, pp. 71277-71283, 2021.
[24] Aslam, M. A., Neutrosophic Rayleigh distribution with some basic properties and application, in *Neutrosophic Sets in Decision Analysis and Operations Research*: IGI Global, 2020, pp. 119-128.
[25] Eliwa, M. S., El-Morshedy, M., and M. Ibrahim, Inverse Gompertz Distribution: Properties and Different Estimation Methods with Application to Complete and Censored Data, *Annals of Data Science,* vol. 6, no. 2, pp. 321-339, 2018.
[26] Lee, E. T., and Wang, J., *Statistical methods for survival data analysis*. John Wiley & Sons, 2003.

DOI: https://doi.org/10.54216/IJNS.250106
Received: December 30, 2023 Revised: February 02, 2024 Accepted: June 11, 2024