



Double Inverse-Gaussian Distributions and Associated Inference

A. Almutairi¹ · M. E. Ghitany¹ · A. Alothman¹ · Ramesh C. Gupta²

Accepted: 18 March 2023 / Published online: 22 April 2023
© The Indian Society for Probability and Statistics (ISPS) 2023

Abstract

In order to have more flexibility, several double distributions have been studied in the literature. In this paper, we propose a double inverse Gaussian distribution using random sign mixture transform and study its associated inferences. The maximum likelihood estimation is performed to estimate the parameters. Extensive simulation studies are carried out to examine the performance of the estimators and the corresponding confidence intervals of the parameters. A real data set example is presented to illustrate the procedure.

Keywords Random sign transform · Random sign mixture transform · Inverse-Gaussian distribution · Maximum likelihood estimation · Monte Carlo simulations

Mathematics Subject Classification Primary 62E15 · Secondary · 62F12

1 Introduction

In the last 40 years, there has been a considerable literature on the double continuous distributions on the real line, some of which explain the word double and some do not. Several of them use the word double as the distribution of the absolute value and some of them use the word reflection. Balakrishnan and Kocherlakota (1985) and Rao and Narasimham (1989) presented the double Weibull distribution and studied its order statistics and linear estimation. Bindu and Sangita (2015) studied the double Lomax distribution as the distribution of the absolute value of the ratio of two independent Laplace distributed variables. Govindarajulu (1966) studied the reflected version of the exponential distribution. The reflected version of the generalized Gamma was studied by Plucinska (1965,

✉ M. E. Ghitany
me.ghitany@ku.edu.kw

¹ Department of Statistics and Operations Research, Faculty of Science, Kuwait University, Kuwait, Kuwait

² Department of Mathematics and Statistics, University of Maine, Orono, ME, USA

1966, 1967) and the reflected version of the gamma distribution was studied by Kantam and Narasimham (1991). Kumar and Jose (2019) called the distribution of the absolute value of the Lindley variable as the double Lindley distribution, see also Ibrahim et al. (2020). Nadarajah et al. (2013) presented a double generalized Pareto distribution and Halvarsson (2020) studied double Pareto type II distribution. Armagan et al. (2013) presented a generalized double Pareto shrinkage distribution and used it as a prior for Bayesian shrinkage estimation and inferences in linear models.

Most of the above cited double continuous distributions have some limitations such as (i) non-existence of moments for some values of the parameters, see for example Nadarajah et al. (2013) and, (ii) non-existence of some MLEs of the parameters, see for example de Zea Bermudez and Kotz (2010a) and de Zea Bermudez and Kotz (2010b).

Recently, Aly (2018) presented a unified approach for developing double continuous/discrete distributions using two well known transforms (representations), namely

(i) *random sign transform (RST)*:

$$Z_1 = (2Y - 1)X,$$

where Y is a Bernoulli r.v. with parameter β and X is a non-negative r.v. independent of Y . The probability density function (p.d.f.) of Z_1 is given by

$$f_{Z_1}(z; \beta, \theta) = \begin{cases} \bar{\beta} f_X(|z|; \theta), & z < 0, \\ \beta f_X(z; \theta), & z \geq 0, \end{cases} \quad (1)$$

where $f_X(\cdot; \theta)$ is the p.d.f. of a non-negative r.v. X with (vector) parameter θ and $\bar{\beta} = 1 - \beta$.

(ii) *random sign mixture transform (RSMT)*:

$$Z_2 = YX_1 - (1 - Y)X_2,$$

where Y is a Bernoulli r.v. with parameter β while X_1, X_2 are independent non-negative r.v.'s independent of Y .

The p.d.f. of Z_2 is given by

$$f_{Z_2}(z; \beta, \theta_1, \theta_2) = \begin{cases} \bar{\beta} f_{X_2}(|z|; \theta_2), & z < 0, \\ \beta f_{X_1}(z; \theta_1), & z \geq 0, \end{cases} \quad (2)$$

where $f_{X_j}(\cdot; \theta_j)$, $j = 1, 2$, are the p.d.f.'s of a non-negative r.v.'s X_1, X_2 with (vector) parameters θ_j , $j = 1, 2$.

If X_1 and X_2 are from the same family of distributions \mathcal{F} , we say that Z_2 has a *double \mathcal{F}* distribution.

Note that RST is a special case of RSMT when X_1, X_2 are independent and identically distributed (i.i.d.), i.e. $X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} X$. Moreover, all the above cited double distributions considered only the case $\beta = \frac{1}{2}$.

The inverse Gaussian distribution, denoted by $IG(\mu, \lambda)$, has a p.d.f.

$$f_X(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi}} x^{-3/2} \exp\left[-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right], \quad x > 0, \quad \mu, \lambda > 0,$$

where μ is the mean and λ is the shape parameter. This distribution is a very versatile life distribution and its various modifications and transformations have been extensively studied in the literature. We refer the reader to Gupta and Akman (1995, 1996, 1997, 1998), Gupta and Kundu (2011) and the references therein. The inverse Gaussian distribution has also been studied under the umbrella of Birnbaum Saunders distribution. For a survey article on Birnbaum Saunders distribution, we refer to Balakrishnan and Kundu (2019).

We follow the procedure presented by Aly (2018) and study the double inverse Gaussian distribution. Specifically, we consider four double inverse-Gaussian distributions:

1. Double inverse Gaussian 1, $DIG-1(\beta, \mu_1, \lambda_1, \mu_2, \lambda_2)$,
2. Double inverse Gaussian 2, $DIG-2(\beta, \mu_1, \mu_2, \lambda) \equiv DIG-1(\beta, \mu_1, \lambda, \mu_2, \lambda)$,
3. Double inverse Gaussian 3, $DIG-3(\beta, \mu, \lambda_1, \lambda_2) \equiv DIG-1(\beta, \mu, \lambda_1, \mu, \lambda_2)$,
4. Double inverse Gaussian 4, $DIG-4(\beta, \mu, \lambda) \equiv DIG-1(\beta, \mu, \lambda, \mu, \lambda)$.

These distributions are bimodal with one mode on each side of the origin.

The contents of this paper are organized as follows. In Sect. 2, we present the statistical properties of the double inverse Gaussian distributions, including the probability density function, cumulative distribution function (c.d.f.), modes, moment generating function (m.g.f.), raw moments, variance, skewness, kurtosis, Tsallis entropy, Shannon entropy and extropy. The maximum likelihood estimation of the parameters and their asymptotic distributions are studied in Sect. 3. Extensive simulation studies are carried out in Sect. 4 to study the performance of the estimators. In Sect. 5, a real data set application is presented to illustrate the procedure. Finally, some conclusion and comments are presented in Sect. 6.

2 Statistical Properties

In this section, we present a comprehensive summary of the basic properties of the $DIG-1(\beta, \mu_1, \lambda_1, \mu_2, \lambda_2)$ distribution. These properties include, the p.d.f., c.d.f., modes, m.g.f., raw moments and associated measures, Tsallis entropy, Shannon entropy and extropy. Corresponding properties for the nested distributions $DIG-2$, $DIG-3$ and $DIG-4$ are obtained as special cases when $(\lambda_1 = \lambda_2 = \lambda)$, $(\mu_1 = \mu_2 = \mu)$ and $(\mu_1 = \mu_2 = \mu, \lambda_1 = \lambda_2 = \lambda)$, respectively.

2.1 Probability Density Function

The p.d.f of DIG-1 distribution is given by

$$f_{Z_2}(z) = \begin{cases} \beta f_{X_2}(|z|; \mu_2, \lambda_2), & z < 0, \\ \beta f_{X_1}(z; \mu_1, \lambda_1), & z \geq 0, \end{cases} \tag{3}$$

where

$$f_{X_j}(x; \mu_j, \lambda_j) = \sqrt{\frac{\lambda_j}{2\pi}} x^{-3/2} \exp\left[-\frac{\lambda_j(x - \mu_j)^2}{2\mu_j^2 x}\right], \tag{4}$$

$x > 0, \quad \mu_j, \lambda_j > 0, \quad j = 1, 2,$

are the p.d.f.’s of inverse Gaussian distributions.

Figure 1 shows the bimodality of the p.d.f. of DIG distributions as a function in β . Also, this figure shows that the left (right) peak gets smaller (larger) as β increases.

The DIG-1 distribution has two modes given by

$$\text{Mode}(Z_2) = -\text{Mode}(X_2) \text{ and } \text{Mode}(X_1), \tag{5}$$

where

$$\text{Mode}(X_j) = \mu_j \left[\sqrt{1 + \left(\frac{3\mu_j}{2\lambda_j}\right)^2} - \frac{3\mu_j}{2\lambda_j} \right], \quad j = 1, 2, \tag{6}$$

are the modes of inverse Gaussian distributions.

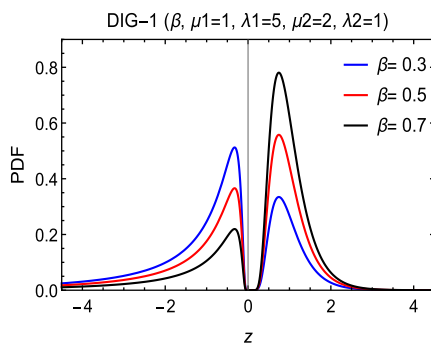


Fig. 1 P.d.f. of DIG-1 distribution

2.2 Cumulative Distribution Function

The c.d.f of DIG-1 distribution is given by

$$F_{Z_2}(z) = \begin{cases} \bar{\beta} [1 - F_{X_2}(|z|; \mu_2, \lambda_2)] & z < 0, \\ \bar{\beta} + \beta F_{X_1}(z; \mu_1, \lambda_1) & z \geq 0, \end{cases} \tag{7}$$

where

$$F_{X_j}(x; \mu_j, \lambda_j) = \Phi\left(\sqrt{\frac{\lambda_j}{x}}\left(\frac{x}{\mu_j} - 1\right)\right) + e^{2\lambda_j/\mu_j} \Phi\left(-\sqrt{\frac{\lambda_j}{x}}\left(\frac{x}{\mu_j} + 1\right)\right), \tag{8}$$

$x > 0, \quad j = 1, 2,$

are the c.d.f.'s of inverse Gaussian distributions and

$$\Phi(a) = P(Z \leq a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz, \quad -\infty < a < \infty, \tag{9}$$

is the c.d.f. of the standard normal distribution.

Figure 2 shows the c.d.f. of the DIG-1 distribution as a function in β . Also, this figure shows that $F_{Z_2}(0) = \bar{\beta}$ and hence $F_{Z_2}(0)$ decreases as β increases.

2.3 Moment Generating Function

The m.g.f. of DIG-1 distribution is given by

$$M_{Z_2}(t) = \beta M_{X_1}(t) + \bar{\beta} M_{X_2}(-t), \quad -\infty < t < \infty, \tag{10}$$

where

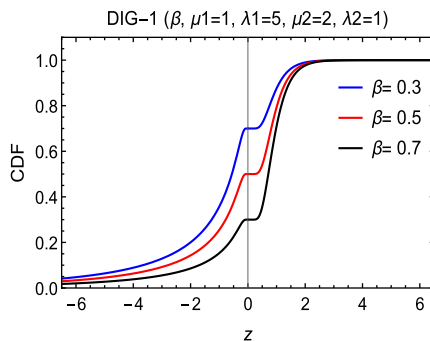


Fig. 2 C.d.f. of DIG-1 distribution

$$M_{X_j}(t) = \exp \left[\frac{\lambda_j}{\mu_j} \left(1 - \sqrt{1 - \frac{2\mu_j^2 t}{\lambda_j}} \right) \right], \quad |t| < \frac{\lambda_j}{2\mu_j^2}, \quad j = 1, 2, \quad (11)$$

are the m.g.f.'s of inverse Gaussian distributions.

2.4 Moments and Associated Measures

The r th moment of DIG-1 distribution is given by

$$E(Z_2^r) = \beta E(X_1^r) + (-1)^r \bar{\beta} E(X_2^r), \quad r \geq 1, \quad (12)$$

where, using the result of Sato and Inoue (1994),

$$E(X_j^r) = \mu_j^r \sum_{i=0}^{r-1} \frac{(r-1+i)!}{r!(r-1-i)!} 2^i \left(\frac{\mu_j}{\lambda_j} \right)^i, \quad j = 1, 2, \quad (13)$$

are the r th moments of inverse Gaussian distributions.

Using the last expressions of $E(Z_2^r)$, the mean, variance, skewness, and kurtosis of DIG-1 distribution are easily obtained. Note that $E(Z_2)$ does not depend on λ_1 and λ_2 .

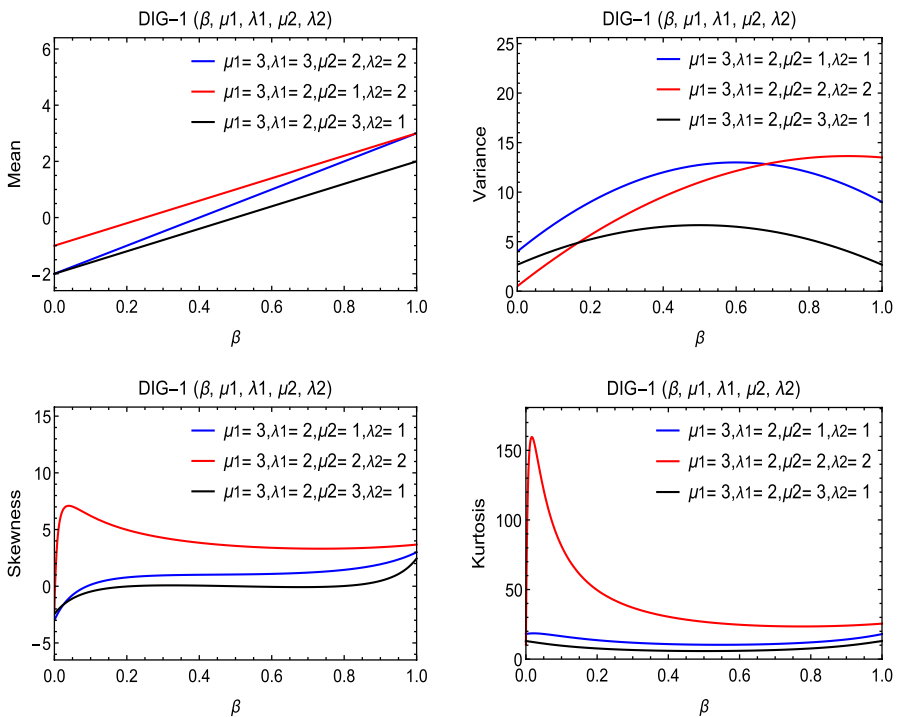


Fig. 3 Mean, variance, skewness and kurtosis of DIG-1 distribution

Figure 3 shows the mean, variance, skewness, and kurtosis of the DIG-1 distribution as a function in β . Also, this figure shows that the skewness can be negative/positive, i.e. the DIG-1 distribution can be skewed to the left/right.

2.5 Tsallis Entropy

Entropies are measures of a system’s variation, instability, or unpredictability. The Tsallis entropy, Tsallis (1988), is an important measure in statistics as index of diversity. It has many applications in areas such as physics, chemistry, biology and economics.

For a continuous r.v. V with p.d.f. $f_V(v)$, the Tsallis entropy of V is defined as

$$\mathcal{T}_\alpha(V) = \frac{1}{\alpha - 1} \left[1 - \int_S f_V^\alpha(v)dv \right] = \frac{1}{\alpha - 1} E[1 - f_V^{\alpha-1}(V)], \quad 0 < \alpha \neq 1,$$

where S is the support of V .

First, we derive the Tsallis entropy of RSMT Z_2 .

$$\begin{aligned} \mathcal{T}_\alpha(Z_2) &= \frac{1}{\alpha - 1} \left\{ 1 - \int_{-\infty}^\infty f_{Z_2}^\alpha(z)dz \right\} \\ &= \frac{1}{\alpha - 1} \left\{ 1 - \int_0^\infty \beta^\alpha f_{X_1}^\alpha(z)dz - \int_{-\infty}^0 \bar{\beta}^\alpha f_{X_2}^\alpha(-z)dz \right\} \\ &= \frac{1}{\alpha - 1} \left\{ 1 - \beta^\alpha \int_0^\infty f_{X_1}^\alpha(z)dz - \bar{\beta}^\alpha \int_0^\infty f_{X_2}^\alpha(x)dx \right\} \\ &= \frac{1}{\alpha - 1} \left\{ 1 - \beta^\alpha E[f_{X_1}^{\alpha-1}(X_1)] - \bar{\beta}^\alpha E[f_{X_2}^{\alpha-1}(X_2)] \right\} \\ &= \frac{1}{\alpha - 1} \left\{ 1 - \beta^\alpha [1 - (\alpha - 1)\mathcal{T}_\alpha(X_1)] - \bar{\beta}^\alpha [1 - (\alpha - 1)\mathcal{T}_\alpha(X_2)] \right\} \\ &= \frac{1}{\alpha - 1} (1 - \beta^\alpha - \bar{\beta}^\alpha) + \beta^\alpha \mathcal{T}_\alpha(X_1) + \bar{\beta}^\alpha \mathcal{T}_\alpha(X_2), \\ &= \mathcal{T}_\alpha(Y) + \beta^\alpha \mathcal{T}_\alpha(X_1) + \bar{\beta}^\alpha \mathcal{T}_\alpha(X_2), \end{aligned} \tag{14}$$

where

$$\mathcal{T}_\alpha(Y) = \frac{1}{\alpha - 1} E[1 - f_Y^{\alpha-1}(Y)] = \frac{1}{\alpha - 1} (1 - \beta^\alpha - \bar{\beta}^\alpha). \tag{15}$$

Note that Tsallis entropy of RSMT $\mathcal{T}(Z_2)$ is a non-linear function in β .

Second, we find the Tsallis entropy of $X \sim IG(\mu, \lambda)$.

$$\begin{aligned}
T_{\alpha}(X) &= \frac{1}{\alpha - 1} \left\{ 1 - \int_0^{\infty} f_X^{\alpha}(x) dx \right\} \\
&= \frac{1}{\alpha - 1} \left\{ 1 - \int_0^{\infty} \left(\frac{\lambda}{2\pi} \right)^{\alpha/2} \frac{1}{x^{3\alpha/2}} \exp \left[-\frac{\alpha\lambda}{2\mu^2 x} (x - \mu)^2 \right] dx \right\} \\
&= \frac{1}{\alpha - 1} \left\{ 1 - \left(\frac{\lambda}{2\pi} \right)^{\alpha/2} e^{\alpha\lambda/\mu} \int_0^{\infty} \frac{1}{x^{3\alpha/2}} \exp \left[-\left(\frac{\alpha\lambda}{2\mu^2} x + \frac{\alpha\lambda}{2x} \right) \right] dx \right\} \\
&= \frac{1}{\alpha - 1} \left\{ 1 - \left(\frac{\lambda}{2\pi} \right)^{\alpha/2} \left(\frac{2\mu^2}{\alpha\lambda} \right)^{-(3\alpha/2-1)} e^{\alpha\lambda/\mu} \right. \\
&\quad \left. \int_0^{\infty} \frac{1}{t^{3\alpha/2}} \exp \left[-\left(t + \frac{(\alpha\lambda/\mu)^2}{4t} \right) \right] dt \right\} \\
&= \frac{1}{\alpha - 1} \left\{ 1 - \left(\frac{\lambda/\mu}{2\pi} \right)^{\alpha/2} \frac{2e^{\alpha\lambda/\mu}}{\mu^{\alpha-1}} K_{3\alpha/2-1}(\alpha\lambda/\mu) \right\}.
\end{aligned} \tag{16}$$

where

$$K_{\nu}(s) = \frac{1}{2}(s/2)^{\nu} \int_0^{\infty} \frac{1}{t^{\nu+1}} \exp \left[-\left(t + \frac{s^2}{4t} \right) \right] dt, \quad -\infty < \nu < \infty, \quad s > 0, \tag{17}$$

is the modified Bessel function of the second kind.

Therefore, Tsallis entropy of DIG-1 distribution is explicitly given by

$$T_{\alpha}(Z_2) = T_{\alpha}(Y) + \beta^{\alpha} T_{\alpha}(X_1) + \bar{\beta}^{\alpha} T_{\alpha}(X_2), \tag{18}$$

where $T_{\alpha}(Y)$ is given by (15) and

$$T_{\alpha}(X_j) = \frac{1}{\alpha - 1} \left\{ 1 - \left(\frac{\lambda_j/\mu_j}{2\pi} \right)^{\alpha/2} \frac{2e^{\alpha\lambda_j/\mu_j}}{\mu_j^{\alpha-1}} K_{3\alpha/2-1}(\alpha\lambda_j/\mu_j) \right\}, \quad j = 1, 2. \tag{19}$$

2.6 Shannon Entropy

Using L'Hospital rule, we have

$$\lim_{\alpha \rightarrow 1} T_{\alpha}(V) = - \int_S \ln(f_V(v)) f_V(v) dv = E[-\ln f_V(V)] = \mathcal{H}(V),$$

which is the Shannon entropy of V , Shannon (1948).

As $\alpha \rightarrow 1$, (14) and (15), are simplified to

$$\mathcal{H}(Z_2) = \mathcal{H}(Y) + \beta \mathcal{H}(X_1) + \bar{\beta} \mathcal{H}(X_2), \tag{20}$$

where

$$\mathcal{H}(Y) = -\beta \ln \beta - \bar{\beta} \ln \bar{\beta}, \tag{21}$$

which agrees with the result obtained by Aly (2018). Note that Shannon entropy of RSMT $\mathcal{H}(Z_2)$ is a non-linear function in β .

Using (16), the Shannon entropy of $X \sim \text{IG}(\mu, \lambda)$ is given by

$$\begin{aligned} \mathcal{H}(X) &= \lim_{\alpha \rightarrow 1} T_\alpha(X) \\ &= \frac{1}{2} - \frac{1}{2} \ln \left(\frac{\lambda}{2\pi\mu^3} \right) - \frac{3}{2} \sqrt{\frac{2\lambda}{\pi\mu}} e^{\lambda/\mu} \frac{\partial}{\partial v} K_\nu(\lambda/\mu)|_{\nu=1/2}, \end{aligned} \tag{22}$$

where $K_\nu(s)$, $s > 0$, is given by (17). The proof follows by using L'Hôpital's rule.

The last expression can be calculated using the Mathematica function $\text{BesselK}^{(1,0)}[1/2, \lambda/\mu] = \frac{\partial}{\partial v} K_\nu(\lambda/\mu)|_{\nu=1/2}$.

Using (20), the Shannon entropy of DIG-1 distribution is given by

$$\mathcal{H}(Z_2) = \mathcal{H}(Y) + \beta \mathcal{H}(X_1) + \bar{\beta} \mathcal{H}(X_2), \tag{23}$$

where $\mathcal{H}(Y)$ is given by (21) and

$$\mathcal{H}(X_j) = \frac{1}{2} - \frac{1}{2} \ln \left(\frac{\lambda_j}{2\pi\mu_j^3} \right) - \frac{3}{2} \sqrt{\frac{2\lambda_j}{\pi\mu_j}} e^{\lambda_j/\mu_j} \frac{\partial}{\partial v_j} K_{\nu_j}(\lambda_j/\mu_j)|_{\nu_j=1/2}, \quad j = 1, 2. \tag{24}$$

2.7 Entropy

The following relation

$$\frac{1}{2} [\mathcal{T}_2(V) - 1] = -\frac{1}{2} \int_S f_V^2(v) dv = E \left[-\frac{1}{2} f_V(V) \right] = \mathcal{J}(V), \tag{25}$$

is known as the entropy of V (Lad et al. 2015).

Using (14) and (15), with $\alpha = 2$, the entropy of RSMT is given by

$$\mathcal{J}(Z_2) = \frac{1}{2} [\mathcal{T}_2(Z_2) - 1] = \beta^2 \mathcal{J}(X_1) + \bar{\beta}^2 \mathcal{J}(X_2). \tag{26}$$

Note that entropy of RSMT $\mathcal{H}(Z_2)$ is a quadratic function in β .

Using (16), with $\alpha = 2$, the entropy of $X \sim \text{IG}(\mu, \lambda)$ is given by

$$\mathcal{J}(X) = \frac{1}{2} [\mathcal{T}_2(X) - 1] = -\frac{\lambda}{2\pi\mu^2} e^{2\lambda/\mu} K_2(2\lambda/\mu), \tag{27}$$

where $K_\nu(z)$ is given by (17).

The entropy of DIG-1 distribution is given by

$$\mathcal{J}(Z_2) = \beta^2 \mathcal{J}(X_1) + \bar{\beta}^2 \mathcal{J}(X_2), \tag{28}$$

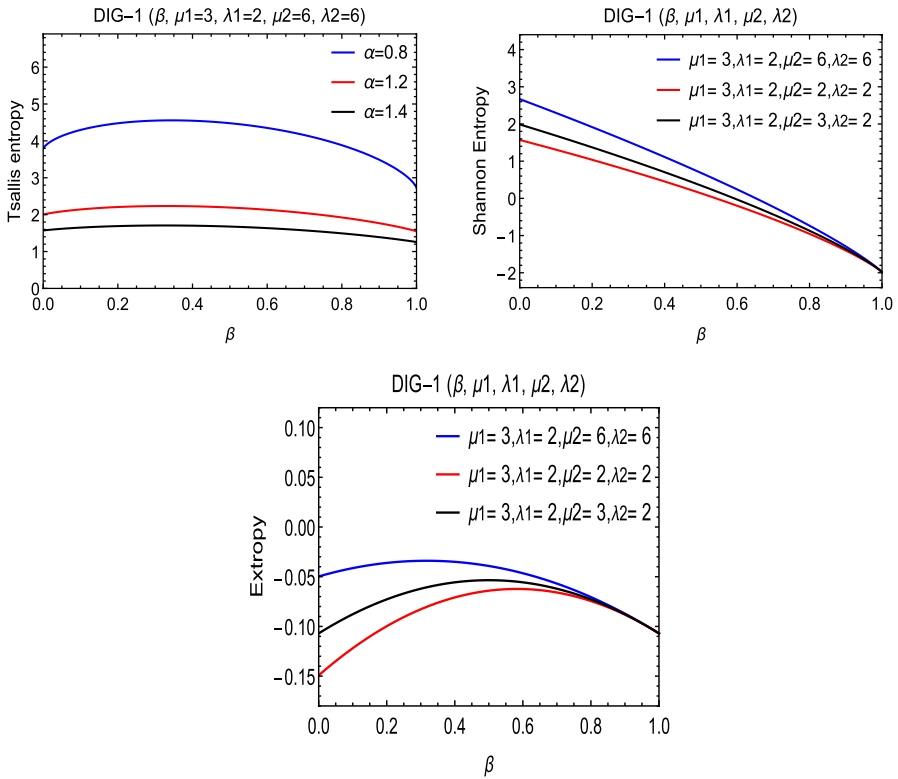


Fig. 4 Tsallis entropy, Shannon entropy and entropy of DIG-1 distribution

where

$$\mathcal{J}(X_j) = - \frac{\lambda_j}{2\pi\mu_j^2} e^{2\lambda_j/\mu_j} K_2(2\lambda_j/\mu_j), \quad j = 1, 2, \tag{29}$$

where $K_\nu(z)$ is given by (17).

Figure 4 shows Tsallis entropy, Shannon entropy and entropy of DIG-1 distribution as a function in β for selected values of the parameters. This figure also shows that Tsallis entropy of DIG-1 distribution decreases as α increases.

3 Maximum Likelihood Estimation

In this section, we derive the MLEs of the parameters of DIG distributions and their asymptotic distributions. These asymptotic distributions turned out to be multivariate normal which can be used to make statistical inference (confidence intervals and hypothesis testing) about the parameters of DIG distributions.

3.1 DIG-1: Maximum Likelihood Estimation

Let $z_{2,1}, z_{2,2}, \dots, z_{2,n}$ be a r.s. from $\text{DIG-1}(\beta, \mu_1, \lambda_1, \mu_2, \lambda_2)$ distribution. The log-likelihood function is given by

$$\ln L_1 = \sum_{i=1}^n \ln[\beta f_{X_1}(z_{2,i}; \mu_1, \lambda_1)] \mathbf{1}_{\{z_{2,i}>0\}} + \sum_{i=1}^n \ln[\bar{\beta} f_{X_2}(|z_{2,i}|; \mu_2, \lambda_2)] \mathbf{1}_{\{z_{2,i}<0\}}, \tag{30}$$

where $\mathbf{1}_A = 1(0)$ if A is true (false) is the indicator function.

The MLEs of $(\beta, \mu_1, \lambda_1, \mu_2, \lambda_2)$ are:

$$\hat{\beta} = \frac{n_1}{n}, \quad \hat{\mu}_1 = \frac{a_1}{n_1}, \quad \hat{\lambda}_1 = \frac{n_1}{c_1 - \frac{n_1^2}{a_1}}, \quad \hat{\mu}_2 = \frac{a_2}{n_2}, \quad \hat{\lambda}_2 = \frac{n_2}{c_2 - \frac{n_2^2}{a_2}}, \tag{31}$$

where

$$n_1 = \sum_{i=1}^n \mathbf{1}_{\{z_{2,i}>0\}}, \quad n_2 = \sum_{i=1}^n \mathbf{1}_{\{z_{2,i}<0\}}, \quad n_1 + n_2 = n, \tag{32}$$

$$a_1 = \sum_{i=1}^n z_{2,i} \mathbf{1}_{\{z_{2,i}>0\}}, \quad a_2 = \sum_{i=1}^n |z_{2,i}| \mathbf{1}_{\{z_{2,i}<0\}}, \tag{33}$$

$$c_1 = \sum_{i=1}^n \frac{1}{z_{2,i}} \mathbf{1}_{\{z_{2,i}>0\}}, \quad c_2 = \sum_{i=1}^n \frac{1}{|z_{2,i}|} \mathbf{1}_{\{z_{2,i}<0\}}. \tag{34}$$

Moreover, the asymptotic distribution of the MLEs is given by:

as $n \rightarrow \infty$,

$$\sqrt{n} \begin{bmatrix} \hat{\beta} - \beta \\ \hat{\mu}_1 - \mu_1 \\ \hat{\lambda}_1 - \lambda_1 \\ \hat{\mu}_2 - \mu_2 \\ \hat{\lambda}_2 - \lambda_2 \end{bmatrix} \xrightarrow{d} MVN \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \beta \bar{\beta} & 0 & 0 & 0 & 0 \\ 0 & \frac{\mu_1^3}{\beta \lambda_1} & 0 & 0 & 0 \\ 0 & 0 & \frac{2\lambda_1^2}{\beta} & 0 & 0 \\ 0 & 0 & 0 & \frac{\mu_2^3}{\beta \lambda_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{2\lambda_2^2}{\beta} \end{bmatrix} \right). \tag{35}$$

where \xrightarrow{d} denotes convergence in distribution and MVN stands for multivariate normal distribution.

3.2 DIG-2: Maximum Likelihood Estimation

Let $z_{2,1}, z_{2,2}, \dots, z_{2,n}$ be a r.s. from $DIG-2(\beta, \mu_1, \mu_2, \lambda)$ distribution. The log-likelihood function is given by

$$\ln L_2 = \sum_{i=1}^n \ln[\beta f_{X_1}\{z_{2,i}; \mu_1, \lambda\}] \mathbf{1}_{\{z_{2,i}>0\}} + \sum_{i=1}^n \ln[\bar{\beta} f_{X_2}(|z_{2,i}|; \mu_2, \lambda)] \mathbf{1}_{\{z_{2,i}<0\}}. \tag{36}$$

The MLE of β is

$$\hat{\beta} = \frac{n_1}{n}, \tag{37}$$

and the MLEs of (μ_1, μ_2, λ) are the solutions of the normal equations:

$$\begin{aligned} \frac{n}{\hat{\lambda}} - \sum_{i=1}^n \frac{1}{\hat{\mu}_1^2 z_{2,i}} \{z_{2,i} - \hat{\mu}_1\}^2 \mathbf{1}_{\{z_{2,i}>0\}} - \sum_{i=1}^n \frac{1}{\hat{\mu}_2^2 |z_{2,i}|} (|z_{2,i}| - \hat{\mu}_2)^2 \mathbf{1}_{\{z_{2,i}<0\}} &= 0, \\ \sum_{i=1}^n \{z_{2,i} - \hat{\mu}_1\} \mathbf{1}_{\{z_{2,i}>0\}} &= 0, \\ \sum_{i=1}^n (|z_{2,i}| - \hat{\mu}_2) \mathbf{1}_{\{z_{2,i}<0\}} &= 0. \end{aligned}$$

It follows that

$$\hat{\mu}_1 = \frac{a_1}{n_1}, \quad \hat{\mu}_2 = \frac{a_2}{n_2}, \quad \hat{\lambda} = \frac{n}{c_1 - \frac{n_1^2}{a_1} + c_2 - \frac{n_2^2}{a_2}}. \tag{38}$$

Moreover, the asymptotic distribution of the MLEs is given by: as $n \rightarrow \infty$,

$$\sqrt{n} \begin{bmatrix} \hat{\beta} - \beta \\ \hat{\lambda} - \lambda \\ \hat{\mu}_1 - \mu_1 \\ \hat{\mu}_2 - \mu_2 \end{bmatrix} \xrightarrow{d} MVN \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \beta\bar{\beta} & 0 & 0 & 0 \\ 0 & 2\lambda^2 & 0 & 0 \\ 0 & 0 & \frac{\mu_1^3}{\beta\lambda} & 0 \\ 0 & 0 & 0 & \frac{\mu_2^3}{\beta\lambda} \end{bmatrix} \right). \tag{39}$$

3.3 DIG-3: Maximum Likelihood Estimation

Let $z_{2,1}, z_{2,2}, \dots, z_{2,n}$ be a r.s. from DIG-3($\beta, \mu, \lambda_1, \lambda_2$) distribution. The log-likelihood function is given by

$$\ln L_3 = \sum_{i=1}^n \ln[\beta f_{X_1}\{z_{2,i}; \mu, \lambda_1\}] \mathbf{1}_{\{z_{2,i}>0\}} + \sum_{i=1}^n \ln[\bar{\beta} f_{X_2}(|z_{2,i}|; \mu, \lambda_2)] \mathbf{1}_{\{z_{2,i}<0\}}. \tag{40}$$

The MLE of β is

$$\hat{\beta} = \frac{n_1}{n}, \tag{41}$$

and the MLEs of $(\mu, \lambda_1, \lambda_2)$ are the solutions of the normal equations:

$$\begin{aligned} \hat{\lambda}_1 \sum_{i=1}^n \{z_{2,i} - \hat{\mu}_1\} \mathbf{1}_{\{z_{2,i}>0\}} + \hat{\lambda}_2 \sum_{i=1}^n (|z_{2,i}| - \hat{\mu}_1) \mathbf{1}_{\{z_{2,i}<0\}} &= 0, \\ \frac{n_1}{\hat{\lambda}_1} - \sum_{i=1}^n \frac{1}{\hat{\mu}^2 z_i} \{z_{2,i} - \hat{\mu}\}^2 \mathbf{1}_{\{z_{2,i}>0\}} &= 0, \\ \frac{n_2}{\hat{\lambda}_2} - \sum_{i=1}^n \frac{1}{\hat{\mu}^2 z_{2,i}} \{z_i - \hat{\mu}\}^2 \mathbf{1}_{\{z_{2,i}<0\}} &= 0. \end{aligned}$$

It follows that

$$\hat{\lambda}_1 = \frac{n_1}{\frac{a_1}{\hat{\mu}^2} - \frac{2n_1}{\hat{\mu}} + c_1}, \quad \hat{\lambda}_2 = \frac{n_2}{\frac{a_2}{\hat{\mu}^2} - \frac{2n_2}{\hat{\mu}} + c_2}, \tag{42}$$

where $\hat{\mu}$ is the solution in μ of the cubic equation:

$$A\mu^3 + B\mu^2 + C\mu + D = 0, \tag{43}$$

where

$$\begin{aligned} A &= n_1^2 c_2 + n_2^2 c_1, \\ B &= -(n_1 c_2 a_1 + n_2 c_1 a_2 + 2n_1 n_2 n), \\ C &= n_1^2 a_2 + n_2^2 a_1 + 2n_1 n_2 (a_1 + a_2), \\ D &= -n a_1 a_2. \end{aligned}$$

The discriminant of the above cubic equation is given by

$$\Delta = 18ABCD - 4B^3D + B^2C^2 - 4AC^3 - 27A^2D^2$$

and it is well known that if this is negative, the cubic equation has a unique real root. This will imply that the MLEs $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are also unique.

Moreover, the asymptotic distribution of the MLEs is given by:
as $n \rightarrow \infty$,

$$\sqrt{n} \begin{bmatrix} \hat{\beta} - \beta \\ \hat{\mu} - \mu \\ \hat{\lambda}_1 - \lambda_1 \\ \hat{\lambda}_2 - \lambda_2 \end{bmatrix} \xrightarrow{d} MVN \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \beta\bar{\beta} & 0 & 0 & 0 \\ 0 & \frac{\mu^3}{\beta\lambda_1 + \beta\lambda_2} & 0 & 0 \\ 0 & 0 & \frac{2\lambda_1^2}{\beta} & 0 \\ 0 & 0 & 0 & \frac{2\lambda_2^2}{\beta} \end{bmatrix} \right). \tag{44}$$

3.4 DIG-4: Maximum Likelihood Estimation

Let $z_{1,1}, z_{1,2}, \dots, z_{1,n}$ be a r.s. from DIG-4(β, μ, λ) distribution. The log-likelihood function is given by

$$\begin{aligned} \ln L_4 &= \sum_{i=1}^n \ln[\beta f_X\{z_{1,i}; \mu, \lambda\}] \mathbf{1}_{\{z_{1,i} > 0\}} + \sum_{i=1}^n \ln[\bar{\beta} f_X(|z_{1,i}|; \mu, \lambda)] \mathbf{1}_{\{z_{1,i} < 0\}} \\ &= n_1 \ln \beta + n_2 \ln \bar{\beta} + \sum_{i=1}^n \ln f_X(|z_{1,i}|; \mu, \lambda). \end{aligned} \tag{45}$$

The MLEs of (β, μ, λ) are:

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{z_{1,i} > 0\}}, \quad \hat{\mu} = \frac{1}{n} \sum_{i=1}^n |z_{1,i}|, \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^n \frac{1}{|z_{1,i}|} - \frac{n}{\hat{\mu}}}. \tag{46}$$

Moreover, the asymptotic distribution of the MLEs is given by:
as $n \rightarrow \infty$,

$$\sqrt{n} \begin{bmatrix} \hat{\beta} - \beta \\ \hat{\mu} - \mu \\ \hat{\lambda} - \lambda \end{bmatrix} \xrightarrow{d} MVN \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \beta\bar{\beta} & 0 & 0 \\ 0 & \frac{\mu^3}{\lambda} & 0 \\ 0 & 0 & 2\lambda^2 \end{bmatrix} \right). \tag{47}$$

4 Simulations

The purpose of this section is to perform simulation studies to evaluate the behaviour of the MLEs of the parameters of the proposed four DIG distributions. Such behaviour will be evaluated in terms of the bias, mean-square error of the MLEs and the coverage probability of the 95% confidence intervals of the parameters. All computations in the simulation studies were done using the R language Version 4.0.5 for Windows.

To generate a random sample of size n , $Z_{2,1}, Z_{2,2}, \dots, Z_{2,n}$, from DIG-1, DIG-2 and DIG-3 distributions, we use the following algorithms:

1. Generate $Y_i \sim \text{Bernoulli}(\beta)$, $i = 1, 2, \dots, n$;
2. Generate $X_{1,i} \sim \text{IG}(\mu_1, \lambda_1)$, $i = 1, 2, \dots, n$;
3. Generate $X_{2,i} \sim \text{IG}(\mu_2, \lambda_2)$, $i = 1, 2, \dots, n$;
4. Set $Z_{2,i} = Y_i X_{1,i} - (1 - Y_i) X_{2,i}$, $i = 1, 2, \dots, n$.

To generate a random sample of size n , $Z_{1,1}, Z_{1,2}, \dots, Z_{1,n}$, from DIG-4 distribution, we use the following algorithm:

1. Generate $Y_i \sim \text{Bernoulli}(\beta)$, $i = 1, 2, \dots, n$;
2. Generate $X_i \sim \text{IG}(\mu, \lambda)$, $i = 1, 2, \dots, n$;
3. Set $Z_{1,i} = (2Y_i - 1) X_i$, $i = 1, 2, \dots, n$.

The sample sizes considered in the simulation studies are $n = 50, 100, \dots, 500$.

The above process of generating random data from DIG distributions is repeated $M = 10,000$ times. In each of the M repetitions, the MLEs of the parameters and their standard errors (S.E.) were calculated using the expressions given in Subsections 3.1 to 3.4.

Measures examined in these simulation studies are:

- (1) Bias of the MLE $\hat{\nu}$ of the parameter $\nu = \beta, \mu_1, \lambda_1, \mu_2, \lambda_2$:

$$\text{Bias}(\hat{\nu}) = \frac{1}{M} \sum_{i=1}^M (\hat{\nu}_i - \nu),$$

where $\hat{\nu}_i$ is the MLE of the parameter ν in the i th simulation repetition.

- (2) Mean square error (MSE) of the MLE $\hat{\nu}$ of the parameter ν :

$$\text{MSE}(\hat{\nu}) = \frac{1}{M} \sum_{i=1}^M (\hat{\nu}_i - \nu)^2.$$

- (3) Coverage probability (CP) of 95% confidence intervals of the parameter ν :

$$\text{CP}(\nu) = \frac{1}{M} \sum_{i=1}^M \mathbf{1}_{\{\nu \in (L_i, U_i)\}},$$

where $L_i = \hat{v}_i - 1.96 \text{ S.E.}(\hat{v}_i)$, $U_i = \hat{v}_i + 1.96 \text{ S.E.}(\hat{v}_i)$, $i = 1, 2, \dots, M$.
The reported figures of the simulation studies support the following conclusions:

1. Figures 5, 6, 7, 8 show that the *absolute* biases of the MLEs of the parameters are small and tend to zero for large n .
2. Figures 9, 10, 11, 12 show that the MSE of the MLEs of the parameters are small and decrease as n increases.
3. Figures 13, 14, 15, 16 show that the coverage probability of 95% confidence intervals of the parameters is close to the nominal level of 95%.

The above conclusions show that the MLEs of the parameters of the DIG distributions are well behaved for point estimation and confidence intervals.

5 Application

In this section, we apply the proposed DIG models to a real data set for illustration. The description of the data is as follows.

In an online final exam at Kuwait university during Covid-19 shut down, students are requested to write down their solutions on paper sheets, scan these sheets as a “pdf” file and send such file to the instructor via Teams Chat. The time of submitting the solution file of each student is recorded automatically on the Teams system. Here, we are interested in modelling the difference between the time (in minutes) spent to submitting the solution file t_i and the two hours exam period of 38 students, i.e. $z_i = t_i - 120$, $i = 1, 2, \dots, 38$.

The data set is given below.

– 18.06, –17.45, –9.90, –8.62, –6.14, –3.47, –2.57, –2.43, – 1.56, 0.84, 1.14, 1.26, 1.34, 1.58, 1.81, 1.82, 1.89, 2.11, 2.23, 2.26, 2.33, 2.36, 2.40, 2.43, 2.52, 2.89, 2.92, 3.25, 3.30, 3.30, 3.47, 3.71, 3.77, 4.02, 4.41, 4.85, 5.20, 7.94 where negative (positive) value means the student submitted the solution file earlier (later) than the two hours exam time.

Table 1 shows the MLE’s of the parameters, their standard errors (S.E.’s) and the maximized log-likelihood of the DIG models. Note that for DIG-3, the discriminant $\Delta = -1.08345 \times 10^{17}$, showing that the MLE $\hat{\mu}$ is unique.

Table 2 shows two goodness-of-fit tests, Anderson-Darling (AD) and Cramer von-Misses (CvM) tests. Clearly, this table shows that all DIG models pass the two tests, i.e., we accept the null hypothesis that the data are drawn from each of the DIG models. However, the test statistics (p -value) for DIG-1 and DIG-2 are much smaller (larger) than those for DIG-3 and DIG-4.

Since DIG-2, DIG-3 and DIG-4 models are nested in DIG-1 model, we can use the likelihood ratio test (LRT) to test each of the following hypotheses:

- (i) $H_0 : \lambda_1 = \lambda_2$ (DIG-2 model) versus $H_1 : \lambda_1 \neq \lambda_2$ (DIG-1 model)

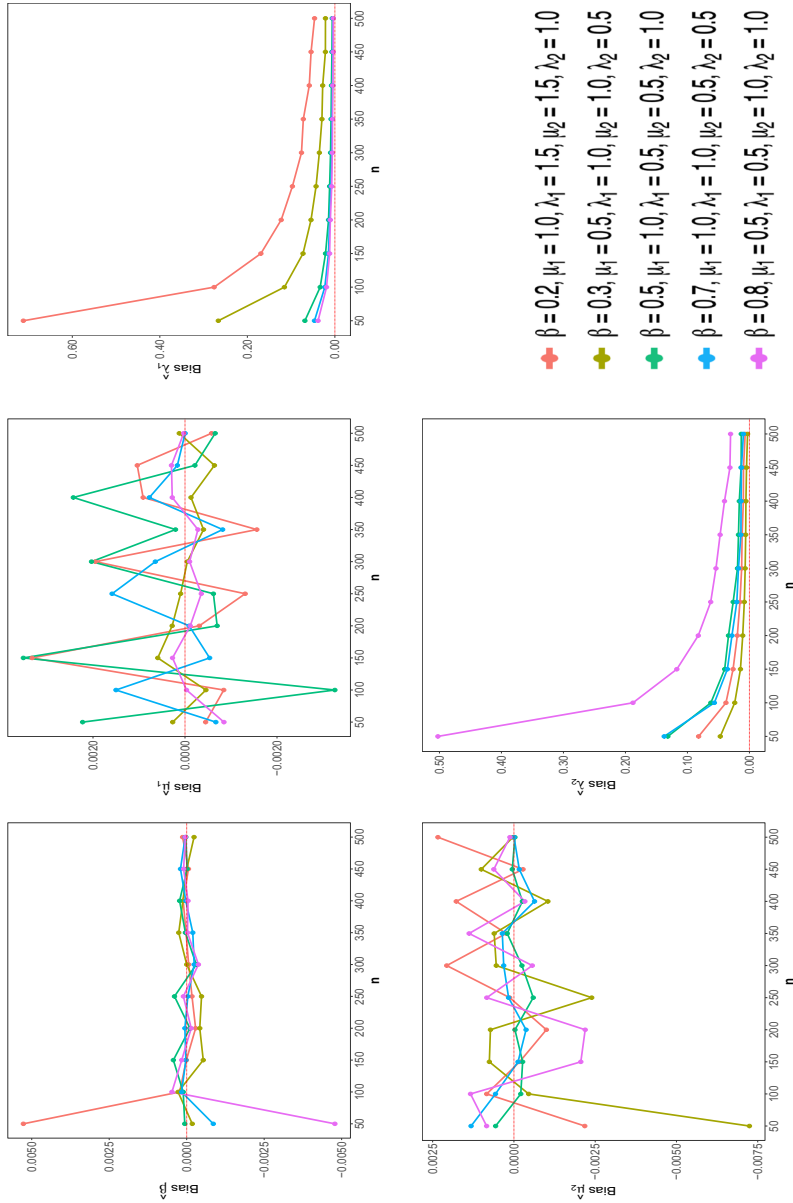


Fig. 5 Bias of the MLEs of the parameters of DIG-1 distribution

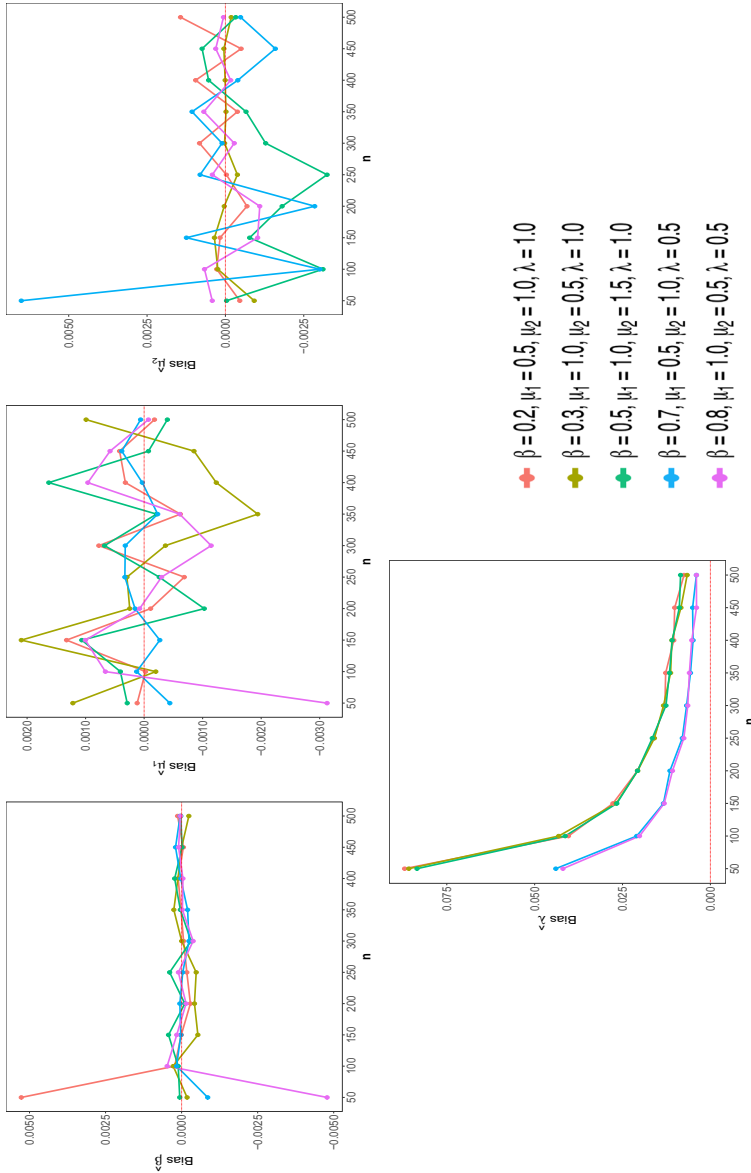


Fig. 6 Bias of the MLEs of the parameters of DIG-2 distribution

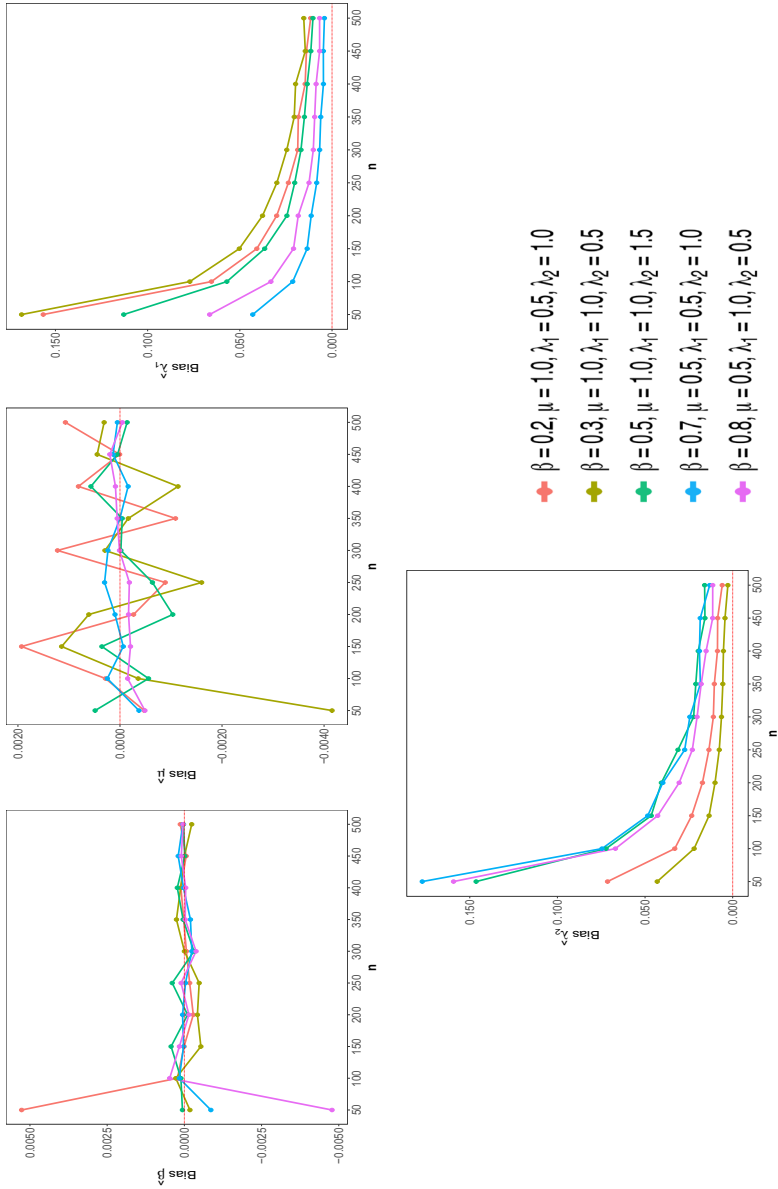


Fig. 7 Bias of the MLEs of the parameters of DIG-3 distribution

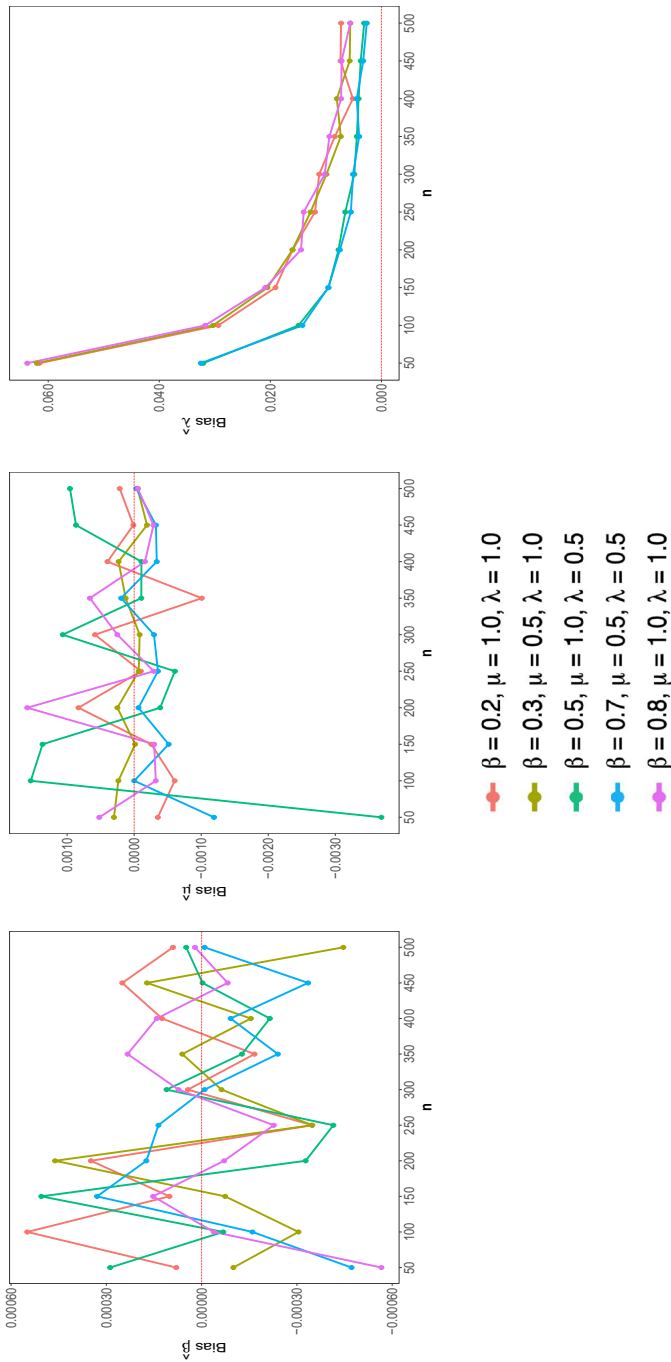


Fig. 8 Bias of the MLEs of the parameters of DIG-4 distribution

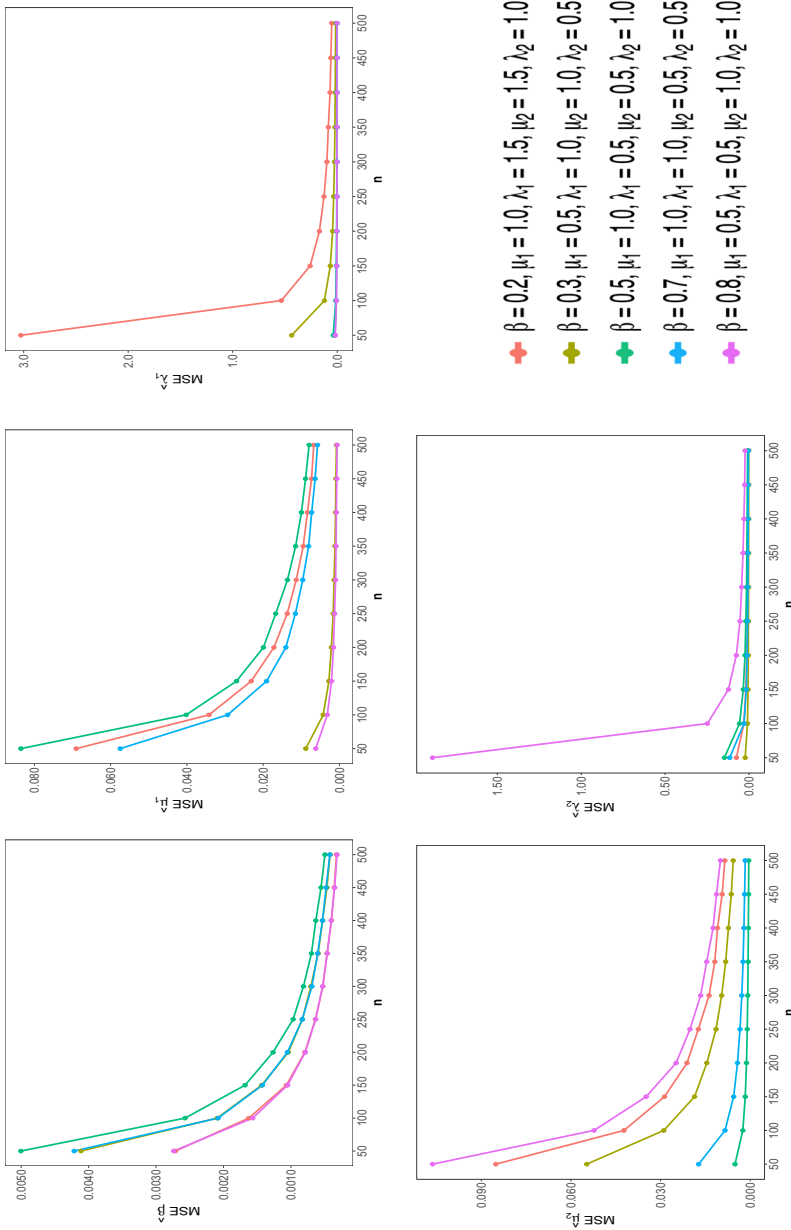


Fig. 9 MSE of the MLEs of the parameters of DIG-I distribution

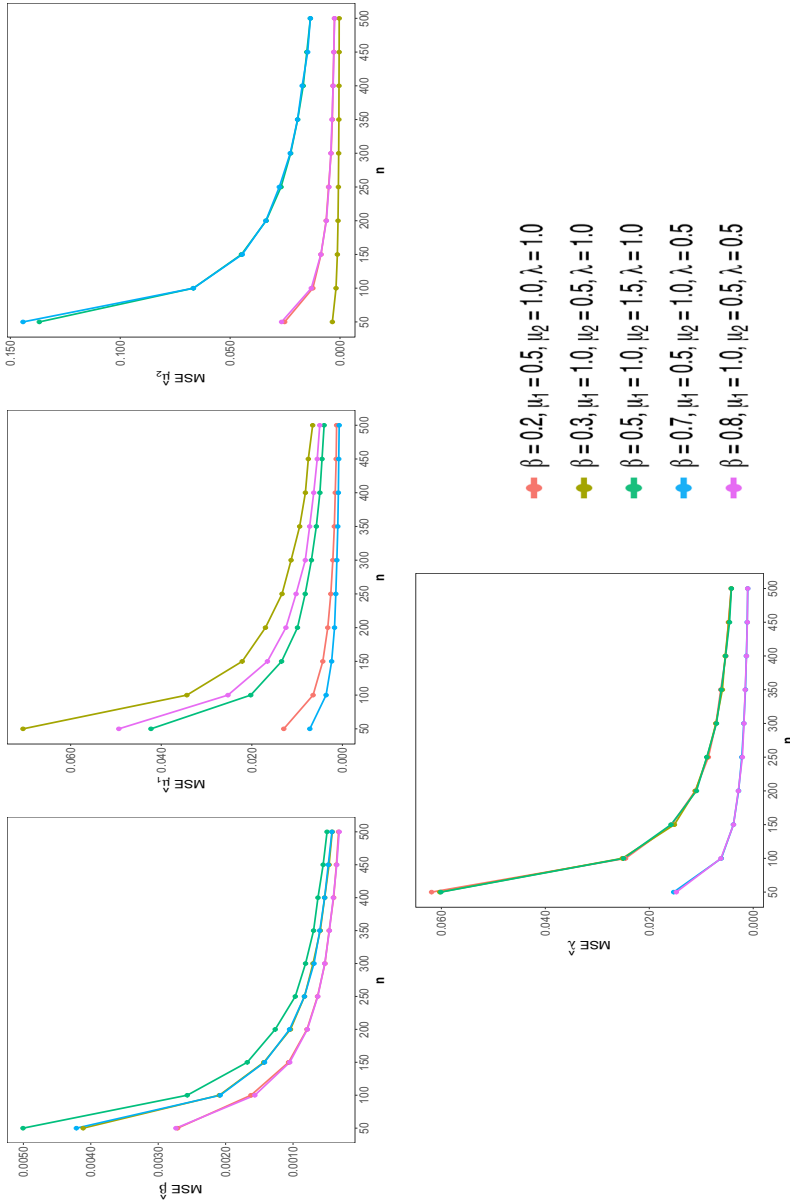


Fig. 10 MSE of the parameters of DIG-2 distribution

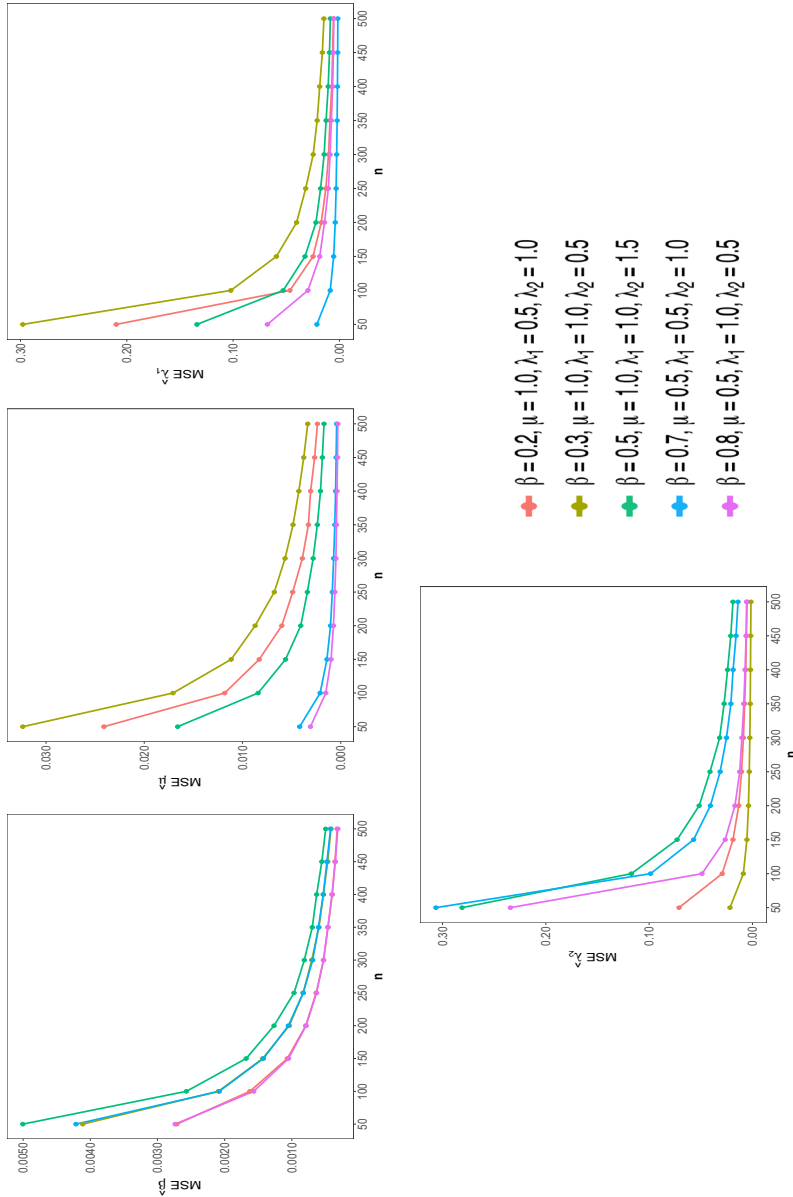


Fig. 11 MSE of the MLEs of the parameters of DIG-3 distribution

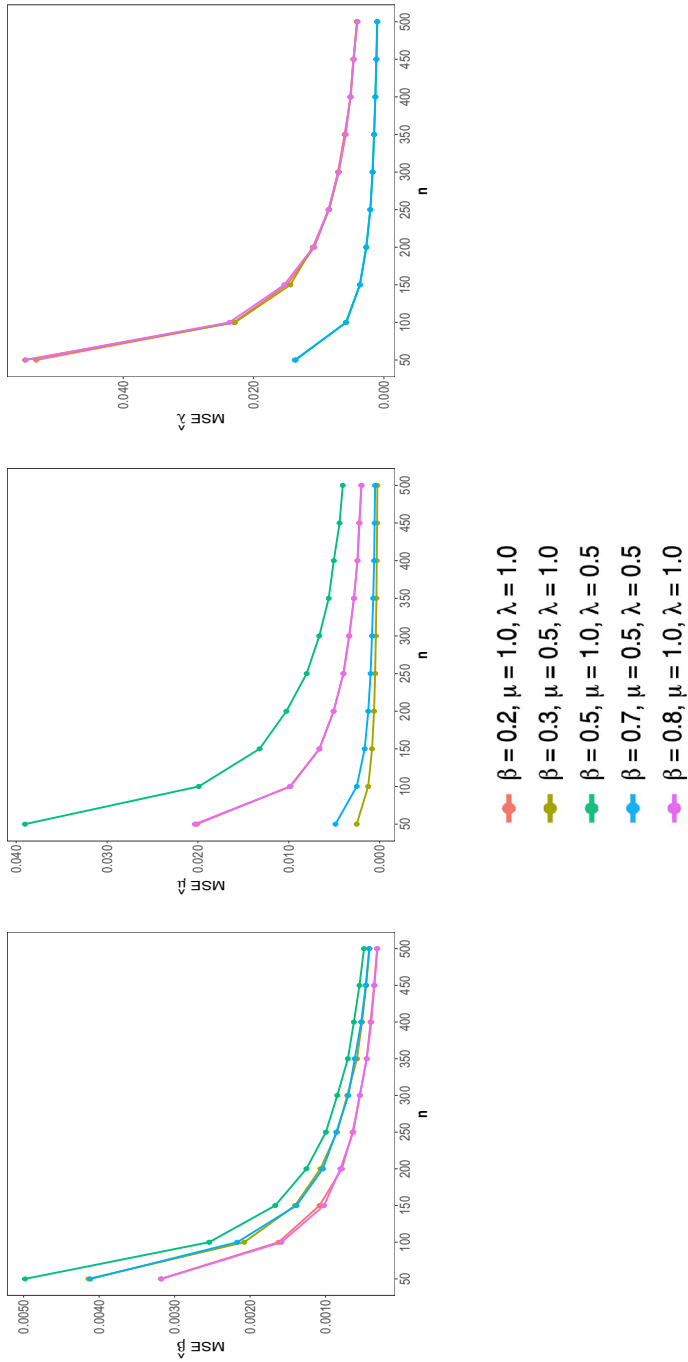


Fig. 12 MSE of the MLEs of the parameters of DIG-4 distribution

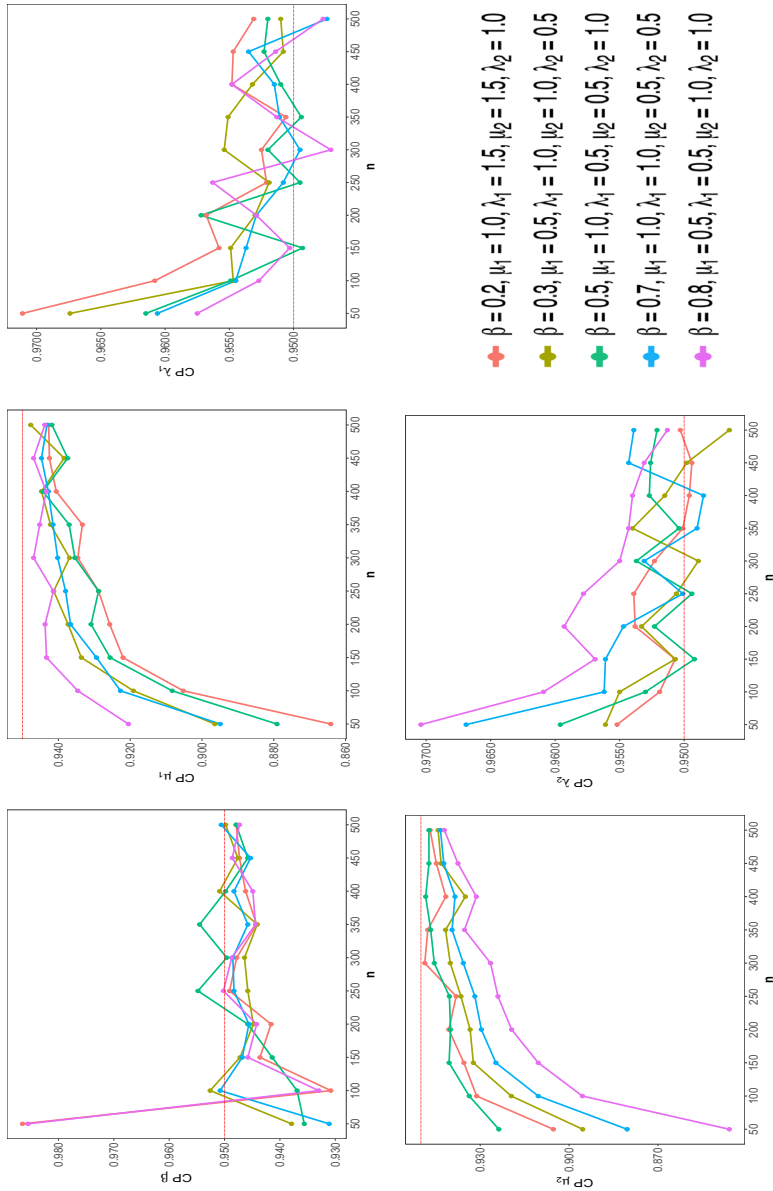


Fig. 13 Coverage probability of the parameters of DIG-1 distribution

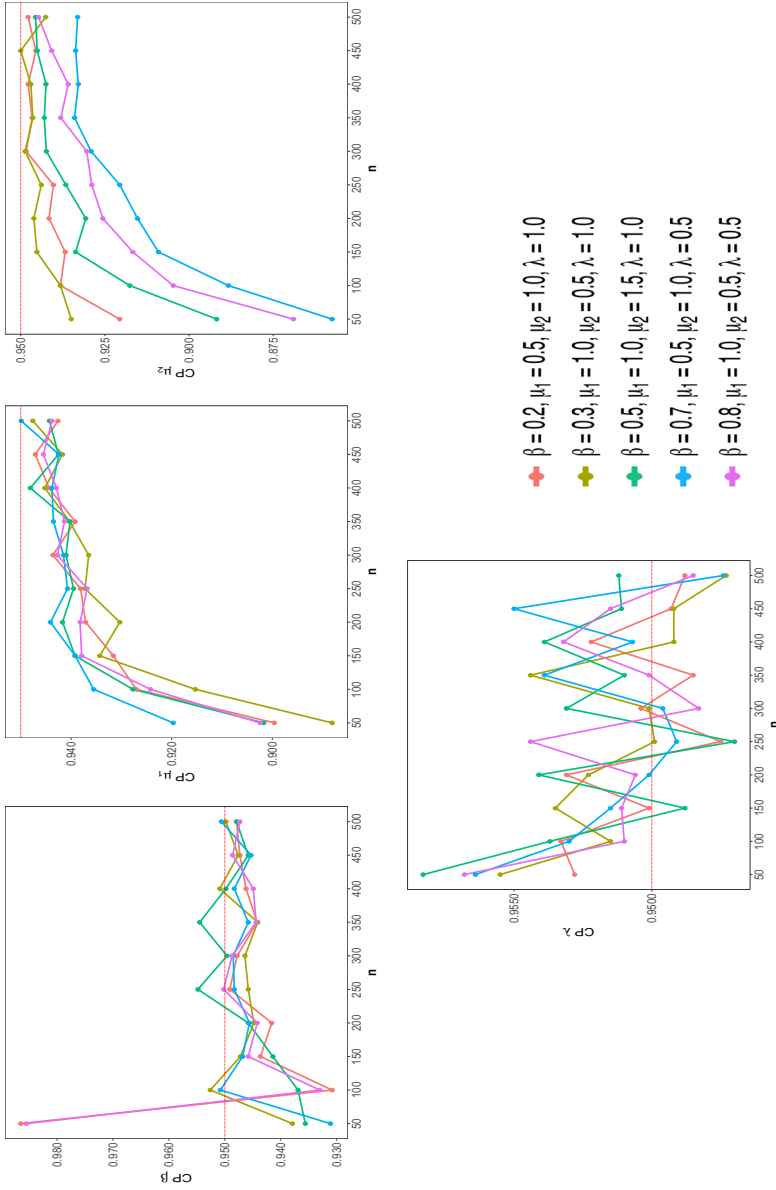


Fig. 14 Coverage probability of the parameters of DIG-2 distribution

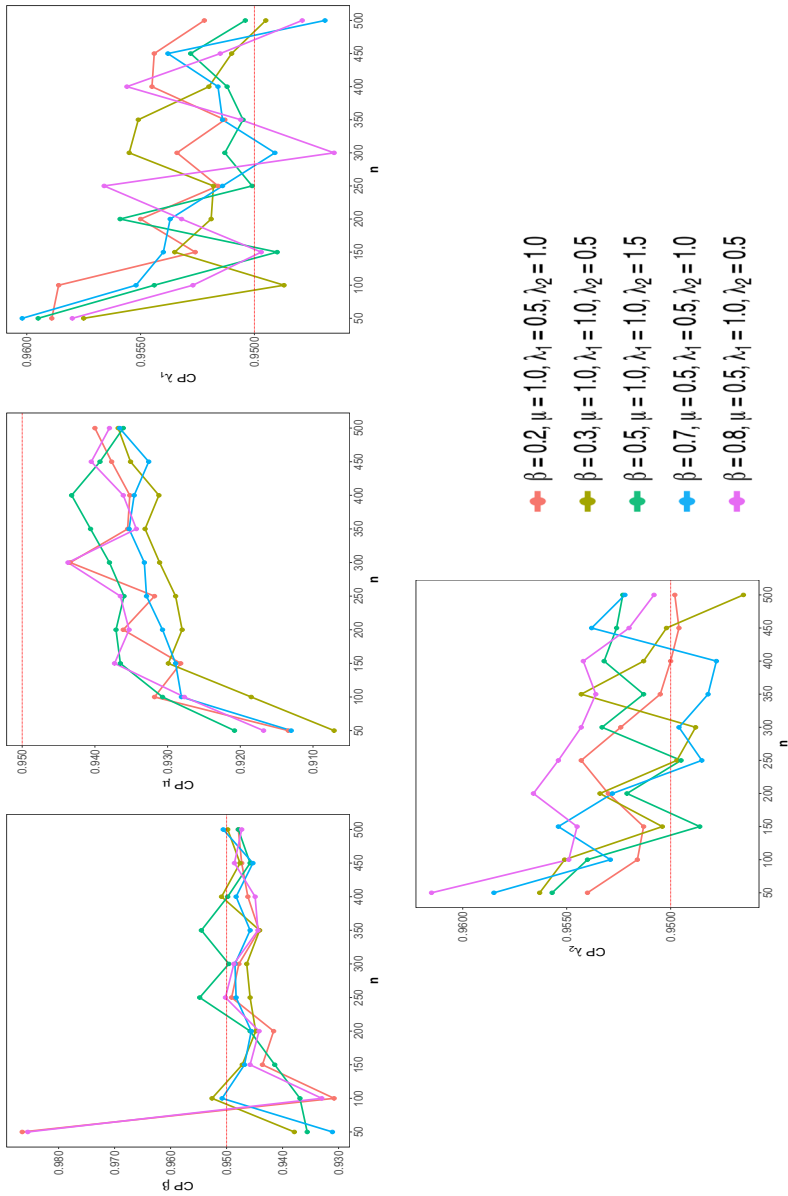


Fig. 15 Coverage probability of the parameters of DIG-3 distribution

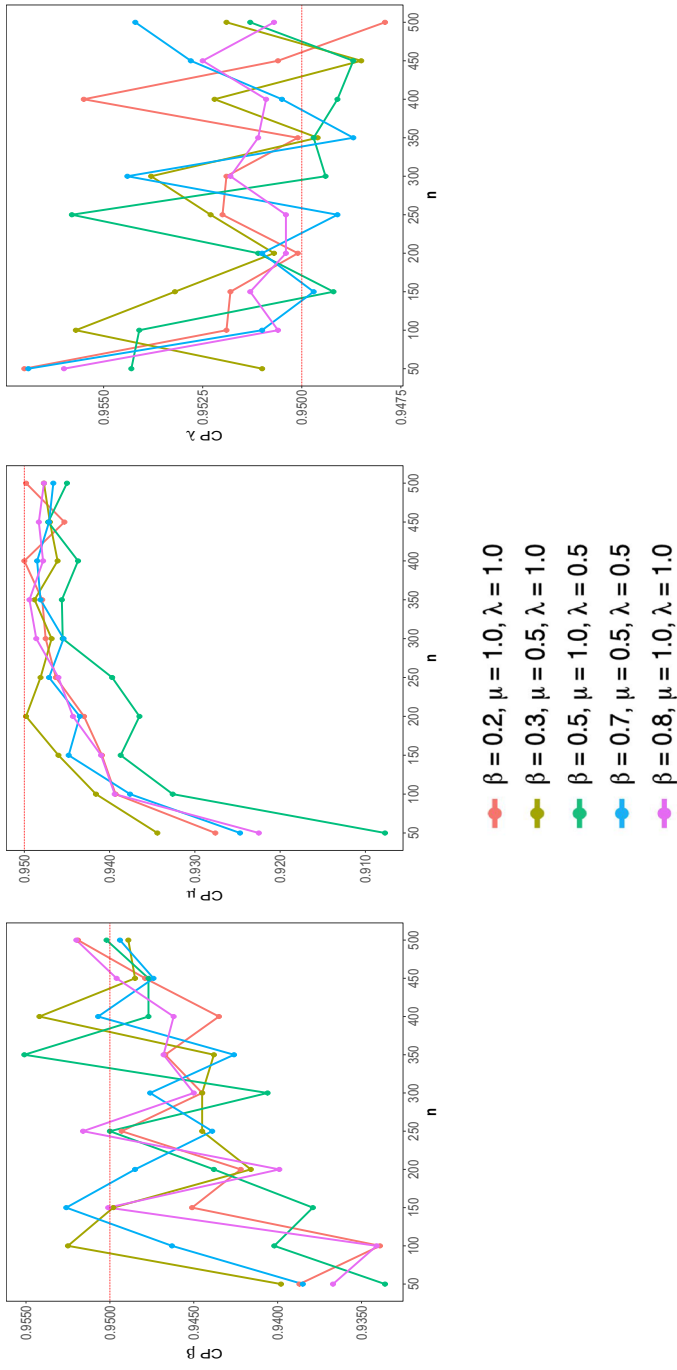


Fig. 16 Coverage probability of the parameters of DIG-4 distribution

Table 1 Summary of fitted DIG distributions

Model	MLE	S.E	$\ln \hat{L}$
DIG-1	$\hat{\beta}=0.763$	0.069	-94.656
	$\hat{\mu}_1=2.874$	0.273	
	$\hat{\lambda}_1=10.945$	2.874	
	$\hat{\mu}_2=7.800$	2.502	
	$\hat{\lambda}_2=8.423$	3.970	
DIG-2	$\hat{\beta}=0.763$	0.069	-94.779
	$\hat{\mu}_1=2.874$	0.283	
	$\hat{\mu}_2=7.800$	2.271	
	$\hat{\lambda}=10.220$	2.345	
DIG-3	$\hat{\beta}=0.763$	0.069	-100.491
	$\hat{\mu}=3.237$	0.322	
	$\hat{\lambda}_1=10.445$	2.743	
	$\hat{\lambda}_2=2.678$	1.262	
DIG-4	$\hat{\beta}=0.763$	0.069	-102.417
	$\hat{\mu}=4.041$	0.504	
	$\hat{\lambda}=6.837$	1.569	

Table 2 Goodness-of-fit tests of DIG distributions

Model	AD statistic	<i>p</i> value	CvM statistic	<i>p</i> value
DIG-1	0.135	0.999	0.019	0.998
DIG-2	0.159	0.998	0.024	0.993
DIG-3	1.271	0.241	0.111	0.533
DIG-4	2.170	0.075	0.284	0.150

- (ii) $H_0 : \mu_1 = \mu_2$ (DIG-3 model) versus $H_1 : \mu_1 \neq \mu_2$ (DIG-1 model)
- (iii) $H_0 : \mu_1 = \mu_2, \lambda_1 = \lambda_2$ (DIG-4 model) versus $H_1 : \mu_1 \neq \mu_2, \lambda_1 \neq \lambda_2$ (DIG-1 model)

Table 3 shows that DIG-2 model cannot be rejected for the given data.

We have seen above that LR test favour the DIG-2 model to be suitable for the given data. This conclusion is also supported by the Probability-Probability (P-P) plots presented in Fig. 17 and the Quantile-Quantile (Q-Q) plots presented in Fig. 18.

Table 3 Likelihood ratio tests for nested DIG distributions

Model	LR statistic	d.f	<i>p</i> value
DIG-2	0.246	1	0.6199
DIG-3	11.670	1	0.0006
DIG-4	15.522	2	0.0004

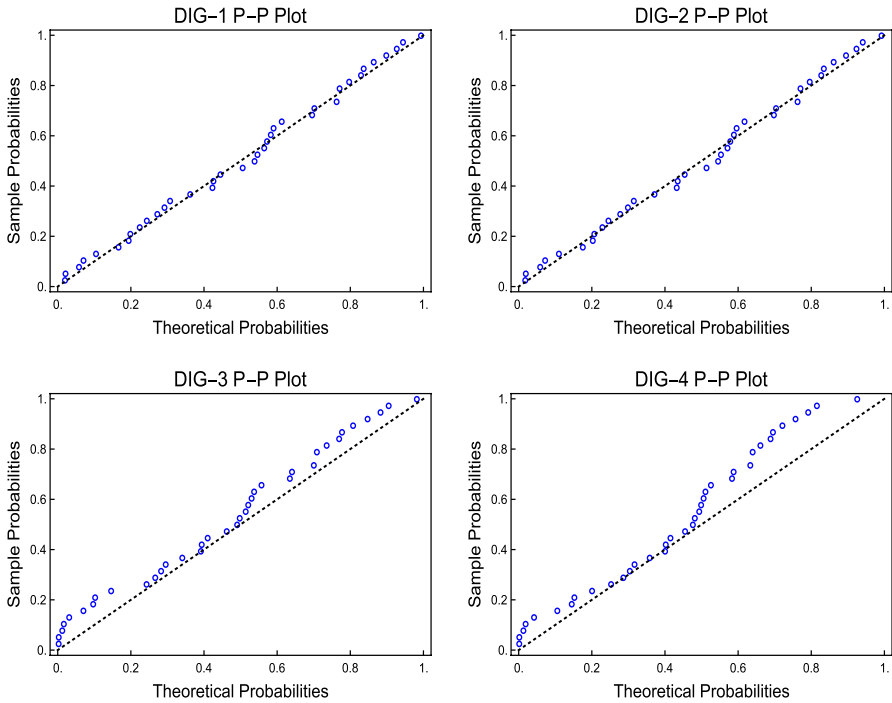


Fig. 17 P-P plots of fitted DIG distributions

6 Conclusion and Comments

Double inverse Gaussian distribution, presented here, has been formed by a procedure proposed by Aly (2018). This procedure is completely different from the procedures adopted in the literature. The unified approach, adopted here, is quite general and can be used to formulate double distributions for various classes of distributions. A natural extension of such double distributions is to include possible covariates to

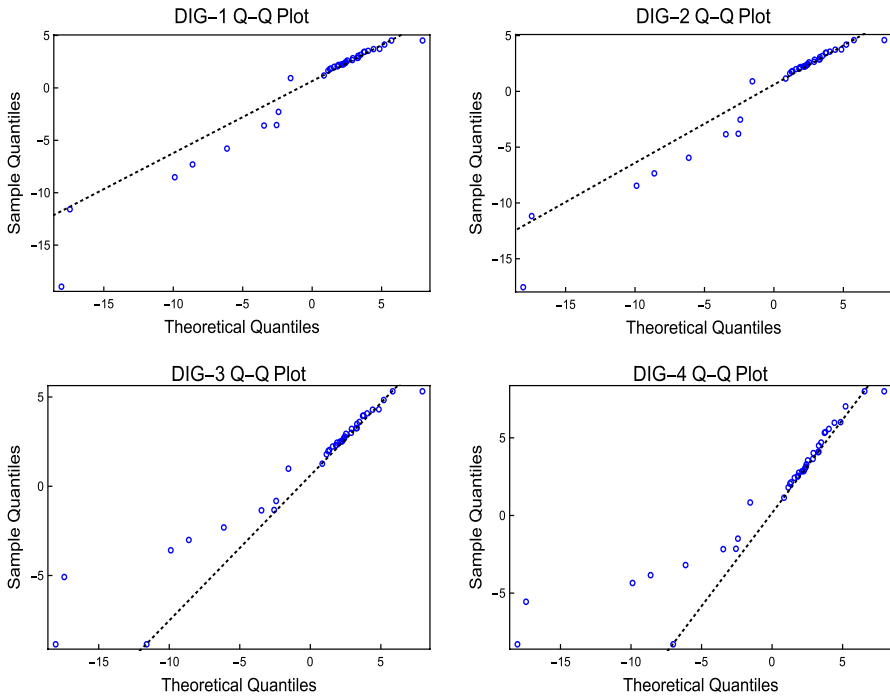


Fig. 18 Q-Q plots of fitted DIG distributions

allow more flexibility for modelling purposes. We hope that the model presented here will be found useful for data analysts.

Acknowledgements We are grateful to the Editor-in-Chief and anonymous referees for their valuable comments which greatly improved the presentation of the paper.

Funding No funding for this research.

Declaration

Conflicts of interest The authors declare no conflict of interest.

References

- Aly E (2018) A unified approach for developing Laplace-type distributions. *J Indian Soc Probab Stat* 19:245–269
- Armagan A, Dunson D, Lee J (2013) Generalized double Pareto shrinkage. *Statistica Sinica* 83:119–143
- Balakrishnan N, Kocherlakota S (1985) On the double Weibull distribution: order statistics and estimation. *Sankhya Indian J Stat Series B* 47:161–178
- Balakrishnan N, Kundu D (2019) Birnbaum-saunders distribution: a review of models, analysis, and applications (with discussion). *Appl Stoch Models Bus Ind* 35:4–125
- Bindu P, Sangita K (2015) Double Lomax distribution and its applications. *Statistica* 75:331–342

- de Zea Bermudez P, Kotz S (2010a) Parameter estimation of the generalized Pareto distribution-Part I. *J Stat Plan Inference* 140:1353–1373
- de Zea Bermudez P, Kotz S (2010b) Parameter estimation of the generalized Pareto distribution-Part II. *J Stat Plan Inference* 140:1374–1388
- Govindarajulu Z (1966) Best linear estimates under symmetric censoring of the parameters of a double exponential population. *J Am Stat Assoc* 61:248–258
- Gupta RC, Akman O (1995) On the reliability studies of a weighted inverse gaussian model. *J Stat Plan Inference* 48:69–83
- Gupta RC, Akman O (1996) Estimation of coefficient of variation in a weighted inverse gaussian model. *Appl Stoch Models Bus Industry* 12:255–263
- Gupta RC, Akman O (1997) Estimation of critical points in the mixture inverse gaussian model. *Stat Papers* 38:445–452
- Gupta RC, Akman O (1998) Statistical inference based on the length-biased data for the inverse-gaussian distribution. *Statistics* 31:325–337
- Gupta RC, Kundu D (2011) Weighted inverse gaussian - a versatile lifetime model. *J Appl Stat* 38:2695–2708
- Halvarsson D (2020) Maximum likelihood estimation of asymmetric double type II Pareto distributions. *J Stat Theor Pract* 22(14):1–34
- Ibrahim M, Mohammed W, Yousof HM (2020) Bayesian and classical estimation for the one parameter double Lindley model. *Pakistan J of Stat Oper Res* 16:409–420
- Kantam RRL, Narasimham VL (1991) Linear estimation in reflected gamma distribution. *Sankhya Indian J Stat Series B* 53:25–47
- Kumar SC, Jose R (2019) On double Lindley distribution and some of its properties. *Am J Math Manag Sci* 38:23–43
- Lad F, Sanfilippo G, Agrò G (2015) Extropy: complementary dual of entropy. *Stat Sci* 30:40–58
- Nadarajah S, Afuecheta E, Chan S (2013) A double generalized Pareto distribution. *Stat Probab Lett* 83:2656–2663
- Plucinska A (1965) On certain problems connected with a division of a normal population into parts. *Zastosow Mat* 8:117–125
- Plucinska A (1966) On a general form of the probability density function and its application to the investigation of the distribution of rheostat resistance. *Zastosow Mat* 9:9–19
- Plucinska A (1967) The reliability of a compound system under consideration of the system elements prices. *Zastosow Mat* 9:123–134
- Rao AVD, Narasimham VL (1989) Linear estimation in double Weibull distribution. *Sankhya Indian J Stat Series B* 51:24–64
- Sato S, Inoue J (1994) Inverse Gaussian distribution and its application. *Electron Commun Japan Part III Fundam Electron Sci* 77:32–42
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423
- Tsallis C (1988) Possible generalization of Boltzmann-Gibbs statistics. *J Stat Phys* 52:479–487

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.