Extended Easily Changeable Kurtosis Distribution

Piotr Sulewski

Pomeranian University in Słupsk

www: sulewski.apsl.edu.pl
e-mail: piotr.sulewski@apsl.edu.pl

MET 2023
Metodologia Badań Statystycznych
3 - 5 lipca 2023, Warszawa

- Instead of kurtosis γ_2, the paper analyzes the kurtosis $\overline{\gamma}_2 = \gamma_2 - 3$, which can be positive or negative.

  For simplicity, the kurtosis of the normal distribution is zero!!!

- Instead of kurtosis γ_2, the paper analyzes the kurtosis $\overline{\gamma}_2 = \gamma_2 - 3$, which can be positive or negative.

  For simplicity, the kurtosis of the normal distribution is zero!!!

- Symmetric distributions do not form such a big family as asymmetric distributions.

- Instead of kurtosis γ_2, the paper analyzes the kurtosis $\overline{\gamma}_2 = \gamma_2 - 3$, which can be positive or negative.

  For simplicity, the kurtosis of the normal distribution is zero!!!

- Symmetric distributions do not form such a big family as asymmetric distributions.

- There is a group of asymmetric distributions, which are symmetrical for certain parameter values, e.q. the truncated normal, Birnbaum-Saunders, skew-normal, beta, two-piece normal and two-piece power normal distributions.

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),
2. distribution with at most two modes (extended normal, extended t, extended Laplace, plasticizing component)

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),
2. distribution with at most two modes (extended normal, extended t, extended Laplace, plasticizing component)
3. distribution with two modes (bimodal exponential power,bimodal power normal)

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),
2. distribution with at most two modes (extended normal, extended t, extended Laplace, plasticizing component)
3. distribution with two modes (bimodal exponential power,bimodal power normal)
4. distribution with an undefined kurtosis (Cauchy,Voigt),

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),
2. distribution with at most two modes (extended normal, extended t, extended Laplace, plasticizing component)
3. distribution with two modes (bimodal exponential power,bimodal power normal)
4. distribution with an undefined kurtosis (Cauchy,Voigt),
5. distribution with a complicated kurtosis formula (bimodal power normal, Tukey, von Mises),

We can divide symmetric distributions into eleven groups, namely:

1. distribution with one mode (normal, Laplace, logistic),
2. distribution with at most two modes (extended normal, extended t, extended Laplace, plasticizing component)
3. distribution with two modes (bimodal exponential power,bimodal power normal)
4. distribution with an undefined kurtosis (Cauchy,Voigt),
5. distribution with a complicated kurtosis formula (bimodal power normal, Tukey, von Mises),
6. distribution with a constant kurtosis value (uniform, hyperbolic secant, semicircle),

7. distribution with discrete kurtosis values (extended t, Bates, Irwin-Hall),

7. distribution with discrete kurtosis values (extended t, Bates, Irwin-Hall),
8. distribution with kurtosis values on infinite interval (generalized normal, normal-exponential-gamma, U-quadratic),

7. distribution with discrete kurtosis values (extended t, Bates, Irwin-Hall),
8. distribution with kurtosis values on infinite interval (generalized normal, normal-exponential-gamma, U-quadratic),
9. distribution with kurtosis values on finite interval (ECK, Q-gaussian, von Mises, U-power),

7. distribution with discrete kurtosis values (extended t, Bates, Irwin-Hall),
8. distribution with kurtosis values on infinite interval (generalized normal, normal-exponential-gamma, U-quadratic),
9. distribution with kurtosis values on finite interval (ECK, Q-gaussian, von Mises, U-power),
10. distribution with an existing discontinuous function $p = f(\overline{\gamma}_2)$, where p is the shape parameter (t, Bates, Irwin-Hall).

7. distribution with discrete kurtosis values (extended t, Bates, Irwin-Hall),
8. distribution with kurtosis values on infinite interval (generalized normal, normal-exponential-gamma, U-quadratic),
9. distribution with kurtosis values on finite interval (ECK, Q-gaussian, von Mises, U-power),
10. distribution with an existing discontinuous function $p = f(\overline{\gamma}_2)$, where p is the shape parameter (t, Bates, Irwin-Hall).
11. distribution with an existing continuous function $p = f(\overline{\gamma}_2)$, where p is the shape parameter (Q-gaussian. ECK).

The $ECK(a > 0, p > -1)$ is unimodal distribution defined in the finite domain with $p = f(\overline{\gamma}_2) = \frac{-5\overline{\gamma}_2 - 6}{2\overline{\gamma}_2}$ and can be used to model kurtosis $\overline{\gamma}_2 \in (-2, 0)$.

The $ECK(a>0, p>-1)$ is unimodal distribution defined in the finite domain with $p = f(\overline{\gamma}_2) = \frac{-5\overline{\gamma}_2-6}{2\overline{\gamma}_2}$ and can be used to model kurtosis $\overline{\gamma}_2 \in (-2,0)$.
The kurtosis of the EECK distribution is $\overline{\gamma}_2^{EECK} \geq -2$, so this distribution, like

- the generalized normal GN $(\overline{\gamma}_2^{GN} \geq -1.2\left(\overline{\gamma}_2^{GN} \neq 0\right))$,
- the normal-exponential-gamma NEG $(\overline{\gamma}_2^{NEG} > 0)$
- the Tukey T defined in infinite domain $(\overline{\gamma}_2^{T} > 0)$,

belongs to the family of symmetrical, unimodal distributions with kurtosis values on infinite interval.

The $ECK(a > 0, p > -1)$ is unimodal distribution defined in the finite domain with $p = f(\overline{\gamma}_2) = \frac{-5\overline{\gamma}_2 - 6}{2\overline{\gamma}_2}$ and can be used to model kurtosis $\overline{\gamma}_2 \in (-2, 0)$.
The kurtosis of the EECK distribution is $\overline{\gamma}_2^{EECK} \geq -2$, so this distribution, like

- the generalized normal GN $(\overline{\gamma}_2^{GN} \geq -1.2\,(\overline{\gamma}_2^{GN} \neq 0))$,
- the normal-exponential-gamma NEG $(\overline{\gamma}_2^{NEG} > 0)$
- the Tukey T defined in infinite domain $(\overline{\gamma}_2^{T} > 0)$,

belongs to the family of symmetrical, unimodal distributions with kurtosis values on infinite interval.
PDF of the NEG has a complicated form, and the analytical formula for kurtosis does not exist.

The $ECK(a > 0, p > -1)$ is unimodal distribution defined in the finite domain with $p = f(\overline{\gamma}_2) = \frac{-5\overline{\gamma}_2 - 6}{2\overline{\gamma}_2}$ and can be used to model kurtosis $\overline{\gamma}_2 \in (-2, 0)$.
The kurtosis of the EECK distribution is $\overline{\gamma}_2^{EECK} \geq -2$, so this distribution, like

- the generalized normal GN $(\overline{\gamma}_2^{GN} \geq -1.2 \, (\overline{\gamma}_2^{GN} \neq 0))$,
- the normal-exponential-gamma NEG $(\overline{\gamma}_2^{NEG} > 0)$
- the Tukey T defined in infinite domain $(\overline{\gamma}_2^{T} > 0)$,

belongs to the family of symmetrical, unimodal distributions with kurtosis values on infinite interval.
PDF of the NEG has a complicated form, and the analytical formula for kurtosis does not exist.
PDF of the T has a simple, closed form for a few exceptional values of the shape parameter, e.g. we get, respectively, for $\lambda = \{1, 0\}$ uniform and logistic distributions.

The analytical formulas for kurtosis of the EECK, GN and T distributions are respectively:

The analytical formulas for kurtosis of the EECK, GN and T distributions are respectively:

$$\overline{\gamma}_2^{EECK} = \frac{\Gamma\left(\frac{1}{q}\right)\Gamma\left(\frac{5}{q}\right)\Gamma\left(p+\frac{3}{q}+1\right)^2}{\Gamma\left(\frac{3}{q}\right)^2\Gamma\left(p+\frac{1}{q}+1\right)\Gamma\left(p+\frac{5}{q}+1\right)} - 3\,(p > -1, q > 0),$$

The analytical formulas for kurtosis of the EECK, GN and T distributions are respectively:

$$\overline{\gamma}_2^{EECK} = \frac{\Gamma\left(\frac{1}{q}\right)\Gamma\left(\frac{5}{q}\right)\Gamma\left(p+\frac{3}{q}+1\right)^2}{\Gamma\left(\frac{3}{q}\right)^2\Gamma\left(p+\frac{1}{q}+1\right)\Gamma\left(p+\frac{5}{q}+1\right)} - 3\,(p > -1, q > 0),$$

$$\overline{\gamma}_2^{GN} = \Gamma\left(\frac{5}{\beta}\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{3}{\beta}\right)^{-2} - 3\,(\beta > 0)$$

The analytical formulas for kurtosis of the EECK, GN and T distributions are respectively:

$$\overline{\gamma}_2^{EECK} = \frac{\Gamma\left(\frac{1}{q}\right)\Gamma\left(\frac{5}{q}\right)\Gamma\left(p+\frac{3}{q}+1\right)^2}{\Gamma\left(\frac{3}{q}\right)^2\Gamma\left(p+\frac{1}{q}+1\right)\Gamma\left(p+\frac{5}{q}+1\right)} - 3\,(p > -1, q > 0),$$

$$\overline{\gamma}_2^{GN} = \Gamma\left(\frac{5}{\beta}\right)\Gamma\left(\frac{1}{\beta}\right)\Gamma\left(\frac{3}{\beta}\right)^{-2} - 3\,(\beta > 0)$$

$$\overline{\gamma}_2^{T} = \frac{\Gamma(2\lambda+1)^2\left[3\Gamma(2\lambda+1)^2-4\Gamma(\lambda+1)\Gamma(3\lambda+1)+\Gamma(4\lambda+1)\right]}{(8\lambda+1)(2\lambda+1)^{-2}\Gamma(4\lambda+1)\left[\Gamma(\lambda+1)^2-\Gamma(2\lambda+1)\right]^2} - 3\,(-0.25 < \lambda)$$

Figure 1 shows the kurtosis as a function of the shape parameters $p > -1$, $\beta > 0$ and $\lambda \in (-0.25, 0)$. The EECK and GN distributions can be used to model the negative and positive kurtosis. The negative values of kurtosis for the EECK and GN distributions are available on $[-2, 0]$ and $[-1.2, 0)$, respectively. It is worth mentioning that the EECK is defined in the finite domain whereas GN and T are defined in infinite domain.

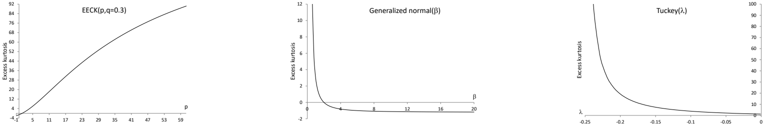


Figure: kurtosis as a function of shape parameter

The new proposal can be extremely useful when you want to seamlessly test the goodness-of-fit tests (GoFTs) ability to detect deviations from normality caused by a negative and positive kurtosis.

The new proposal can be extremely useful when you want to seamlessly test the goodness-of-fit tests (GoFTs) ability to detect deviations from normality caused by a negative and positive kurtosis.

Special cases of the $EECK(p > -1, q > 0)$ distribution are: the uniform, triangle and obviously $ECK(a > 0, p > -1)$. The $EECK(p > -1, q > 0)$ tends to the normal distribution

Name	Modes	Range of $\overline{\gamma}_2$	Name	Modes	Range of $\overline{\gamma}_2$
Arcsine	2	-1.5	Normal-exponential-gamma	1	$(-\infty, 3)$
Bates	1	$[-1.2, 0)$	Plasticizing component	2	$(-2, 0)$
Bimodal exponential power	$1, 2$	$[-3, 3]$	Q-gaussian	1	$[-0.857, 0]$
Bimodal normal	2	$-4/3$	Rademacher	N/A	-2
Bimodal Laplace	2	$1/3$	Raised cosine	1	$\frac{6(90-\pi^4)}{5(\pi^2-6)^2}$
Bimodal power normal	$1, 2$	$(-2, 0) \vee (0, 10.97)$	Sine	1	$\frac{2(96-\pi^4)}{(\pi^2-8)^2}$
Cauchy	1	-	Semicircle	1	-1
Degenerate	1	-	t	1	$(0, 6]$
ECK	1	$(-2, 0)$	Triangle	1	-0.6
Extended Normal	$1, 2$	$[-4/3, 0]$	Tukey[1]	1	$(0, \infty)$
Extended Laplace	$1, 2$	$(-1/3, 3]$	Tukey[2]	1	$[-1.25, 10.59]$
Extended t	$1, 2$	$[-4/3, 6]$	Uniform	∞	$-6/5$
Generalized normal	1	$[-1.2, 0) \vee (0, \infty)$	U-power	2	$[-2, -1.81]$
Hyperbolic secant	1	2	U-quadratic	2	$(0, \infty)$
Irwin–Hall	1	$[-1.2, 0)$	U-shaped	2	-1.5
Laplace	1	3	Voigt	1	-
Logistic	1	$6/5$	Von Mises	1	$[-1.2, 1.069]$
Normal	1	0	Wigner semicircle	1	-1

[1]infinite domain, [2]finite domain,

Source: Own material.

Table: Symmetric distributions with range of kurtosis and modality

Definition 1 The Eta function for $p > -1$ and $q > 0$ is defined as

$$H(p,q) = \int_{-1}^{1} [1 - |x|^q]^p \, dx = \frac{2B\left(\frac{1}{q}, p+1\right)}{q} = \frac{2\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)}{\Gamma\left(p+\frac{1}{q}+1\right)}$$

where $B(u,v)$ is the beta function.

Definition 1 The Eta function for $p > -1$ and $q > 0$ is defined as

$$H\left(p,q\right) = \int_{-1}^{1}\left[1-\left|x\right|^{q}\right]^{p}dx = \frac{2B\left(\frac{1}{q},p+1\right)}{q} = \frac{2\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)}{\Gamma\left(p+\frac{1}{q}+1\right)}$$

where $B(u,v)$ is the beta function.
Calculations were performed by the formula

$$\int_0^1 x^{a-1}\left(1-x^b\right)^{c-1}dx = \frac{B\left(\frac{a}{b},c\right)}{b}$$

Exemplary values of the Eta function: $H\left(1,1\right)=1$, $H\left(0,1\right)=2$, $H\left(-0.5,1\right)=4$, $H\left(1,0.5\right)=\frac{2}{3}$, $H(0.5,1)=\frac{4}{3}$.

Definition 2 The distribution of the random variable X with PDF given by

$$f(x;p,q) = \frac{[1-|x|^q]^p}{H(p,q)}, x \in \begin{cases} (-1,1) & \text{if } -1 < p < 0 \\ [-1,1] & \text{if } p \geq 0 \end{cases} \tag{1}$$

is called the extended easily changeable kurtosis (EECK) distribution, where $p > -1$ and $q > 0$ are the shape parameters.

Definition 2 The distribution of the random variable X with PDF given by

$$f(x;p,q) = \frac{\left[1-|x|^q\right]^p}{H(p,q)}, x \in \begin{cases} (-1,1) & \text{if } -1 < p < 0 \\ [-1,1] & \text{if } p \geq 0 \end{cases} \quad (1)$$

is called the extended easily changeable kurtosis (EECK) distribution, where $p > -1$ and $q > 0$ are the shape parameters.
The $EECK(p > -1, q > 0)$ is symmetric around zero, since $f(x;p,q) = f(-x;p,q)$.
The $EECK(p > -1, q = 2)$ is the $ECK(a = 1, p > -1)$ (Sulewski, 2022) .

The variance of the new proposal equals

$$\mu_2 = \frac{(1+pq)\,\Gamma\left(\frac{3}{q}\right)\Gamma\left(p+\frac{1}{q}\right)}{(3+pq)\,\Gamma\left(\frac{1}{q}\right)\Gamma\left(p+\frac{3}{q}\right)}$$

therefore the $EECK(p,q)$ distribution tends to the normal distribution $N\left(0,\sqrt{\mu_2}\right)$ with PDF $\phi\left(x;0,\sqrt{\mu_2}\right)$

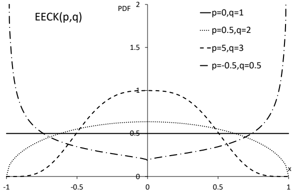
The variance of the new proposal equals

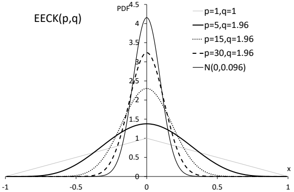
$$\mu_2 = \frac{(1+pq)\,\Gamma\left(\frac{3}{q}\right)\Gamma\left(p+\frac{1}{q}\right)}{(3+pq)\,\Gamma\left(\frac{1}{q}\right)\Gamma\left(p+\frac{3}{q}\right)}$$

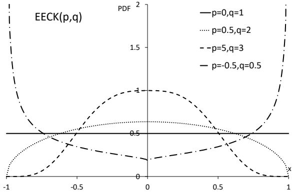
therefore the $EECK(p,q)$ distribution tends to the normal distribution $N\left(0,\sqrt{\mu_2}\right)$ with PDF $\phi\left(x;0,\sqrt{\mu_2}\right)$
Let $M(p,q)$ be the similarity measure of these distributions. We have for $p>-1, q>0$ (Sulewski, 2020)

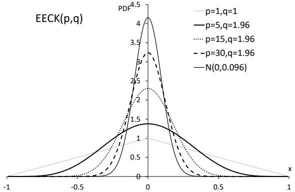
$$M(p,q) = \int_{-1}^{1} \min\left\{ f(x;p,q), \phi\left[x;0,\sqrt{\frac{(1+pq)\Gamma\left(\frac{3}{q}\right)\Gamma\left(p+\frac{1}{q}\right)}{(3+pq)\Gamma\left(\frac{1}{q}\right)\Gamma\left(p+\frac{3}{q}\right)}}\right]\right\}$$

The similarity measure M takes values on (0,1) and if PDFs are identical then $M=1$. For example $M(33,1)=0.871$, $M(33,2)=0.995$, $M(33,2.5)=0.961$. It has the highest values for $q=1.96$. We have $M(50,1.96)=0.999$.

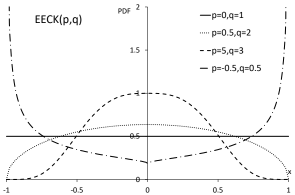

EECK(p,q)
PDF
p=0,q=1
p=0.5,q=2
p=5,q=3
p=-0.5,q=0.5
x

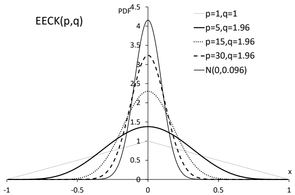

EECK(p,q)
PDF
p=1,q=1
p=5,q=1.96
p=15,q=1.96
p=30,q=1.96
N(0,0.096)
x



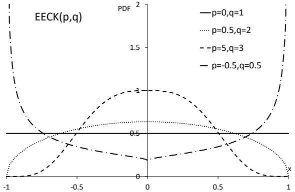


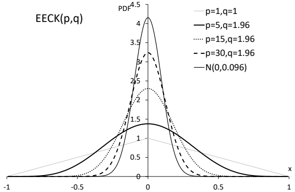
- $EECK(p > -1, q > 0)$ is symmetrical



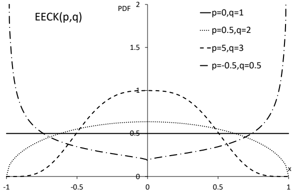


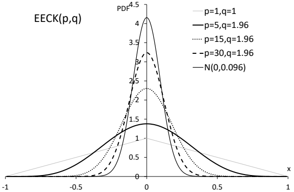
- $EECK(p > -1, q > 0)$ is symmetrical
- The $EECK(p = 0, q > 0)$ is the $U(-1, 1)$



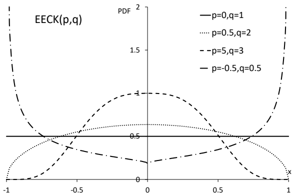


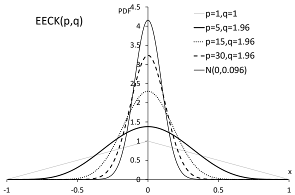
- $EECK(p > -1, q > 0)$ is symmetrical
- The $EECK(p = 0, q > 0)$ is the $U(-1, 1)$
- The $EECK(p > 0, q > 0)$ is unimodal with mode equals 0



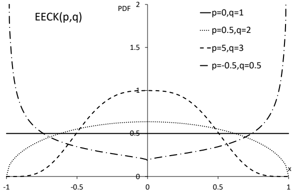


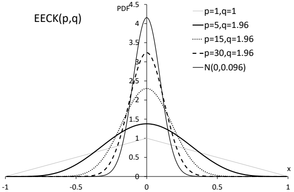
- $EECK(p > -1, q > 0)$ is symmetrical
- The $EECK(p = 0, q > 0)$ is the $U(-1, 1)$
- The $EECK(p > 0, q > 0)$ is unimodal with mode equals 0
- The $EECK(-1 < p < 0, q > 0)$ is pseudo $(-1 < x < 1)$ bimodal with bathtub (anti-mode) shape





- $EECK(p > -1, q > 0)$ is symmetrical
- The $EECK(p = 0, q > 0)$ is the $U(-1, 1)$
- The $EECK(p > 0, q > 0)$ is unimodal with mode equals 0
- The $EECK(-1 < p < 0, q > 0)$ is pseudo $(-1 < x < 1)$ bimodal with bathtub (anti-mode) shape
- The $EECK(p = 1, q = 1)$ is the triangle distribution





- $EECK(p > -1, q > 0)$ is symmetrical
- The $EECK(p = 0, q > 0)$ is the $U(-1, 1)$
- The $EECK(p > 0, q > 0)$ is unimodal with mode equals 0
- The $EECK(-1 < p < 0, q > 0)$ is pseudo $(-1 < x < 1)$ bimodal with bathtub (anti-mode) shape
- The $EECK(p = 1, q = 1)$ is the triangle distribution

Theorem 1. If $X \sim EECK(p > -1, q > 0)$ with PDF $f(x; p, q)$ (1) then CDF of X is given by

$$F(x; p, q) = 0.5 + x \frac{{}_2F_1\left(-p, \frac{1}{q}, 1 + \frac{1}{q}, |x|^q\right)}{H(p, q)} \tag{2}$$

where ${}_2F_1(a, b, c, x)$ is the Gaussian hypergeometric function.

Theorem 1. If $X \sim EECK(p > -1, q > 0)$ with PDF $f(x; p, q)$ (1) then CDF of X is given by

$$F(x; p, q) = 0.5 + x \frac{{}_2F_1\left(-p, \frac{1}{q}, 1 + \frac{1}{q}, |x|^q\right)}{H(p, q)} \tag{2}$$

where ${}_2F_1(a, b, c, x)$ is the Gaussian hypergeometric function.

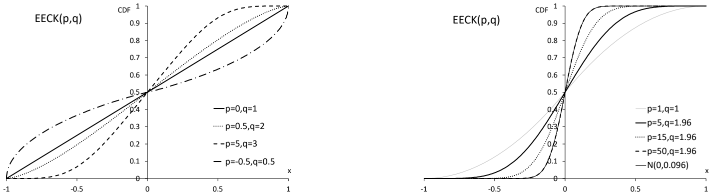


Figure: CDF of the $EECK(a, p)$ distribution for various parameter values

Theorem 2. The $EECK(p > -1, q > 0)$ distribution with PDF given by (1) is identifiable in the parameter space $v = (p, q)$.

Theorem 2. The $EECK(p > -1, q > 0)$ distribution with PDF given by (1) is identifiable in the parameter space $v = (p, q)$.
Proof Let $v_1 = (p_1, q_1)$ and $v_2 = (p_2, q_2)$. Let us suppose that $f_{v_1}(x) = f_{v_2}(x)$ for all x from support. This condition implies that

$$\frac{q_1(1-|x|^{q_1})^{p_1}}{2B\left(\frac{1}{q_1},p_1+1\right)} = \frac{q_2(1-|x|^{q_2})^{p_2}}{2B\left(\frac{1}{q_2},p_2+1\right)}$$

If we apply log to both sides we obtain the system of three equations

$$\log\left(\frac{q_1}{q_2}\right) = 0, p_1\log(1-|x|^{q_1}) - p_2\log(1-|x|^{q_2}) = 0, \log\left[\frac{B\left(\frac{1}{q_2},p_2+1\right)}{B\left(\frac{1}{q_1},p_1+1\right)}\right] = 0$$

From the first equation is $q_1 = q_2$ and then from the second one is $p_1 = p_2$. The proof is complete.

Theorem 3. Let $X \sim EECK(p > -1, q > 0)$.

- If $p = 0$ then modal values $x_m \in [-1, 1]$ (case of uniform distribution).

Theorem 3. Let $X \sim EECK(p > -1, q > 0)$.

- If $p = 0$ then modal values $x_m \in [-1, 1]$ (case of uniform distribution).
- If $p > 0$ then $x_m = 0$.

Theorem 3. Let $X \sim EECK(p > -1, q > 0)$.

- If $p = 0$ then modal values $x_m \in [-1, 1]$ (case of uniform distribution).
- If $p > 0$ then $x_m = 0$.
- If $-1 < p < 0$ then the $EECK(p, q)$ distribution is pseudo bimodal with modes $x_m(-1), x_m(1)$.

Theorem 3. Let $X \sim EECK(p > -1, q > 0)$.

- If $p = 0$ then modal values $x_m \in [-1, 1]$ (case of uniform distribution).
- If $p > 0$ then $x_m = 0$.
- If $-1 < p < 0$ then the $EECK(p, q)$ distribution is pseudo bimodal with modes $x_m(-1), x_m(1)$.
- The $f(x; p > 0, q)$ is monotonically increasing on the interval $(-1, 0)$ and monotonically decreasing on the interval $(0, 1)$.

Theorem 3. Let $X \sim EECK(p > -1, q > 0)$.

- If $p = 0$ then modal values $x_m \in [-1, 1]$ (case of uniform distribution).
- If $p > 0$ then $x_m = 0$.
- If $-1 < p < 0$ then the $EECK(p, q)$ distribution is pseudo bimodal with modes $x_m(-1), x_m(1)$.
- The $f(x; p > 0, q)$ is monotonically increasing on the interval $(-1, 0)$ and monotonically decreasing on the interval $(0, 1)$.
- The $f(x; -1 < p < 0, q)$ is monotonically decreasing on the interval $(-1, 0)$ and monotonically increasing on the interval $(0, 1)$.

Theorem 4. Let $X \sim ECK(p > -1, q > 0)$. The inflection points of the $f(x; p, q)$ for $p > 1 \wedge q > 1$ or $-1 < p < 1 \wedge 0 < q < 1$ are given by means of the following formulas

$$x_1 = -\left(\frac{1-q}{1-pq}\right)^{\frac{1}{q}}, x_2 = \left(\frac{1-q}{1-pq}\right)^{\frac{1}{q}}. \tag{3}$$

Theorem 5. Let $X \sim EECK(p > -1, q > 0)$. The u-th $(0 < u < 1)$ quantile x_u is the solution of the following equation

$$(0.5 - u)\, H\,(p, q) + {}_2F_1\left(-p, \frac{1}{q}, 1 + \frac{1}{q}, |x_u|^q\right) x_u = 0, \qquad (4)$$

where ${}_2F_1\,(a, b, c, x)$ is the Gaussian hypergeometric function and $H(p, q)$ is the eta function.

Theorem 5. Let $X \sim EECK(p > -1, q > 0)$. The u-th $(0 < u < 1)$ quantile x_u is the solution of the following equation

$$(0.5 - u)\, H(p, q) + {}_2F_1\left(-p, \frac{1}{q}, 1 + \frac{1}{q}, |x_u|^q\right) x_u = 0, \qquad (4)$$

where ${}_2F_1(a, b, c, x)$ is the Gaussian hypergeometric function and $H(p, q)$ is the eta function.

The proposed distribution is symmetrical then $x_u = -x_{1-u}$, obviously and $x_{0.5} = 0$.

The quantile x_u can be computed by numerical methods.

Theorem 6. The k-th $(k = 0, 1, 2, ...)$ non-central moments of the $EECK(p > -1, q > 0)$ distribution are given by

$$\alpha_k = \frac{\left[1+(-1)^k\right]B\left(\frac{k+1}{q},p+1\right)}{qH(p,q)} = \frac{\left[1+(-1)^k\right]B\left(\frac{k+1}{q},p+1\right)}{2B\left(\frac{1}{q},p+1\right)}$$

Theorem 6. The k-th $(k = 0, 1, 2, ...)$ non-central moments of the $EECK(p > -1, q > 0)$ distribution are given by

$$\alpha_k = \frac{\left[1+(-1)^k\right] B\left(\frac{k+1}{q}, p+1\right)}{qH(p,q)} = \frac{\left[1+(-1)^k\right] B\left(\frac{k+1}{q}, p+1\right)}{2B\left(\frac{1}{q}, p+1\right)}$$

Theorem 7. The non-central moments $\alpha_k (k = 1, 3, ...)$, variance μ_2 and kurtosis $\overline{\gamma}_2$ of the $EECK(p > -1, q > 0)$ distribution are given by

$$\alpha_k = 0\,(k = 1, 3, ...)\,, \mu_2 = \frac{(1+pq)\Gamma\left(\frac{3}{q}\right)\Gamma\left(p+\frac{1}{q}\right)}{(3+pq)\Gamma\left(\frac{1}{q}\right)\Gamma\left(p+\frac{3}{q}\right)}$$

$$\overline{\gamma}_2 = \frac{(pq+3)^2\Gamma\left(\frac{1}{q}\right)\Gamma\left(\frac{5}{q}\right)\Gamma\left(p+\frac{3}{q}\right)^2}{(pq+1)(pq+5)\Gamma\left(p+\frac{1}{q}\right)\Gamma\left(p+\frac{5}{q}\right)\Gamma\left(\frac{3}{q}\right)^2} - 3$$

Figure shows the kurtosis $\bar{\gamma}_2$ as a function of the shape parameter p for $q = 0.4, 0.6, 0.8, 1$ (left) and for $q = 2, 4, 6, 8$ (right). The kurtosis, according to the definition, varies in the range $[-2, \infty)$. The smaller q value, the higher kurtosis and the parameter p has a greater effect on the kurtosis.

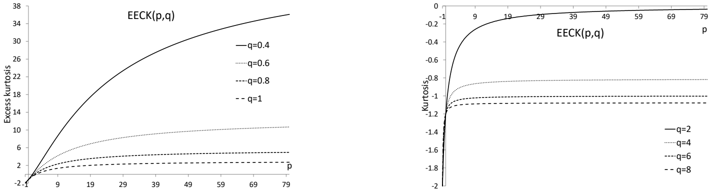


Figure: kurtosis $\bar{\gamma}_2$ as a function of the shape parameter p

Figure shows the kurtosis $\bar{\gamma}_2$ as a function of the shape parameter q for $p = 0.3, 0.5, 0.7, 0.9$ (left) and for $p = 0.25, 0.75, 1, 10$ (right). For $p \in (-1, 0)$ the kurtosis tends from -2 to -1.2 when $q \to \infty$. For $p > 0$ kurtosis tends from ∞ to -1.2 when $q \to \infty$.

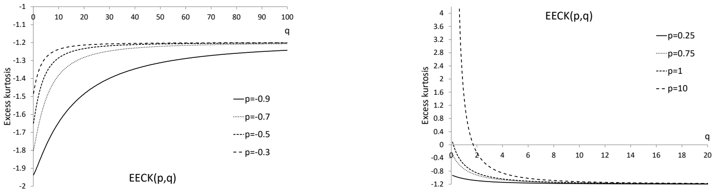
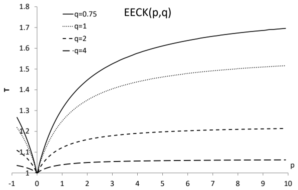


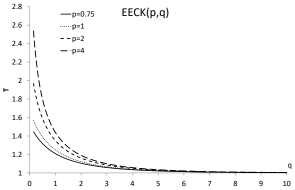
Figure: kurtosis $\bar{\gamma}_2$ as a function of the shape parameter q

Moors proposed a measure based on quantiles in the form

$$T = \frac{x_{7/8} - x_{5/8} + x_{3/8} - x_{1/8}}{x_{6/8} - x_{2/8}}$$

where x_u is the solution of quantile equation. The measure T is a quantile alternative for kurtosis and exists even for distribution for which no moments exist. The $T(p)$ function decreases for $p\,(-1, 0)$ and increases for $p > 0$ mainly for its initial values. The $T(q)$ function tends to one.





Let $X \sim EECK(p > -1, q > 0), R \sim U(0,1)$. The algorithm for generating n values of X, using the inverse CDF method, is as follows:

1. Repeat steps 1.1-1.4 n times:
   1.1 Let $R \sim U(0,1)$,
   1.2 Let $x = -1 + 0.01$,
   1.3 If $CDF(x; p, q) < R$, then $x = x + 0.01$,
   1.4 Return x,

It is obviously a universal algorithm for any distribution with $CDF(x; par)$, where par is the vector of distribution parameters.

The quantile function of the $EECK(p,q)$ does not have an analytical form, PDF is non-negative on the interval $[-1,1]$ and bounded by constant $d=f(0;p\geq 0,q)$, then we can use the von Neumann method, which in this case is much faster than the inverse CDF method.

The quantile function of the $EECK(p, q)$ does not have an analytical form, PDF is non-negative on the interval $[-1, 1]$ and bounded by constant $d = f(0; p \geq 0, q)$, then we can use the von Neumann method, which in this case is much faster than the inverse CDF method.

The algorithm for generating n values of X, using the von Neumann method, is as follows:

1. If $-1 < p < 0$ then use the inverse CDF method
2. If $p \geq 0$ then $d = f(0; p, q)$
3. Repeat steps 3.1-3.3 n times:
   3.1 Let $R_1 \sim U(-1, 1)$, $R_2 \sim U(0, d)$,
   3.2 If $f(R_1; p, q) < R_2$ then goto Step 3.1 else $x = R_1$
   3.3 Return x,

Theorem 9. The Fisher information matrix $I_{i,j}(i,j=1,2)$ for the $EECK(p>-1,q>0)$ distribution is given by

$$I_{11}=$$
$$\left[A-B+\widetilde{H}(p)-\widetilde{H}\left(p+\frac{1}{q}\right)\right]^2+\Psi_1(p+1)-\Psi_1\left(p+\frac{1}{q}+1\right)$$

$$I_{12}=I_{21}=\frac{(A-B)(C-A)}{q^2}-\frac{(A-B)\Gamma\left(p+\frac{1}{q}+1\right)}{\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)}+$$
$$+\frac{(C-A)\left[\widetilde{H}(p)-\widetilde{H}\left(p+\frac{1}{q}\right)\right]}{q^2}+\frac{\Gamma\left(p+\frac{1}{q}+1\right)}{p\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)}$$

$$I_{22}=\frac{(C-A)^2}{q^4}-\frac{2(C-A)\Gamma\left(p+\frac{1}{q}+1\right)}{q^3\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)}+\frac{pq^2(pq+1)\Gamma\left(2-\frac{1}{q}\right)\Gamma\left(p+\frac{1}{q}\right)}{(p-1)(pq-1)\Gamma\left(p-\frac{1}{q}\right)\Gamma\left(\frac{1}{q}\right)}$$

where $\widetilde{H}(z)=\sum_{k=1}^{z}\frac{1}{k}$ is the harmonic function, $\Psi_n(z)$ is the n^{th} derivative of the digamma function

$$\Psi(z),\ A=\Psi\left(p+\frac{1}{q}+1\right),B=\Psi(p+1),C=\Psi\left(\frac{1}{q}+1\right)$$

Let $x_1^*, x_2^*, ..., x_n^*$ be a random sample size n from the $EECK(p > -1, q > 0)$ distribution. Our target is to estimate the unknown values of the parameters p, q.

Let $x_1^*, x_2^*, ..., x_n^*$ be a random sample size n from the $EECK(p > -1, q > 0)$ distribution. Our target is to estimate the unknown values of the parameters p, q.

The likelihood function is given by

$$L = \prod_{i=1}^{n} f\left(x_i^*; p, q\right) = \frac{\Gamma\left(p+\frac{1}{q}+1\right)}{2\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)} \prod_{i=1}^{n} \left(1 - |x_i^*|^q\right)^p$$

then the log-likelihood function is defined as

$$l = n\ln\left[\Gamma\left(p+\frac{1}{q}+1\right)\right] - n\ln\left[2\Gamma\left(p+1\right)\right] - n\ln\left[\Gamma\left(\frac{1}{q}+1\right)\right] + p\sum_{i=1}^{n}\ln\left(1-|x_i^*|^q\right)$$

and

Let $x_1^*, x_2^*, ..., x_n^*$ be a random sample size n from the $EECK(p > -1, q > 0)$ distribution. Our target is to estimate the unknown values of the parameters p, q.
The likelihood function is given by

$$L = \prod_{i=1}^{n} f\left(x_i^*; p, q\right) = \frac{\Gamma\left(p+\frac{1}{q}+1\right)}{2\Gamma(p+1)\Gamma\left(\frac{1}{q}+1\right)} \prod_{i=1}^{n} \left(1 - |x_i^*|^q\right)^p$$

then the log-likelihood function is defined as

$$l = n\ln\left[\Gamma\left(p + \frac{1}{q} + 1\right)\right] - n\ln\left[2\Gamma\left(p+1\right)\right] - n\ln\left[\Gamma\left(\frac{1}{q}+1\right)\right] + p\sum_{i=1}^{n} \ln\left(1 - |x_i^*|^q\right)$$

and

$$\frac{dl}{dp} = n\Psi\left(p + \frac{1}{q} + 1\right) - n\Psi\left(p+1\right) + \sum_{i=1}^{n} \ln\left(1 - |x_i^*|^q\right) = 0$$

$$\frac{dl}{dq} = \frac{-n}{q^2}\Psi\left(p + \frac{1}{q} + 1\right) + \frac{n}{q^2}\Psi\left(\frac{1}{q}+1\right) - \frac{npq\left|x_i^*\right|^{q-1}}{1-\left|x_i^*\right|^q} = 0$$

where Ψ is the digamma function.

The maximum likelihood estimates (MLEs) are solutions of the system equations. We have

$$\frac{1}{n}\sum_{i=1}^{n}\ln\left(1-\left|x_i^*\right|^q\right)=\Psi\left(p+1\right)-\Psi\left(p+\frac{1}{q}+1\right), \tag{5}$$

$$\Psi\left(\frac{1}{q}+1\right)-\Psi\left(p+\frac{1}{q}+1\right)=-\frac{pq^3\left|x_i^*\right|^{q-1}}{1-\left|x_i^*\right|^q}. \tag{6}$$

The maximum likelihood estimates (MLEs) are solutions of the system equations. We have

$$\frac{1}{n}\sum_{i=1}^{n}\ln\left(1-|x_i^*|^q\right)=\Psi\left(p+1\right)-\Psi\left(p+\frac{1}{q}+1\right), \qquad (5)$$

$$\Psi\left(\frac{1}{q}+1\right)-\Psi\left(p+\frac{1}{q}+1\right)=-\frac{pq^3\,|x_i^*|^{q-1}}{1-\left|x_i^*\right|^q}. \qquad (6)$$

Solving this system equations with numerical method we have obtain $\hat{p}, \hat{q}$. We can also maximize the log-likelihood function to obtain the MLEs of the p, q parameters.

The simulation study was performed with 10^3 samples using sample sizes of $100, 150, 200$. The samples were drawn from the $EECK(p, 3)$, where $p = 1, 2, 3$ (see Table left) and from the $EECK(3, q)$, where $q = 1, 2, 3$ (see Table right).

The simulation study was performed with 10^3 samples using sample sizes of $100, 150, 200$. The samples were drawn from the $EECK(p, 3)$, where $p = 1, 2, 3$ (see Table left) and from the $EECK(3, q)$, where $q = 1, 2, 3$ (see Table right).

p	n	$\hat{p}$ Bias	$\hat{p}$ RMSE	$\hat{q}$ Bias	$\hat{q}$ RMSE
1	100	0.555	2.820	0.443	3.593
	150	0.296	1.802	0.186	2.992
	200	0.110	1.172	-0.081	2.279
2	100	0.965	4.408	0.379	2.313
	150	0.724	2.739	0.339	1.908
	200	0.338	1.468	0.110	1.397
3	100	1.255	3.875	0.336	1.701
	150	0.892	3.126	0.269	1.441
	200	0.712	2.289	0.259	1.261

q	n	$\hat{p}$ Bias	$\hat{p}$ RMSE	$\hat{q}$ Bias	$\hat{q}$ RMSE
1	100	0.266	5.510	0.485	3.361
	150	0.011	1.202	0.231	2.169
	200	-0.125	0.320	-0.034	1.290
2	100	0.173	1.168	0.531	3.600
	150	0.046	0.873	0.176	3.234
	200	-0.020	0.711	-0.044	2.729
3	100	0.264	2.584	0.439	5.733
	150	0.149	1.586	0.209	5.283
	200	0.047	1.005	0.029	4.544

As it was mentioned in Introduction, the shape parameter of the EECK distribution cannot be represented as a function of $\overline{\gamma}_2$, as is for the ECK distribution (Sulewski, 2022). Recall, however, that the ECK kurtosis takes values on interval $(-2, 0)$, while the EECK kurtosis has values on interval $[-2, \infty)$. Using e.g. Mathcad, you can easily calculate the argument of a function knowing its value.

As it was mentioned in Introduction, the shape parameter of the EECK distribution cannot be represented as a function of $\overline{\gamma}_2$, as is for the ECK distribution (Sulewski, 2022). Recall, however, that the ECK kurtosis takes values on interval $(-2, 0)$, while the EECK kurtosis has values on interval $[-2, \infty)$. Using e.g. Mathcad, you can easily calculate the argument of a function knowing its value. Let $x_{(1)}, x_{(2)}, ..., x_{(n)}$ be an ordered random sample of size n. Seven GoFTs were selected to be subjects of the Monte Carlo simulation. Five of them as being very popular GoFTs have been implemented in the R software. These tests are: Shapiro-Wilk (SW), Kolmogorov-Smirnov (KS), Cramer-von Mises (CVM), Anderson-Darling (AD) and Shapiro - Francia (SF). Two tests not implemented yet, probably for their novelty, are: H_n (Torabi, 2016) and LF_m (Sulewski, 2020) tests.

The H_n test statistic is defined as

$$H_n = \frac{1}{n}\sum_{i=1}^{n} h\left[\frac{1+\Phi\left(\frac{x_{(i)}-\overline{x}}{s},0,1\right)}{1+\frac{i}{n}}\right], h\left(x\right) = \left(\frac{x-1}{x+1}\right)^2, \quad (7)$$

where $\overline{x}$ and s^2 are the sample mean and sample variance, respectively.

The H_n test statistic is defined as

$$H_n = \frac{1}{n}\sum_{i=1}^{n} h\left[\frac{1+\Phi\left(\frac{x_{(i)}-\overline{x}}{s},0,1\right)}{1+\frac{i}{n}}\right], h(x) = \left(\frac{x-1}{x+1}\right)^2, \qquad (7)$$

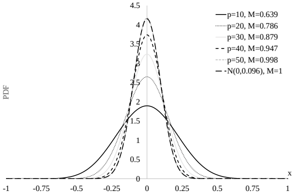
where $\overline{x}$ and s^2 are the sample mean and sample variance, respectively.

The LF_m test statistic is given by

$$LF_m = \max\left|\frac{i-\overline{\alpha}}{n-\overline{\alpha}-\overline{\beta}+1} - \Phi\left(\frac{x_{(i)}-\overline{x}}{s},0,1\right)\right|, (\overline{\alpha},\overline{\beta}\geq 1). \qquad (8)$$

If an alternatively distribution is both symmetric and of negative (positive) kurtosis $\overline{\alpha}=\overline{\beta}=0$ ($\overline{\alpha}=\overline{\beta}=1$) are recommended.

Figure (left) shows PDF of the $N(0, 0.096)$ and $EECK(p, 1.96)$ distributions. For the presented values of the shape parameters, an kurtosis of the EECK is negative. If p increases, the similarity measure also increases. Figure (right) shows PDF of the $N(0, 0.259)$ and $EECK(p, 1.3)$ distributions.



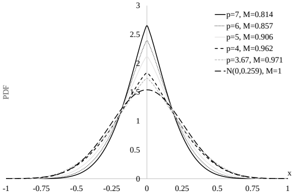


Table 2 (Table 3) shows the modeling of negative (positive) kurtosis, i.e. for a given value of $\overline{\gamma}_2$ of the $EECK(p, 1.96)$ ($EECK(p, 1.3)$) the value of the shape parameter p is caclulated.

	EECK(p,1.96)									N(0,0.096)
$\overline{\gamma}_2$	-1	-0.75	-0.5	-0.4	-0.3	-0.2	-0.1	-0.05	-0.025	0
p	0.49	1.451	3.3	4.627	6.731	10.58	19.882	32.188	45.316	-

Source: Own material.

Table: Modeling of negative kurtosis $\overline{\gamma}_2$. $EECK(p, 1.96)$

	EECK(p,1.3)									N(0,0.259)
$\overline{\gamma}_2$	0.5	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0
p	8.261	6.95	5.891	5.018	4.286	3.963	3.81	3.721	3.693	3.669	-

Source: Own material.

Table: Modeling of positive kurtosis $\overline{\gamma}_2$. $EECK(p, 1.3)$

Phase 1: In this phase the aim is to investigate to what degree selected GoFTs are able to distinct between the normal and proposed distributions. In other words the aim is to determine powers of GoFTs being under discussion when samples come from $EECK(p,q)$ general populations. Critical values $cv_{0.05}$ ascribed to GoFTs (where $\alpha = 0.05$ is the test significance level) were estimated with the Monte Carlo method.

n	20		40		60	
s	0.096	0.259	0.096	0.259	0.096	0.259
LF	0.19177		0.13841		0.11385	
CVM	0.12278		0.12445		0.12490	
AD	0.72300		0.73751		0.74215	
SW	0.98287		0.98860		0.99140	
SF	0.98464		0.99003		0.99248	
H_n	0.00077		0.00038		0.00025	
LF_m	0.16195		0.12388		0.10450	

GoFT		$\bar{\gamma}_2$									
		-1	-0.75	-0.5	-0.4	-0.3	-0.2	-0.1	-0.05	-0.025	0
		p									
	n	0.49	1.451	3.3	4.627	6.731	10.58	19.882	32.188	45.316	-
LF	20	0.063	0.046	0.044	0.045	0.045	0.048	0.048	0.049	0.050	0.050
	40	0.099	0.058	0.048	0.045	0.047	0.047	0.049	0.049	0.050	0.050
	60	0.148	0.074	0.052	0.049	0.047	0.047	0.049	0.051	0.050	0.051
CVM	20	0.074	0.047	0.044	0.042	0.044	0.047	0.047	0.049	0.050	0.050
	40	0.144	0.069	0.049	0.045	0.045	0.046	0.048	0.049	0.049	0.050
	60	0.237	0.095	0.055	0.049	0.046	0.046	0.049	0.049	0.049	0.051
AD	20	0.079	0.047	0.041	0.040	0.042	0.045	0.047	0.048	0.050	0.050
	40	0.178	0.075	0.048	0.043	0.044	0.044	0.047	0.049	0.048	0.050
	60	0.311	0.109	0.057	0.048	0.045	0.045	0.048	0.049	0.048	0.050
SW	20	0.083	0.043	0.036	0.038	0.039	0.041	0.045	0.048	0.048	0.049
	40	0.223	0.071	0.040	0.036	0.036	0.038	0.043	0.046	0.049	0.051
	60	0.429	0.115	0.047	0.039	0.037	0.037	0.043	0.045	0.046	0.051
SF	20	0.034	0.022	0.025	0.030	0.033	0.039	0.045	0.049	0.050	0.049
	40	0.083	0.025	0.019	0.021	0.025	0.032	0.041	0.045	0.050	0.052
	60	0.195	0.040	0.019	0.020	0.023	0.028	0.040	0.045	0.047	0.051
H_n	20	0.075	0.049	0.043	0.044	0.044	0.046	0.047	0.048	0.049	0.049
	40	0.154	0.074	0.051	0.046	0.047	0.046	0.048	0.049	0.049	0.051
	60	0.259	0.105	0.059	0.053	0.049	0.050	0.051	0.051	0.050	0.053
LF_m	20	0.082	0.056	0.050	0.049	0.048	0.050	0.049	0.049	0.051	0.051
	40	0.125	0.073	0.054	0.050	0.051	0.049	0.051	0.050	0.050	0.051
	60	0.181	0.087	0.059	0.053	0.051	0.049	0.050	0.051	0.050	0.051

Table: Powers of tests at $\alpha = 0.05$, when the $EECK(p, 1.96)$ is the actual population distribution. The case of negative kurtosis values

GoFT	$\bar{\gamma}_2$	0.5	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0
	p											
	n	8.261	6.95	5.891	5.018	4.286	3.963	3.810	3.721	3.693	3.669	-
LF	20	0.072	0.069	0.064	0.060	0.055	0.056	0.053	0.054	0.054	0.053	0.050
	40	0.089	0.081	0.074	0.070	0.061	0.060	0.058	0.056	0.055	0.057	0.051
	60	0.105	0.094	0.084	0.075	0.065	0.062	0.060	0.060	0.060	0.060	0.052
CVM	20	0.081	0.077	0.071	0.063	0.059	0.057	0.056	0.056	0.054	0.055	0.052
	40	0.101	0.091	0.080	0.073	0.064	0.060	0.059	0.056	0.056	0.056	0.049
	60	0.122	0.108	0.093	0.082	0.068	0.063	0.062	0.061	0.060	0.061	0.051
AD	20	0.082	0.077	0.071	0.062	0.057	0.054	0.055	0.054	0.052	0.053	0.052
	40	0.101	0.090	0.079	0.071	0.062	0.058	0.056	0.054	0.053	0.053	0.049
	60	0.124	0.107	0.092	0.081	0.066	0.061	0.060	0.058	0.057	0.059	0.051
SW	20	0.081	0.073	0.067	0.060	0.052	0.051	0.049	0.048	0.047	0.050	0.050
	40	0.098	0.085	0.074	0.060	0.052	0.047	0.045	0.044	0.044	0.048	0.050
	60	0.114	0.095	0.079	0.064	0.050	0.045	0.044	0.042	0.042	0.046	0.051
SF	20	0.102	0.092	0.081	0.072	0.062	0.058	0.055	0.055	0.052	0.057	0.049
	40	0.127	0.111	0.093	0.074	0.061	0.053	0.049	0.048	0.049	0.053	0.049
	60	0.148	0.125	0.102	0.078	0.058	0.050	0.047	0.045	0.044	0.051	0.052
H_n	20	0.078	0.073	0.068	0.061	0.058	0.056	0.055	0.054	0.053	0.055	0.052
	40	0.094	0.084	0.076	0.069	0.061	0.058	0.056	0.055	0.054	0.054	0.049
	60	0.118	0.105	0.091	0.080	0.067	0.063	0.061	0.060	0.059	0.061	0.053
LF_m	20	0.082	0.078	0.073	0.066	0.061	0.060	0.057	0.057	0.057	0.057	0.050
	40	0.100	0.091	0.082	0.076	0.066	0.064	0.062	0.060	0.059	0.060	0.050
	60	0.116	0.104	0.092	0.081	0.069	0.066	0.064	0.064	0.063	0.063	0.051

Table: Powers of tests at $\alpha = 0.05$, when the $EECK(p, 1.3)$ is the actual population distribution. The case of positive kurtosis values

The conclusions from tables are very interesting.

The conclusions from tables are very interesting.

- For $n = 20$, the LF, CvM, AD, SW, Hn tests detect only $\overline{\gamma}_2 = -1$, LFm - $\overline{\gamma}_2 = -0.75$; LFm, SF tests detect even $\overline{\gamma}_2 = 0.001$.

The conclusions from tables are very interesting.

- For $n = 20$, the LF, CvM, AD, SW, Hn tests detect only $\overline{\gamma}_2 = -1$, LFm - $\overline{\gamma}_2 = -0.75$; LFm, SF tests detect even $\overline{\gamma}_2 = 0.001$.
- For $n = 40$, the LF, CvM, AD, SW, Hn and LFm tests detect only $\overline{\gamma}_2 = -0.75$; LF, CvM, AD, Hn, and LFm tests detect even $\overline{\gamma}_2 = 0.001$.

The conclusions from tables are very interesting.

- For $n = 20$, the LF, CvM, AD, SW, Hn tests detect only $\overline{\gamma}_2 = -1$, LFm - $\overline{\gamma}_2 = -0.75$; LFm, SF tests detect even $\overline{\gamma}_2 = 0.001$.
- For $n = 40$, the LF, CvM, AD, SW, Hn and LFm tests detect only $\overline{\gamma}_2 = -0.75$; LF, CvM, AD, Hn, and LFm tests detect even $\overline{\gamma}_2 = 0.001$.
- For n=60, the AD, Hn and LFm tests detect only $\overline{\gamma}_2 = -0.5$; LF and CVM tests detect only $\overline{\gamma}_2 = -0.75$; LF, CvM, AD, Hn, and LFm tests detect even $\overline{\gamma}_2 = 0.001$.

In Phase 1, we showed that the considered GoFTs detect positive kurtosis better than negative one.

Phase 2. In this phase the aim is to investigate to what degree an undetected kurtosis impacts the performance of two basic tests related to parameters of the Normal distribution, namely Student t test and Fisher – Snedecor F test.

Phase 2. In this phase the aim is to investigate to what degree an undetected kurtosis impacts the performance of two basic tests related to parameters of the Normal distribution, namely Student t test and Fisher – Snedecor F test.

Let $x_{1,1}, x_{1,2}, ..., x_{1,n}$ and $x_{2,1}, x_{2,2}, ..., x_{2,n}$ be two samples of sizes n drawn from particular general populations. Let us remember that t and F test statistics have the following forms:

$$\dot{t} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_{x1}^2 + s_{x2}^2}{n}}}, \dot{F} = \frac{s_{x1}^2}{s_{x2}^2}, \tag{9}$$

where $\overline{x_1}, \overline{x_2}$ are the sample means and s_{x1}, s_{x2} are the sample standard deviations.

The course of action was as follows:

The course of action was as follows:

1. $m = 10^5$ pairs of samples both of size $n = 60$ were drawn from $EECK(4, 627, 1.96)$ (for negative kurtosis) and $EECK(3.669, 1.3)$ (for positive kurtosis) general populations.

The course of action was as follows:

1. $m = 10^5$ pairs of samples both of size $n = 60$ were drawn from $EECK(4, 627, 1.96)$ (for negative kurtosis) and $EECK(3.669, 1.3)$ (for positive kurtosis) general populations.
2. These pairs of samples were consecutively, converted into pairs of $\dot{t}_v$ statistics and $\dot{F}_v$ statistics, $v = 1, 2, ..., m$.

The course of action was as follows:

1. $m = 10^5$ pairs of samples both of size $n = 60$ were drawn from $EECK(4, 627, 1.96)$ (for negative kurtosis) and $EECK(3.669, 1.3)$ (for positive kurtosis) general populations.
2. These pairs of samples were consecutively, converted into pairs of $\dot{t}_v$ statistics and $\dot{F}_v$ statistics, $v = 1, 2, ..., m$.
3. Sets of values of $\dot{t}_v$ and $\dot{F}_v$ statistics were stored in two matrices named T and F.

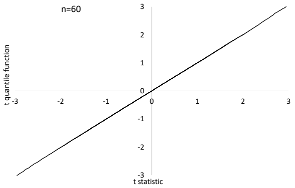
The course of action was as follows:

1. $m = 10^5$ pairs of samples both of size $n = 60$ were drawn from $EECK(4, 627, 1.96)$ (for negative kurtosis) and $EECK(3.669, 1.3)$ (for positive kurtosis) general populations.
2. These pairs of samples were consecutively, converted into pairs of $\dot{t}_v$ statistics and $\dot{F}_v$ statistics, $v = 1, 2, ..., m$.
3. Sets of values of $\dot{t}_v$ and $\dot{F}_v$ statistics were stored in two matrices named T and F.
4. The matrices were sorted in ascending order and served to determine two empirical CDFs namely $\Theta_t\left(\dot{t}_v\right)$ and $\Theta_F\left(\dot{F}_v\right)$.

The course of action was as follows:

1. $m = 10^5$ pairs of samples both of size $n = 60$ were drawn from $EECK(4, 627, 1.96)$ (for negative kurtosis) and $EECK(3.669, 1.3)$ (for positive kurtosis) general populations.
2. These pairs of samples were consecutively, converted into pairs of $\dot{t}_v$ statistics and $\dot{F}_v$ statistics, $v = 1, 2, ..., m$.
3. Sets of values of $\dot{t}_v$ and $\dot{F}_v$ statistics were stored in two matrices named T and F.
4. The matrices were sorted in ascending order and served to determine two empirical CDFs namely $\Theta_t\left(\dot{t}_v\right)$ and $\Theta_F\left(\dot{F}_v\right)$.
5. Probability papers were employed to check whether the above empirical CDFs fit the Student and Fisher-Snedecor distributions.

Figures show empirical CDFs plotted on the Student and Snedecor probability papers, when samples were drawn from $EECK(4, 627, 1.96)$ and $EECK(3.669, 1.3)$, appropriately. These probability papers were constructed in the same way as the Normal probability is constructed.



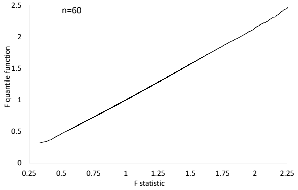
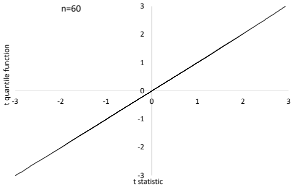


Figure: Empirical CDFs plotted on the Student and Snedecor probability paper. Case of negative kurtosis values



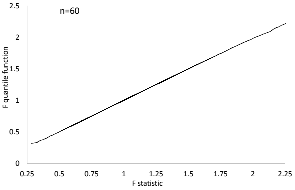


Figure: Empirical CDFs plotted on the Student and Snedecor probability paper. Case of positive kurtosis values

It turns out that the empirical distribution in question perfectly fit straight lines that relevant theoretical distributions. Thus, we can conclude that Student and Fisher-Snedecor tests may be applied even as population distributions are of negative or positive kurtosis.

Symmetric distributions have limited use in fitting the distributions to data (e.g. normal distribution). However, the situation looks much better when we use their mixture (e.g. compound normal distribution).

Symmetric distributions have limited use in fitting the distributions to data (e.g. normal distribution). However, the situation looks much better when we use their mixture (e.g. compound normal distribution).

For the purposes of this subsection, we extend the domain of the $EECK(p,q)$ from $[-1,1]$ to $[-a,a] (a \in R)$. PDF of the modified $EECK(p,q)$ distribution denoted as $EECK2(x,a,p,q)$ has the form

$$EECK2\left(x;a,p,q\right)=\frac{\int_{-a}^{a}\left[1-\left(\frac{|x|}{a}\right)^{q}\right]^{p}dx}{2\int_{0}^{a}\left[1-\left(\frac{|u|}{a}\right)^{q}\right]^{p}du}$$

$$x\in\begin{cases}(-a,a) & \text{if } -1<p<0\\ [-a,a] & \text{if } p\geq 0\end{cases}$$

We present real data examples to demonstrate a flexibility of the $EECK(p > -1, q > 0)$ distribution in the mixed variant. PDF of the compound EECK (CEECK) distribution is given by

$$CEECK\left(x; a, p_1, q_1, p_2, q_2, \omega\right) = \\ \omega EECK2\left(x; a, p_1, q_1\right) + \left(1 - \omega\right) EECK2\left(x; a, p_2, q_2\right)$$

We present real data examples to demonstrate a flexibility of the $EECK(p > -1, q > 0)$ distribution in the mixed variant. PDF of the compound EECK (CEECK) distribution is given by

$$CEECK\left(x; a, p_1, q_1, p_2, q_2, \omega\right) = \\ \omega EECK2\left(x; a, p_1, q_1\right) + \left(1 - \omega\right) EECK2\left(x; a, p_2, q_2\right)$$

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.

We present real data examples to demonstrate a flexibility of the $EECK(p > -1, q > 0)$ distribution in the mixed variant. PDF of the compound EECK (CEECK) distribution is given by

$$CEECK\left(x; a, p_1, q_1, p_2, q_2, \omega\right) = \omega EECK2\left(x; a, p_1, q_1\right) + (1 - \omega)\, EECK2\left(x; a, p_2, q_2\right)$$

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.
The KS GoFT was used for model fitting, while the AIC, BIC and HQIC were used for model comparisons.

The first data set presents temperature dynamics of beaver Castor canadensis in north-central Wisconsin (Reynolds, 1994). Body temperature was measured by telemetry every 10 minutes from one period of less than a day. The data consists of 114 observations of the variable measured body temperature in degrees Celsius and are available in the R software with code beaver1[3].

The first data set presents temperature dynamics of beaver Castor canadensis in north-central Wisconsin (Reynolds, 1994). Body temperature was measured by telemetry every 10 minutes from one period of less than a day. The data consists of 114 observations of the variable measured body temperature in degrees Celsius and are available in the R software with code beaver1[3].
The second data set contains statistics, in arrests per 100,000 residents for assault in each of the 50 US states in 1973 (McNeil, 1977). The data consisting of 50 observations are available in the R software with code USArrests[2].

The models selected for comparison with the CEECK are:

The models selected for comparison with the CEECK are:

- the compound ECK (CECK):

$$f_{CECK}(x) = \omega \frac{\left(1-\frac{x^2}{a^2}\right)^{p_1}}{aB(0.5,p_1+1)} + (1-\omega)\frac{\left(1-\frac{x^2}{a^2}\right)^{p_2}}{aB(0.5,p_2+1)}$$

The models selected for comparison with the CEECK are:

- the compound ECK (CECK):

$$f_{CECK}(x) = \omega \frac{\left(1-\frac{x^2}{a^2}\right)^{p_1}}{aB(0.5,p_1+1)} + (1-\omega)\frac{\left(1-\frac{x^2}{a^2}\right)^{p_2}}{aB(0.5,p_2+1)}$$

- the compound normal (CN):

$$f_{CN}(x) = \omega\phi(x;a_1,b_1) + (1-\omega)\phi(x;a_2,b_2)$$

The models selected for comparison with the CEECK are:

- the compound ECK (CECK):

$$f_{CECK}(x) = \omega \frac{\left(1-\frac{x^2}{a^2}\right)^{p_1}}{aB(0.5,p_1+1)} + (1-\omega)\frac{\left(1-\frac{x^2}{a^2}\right)^{p_2}}{aB(0.5,p_2+1)}$$

- the compound normal (CN):

$$f_{CN}(x) = \omega\phi(x; a_1, b_1) + (1-\omega)\phi(x; a_2, b_2)$$

- the compound Laplace (CL):

$$f_{CL}(x) = \frac{\omega}{2b_1}\exp\left[\exp\left(-\frac{|x-a_1|}{b_1}\right)\right] + \frac{1-\omega}{2b_2}\exp\left[\exp\left(-\frac{|x-a_2|}{b_2}\right)\right],$$

The models selected for comparison with the CEECK are:

- the compound ECK (CECK):

$$f_{CECK}(x) = \omega \frac{\left(1-\frac{x^2}{a^2}\right)^{p_1}}{aB(0.5,p_1+1)} + (1-\omega)\frac{\left(1-\frac{x^2}{a^2}\right)^{p_2}}{aB(0.5,p_2+1)}$$

- the compound normal (CN):

$$f_{CN}(x) = \omega\phi(x; a_1, b_1) + (1-\omega)\phi(x; a_2, b_2)$$

- the compound Laplace (CL):

$$f_{CL}(x) = \frac{\omega}{2b_1}\exp\left[\exp\left(-\frac{|x-a_1|}{b_1}\right)\right] + \frac{1-\omega}{2b_2}\exp\left[\exp\left(-\frac{|x-a_2|}{b_2}\right)\right],$$

- the compound Cauchy (CC):

$$f_{CC}(x) = \frac{\omega}{\pi b_1\left[1+\left(\frac{x-a_1}{b_1}\right)^2\right]} + \frac{1-\omega}{\pi b_2\left[1+\left(\frac{x-a_2}{b_2}\right)^2\right]}$$

The models selected for comparison with the CEECK are:

- the compound ECK (CECK):

$$f_{CECK}(x) = \omega \frac{\left(1-\frac{x^2}{a^2}\right)^{p_1}}{aB(0.5,p_1+1)} + (1-\omega)\frac{\left(1-\frac{x^2}{a^2}\right)^{p_2}}{aB(0.5,p_2+1)}$$

- the compound normal (CN):

$$f_{CN}(x) = \omega\phi(x; a_1, b_1) + (1-\omega)\phi(x; a_2, b_2)$$

- the compound Laplace (CL):

$$f_{CL}(x) = \frac{\omega}{2b_1}\exp\left[\exp\left(-\frac{|x-a_1|}{b_1}\right)\right] + \frac{1-\omega}{2b_2}\exp\left[\exp\left(-\frac{|x-a_2|}{b_2}\right)\right],$$

- the compound Cauchy (CC):

$$f_{CC}(x) = \frac{\omega}{\pi b_1\left[1+\left(\frac{x-a_1}{b_1}\right)^2\right]} + \frac{1-\omega}{\pi b_2\left[1+\left(\frac{x-a_2}{b_2}\right)^2\right]}$$

- the compound logistic (CLOG):

$$\omega\exp\left(\frac{x-a_1}{\ }\right) \quad (1-\omega)\exp\left(\frac{x-a_2}{\ }\right)$$

Model	MLEs	-l	AIC	BIC	HQIC	KS(p-value)
CEECK	$\widehat{a}$=4.151,$\widehat{p}_1$=2.039,$\widehat{q}_1$=0.700,$\widehat{p}_2$=8958.252,$\widehat{q}_2$=5.371,$\widehat{\omega}$=0.660	154.785	321.570	337.987	328.233	**0.040(0.978)**
CECK	$\widehat{a}$=5.010,$\widehat{p}_1$=5.413,$\widehat{p}_2$=55.361,$\widehat{\omega}$=0.484	155.1654	318.331	329.275	322.773	0.041(0.964)
CN	$\widehat{a}_1$=-2.700,$\widehat{b}_1$=0.042,$\widehat{a}_2$=0.071,$\widehat{b}_2$=0.906,$\widehat{\omega}$=0.026	154.745	319.489	333.170	325.041	0.084(0.324)
CL	$\widehat{a}_1$=-0.580,$\widehat{b}_1$=0.722,$\widehat{a}_2$=0.144,$\widehat{b}_2$=0.663,$\widehat{\omega}$=0.201	155.393	320.786	334.467	326.338	0.087(0.277)
CC	$\widehat{a}_1$=-0.581,$\widehat{b}_1$=0.353,$\widehat{a}_2$=0.226,$\widehat{b}_2$=0.369,$\widehat{\omega}$=0.294	163.386	336.771	350.452	342.324	0.103(0.142)
CLOG	$\widehat{a}_1$=0.045,$\widehat{b}_1$=0.492,$\widehat{a}_2$=-2.700,$\widehat{b}_2$=0.026,$\widehat{\omega}$=0.975	**152.318**	**314.636**	**328.317**	**320.188**	0.050(0.877)

Source: Own material.

Table: Results of estimation for the first data set

Model	MLEs	-l	AIC	BIC	HQIC	KS(p-value)
CEECK	$\widehat{a}$=4.151,$\widehat{p}_1$=2.039,$\widehat{q}_1$=0.700,$\widehat{p}_2$=8958.252,$\widehat{q}_2$=5.371,$\widehat{\omega}$=0.660	154.785	321.570	337.987	328.233	**0.040(0.978)**
CECK	$\widehat{a}$=5.010,$\widehat{p}_1$=5.413,$\widehat{p}_2$=55.361,$\widehat{\omega}$=0.484	155.1654	318.331	329.275	322.773	0.041(0.964)
CN	$\widehat{a}_1$=-2.700,$\widehat{b}_1$=0.042,$\widehat{a}_2$=0.071,$\widehat{b}_2$=0.906,$\widehat{\omega}$=0.026	154.745	319.489	333.170	325.041	0.084(0.324)
CL	$\widehat{a}_1$=-0.580,$\widehat{b}_1$=0.722,$\widehat{a}_2$=0.144,$\widehat{b}_2$=0.663,$\widehat{\omega}$=0.201	155.393	320.786	334.467	326.338	0.087(0.277)
CC	$\widehat{a}_1$=-0.581,$\widehat{b}_1$=0.353,$\widehat{a}_2$=0.226,$\widehat{b}_2$=0.369,$\widehat{\omega}$=0.294	163.386	336.771	350.452	342.324	0.103(0.142)
CLOG	$\widehat{a}_1$=0.045,$\widehat{b}_1$=0.492,$\widehat{a}_2$=-2.700,$\widehat{b}_2$=0.026,$\widehat{\omega}$=0.975	**152.318**	**314.636**	**328.317**	**320.188**	0.050(0.877)

Source: Own material.

Table: Results of estimation for the first data set

Model	MLEs	-l	AIC	BIC	HQIC	KS(p-value)
CEECK	$\widehat{a}$=2.040,$\widehat{p}_1$=200.766,$\widehat{q}_1$=385.145,$\widehat{p}_2$=378.908,$\widehat{q}_2$=23.014,$\widehat{\omega}$=0.203	**63.951**	**139.901**	**151.373**	**144.270**	0.088(0.687)
CECK	$\widehat{a}$=9.802,$\widehat{p}_1$=32.874,$\widehat{p}_2$=63.054,$\widehat{\omega}$=0.356	70.964	149.929	157.577	152.841	0.136(0.215)
CN	$\widehat{a}_1$=-0.637,$\widehat{b}_1$=0.567,$\widehat{a}_2$=1.089,$\widehat{b}_2$=0.473,$\widehat{\omega}$=0.631	66.001	142.002	151.562	145.643	0.072(0.853)
CL	$\widehat{a}_1$=0.999,$\widehat{b}_1$=0.415,$\widehat{a}_2$=-0.693,$\widehat{b}_2$=0.473,$\widehat{\omega}$=0.392	66.965	143.930	153.490	147.570	**0.061(0.944)**
CC	$\widehat{a}_1$=-0.650,$\widehat{b}_1$=0.404,$\widehat{a}_2$=1.025,$\widehat{b}_2$=0.223,$\widehat{\omega}$=0.655	73.122	156.244	165.805	159.885	0.081(0.767)
CLOG	$\widehat{a}_1$=-0.617,$\widehat{b}_1$=0.355,$\widehat{a}_2$=1.099,$\widehat{b}_2$=0.262,$\widehat{\omega}$=0.648	66.684	143.369	152.929	147.009	0.072(0.880)

Source: Own material.

Table: Results of estimation for the second data set

Figure 7 presents histograms, estimated PDFs of the analyzed models. The CEECK model is distinguished in terms of the KS GoFT for the first data set. This model has the smallest values of the $-l$, AIC, BIC and HQIC.

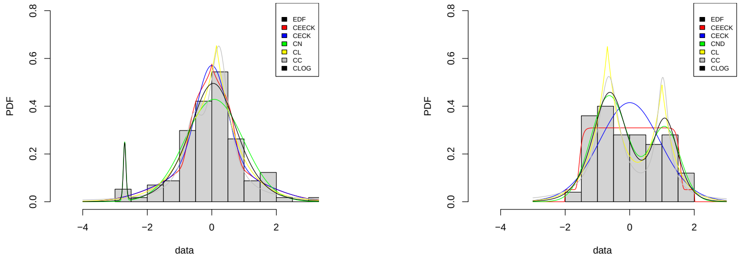


Figure: Histograms and estimated PDF of analyzed models for first (left) and second (right) data sets

- The article presents the extended easily changeable kurtosis (EECK) distribution, the special cases of which are the ECK, uniform and triangle distributions.

- The article presents the extended easily changeable kurtosis (EECK) distribution, the special cases of which are the ECK, uniform and triangle distributions.
- The new proposal tends to the normal distribution.

- The article presents the extended easily changeable kurtosis (EECK) distribution, the special cases of which are the ECK, uniform and triangle distributions.
- The new proposal tends to the normal distribution.
- The EECK, like the ECK, belongs to the family of symmetric, unimodal distributions, defined in the finite domain with kurtosis values on infinite interval.

- The article presents the extended easily changeable kurtosis (EECK) distribution, the special cases of which are the ECK, uniform and triangle distributions.
- The new proposal tends to the normal distribution.
- The EECK, like the ECK, belongs to the family of symmetric, unimodal distributions, defined in the finite domain with kurtosis values on infinite interval.
- the new proposal can be extremely useful when we want to seamlessly test GoFT's ability to detect deviations from normality by modeling of negative or positive kurtosis.

- The article presents the extended easily changeable kurtosis (EECK) distribution, the special cases of which are the ECK, uniform and triangle distributions.
- The new proposal tends to the normal distribution.
- The EECK, like the ECK, belongs to the family of symmetric, unimodal distributions, defined in the finite domain with kurtosis values on infinite interval.
- the new proposal can be extremely useful when we want to seamlessly test GoFT's ability to detect deviations from normality by modeling of negative or positive kurtosis.
- Student and Fisher-Snedecor tests may be applied even as population distributions are of negative or positive kurtosis.

- Real data example demonstrates that the $EECK(p, q)$ distribution in the mixed variant is flexible and competitive model that deserves to be added to the existing distributions in data modeling.

- Real data example demonstrates that the $EECK(p, q)$ distribution in the mixed variant is flexible and competitive model that deserves to be added to the existing distributions in data modeling.
- The information presented in the article shows that the proposed distribution deserves to be added to the symmetric distribution family.

[1] AZZALINI, A. (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **2(12)**, 171—178.

[2] BIRNBAUM, Z,W. and SAUNDERS, S.C. (1969). A new family of life distributions, *Journal of Applied Probability*, **6(2)**, 319—327.

[3] BOLFARINE, H.; MARTÍNEZ-FLOREZ, G. and SALINAS, H.S. (2018). Bimodal symmetric-asymmetric power-normal families, *Communications in Statistics-Theory and Methods*, **47(2)**, 259–276.

[4] CORTES, M.A., ELAL-OLIVERO, D. and OLIVARES-PACHECO, J.F. (2018). A new class of distributions generated by the extended bimodal-normal distribution, *Hindawi Journal of Probability and Statistics*, **2018**, 1–10.

[5] GIBBONS, J.F. and MYLROIE, S. (1985). Estimation of impurity profiles in ion-implanted amorphous targets using joined half-Gaussian distributions, *Applied Physics Letters*, **11(22)**, 568–569.

[6] GRADSHTEYN, I.S. and RYZHIK, I.M. (2014). Table of integrals, series, and products, *Academic press.*

[7] HASSAN, M.Y. and EL-BASSIOUNI, M.Y. (2016). Bimodal skew-symmetric normal distribution, *Communications in Statistics-Theory and Methods*, **45(5)**,1527–1541.

[8] HASSAN, M.Y. and HIJAZI, R.H. (2010). A bimodal exponential power distribution, *Pakistan Journal of Statistics*, **2**(26), 379–396.

[9] MOORS, J.J.A. (1988). A quantile alternative for kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)*, **1(37)**,25–32.

[10] MCNEIL, D. R. (1977). Interactive Data Analysis, *New York, Wiley.*

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

[12] Reynolds, P.S. (1994). Time-series analysis of beaver body temperatures. Chapter 11 of Lange, N., Billard, L., Conquest, L., Ryan, L., Brillinger, D. and Greenhouse J. (Eds.), Case Studies in Biometry, *Proc. Berekeley Symposium, Statist. Probability, Wiley, New York.*

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

[14] Sulewski, P. (2021). Two-piece power normal distribution, *Communications in Statistics - Theory and Methods*, **50(11)**, 2619–2639.

[15] Sulewski, P. (2020). Modified Lilliefors goodness-of-fit test for normality, *Communications in Statistics - Simulation and Computation*, **51(3)**, 1199–1219.

[16] Sulewski, P. (2022). Normal distribution with plasticizing component, *Communications in Statistics – Theory and Method*, **51(11)**, 3806–3835.

[17] Sulewski, P. (2022). Easily Changeable Kurtosis Distribution, *Austrian Journal of Statistics*, https://www.ajs.or.at/index.php/ajs/article/view/1434.

[18] 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.

[19] Torabi, H.; Montazeri, N.H. and Grane, A. (1961). A test of normality based on the empirical distribution function, *SORT*, **1(40)**, 1, 55–88.

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

[21] Wolfram (1988). Hypergeometric0F1, Wolfram Language function, *https://reference.wolfram.com/language/ref/Hypergeometric0F1.html.*

[22] Von Neumann, J. (1951). Various techniques used in connection with random digits, *National Bureau of Standards Applied Mathematics Series*, **12**, 36–38.


THANK YOU