REVSTAT – Statistical Journal
Volume 21, Number 4, October 2023, 535–556
https://doi.org/10.57805/revstat.v21i4.429

New Members of the Johnson Family of Probability Distributions: Properties and Application

Author: Piotr Sulewski
– Institute of Exact and Technical Sciences, Pomeranian University, Poland
piotr.sulewski@apsl.edu.pl

Received: June 2021 Revised: February 2022 Accepted: February 2022

Abstract:

- In this paper we introduce new non-mixed bimodal distributions belonging to the Johnson family of distributions (JFD) named SC and SD, where SC is a special case of SD. The SD is the only distribution among the JFD that can be both unimodal and bimodal distribution. Properties of the SD are methodically studied. The SC and SD are compared with the seven (except the normal distribution) members of the JFD for flexibility and applicability. In order to test for flexibility, a special measure called skewness-kurtosis-square is defined. The best dispersion of points with coordinates (skewness, kurtozis) occurs for the SD and for the well-known SU and SB. Two real datasets were used to test the applicability. EN turned out to be better than its competitors in terms of information criteria and results of three goodness-of-fit tests.

Keywords:

- *normal distribution; flexibility of distribution; departure from normality.*

AMS Subject Classification:

- 60E05, 65C20.

1. INTRODUCTION

Routinely statisticians start on analysis of data estimating PDF (commonly called histogram) or plotting empirical CDF on an appropriate probability paper. It depends on the sample size they possess. It may happen that a solid evidence emerges from the obtained figures suggesting bimodality of the population distribution as it is shown in Figure 1.

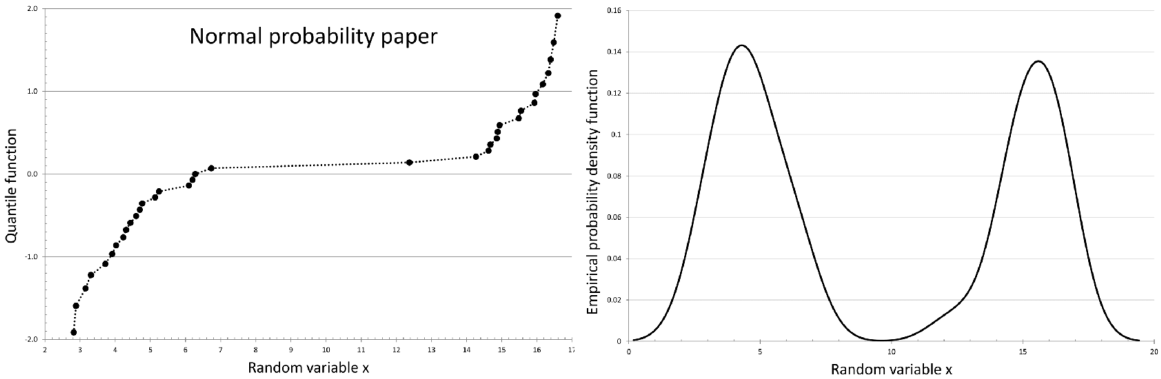

Figure 1: Bimodality of the population distribution.

In such a situation the statisticians as a rule employ the mixed (also called compound) theoretical distribution that has a general form:

$$f_m(x;\omega,\Theta_1,\sigma_1,\Theta_2,\sigma_2) = \omega f_1(x;\Theta_1,\sigma_1) + (1-\omega) f_2(x;\Theta_2,\sigma_2). \tag{1.1}$$

Let us denote such a distribution as mixed bimodal distribution (MBD). In (1.1) ω is the fraction parameter whereas Θ_1, σ_1 and Θ_2, σ_2 are pairs of location — scale or scale — shape parameters depending on sorts of distributions being mixed.

The MBD can be made bimodal and fitted to data of the sort we say about. However, employing MBD the statisticians unambiguously state that the population is nonhomogeneous. Wide applicability of the MBD comes from its clarity and interpretability of parameters. Nevertheless, it is hard to believe that non-homogeneity is a sole cause of distribution bimodality. It is hard to believe because many factors other than intentional or unintentional mixing sample items play a role in shaping population distribution. Thus, we see that a vital necessity arises to develop non-mixed bimodal distribution (nMBD) arises. What makes our task more difficult is that parameters of such distribution should be relatively clearly interpretable. In order that nMBD be a worthy challenger to MBD. However, we are fully aware that the parameters in question will never be so clearly interpretable as parameters of the MBD are, which is easy to explain. The MBD comes into being due to one factor, which causes mixing in particular proportion items belonging to two different subpopulations. In contrast nMBD comes into being due to many factors. And effects of "activity" of all these many factors have to be expressed also by means of only five parameters as estimators of parameters consume information.

Let us recall the Johnson family of probability distributions (JFD). All the distributions that belong to JFD have the following general form

$$F(x) = \Phi\left[c + \rho\varphi\left(\frac{x-a}{b}\right); 0, 1\right], \tag{1.2}$$

where $\varphi(x)$ can be any non-decreasing function of x, $\Phi(x; u, v)$ is CDF of $N(u, v)$. Literature related to JFD is numerous (see e.g. [14], [15], [3]).

Before defining the new members of the JFD, it is worth taking a look at the distributions that belong to this family.

The normal (N) distribution with location parameter $a \in R$ and scale parameter $b > 0$ is defined as

$$F_N(x; a, b) = \Phi\left[\varphi\left(\frac{x-a}{b}\right); 0, 1\right]. \tag{1.3}$$

We obtain (1.3) from (1.2) by considering $c = 0$, $\delta = 1$, $\varphi(y) = y$ and $y = (x-a)/b$.

The Birnbaum–Saunders (BS) distribution with location parameter $a \in R$, scale parameter $b > 0$ and shape parameter $\alpha > 0$, is defined as [7]

$$F_{BS}(x; \alpha, a, b) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x-a}{b}} - \sqrt{\frac{b}{x-a}}\right); 0, 1\right] \quad (x > a). \tag{1.4}$$

We obtain (1.4) from (1.2) by considering $c = 0$, $\delta = 1/\alpha$, $\varphi(y) = \sqrt{y} - \sqrt{1/y}$ and $y = (x-a)/b$.

The generalization of the Birnbaum–Saunders (GBS) distribution with location parameter $a \in R$, scale parameter $b > 0$ and shape parameters $\alpha > 0$, $\beta > 0$, is defined as [18]

$$F_{GBS}(x; \alpha, a, b, \beta) = \Phi\left[\frac{1}{\alpha}\left(\left(\frac{x-a}{\beta}\right)^{\beta} - \left(\frac{\beta}{x-a}\right)^{\beta}\right); 0, 1\right] \quad (x > a). \tag{1.5}$$

The BS distribution is a special case of the GBS distribution for $\beta = 0.5$. We obtain (1.5) from (1.2) by considering $c = 0$, $\delta = 1/\alpha$, $\varphi(y) = y^{\beta} - y^{-\beta}$ and $y = (x-a)/b$.

The Four-Parameter BS (FBS) distribution with location parameter $a \in R$, scale parameter $b > 0$, shape parameter $\delta > 0$ and non-centrality parameter $c \in R$, is given by [2]

$$F_{FBS}(x; c, \delta, a, b) = \Phi\left[c + \delta\left(\sqrt{\frac{x-a}{b}} - \sqrt{\frac{b}{x-a}}\right); 0, 1\right] \quad (x > a). \tag{1.6}$$

Formula (1.4) is a special case of (1.6) for $c = 0$, $\delta = 1/\alpha$. We obtain (1.6) from (1.2) by considering $\varphi(y) = \sqrt{y} - \sqrt{1/y}$ and $y = (x-a)/b$.

The sinh-normal (SN) distribution with the location parameter $a \in R$, the scale parameter $b > 0$ and the shape parameter $\alpha > 0$, is given by [19]

$$F_{SN}(x; \alpha, a, b) = \Phi\left[\frac{2}{\alpha}\sinh\left(\frac{x-a}{b}\right); 0, 1\right]. \tag{1.7}$$

This distribution is symmetric about the location parameter $a \in R$. We obtain (1.7) from (1.2) by considering $c = 0$, $\delta = 2/\alpha$, $\varphi(y) = \sinh(y)$ and $y = (x-a)/b$.

The lognormal or SL distribution with location parameter $a \in R$, scale parameter $b > 0$ and shape parameters $c_1 \in R$, $\delta > 0$, is defined as [14]

$$F_{SL}(x; c_1, \delta, a, b) = \Phi\left[c_1 + \delta \ln\left(\frac{x-a}{b}\right); 0, 1\right] \quad (x > a). \tag{1.8}$$

Formula (1.8) can be written using three parameters, namely:

$$F_{SL}(x; c_1, \delta, a, b) = \Phi[c + \delta \ln(x-a); 0, 1] \quad (x > a), \tag{1.9}$$

where $c = c_1 - \delta \ln(b)$, $c \in R$. We obtain (1.9) from (1.2) by considering $b = 1$, $\varphi(y) = \ln(y)$ and $y = (x - a)$. Please notice that the lognormal distribution with the CDF [11] $\breve{F}_{SL}(x; e_1, e_2) = \Phi\left[\frac{\ln(x)-e_1}{e_2}; 0, 1\right]$ $(x > 0)$ widely used in practice can be treated as a special case of (1.8) when $a = 0$, $\delta = 1/e_2$ and $c = -e_1/e_2$.

The SB distribution with the location parameter $a \in R$, the scale parameter $b > 0$ and the shape parameters $c \in R$, $\delta > 0$, is defined as [14]

$$F_{SB}(x; c, \delta, a, b) = \Phi\left[c + \delta \ln\left(\frac{x-a}{b+a-x}\right); 0, 1\right] \quad (a < x < a+b). \tag{1.10}$$

We obtain (1.10) from (1.2) by considering $\varphi(y) = \ln(y) - \ln(1-y)$ and $y = (x-a)/b$. Let $a = 0$, $b = 1$, $\delta = 1/e_2$ and $c = -e_1/e_2$, then we obtain the special case of (1.10) widely used in practice defined as

$$\breve{F}_{SB}(x; e_1, e_2) = \Phi\left[\frac{\ln\left(\frac{x}{1-x}\right) - e_1}{e_2}; 0, 1\right].$$

The SU distribution with the location parameter $a \in R$, the scale parameter $b > 0$ and the shape parameters $c \in R$, $\delta > 0$, is defined as [14]

$$F_{SU}(x; c, \delta, a, b) = \Phi\left[c + \delta\, \mathrm{asinh}\left(\frac{x-a}{b}\right); 0, 1\right]. \tag{1.11}$$

We obtain (1.11) from (1.2) by considering $\varphi(y) = \mathrm{asinh}(y)$ and $y = (x-a)/b$.

This paper introduces two new members of the JFD, namely SC and SD. In the SU distribution Johnson employed $\mathrm{asinh}(x) = \ln\left(x + \sqrt{1+x^2}\right)$. In the SC and SD distributions we will employ $\sinh(x) = \frac{\exp(x)-\exp(-x)}{2}$.

The SC distribution with location parameter $a \in R$, the scale parameter $b > 0$ and shape parameter $c \in R$, is defined as

$$F_{SC}(x; c, a, b) = \Phi\left[c + 2\sinh\left(\frac{x-a}{b}\right); 0, 1\right]. \tag{1.12}$$

Please notice that ρ parameter appearing in (1.2) has been in (1.12) replaced with a constant equal to 2. This constant compensates denominator in definition of the $\sinh(x)$ function. For $c = 0$ in (1.12) and $\alpha = 1$ in (1.7), the SN distribution is equivalent to the SC distribution. We obtain (1.12) from (1.2) by considering $\delta = 2$, $\varphi(y) = \sinh(y)$ and $y = (x-a)/b$.

The SD distribution with multipurpose parameters $a_1, a_2 \in R$, $b_1, b_2 > 0$ and semi-fraction parameter $c > 0$ (see Figure 2), is defined as

$$F_{SD}(x; c, a_1, b_1, a_2, b_2) = \Phi\left[c - \exp\left(\frac{a_1 - x}{b_1}\right) + \exp\left(\frac{x - a_2}{b_2}\right); 0, 1\right]. \tag{1.13}$$

The SD distribution is obtained from SC by adding the second exponential function. The CDF $F_{SD}(x; c, a_1, b_1, a_1, b_1)$ is equal to the CDF $F_{SC}(x; c, a_1, b_1)$, so the SC is a special case of the SD.

Although SD cannot be acknowledged as a special case of (1.2), it seems reasonable to treat SD as a member of the JFD, which can be justified by appearing as a generalization or extension of an element of the SN distribution family obviously belonging to the JFD family. The mentioned element is the SC distribution. The F_{SD} involves two exponential components that can be independently movable on the x axis. Owing to this we are able to obtain bimodal distribution provided we locate the components sufficiently far from one another on the x axis as it is exemplified in Figure 2 (left). This figure shows examples of CDFs of the SD distribution plotted on the Normal probability paper. The reader is prompted to compare Figure 2 (left) with Figure 1 (left). Figure 2 (right) shows examples of PDFs of the SD distribution with exemplifying a role of c parameter. No doubt, c parameter can be called the semi-fraction parameter.

Let

$$P(x; c, a_1, b_1, a_2, b_2) = c - \exp\left(\frac{a_1 - x}{b_1}\right) + \exp\left(\frac{x - a_2}{b_2}\right).$$

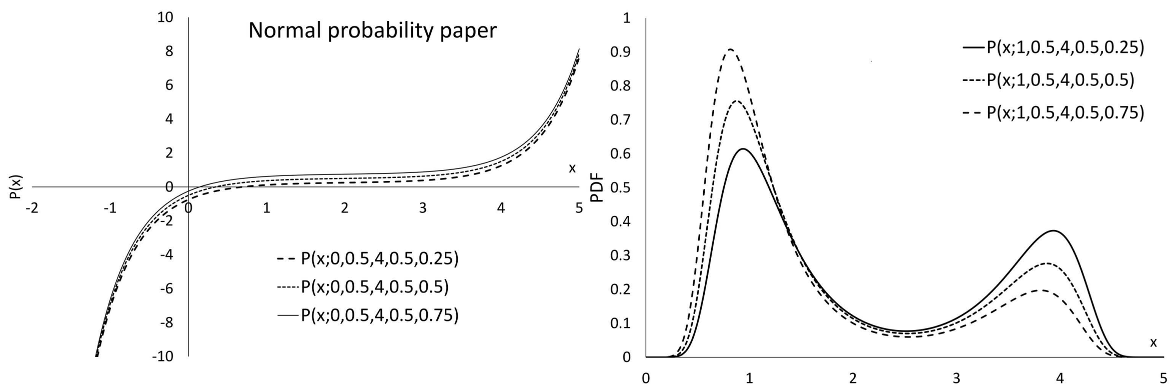

Figure 2: Bimodality in the SD distribution.

If we subject a random variable to a linear transformation, the skewness and kurtosis retain their values. This fact was also confirmed by a simulation study. To simplify the study of the skewness and kurtosis of the SD distribution, let us standardize a random variable x: $z = \frac{x - a_1}{b_1} \Rightarrow x = b_1 z + a_1$. As a result of simple transformation the CDF (1.13) has the form

$$F_{SD}(x; c, a_0, b_0) = \Phi\left[c - \exp(-z) + \exp\left(\frac{x - a_0}{b_0}\right); 0, 1\right],$$

where $a_0 = \frac{a_2 - a_1}{b_1}$, $b_0 = \frac{b_2}{b_1}$.

Let (γ_1, γ_2) be coordinate of a point described by skewness and kurtosis, respectively. The normal distribution (ND) is characterized by only one point $(0, 3)$, obviously. For every distribution from the JFD (except the ND), values of (γ_1, γ_2) are calculated for 10^4 randomly determined values of parameters influencing skewness and kurtosis in the Malakhov area (MA) $10 \geq \gamma_2 \geq \gamma_1^2 + 1$ (Table 1). In the MA $\gamma_2 \in [1, 10]$, then $\gamma_1 \in [-3, 3]$ (see Figure 3). The parameter value ranges are selected to maximize MA filling according to the SKS measure (1.14). To make the calculations more reliable (without the so called outliers) the normalization conditions were checked.

Table 1: Ranges of parameter values influencing skewness and kurtosis as well as skewness and kurtosis values for the JDF in the Malakhov area $10 \geq \gamma_2 \geq \gamma_1^2 + 1$.

JFD	Parameter ranges	Skewness range	Kurtosis range
BS	$\alpha \in (0, 0.46)$	$\gamma_1 \in (0, 2.22)$	$\gamma_2 \in (3, 5.92)$
GBS	$\alpha \in (0, 7)$, $\beta \in (0, 1.75)$	$\gamma_1 \in (0.07, 2.38)$	$\gamma_2 \in (2.03, 10)$
FBS	$c \in (-6.25, 6.25)$, $\delta \in (0, 2.75)$	$\gamma_1 \in (0.44, 2.23)$	$\gamma_2 \in (3.27, 10)$
SN	$\alpha \in (0.1, 180.4)$	$\gamma_1 = 0$	$\gamma_2 \in (1.15, 3)$
SL	$\delta \in (1.88, 100)$	$\gamma_1 \in (0.03, 1.89)$	$\gamma_2 \in (3, 9.98)$
SB	$c \in (-3.35, 3.39)$, $\delta \in (0.1, 1.2)$	$\gamma_1 \in (-2.84, 2.91)$	$\gamma_2 \in (1.13, 10)$
SU	$c \in (-2.05, 2.05)$, $\delta \in (1.31, 1.9)$	$\gamma_1 \in (-1.8, 1.79)$	$\gamma_2 \in (4.76, 10)$
SC	$c \in (-89.94, 89.97)$	$\gamma_1 \in (-0.69, 0.69)$	$\gamma_2 \in (2.52, 3.90)$
SD	$c \in (-4.1, 4.1)$, $a_0 \in (-4.3, 4.3)$, $b_0 \in (0.1, 0.9)$	$\gamma_1 \in (-2.79, 2.46)$	$\gamma_2 \in (1.26, 10)$

Figure 3 presents sets of points (γ_1, γ_2) and the MP $\gamma_2 = \gamma_1^2 + 1$ in the MA $10 \geq \gamma_2 \geq \gamma_1^2 + 1$ related to the SB, SU, SC, SD distributions. The SD and SU distributions are the best filling the MA. The SD distribution has common areas of skewness and kurtosis with the SB and SU distributions. Sets of points (γ_1, γ_2) and the MP $\gamma_2 = \gamma_1^2 + 1$ in the MA $10 \geq \gamma_2 \geq \gamma_1^2 + 1$ related to the BS, GBS, FBS, SL distributions are presented in the supplementary material.

In addition to visual assessment, the skewness-kurtosis-square (SKS) measure [22] is used to compare the flexibility of distributions. Colored circles of diameter and coordinates of their centers determined by skewness γ_1 and kurtosis γ_2 are placed within the MA that is described by inequality $\gamma_2 \geq \gamma_1^2 + 1$ [17]. Then colored area fraction is calculated. Squares of sides equal to η seem a reasonable alternative to circles since they simplify calculation of the total colored area. Obviously, when some squares overlap, only one is taken into account. The SKS measure is given by [22]

$$SKS = \frac{SI}{ST}, \tag{1.14}$$

where ST denotes a total number of squares within the MA, SI — a number of squares to which the point (γ_1, γ_2) has fallen. The SKS measure takes values in $[0, 1]$. The maximum value denotes a perfect dispersal of points (γ_1, γ_2) in the MA. The R codes for calculating the SKS measure are presented in the supplementary material.

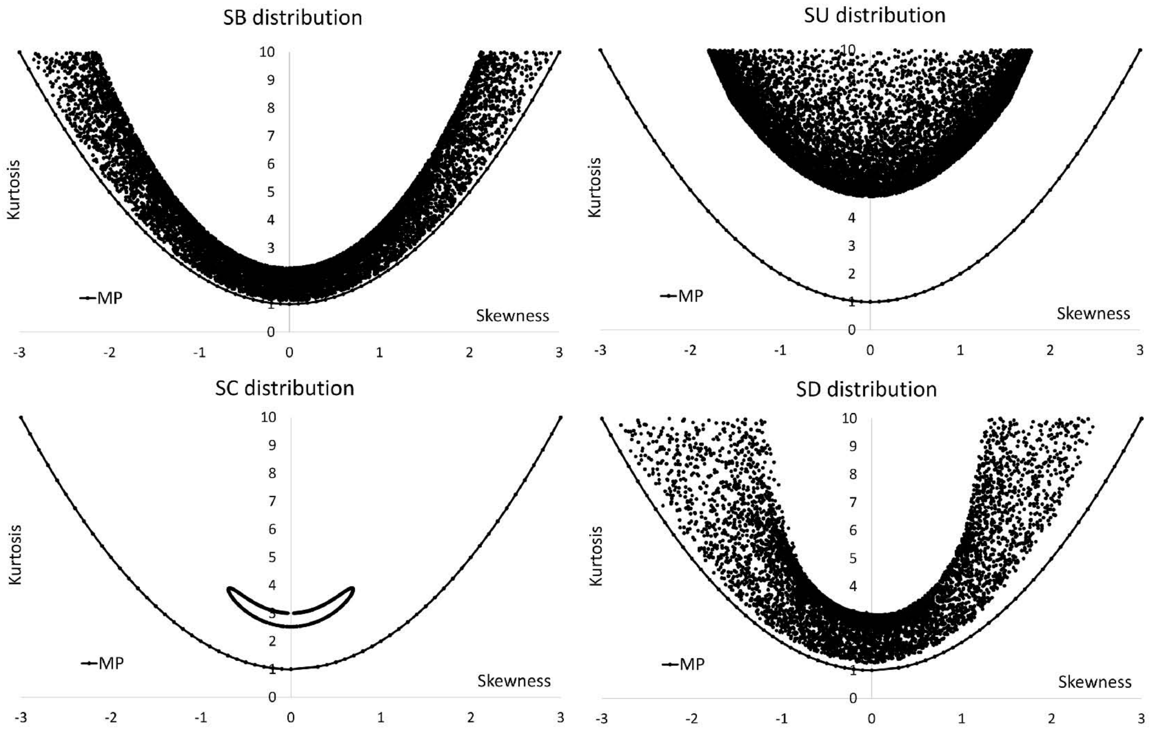

Figure 3: Skewness and kurtosis for the SB, SU, SC, SD distributions.

Table 2 presents values of SKS measures (1.14) obtained for square side $\eta = 0.05, 0.1, 0.15, 0.20$. The best dispersion of points (γ_1, γ_2), taking into account the accuracy expressed by η, occurs for the SD, SU and SB distribution (see bold).

Table 2: SKS measure values for JFD in the MA $10 \geq \gamma_2 \geq \gamma_1^2 + 1$ for square side η.

JFD	$\eta = 0.05$	$\eta = 0.1$	$\eta = 0.15$	$\eta = 0.2$
BS	0.0070	0.0140	0.0203	0.0271
GBS	0.0778	0.1051	0.1277	0.1333
FBS	0.0561	0.0694	0.0794	0.0906
SL	0.0111	0.0237	0.0346	0.0458
SB	**0.2410**	0.3401	0.3813	0.4052
SU	0.2375	0.3433	0.3693	0.3740
SC	0.0092	0.0175	0.0274	0.0323
SD	0.2185	**0.4102**	**0.4994**	**0.5469**

New distributions, modelled on the SL, SB, SU distributions, was named as SC and SD distributions. The SC is a special cases of the SD, so the remainder of the paper is devoted to the SD distribution. The lognormal distribution is defined with the log function and the SD distribution is defined with the exp function, therefore the SD distribution is also called the expnormal (EN) distribution.

This paper is organized as follows. Section 2 presents properties of the SD distribution. The unknown parameters are estimated in Section 3 and entropies are calculated in Section 4. Examples are presented in Section 5. Section 6 deals with conclusions. Due to the size of the paper, the selected figures and tables as well as the main R codes have been transferred to the supplementary material.

2. MAIN PROPERTIES OF INTRODUCED DISTRIBUTION

2.1. Distribution and density function

Definition 2.1. The distribution of the random variable X with PDF given by

$$f(x;a_1,b_1,a_2,b_2,c)=\left(\frac{1}{b_1}e^{-z_1(x)}+\frac{1}{b_2}e^{z_2(x)}\right)\phi[c-\exp(-z_1(x))+\exp(z_2(x));0,1], \tag{2.1}$$

where $\phi(x;u,v)$ is PDF of $N(u,v)$, $z_1(x)=\frac{x-a_1}{b_1}$ and $z_2(x)=\frac{x-a_2}{b_2}$, is called the expnormal (EN) distribution. In (2.1) $a_1,a_2\in R$ are position parameters, $b_1,b_2>0$ are scale parameters and $c\in R$ is the semi-fraction parameter (see Figure 2). For these parameter values, the main argument of ϕ in (2.1) is an increasing function, hence

$$\int_{-\infty}^{+\infty} f(x;a_1,b_1,a_2,b_2,c)=1.$$

PDF of the EN distribution is calculated using the R function dEN (see supplementary material).

If $a_1=a_2,\ b_1=b_2,\ c=0$, then $EN(a_1,b_1,a_2,b_2,c)$ is very similar to the $N\left(a_1,\frac{b_1}{2}\right)$. According to the similarity measure between two distributions defined in [23], we have for $a_1\in R$, $b_1>0$.

$$\int_{-\infty}^{+\infty} \min\left[f(x;a_1,b_1.a_1,b_1,0),\phi\left(x;a_1,\frac{b_1}{2}\right)\right]=0.966. \tag{2.2}$$

Thus the $EN(0,2,0,2,0)$ is similar to the $N(0,1)$ in 96.6%. The distribution with multipurpose parameter $a_1,\ b_1,\ a_2,\ b_2=b_1$ is symmetrical for $c=0$ (see Table 4 and Figure 4, series D1,D2). If $X\sim EN(a_1,\ b_1,\ a_2,\ b_2=b_1,\ c=0)$ then $E(X)=\frac{a_1+a_2}{2}$. In this case the modes are at the same height. The mean value formula is also confirmed by numerical methods. The $EN(a_1,b_1,a_2,b_2,\ c>0)$ is positively skewed (Figure 4, series A1, A2, E1, E2) and the $EN(a_1,b_1,a_2,b_2,\ c\le 0)$ is negatively skewed (Figure 4, series B1, B2, F1, F2). The EN distribution can be unimodal (Figure 4, series A1, B1, D1, E1, F1) and bimodal (Figure 4, series A2, B2, D2, E2, F2). See Table 4 for more information.

Table 3 presents the division of distributions by their skewness and excess kurtosis [22]. The ND obviously does not belong to this family. Selecting appropriate parameter values of the EN distribution, we can obtain skewness and excess kurtosis values belonging to the analyzed groups A1–B2 and D1–F2 (Table 4).

Table 3: Groups of distributions according to their skewness and excess kurtosis [23]. Denote: * unimodal distribution, ** bimodal distribution.

Group	**Skewness**	**Ex. kurtosis**	**Group**	**Skewness**	**Ex. kurtosis**
A1*	positive	positive	D1*	zero	negative
A2**	positive	positive	D2**	zero	negative
B1*	negative	positive	E1*	positive	negative
B2**	negative	positive	E2**	positive	negative
C1*	zero	positive	F1*	negative	negative
C2**	zero	positive	F2**	negative	negative

Table 4: The $EN(a_1, b_1, a_2, b_1, 0)$ distribution with parameter values for groups A1–B2 and D1–F2.

a_1	b_1	a_2	b_2	**c**	**Skewness**	**Ex. kurtosis**	**Group**
0	1	1	1.25	1	0.740	0.268	A1
−1	1	3	1	1	1.239	0.608	A2
1	2	0	1	0	−0.527	0.151	B1
−4	0.5	1	1	−1	−1.298	0.334	B2
0	2	0	2	0	0	−0.479	D1
0	0.5	1	0.5	0	0	−1.024	D2
0	1	1	1	1	0.584	−0.13	E1
−1	1	3	1	0.5	0.601	−0.961	E2
0	1	1	1	−1	−0.584	−0.13	F1
−1	1	3	1	−0.5	−0.601	−0.961	F2

Figure 4 plots the PDF of the $EN(a_1, b_1, a_2, b_2, c)$ for groups of parameters presented in Table 4.

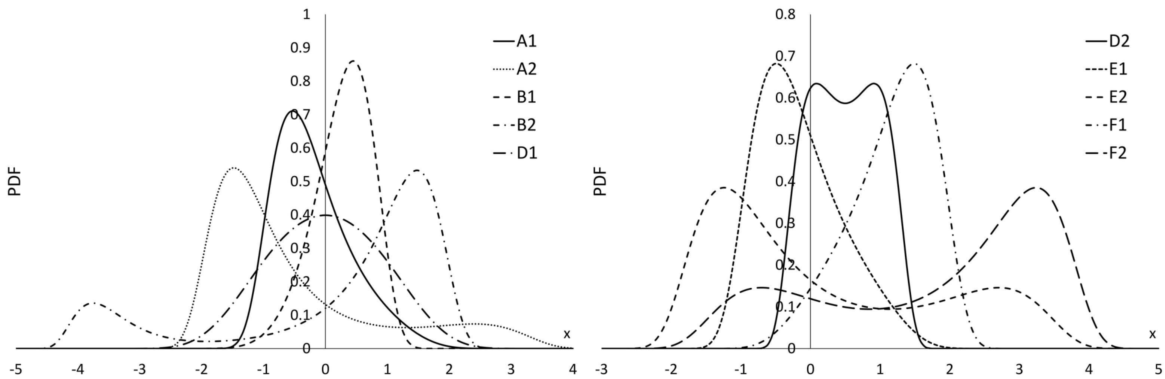

Figure 4: PDF of the $EN(a_1, b_1, a_2, b_1, 0)$ for groups from Table 4.

Theorem 2.1. *Let $X \sim EN(a_1, b_1, a_2, b_2, c)$, then the CDF of X is given by*

$$F(x; a_1, b_1, a_2, b_2, c) = \Phi\left[c - \exp\left(-\frac{x-a_1}{b_1}\right) + \exp\left(\frac{x-a_2}{b_2}\right); 0, 1\right]. \tag{2.3}$$

Proof: Obtaining (2.3) based on (2.1) is trivial. □

CDF of the EN distribution is calculated using the R function pEN (see supplementary material).

Figure 5 (left) plots the CDF of the $EN(a_1, b_1, a_2, b_2, c)$ for groups A1, A2, B1, B2. The CDF of the $EN(a_1, b_1 > 0, a_2, b_2 > 0, c)$ on the normal Q-Q plot is monotonically increasing curve (Figure 5, right).

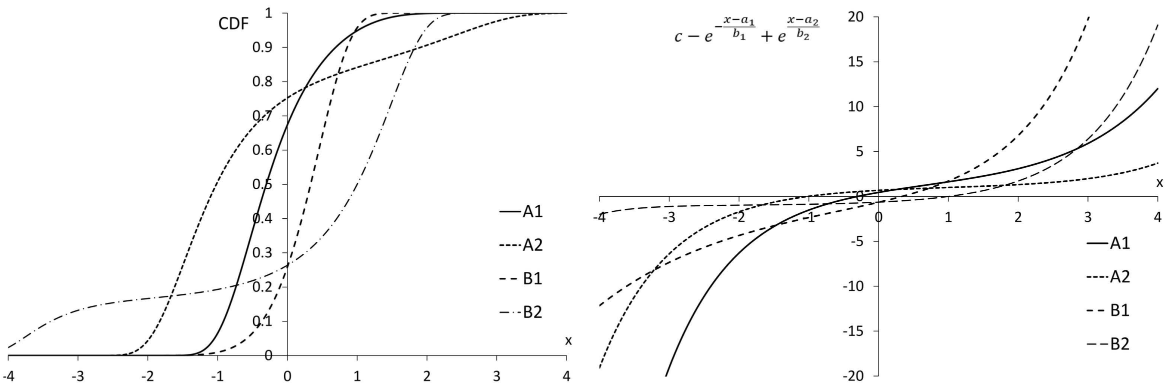

Figure 5: CDF of the $EN(a_1, b_1, a_2, b_1, 0)$ (left) and the normal Q-Q plot (right).

Theorem 2.2. *The $EN(a_1, b_1, a_2, b_2, c)$ with the PDF given by (14) is identifiable in a parameter space $v = (a_1, b_1, a_2, b_2, c)$.*

Proof: Let $v_1 = (a_{11}, b_{11}, a_{21}, b_{21}, c_1)$ and $v_2 = (a_{12}, b_{12}, a_{22}, b_{22}, c_2)$. Let us suppose that $f_{v_1}(x) = f_{v_2}(x)$ for all x. This condition based on (2.3) implies that

$$\Phi\left[c_1 - \exp\left(-\frac{x - a_{11}}{b_{11}}\right) + \exp\left(\frac{x - a_{21}}{b_{21}}\right); 0, 1\right] = \\ = \Phi\left[c_2 - \exp\left(-\frac{x - a_{12}}{b_{12}}\right) + \exp\left(\frac{x - a_{22}}{b_{22}}\right); 0, 1\right].$$

The function Φ is an increasing function which implies that

$$c_1 - \exp\left(-\frac{x - a_{11}}{b_{11}}\right) + \exp\left(\frac{x - a_{21}}{b_{21}}\right) = c_2 - \exp\left(-\frac{x - a_{12}}{b_{12}}\right) + \exp\left(\frac{x - a_{22}}{b_{22}}\right)$$

or

$$c_1 - c_2 + \exp\left(-\frac{x - a_{12}}{b_{12}}\right) - \exp\left(-\frac{x - a_{11}}{b_{11}}\right) + \exp\left(\frac{x - a_{21}}{b_{21}}\right) - \exp\left(\frac{x - a_{22}}{b_{22}}\right) = 0.$$

As a result of simple transformation $a_{11} = a_{12}$, $b_{11} = b_{12}$, $a_{21} = a_{22}$, $b_{21} = b_{22}$, $c_1 = c_2$. □

2.2. Hazard function

Proposition 2.1. *Let $X \sim EN(a_1, b_1, a_2, b_2, c)$. The hazard function associated with the EN distribution is*

$$h(x) = \frac{\left(\frac{1}{b_1} e^{-\frac{x-a_1}{b_1}} + \frac{1}{b_2} e^{\frac{x-a_2}{b_2}}\right) \phi\left[c - \exp\left(-\frac{x-a_1}{b_1}\right) + \exp\left(\frac{x-a_2}{b_2}\right); 0, 1\right]}{1 - \Phi\left[c - \exp\left(-\frac{x-a_1}{b_1}\right) + \exp\left(\frac{x-a_2}{b_2}\right); 0, 1\right]}. \tag{2.4}$$

The limits of the EN hazard function as $x \to -\infty$ and $x \to \infty$ are respectively 0 and ∞ (Figure 6).

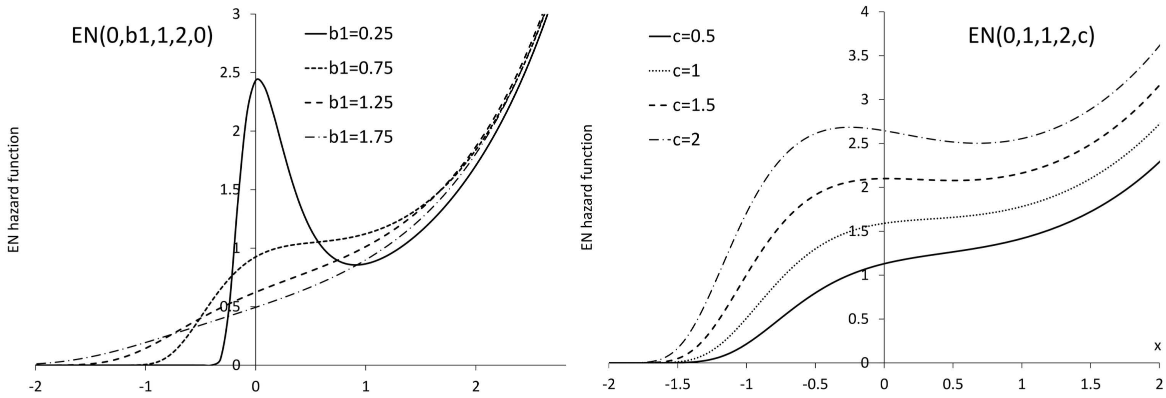

Figure 6: The EN hazard function for various values of parameters.

2.3. Quantiles

Proposition 2.2. *Let $X \sim EN(a_1, b_1, a_2, b_2, c)$. The p-th $(0 < p < 1)$ quantiles are the solution of the following equation*

$$c - \exp\left(-\frac{x_p - a_1}{b_1}\right) + \exp\left(\frac{x_p - a_2}{b_2}\right) - \Phi^{-1}(p) = 0.$$

The value of x_p is obtained by the numerical method, e.g. using the R software. Quantile function of the EN distribution is calculated using the R function qEN (see supplementary material).

2.4. Moments and moment generating function

Proposition 2.3. *Let $X \sim EN(a_1, b_1, a_2, b_2, c)$. The k-th, $k \in Z$ non-central moments from (14) are given by*

$$\alpha_k = \int_{-\infty}^{+\infty} x^k \left(\frac{1}{b_1} e^{-z_1} + \frac{1}{b_2} e^{z_2}\right) \phi[c - \exp(-z_1) + \exp(z_2); 0, 1], \tag{2.5}$$

where $z_1 = \frac{x-a_1}{b_1}$ and $z_2 = \frac{x-a_2}{b_2}$, $\phi(x; a, b)$ is PDF of $N(a, b)$

Thus the variance μ_2, skewness γ_1 and kurtosis γ_2 of the EN distribution are defined as

$$\mu_2 = \alpha_2 - \alpha_1^2, \quad \gamma_1 = \frac{\alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3}{\mu_2^{1.5}}, \quad \gamma_2 = \frac{\alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1^2\alpha_2 - 3\alpha_1^4}{\mu_2^2}.$$

Table 5 provides the mode x_{mod}, mean α_1, variance μ_2, skewness γ_1 and kurtosis γ_2 of the EN distribution for various parameter combinations.

Table 5: Mode, mean, variance, skewness and kurtosis of the $EN(a_1, b_1, a_2, b_2, c)$.

a_1	b_1	a_2	b_2	c	x_{mod}	α_1	μ_2	γ_1	γ_2
0	2	0	2	0	0	0	0.837	0	2.521
1					0.5	0.5	0.538	0	2.643
2					1	1	0.34	0	2.745
3					1.5	1.5	0.212	0	2.826
0	0.5	1	1	0	−0.005	0.444	0.334	0.194	2.532
	1				0.5	0.5	0.474	0	2.245
	1.5				0.95	0.52	0.622	−0.216	2.455
	2				1.077	0.523	0.775	−0.469	2.812
−2	1	−2	1	0	−2	−2	0.209	0	2.521
		−1			−1.5	−1.5	0.474	0	2.245
		1			−1.946,0.946	−0.5	1.791	0	1.751
		2			1.981	0	3.009	0	1.578
0	0.5	0.5	0.25	1	−0.215	−0.025	0.096	0.019	2.156
			0.5		−0.245	−0.08	0.096	0.056	2.87
			0.75		−0.281	−0.114	0.101	0.089	3.686
			1		−0.308	−0.137	0.107	0.12	4.583
0	1	1	2	0.5	−0.377	0.072	0.597	0.534	3.594
				1	−0.615	−0.274	0.427	0.48	4.583
				1.5	−0.815	−0.561	0.293	0.367	5.503
				2	−0.988	−0.798	0.199	0.252	6.006
−2	2	2	1	−1	2.508	1.721	1.633	−2.662	5.457
				0	−1.833,1.99	0.226	3.357	−1.557	2.014
				1	−2.949,1.273	−1.604	3.253	1.832	2.271
				2	−3.762	−3.097	1.642	2.352	5.387

Table 5 shows that the PDF of EN distribution may be unimodal or bimodal. The EN is a symmetric distribution for $c = 0$ and $b_1 = b_2$. If $c > 0$ or $c = 0$ and $b_1 < b_2$, then the EN distribution is positively skewed. If $c < 0$ or $c = 0$ and $b_1 > b_2$ — negatively skewed.

Equidispersion occurs when the variance is equal to the mean ([1]). Overdispersion is a situation in which the variance exceeds the mean, underdispersion is the opposite. The mean of the $EN(a_1, b_1, a_2, b_2, 0)$ — as mentioned earlier — equals $\frac{a_1+a_2}{2}$, so the $EN(a_1, b_1, a_2 \leq -a_1, b_1, 0)$ has underdispersion property. Figure 7 shows the regions in which the $EN(a_1, b_1, 0, 1, 2)$ and $EN(a_1, b_1, 0, 2, 1)$ distributions are overdispersed and underdispersed for selected parameter values. The regions for the $EN(a_1, b_1, 1, 1, -2)$ and $EN(a_1, b_1, 1, 1, 0)$ as well as for the $EN(0, b_1, 0, 1, c)$ and $EN(0, b_1, 0, 2, c)$ are presented in the supplementary material. The curve connects the points where the distribution is equidispersed. It is interesting to point out that the relationship between a_1 and b_1 in the $EN(a_1, b_1, 0, b_2, c > 0)$ remains linear for $b_2 = 1$, $c = 2$ and $b_2 = 2$, $c = 1$ (see Figure 7).

Proposition 2.4. *The moment generating function (MGF) of the EN distribution, based on (2.1), is given by*

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} \left(\frac{1}{b_1} e^{-z_1} + \frac{1}{b_2} e^{z_2} \right) \phi[c - \exp(-z_1) + \exp(z_2); 0, 1], \tag{2.6}$$

where $z_1 = \frac{x-a_1}{b_1}$ *and* $z_2 = \frac{x-a_2}{b_2}$.

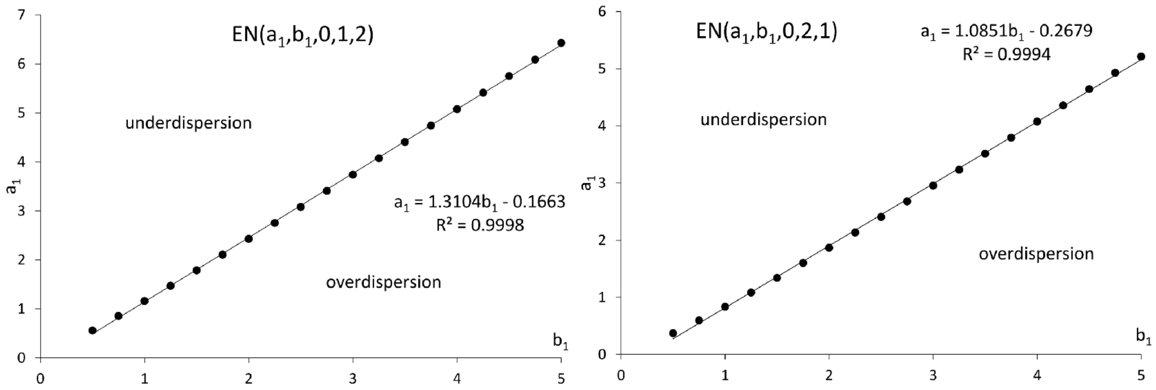

Figure 7: Dispersion regions for the $EN(a_1, b_1, 0, 1, 2)$ and $EN(a_1, b_1, 0, 2, 1)$.

2.5. Moments of order statistics

Proposition 2.5. *Let the random variable $X_{i,n}$ be the i-th order statistic $X_{1.n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ in a sample of size n from the $EN(a_1, b_1, a_2, b_2, c)$. The PDF of $X_{k,n}$ is given by*

$$f_{i.n}(x;*) = \frac{n!}{(i-1)!(n-i)!} f(x;*) F(x;*)^{i-1} [1 - F(x;*)]^{n-i},$$

where $* = (a_1, b_1, a_2, b_2, c)$, and $f(x;*)$, $F(x;*)$ are respectively given by (2.1) and (2.3).

Figure 8 plots the PDF of $X_{i,20}$ for some parameter values of the EN distribution. The k-th moment of the i-th order statistic $X_{k,n}$ is defined as

$$E\left(X_{i.n}^k\right) = \int_{-\infty}^{+\infty} x^k f_{i,n}(x).$$

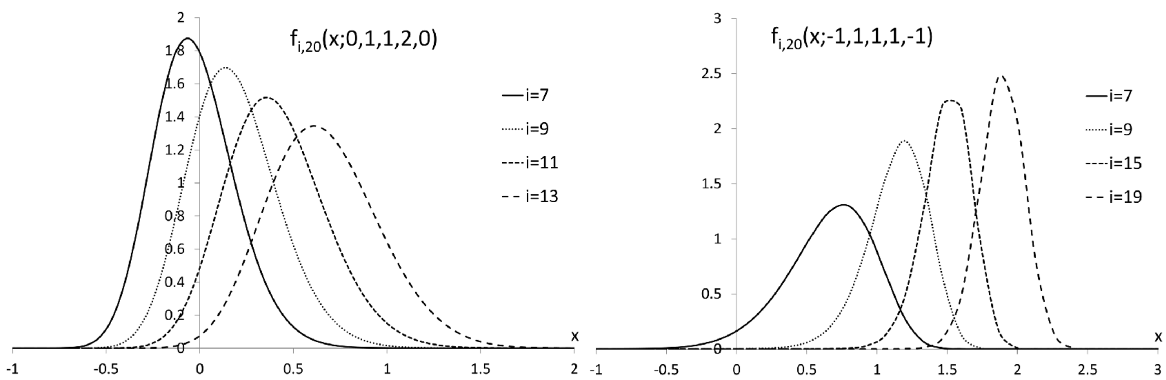

Figure 8: The PDF of the $X_{i,20}$ of the EN distribution.

2.6. Random numbers generator

Proposition 2.6. *Let* $X \sim EN(a_1, b_1, a_2, b_2, c)$, $R \sim Uniform(0,1)$. *The formula for generating* X *value, using the quantile function qEN of the EN distribution, is given by*

$$X = qEN(R; a_1, b_1, a_2, b_2, c).$$

The R codes for generating n values of X in increasing order are in the supplementary material as function rEN .

3. ESTIMATION PROCEDURES

Let $x_1^*, x_2^*, ..., x_n^*$ be a random sample of size n from the $EN(a_1, b_1, a_2, b_2, c)$. Our aim is to estimate the unknown parameter vector $\Theta = (a_1, b_1, a_2, b_2, c)$. The log-likelihood function based on (2.1) is given by

$$l(\Theta) = \sum_{i=1}^{n} \ln\left(\frac{1}{b_1} e^{-z_{1i}^*} + \frac{1}{b_2} e^{z_{2i}^*}\right) + \sum_{i=1}^{n} \ln\left[\phi\left(c - e^{-z_{1i}^*} + e^{z_{2i}^*}\right)\right], \tag{3.1}$$

where $z_{1i}^* = \frac{x_i^* - a_1}{b_1}$, $z_{2i}^* = \frac{x_i^* - a_2}{b_2}$. Solving the system of five complicated nonlinear equations in the form

$$\frac{dl(\Theta)}{da_1} = 0, \frac{dl(\Theta)}{db_1} = 0, \frac{dl(\Theta)}{da_2} = 0, \frac{dl(\Theta)}{db_2} = 0, \frac{dl(\Theta)}{dc} = 0$$

is not possible analytically. We had better maximize the log-likelihood function (3.1) in mathematical computing environments such as Excel, R and Mathcad. The MLEs of parameters a_1, b_1, a_2, b_2, c were calculated in R software using "optim" function.

The ordinary least square estimators (OLSEs) can be obtained by minimizing

$$O(\Theta) = \sum_{i=1}^{n} \left[F(x_i; a_1, b_1, a_2, b_2, c) - \frac{i}{n+1}\right]^2,$$

where $F(x; \Theta)$ is the CDF of the EN distribution (2.3).

The weighted least square estimators (WLSEs) can be obtained by minimizing

$$W(\Theta) = (n+1)^2(n+2) \sum_{i=1}^{n} \frac{1}{i(n-i+1)} \left[F(x_i; a_1, b_1, a_2, b_2, c) - \frac{i}{n+1}\right]^2,$$

where $F(x; \Theta)$ is the CDF of the EN distribution (2.3).

A simulation study is conducted to assess the properties of the MLEs, OLSEs, WLSEs of the parameter vector $\Theta = (a_1, b_1, a_2, b_2, c)$ using sample sizes of 50, 500 and 1000. In each case, 10^4 samples from the EN distribution with the specified parameters are drawn (see Figure 9).

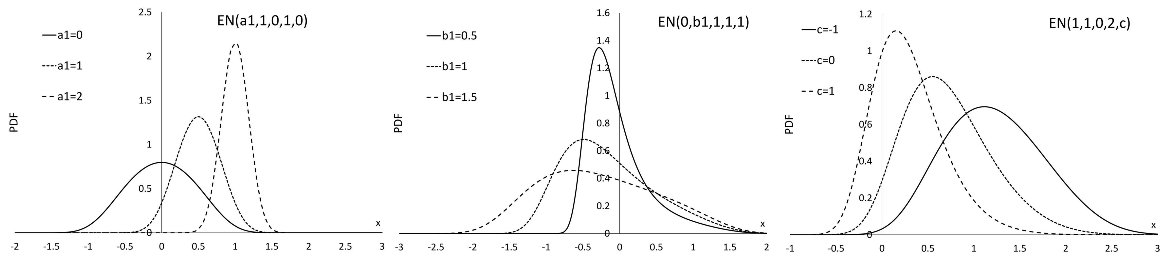

Figure 9: PDF of the EN distribution used in the estimation procedures (EPs).

The biases and the root mean squared errors (RMSEs) of the MLEs, OLSEs, WLSEs for the $EN(a_1, 1, 0, 1, 0)$ are presented in Table 6. The biases and the root mean squared errors (RMSEs) of the MLEs, OLSEs, WLSEs for the $EN(0, b_1, 1, 1, 1)$ and $EN(1, 1, 0, 2, c)$ are presented in the supplementary material.

Table 6: Biases and RMSEs of the MLEs (denoted as 1), OLSEs (denoted as 2), WLSEs (denoted as 3) for the $EN(a_1, 1, 0, 1, 0)$.

a_1	EP	n	Bias					RMSE				
			$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{a}_2$	$\widehat{b}_2$	$\widehat{c}$	$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{a}_2$	$\widehat{b}_2$	$\widehat{c}$
0	1	50	0.53	0.16	−0.65	0.15	−0.01	2.48	1.35	2.99	1.44	1.82
	2		0.56	0.47	−0.76	0.44	−0.12	1.62	1.38	1.97	1.42	1.15
	3		0.77	0.57	−1.06	0.57	−0.18	2.04	1.62	2.50	1.51	1.33
	1	500	0.10	0.03	−0.10	0.02	0.01	0.61	0.33	0.65	0.35	1.00
	2		0.24	0.14	−0.26	0.14	0.00	0.76	0.50	0.87	0.55	0.49
	3		0.14	0.09	−0.16	0.09	−0.01	0.57	0.37	0.69	0.42	0.44
	1	1e3	0.07	0.02	−0.05	0.01	0.03	0.47	0.24	0.47	0.25	0.84
	2		0.15	0.09	−0.15	0.09	0.01	0.52	0.36	0.58	0.39	0.37
	3		0.07	0.04	−0.06	0.04	0.01	0.33	0.23	0.37	0.25	0.34
1	1	50	0.41	0.13	−0.46	0.07	0.06	2.38	1.37	2.79	1.42	1.87
	2		0.28	0.24	−0.37	0.25	−0.05	1.15	0.94	1.43	1.14	1.07
	3		0.51	0.34	−0.62	0.34	−0.08	1.53	1.07	1.79	1.21	1.36
	1	500	0.29	0.10	−0.29	0.10	0.03	1.29	0.60	1.35	0.63	1.15
	2		0.15	0.07	−0.14	0.06	0.02	0.67	0.40	0.72	0.41	0.58
	3		0.22	0.09	−0.20	0.08	0.04	0.83	0.45	0.87	0.47	0.69
	1	1e3	0.19	0.07	−0.14	0.05	0.06	0.86	0.41	0.85	0.42	0.91
	2		0.10	0.05	−0.10	0.04	0.01	0.52	0.31	0.58	0.32	0.48
	3		0.16	0.07	−0.13	0.06	0.04	0.64	0.34	0.65	0.36	0.57
2	1	50	−0.02	−0.05	0.08	−0.08	0.23	1.63	1.32	1.72	0.97	2.34
	2		0.09	0.19	−0.13	0.26	0.04	0.93	0.84	1.43	1.60	1.45
	3		0.16	0.20	−0.20	0.24	−0.02	1.05	0.92	1.26	1.11	1.53
	1	500	0.11	0.01	−0.09	0.01	0.02	1.07	0.47	0.97	0.44	1.86
	2		0.05	0.03	−0.03	0.04	0.08	0.35	0.20	0.37	0.25	0.84
	3		0.06	0.03	−0.04	0.03	0.05	0.50	0.25	0.48	0.25	0.81
	1	1e3	0.08	0.01	−0.06	0.01	0.02	0.88	0.38	0.75	0.34	1.63
	2		0.03	0.02	−0.02	0.02	0.05	0.26	0.14	0.27	0.17	0.61
	3		0.04	0.02	−0.02	0.01	0.03	0.41	0.20	0.39	0.20	0.65

We observe in Table 6 that the estimates approach true values and RMSEs decrease as the sample size increases implying the consistency of the estimates. For $EN(0,1,0,1,0)$ and $EN(1,1,0,1,0)$ biases are the smallest for $\widehat{c}$ and the greatest for $\widehat{a}_2$ as well as RMSEs are the smallest for $\widehat{b}_1$ and the greatest for $\widehat{a}_2$ (see Table 6). The smallest biases are for maximum likelihood estimate (MLE) related to the $EN(0,1,0,1,0)$.

To examine the accuracy of the coverage probability of the asymptotic confidence intervals (CIs) using MLEe, another simulation study was performed with 10^4 samples using sample sizes of $50, 100, 250, 500$ and 1000. The study focused on the parameters a_1, b_1, a_2, b_2, c and samples were drawn from the $EN(0,1,1,1.25,1)$ (see Table 4). The coverage probabilities of the obtained 95% CIs for $a_1 = 0$, $b_1 = 1$, $a_2 = 1$, $b_2 = 1.25$, $c = 1$ reported in Table 7 are very close to the nominal level. The results suggest that the obtained standard errors and hence the asymptotic CIs are reliable.

Table 7: Coverage probabilities for the standard asymptotic 95% CIs.

Sample size n	a_1	b_1	a_2	b_2	c
50	0.9531	0.9521	0.9495	0.9496	0.9500
100	0.9511	0.9517	0.9422	0.9495	0.9455
250	0.9484	0.9507	0.9513	0.9529	0.9495
500	0.9509	0.9522	0.9519	0.9543	0.9522
1000	0.9449	0.9461	0.9495	0.9499	0.9472

4. SHANNON, RENYI AND TSALLIS ENTROPIES

Let $f(x, a_1, b_1, a_2, b_2, c)$ be a PDF of the EN distribution (2.1). The Shannon entropy of the EN distribution is given by [26]

$$S(a_1, b_1, a_2, b_2, c) = -\int_{-\infty}^{+\infty} f(x; a_1, b_1.a_1, b_1, c) \ln f(x; a_1, b_1.a_1, b_1, c) dx.$$

The Renyi entropy of order α for the EN distribution is defined as [21]

$$R_\alpha(a_1, b_1, a_2, b_2, c) = \frac{1}{1-\alpha} \ln\left(\int_{-\infty}^{+\infty} f(x; a_1, b_1.a_1, b_1, c)^\alpha dx \right) \quad (\alpha > 0, \alpha \neq 1).$$

The Tsallis entropy of order α for the EN distribution has the form [29]

$$T_\alpha(a_1, b_1, a_2, b_2, c) = \frac{1}{\alpha - 1} \int_{-\infty}^{+\infty} f(x; a_1, b_1.a_1, b_1, c)^\alpha dx - 1 \quad (\alpha > 0, \alpha \neq 1).$$

Renyi and Tsallis entropies converge to the Shannon entropy. Table 8 presents values of the Shannon, Renyi and Tsallis entropies for parameter values from groups A1–B2 and D1–F2 (see Table 4).

Table 8: Shannon (S), Renyi (R_α) and Tsallis (T_α) entropies. Groups of parameter values A1–B2, D1–F2.

Group	S	R_α			T_α		
		$\alpha = 0.5$	$\alpha = 2$	$\alpha = 3$	$\alpha = 0.5$	$\alpha = 2$	$\alpha = 3$
A1	0.89	1.06	0.727	0.65	−4.39	−0.52	−0.86
A2	1.43	1.64	1.17	1.02	−5.53	−0.69	−0.94
B1	0.65	0.82	0.50	0.43	−4.02	−0.39	−0.79
B2	1.47	1.69	1.19	1.04	−5.66	−0.70	−0.94
D1	1.32	1.45	1.21	1.15	−5.14	−0.70	−0.95
D2	0.65	0.71	0.59	0.57	−3.86	−0.45	−0.84
E1	0.89	1.03	0.75	0.68	−4.35	−0.53	−0.87
E2	1.66	1.76	1.51	1.40	−5.83	−0.78	−0.97
F1	0.89	1.03	0.75	0.68	−4.35	−0.53	−0.87
F2	1.66	1.76	1.51	1.40	−5.83	−0.78	−0.97

5. APPLICATION

The aim of this Section is to demonstrate the flexibility and applicability of the EN distribution. This section is composed of two real data examples. As mentioned in Introduction, the EN distribution is bimodal, so the analyzed real data are also bimodal. In papers devoted to probability distributions, Johnson distributions such as SB and SU are used very rarely in the examples, perhaps because of their unimodality. The other models selected for comparison with the new proposal are:

a) compound normal (CN) with PDF:

$$f(x; a_1, b_1, a_2, b_2, c) = \omega\phi(x; a_1, b_1) + (1-\omega)\phi(x; a_2, b_2);$$

b) compound Gumbel (CG) with PDF:

$$f_G(x; a, b) = \frac{1}{b}\exp\left[\frac{a-x}{b} - \exp\left(\frac{a-x}{b}\right)\right],$$

$$f(x; a_1, b_1, a_2, b_2, c) = \omega f_G(x; a_1, b_1) + (1-\omega) f_G(x; a_2, b_2);$$

c) two-piece power normal (TPPN) [22] with PDF:

$$\sigma = \sigma_1 I(x < \theta) + \sigma_2 I(x \geq \theta),$$

$$f(x; \theta, \sigma_1, \sigma_2, c) = \frac{c}{\sigma\sqrt{2\pi}}\left|\frac{x-\theta}{\sigma}\right|^{c-1}\exp\left[-0.5\left|\frac{x-\theta}{\sigma}\right|^{2c}\right];$$

d) bimodal skew-symmetric normal (BSSN) [12] with PDF:

$$f(x; \theta_1, \theta_2, c, d) = \frac{2c^{1.5}\left[d + (x-\theta_2)^2\right]\exp\left[-c(x-\theta_1)^2\right]}{\sqrt{\pi}\left[1 + 2c\left[d + (\theta_2-\theta_1)^2\right]\right]};$$

e) flexible generalized skew-normal of order 3 (FGSN) [16] with PDF:

$$u = \frac{x-a}{b},$$
$$f(x;a,b,\alpha_0,\alpha_1) = \frac{2}{b}\phi(u;0,1)\Phi\big(\alpha_0 u + \alpha_1 u^3;0,1\big);$$

f) bimodal asymmetric power-normal (BAPN) [8] with PDF:

$$u = \frac{x-\theta}{\sigma},$$
$$f(x;\alpha,\beta,\theta,\sigma) = \frac{\alpha 2^{\alpha}}{2^{\alpha}-1}\phi(u;0,1)\Phi(u;0,1)^{\alpha-1}\Phi(\beta u;0,1);$$

g) normal distribution with plasticizing component (NDPC) [24] with PDF:

$$u = \frac{x-a_2}{b_2}, \quad f_{pc}(x;a_2,b_2,c) = \frac{c}{b_2}|u|^{c-1}\phi(|u|^c;0,1),$$
$$f(x;a_1,b_1,a_2,b_2,c,\omega) = \omega\phi(x;a_1,b_1) + (1-\omega)f_{pc}(x;a_2,b_2,c).$$

The estimation of the model parameters is carried out by the maximum likelihood method. To avoid local maxima of the logarithmic likelihood function, the optimization routine is run 100 times with several different starting values that are widely scattered in the parameter space.

Table 9 presents the MLEs, confidence interval (CI), log-likelihood function l, AIC, BIC and HQIC for the first data sets. Models are sorted by AIC values.

Following the bootstrap method proposed in [5], [4] and [20], we used the obtained estimates $\widehat{\Theta}$ (Table 9) to derive the 95% bootstrap CIs for the parameters of distributions. We generated 10^4 samples of size n from the given distribution with values of the parameters equal to $\widehat{\Theta}$. For each obtained sample, we obtained the MLEs $\widehat{\Theta}_i^*\big(i=1,2,...,10^4\big)$ using the true values of estimates as starting values for the maximum likelihood estimation. For the 95% bootstrap CIs, we took the 250-th and 9750-th ordered estimates.

Table 10 shows p-values (sorted by p-value of the KS test) for mentioned GoFTs calculated as follows. First, we obtain the values of the Kolmogorov–Smirnov (KS), Anderson-Darling (AD) and Cramer-von Mises (CvM) test statistics (denoted ST) for true values of parameters $\widehat{\Theta}$ based on the sample $x_1, x_2, ..., x_n$. In the next step we simulate 10^4 samples $x_1', x_2', ..., x_n'$ from the given distribution with true values of parameters $\widehat{\Theta}$. For each sample, we calculate the values of the KS, AD and CvM test statistics (denoted ST^S). Finally, the p-value is calculated as $p \approx \#\big\{i: ST_i^S \geq ST\big\}10^{-4}$.

5.1. Example 1

The first real data present waiting time between eruptions and the duration of the eruption for the Old Faithful geyser in Yellowstone National Park, Wyoming, USA ([13]).

The data consist of 272 observations of the variable "eruptions numeric Eruption time in mins" and are available in the R software with code faithful[1].

As shown in Table 9 the EN model is definitely the best in terms of the $-l$, AIC, BIC and HQIC values. The AIC ranking is the same as the BIC and HQIC rankings. The EN model is definitely distinguished by the p-values (see Table 10). The p-value ranking for the KS test is, with only one exception, the same as the p-value rankings for the AD and CvM tests. The information criteria ranking is not the same as the p-value ones. It is worth noting that the rankings are similar for most models, with the biggest difference in the rankings for the TPPN model.

Table 9: Results of estimation. Information criteria. Example 1.

Model		$\widehat{\Theta}$	95%CI	$-l$	AIC	BIC	HQIC
EN	$\widehat{a}_1$	-1.453	$[-1.486, -1.416]$	224.331	456.663	471.086	462.453
	$\widehat{b}_1$	0.185	$[0.147, 0.224]$				
	$\widehat{a}_2$	0.820	$[0.729, 0.896]$				
	$\widehat{b}_2$	0.481	$[0.405, 0.563]$				
	$\widehat{c}$	-0.427	$[-0.602, -0.264]$				
NDPC	$\widehat{a}_1$	0.508	$[0.353, 0.611]$	227.238	466.476	488.111	475.161
	$\widehat{b}_1$	0.375	$[0.275, 0.444]$				
	$\widehat{a}_2$	-0.173	$[-0.210, -0.137]$				
	$\widehat{b}_2$	1.219	$[1.182, 1.256]$				
	$\widehat{c}$	4.795	$[4.186, 5.786]$				
	$\widehat{\omega}$	0.342	$[0.235, 0.432]$				
CN	$\widehat{a}_1$	0.688	$[0.631, 0.745]$	240.394	490.788	508.817	498.026
	$\widehat{b}_1$	0.383	$[0.341, 0.423]$				
	$\widehat{a}_2$	-1.287	$[-1.328, -1.245]$				
	$\widehat{b}_2$	0.206	$[0.175, 0.237]$				
	$\widehat{\omega}$	0.652	$[0.597, 0.706]$				
TPPN	$\widehat{\theta}_1$	-0.454	$[-0.537, -0.370]$	244.651	497.301	511.724	503.092
	$\widehat{\sigma}_1$	0.921	$[0.835, 1.007]$				
	$\widehat{\sigma}_2$	1.357	$[1.267, 1.448]$				
	$\widehat{c}$	3.166	$[2.891, 3.549]$				
CG	$\widehat{a}_1$	-1.367	$[-1.405, -1.307]$	250.318	510.636	528.665	517.874
	$\widehat{b}_1$	0.180	$[0.145, 0.213]$				
	$\widehat{a}_2$	0.532	$[0.456, 0.604]$				
	$\widehat{b}_2$	0.362	$[0.218, 0.411]$				
	$\widehat{\omega}$	0.362	$[0.305, 0.427]$				
FGSN	$\widehat{a}$	0.191	$[0.153, 0.236]$	271.813	551.626	566.049	557.416
	$\widehat{b}$	1.016	$[0.930, 1.102]$				
	$\widehat{\alpha}_0$	4.148	$[3.389, 5.351]$				
	$\widehat{\alpha}_1$	-3.406	$[-4.942, -2.460]$				
BSSN	$\widehat{\theta}_1$	-0.212	$[-0.265, -0.155]$	277.255	562.509	576.932	568.300
	$\widehat{\theta}_2$	1.402	$[1.279, 1.625]$				
	$\widehat{c}$	-0.323	$[-0.372, -0.265]$				
	$\widehat{d}$	0.003	$[-0.045, 0.021]$				
BAPN	$\widehat{\alpha}$	16.160	$[14.243, 18.776]$	464.240	936.479	953.675	943.203
	$\widehat{\beta}$	0.048	$[-0.011, 0.108]$				
	$\widehat{\theta}$	-0.070	$[-0.090, -0.040]$				
	$\widehat{\sigma}$	0.543	$[0.520, 0.565]$				

Table 10: The KS, AD and CvM tests. Example 1.

Model	**KS test**		**AD test**		**CvM test**	
	TS	p-value	TS	p-value	TS	p-value
EN	0.0306	0.935	0.2845	0.9546	0.0414	0.9316
CN	0.049	0.4644	1.063	0.322	0.124	0.4741
NDPC	0.0514	0.4108	1.1111	0.3066	0.1724	0.3336
CG	0.0639	0.1858	2.129	0.0801	0.2636	0.1775
BSSN	0.0751	0.0814	3.896	0.0118	0.5454	0.0302
FGSN	0.0832	0.0404	3.289	0.0195	0.4366	0.0537
BAPN	0.1331	0.0001	8.7874	0.0002	1.0307	0.0019
TPPN	0.1495	0	7.201	0	1.516	0

Concluding, the EN model fits better than the other models analyzed in this case.

The second real data present Intercountry Life-Cycle Savings Data ([27], [6]). A detailed analysis of this example done identically to Example 1 is presented in the supplementary material.

6. CONCLUSIONS

Heterogeneity is not the only one cause of population distribution's bimodality. The population distribution is shaped by many factors. Therefore, the aim of the paper was to introduce into a family of the mixed bimodal distributions two distant relatives more. The relatives in question are distant since they are not of mixture form. So, they was denoted as non-mixed bimodal distributions. It is author's duty to give potential user of non-mixed bimodal distributions warning. Parameters of non-mixed bimodal distributions are not so clearly interpretable as parameters of mixed bimodal distributions are. Interpretability complication may, in turn, complicate conclusions when statistical reasoning procedure involves non-mixed bimodal distributions.

As a result of considerations presented in this paper two probability distributions denoted SC and SD came into existence. The distributions are members of the Johnson family of distribution. The SC and SD were tested in great depth, first for flexibility then for applicability.

In order to test for flexibility the Malachov plot was applied. The Malachov plot is a rectangular coordinate system with skewness (γ_1) as the abscissa and kurtozis (γ_2) as the ordinate. Points located below Malachov parabola $\gamma_2 = \gamma_1^2 + 1$ are related to obtainable γ_1/γ_2 combinations. The more flexible distribution is the wider points are scattered on the Malachov plot. In this paper the above fact served as a basis for definition of numerical flexibility measure being a fraction of an area "occupied" by particular distribution. The skewness-kurtosis-square measure was denoted SKS. Points are dimensionless entities, for a purpose of SKS measure, they were replaced with micro-squares. The best dispersion of points (γ_1, γ_2) occurs for the SD, SU and SB distribution.

After having flexibility testing completed the EN distribution was tested for applicability. For the purpose of applicability testing two real data sets were used. Empirical pdf's estimated from these data sets display bimodality. The EN had seven competitors with respect to applicability. These were already existing distributions that all have a property of bimodality. The competition consisted in fitting distributions to the data sets. Two types of rankings were performed. First the EN and its competitors were ranked with respect to information criteria. The criteria were AIC, BIC and HQIC ones. Then the EN and its competitors were ranked with respect to results of goodness-of-fit tests. The results were measured with p-values. The goodness-of-fit test involved in rankings were Kolmogorov-Smirnow, Anderson-Darling and Crmaer-von Mises ones. Altogether there were three information criteria rankings and three p-value rankings performed. It is interesting that all three information criteria rankings gave quite the same results. What makes a matter of rankings more interesting is that all three p-value rankings gave quite the same results too! So, one can say about one joint information criteria ranking and one joint p-value ranking. These rankings considerably differed from each other. In its essence this fact is not even strange since criteria differ considerably too. It is of special interest that the EN ranks high in all the rankings.

The content of the paper shows that the EN (including SC) as a new member of the Johnson family of distributions and simultaneously as a new distribution from the non-mixed bimodal distribution category, is a competitive model that deserves to be added to the existing distributions in modeling data.

ACKNOWLEDGMENTS

The author is thankful to the two Referees whose helpful comments and suggestions contributed to improve the quality of the paper.

REFERENCES

[1] Alzaatreh A.; Famoye, F. and Lee, C. (2014). The gamma-normal distribution: properties and application, *Computational Statistics & Data Analysi*, **69**, 67–80.

[2] Athayde, E.; Azevedo, C.; Leiva, V. and Sanhueza, A. (2012). About Birnbaum–Saunders distributions based on the Johnson system, *Communications in Statistics – Theory and Methods*, **41**(11), 2061–2079.

[3] Babak, V.; Eremenko, V. and Zaporozhets, A. (2019). Research of diagnostic parameters of composite materials using Johnson distribution, *Int. J. Comput*, **18**(4), 483–494.

[4] Balakrishnan, N. and Ristić, M.M. (2016). Multivariate families of gamma-generated distributions with finite or infinite support above or below the diagonal, *Journal of Multivariate Analysis*, **143**, 194–207.

[5] Balakrishna, N. and Shiji, K. (2014). On a class of bivariate exponential distributions, *Statistics & Probability Letter*, **85**, 153–180.

[6] Belsley, D.A.; Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics*, Wiley, New York.

[7] Birnbaum, Z.W. and Saunders, S.C. (1969). A new family of life distributions, *Journal of Applied Probability*, **6**(2), 319–327.

[8] Bolfarine, H.; Martínez-Flórez, G. and Salinas, H.S. (2018). Bimodal symmetric-asymmetric power-normal families, *Communications in Statistics – Theory and Methods*, **47**(2), 259–276.

[9] Castro R. (2013). Goodness-of-fit (GoF) Tests. Available at `http://www.win.tue.nl/~rmcastro/AppStat2013/files/lectures23.pdf`

[10] Díaz-García, J.A. and Domínguez-Molina, J.R. (2005). Generalised Birnbaum–Saunders and sinh-normal distributions, *Comunicación Técnica*, **I**, 5.

[11] Gaddum, J.H. (1945). *Lognormal distributions*, **1**(56), 463–466.

[12] Hassan, M.Y. and El-Bassiouni, M.Y. (2016). Bimodal skew-symmetric normal distribution, *Communications in Statistics – Theory and Methods*, **45**(5), 1527–1541.

[13] Härdle, W. (1991). *Smoothing Techniques with Implementation in S*, Springer, New York.

[14] Johnson, N.L. (1949). System of frequency curves generated by methods of translation, *Biometrika*, **36**, 149–176.

[15] Lu, Y.; Ramirez, O.A.; Rejesus, R.M.; Knight, T.O. and Sherrick, B.J. (2008). Empirically evaluating the flexibility of the Johnson family of distributions: A crop insurance application, *Agricultural and Resource Economics Review*, **37**, 79–91.

[16] Ma, Y. (2004). Flexible class of skew-symmetric distributions, *Journal of Statistics*, **31**(3), 456–468.

[17] Malakhov, A.N. (1978). *A Cumulant Analysis of Random Non-Gaussian Processes and Their Transformations*, Soviet Radio, Moscow (in Russian).

[18] Owen, W.J. (2004). Another look at the Birnbaum–Saunders distribution. Available at `https://www.researchgate.net/publication/267833513_Another_Look_at_the_Birnbaum_Saunders_Distribution`

[19] Rieck, J.R. and Nedelman, J.R. (1991). A log-linear model for the Birnbaum–Saunders distribution, *Technometrics*, **33**(1), 51–60.

[20] Ristić, S.; Popovic, M.M.B.V.; Zografos K. and Balakrishnan N. (2018). Discrimination among bivariate beta-generated distributions, *Statistics*, **52**(2), 303–320.

[21] Renyi, A. (1961). On measures of entropy and information, *Proc. Berekeley Symposium, Statist. Probability*, **1**, 547–561.

[22] Sulewski, P. (2019). Two-Piece Power Normal Distribution, *Communications in Statistics – Theory and Method*, DOI:10.1080/03610926.2019.1674871.

[23] Sulewski, P. (2019). Modified Lilliefors goodness-of-fit test for normality, *Communications in Statistics – Theory and Method*, DOI:10.1080/03610918. 2019.1664580.

[24] Sulewski, P. (2020). Normal distribution with plasticizing component, *Communications in Statistics – Theory and Method*, DOI:10.1080/03610926.2020.1837881.

[25] Stute W. (1993). Bootstrap based goodness-of-fit-tests, *Metrika*, **40**, 243–256.

[26] Shannon, C.E. (1948). A mathematical theory of communication, *Bell System Technical Journa*, **27**, 379–423.

[27] Sterling, A. (1977). Unpublished BS Thesis, Massachusetts Institute of Technology.

[28] Tabass, M.S.; Borzadaran, G.M. and Amini, M. (2016). Renyi entropy in continuous case is not the limit of discrete case, *Mathematical Sciences and Applications e-Notes*, **4**(1), 113–117.

[29] Tsallis, C. (1988). Possible generalizations of Boltzmann-Gibbs statistics, *Journal of Statistical Physics*, **52**, 479–487.