General Mathematics Vol. 29, No. 1 (2021), 13–29
DOI:10.2478/gm-2021-0002

sciendo

Epanechnikov-exponential distribution: properties and applications [1]

Ahmad Alkhazaalh, Loai Al-Zoubi

Abstract

A new continuous distribution is proposed using Epanechnikov kernel function and the exponential distribution. This distribution is called the Epanechnikov-exponential distribution. Some properties of this distribution are studied. A simulation study is conducted to calculate the mean and the standard deviation of this distribution and to investigate the behavior of MLE to conserve the consistency property. An application to a real data set is conducted, it showed that he new distribution is more flexible than the exponential distribution.

2010 Mathematics Subject Classification: 62-XX, 62Exx.
Key words and phrases: Epanechnikov exponential distribution, Epanechnikov distribution, Moments, Skewness, Kurtosis, Entropy, Order statistics, quantile function.

1 Introduction

Kernel functions are used and applied in theory of functions that achieve, in a fixed domain, specific differential equation. One of these functions is the Epanechnikov kernel function (EKF) [11] which is defined as:

$$k(u) = \frac{3}{4} max\{1 - u^2, 0\} \tag{1}$$

Epanechnikov kernel function is a continuous function and has the least mean square error (MSE). It is a non-negative function and satisfies the normalization $\left(\int_{-\infty}^{\infty} k(u)du = 1 \right)$ and the symmetric $(k(u) = k(-u))$ properties. Thus, from the normalization property we can conclude that the EKF is a probability density function (pdf) [32].

[1] *Received 28 September, 2020*
Accepted for publication (in revised form) 23 November, 2020

Many researchers used the distribution theory to generalize some distributions. [2] proposed the transmuted Janardan distribution using the quadratic transmutation map developed by [25]. The same map is used to generalize Burr type XII distribution by [17], the two-parameter Lindley distribution [1], the New-Weibull Pareto Distribution by [4], Ishita distribution [13]. [7] generalized two distributions through her master thesis, namely: the transmuted Gamma-Gompertz and transmuted Generalized Type-II Half-Logistics Distributions. [6] worked out the transmuted reciprocal and the transmuted two-parameter weighted exponential distributions. Whereas [20] used the quadratic transmutation map to generalize the power function and Type-I half logistic distributions. [5] used the same map to develop Mukherjee-Islam distribution into transmuted Mukherjee-Islam distribution. [3] used the Topp-Leone family of distributions to generate the Topp-Leone Mukherjee-Islam distribution. They studied its properties and applied it to some data sets. [31] used the quadratic transmutation map to propose the transmuted Aradhana distribution. In this paper we will use the EKF and the exponential distribution to generate a new distribution called the Epanechnikov-exponential distribution (EED).

2 Epanechnikov-Exponential Distribution

The exponential distribution (ED) is the probability distribution of the time between events in a Poisson point process. The probability density and cumulative distribution functions of the exponential random variable ($X \sim Exp(\lambda)$) are defined by [16] as:

$$f(x) = \lambda e^{-\lambda x}, x > 0, \quad \lambda > 0, \tag{2}$$

$$F(x) = 1 - e^{-\lambda x}, x > 0 \tag{3}$$

The mean of ED is $\mu = E(X) = \frac{1}{\lambda}$, the variance is $\sigma^2 = var(X) = \frac{1}{\lambda^2}$, skewness = 2 and kurtosis = 6.

The Epanechnikov density function is defined by [32] as:

$$k(x) = \frac{3}{4}\big(1 - x^2\big), |x| \leq 1, \tag{4}$$

Using the cumulative distribution function technique of transformation, the CDF and the pdf of the Epanechnikov-exponential distribution (EED) are given by the following theorem.

Theorem 1 *The random variable X is said to have an EED if its CDF and pdf are, respectively; given by:*

$$G(x) = \frac{1}{2}\big(e^{-3\lambda x} - 3e^{-2\lambda x}\big) + 1, x \geq 0, \tag{5}$$

$$g(x) = \frac{3\lambda}{2}\big(2e^{-2\lambda x} - e^{-3\lambda x}\big), x > 0. \tag{6}$$

Proof.

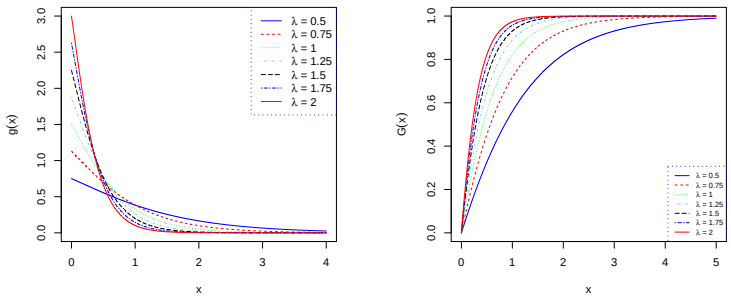
$$\begin{aligned} G(x) &= 2\int_0^{F(x)} k(u)du \\ &= 2\int_0^{1-e^{-\lambda x}} \frac{3}{4}(1-u^2)du \\ &= \frac{3}{2}\Big(u-\frac{1}{3}u^3\Big)\Big|_0^{1-e^{-\lambda x}} \\ &= \frac{\mathrm{e}^{-3\lambda x}-3\mathrm{e}^{-2\lambda x}}{2}+1. \end{aligned}$$

Deriving $G(x)$ with respect to x, we get

$$g(x) = \frac{3\lambda}{2}\big(2e^{-2\lambda x}-e^{-3\lambda x}\big)$$

(a) The pdf of EED for different values of λ.

(b) The cdf of EED for different values of λ.



3 Reliability Analysis

The reliability function is defined to be the probability that an item′s life lasts longer than t units [26]. Therefore, mathematically it is defined as:

$$R(t) = p(T \geq t) = 1 - G(t)$$

Thus, the reliability of the EED is defined as:

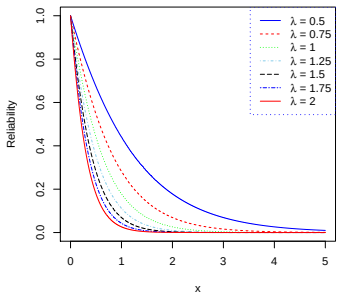
$$= \frac{1}{2}\big(3e^{-2\lambda t}-e^{-3\lambda t}\big) \tag{7}$$

The hazard rate is the probability that the life of the item ends in the next moment if it remains alive till time t. It is a practical approach of explaining the distribution

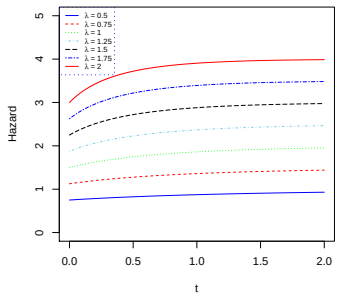
of 'time to event' as it has an accepted understanding that relates to the aging of a population [27]. It is defined as the ratio of the pdf to the reliability, ($h(t) = \frac{g(t)}{R(t)}$). Other rate functions are: cumulative hazard rate ($CH = -ln(R(t))$, reversed hazard rate ($RH = \frac{g(t)}{G(t)}$ and the odds ratio ($O = \frac{G(t)}{1-G(t)}$ functions. They are defined for the EED as:

$$\begin{aligned} h(t) &= \frac{3\lambda\left(2e^{-2\lambda t} - e^{-3\lambda t}\right)}{\left(3e^{-2\lambda t} - e^{-3\lambda t}\right)} \\ CH &= -ln\left(\frac{1}{2}\left(3e^{-2\lambda t} - e^{-3\lambda t}\right)\right) \\ RH &= \frac{3\lambda\left(2e^{-2\lambda t} - e^{-3\lambda t}\right)}{\left(e^{-3\lambda t} - 3e^{-2\lambda t}\right) + 2} \\ O &= \frac{\left(e^{-3\lambda t} - 3e^{-2\lambda t}\right) + 2}{\left(3e^{-2\lambda t} - e^{-3\lambda t}\right)} \end{aligned}$$

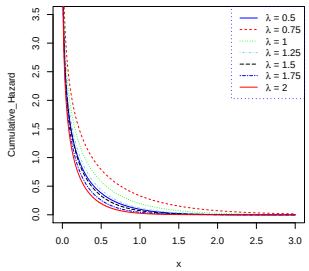
(a) The reliability of EED for different values of λ.



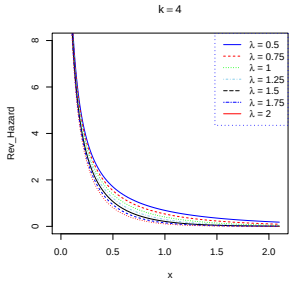
(b) The hazard rate of EED for different values of λ.



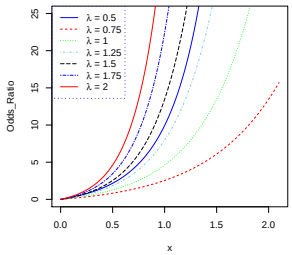
(a) The cumulative hazard rate of EED for different values of λ.



(b) The reversed hazard rate of EED for different values of λ.



(c) The odds ratio of EED for different values of λ.



4 Properties of EED

4.1 The moments

The r^{th} moment of a continuous random variable X is defined by [28] as

$$E(X^r) = \int_{-\infty}^{\infty} x^r g(x) dx$$

Thus the r^{th} moment of the EED distribution is given by the following theorem:

Theorem 2 *Let X be a random variable that follows an EED, therefore the r^{th} moment is defined as:*

$$E(X^r) = \frac{3r!}{(6\lambda)^{r+1}}\left(3^{r+1} - 2^{r+1}\right) \tag{8}$$

Proof.

$$\begin{aligned} E(X^r) &= \int_0^{\infty} x^r \Big[\frac{3\lambda}{2}\left(2e^{-2\lambda x} - e^{-3\lambda x}\right)\Big] dx \\ &= \int_0^{\infty} \frac{3\lambda}{2}\left(2x^r e^{-2\lambda x} - x^r e^{-3\lambda x}\right) dx \end{aligned}$$

using the gamma function and the u-substitution we get:

$$\begin{aligned} E(X^r) &= \frac{3\lambda\Gamma(r+1)}{(2\lambda)^{r+1}} - \frac{3\lambda\Gamma(r+1)}{2(3\lambda)^{r+1}} \\ &= \frac{r!}{\lambda^r}\left(\frac{3^{r+1} - 2^r}{2(6)^r}\right) \end{aligned}$$

The first four moments can be calculated by substituting $r = 1$, 2, 3 and 4 in (8). Thus

$$\mu = E(X) = \frac{7}{12\lambda} \tag{9}$$

$$E(X^2) = \frac{23}{36\lambda^2} \tag{10}$$

$$E(X^3) = \frac{73}{72\lambda^3} \tag{11}$$

$$E(X^4) = \frac{227}{108\lambda^4} \tag{12}$$

The variance ($var(X) = \sigma^2$, standard deviation (σ), skewness ($sk(X)$) and kurtosis ($kur(X)$) of the random variable X that follow EED are given by:

$$\begin{aligned}
\sigma^2 &= var(X) = E(X^2) - (E(X))^2 = \frac{43}{144\lambda^2} \\
\sigma &= \frac{\sqrt{43}}{12\lambda} \\
Sk(X)) &= \frac{E(X^3) - 3(E(X))E(X^2) + 2(E(X))^3}{(var(X))^{\frac{3}{2}}} = 1.7945 \\
kur(X) &= \frac{E(X^4) - 4(E(X))(E(X^3)) + 6(E(X))^2E(X^2) - 3(E(X))^4}{var(X)^2} = 7.7734
\end{aligned}$$

The values of the Mean and standard deviation of the EED random variable X for different values of the parameter λ were computed and listed in Table (1). They are calculated using R software [22]. The table shows that the values of the mean and the standard deviation of the EED random variable decrease by increasing the value of the parameter, (λ).

Table 1: The mean, standard deviation of EED for different values of λ

λ	μ	σ	λ	μ	σ	λ	μ	σ
0.01	10.5507	13. 7253	2. 00	0.292	0.273	4. 00	0.146	0.137
0.10	5. 824	5. 432	2. 10	0.278	0.260	4. 10	0.142	0.133
0.20	2. 917	2. 732	2. 20	0.265	0.248	4. 20	0.139	0.130
0.30	1. 944	1. 822	2. 30	0.254	0.238	4. 30	0.136	0.127
0.40	1. 458	1. 366	2. 40	0.243	0.228	4. 40	0.133	0.124
0.50	1. 167	1. 093	2. 50	0.233	0.219	4. 50	0.130	0.121
0.60	0.972	0.911	2. 60	0.224	0.210	4. 60	0.127	0.119
0.70	0.833	0.781	2. 70	0.216	0.202	4. 70	0.124	0.116
0.80	0.729	0.683	2. 80	0.208	0.195	4. 80	0.122	0.114
0.90	0.648	0.607	2. 90	0.201	0.188	4. 90	0.119	0.112
1. 00	0.583	0.546	3. 00	0.194	0.182	5. 00	0.117	0.109
1. 10	0.530	0.497	3. 10	0.188	0.176	5. 10	0.114	0.107
1. 20	0.486	0.455	3. 20	0.182	0.171	5. 20	0.112	0.105
1. 30	0.449	0.420	3. 30	0.177	0.166	5. 30	0.110	0.103
1. 40	0.417	0.390	3. 40	0.172	0.161	5. 40	0.108	0.101
1. 50	0.389	0.364	3. 50	0.167	0.156	5. 50	0.106	0.099
1. 60	0.365	0.342	3. 60	0.162	0.152	5. 60	0.104	0.098
1. 70	0.343	0.321	3. 70	0.158	0.148	5. 70	0.102	0.096
1. 80	0.324	0.304	3. 80	0.154	0.144	5. 80	0.101	0.094
1. 90	0.307	0.288	3. 90	0.150	0.140	5. 90	0.099	0.093
6. 00	0.097	0.091	8. 00	0.073	0.068	10.00	0.058	0.055
6. 10	0.096	0.090	8. 10	0.072	0.067	10.10	0.058	0.054
6. 20	0.094	0.088	8. 20	0.071	0.067	10.20	0.057	0.054

6. 30	0.093	0.087	8. 30	0.070	0.066	10.30	0.057	0.053
6. 40	0.091	0.085	8. 40	0.069	0.065	10.40	0.056	0.053
6. 50	0.090	0.084	8. 50	0.069	0.064	10.50	0.056	0.052
6. 60	0.088	0.083	8. 60	0.068	0.064	10.60	0.055	0.052
6. 70	0.087	0.082	8. 70	0.067	0.063	10.70	0.055	0.051
6. 80	0.086	0.080	8. 80	0.066	0.062	10.80	0.054	0.051
6. 90	0.085	0.079	8. 90	0.066	0.061	10.90	0.054	0.050
7. 00	0.083	0.078	9. 00	0.065	0.061	11. 00	0.053	0.050
7. 10	0.082	0.077	9. 10	0.064	0.060	11. 10	0.053	0.049
7. 20	0.081	0.076	9. 20	0.063	0.059	11. 20	0.052	0.049
7. 30	0.080	0.075	9. 30	0.063	0.059	11. 30	0.052	0.048
7. 40	0.079	0.074	9. 40	0.062	0.058	11. 40	0.051	0.048
7. 50	0.078	0.073	9. 50	0.061	0.058	11. 50	0.051	0.048
7. 60	0.077	0.072	9. 60	0.061	0.057	11. 60	0.050	0.047
7. 70	0.076	0.071	9. 70	0.060	0.056	11. 70	0.050	0.047
7. 80	0.075	0.070	9. 80	0.060	0.056	11. 80	0.049	0.046
7. 90	0.074	0.069	9. 90	0.059	0.055	11. 90	0.049	0.046

4.2 Moment and characteristic generating function

The moment generating function can be used to derive the moments of a random variable. It is defined by $M_X(t) = E\big(e^{tX}\big)$. For the EED random variable, it is defined by the following theorem:

Theorem 3 *Let X be an EED random variable, then its moment generating function is defined as:*

$$M_X(t) = \frac{3\lambda}{2\lambda - t} - \frac{3\lambda}{2(3\lambda - t)}, \quad \lambda > \frac{t}{2} \tag{13}$$

Proof.

$$\begin{aligned} M_X(t) &= E\big(e^{tX}\big) = \int_0^\infty e^{tx}\Big(\frac{3\lambda}{2}\big(2e^{-2\lambda x} - e^{-3\lambda x}\big)\Big)dx \\ &= \int_0^\infty \frac{3\lambda}{2}\big(2e^{-(2\lambda - t)x} - e^{-(3\lambda - t)x}\big)dx \\ &= \frac{3\lambda}{2\lambda - t} - \frac{3\lambda}{2(3\lambda - t)}, \quad \lambda > \frac{t}{2} \end{aligned}$$

The characteristic function is another function that can be used to generate the moments. It is defined to be $\phi(t) = E\big(e^{itx}\big)$. Thus, from (13) the characteristic function is defined to be

$$\phi(t) = E\big(e^{itX}\big) = \frac{3\lambda}{2\lambda - it} - \frac{3\lambda}{2(3\lambda - it)}.$$

5 Maximum Likelihood Estimates

Maximum likelihood method of estimation [12] is a well-known method of estimating the distribution parameters. For the EED, the maximum likelihood estimator of the distribution parameter is given by the following theorem:

Theorem 4 *Let $X_1, X_2, \cdots, X_n$ be an independent and identically distributed random sample with pdf defined in (6). The maximum likelihood estimate of the distribution parameter is the solution of the following equation*

$$\sum_{i=1}^{n} x_i \left(\frac{4e^{-2\lambda x_i} - 3e^{-3\lambda x_i}}{2e^{-2\lambda x_i} - e^{-3\lambda x_i}} \right) = \frac{n}{\lambda} \tag{14}$$

Proof. Let $X_1, X_2, \cdots, X_n$ be an independent and identically distributed random sample with pdf

$$g(x_i) = \frac{3\lambda}{2}\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right), x_i > 0.$$

Then, the joint pdf of $X_1, X_2, \cdots, X_n$ is given by

$$g(x_1, \cdots, x_n|\lambda) = \prod_{i=1}^{n} \left(\frac{3\lambda}{2}\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right) \right).$$

The function $ln(x)$ is a monotonically increasing function, thus log-likelihoods have the same relations of order as the likelihoods. The likelihood function L is defined as:

$$L = L(\lambda|x_1, \cdots, x_n) = \prod_{i=1}^{n} \left(\frac{3\lambda}{2}\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right) \right).$$

The log-likelihood function is, then

$$\begin{aligned} \ell = ln(L) &= ln\left(\prod_{i=1}^{n} \left(\frac{3\lambda}{2}\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right) \right) \right) \\ &= ln\left(\left(\frac{3\lambda}{2}\right)^n \prod_{i=1}^{n} \left(\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right) \right) \right) \\ &= nln\left(\frac{3}{2}\right) + nln(\lambda) + \sum_{i=1}^{n} ln\left(2e^{-2\lambda x_i} - e^{-3\lambda x_i}\right) \end{aligned}$$

Now, deriving the log-likelihood function with respect to the parameter (λ), we get

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i \left(\frac{4e^{-2\lambda x_i} - 3e^{-3\lambda x_i}}{2e^{-2\lambda x_i} - e^{-3\lambda x_i}} \right)$$

Equating this equation to 0, we get

$$\sum_{i=1}^{n} x_i \left(\frac{4e^{-2\lambda x_i} - 3e^{-3\lambda x_i}}{2e^{-2\lambda x_i} - e^{-3\lambda x_i}} \right) = \frac{n}{\lambda}$$

There is no exact solution for this equation.

6 Quantile Function

The quantile q of the random variable, say X, is the solution of the equation $q = G(x_q) = p(X \leq x_q)$ [24]. Thus for a random variable that follows EED it is can be derived as:

$$\begin{aligned} q = G(x_q) = p(X \leq x_q) &= \frac{1}{2}\left(e^{-3\lambda x_q} - 3e^{-2\lambda x_q}\right) + 1 \\ 2(q-1) &= e^{-3\lambda x_q} - 3e^{-2\lambda x_q} \\ 0 &= \left(e^{-\lambda x_q}\right)^3 - 3\left(e^{-\lambda x_q}\right)^2 - 2(q-1) \end{aligned} \tag{15}$$

This is a cubic equation which can be solved numerically. So, we have used online calculator [29] to solve this equation for three values of $q = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$. In all cases we assumed $\lambda = \frac{1}{2}$.

For $q = \frac{1}{4}$, we have

$$\begin{aligned} e^{-0.5x_q} &= -0.64178 \Rightarrow x_q \text{ there is no value of} x_q \text{ achieves this equation.} \\ e^{-0.5x_q} &= 0.83174 \Rightarrow x_q = 0.3652704 \text{ (which is an accepted root as it is in the range of the random variable.)} \\ e^{-0.5x_q} &= 2.81003 \Rightarrow x_q = -2.066396 \text{ which is not in the range of X.} \end{aligned}$$

Assuming $q = \frac{1}{2}$, we have

$$\begin{aligned} e^{-0.5x_q} &= -0.532089 \Rightarrow x_q \text{ there is no value of} x_q \text{ achieves this equation.} \\ e^{-0.5x_q} &= 0.652704 \Rightarrow x_q = 0.853263 \text{ (which is an accepted root as it is in the range of the random variable.)} \\ e^{-0.5x_q} &= 2.879385 \Rightarrow x_q = -2.11515 \text{ which is not in the range of X.} \end{aligned}$$

Assuming $q = \frac{3}{4}$, we have

$$\begin{aligned} e^{-0.5x_q} &= -0.384367 \Rightarrow x_q \text{ there is no value of} x_q \text{ achieves this equation.} \\ e^{-0.5x_q} &= 0.442125 \Rightarrow x_q = 1.632325 \text{ (which is an accepted root as it is in the range of the random variable.)} \\ e^{-0.5x_q} &= 2.942242 \Rightarrow x_q = -2.158344 \text{ which is not in the range of X.} \end{aligned}$$

7 Mean and Median Absolute Deviation

Mean absolute deviation (MAD) [8] of a data set is the average distance between each data value and the mean. Mean absolute deviation is a way to describe variation in a data set. Mean absolute deviation helps us get a sense of how "spread out" the values in a data set are. The mean deviation for an EED random variable, X, with

mean $\mu = E(X)$ can be derived as:

$$\begin{aligned}
D(\mu) &= E(X-\mu) = \int_0^{\infty} |x-\mu| g(x)dx \\
&= \int_0^{\mu} (x-\mu) g(x)dx + \int_{\mu}^{\infty} (\mu - x) g(x)dx \\
&= \int_0^{\mu} (x-\mu) \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&\quad + \int_{\mu}^{\infty} (\mu - x) \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&= \int_0^{\mu} x \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx - \int_0^{\mu} \mu \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&\quad + \int_{\mu}^{\infty} \mu \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx - \int_{\mu}^{\infty} x \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&= \frac{1}{3\lambda}(3\lambda+1)e^{-3\lambda\mu} - \frac{1}{2\lambda}(6\lambda\mu+3)e^{-2\lambda\mu} + \mu(3e^{-2\lambda\mu} - e^{-3\lambda\mu}) + \frac{3}{4\lambda} - \mu
\end{aligned}$$

The median absolute deviation(MAD) [30] is a robust measure of how spread out a set of data is. The variance and standard deviation are also measures of spread, but they are more affected by extremely high or extremely low values and non normality. If your data is normal, the standard deviation is usually the best choice for assessing spread. However, if your data is not normal, the MAD is one statistic you can use instead. The median deviation for an EED random variable, X, with mean $\mu = E(X)$ can be derived as:

$$\begin{aligned}
D(M) &= E(X-M) = \int_0^{\infty} |x-M| g(x)dx \\
&= \int_0^{M} (x-M) g(x)dx + \int_{M}^{\infty} (M - x) g(x)dx \\
&= \int_0^{M} (x-M) \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&\quad + \int_{M}^{\infty} (M - x) \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx
\end{aligned}$$

$$\begin{aligned}
&= \int_0^{M} x \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx - \int_0^{M} M \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&\quad + \int_{M}^{\infty} M \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx - \int_{M}^{\infty} x \left[\frac{3\lambda}{2} \left(2e^{-2\lambda x} - e^{-3\lambda x} \right) \right] dx \\
&= \frac{1}{3\lambda}(3\lambda+1)e^{-3\lambda M} - \frac{1}{2\lambda}(6\lambda M+3)e^{-2\lambda M} + M(3e^{-2\lambda M} - e^{-3\lambda M}) + \frac{3}{4\lambda} - M
\end{aligned}$$

8 Rényi Entropy

The Rényi entropy [23] is defined as $E_R = \frac{1}{1-\rho} ln\Big[\int_0^\infty (g(x))^\rho dx\Big]$. Therefore, for the EED it is defined by the following theorem:

Theorem 5 *The Rényi entropy for a random variable X that follows an EED with pdf $g(x)$ defined in (6) is defined as:*

$$E_R = \frac{1}{1-\rho} ln\left[\sum_{k=0}^{\rho} \binom{\rho}{k} (-1)^{\rho-k} (2)^k \left(\frac{3\lambda}{2}\right)^\rho \frac{1}{(3\rho-k)\lambda}\right], \qquad \rho > \frac{k}{3}, \rho \neq 1$$

Proof.

$$\begin{aligned}
E_R &= \frac{1}{1-\rho} ln\left[\int_0^\infty \left(\frac{3\lambda}{2}\big(2e^{-2\lambda x} - e^{-3\lambda x}\big)\right)^\rho dx\right] \\
&= \frac{1}{1-\rho} ln\left[\int_0^\infty \left(\frac{3\lambda}{2}\right)^\rho \sum_{k=0}^{\rho} \binom{\rho}{k} (2e^{-2\lambda x})^k (-e^{-3\lambda x})^{\rho-k} dx\right] \\
&= \frac{1}{1-\rho} ln\left[\sum_{k=0}^{\rho} \binom{\rho}{k} (-1)^{\rho-k} (2)^k \left(\frac{3\lambda}{2}\right)^\rho \int_0^\infty e^{-2k\lambda x} e^{-3(\rho-k)\lambda x} dx\right] \\
&= \frac{1}{1-\rho} ln\left[\sum_{k=0}^{\rho} \binom{\rho}{k} (-1)^{\rho-k} (2)^k \left(\frac{3\lambda}{2}\right)^\rho \int_0^\infty e^{-(3\rho-k)\lambda x} dx\right] \\
&= \frac{1}{1-\rho} ln\left[\sum_{k=0}^{\rho} \binom{\rho}{k} (-1)^{\rho-k} (2)^k \left(\frac{3\lambda}{2}\right)^\rho \frac{1}{(3\rho-k)\lambda}\right], \qquad \rho > \frac{k}{3}
\end{aligned}$$

9 Order Statistics

Consider the random variables $X_1, X_2, \cdots, X_n$. They almost assume to be independent and identically distributed [10]. Arrange them in an ascending order, so we get $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, where

$$X_{(1)} = min(X_1, X_2, \cdots, X_n),$$

$X_{(i)}$ is the i^{th} order statistic $(i = 1, ..., n)$, and

$$X_{(n)} = max(X_1, X_2, \cdots, X_n)$$

Assuming that these random variables are following EED. The order statistics topic deals with the properties of these random variable and the applications of them [10]. Therefore, the pdf of $X_{(1)}$, $X_{(n)}$ and the general case $X_{(i)}$, $1 < i < n$ are defined by

[10] as:

$$\begin{aligned}
g_1(x) &= ng(x)[1-G(x)]^{n-1} \\
&= n\left(\frac{3\lambda}{2}\left(2e^{-2\lambda x}-e^{-3\lambda x}\right)\right)\left(\frac{1}{2}\left(e^{-3\lambda x}-3e^{-2\lambda x}\right)\right)^{n-1} \\
&= 3\lambda n\left(\frac{1}{2}\right)^n\left(2e^{-2\lambda x}-e^{-3\lambda x}\right)\left(e^{-3\lambda x}-3e^{-2\lambda x}\right)^{n-1} \\
g_n(x) &= ng(x)[G(x)]^{n-1} \\
&= n\left(\frac{3\lambda}{2}\left(2e^{-2\lambda x}-e^{-3\lambda x}\right)\right)\left(\frac{1}{2}\left(e^{-3\lambda x}-3e^{-2\lambda x}\right)+1\right)^{n-1} \\
&= 3\lambda n\left(\frac{1}{2}\right)^n\left(2e^{-2\lambda x}-e^{-3\lambda x}\right)\left(e^{-3\lambda x}-3e^{-2\lambda x}+2\right)^{n-1} \\
g_i(x) &= \frac{n!}{(i-1)!}g(x)[G(x)]^{i-1}[1-G(x)]^{n-i} \\
&= \frac{3\lambda n!}{2^n(i-1)!}\left(2e^{-2\lambda x}-e^{-3\lambda x}\right)\left(e^{-3\lambda x}-3e^{-2\lambda x}+1\right)^{i-1}\left(e^{-3\lambda x}-3e^{-2\lambda x}\right)^{n-i}
\end{aligned}$$

10 Applications

Two real data sets have been used to study the performance of the EED in comparison with the exponential using the log-likelihood, Akaike Information Criteria (AIC) and Bayesian information criteria (BIC).

The first data represents the survival times of 121 patients with breast cancer computed from a large hospital in a period from 1929-1938 ([18]). The observations are as follows:

0.3	0.3	4. 0	5. 0	5. 6	6. 2	6. 3	6. 6	6. 8	7. 4
7. 5	8. 4	8. 4	10.3	11. 0	11. 8	12. 2	12. 3	13. 5	14. 4
14. 4	14. 8	15. 5	15. 7	16. 2	16. 3	16. 5	16. 8	17. 2	17. 3
17. 5	17. 9	19. 8	20.4	20.9	21. 0	21. 0	21. 1	23. 0	23. 4
23. 6	24. 0	24. 0	27. 9	28. 2	29. 1	30.0	31. 0	31. 0	32. 0
35. 0	35. 0	37. 0	37. 0	37. 0	38. 0	38. 0	38. 0	39. 0	39. 0
40.0	40.0	40.0	41. 0	41. 0	41. 0	42. 0	43. 0	43. 0	43. 0
44. 0	45. 0	45. 0	46. 0	46. 0	47. 0	48. 0	49. 0	51. 0	51. 0
51. 0	52. 0	54. 0	55. 0	56. 0	57. 0	58. 0	59. 0	60.0	60.0
60.0	61. 0	62. 0	65. 0	65. 0	67. 0	67. 0	68. 0	69. 0	78. 0
80.0	83. 0	88. 0	89. 0	90.0	93. 0	96. 0	103. 0	105. 0	109. 0
109. 0	111. 0	115. 0	117. 0	125. 0	126. 0	127. 0	129. 0	129. 0	139. 0
154. 0.									

The second data set represent the total milk production in the first birth of 107 cows from SINDI race. These cows are property of the Carnaúba farm which belongs to the Agropecuária Manoel Dantas Ltda (AMDA), located in Taperoá City, Paraíba (Brazil). This data is presented by [9]. These data are

Table 2: Performance of the EED using the survival times of 121 patients with breast cancer

Distribution	$\hat{\lambda}$	SE	LogLik	AIC	BIC
EED	0.01276058	0.001073257	-582. 8024	1167. 605	1170.4
Exp	0.02159879	0.001959307	-585. 1278	1172. 256	1175. 051

0.4365, 0.4260, 0.5140, 0.6907, 0.7471, 0.2605, 0.6196, 0.8781,
0.4990, 0.6058, 0.6891, 0.5770, 0.5394, 0.1479, 0.2356, 0.6012,
0.1525, 0.5483, 0.6927, 0.7261, 0.3323, 0.0671, 0.2361, 0.4800,
0.5707, 0.7131, 0.5853, 0.6768, 0.5350, 0.4151, 0.6789, 0.4576,
0.3259, 0.2303, 0.7687, 0.4371, 0.3383, 0.6114, 0.3480, 0.4564,
0.7804, 0.3406, 0.4823, 0.5912, 0.5744, 0.5481, 0.1131, 0.7290,
0.0168, 0.5529, 0.4530, 0.3891, 0.4752, 0.3134, 0.3175, 0.1167,
0.6750, 0.5113, 0.5447, 0.4143, 0.5627, 0.5150, 0.0776, 0.3945,
0.4553, 0.4470, 0.5285, 0.5232, 0.6465, 0.0650, 0.8492, 0.8147,
0.3627, 0.3906, 0.4438, 0.4612, 0.3188, 0.2160, 0.6707, 0.6220,
0.5629, 0.4675, 0.6844, 0.3413, 0.4332, 0.0854, 0.3821, 0.4694,
0.3635, 0.4111, 0.5349, 0.3751, 0.1546, 0.4517, 0.2681, 0.4049,
0.5553, 0.5878, 0.4741, 0.3598, 0.7629, 0.5941, 0.6174, 0.6860,
0.0609, 0.6488, 0.2747

Table 3: Performance of the EED using the milk production in the first birth of 107 cows from SINDI race.

Distribution	$\hat{\lambda}$	SE	LogLik	AIC	BIC
EED	1. 28906	0.1147328	-20.63622	43. 27244	45. 94526
Exp	2. 132867	0.206192	-25. 95077	53. 90155	56. 57438

The goal of generalizing or covering an ordinary distribution is to make it more flexible. As presented in Tables (2) and (3) for the both data sets; the EED is more flexible than the exponential distribution because it has the lower AIC and BIC values and highest log-likelihood value.

11 Simulation Study

To study the performance of the Epanechnikov-exponential distribution and to investigate the behavior of MLE. We have used the R software ([22]) to run the simulation study. We generate 10, 000 samples of the Epanechnikov-exponential distribution, with sizes, n=30, 80, 100, 200, 350, and by choosing λ=0.5, 1, 1. 5, 2, 2. 5, 3. For each parameter and each sample, we calculate the average absolute bias $Bias(\hat{\lambda})$, which is defined by $Bias(\hat{\lambda}) = \frac{1}{N}\sum_{j=1}^{N}|\hat{\lambda} - \lambda|$. The mean square error of the estimates $MSE(\hat{\lambda})$, $MSE(\hat{\lambda})$, $MSE(\hat{\lambda}) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\lambda}-\lambda)^2$ and the mean relative estimates $MRE(\hat{\lambda})$, $MRE(\hat{\lambda}) = \frac{1}{N}\sum_{j=1}^{N}\frac{|\hat{\lambda}-\lambda|}{\lambda}$. Table (4) shows the $Bias(\hat{\lambda})$, $MSE(\hat{\lambda})$ and the $MRE(\hat{\lambda})$. It can be observed that the values of $Bias$, MSE and MRE decreases as the sample sizes increases. From these results we can conclude that the estimation method via MLE conserves the consistency property for all values of the parameters.

Table 4: Simulation results for $Bias$, MSE and MRE for different values of λ and n.

N	λ	0.5	1	1. 5	2	2. 5	3
	$\hat{\lambda}$	0.51543	1. 031	1. 5442	2. 05937	2. 57627	3. 08838
	$Bias$	0.07049	0.14351	0.21209	0.28119	0.35112	0.42339
30	MSE	0.00852	0.03473	0.07605	0.13412	0.20998	0.30303
	MRE	0.14097	0.14351	0.14139	0.1406	0.14045	0.14113
	$\hat{\lambda}$	0.50562	1. 01221	1. 51789	2. 02342	2. 53395	3. 03933
80	$Bias$	0.04201	0.0846	0.12622	0.17	0.21319	0.25579
	MSE	0.00285	0.01167	0.02582	0.04652	0.07361	0.1073
	MRE	0.08403	0.0846	0.08415	0.085	0.08528	0.08526
	$\hat{\lambda}$	0.50462	1. 01109	1. 51376	2. 01721	2. 52424	3. 02787
100	$Bias$	0.03788	0.07599	0.11344	0.15135	0.18992	0.22757
	MSE	0.00231	0.00927	0.02078	0.03689	0.05808	0.08386
	MRE	0.07576	0.07599	0.07562	0.07568	0.07597	0.07586
	$\hat{\lambda}$	0.5021	1. 00422	1. 50674	2. 00824	2. 51195	3. 01724
200	$Bias$	0.02644	0.05331	0.07918	0.10675	0.1318	0.16069
	MSE	0.0011	0.0045	0.00995	0.01807	0.02764	0.04113
	MRE	0.05287	0.05331	0.05279	0.05337	0.05272	0.05356
	$\hat{\lambda}$	0.50176	1. 00334	1. 50409	2. 00535	2. 50753	3. 00985
350	$Bias$	0.02153	0.04291	0.0651	0.08617	0.10743	0.1316
	MSE	0.00073	0.00293	0.00667	0.01164	0.01835	0.02727
	MRE	0.04307	0.04291	0.0434	0.04308	0.04297	0.04387

12 Conclusion

A new distribution is proposed. This distribution is a composition of Epanechnikov kernel function and the exponential distribution. We called this distribution the Epanechnikov exponential distribution (EED). We have studied many properties of this distribution. Two simulation studies are conducted, they show that the variability is increased as the parameter values decreased and the Bias, MSE and MRE are decreased as the sample size increased. Two real data sets are applied to EED. They showed that the performance of EED is better of the exponential distribution.

References

[1] M. Al-khazaleh, A. Al-Omari, A. M. Al-khazaleh, *Transmuted Two-Parameter Lindley Distribution*, Journal of Statistics Applications and Probability, vol. 5, no. 3, 2016, 1-11.

[2] A. I. Al-Omari, A. M. Al-khazaleh, L. M. AlZoubi, *Transmuted Janardan Distribution: A Generalization of the Janardan Distribution*, Journal of Statistics Applications and Probability, vol. 5, no. 2, 2017, 1-11.

[3] A. Al-Omari, M. Gharaibeh, *Topp-leone mukherjee-islam distribution: properties and applications*, Pakistan Journal of Statistics, vol. 34, no. 6, 2018, 479-494.

[4] A. Al-Omari, A. M. Al-khazaleh, L. AlZoubi, *A Generalization of the New-Weibull Pareto Distribution*, Revista Investigacion Operacionl, vol. 41, no. 1, 2020, 138-146.

[5] L. Al-Zoubi, *Transmuted Mukherjee-Islam Distribution: A Generalization of Mukherjee-Islam Distribution*, Journal of Mathematics Research, vol. 9, no. 4, 2017, 135-144.

[6] H. A. Alsikeek, *Quadratic Transmutation Map for Reciprocal Distribution and Two-Tarameter Weighted Exponential Distribution*, Masterś thesis, Department of Mathematics, Al al-Bayt University, Mafraq, Jordan, 2018.

[7] R. M. AzZwideen, *Transmutation Maps for Gamma-Gompertz and Generalized Type-II Half-Logistics Distributions*, Master thesis, Department of Mathematics, Al al-Bayt University, Mafraq, Jordan, 2017.

[8] P. Bloomfield, W. S. Steiger, *Least Absolute Deviations: Theory, Applications and Algorithms*, Birkhiiuser, Boston, 1983.

[9] G. M. Cordeiro, R. D. S. Brito, *The beta power distributions*, Brazilian Journal of Probability and Statistics, vol. 26, no. 1, 2012, 88-112.

[10] H. A. David, *Order statistics*, John Wiley and Sons, New York, 1981.

[11] V. A. Epanechnikov, *Non-parametric estimation of a multivariate probability density*, Theory of Probability and Its Applications, vol. 14, 1969, 153-158.

[12] R. A. Fisher, *On an absolute criterion for fitting frequency curves*, Statistical Sciences, vol. 12, no. 1, 1997, 39-41.

[13] M. M. Gharaibeh, A. I. Al-Omari, *Transmuted Ishita Distribution and its Applications*, Journal of Statistical Applications and Probability, vol. 8, no. 2, 2019 1-14.

[14] A. C. Guidoum, *Kernel estimator and bandwidth selection for density and its derivatives, kedd Package*, version 1, 2015.

[15] A. D. Herbert, H. N. Nagaraja, *Order Statistics*, 3rd Edition, Wiley Series in Probability and Statistics, 2003.

[16] G. Jowett, *The Exponential Distribution and Its Applications*, The Incorporated Statistician, vol. 8, no. 2, 1958, 89-95.

[17] A. M. Khazaleh, *Transmuted burr type XII distribution: A generalization of the burr type XII distribution*, International Mathematical Forum, vol. 11, no. 12, 2016, 547-556.

[18] E. T. Lee, *Statistical Methods for Survival Data Analysis*, 2nd Edition, John Wiley and Sons Ltd. , New York, USA, 2003.

[19] R Development Core Team , *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020.

[20] F. S. Rabaiah, *Generalizations of Power Function and Type-I Half Logistic Distributions Using Quadratic Transmutation Map*, Master thesis, Department of Mathematics, Al al-Bayt University, Mafraq, Jordan, 2018.

[21] M. W. A. Ramos, G. M. Cordeiro, P. R. D. Marinho, C. R. B. Dias, G. G. Hamedani, *The Zografos-Balakrishnan log-logistic distribution: Properties and applications*, J. Stat. Theory Applic., vol. 12, 2013, 225-244.

[22] R Core Team, *A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, URL http://www. R-project. org/, 2019.

[23] A. Rényi, *On measures of entropy and information*, Proceedings of the 4^{th} Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1961, 547-561.

[24] E. Parzen, *Quantile Probability and Statistical Data Modeling*, Statistical Science, vol. 19, no. 4, 2004, 652-662.

[25] W. T. Shaw, I. R. Buckley, *The alchemy of probability distributions: Beyond gram-charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map*, Technical report, 2007.

[26] N. Unnikrishnan Nair, P. G. Sankaran, N. Balakrishnan, *Chapter 2 - Basic Reliability Concepts*, Editors: N. Unnikrishnan Nair, P. G. Sankaran, N. Balakrishnan, Reliability Modelling and Analysis in Discrete Time, Academic Press, 2018, 43-106.

[27] S. Borovkova, *Analysis of survival data*, Nieuw Archief voor Wiskunde, vol. 5, no. 3, 2002, 302-307.

[28] P. Sahoo, *Probability and Mathematical Statistics*, First Edition, University of Louisville, 2013.

[29] https://mathforyou. net/en/online/equation/cubic/

[30] P. J. Rousseeuw, C. Croux, *Alternatives to the median absolute deviation*, Journal of the American Statistical Association, vol. 88, no. 424, 1993, 1273–1283.

[31] M. Gharaibeh, *Transmuted Aradhana distribution: Properties and applications*, Jordan Journal of Mathematics and Statistics (JJMS), vol. 13, no. 2, 2020, 287-304.

[32] Z. Zambom, R. Dias, *A Review of Kernel Density Estimation with Applications to Econometrics*, International Econometric Review (IER), Econometric Research Association, vol. 5, no. 1, 2013, 20-42.

Ahmad Mohammad Alkhazaalh
Al al-Bayt University
Faculty of Science
Department of Mathematics
Jordan, Mafraq, 25113
e-mail: ahmed_2005kh@yahoo.com

Loai Mahmoud Al-Zoubi
Al al-Bayt University
Faculty of Science
Department of Mathematics
Jordan, Mafraq, 25113
e-mail: loai67@aabu.edu.jo