Asian Journal of Probability and Statistics

Volume 27, Issue 1, Page 122-139, 2025; Article no.AJPAS.129307
ISSN: 2582-0230

On the Inverted Beta Generated Slash Distribution: Properties and Application

Nandita Borah [a*] and Sahana Bhattacharjee [a]

[a] *Department of Statistics, Gauhati University, Guwahati, India.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/ajpas/2025/v27i1708

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/129307

Received: 03/11/2024
Accepted: 07/01/2025
Published: 18/01/2025

Original Research Article

Abstract

In this paper, we propose a new distribution called inverted beta generated slash distribution which is the inverted form of the beta generated slash distribution. The beta generated slash distribution is first introduced and then the inverted form of this distribution is established which is named as inverted beta generated slash distribution. The explicit expressions for pdf, cdf, moments, skewness, kurtosis, median are derived and their variation with different values of the parameters are studied. The hazard rate function assumes different shapes depending on the values of the parameters. A few additional properties such as moment generating function, Mills Ratio, Lorenz and Bonferroni curves, order statistics, hazard rate function of the proposed distribution are also explored. The method of maximum likelihood is used to estimate the unknown parameters of this distribution and a simulation study is conducted to check the performance of these estimates. The MLE's are found to be consistent and precise in estimating the true value of the parameters. Finally, the proposed distribution is applied to a data set to check the flexibility of the model and the goodness-of-fit of the proposed distribution is compared with three other competing distribution to show its flexibility and advantage particularly in modeling heavy-tailed data sets.

**Corresponding author: E-mail: nanditaborahnandini@gmail.com;*

Cite as: *Borah, Nandita, and Sahana Bhattacharjee. 2025. "On the Inverted Beta Generated Slash Distribution: Properties and Application". Asian Journal of Probability and Statistics 27 (1):122-39. https://doi.org/10.9734/ajpas/2025/v27i1708.*

Keywords: *Inverted beta generated slash distribution; slash distribution; additive property; estimation; simulation; real-life application.*

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16

1 Introduction

The beta generated slash distribution proposed by Bhattacharjee and Borah(2024) is an extension of beta distribution through the slash construction idea. Slash distribution is defined by Rogars and Tukey (1972) as the ratio of standard normal variate to the uniform random variate which can be stochastically represented as:

$$Y = \frac{X}{U^{\frac{1}{q}}}$$

where X is a standard normal variate and $U \sim$ *Uniform(0,1)*. The shape parameter $q > 0$ controls the kurtosis of the distribution. Since the last decade, various population distributions have been explored with the help of slash construction idea. For example, the slashed versions for the epsilon half-normal by Gui et al.(2013), the logistic distribution by Punathumparambath(2011) and the Birnbaum−Saunders distribution by Gomez et al.(2009) etc. The modified version of the slash distribution has been proposed by Reyes et al.(2013). The beta generated slash distribution (BGSl) is an extension of beta distribution through the slash construction idea. We shall say that X follows the BGSl distribution with parameters *a, b* and *q* or $X \sim$ *BGSl(a,b,q)* if it can be stochastically expressed as

$$X = \frac{Z}{U^{\frac{1}{q}}}$$

where $Z \sim$ *beta(a,b)* and $U \sim$ *Uniform(0,1)* and are distributed independently of each other. Transforming a known probability distribution into a new one is a cornerstone of statistical modeling andanalysis. Among the various transformation techniques, the inverse transformation method, based on the standard inverse or ratio function, holds particular significance.

An inverted distribution, also known as a reversed distribution, is a distribution of the reciprocal of a random variable. If the probability distribution of any random variable X is given, then the distribution of the reciprocal $Y = \frac{1}{X}$ can be easily obtained. If X is a continuous random variable with probability density function f(x) and cumulative density function F(x), then the cumulative distribution function of the reciprocal G(y) is given by:

$$G(Y) = P(Y \leq y) = P\left(X \geq \frac{1}{Y}\right) = 1 - P\left(X < \frac{1}{Y}\right) = 1 - F\left(\frac{1}{Y}\right)$$

The probability density function of Y is obtained as the derivative of cumulative distribution function as shown below:

$$g(y) = \frac{1}{y^2} f\left(\frac{1}{y}\right)$$

The inverted distributions are widely used in many fields such as medical research, econometrices, biological sciences etc. It is also used in financial literature, environmental studies, survival and reliability theory. While there is no single inventor of the concept of inverse distribution in statistics, it evolved from the cumulative work of mathematicians and statisticians over centuries. In the 18th century, While there is no single "inventor" of the concept of inverse distributions, it evolved from the cumulative work of mathematicians and statisticians over centuries. Its origin can be traced back to the foundational work in probability theory by Thomas Bayes, whose development of Bayesian inference laid the groundwork for using inverse distributions like the inverse gamma as priors. Vilfredo Pareto's study of income distributions introduced the Pareto distribution (Pareto(1964)), later extended the inverted forms to analyze wealth concentration. Similarly, advancements in reliability engineering and survival analysis, building on Waloddi Weibull's work on the Weibull distribution (Weibull(1951)), formalized inverted distributions like the inverted Weibull and inverted exponential for modeling decreasing failure rates. The 20th century saw the integration of these concepts into diverse fields, including

Bayesian statistics and reliability analysis, highlighting the versatility of inverse distributions in modeling rare events, system failures and parameter uncertainty. The statistical literature contains several extensions of inverted distributions, for example, Inverse exponential distribution by Abouammoh and Alshingiti(2009) has garnered attention as a viable alternative to the standard exponential distribution. Additionally, Inverse Lindley distribution was introduced by Sharma et al.(2015) and demonstrated its applicability in analyzing stress-strength reliability, highlighting its utility in reliability and risk analysis contexts. Abd AL-Fattah et al.(2017) proposed Inverted Kumaraswamy distribution and analyzed its characteristics. Tahir et al.(2018) presented the inverted Nadarajah-Haghighi distribution, while Hassan et al.(2022) examined the inverted exponentiated Lomax distribution, contributing to the expanding research on inverse distributions and their applications. Recently the statistical properties and estimation of inverted Top-Leonne distribution was proposed by Saeed et al.(2023). The inverted unit-teisser distribution along with some Recent Advances in Statistical Modeling and Simulations with Application was propsoed by Krishna et al.(2022). These developments underscore the growing interest in inverse distributions for capturing complex real-world phenomena.

A comprehensive review of research on inverted distributions reveals their significant utility in modeling real-world data, particularly when extreme values or reciprocal relationships are involved. In scenarios where heavy-tailed behavior dominates, inverted distributions are essential for accurate modeling, as they effectively capture the influence of extreme or rare events. Models exhibiting skewness offer an improved approach to analyzing heavy-tailed data, and inverted distributions are a valuable class within this framework.

Inverted distributions excel in scenarios involving heavy-tailed lifetime data, which frequently arise in practical applications and require a probabilistic model capable of describing such behavior. Moreover, these models are well-suited to datasets with outliers, as they enhance kurtosis and adapt to the data's variability. While much of the existing research focuses on inverted versions of random variables with support over $(0,\infty)$ and $(-\infty, \infty)$ limited attention has been given to inverted distributions for bounded random variables. This gap in the literature motivates the development of inverted distributions specifically designed for finite bounds. Such models hold particular relevance for lifetime data, where the values are naturally constrained within a specific range. In this study, we aim to construct and analyze the properties of inverted distributions for finitely bounded random variables, thereby expanding their applicability in modeling lifetime and other bounded datasets.

Traditional regression models, widely used across fields such as biology, sociology, economics, psychology, epidemiology, and marketing, often assume error structures that align with the normal probability distribution. However, this assumption may not always hold, especially when the data exhibit asymmetry or skewness. In such cases, inverted distributions offer a more appropriate alternative by accommodating non-standard error structures. These distributions also provide the flexibility required to model extreme events and account for outliers, making them particularly suitable for robust statistical analysis.

The paper is organised as follows. Section 2 introduces the density function of the proposed distribution. Expressions for pdf, cdf, various descriptive statistics are derived and and behaviour of the curve of the proposed distribution for varying values of the parameters graphically are shown in section 3. The maximum likelihood estimation of the parameters of the distribution are dealt with in section 4. In section 5, some stochastic simulations are performed to illustrate the behaviour of the parameters of the proposed distribution. In section 6, the proposed model is applied to data set on failure times to exhibit the potential of the distribution in modeling real-life data sets. The findings of the paper are finally summarized in section 7.

2 Definition and Derivation of the Inverted Beta Generated Slash Distribution

In this section, the pdf and cdf of the proposed distribution have been derived. These results have been presented under Theorem 1 as shown below:

Theorem 2.1. *If a random variable X∼ BGSl(a,b,q) then Y=$\frac{1}{X}$ ∼ IBGSl(a,b,q), whose probability density fucntion is given by:*

$$g(y;a,b,q) = \begin{cases} \frac{qy^{q-1}}{\beta(a,b)}\beta(a+q,b), & 0\leq y<1 \\ \\ \frac{qy^{q-1}}{\beta(a,b)}\beta(\frac{1}{y};a+q,b), & 1\leq y<\infty \end{cases}$$

And the cumulative distribution function is given by:

$$G(y;a,b,q) = \begin{cases} \frac{\beta(a+q,b)y^q}{\beta(a,b)}, & 0\leq y<1 \\ \\ \frac{\beta(\frac{1}{y},a+q,b)y^q-\beta(a+q,b)-\beta(\frac{1}{y},a,b)}{\beta(a,b)}+1, & 1\leq y<\infty \end{cases}$$

Proof. Let $X \sim BGSl(a,b,q)$ which can be stochastically expressed as:

$$X = \frac{Z}{U^{\frac{1}{q}}}$$

where $Z \sim$ *beta(a,b).* The pdf of Z is given by:

$$f(z;a,b) = \frac{z^{a-1}(1-z)^{b-1}}{\beta(a,b)}, \qquad 0\leq z\leq 1$$

Suppose

$$W = U \implies Z = XW^{\frac{1}{q}}$$

∴

$$\begin{aligned} f_X(x,w) &= f_{x,u}(xw^{\frac{1}{q}},w)|J| \\ &= \frac{1}{\beta(a,b)}z^{a-1}(1-z)^{b-1}w^{\frac{1}{q}} \\ &= \frac{1}{\beta(a,b)}z^{a-1}w^{\frac{a}{q}}(1-xw^{\frac{1}{q}})^{b-1} \end{aligned}$$

Hence, the marginal distribution function of X is given by:

$$f(x) = \begin{cases} f_1(x), & 0\leq x<1 \\ f_2(x), & 1\leq x<\infty \end{cases} \tag{2.1}$$

where

$$\begin{aligned} f_1(x) &= \frac{x^{a-1}}{\beta(a,b)}\int_0^1 w^{\frac{a}{q}}(1-xw^{\frac{1}{q}})^{b-1}dw \\ &= \frac{q}{\beta(a,b)x^{q+1}}\beta(x;a+q,b) \end{aligned} \tag{2.2}$$

$\beta(x;a+b,q)$ being the incomplete beta function.

$$\begin{aligned} f_2(x) &= \frac{x^{a-1}}{\beta(a,b)}\int_0^{\frac{1}{x^q}} w^{\frac{a}{q}}(1-xw^{\frac{1}{q}})^{b-1}dw \\ &= \frac{q}{\beta(a,b)x^{q+1}}\beta(a+q,b) \end{aligned} \tag{2.3}$$

and cdf of X is

$$F(X) = \begin{cases} F_1(\text{X}), & 0 \leq y < 1 \\ F_2(\text{X}), & 1 \leq y < \infty \end{cases} \tag{2.4}$$

Where

$$\begin{aligned} F_1(X) &= P(X \leq \ x) \\ &= \int_0^x \frac{q}{\beta(a,b)x^{q+1}} \beta(x; a+q, b) dx \\ &= \frac{q}{\beta(a,b)} \int_0^x \beta(y; a+q, b) x^{-(q+1)} dx \\ &= \frac{\beta(x; a, b)}{\beta(a,b)} - x^{-q} \frac{\beta(x; a+q, b)}{\beta(a,b)} \end{aligned} \tag{2.5}$$

$$\begin{aligned} F_2(X) &= P(X \leq \ x) \\ &= \int_0^1 f_1(x) dx + \int_1^x f_2(x) dx \\ &= 1 - \frac{\beta(a+q,b)}{\beta(a,b)} + \frac{\beta(a+q,b)}{\beta(a,b)} (1 - x^{-q}) \end{aligned} \tag{2.6}$$

Finally, the pdf of $Y = \frac{1}{X}$ is obtained as

$$\begin{aligned} g(y) &= I_{(0,1)}(y) \frac{1}{y^2} f\left(\frac{1}{y}\right) + (1 - I_{(0,1)})(y) \frac{1}{y^2} f\left(\frac{1}{y}\right) \\ &= I_{(0,1)}(y) \frac{1}{y^2} \left[\frac{q}{\beta(a,b)\left(\frac{1}{y}\right)^{q+1}} \beta(a+q,b) \right] \\ &+ (1 - I_{(0,1)})(y) \frac{1}{y^2} \left[\frac{q}{\beta(a,b)\left(\frac{1}{y}\right)^{q+1}} \beta\left(\frac{1}{y}; a+q, b\right) \right] \\ &= I_{(0,1)}(y) \left[\frac{qy^{q-1}}{\beta(a,b)} \beta(a+q,b) \right] \\ &+ (1 - I_{(0,1)})(y) \left[\frac{qy^{q-1}}{\beta(a,b)} \beta\left(\frac{1}{y}; a+q, b\right) \right] \end{aligned} \tag{2.7}$$

$$I_{(0,1)}(y) = \begin{cases} 1, & if \quad 0 \leq y < 1 \\ 0, & if \quad 1 \leq y < \infty \end{cases}$$

Again the cdf of Y is obtained as :

$$G(y) = \begin{cases} G_1(y), & 0 \leq y < 1 \\ G_2(y), & 1 \leq y < \infty \end{cases} \tag{2.8}$$

Where

$$\begin{aligned} G_1(y) &= P(Y \leq \ y) \\ &= \int_0^t \frac{qy^{q-1}}{\beta(a,b)} \beta(a+q,b) \\ &= \frac{\beta(a+q,b)y^q}{\beta(a,b)} \end{aligned} \tag{2.9}$$

$$\begin{aligned} G_2(y) &= P(Y \leq \ y) \\ &= \int_0^1 g_1(y)dy + \int_1^y g_2(y)dy \\ &= \frac{\beta(\frac{1}{y}, a+q, b)y^q - \beta(a+q,b) - \beta(\frac{1}{y}, a, b)}{\beta(a,b)} + 1 \end{aligned} \tag{2.10}$$

□

2.1 Location - Scale form of IBGSl Distribution

Another form of IBGSl distribution is the location - scale form. By applying the well known location - scale transformation, we get the location - scale transformed *IBGSL* variate as (refer to Genc et al., (2014) for details)

$$T = \mu + \sigma Y \tag{2.11}$$

where $Y \sim$*IBGSl(a,b,q)*, $0 < \mu < \infty$ and $\sigma > 0$. μ and σ are the location and scale parameters respectively. The location- scale form of *IBGSl* distribution has the following pdf:

$$f(t; a, b, q) = \begin{cases} \frac{q(t-\mu)^{(q-1)}}{\beta(a,b)\sigma^q}\beta(a+q,b), & \mu < T < \mu + \sigma \\ \\ \frac{q(t-\mu)^{(q-1)}}{\beta(a,b)\sigma^q}\beta\left(\frac{\sigma}{t-\mu}, a+q, b\right), & \mu + \sigma \leq T < \infty \end{cases} \tag{2.12}$$

where a,b,q,μ,σ are the parameter vector. We denote it by $T \sim IBGSlLS(a, b, q, \mu, \sigma)$.

The density plots of the pdf (2.7) for some seleted values of the parameters are given in Fig. 1. These plots show the greater flexibility of the newly proposed distribution for different values of the parameters *a, b* and *q*.

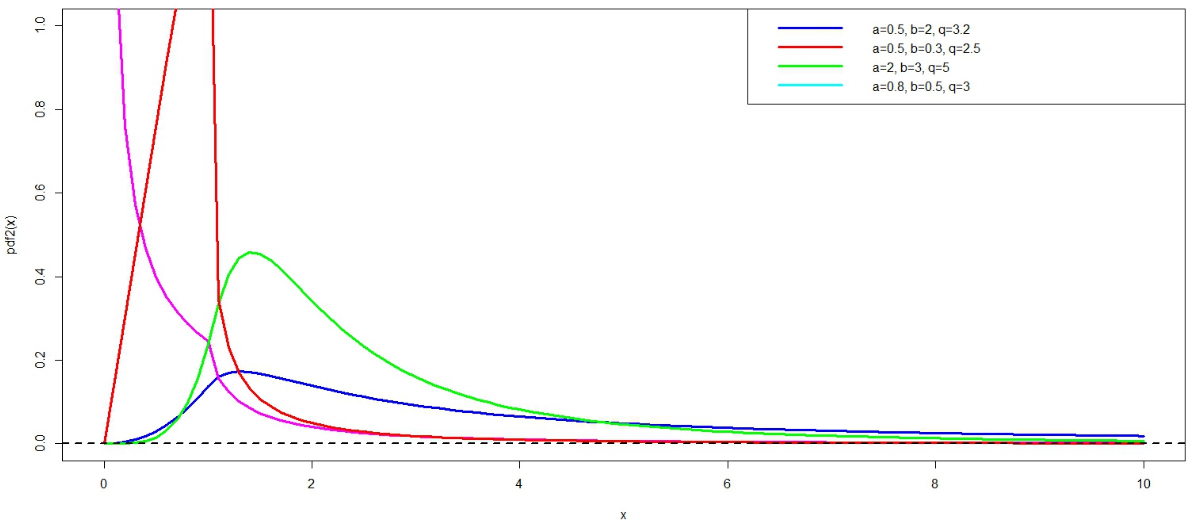

Fig. 1. Probability density function plots of the IBGSl(a,b,q) distribution

The cdf plot of *IBGSl* distribution for different values of the parameters are shown in the following figure.

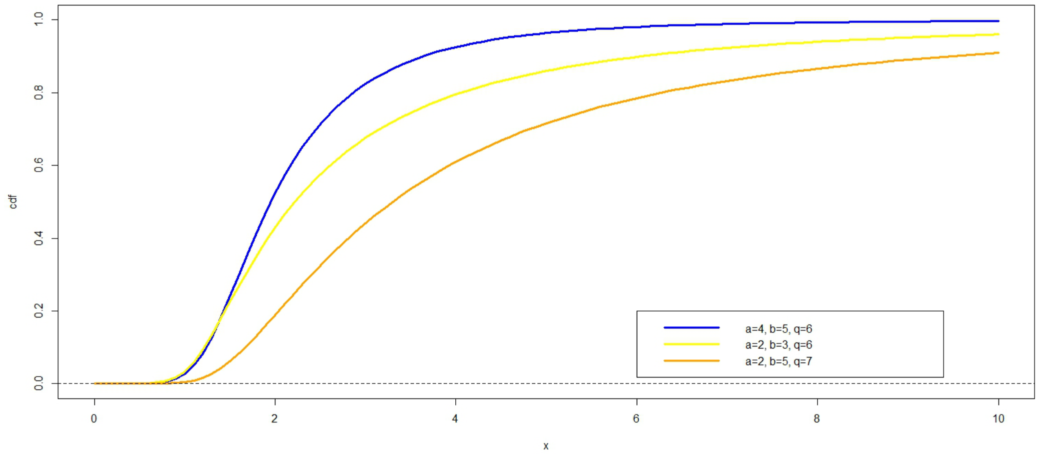

Fig. 2. Cumulative distribution function plot of the IBGSl(a,b,q) distribution

3 Properties of *IBGSl* Distribution

In this section, various descriptive statistics such as moments, skewness, kurtosis, median, mode along with some inequality measures and uncertainity measures of *IBGSl* distribution are derived.

3.1 Moments

If $Y \sim IBGSl(a,b,q)$, then the r^{th} raw moment of Y is given by:

$$\mu_r^{/} = E(Y^r) = \int_0^{\infty} y^r f(y)dy$$

In particular,

$$\begin{aligned}
\mu_1^{/} &= \frac{a+b-1}{(a-1)} \frac{q}{(q+1)}, a>1 \\
\mu_2^{/} &= \frac{(a+b-1)(a+b-2)}{(a-1)(a-2)} \frac{q}{(q+2)}, a>2 \\
\mu_3^{/} &= \frac{(a+b-1)(a+b-2)(a+b-3)}{(a-1)(a-2)(a-3)} \frac{q}{(q+3)}, a>3 \\
\mu_4^{/} &= \frac{(a+b-1)(a+b-2)(a+b-3)(a+b-4)}{(a-1)(a-2)(a-3)(a-4)} \frac{q}{(q+4)}, a>4
\end{aligned}$$

The measures of skewness and kurtosis,denoted by γ_1 and γ_2, respectively are defined as

$$\begin{aligned}
\gamma_1 &= \frac{\mu_3^{/} - 3\mu_2^{/}\mu_1^{/} + 2{\mu_1^{/}}^3}{\left(\mu_2^{/} - {\mu_1^{/}}^2\right)^{\frac{3}{2}}} \\
\gamma_2 &= \frac{\mu_4^{/} - 4\mu_3^{/}\mu_1^{/} + 6\mu_2^{/}{\mu_1^{/}}^2 - 3{\mu_1^{/}}^4}{\left(\mu_2^{/} - {\mu_1^{/}}^2\right)^2}
\end{aligned}$$

The skewness and kurtosis values for some selected values of a and b are displayed in Table 1.

Table 1. Skewness and Kurtosis values for different parameters

(a,b,q)	skewness	kurtosis
(4.5,0.3,1)	0.66	10.69
(4.5,0.6,1)	1.21	17.43
(4.5,0.8,1)	1.35	21.10
(5,0.8,1)	0.92	9.05
(5,0.8,5)	2.90	43.86
(5,0.8,10)	4.15	66.29

From Table 1, it is observed that both the skewness and kurtosis increases with an increase in q.

3.2 Median

The median (M) of any continous probability distribution is the point that splits the area under the probability density curve into two equal parts. In the case of IBBGSl distribution, the area under the curve differ between the intervals $[0,1)$ and $[1,\infty)$. Therefore the median of this distribution will fall within either of these two ranges. The following algorithm are used for calculation of median:

1. Compute F(1)=$\int_0^1 f_1(y)dy$.
2. If F(1) $\geq$0.5 then the median will lie in $[0,1)$ and M is obtained by solving the following equation:

$$\int_0^M f_1(y)dy = 0.5$$
$$\implies \frac{\beta\left(a+q,b\right)M^q}{\beta(a,b)} = 0.5$$

3. If F(1) <0.5 then the median will lie in $[1,\infty)$and M is obtained by solving the following equation:

$$\int_0^1 f_1(y)dy + \int_1^M f_2(y)dy = 0.5$$
$$\implies \frac{\beta\left(\frac{1}{M},a,b\right)-\beta\left(\frac{1}{M},a+q,b\right)M^q}{\beta(a,b)} = 0.5 \tag{3.1}$$

The median values for different set of parameters are given in Table 2:

Table 2. Median values for diffrent set of parameters

Parameters	Median
(0.5,0.3,1.5)	0.654
(1,2,2.5)	0.305
(2,0.3,0.5)	6.764
(0.5,0.3,0.5)	8.264

3.3 Moment Generating Function

The moment generating function of *IBGSL* distribution is given by:

$$\begin{aligned} M_Y(t) &= E\left(e^{ty}\right) \\ &= \int_0^1 e^{ty} f_1(y)dy + \int_1^\infty e^{ty} f_2(y)dy \\ &= 1 + \frac{q}{\beta(a,b)} \sum_{k=1}^{\infty} \frac{t^k}{k!(k+q)} \beta(a-k,b) \end{aligned} \tag{3.2}$$

3.4 Additive Property of IBGSL distribution

Theorem 3.1. *IBGSl distribution does not satisfy the additive property i.e., if* $X \sim IBGSl(a_1,\ b_1,q_1)$ *and* $Y \sim IBGSl(a_2,b_2,q_2)$*, then* $(X+Y)$ *does not follow the IBGSl distribution.*

The m.g.f. of *IBGSl(a,b,q)* distribution is given by:

$$\begin{aligned} M_Y(t) &= E\left(e^{ty}\right) \\ &= \int_0^1 e^{ty} f_1(y)dy + \int_1^\infty e^{ty} f_2(y)dy \\ &= 1 + \frac{q}{\beta(a,b)} \sum_{k=1}^{\infty} \frac{t^k}{k!(k+q)} \beta(a-k,b) \end{aligned}$$

Let $Z = X+Y$ where $X \sim IBGSl(a_1,\ b_1,q_1)$ and $Y \sim IBGSl(a_2,b_2,q_2)$ and are independently distributed of each other. Then the m.g.f. of Z is

$$\begin{aligned} M_Z(t) &= M_{X+Y}(t) \\ &= M_X(t)M_Y(t) \\ &= \left(1 + \frac{q_1}{\beta(a_1,b_1)} \sum_{k=1}^{\infty} \frac{t^k}{k!(k+q_1)} \beta(a_1-k,b_1)\right) \times \\ &\left(1 + \frac{q_2}{\beta(a_2,b_2)} \sum_{k=1}^{\infty} \frac{t^k}{k!(k+q_2)} \beta(a_2-k,b_2)\right) \end{aligned} \tag{3.3}$$

which is not the m.g.f. of *IBGSl* distribution.

Thus, $X+Y$ does not follow *IBGSl* distribution or in other words, the IBGSl distribution does not satisfy the additive property.

3.5 Mean Deviation about mean

The mean deviation about mean of a population measure the amount of scatter in a population to some extent. For a random variable Y with pdf *g(y)*, cdf *G(Y)*, mean μ = E(Y), the mean deviation about mean are defined by:

$$\begin{aligned}
\delta_1(y) &= \int_0^\infty |y-\mu| g(y)dy \\
&= \int_0^\infty (\mu - y)g(y)dy + \int_\mu^\infty (y-\mu)g(y)dy \\
&= \mu G(\mu) - \int_0^\infty yf(y)dy - \mu\left[1 - G(\mu)\right] + \int_\mu^\infty yg(y)dy \\
&= 2\mu G(\mu) - 2\mu + 2\int_\mu^\infty yg(y)dy \\
&= 2\mu G(\mu) - 2\int_0^\mu yg(y)dy \qquad (3.4)
\end{aligned}$$

Hence the mean deviation for *IBGSl (a,b,q)* are given by:

$$\begin{aligned}
\delta_1(y) = I_{(0,1)}(y)&\left[\frac{2q\beta(a+q,b)\mu^q}{\beta(a,b)(1+q)}\left\{\frac{\beta(a+q,b)}{\beta(a,b)} - \mu\right\}\right] \\
+(1 - I_{(0,1)}(y)&\left[\frac{2q}{\beta(a,b)}\left\{\frac{\beta(a-1,b) - \beta(a+q,b)}{(1-q)}\right\}\right. \\
&\left\{\frac{\beta\left(\frac{1}{\mu}, a+q, b\right) - \beta(a+q,b) + \beta(a,b) - \beta\left(\frac{1}{\mu}, a, b\right)}{\beta(a,b)}\right\} \\
-\frac{2q}{\beta(a,b)(q+1)}&\left.\left\{\beta\left(\frac{1}{\mu}, a+q, b\right)\mu^{q+1} - \beta(a+q,b) + \beta(a-1,b)\right\}\right] \qquad (3.5)
\end{aligned}$$

3.6 Mills Ratio

The Mills Ratio is the ratio of complementary cumulative distribution function to the probability density function. Mills ratio can be used in regression analysis to take account of a possible selection bias. Mills Ratio for *IBGSl(a,b,q)* distribution is :

$$\begin{aligned}
m(y) &= I_{(0,1)}(y)\frac{1-G_1(y)}{g_1(y)} + (1 - I_{(0,1)}(y))\frac{1-G_2(y)}{g_2(y)} \\
&= I_{(0,1)}(y)\left\{\frac{1 - \frac{\beta(a+q,b)y^q}{\beta(a,b)}}{\frac{qy^{q-1}}{\beta(a,b)}\beta(a+q,b)}\right\} \\
&+ (1 - I_{(0,1)})(y)\left\{\frac{\frac{\beta(a+q,b) - \beta(\frac{1}{y}, a+q, b)y^q + \beta(\frac{1}{y}, a, b)}{\beta(a,b)}}{\frac{qy^{q-1}}{\beta(a,b)}\beta\left(\frac{1}{y}; a+q, b\right)}\right\} \\
&= I_{(0,1)}(y)\left\{\frac{\beta(a,b) - \beta(a+q,b)y^q}{qy^{q-1}\beta(a+q,b)}\right\} \\
&+ (1 - I_{(0,1)})(y)\left\{\frac{\beta(a+q,b) - \beta(\frac{1}{y}, a+q, b) + \beta(\frac{1}{y}, a, b)}{qy^{(q-1)}\beta(\frac{1}{y}, a+q, b)}\right\} \qquad (3.6)
\end{aligned}$$

3.7 Order Statistics

Consider a random sample y_1, y_2,......,y_n of size n drawn from *IBGSl(a,b,q)* distribution. Further, let $y_{(1)} <$ $y_{(2)} <$ .... $< y_{(n)}$ denote the order statistics corresponding to this sample. Then the probability density function of the k^{th} order statistic is given by

$$f_{(k)}(y) = \frac{n!}{(k-1)!(n-k)!}\left[F(y)\right]^{k-1}\left[1-F(y)\right]^{n-k} f(y)$$

Hence the density of k^{th} order statistic for IBGSl(a,b,q) distribution is

$$\begin{aligned} g_{(k)}(y) = I_{(0,1)}(y)&\left[\frac{n!}{(k-1)!(n-k)!}\left\{\frac{\beta(a+q,b)y^q}{\beta(a,b)}\right\}^{k-1}\left\{1-\frac{\beta(a+q,b)y^q}{\beta(a,b)}\right\}^{n-k}\frac{qy^{q-1}}{\beta(a,b)}\beta(a+q,b)\right] \\ &+(1-I_{(0,1)})(y)\left[\left\{\frac{\beta(\frac{1}{y},a+q,b)y^q-\beta(a+q,b)-\beta(\frac{1}{y},a,b)}{\beta(a,b)}\right\}^{k-1}\right. \\ &\left.\left\{\beta(a+q,b)+\beta(\frac{1}{y},a,b)-\frac{1}{y},a+q,b)y^q\beta(a,b)\right\}^{n-k}\frac{qy^{q-1}}{\beta(a,b)}\beta\left(\frac{1}{y};a+q,b\right)\right] \end{aligned} \tag{3.7}$$

In particular, the p.d.f of the smallest order statistic $y_{(1)}$ is

$$\begin{aligned} g_{(1)}(y) = I_{(0,1)}(y)&\left[n\left\{1-\frac{\beta(a+q,b)y^q}{\beta(a,b)}\right\}^{n-1}\frac{qy^{q-1}}{\beta(a,b)}\beta(a+q,b)\right] \\ &+(1-I_{(0,1)})(y)\left[\left\{\frac{\beta(a+q,b)+\beta(\frac{1}{y},a,b)-\beta(\frac{1}{y},a+q,b)y^q}{\beta(a,b)}\right\}^{n-1}\frac{qy^{q-1}}{\beta(a,b)}\beta\left(\frac{1}{y};a+q,b\right)\right] \end{aligned} \tag{3.8}$$

The pdf of the largest order statistic $y_{(n)}$ is

$$\begin{aligned} g_{(n)}(y) = I_{(0,1)}(y)&\left[n\left\{\frac{\beta(a+q,b)y^q}{\beta(a,b)}\right\}^{n-1}\frac{qy^{q-1}}{\beta(a,b)}\beta(a+q,b)\right] \\ &+(1-I_{(0,1)})(y)\left[\left\{\frac{\beta(\frac{1}{y},a+q,b)y^q-\beta(a+q,b)-\beta(\frac{1}{y},a,b)}{\beta(a,b)}\right\}^{n-1}\frac{qy^{q-1}}{\beta(a,b)}\beta\left(\frac{1}{y};a+q,b\right)\right] \end{aligned} \tag{3.9}$$

3.8 Lorenz and Bonferroni Curve

The Bonferroni and Lorenz Curve are the most used tools in income inequality measurement. These two curves are widely used in the field of reliability, demography, medicine and insurance. The Bonferroni and Lorenz curves are defined as:

$$L(G(y)) = I_{(0,1)}(y)\left[\frac{1}{\mu}\int_0^y tg_1(t)dt\right]+(1-I_{(0,1)}(y))\left[\frac{1}{\mu}\int_1^y tg_2(t)dt\right] \tag{3.10}$$

$$\begin{aligned} B(G(y)) &= I_{(0,1)}(y)\left[\frac{1}{\mu G_1(y)}\int_0^y tg_1(t)dt\right]+(1-I_{(0,1)}(y))\left[\frac{1}{\mu G_2(y)}\int_1^y tg_2(t)dt\right] \\ &= I_{(0,1)}(y)\left[\frac{L(G_1(y))}{G_1(y)}\right]+(1-I_{(0,1)}(y))\left[\frac{L(G_2(y))}{G_2(y)}\right] \end{aligned} \tag{3.11}$$

After simplifications,

$$\begin{aligned} L(G(y)) = I_{(0,1)}(y)&\left[\frac{\beta(a+q,b)(a-1)y^{q+1}}{\beta(a,b)(a+b-1)}\right] \\ -(1-I_{(0,1)}(y))&\left[\frac{(a-1)\left\{y^{q+1}\beta\left(\frac{1}{y},a+q,b\right)-\beta(a+q,b)+\beta(a-1,b)\right\}}{\beta(a,b)(a+b-1)}\right. \end{aligned} \tag{3.12}$$

$$B(G(y)) = I_{(0,1)}(y)\left[\frac{y(a-1)}{a+b-1}\right]$$

$$+(1-I_{(0,1)}(y))\left[\frac{(a-1)\left\{y^{q+1}\beta(\frac{1}{y},a+q,b)-\beta\left(a+q,b\right)+\beta\left(a-1,b\right)\right\}}{(a+b-1)\left\{\beta\left(\frac{1}{y},a+q,b\right)-\beta(a+q,b)-\beta(\frac{1}{y},a,b)+\beta(a,b)\right\}}\right] \tag{3.13}$$

3.9 Hazard Rate Function

The hazard rate function is a very important tool in understanding about the failure mechanism of a lifetime distribution. Hazard rate function can be used to postulate life distributions in the presence of several competing risk factors. It measures the instantaneous rate at which a system or component is likely to fail, given that it has survived up to a certain time. The hazard rate function of *IBGSl(a,b,q)* distribution is obtained by using the following formula:

$$\begin{aligned} h(y) &= I_{(0,1)}(y)\frac{g_1(y)}{1-G_1(y)} + (1-I_{(0,1)}(y))\frac{g_2(y)}{1-G_2(y)} \\ &= I_{(0,1)}(y)\left[\frac{qy^{q-1}\beta\left(a+q,b\right)}{\beta\left(a,b\right)-y^q\beta\left(a+q,b\right)}\right] \\ &+ (1-I_{(0,1)}(y))\left[\frac{qy^{q-1}\beta\left(\frac{1}{y},a+q,b\right)}{\beta\left(a+q,b\right)+\beta\left(\frac{1}{y},a,b\right)-\beta\left(\frac{1}{y},a+q,b\right)}\right] \end{aligned} \tag{3.14}$$

The hrf plot of *IBGSl* distribution for different values of parameters, is plotted in Fig. 3.

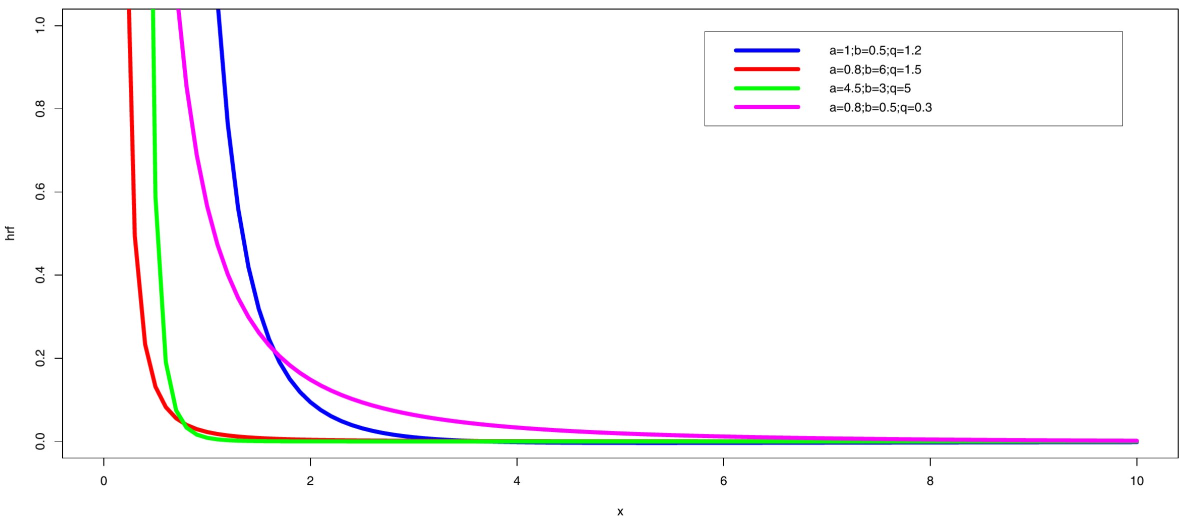

Fig. 3. Hazard rate function plots of the IBSl(a,b,q) distribution

4 Estimation

In this section, we discuss the maximum likelihood method of estimation for the unknown model parameters of *IBGSl(a,b,q)*. Let $y_1, y_2, \ldots\ldots, y_n$ be a random sample of size n from *IBGSl(a,b,q)* distribution. Then the log -

likelihood function is obtained as:

$$\begin{aligned} L(a,b,q,\boldsymbol{y}) &= \prod_{i=1}^{n} f(y_i,a,b,q) \\ &= L_1(a,b,q,y) * L_2(a,b,q,y) \end{aligned}$$

$$\begin{aligned} L_1(a,b,q,\boldsymbol{y}) &= \prod_{i=1}^{n} g_1^{I_{(0,1)}(y_i)} \\ &= g_1^{\sum_{i=1}^{n} I_{(0,1)}(y_i)} \end{aligned}$$

$$\log L_1(a,b,q,\boldsymbol{y}) = \sum_{i=1}^{n} I_{(0,1)}(y_i)\left[\log q + \log\beta(a+q,b) - \log\beta(a,b) + (q+1)\log y_i\right] \tag{4.1}$$

Again,

$$\begin{aligned} L_2(a,b,q,\boldsymbol{y}) \quad &= \textstyle\prod_{i=1}^{n} g_1^{1-I_{(0,1)}(y_i)} \\ &= g_2^{\sum_{i=1}^{n}\left\{1-I_{(0,1)}(y_i)\right\}} \\ &= g_2^{\left\{n-\sum_{i=1}^{n} I_{(0,1)}(y_i)\right\}} \end{aligned}$$

$$\log L_2(a,b,q,\boldsymbol{y}) = \sum_{i=1}^{n}\left(n - I_{(0,1)}(y_i)\right)\left[\log q + \log\beta\left(\frac{1}{y},a+q,b\right) - \log\beta(a,b) + (q+1)\log y_i\right] \tag{4.2}$$

$$\begin{aligned} \log L(a,b,q,\boldsymbol{y}) = &\sum_{i=1}^{n} I_{(0,1)}(y_i)\left[\log q + \log\beta(a+q,b) - \log\beta(a,b) + (q+1)\log y_i\right] \\ &+ \sum_{i=1}^{n}\left(n - I_{(0,1)}(y_i)\right)\left[\log q + \log\beta\left(\frac{1}{y},a+q,b\right) - \log\beta(a,b) + (q+1)\log y_i\right] \end{aligned} \tag{4.3}$$

The maximum likelihood estimates (MLE) of the parameters are computed by solving the maximum likelihood equations, which are given by

$$\begin{aligned} \frac{\partial \log L}{\partial a} \implies &\sum_{i=1}^{n} I_{(0,1)}(y_i)\left[\frac{1}{\beta(a+q,b)}\frac{d}{da}\beta(a+q,b) - \{\psi_0(a)-\psi_0(a+b)\}\right] \\ &+ \left(n - I_{(0,1)}(y_i)\right)\left[\frac{1}{\beta\left(\frac{1}{y_i},a,b\right)}\frac{d}{da}\beta\left(\frac{1}{y_i},a,b\right) - \{\psi_0(a)-\psi_0(a+b)\}\right] = 0 \end{aligned} \tag{4.4}$$

$$\begin{aligned} \frac{\partial \log L}{\partial b} \implies &\sum_{i=1}^{n} I_{(0,1)}(y_i)\left[\frac{1}{\beta(a+q,b)}\frac{d}{db}\beta(a+q,b) - \{\psi_0(b)-\psi_0(a+b)\}\right] \\ &+ \left(n - I_{(0,1)}(y_i)\right)\left[\frac{1}{\beta\left(\frac{1}{y_i},a,b\right)}\frac{d}{db}\beta\left(\frac{1}{y_i},a,b\right) - \{\psi_0(b)-\psi_0(a+b)\}\right] = 0 \end{aligned} \tag{4.5}$$

$$\frac{\partial \log L}{\partial q} \implies \sum_{i=1}^{n} I_{(0,1)}(y_i) \left[\frac{1}{q} + \frac{1}{\beta(a+q,b)} \frac{d}{dq}\beta(a+q,b) + logy_i\right]$$

$$+ \left(n - I_{(0,1)}(y_i)\right) \left[\frac{1}{q} + \frac{1}{\beta\left(\frac{1}{y_i}, a+q, b\right)} \frac{d}{dq}\beta\left(\frac{1}{y_i}, a+q, b\right) + logy_i\right] = 0 \tag{4.6}$$

The above maximum likelihood equations are not in closed form and so, they are difficult to be solved analytically. Hence, we shall use a suitable numerical technique to solve the above equations for a, b and q. Here all the calculations have been carried out using the R software version 3.6.3. The maxLik package is used to obtain the maximum likelihood estimates of the parameters, the rootSolve package is used to generate random variables from *IBGSl* distribution and zipfR package is used to evaluate the incomplete beta function.

5 Simulation

In this section, generation of random numbers from *IBGSl(a,b,q)* distribution is discussed. For different values of *a,b* and *q*, we generate random samples of size 30, 100, 300,800 and 1000 from *IBGSl(a,b,q)*. Finally, the average values of bias and mean squared error (MSE) of these estimates are calculated by using the Monte Carlo approximation technique, taking $N = 1{,}000$ replicates. The algorithm used in this simulation study is shown below:

1. Simulate $X \sim BGSl(a, b, q)$
2. Compute $Y = \frac{1}{X}$

Y thus generated is a random number from the *IBGSl(a,b,q)* distribution. To calculate the average bias and MSE of the likelihood estimates, we use the formulae as shown below :

Let the true value of the parameter a be a^* and estimate be $\hat{a}$. Then the bias and mean square error (MSE) of $\hat{a}$ in estimating a is given by:

$$Bias(\hat{a}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{a}_i - a*) \tag{5.1}$$

$$MSE(\hat{a}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{a}_i - a*)^2 \tag{5.2}$$

where N is the number of replications and $\hat{a}_i$ is the MLE of $\hat{a}$ obtained in the i^{th} replicate. Similarly, the bias and MSE of b and q are calculated. It is well known that an estimate is consistent if the bias and MSE decreases (approaches to zero) with an increase in the sample size. Table 3 shows the results of the simulation studies. It is seen that the parameters are well estimated and the bias and MSE of all the estimators approaches towards zero with an increase in the sample size. Hence, the estimates of the parameters can be believed to be consistent.

6 Application

To show the flexibility of the proposed distribution over some existing distributions in modeling heavy - tailed data we apply these distributions to a real life data set. Data set representing vinyl chloride data from clean up gradient monitoring wells in mg/l used by Bhaumik et al.(2009) has been considered . The data set comprises of the observations:
5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 2.0, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1.0, 0.2, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 6.8, 1.2, 0.4, 0.2.

Table 3. Average bias and RMSE for the estimates of *IBGSl* distribution

Parameters	n	$\hat{a}$		$\hat{b}$		$\hat{q}$	
		Bias($\hat{a}$)	RMSE($\hat{a}$)	Bias($\hat{b}$)	RMSE($\hat{b}$)	Bias($\hat{q}$)	RMSE($\hat{q}$)
(a=0.5, b=0.8, q=2.5)	30	0.00650	0.00118	0.00507	0.00092	-0.00180	0.00033
	100	0.00308	0.00030	0.00135	0.00013	-0.00187	0.00019
	300	0.00307	0.00018	0.00108	0.00005	-0.00191	0.00011
	500	0.00305	0.00014	0.00095	0.00004	-0.00193	0.00008
	800	0.00242	0.00007	0.00099	0.00003	-0.00196	0.00006
	1000	0.00180	0.00004	0.00033	0.00001	-0.00198	0.00004
(a=1.5, b=3, q=5)	30	0.03401	0.00621	0.00057	0.00034	-0.00419	0.00076
	100	0.00448	0.00044	-0.00276	0.00027	-0.00418	0.00041
	300	0.00236	0.00013	-0.00282	0.00016	-0.00416	0.00024
	500	0.00192	0.00007	-0.00286	0.00012	-0.00415	0.00018
	800	0.00174	0.00006	-0.00290	0.00010	-0.00414	0.00014
	1000	0.00110	0.00003	-0.00293	0.00001	-0.00413	0.00013

The histogram of the data set exhibits a right skewed behavior, which may be aptly modelled by the proposed distribution.

Using this data set, we compare the *IBGSl* distribution with Beta Generated Slash distribution (BGSl), Inverted Weibull (IW) distribution ,Inverted Gamma (IG) and Beta Moyal Slash Distribution(BMSl) distribution. In order to compare the distributions we calculate the log - likelihood, Akaike Information Criterion (AIC), the Corrected Akaike Information Criterion (AICC), Kolmogorov-Smirnov (K-S) statistic and p-value. The model with minimum AIC, AICC, KS statistic and p-value is chosen as the best model to fit the data.
The values of log - likelihood, AIC, AICC ,KS and p-value are shown in Table 4.

Table 4. Estimated parameters, AIC, AICC , K-S statistic and p-values of *IBGSl(a,b,q)* and other competing distributions fitted to the data set

Distribution	MLE	log-likelihood	AIC	AICC	K-S statistic	p-value
IBGSl(a,b,q)	â=2.098 $\hat{b}$=3.797 q̂=0.9000848	56.765	119.530	120.189	0.106	0.025
BGSl(a,b,q)	â=1.227 $\hat{b}$=1.150 q̂=1.012	57.795	121.253	121.124	0.774	0.421
IW	â=0.880 0 $\hat{b}$=0.653	58.626	121.591	122.316	0.625	0.305
IG	â=0.900 $\hat{b}$=0.515	2008.242	4020.48	4021.129	0.771	0.264
BMSl	â=0.20058 $\hat{b}$=0.795	242.658	3587.63	3452.321	0.335	0.120

From the Table 4, it is been seen that the *IBGSl* distribution has minimum likelihood and AIC, AICC , K-S statistic and p-value. Hence the *IBGSl* distribution performs well than the other competing distributions. Furthermore, Fig. 4 and Fig. 5 shows the histogram of the data sets along with fitted pdfs and the empirical cdf versus fitted cdfs for the data representing vinyl chloride from clean up gradient monitoring wells in mg/l. These figures confirms that the goodness-of-fit of IBGSL distribution with respect to all the fitted distributions.

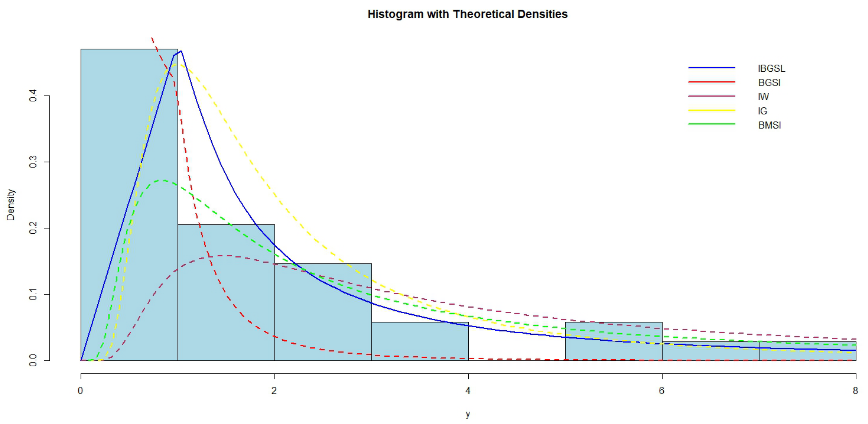

Fig. 4. Plots of the estimated densities with histogram

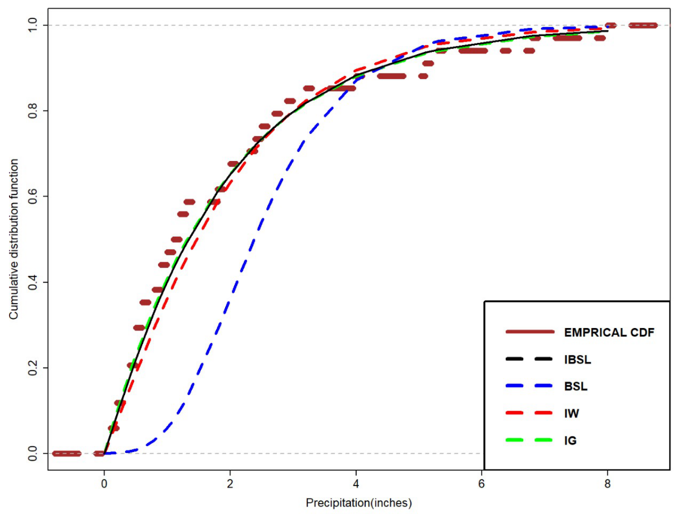

Fig. 5. CDF plot of the observed data and fitted distributions

7 Conclusion

This paper introduces the inverted beta generated slash distribution having three parameters, which is an inverted form of beta generated slash distribution. The various distributional aspects such as moments, skewness, kurtosis, median, moment generating function, mean deviation, mills ratio, order statistics, Lorenz and Bonferroni curves are studied. The method of maximum likelihood is used to estimate the parameters and a simulation study is performed to study the finite sample behaviour of the ML estimates. The MLE's are found to be consistent and precise in estimating the true value of the parameters. To show the application of the proposed distribution, it is applied to a dataset representing vinyl chloride from clean up gradient monitoring wells in mg/l. The fit of the proposed distribution is compared with beta generated slash distribution (BGSl), Inverted Weibull (IW) distribution, Inverted Gamma (IG) distribution and Beta Moyal Slash distribution using log - likelihood measure, AIC, AICC, K-S statistic and p-value. It is observed that the *IBGSl* distribution is a better fit to the data as compared to the others. As a scope for future research,the enhanced flexibility of the proposed distribution by introducing additional parameters or via suitable mathematical transformation can be explored and its utitlity can be assessed through some real-life example.

Disclaimer (Artificial Intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

Competing Interests

Authors have declared that no competing interests exist.

References

Abouammoh, A. M., & Alshingiti, A. M. (2009). Reliability estimation of generalized inverted exponential distribution. Journal of statistical computation and simulation, 79(11), 1301-1315.

Abd AL-Fattah, A. M., El-Helbawy, A. A., & Al-Dayian, G. R. (2017). Inverted Kumaraswamy distribution: Properties and Estimation. Pakistan Journal of Statistics, 33(1).

Abd EL-Kader R. I. (2013). A general class of some inverted distributions (Doctoral dissertation, Ph. D. Thesis, AL-Azhar University, Girls' Branch, Egypt, Cairo).

Abd EL-Kader R. I., Al Dayian G. R., & Al Gendy S. A. (2003). Inverted Pareto Type I distribution: properties and estimation. Journal of Faculty of Commerce. Al-Azhar University, Girl's branch. 21:19-40.

Bhattacharjee, S., & Borah, N. (2024). Beta generated slash distribution: derivation, properties, and application to lifetime data. Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics, 73(3), 641-663.

Bhaumik, D. K., Kapur, K., & Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. Technometrics, 51(3), 326-334.

Conway, D. A. (2014). Farlie-Gumbel-Morgenstern Distributions. Wiley StatsRef: Statistics Reference Online.

Von Rosen, D. (1997). On moments of the inverted Wishart distribution. Statistics: A Journal of Theoretical and Applied Statistics, 30(3), 259-278.

Genc, A. A., Korkmaz, M. Ç., & Kus, C. (2014). The beta Moyal-slash distribution. Journal of Selcuk University Natural and Applied Science, 3(4), 88-104.

Gómez, H. W., Olivares-Pacheco, J. F., & Bolfarine, H. (2009). An extension of the generalized Birnbaum–Saunders distribution. Statistics & Probability Letters, 79(3), 331-338.

Gui, W., Chen, P. H., & Wu, H. (2013). An epsilon half-normal slash distribution and its applications to nonnegative measurements.

Hassan, A. S., Almetwally, E. M., Gamoura, S. C., & Metwally, A. S. (2022). Inverse exponentiated Lomax power series distribution: Model, estimation, and application. Journal of Mathematics, 2022(1), 1998653.

Hassan, N. J., Nasar, A. H., & Hadad, J. M. (2020). Distributions of the ratio and product of two independent Weibull and Lindley random variables. Journal of Probability and Statistics, 2020(1), 5693129.

Krishna, A., Maya, R., Chesneau, C., & Irshad, M. R. (2022). The unit Teissier distribution and its applications. Mathematical and Computational Applications, 27(1), 12.

Kolev, N., & Paiva, D. (2009). Copula-based regression models: A survey. Journal of statistical planning and inference, 139(11), 3847-3856.

Patton, A. (2013). Copula methods for forecasting multivariate time series. Handbook of economic forecasting, 2, 899-960.

Pareto, V. (1964). Cours d'économie politique (Vol. 1). Librairie Droz.

Punathumparambath, B. (2011). A new family of skewed slash distributions generated by the normal kernel. Statistica, 71(3), 345-353.

Reyes, J., Gómez, H. W., & Bolfarine, H. (2013). Modified slash distribution. Statistics, 47(5), 929-941.

Rogers, W. H., & Tukey, J. W. (1972). Understanding some long-tailed symmetrical distributions. Statistica Neerlandica, 26(3), 211-226.

El-Saeed, A. R., Hassan, A. S., Elharoun, N. M., Al Mutairi, A., Khashab, R. H., & Nassr, S. G. (2023). A class of power inverted Topp-Leone distribution: Properties, different estimation methods & applications. Journal of Radiation Research and Applied Sciences, 16(4), 100643.

Shaked, M., & Shanthikumar, J. G. (1994). Stochastic orders and their applications.

Shannon, C. E. (1948). A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423.

Sharma, V. K., Singh, S. K., Singh, U., & Agiwal, V. (2015). The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data. Journal of Industrial and Production Engineering, 32(3), 162-173.

Tahir, M. H., Cordeiro, G. M., Ali, S., Dey, S., & Manzoor, A. (2018). The inverted Nadarajah–Haghighi distribution: estimation methods and applications. Journal of Statistical Computation and Simulation, 88(14), 2775-2798.

Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of applied mechanics.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/129307