Natural and Applied Sciences International Journal (NASIJ)
eISSN: 2788-4619 (online)
https://www.ideapublishers.org/index.php/nasij
Vol. 4, No. 2 (July-December), 2023, pp. 41-62
https://doi.org/10.47264/idea.nasij/4.2.3 Research Article

Statistical properties and different estimation methods of Inverse Unit Gompertz Distribution with applications on health data sets

Shakila Bashir*[1] | Ammara Tayyab[1] | Nadia Mushtaq[1] | Itrat Batool Naqvi[1] | Khristina Maksudovna Vafaeva[2]

1. Department of Statistics, Forman Christian College (A Chartered University), Lahore, Pakistan.
2. Gipronickel Institute, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia.

* Corresponding Author Email: shakilabashir@fccollege.edu.pk

Abstract:
Continuous probability distributions are always helpful in lifetime data and health-related data sets. Various techniques exist to develop new probability distributions, adding new parameters and applying different transformations. Adding new parameters is not always good; rather, it can also have complex expressions for the function and properties. This research aimed to develop a model without adding new parameters, which will work more efficiently than the existing models. This study proposes a new probability density function by taking the inversion of a random variable whose probability density function is Unit Gompertz Distribution. The newly proposed distribution is called an Inverse Unit Gompertz Distribution (IUGD). Various properties include reliability/survivorship measures, odd function, elasticity, and Mills ratio. Different statistical properties such as moments, quantile function, and Lorenz and Bonferroni curves for IUGD are developed. Five estimation methods are discussed for unknown parameters of the IUGD, and simulations have been conducted. Finally, IUGD is applied to two real-life data sets, i.e., COVID-19 death rates in the Netherlands and the pain relief time of individuals who received analgesics experienced. IUGD is flexible compared to other competing densities. Moreover, the proposed density can be used for health-related data sets to take accurate precautions and treatments.

Article History

Received:
20-Aug-2023

Revised:
14-Sep-2023

Re-revised:
23-Oct-2023

Accepted:
28-Oct-2023

Published:
9-Nov-2023

Keywords: COVID-19, IUGD, Gompertz Distribution, Monte Carlo simulation, unit distribution, inversion, reliability measures, probability density function.

How to Cite: Bashir, S., Tayyab, A., Mushtaq, N., Naqvi, I. B. & Vafaeva, K. M. (2023). Statistical properties and different estimation methods of Inverse Unit Gompertz Distribution with applications on health data sets. *Natural and Applied Sciences International Journal (NASIJ)*, *4*(2), 41-62. https://doi.org/10.47264/idea.nasij/4.2.3

1. Introduction

Gompertz (1825) first presented the Gompertz Distribution (GD), especially in relation to actuarial tables and human mortality. Since then, demographers and actuaries have shown significant interest in this distribution. The GD is an extension of the exponential distribution and has many uses in practical situations, especially in actuarial and medical research. Intriguing similarities between it and well-known distributions are the exponential, double exponential, Weibull, extreme value (Gumbel Distribution), and generalized logistic distribution. The GD's exponentially rising failure rate for system lifetimes is remarkable. Several writers have significantly contributed to this distribution's statistical methods and characterization recently.

Garg *et al*. (1970) studied the Gompertz distribution's statistical characteristics and estimated the distribution's parameters using the maximum likelihood approach. Data on the effect of prolonged oral contraceptives on mouse mortality served as the basis for their inquiry. On the other hand, Jaheen (2003) looked into the GD from a different angle using a Bayesian analysis approach and centred work on progressive type-II censoring and record values. Many researchers have tried a lot of generalizations of the GD. There have been several significant advancements in the area. Bemmaor (1994) proposed the shifted GD. Roy and Adnan (2012) proposed the wrapped generalized GD. Abu-Zinadah (2014) proposed estimation methods to estimate the shape parameter of the exponential GD. Dey *et al.* (2018) developed some properties and estimations for the GD, Lee and Seo (2020) investigated different estimation approaches for the GD under censoring, Eliwa et al. (2020) proposed a the discrete Gompertz-G family of distributions, and El-Morshedy *et al.* (2020) proposed a generalization of the inverse GD.

Mazucheli *et al.* (2019) introduced a novel distribution known as the Unit Gompertz Distribution (UGD). This distribution was derived from the GD through a transformation $Y = e^{-T}$ involving a variable T which follows the GD. Equation (1) is a representation of the UG distribution's Cumulative Distribution Function (CDF):

$$F\left(y \mid \alpha, \beta\right) = \exp\left[-\alpha\left(y^{-\beta} - 1\right)\right] \tag{1}$$

The corresponding Probability Density Function (PDF) is given in Equation (2).

$$f\left(y \mid \alpha, \beta\right) = \alpha\beta y^{-(\beta+1)} \exp\left[-\alpha\left(y^{-\beta} - 1\right)\right], \quad 0 < y < 1 \tag{2}$$

Where $y > 0, \alpha > 0$ and $\beta > 0$ are scale and shape parameters, respectively.

Various work has been by researchers including Jha *et al.* (2019) reliability estimation on UGD, Jha *et al.* (2020) reliability estimation under censoring for UGD, Lee *et al.* (2020) discussed different estimation methods for the GD under progressive type-II censoring, Khaleel *et al.* (2020) investing distribution, stress exponential GD, Kumar *et al.* (2020) described the inference for the UGD based on record values theory, Arshad *et al.* (2021) discussed Bayesian inference on UGD, Bantan *et al.* (2021) discussed theory and applications on UGD, a stress strength reliability estimation for the UGD discussed by Alsadat *et al.* (2023), and Adegoke *et al.* (2023) proposed Topp-Leone Inverse GD with different estimation methods and applications.

The UG distribution outperforms well-renowned life distributions like other unit interval distributions regarding goodness of fit. Due to various hazard rate shapes, t can capture multiple hazard rate shapes. The UG distribution is particularly suitable for modelling skewed data that other common distributions cannot adequately describe. It finds applications in diverse fields such as environmental studies, industrial reliability, and survival analysis. The UG distribution has also been used to solve modern reliability estimation issues. For instance, Jha and Dey (2020) studied the application of the UG distribution to evaluate the reliability of multicomponent stress strength under progressive type II censoring. The use of the record values and the inter-record times in conjunction with the UG distribution for the inference was also covered.

More generally, numerous methods exist for transforming an existing distribution into a new one. The inverse transformation method, which uses the common inverse function (sometimes called the ratio function), is one of the most well-known techniques. In a more particular context, we use the inverse of random variable X as 1/Y, where Y is a random variable that follows an existing distribution. The Inverse Unit Gompertz distribution (IUGD) is a novel inverse distribution that is introduced in this article using the information mentioned earlier as its basis. The IUGD is conceptualized as a basic two-parameter distribution with a support range $[1,\infty)$. The proposed distribution is continuous in nature and valuable in fields such as lifetime data sets, reliability engineering (to improve the quality of products), health and medicine (survival times, deaths, proper medication etc.), economy and revenue (to check the wealth condition of a nation, policy making) etc.

The structure of this paper is outlined as: the Section 2 presents methodology including the data, tools, methods, and material where the formulation of the IUGD with graphical presentation is presented; the Section 3 presents the reliability measures of the proposed density (which is useful for life data analysis); the Section 4, is about mathematical framework including some important curves (which useful to check economic conditions) and order statistics of the IUGD is proposed; the Sections 5, discusses different estimation methods to estimate the parameters for IUGD; the Section 6 comprehensively explains simulation study; the Section 7 presents the application of the IUGD on life-time data sets; and Section 8 concludes the study.

2. Methodology

2.1. Formulation of the proposed distribution

The formulation of the IUGD involves the Cumulative Distribution (CDF) by taking $X = 1/Y$, where Y conforms to the UGD with support $[0,1)$. Consequently, the support of the IUGD extends from 1 to positive infinity, encompassing the range $[1,\infty)$. The expression for calculating the IUGD's Cumulative Distribution Function (CDF) is given by the Equation (3).

$$F(x) = 1 - \exp\left[-\alpha\left(x^{\beta} - 1\right)\right] \tag{3}$$

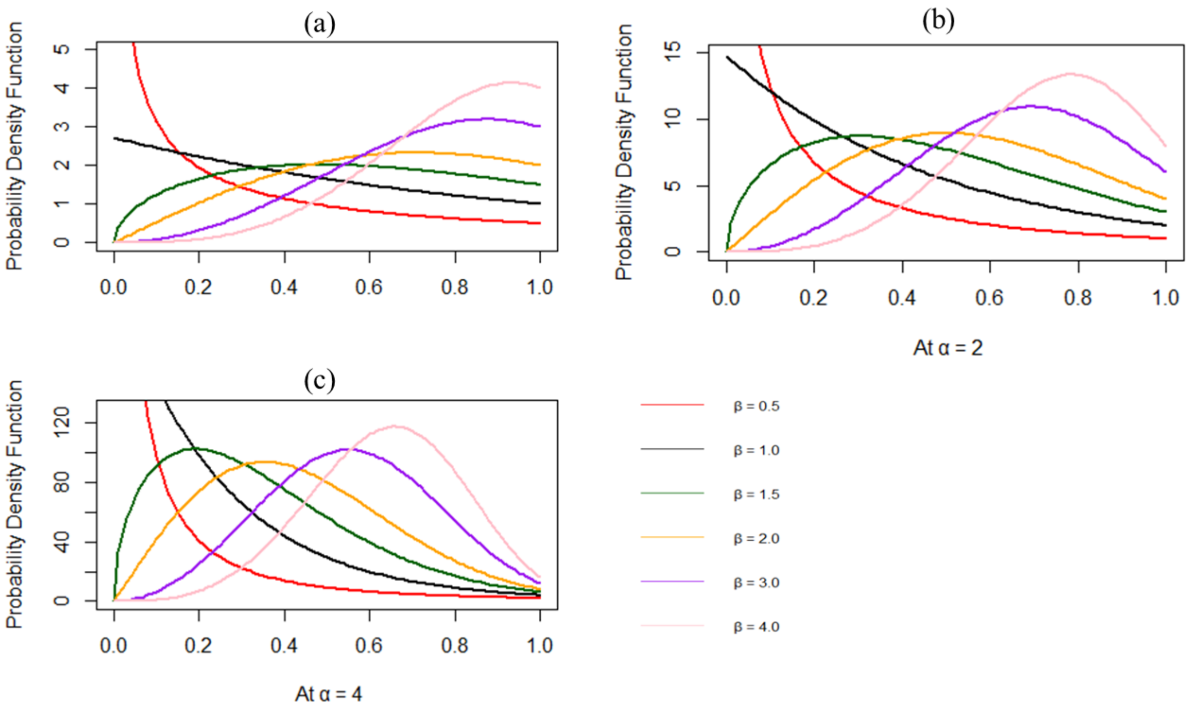
The PDF of IUGD is established as given in Equation (4).

$$f(x) = \alpha\beta x^{\beta-1} \exp\left[-\alpha\left(x^{\beta} - 1\right)\right], \quad x \geq 1 \tag{4}$$

For $x < 1$, we naturally choose to set $f(x) = 0$.

The PDF in Equation (4) contains significant information about the IUGD's modelling capabilities. Its capacity to define the numerous shapes and forms the PDF can assume is particularly significant.

Figure 1: Density plots of IUGD for different values of α and β



The density plots in Figure 1 clearly indicate that the IUGD has a single peak i.e., unimodal distribution and takes on several shapes depending on the value of α, β . The parameter α defines whether the curve shrinks or expands, whereas the parameter β dictates the shape of the curve. The IUGD can have several shapes, such as an L-shape, symmetry, or positively and negatively skewed forms, as seen in Figure 1.

From Figure 1(a), it is observed that for $\alpha = 1$, and for different values of β, the shape of the density plot is right and left skewed, reverse U shape, and peaked and flat as well. From Figure 1(b), it can be seen that for $\alpha = 2$, and for different values of β, the shape of the density plot is right and left skewed, but more peaked as compared to plot in (a). From Figure 1(c), it is perceived that for $\alpha = 4$, and for different values of β, the density plot is right skewed for smaller values of β and as β increased the shapes are approaching to normality.

3. Reliability measures of IUGD

In this section, a few reliability measures such as survival function, hazard rate function (failure rate), cumulative hazard function, reversed hazard function, odd function, mills ratio, and elasticity for IUGD have been developed.

The Survival Function (SF) for the IUGD is given in Equation (5).

$$S(x) = \exp\left[-\alpha\left(x^{\beta} - 1\right)\right], \quad x \geq 1 \tag{5}$$

Additionally, $S(x) = 1$ for $x < 1$.

The corresponding Hazard Rate Function (HRF) is given in Equation (6).

$$h(x) = \alpha\beta x^{\beta-1}, \quad x \geq 1 \tag{6}$$

The IUGD shows a hazard rate function in the figure above that can either decrease, increase, or remain constant for different values of $\alpha \; and \; \beta$.

The CHF is defined as follows.

$$H(x) = \int_{1}^{x} h(t)\,dt = -\log\left(S(x)\right) \tag{7}$$

So, for IUGD the CHF is defined as follows.

$$H(x) = \alpha\left(x^{\beta} - 1\right) \tag{8}$$

The reversed hazard rate for IUGD is given in Equation (9).

$$r(x)=\frac{\alpha\beta x^{\beta-1}\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{9}$$

The odd function for IUGD is given in Equation (10).

$$O(x)=\frac{F(x)}{S(x)}=\frac{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{10}$$

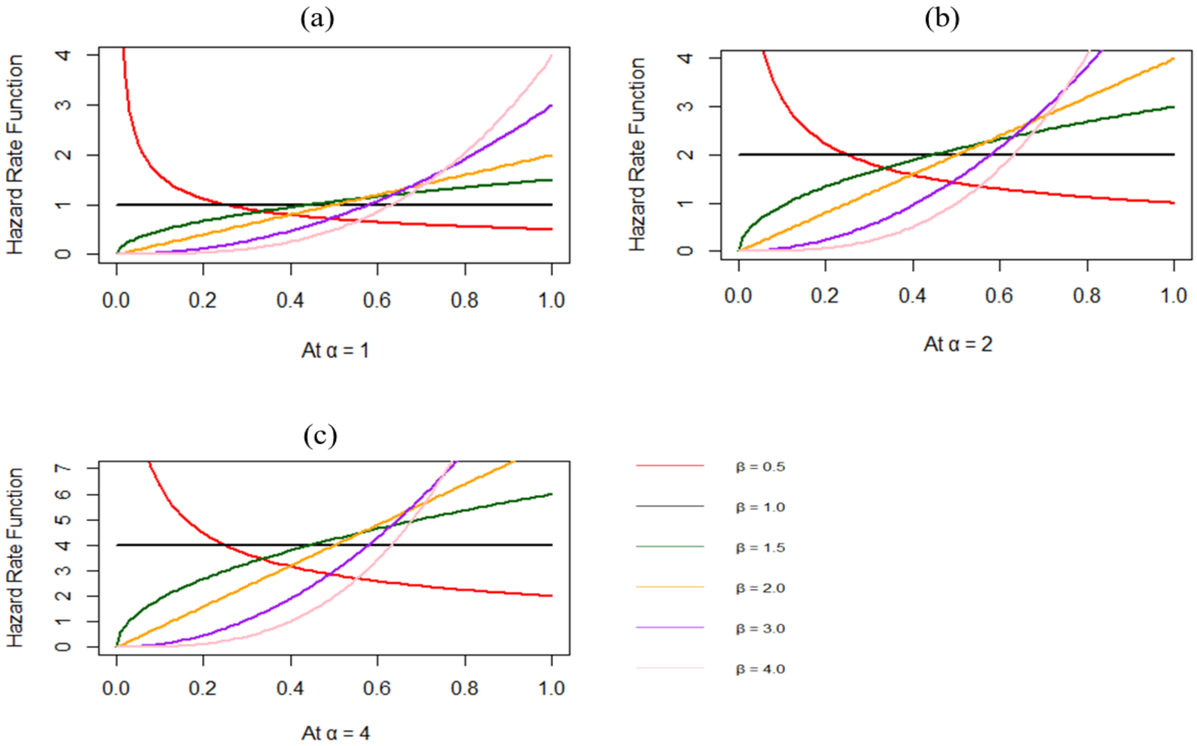
The Mills ratio for IUGD is given in Equation (12).

$$M(x)=\frac{S(x)}{f(x)}=\frac{1}{\alpha\beta x^{\beta-1}} \tag{11}$$

The IUGD's elasticity is given below in Equation (13).

$$\varepsilon(x)=\frac{xf(x)}{F(x)}=\frac{\alpha\beta x^{\beta}\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{12}$$

Figure 2: Hazard rate plots of IUGD for different values of α and β



From Figure 2 (a), (b) and (c), it can be seen that for $\beta = 1$, the HRF for IUGD is constant, for $\beta < 1$, the HRF is decreasing and for $\beta > 1$, it is increasing. For smaller values of β it is determined that initially the failure rate is high and then it is decreasing with time, but for larger values of β initially the failure rate is low and then increases with time. Generally, the IUGD shows a hazard rate function in the Figure above that can either decrease, increase, or remain constant for different values of α and β. The CHF is defined as follows:

$$H(x) = \int_1^x h(t)\,dt = -\log(S(x)) \tag{13}$$

So, for IUGD the CHF is defined as follows.

$$H(x) = \alpha\left(x^{\beta} - 1\right) \tag{14}$$

The reversed hazard rate for IUGD is given in Equation (15).

$$r(x) = \frac{\alpha\beta x^{\beta-1}\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{15}$$

The odd function for IUGD is given by Equation (16).

$$O(x) = \frac{F(x)}{S(x)} = \frac{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{16}$$

The Mills ratio for IUGD is given by Equation (17).

$$M(x) = \frac{S(x)}{f(x)} = \frac{1}{\alpha\beta x^{\beta-1}} \tag{17}$$

The IUGD's elasticity is given by Equation (18).

$$\varepsilon(x) = \frac{xf(x)}{F(x)} = \frac{\alpha\beta x^{\beta}\exp\left[-\alpha\left(x^{\beta}-1\right)\right]}{1-\exp\left[-\alpha\left(x^{\beta}-1\right)\right]} \tag{18}$$

4. Some mathematical properties of IUGD

This section focuses on determining some mathematical properties of IUGD, such as the moments, quantile function, median, mode, Bonferroni and Lorenz Curves, and order statistics.

4.1. Moments

Moments are crucial in statistical analysis and are extremely important, especially in practical work. Moments enable the analysis of key aspects and traits of a distribution, including tendency, dispersion, skewness, and kurtosis. The rth moments for IUGD are defined as.

$$\mu_r' = \int_1^{\infty} x^r \alpha\beta x^{\beta-1} \exp\left[-\alpha\left(x^{\beta}-1\right)\right] dx$$
$$\mu_r' = \frac{e^{\alpha}}{\alpha^{\frac{r}{\beta}}} \Gamma\left(1+\frac{r}{\beta}, \alpha\right) \tag{19}$$

Therefore, the mean and variance for IUGD are given as follows.

$$\mu = \frac{e^{\alpha}}{\alpha^{\frac{1}{\beta}}} \Gamma\left(1+\frac{1}{\beta}, \alpha\right) \tag{20}$$

$$\sigma^2 = \frac{e^{\alpha}}{\alpha^{\frac{2}{\beta}}} \left\{ \Gamma\left(1+\frac{2}{\beta}, \alpha\right) - e^{\alpha}\left(\Gamma\left(1+\frac{1}{\beta}, \alpha\right)\right)^2 \right\} \tag{21}$$

The quantile function of IUGD is obtained by using $x = Q(p) = F^{-1}(p)$

$$x = \exp\left[\frac{1}{\beta}\left\{\log\left(\alpha - \log(1-p)\right) - \log\alpha\right\}\right] \tag{22}$$

The median for IUGD is as follows.

$$m = \exp\left[\frac{1}{\beta}\left\{\log\left(\alpha + \log 2\right) - \log\alpha\right\}\right] \tag{23}$$

The first derivative of the logarithm of the IUGD probability density function is given as:

$$\frac{d}{dx}\log f(x) = \frac{\beta-1}{x} - \alpha\beta^{\beta-1} \tag{24}$$

Therefore, by finding the equation's root, the mode of the IUGD, can be found as follows.

$$\frac{d}{dx}\log f(x) = \frac{\beta-1}{x} - \alpha\beta^{\beta-1} = 0 \tag{25}$$

Hence, if $x = x_0$

$$x_0 = \left(\frac{\beta-1}{\alpha\beta}\right)^{\frac{1}{\beta}} \tag{26}$$

The second derivative is given as:

$$\frac{d^2}{dx^2}\log f(x) = -\frac{(\beta-1)}{x^2} - \alpha\beta(\beta-1)x^{\beta-2} < 0 \tag{27}$$

It follows from this $x_0 = \left(\frac{\beta-1}{\alpha\beta}\right)^{\frac{1}{\beta}}$ is the only critical point where $f(x|\alpha,\beta)$ is maximized.

4.2. Order statistics

Let $X_{(1)} \le X_{(2)} \dots \le X_{(n)}$ represent the associated order statistics for a sample of size $X_1, X_2, \dots, X_n$ taken at random from the IUGD. The CDF and PDF of the r^{th} order statistic, designated as $X_{(r)}$ is found by using Equation (26).

$$F_r(x) = \sum_{i=k}^{n}\binom{n}{i}\left[1-\exp\left\{-\alpha\left(x^{\beta}-1\right)\right\}\right]^{i}\left[\exp\left\{-\alpha\left(x^{\beta}-1\right)\right\}\right]^{n-i} \tag{28}$$

And,

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!}\alpha\beta x^{\beta-1}\left[1-\exp\left\{-\alpha\left(x^{\beta}-1\right)\right\}\right]^{r-1}\left[\exp\left\{-\alpha\left(x^{\beta}-1\right)\right\}\right]^{n-r+1} \tag{29}$$

4.3. Bonferroni and Lorenz Curves

The Bonferroni and Lorenz curves are used in economics, reliability, demography, insurance, and medicine. These curves have been used in a variety of other fields in addition to economics, where they have typically been used to analyze poverty and income distribution. The Bonferroni and Lorenz curves are defined as

$$B(p) = \frac{1}{p\mu}\int_0^q xf(x)\,dx \tag{30}$$

$$L(p) = \frac{1}{\mu}\int_0^q xf(x)\,dx \tag{31}$$

Where

$$\mu = E(X) \text{ and } q = F^{-1}(p) \tag{32}$$

The Bonferroni and Lorenz curves for IUGD are as follows.

$$B(p) = \frac{1}{p}\left[1 - \frac{\Gamma\left(1+\frac{1}{\beta}, \alpha q^{\beta}\right)}{\Gamma\left(1+\frac{1}{\beta}, \alpha\right)}\right] \tag{33}$$

And,

$$L(p) = 1 - \frac{\Gamma\left(1+\frac{1}{\beta}, \alpha q^{\beta}\right)}{\Gamma\left(1+\frac{1}{\beta}, \alpha\right)} \tag{34}$$

5. Estimation and inference

This section introduces a few traditional methods for estimating the IUGD's parameters. In this study, five estimation methods are examined. The definitions of the functions that need to be optimized and the specific estimation context are explained below.

5.1. Maximum Likelihood Estimation (MLE)

Let $x = (x_1, x_2, ..., x_n)$ denote an unknown parameter vector $\theta = (\alpha, \beta)$ representing a random sample of size n taken from the IUGD. The likelihood function for θ is expressed as follows:

$$L(x \mid \theta) = \alpha^n \beta^n \prod_{i=1}^{n} x_i^{\beta-1} \exp\left[-\alpha\left(x_i^{\beta} - 1\right)\right] \tag{35}$$

So, the log-likelihood function can be written as:

$$l(x \mid \theta) = n\log\alpha + n\log\beta + (\beta - 1)\sum_{i=1}^{n}\log x_i - \alpha\sum_{i=1}^{n} x_i^{\beta} + n\alpha \tag{36}$$

The score vector U_{θ} is obtained by taking the partial derivatives of the log-likelihood function with respect to α and β, whose components are as follows.

$$U_{\alpha} = n\left(1 + \frac{1}{\alpha}\right) - \sum_{i=1}^{n} x_i^{\beta} \tag{37}$$

And,

$$U_{\beta}=\frac{n}{\beta}+\sum_{i=1}^{n}\log x_{i}-\alpha\sum_{i=1}^{n}x_{i}^{\beta}\log x_{i} \tag{38}$$

The equations $U_{\alpha}=U_{\beta}=0$ are simultaneously solved to get the maximum likelihood estimates, written as $\hat{\theta}=\left(\hat{\alpha},\hat{\beta}\right)$ for the parameter vector $\theta=\left(\alpha,\beta\right)$. We can see that the following procedures can be used to find the maximum likelihood estimator for α :

$$\hat{\alpha}=\frac{n}{\sum_{i=1}^{n}x_{i}^{\beta}-n} \tag{39}$$

The second derivatives of the log-likelihood function with regard to the unknown parameters are represented by the Hessian matrix, which is given by:

$$H=\begin{bmatrix}-\frac{n}{\alpha^{2}} & -\sum_{i=1}^{n}x_{i}^{\beta}\log x_{i}\\ -\sum_{i=1}^{n}x_{i}^{\beta}\log x_{i} & -\frac{n}{\beta^{2}}-\alpha\sum_{i=1}^{n}x_{i}^{\beta}\left(\log x_{i}\right)^{2}\end{bmatrix} \tag{40}$$

As a result, by using the negative expectation of the Hessian matrix, the expected Fisher information matrix may be calculated.

5.2. Anderson–Darling estimation

The study considered a series of values $x_{(1)},\ldots,x_{(n)}$ from a random variable that follows the Independent and Identically Distributed IUGD pattern. The Anderson-Darling (AD) estimation method is then used to estimate the parameters, which is accomplished by minimising the subsequent function:

$$A=-n-\frac{1}{n}\sum_{i=1}^{n}\left(2i-1\right)\left\{\log\left[1-\exp\left\{-\alpha\left(x_{(i)}^{\beta}-1\right)\right\}\right]-\alpha\left(x_{(i)}^{\beta}-1\right)\right\} \tag{41}$$

5.3. Cramér–Von Mises Estimation (CVM)

The CVM method is used to approximate the values of parameters by minimizing the resulting function.

$$C=\frac{1}{12n}+\sum_{i=1}^{n}\left[F\left(x_{(i)}\right)-\frac{2i-1}{2n}\right]^{2}=\frac{1}{12n}+\sum_{i=1}^{n}\left[1-\exp\left\{-\alpha\left(x_{(i)}^{\beta}-1\right)\right\}-\frac{2i-1}{2n}\right]^{2} \tag{42}$$

5.4. Least-Squares Estimation (LSE)

The LSE is used to create an approximation for the values of the parameters:

$$V=\sum_{i=1}^{n}\left[F\left(x_{(i)}\right)-\frac{i}{n+1}\right]^{2}=\sum_{i=1}^{n}\left[1-\exp\left\{-\alpha\left(x_{(i)}^{\beta}-1\right)\right\}-\frac{i}{n+1}\right]^{2} \quad (43)$$

5.5. Weighted Least-Squares Estimation (WLSE)

It involves minimising the following function, is used to determine the values of parameters:

$$W=\sum_{i=1}^{n}\frac{\left(n+1\right)^{2}\left(n+2\right)}{i\left(n-i+1\right)}\left[1-\exp\left\{-\alpha\left(x_{(i)}^{\beta}-1\right)\right\}-\frac{i}{n+1}\right]^{2} \quad (44)$$

6. Simulation study

In this section, Monte Carlo simulation study is conducted to determine the performance of the proposed estimation methods for IUGD. A sample of sizes 20, 50, 100 and 300 has been taken from 10,000 size from IUGD for different parameter settings. Average biases, biases, Mean Square Errors (MSE) and Mean Relative Estimates (MRE) are calculated.

Table-1: Average biases, biases, MSE and MRE for MLE, AD, CVM, OLS and WLS

	Methods	$\alpha = 1.7$				$\beta = 1.6$			
		20	50	100	300	20	50	100	300
MLE	Average Biase	0.2204	0.2386	0.2468	0.2541	5.6684	5.2334	5.0754	4.9525
	Bias	1.4797	1.4614	1.4532	1.4459	4.0684	3.6334	3.4754	3.3525
	MSE	2.1937	2.1376	2.1130	2.0912	18.0524	13.6876	12.3057	11.3114
	MRE	0.8704	0.8596	0.8548	0.8506	2.5427	2.2709	2.1721	2.0953
AD	Average Biase	0.1959	0.1933	0.1934	0.1940	5.9876	5.9836	5.9581	5.9065
	Bias	1.5042	1.5068	1.5066	1.5060	4.3876	4.3836	4.3581	4.3065
	MSE	2.2651	2.2714	2.2704	2.2684	20.4546	19.6407	19.2086	18.6288
	MRE	0.8848	0.8863	0.8862	0.8859	2.7422	2.7398	2.7238	2.6916
CVM	Average Biase	0.1511	0.1592	0.1610	0.1628	7.6622	7.2422	7.0834	6.9906
	Bias	1.5489	1.5408	1.5390	1.5372	6.0622	5.6422	5.4834	5.3906
	MSE	2.4010	2.3749	2.3691	2.3631	40.4417	33.0048	30.5868	29.2048
	MRE	0.9111	0.9064	0.9053	0.9042	3.7888	3.5264	3.4271	3.3691
OLS	Average Biase	0.1691	0.1677	0.1653	0.1647	7.2186	6.9556	6.9886	6.9454
	Bias	1.5309	1.5323	1.5347	1.5353	5.6186	5.3556	5.3886	5.3454
	MSE	2.3459	2.3487	2.3557	2.3573	34.9889	29.6281	29.4835	28.7248
	MRE	0.9005	0.9014	0.9028	0.9031	3.5116	3.3473	3.3679	3.3409
WLS	Average Biase	0.1347	0.1159	0.1058	0.0928	7.7191	8.1148	8.4194	8.8496
	Bias	1.5653	1.5841	1.5942	1.6072	6.1191	6.5148	6.8194	7.2496
	MSE	2.4513	2.5095	2.5417	2.5832	40.7508	43.8046	47.2925	52.8292
	MRE	0.9208	0.9318	0.9378	0.9454	3.8244	4.0718	4.2621	4.5310

From Table-1, it is observed that as sample size increases the MSE decreases for all estimation methods and overall MSE performs better for small, medium and large samples as compared to other methods.

From Table-2, it is perceived that as sample size increases the MSE decreases for all estimation methods and overall WLS performs better, secondly OLS and thirdly CVM perform better as compared to other methods in the case of small samples. In contrast, for medium and large samples MLE perform better than others.

Table-2: Average biases, biases, MSE and MRE for MLE, AD, CVM, OLS and WLS

Methods		$\alpha = 2.0$				$\beta = 2.5$			
		20	50	100	300	20	50	100	300
MLE	Average Biase	0.2086	0.2314	0.2391	0.2461	9.3950	8.4579	8.1944	7.9761
	Bias	1.7914	1.7686	1.7609	1.7539	6.8950	5.9579	5.6944	5.4761
	MSE	3.2127	3.1300	3.1017	3.0766	52.1271	37.0366	33.1625	30.2352
	MRE	0.8957	0.8843	0.8804	0.8769	2.7580	2.3832	2.2778	2.1905
AD	Average Biase	0.2013	0.2031	0.2005	0.2022	9.7358	9.5612	9.5515	9.4215
	Bias	1.7987	1.7969	1.7995	1.7978	7.2358	7.0612	7.0515	6.9215
	MSE	3.2381	3.2300	3.2390	3.2324	56.0717	51.1504	50.4173	48.2058
	MRE	0.8993	0.8985	0.8998	0.8989	2.8943	2.8245	2.8206	2.7686
CVM	Average Biase	0.1541	0.1638	0.1660	0.1666	12.6302	11.9646	11.6837	11.5129
	Bias	0.1638	11.9646	1.8362	9.4646	3.3724	93.2177	0.9181	3.7858
	MSE	0.1660	11.6837	1.8340	9.1837	3.3638	85.9385	0.9170	3.6735
	MRE	0.1666	11.5129	1.8334	9.0129	3.3615	81.6857	0.9167	3.6052
OLS	Average Biase	0.1748	11.7940	1.8252	9.2940	3.3334	96.7782	0.9126	3.7176
	Bias	0.1693	11.6050	1.8307	9.1050	3.3524	86.2471	0.9154	3.6420
	MSE	0.1682	11.5104	1.8318	9.0104	3.3560	82.6095	0.9159	3.6042
	MRE	0.1681	11.4773	1.8319	8.9773	3.3561	81.1134	0.9160	3.5909
WLS	Average Biase	0.1359	13.1079	1.8641	10.6079	3.4759	125.0550	0.9321	4.2432
	Bias	0.1164	13.6394	1.8836	11.1394	3.5482	128.9470	0.9418	4.4558
	MSE	0.1072	14.1207	1.8928	11.6207	3.5830	137.9445	0.9464	4.6483
	MRE	0.0945	14.9570	1.9055	12.4570	3.6312	156.3502	0.9528	4.9828

From Table-3, it is seen that as sample size increases MSE decreases for all estimation methods and overall WLS performs better, secondly CVM, thirdly OLS, perform better as compared to other methods in case of small samples. In contrast, for medium and large samples MLE perform better than others.

Table-3: Average biases, biases, MSE and MRE for MLE, AD, CVM, OLS and WLS

Methods		$\alpha = 2.5$				$\beta = 3.2$			
		20	50	100	300	20	50	100	300
MLE	Average Biase	0.2140	13.3902	2.2860	10.1902	5.2298	116.7323	0.9144	3.1844
	Bias	0.2353	12.0909	2.2647	8.8909	5.1311	82.7379	0.9059	2.7784
	MSE	0.2429	11.6641	2.2571	8.4641	5.0956	73.2132	0.9028	2.6450
	MRE	0.2509	11.3702	2.2491	8.1702	5.0588	67.3544	0.8996	2.5532
AD	Average Biase	0.2129	13.8791	2.2871	10.6791	5.2339	124.5790	0.9148	3.3372
	Bias	0.2112	13.6428	2.2888	10.4428	5.2401	112.5336	0.9155	3.2634
	MSE	0.2133	13.4095	2.2867	10.2095	5.2298	105.9245	0.9147	3.1905
	MRE	0.2166	13.2530	2.2835	10.0530	5.2146	101.8686	0.9134	3.1416
CVM	Average Biase	0.1629	18.3168	2.3371	15.1168	5.4640	256.6336	0.9348	4.7240
	Bias	0.1674	17.5347	2.3326	14.3347	5.4419	213.9436	0.9330	4.4796
	MSE	0.1685	17.1659	2.3315	13.9659	5.4362	198.8991	0.9326	4.3644
	MRE	0.1707	16.9683	2.3293	13.7683	5.4257	190.6353	0.9317	4.3026
OLS	Average Biase	0.1773	17.4904	2.3227	14.2904	5.3972	232.6027	0.9291	4.4658
	Bias	0.1735	16.9862	2.3265	13.7862	5.4134	197.6054	0.9306	4.3082
	MSE	0.1729	16.8772	2.3271	13.6772	5.4158	190.8747	0.9308	4.2741
	MRE	0.1716	16.9039	2.3284	13.7039	5.4215	189.0377	0.9314	4.2825
WLS	Average Biase	0.1379	19.2586	2.3621	16.0586	5.5806	287.3942	0.9448	5.0183
	Bias	0.1191	20.4735	2.3809	17.2735	5.6690	313.1634	0.9524	5.3980
	MSE	0.1087	21.1288	2.3913	17.9288	5.7186	327.7778	0.9565	5.6028
	MRE	0.0951	22.4446	2.4049	19.2446	5.7835	373.1116	0.9620	6.0139

From Table-4, it is determined that as sample size increases the MSE is decreases for all estimation methods and overall WLS performs better, secondly CVM, thirdly OLS, fourthly AD perform better as compared to MLE methods in case of small samples. In contrast, for medium and large samples MLE performs better than others.

Table-4: Average biases, biases, MSE and MRE for MLE, AD, CVM, OLS and WLS

Methods		$\alpha = 3.7$				$\beta = 2.8$			
		20	50	100	300	20	50	100	300
MLE	Average Biase	0.2258	15.2069	3.4742	12.4069	12.0747	172.5161	0.9390	4.4311
	Bias	0.2521	13.5888	3.4479	10.7888	11.8902	121.8663	0.9319	3.8532
	MSE	0.2592	13.0755	3.4408	10.2755	11.8406	108.6256	0.9299	3.6698
	MRE	0.2715	12.4892	3.4285	9.6892	11.7555	94.8371	0.9266	3.4604
AD	Average Biase	0.2201	16.1364	3.4799	13.3364	12.1130	193.7817	0.9405	4.7630
	Bias	0.2250	15.5248	3.4750	12.7248	12.0773	167.4148	0.9392	4.5446
	MSE	0.2291	15.2325	3.4709	12.4325	12.0481	157.2098	0.9381	4.4402
	MRE	0.2359	14.7810	3.4641	11.9810	12.0010	145.1179	0.9363	4.2789
CVM	Average Biase	0.1659	22.7166	3.5342	19.9166	12.4922	446.3039	0.9552	7.1131
	Bias	0.1741	20.9032	3.5259	18.1032	12.4330	341.4767	0.9530	6.4654
	MSE	0.1753	20.4110	3.5247	17.6110	12.4238	316.9460	0.9526	6.2896
	MRE	0.1763	20.2722	3.5237	17.4722	12.4168	307.2560	0.9524	6.2401
OLS	Average Biase	0.1845	20.3908	3.5155	17.5908	12.3610	348.7379	0.9501	6.2824
	Bias	0.1815	20.0467	3.5186	17.2467	12.3810	310.3754	0.9510	6.1595
	MSE	0.1792	20.1288	3.5208	17.3288	12.3967	306.3855	0.9516	6.1889
	MRE	0.1784	20.0706	3.5216	17.2706	12.4016	300.3531	0.9518	6.1681
WLS	Average Biase	0.1419	23.5610	3.5582	20.7610	12.6616	483.6908	0.9617	7.4146
	Bias	0.1213	24.7377	3.5787	21.9377	12.8073	504.6956	0.9672	7.8349
	MSE	0.1103	25.6862	3.5897	22.8862	12.8859	535.5336	0.9702	8.1736
	MRE	0.0965	27.2586	3.6035	24.4586	12.9852	602.4628	0.9739	8.7352

7. Applications

The use of actual data studies in this section illustrates the IUGD's applicability. The P, Exponential, T, Lindley, Inverse Lindley, and Inverse Rayleigh distributions are used to make comparison easier. This involves supplying information about the related Probability Density Functions (PDFs) and the parameters that go along with them. Using the Maximum Likelihood Estimation (MLE) approach, these parameters with Standard Errors (SE) are estimated. After obtaining the parameter estimations, the model's fit is evaluated using several informational criteria and statistical tests. These measurements and tests aid in determining the goodness of fit and shed light on the model's suitability.

Data I:

The first dataset comprises COVID-19-related death rates observed in the Netherlands between March 31, 2020, and April 30, 2020. These percentages represent the number of fatalities brought on by the virus at that time.

The data are given as follows:

14.918, 10.656, 12.274, 10.289, 10.832, 7.099, 5.928, 13.211, 7.968, 7.584, 5.555, 6.027, 4.097, 3.611, 4.960, 7.498, 6.940, 5.307, 5.048, 2.857, 2.254, 5.431, 4.462, 3.883, 3.461, 3.647, 1.974, 1.273, 1.416, 4.235.

The dataset's descriptive statistics are generated by the calculations listed below.

The table-5 describes the descriptive statistics for the first dataset comprised of the COVID-19-related death rates that were observed in the Netherlands between March 31, 2020, and April 30, 2020. These are the percentages of the number of fatalities brought on by the virus at that time.

Table 5: Summary statistics

Minimum	Median	Mean	Maximum	Skewness	Kurtosis
1.273	5.369	6.157	14.918	0.879	0.175

Table-6: The PDFs for the distributions under discussion

Distribution	PDF	Support
IUGD	$\alpha\beta x^{\beta-1}\exp\left[-\alpha\left(x^{\beta}-1\right)\right]$	$\left[1,+\infty\right)$
P	$\frac{\alpha}{x^{1+\alpha}}$	$\left[1,+\infty\right)$
Exponential	$\frac{1}{\alpha}e^{-x/\alpha}$	$\left[0,+\infty\right)$
T	$\alpha\left(e^{\alpha x}-1\right)e^{\left(\alpha x-e^{\alpha x}+1\right)}$	$\left[0,+\infty\right)$
Lindley	$\frac{(x+1)\alpha^{2}}{1+\alpha}e^{-\alpha x}$	$\left[0,+\infty\right)$
Inverse Lindley	$\frac{\left(x+1\right)\alpha^{2}}{\left(1+\alpha\right)x^{3}}e^{-\alpha/x}$	$\left(0,+\infty\right)$
Inverse Rayleigh	$\frac{2\alpha e^{-\alpha/x^{2}}}{x^{3}}$	$\left(0,+\infty\right)$

Table-7 gives the modelling of the proposed density on COVID-19 data, and the values shows that the proposed density is more efficient and applicable.

Table-7: The goodness-of-fit results for the first dataset

	IUGD	P	Exponential	T	Lindley	Inverse Lindley	Inverse Rayleigh
$\hat{l}$	-76.1031	-94.3841	-84.5253	-79.5285	-79.9640	-85.2373	-83.2789
AIC	156.2061	190.7682	171.0505	161.0571	161.9281	172.4747	168.5577
BIC	159.0085	192.1694	172.4517	162.4583	163.3292	173.8759	169.9589
CAIC	156.6506	190.9111	171.1934	161.1999	162.0709	172.6175	168.7006
HQIC	157.1026	191.2165	171.4988	161.5053	162.3763	172.9229	169.0060
K-S	0.0821	0.3627	0.2634	0.2203	0.1794	0.2428	0.2178
K-S p-value	0.9773	0.0005	0.0252	0.0929	0.2569	0.0483	0.0994
α	0.0382	0.6071	6.1565	0.1462	0.2885	4.9595	11.8488
SE of α	0.0261	0.1108	1.1240	0.0101	0.0377	0.7919	2.1633
β	1.7194	---	---	---	---	---	---
SE of β	0.2955	---	---	---	---	---	---

Data II:

This dataset displays the lengths of pain relief that 20 individuals who received analgesics experienced. The following information is provided: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0. The following calculations result in the dataset's descriptive statistics.

Table-8 provides the descriptive statistics for the dataset and displays the lengths of pain relief that 20 individuals who received analgesics experienced.

Table-8: Summary statistics

Minimum	Median	Mean	Maximum	Skewness	Kurtosis
1.1	1.7	1.9	4.1	1.862	4.185

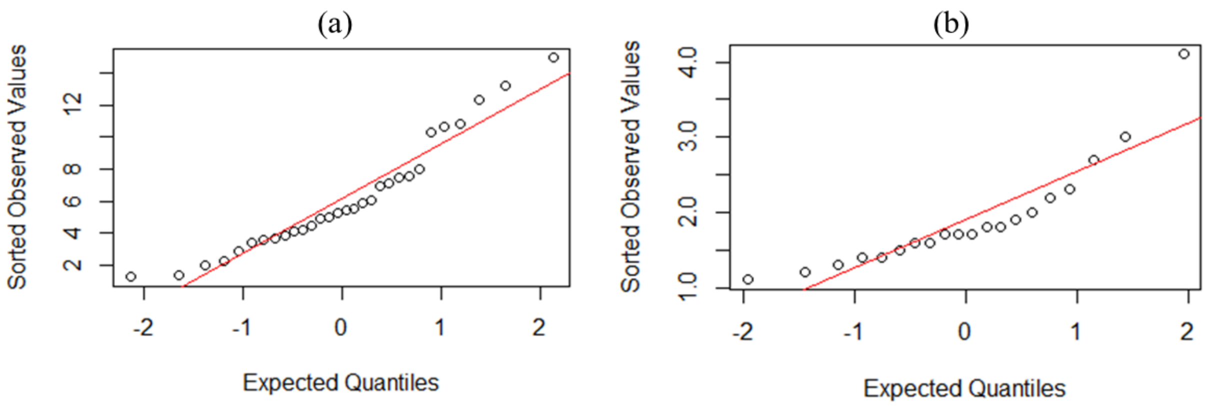
Table-9 depicts the modelling of the proposed density on the relief time data and shows that the IUGD is flexible over health-related data sets. It is clear from the results in Table-7 and Table-9, that the new distribution performs better than the others, making it ideal for modelling datasets related to health.

Table-9: The goodness-of-fit results for the second dataset

	IUGD	P	Exponential	T	Lindley	Inverse Lindley	Inverse Rayleigh
$\hat{l}$	-16.8416	-21.2071	-32.8371	-21.8751	-30.2495	-31.7572	-21.1825
AIC	37.6833	44.4143	67.6742	45.7501	62.4991	65.5144	44.3650
BIC	39.6747	45.4100	68.6699	46.7459	63.4948	66.5101	45.3607
CAIC	38.3891	44.6365	67.8964	45.9724	62.7213	65.7366	44.5872
HQIC	38.0720	44.6087	67.8685	45.9445	62.6935	65.7088	44.5594
K-S	0.1277	0.2850	0.4395	0.1921	0.3911	0.3695	0.2566
K-S p-value	0.9001	0.0775	0.0009	0.4516	0.0044	0.0085	0.1436
α	0.3647	1.6971	1.9000	0.5356	0.8161	2.2547	2.7607
SE of α	0.2575	0.3795	0.4249	0.0453	0.1361	0.4089	0.6173
β	1.8950	---	---	---	---	---	---
SE of β	0.5997	---	---	---	---	---	---

Figure 3 given below, shows the fitting of the IUGD to the probability-probability (P-P) plots, showing its appropriateness for both datasets, further supporting the suitability of the novel distribution.

Figure 3: PP Plots (a) Data I and (b) Data II



8. Conclusion

In this research paper, a two-parameter distribution named Inverse Unit Gompertz Distribution (IUGD) is proposed as an alternative to the P, Exponential, T, Lindley, Inverse Lindley and Inverse Rayleigh distributions. The support for this newly proposed distribution extends from 1 to positive infinity. The construction of this distribution involved utilizing the Unit Gompertz distribution, which served as the standard distribution, and applying the inverse of random variable transformation. The analysis conducted confirmed that the probability density function (PDF) of the new distribution exhibits unimodal behaviour, and the Hazard Rate Function

(HRF) demonstrates either increasing, decreasing or constant characteristics. Mathematical properties, such as moments, quantile function, median, mode, order statistics, Bonferroni, and Lorenz curves are thoroughly investigated. The Lorenz and Bonferroni curves can be used to see the wealth condition of a nation. The reliability measures for the IUGS are derived which are helpful to conduct the life data analysis. A few estimation methods are discussed to estimate the parameters for the IUGD, and a simulation study is conducted for small, medium, and large samples taken from 10,000 sizes. From the simulation study, it was observed that for small samples, the MLE does not perform well. Instead, WLS and CVM are better, while for the large samples MLE performs best as compared to others. The practical performance of the Inverse Unit Gompertz Distribution (IUGD) is demonstrated through the analysis of the real-life dataset. The IUGD outperforms the competing distributions such as P, Exponential, T, Lindley, Inverse Lindley, and Inverse Rayleigh distributions, according to several criteria and goodness-of-fit tests. The outcomes of these analyses show that the IUGD is a highly advantageous and effective option for modelling and analyzing datasets. Finally, the proposed density is helpful for the lifetime data sets (living and non-living objects). Therefore, by using the proposed model on a lifetime of products, the performance of the products can be improved, and necessary repairs can be made; for the lifetime of living objects, proper medication can be done, and treatment ways can be improved.

Declaration of conflict of interest:

The author(s) declared no potential conflicts of interest(s) with respect to the research, authorship, and/or publication of this article.

Funding:

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD:

Shakila Bashir https://orcid.org/0000-0003-4701-6977

Ammara Tayyab https://orcid.org/0009-0009-0330-9437

Nadia Mushtaq https://orcid.org/0000-0002-0652-0029

Itrat Batool Naqvi https://orcid.org/0009-0000-9907-2175

Khristina M. Vafaeva https://orcid.org/0000-0002-7422-5494

Publisher’s Note/Disclaimer:

IDEA PUBLISHERS (IDEA Publishers Group) stands neutral with regard to jurisdictional claims in the published maps and the institutional affiliations.

References

Abu-Zinadah, H. (2014). Six methods of estimations for the shape parameter of Exponentiated Gompertz Distribution. *Applied Mathematical Sciences*. 8, 88, 4349–4359. http://dx.doi.org/10.12988/ams.2014.46503

Adegoke, T. M., Oladoja, O. M., Bashiru, S. O., Mustapha, A. A., Aderupatan, D. E., & Nzei, L. C. (2023). Topp-Leone Inverse Gompertz Distribution: properties and different estimations techniques and applications. *Pakistan Journal of Statistics*, *39*(4), 433-456. https://www.pakjs.com/wp-content/uploads/2023/08/39402.pdf

Alsadat, N., Hassan, A., Elgarhy, M., Chesneau, C., & Mohamed, R. (2023). An efficient stress–strength reliability estimate of the Unit Gompertz Distribution using ranked set sampling. *Symmetry*, 15, 1121. https://doi.org/10.3390/sym15051121

Arshad, M., Azhadc, Q.J., Gupta, N., & Pathake, A.K. (2021). Bayesian inference of Unit Gompertz distribution based on dual generalized order statistics. *Communications in Statistics: Simulation and Computation*, *52*(8), 3657-3675. https://doi.org/10.1080/03610918.2021.1943441

Bantan, R., Jamal, F., Chesneau, C., & Elgarhy, M. (2021). Theory and applications of the Unit Gamma/Gompertz Distribution. *Mathematics*, 9, 1850. https://doi.org/10.3390/math9161850

Bemmaor, A. C. (1994). Modelling the diffusion of new durable goods: word-of-mouth effect versus consumer heterogeneity. In G. Laurent, G. L. Lilien, & B. Pras (eds.), *Research traditions in marketing* (pp. 201-229). https://link.springer.com/chapter/10.1007/978-94-011-1402-8_6

Dey, S., Moala, F., & Kumar, D. (2018). Statistical properties and different methods of estimation of Gompertz distribution with application. *Journal of Statistics and Management Systems*, *21*(5), 839-876. https://doi.org/10.1080/09720510.2018.1450197

Eliwa, M. S., Alhussain, Z. A., & El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over- and under-dispersed data with properties, estimation, and applications. *Mathematics*, *8*(3), 358. https://doi.org/10.3390/math8030358

El-Morshedy, M., El-Faheem, A., & El-Dawoody, M. (2020). Kumaraswamy inverse Gompertz distribution: Properties and engineering applications to complete, type-II right censored and upper record data. *PLoS ONE 15*(12), e0241970. https://doi.org/10.1371/journal.pone.0241970

Garg, M., Rao, B., & Redmond, C. (1970). Maximum-likelihood estimation of the parameters of the Gompertz survival function. *Journal of the Royal Statistical Society, Series C (Applied Statistics), 19*(2), 152-159. https://doi.org/10.2307/2346545

Jaheen, Z. (2003). A Bayesian analysis of record statistic from the Gompertz model. *Applied Mathematics and Computation, 145*(2), 307-320. https://doi.org/10.1016/S0096-3003(02)00489-7

Jha, M. K., Dey, S., Alotaibi, R. M., Alomani, G., & Tripathi, Y. M. (2020). Reliability estimation of a multicomponent stress-strength model for unit Gompertz distribution under progressive Type II censoring. *Qual. Reliab. Eng. Inter*. 36, 965–987. https://doi.org/10.1002/qre.2610

Jha, M.K., Dey, S., & Tripathi, Y. (2019). Reliability estimation in a multicomponent stress-strength based on unit-Gompertz distribution. *Inter. J. Qual. Reliab. Manag.*, *37*, 428–450. https://doi.org/10.1108/IJQRM-04-2019-0136

Khaleel, M., Hashim, N., & Abdal-Hameed, M. (2020). Marshall Olkin exponential Gompertz distribution: Properties and applications. *Periodicals of Engineering and Natural Sciences,* 8, 1, 298-312. https://doi.org/10.21533/pen.v8i1.1152.g513

Kumar, D., Dey, S., Ormoz, E., & MirMostafaee, S. (2020). Inference for the unit-Gompertz model based on record values and inter-record times with an application. *Rendiconti del Circolo Matematico di Palermo Series 2,* 69, 1295–1319. https://doi.org/10.1007/s12215-019-00471-8

Lee, K., & Seo, J. I. (2020). Different Approaches to Estimation of the Gompertz Distribution under the Progressive Type-II Censoring Scheme. *Journal of probability and Statistics, 2020*, 3541946. https://doi.org/10.1155/2020/3541946

Mazucheli, J., Menezes, A. F., & Dey, S. (2019). Unit-Gompertz Distribution with Applications. *Statistica*, *79*(1), 25–43. https://doi.org/10.6092/issn.1973-2201/8497

Roy, S., & Adnan, M. A. S. (2012). Wrapped Generalized Gompertz distribution: an application to Ornithology. *Journal of Biometrics & Biostatistics, 3*(6), 1-4. https://doi.org/10.4172/2155-6180.100015