Aktüerya Derneği

İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 12, 2019, 2, 43-57

Geliş/Received:15.07.2019, Kabul/Accepted: 25.09.2019

www.istatistikciler.org

Araştırma Makalesi / Research Article

A comparative study for estimation of the parameters of the folded exponential power distribution

Şükrü Acıtaş
Eskişehir Technical University
Faculty of Science
Department of Statistics
26470 Eskişehir, TURKEY
sacitas@eskisehir.edu.tr
0000-0002-4131-0086

Abstract

Folded distributions are commonly used for the data set which is obtained without regarding the algebraic signs of the measurements. Therefore, they have extensive applications in different fields such as engineering, finance, insurance and so on. Folded exponential power (FEP) distribution is a newly proposed distribution which has modeling flexibility and easy usage [1]. In this study, we therefore consider different parametric methods for estimating the unknown parameters of FEP distribution. Maximum likelihood (ML), ordinary and weighted least squares (LS and WLS), Cramer von Mises (CVM) and maximum product of spacings (MPS) methods are used during the estimation process. The performances of the considered estimators are compared in a Monte-Carlo simulation study via bias and mean squared error (MSE) criteria. Results show that MPS method outperforms its rivals. Two real life applications taken from the literature are also considered.

Keywords: Bias, Efficiency, Estimation, Folded exponential power distribution.

Öz

Katlanmış üstel güç dağılımının parametrelerinin tahmini için karşılaştırmalı bir çalışma

Katlanmış dağılımlar, ölçümlerin cebirsel işaretlerinin önemli olmadığı veri setleri için yaygın olarak kullanılırlar. Bu nedenle, mühendislik, finans, ekonomi v.b. bir çok alanda kapsamlı uygulamaları vardır. Modelleme esnekliği ve kolay kullanımı olan katlanmış üstel güç (FEP) dağılımı yeni önerilmiştir [1]. Bu çalışmada bu nedenle, FEP dağılımının bilinmeyen parametrelerinin farklı parametrik yöntemlerle tahmin edilmesi ele alınmıştır. En çok olabilirlik (ML), sıradan ve ağırlıklandırılmış en küçük kareler (LS ve WLS), Cramer von Mises (CVM) ve aralıkların çarpımının maksimumu (MPS) metotları tahmin sürecinde kullanılmıştır. Ele alınan tahmin edicilerin performansı, Monte-Carlo simülasyon çalışmasında yan ve hata kareler ortalaması (MSE) kriterleri kullanılarak karşılaştırılmıştır. Sonuçlar, MPS yönteminin rakiplerinden daha iyi bir performansa sahip olduğunu göstermiştir. Literatürden alınan iki gerçek hayat uygulaması ele alınmıştır.

Anahtar sözcükler: *Yan, Etkinlik, Tahmin, Katlanmış üstel güç dağılımı.*

1. Introduction

Folded distributions have many applications in different fields of science such as engineering, economics, medicine and so on [2 - 6]. Folded distributions are useful when the algebraic signs of the measurements are ignored. Some examples for these kind of measurements can be given as differences, deviations, lengths and angels. Therefore, the distribution of these measurements are mostly described by folded distributions which are defined as the distribution of absolute measurements.

Suppose that random variable X has probability density function (pdf) and cumulative distribution function (cdf) $g(\cdot)$ and $G(\cdot)$, respectively and define $Y = |X|$. The distribution of Y is a folded distribution and its pdf and cdf are then given by

$$f(y) = g(y) + g(-y) \quad \text{and} \quad F(y) = G(y) - G(-y), \tag{1}$$

respectively. If one take $g(\cdot)$ as the pdf of $N(\mu, \sigma^2)$, Y has folded normal (FN) distribution which is proposed by Leone et al. [5]. In the same manner, folded t distribution is suggested by Psarakis and Panareteos [7]. Cooray et al. [8] introduce folded logistic (FL) distribution. Nadarajah and Bakar [1] also consider some folded distributions to model the log-transformed Norwegian data. Among these, they originate three new distributions namely folded generalized t, folded Gumbel and folded exponential power (FEP) distributions and study some statistical properties and maximum likelihood (ML) estimation of the parameters.

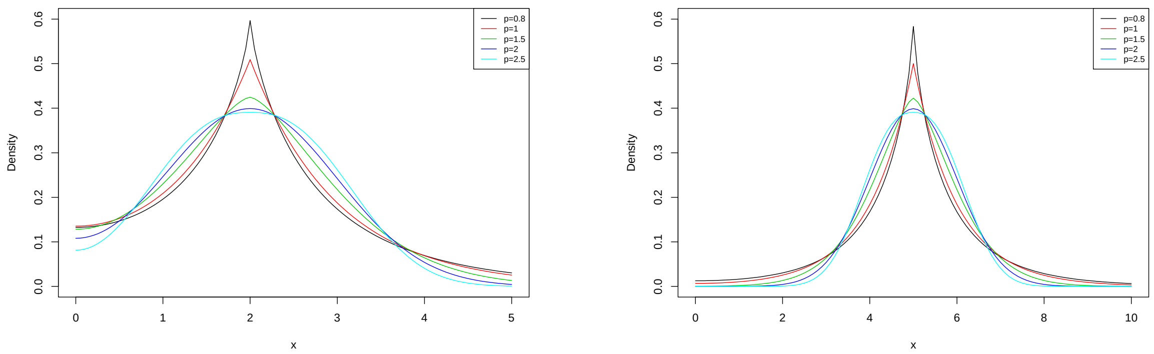
In this study, we consider FEP distribution because of its simplicity and modeling flexibility. The pdf and cdf of FEP distribution are given by

$$f(x;\mu,\sigma,p) = \frac{1}{2p^{1/p}\Gamma(1+1/p)\sigma}\left[\exp\left\{-\frac{|x-\mu|^p}{p\sigma^p}\right\} + \exp\left\{-\frac{|x+\mu|^p}{p\sigma^p}\right\}\right] \tag{2}$$

and

$$F(x;\mu,\sigma,p) = \frac{1}{2\Gamma(1/p)\sigma}\left\{\text{sign}(x-\mu)\gamma\left(\frac{1}{p},\frac{|x-\mu|^p}{p\sigma^p}\right) - \text{sign}(-x-\mu)\gamma\left(\frac{1}{p},\frac{|x+\mu|^p}{p\sigma^p}\right)\right\}, \tag{3}$$

respectively. Here, $x > 0$, $\mu > 0$, $\sigma > 0$ and $p > 0$. It should be mentioned that p stands for the shape parameter and μ, σ have their standard meanings, i.e the former and the latter are location and scale parameters, respectively. $\Gamma(\cdot)$ is the well-known gamma function and $\gamma(\cdot)$ denotes the incomplete gamma function which is defined as $\gamma(a,x) = \int_0^x t^{a-1}\exp(-t)dt, \quad a > 0, \quad x > 0.$



(a) $\mu = 2$ and $\sigma = 1$ (b) $\mu = 5$ and $\sigma = 1$

Figure 1. The pdf plots of FEP distribution for some certain values of the parameters.

FEP distribution is obtained using exponential power (EP), which is also known as generalized normal distribution [9], therefore it inherits properties of EP. For example, FEP distribution is platykurtic when $p > 2$ otherwise it is leptokurtic, see Figure 1 in which density plots of FEP distribution are provided for some representative values of the parameters. FEP reduces to folded Laplace [10] and folded normal distributions for $p = 1$ and $p = 2$, respectively. Furthermore, half exponential power distribution [11] is obtained when $\mu = 0$ and $\sigma = 1$. It is clear that half normal distribution is a special case of FEP distribution

if $\mu = 0$, $\sigma = 1$ and $p = 2$. We refer to Nadarajah and Bakar [1] for further information about FEP distribution.

As mentioned above, FEP distribution has attractive properties and modelling flexibility. However, there are no previous studies concerning different parametric estimation methods for the parameters of FEP distribution to the best of our knowledge. We therefore consider least squares (LS), weighted LS, Cramer von Mises (CVM) and maximum product of spacings (MPS) methods for estimating the parameters of FEP distribution. The performances of these estimation methods are compared using different settings of parameter values and sample sizes according to the bias, mean squared error (MSE) and deficiency (Def) criteria. We aim to evaluate the best method for estimating the parameters of FEP distibution. There are plenty of studies in which performances of the different estimation methods are compared in the context of different distributions, see for example [12 - 14].

The rest of the paper is organized as follows. Brief descriptions of the estimation methods are given in Section 2. Section 3 consists of a Monte-Carlo simulation study and its results. Real life data applications are considered in Section 4. The paper is finalized with some concluding remarks.

2. Methods of estimation

In this section, brief descriptions of the ML, LS, WLS, CVM and MPS methods are provided. We first provide a common notation which will be used in descriptions of the methods. Let $x_1, x_2, \ldots, x_n$ be a random sample drawn from FEP distribution. The order statistics of this sample are then denoted by $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$.

2.1. ML method

The ML estimators of the parameters μ, σ and p are obtained by maximizing the following loglikelihood function:

$$\log L = -n\log 2 - \frac{n}{p}\log p - n\log\sigma - n\log\Gamma\left(1+\frac{1}{p}\right) + \sum_{i=1}^{n}\log\left(\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\} + \exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}\right). \tag{4}$$

After taking partial derivatives of the $\log L$ function with respect to the parameters of interest and setting them equal to zero, we obtain the following likelihood equations [1]:

$$\frac{\partial\log L}{\partial\mu} = \sum_{i=1}^{n}\frac{|x_i-\mu|^{p-1}\operatorname{sign}(x_i-\mu)\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}}{\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}+\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}} - \sum_{i=1}^{n}\frac{|x_i+\mu|^{p-1}\operatorname{sign}(x_i+\mu)\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}}{\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}+\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}} = 0, \tag{5}$$

$$\frac{\partial\log L}{\partial\sigma} = -n\sigma^p + \sum_{i=1}^{n}\frac{|x_i-\mu|^p\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\} + |x_i+\mu|^p\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}}{p^2\sigma^p\left[\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}+\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}\right]} = 0, \tag{6}$$

$$\frac{\partial\log L}{\partial p} = \frac{n}{p^2}\log p - \frac{n}{p^2} + \frac{n}{p^2}\psi\left(1+\frac{1}{p}\right) + \sum_{i=1}^{n}\frac{|x_i-\mu|^p\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}\left(1-p\log|x_i-\mu|+p\log\sigma\right)}{\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}+\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}} + \sum_{i=1}^{n}\frac{|x_i+\mu|^p\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}\left(1-p\log|x_i+\mu|+p\log\sigma\right)}{p^2\sigma^p\left[\exp\left\{-\frac{|x_i-\mu|^p}{p\sigma^p}\right\}+\exp\left\{-\frac{|x_i+\mu|^p}{p\sigma^p}\right\}\right]} = 0 \tag{7}$$

where $\psi(\cdot)$ is the digamma function.

It is clear that equations (5)-(7) cannot be solved explicitly since they involve nonlinear functions of the parameters. Therefore, some numerical methods should be employed to obtain the ML estimates.

2.2. LS method

The LS estimators of the parameters μ, σ and p are obtained by minimizing the following function with respect to the parameters of interest [15]:

$$S = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right)^2 . \tag{8}$$

Here, $F(\cdot)$ is the cdf of FEP distribution. LS estimators can also be obtained by solving the following equations:

$$\frac{\partial S}{\partial \mu} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_1(x_{(i)}; \mu, \sigma, p) = 0, \tag{9}$$

$$\frac{\partial S}{\partial \sigma} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_2(x_{(i)}; \mu, \sigma, p) = 0, \tag{10}$$

$$\frac{\partial S}{\partial p} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_3(x_{(i)}; \mu, \sigma, p) = 0 \tag{11}$$

where

$$\Delta_1(x_{(i)}; \mu, \sigma, p) = \frac{1}{2\Gamma(1/p)\sigma} \left[\mathrm{sign}(x_{(i)} - \mu)\gamma'_{\mu} \left(\frac{1}{p}, \frac{|x_{(i)} - \mu|^p}{p\sigma^p} \right) - \mathrm{sign}(-x_{(i)} - \mu)\gamma'_{\mu} \left(\frac{1}{p}, \frac{|x_{(i)} + \mu|^p}{p\sigma^p} \right) \right], \tag{12}$$

$$\Delta_2(x_{(i)}; \mu, \sigma, p) = -\frac{1}{\sigma} F(x_{(i)}; \mu, \sigma, p) + \frac{1}{2\Gamma(1/p)\sigma} \left[\mathrm{sign}(x_{(i)} - \mu)\gamma'_{\sigma} \left(\frac{1}{p}, \frac{|x_{(i)} - \mu|^p}{p\sigma^p} \right) - \mathrm{sign}(-x_{(i)} - \mu)\gamma'_{\sigma} \left(\frac{1}{p}, \frac{|x_{(i)} + \mu|^p}{p\sigma^p} \right) \right] \tag{13}$$

and

$$\Delta_3(x_{(i)}; \mu, \sigma, p) = \frac{\psi\left(\frac{1}{p}\right)}{p^2} F(x_{(i)}; \mu, \sigma, p) + \frac{1}{2\Gamma(1/p)\sigma} \left[\mathrm{sign}(x_{(i)} - \mu)\gamma'_{p} \left(\frac{1}{p}, \frac{|x_{(i)} - \mu|^p}{p\sigma^p} \right) - \mathrm{sign}(-x_{(i)} - \mu)\gamma'_{p} \left(\frac{1}{p}, \frac{|x_{(i)} + \mu|^p}{p\sigma^p} \right) \right]. \tag{14}$$

Here, $\gamma_{\mu}'(\cdot)$, $\gamma_{\sigma}'(\cdot)$ and $\gamma_{p}'(\cdot)$ denote the partial derivatives of incomplete gamma function with respect to the parameters μ, σ and p, respectively.

Equations (9)-(11) include nonlinear functions of the parameters and thus numerical methods should be performed.

2.3. WLS method

The WLS estimators of the parameters μ, σ and p are obtained by minimizing the following function:

$$S_w = \sum_{i=1}^{n} w_i \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right)^2 \tag{15}$$

where w_i denotes the weights which are computed by

$$w_i = \frac{1}{Var(F(x_{(i)})} = \frac{(n+1)^2(n+2)}{i(n-i+1)}, \quad i = 1, 2, \ldots, n.$$

After incorporating the cdf of FEP distribution into the function S_w and taking the partial derivatives with respect to the parameters of interest, we obtain the following nonlinear equations:

$$\frac{\partial S_w}{\partial \mu} = \sum_{i=1}^{n} w_i \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_1(x_{(i)}; \mu, \sigma, p) = 0, \tag{16}$$

$$\frac{\partial S_w}{\partial \sigma} = \sum_{i=1}^{n} w_i \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_2(x_{(i)}; \mu, \sigma, p) = 0, \tag{17}$$

$$\frac{\partial S_w}{\partial p} = \sum_{i=1}^{n} w_i \left(F(x_{(i)}; \mu, \sigma, p) - \frac{i}{n+1} \right) \Delta_3(x_{(i)}; \mu, \sigma, p) = 0 \tag{18}$$

where Δ_1, Δ_2 and Δ_3 are given in equations (12), (13) and (14), respectively. Numerical methods should be utilized since equations (16)-(18) cannot be solved explicitly.

2.4. CVM method

CVM method is based on minimizing the distance between the estimated cdf and empirical distribution function. Therefore, CVM estimators are in the class of minimum distance estimators which are also known as maximum goodness of fit test, see i.e. [14] and references therein.

The CVM estimators of the parameters of FEP distribution are obtained by minimizing the following function with respect to the parameters μ, σ and p:

$$C = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{2i-1}{2n} \right)^2 \tag{19}$$

where $F(\cdot)$ is the cdf of FEP distribution. It is clear that this minimization problem yields following nonlinear equations:

$$\frac{\partial C}{\partial \mu} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{2i-1}{2n} \right) \Delta_1(x_{(i)}; \mu, \sigma, p) = 0, \tag{20}$$

$$\frac{\partial C}{\partial \sigma} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{2i-1}{2n} \right) \Delta_2(x_{(i)}; \mu, \sigma, p) = 0, \tag{21}$$

$$\frac{\partial C}{\partial p} = \sum_{i=1}^{n} \left(F(x_{(i)}; \mu, \sigma, p) - \frac{2i-1}{2n} \right) \Delta_3(x_{(i)}; \mu, \sigma, p) = 0. \tag{22}$$

Here, Δ_1, Δ_2 and Δ_3 are same as given in equations (12), (13) and (14), respectively. As it is clear from equations (20)-(22), the explicit solutions cannot be obtained and thus numerical methods should be employed.

2.5. MPS method

The MPS method is originated by Cheng and Amin [16] and Ranneby [17]. The MPS estimator is obtained by maximizing the geometric mean of spacings. The spacings are defined as follows:

$$D_i(\mu, \sigma, p) = F(x_{(i)}; \mu, \sigma, p) - F(x_{(i-1)}; \mu, \sigma, p), \quad i = 1, 2, \dots, n+1 \tag{23}$$

where $F(\cdot)$ is the cdf of FEP distribution, $x_{(0)} = -\infty$ and $x_{(n+1)} = +\infty$. Then, the MPS estimators of μ, σ and p are defined as maximizers of the following function:

$$D = \frac{1}{n+1} \sum_{i=1}^{n} \log D_i(\mu, \sigma, p). \tag{24}$$

The corresponding partial derivatives are obtained as follows:

$$\frac{\partial D}{\partial \mu} = \frac{1}{n+1} \sum_{i=1}^{n} \frac{\Delta_1(x_{(i)}; \mu, \sigma, p) - \Delta_1(x_{(i-1)}; \mu, \sigma, p)}{D_i(\mu, \sigma, p)} = 0, \tag{25}$$

$$\frac{\partial D}{\partial \sigma} = \frac{1}{n+1} \sum_{i=1}^{n} \frac{\Delta_2(x_{(i)}; \mu, \sigma, p) - \Delta_2(x_{(i-1)}; \mu, \sigma, p)}{D_i(\mu, \sigma, p)} = 0, \tag{26}$$

$$\frac{\partial D}{\partial p} = \frac{1}{n+1} \sum_{i=1}^{n} \frac{\Delta_3(x_{(i)}; \mu, \sigma, p) - \Delta_3(x_{(i-1)}; \mu, \sigma, p)}{D_i(\mu, \sigma, p)} = 0. \tag{27}$$

See equations (12), (13) and (14) for definitions of Δ_1, Δ_2 and Δ_3, respectively. It is clear that MPS estimators should also be obtained using numerical methods.

It should also be noted that if there are ties in the data set, i.e. $x_{(i)} = x_{(i-1)}$, $D_i(\mu, \sigma, p)$ is replaced by the corresponding log density [16].

3. Simulation study

This section consists of a Monte-Carlo simulation study to determine the estimator having the best performance. The bias and MSE criteria are used to compare the performances of the ML, LS, WLS, CVM and MPS estimators. The bias and MSE are respectively formulated as follows:

$$\text{Bias}(\hat{\theta}) = E(\theta - \hat{\theta}) \quad \text{and} \quad \text{MSE}(\hat{\theta}) = E(\theta - \hat{\theta})^2.$$

Here, θ shows true parameter value and stands for μ, σ or p. $\hat{\theta}$ denotes the estimator of θ obtained based on one of the methods given in the previous section. We also compute the joint efficiencies of the estimators using deficiency criterion which is formulated by

$$\text{Def} = MSE(\hat{\mu}) + MSE(\hat{\sigma}) + MSE(\hat{p}).$$

For further information on Def, see Kantar and Senoglu [13].

All computations are done using R software. Since the estimators cannot be obtained explicitly, we use “optim” function with “Nelder-Mead” algorithm. Random numbers are generated from FEP distribution using the “rpe” function in the “LaplacesDemon” package [18]. Without loss of generality, the true values of the parameters are taken as follows:

(i) $\mu = 2$ or $\mu = 5$, (ii) $\sigma = 1$ and (iii) $p = 0.8$, 1, 1.5, 2 and 2.5.

The sample size is taken to be $n = 100$ and 200. Therefore, we have 20 different simulation schemes. The results are tabulated in Table 1 – Table 4.

Conclusions obtained from Table 1:

All estimators of μ and σ have negligible biases. In other words, they are almost unbiased. However, estimators of p, except MPS, have bias. MPS estimator of p have lower bias values.

The MSE values of ML, LS, WLS, CVM and MPS estimators of μ are low and more or less the same. This conclusion is also true for those of σ. Therefore, all of the estimators can be preferred for estimating μ and σ. For p, we have similar conclusions given for bias. In other words, the MSEs of the estimators of p are relatively large compared to those of μ and σ. Furthermore, WLS has the worst performance. MPS estimator of p is the best since it has the minimum value of MSE among the others.

As indicated previously, Def is a measure of joint efficiency. Therefore, it can be seen as total efficiency measure for ML, LS, WLS, CVM and MPS methods. From this point of view, MPS outperforms its rivals. ML method is following MPS and also giving promising results.

Conclusions obtained from Table 2:

It is clear from bias values of μ given in second column of Table 2 that all estimators are almost unbiased. However, it should be noticed that ML method has larger bias values. All estimators of σ have relatively smaller biases. It should also be mentioned that CVM and MPS underestimates σ in most of the cases. We have comments similar to those obtained from Table 1 for the estimators of p in terms of bias criterion. They have larger biases and LS and WL are more preferable in most of the cases.

The MSEs of all estimators of μ are close to each other and all of them are preferable. Similarly, all estimators of σ can also be used confidentially. However, we cannot say the same for estimating p since the corresponding MSEs are large. Among all estimators, MPS estimator of p is more preferable.

According to the Def criterion, MPS is the best estimator and it is followed by ML. There are no remarkable efficiencies of WLS and CVM but the performance of LS method is promising in some cases, i.e. $p = 0.8$.

Conclusions obtained from Table 3 and Table 4 :

The conclusions obtained from Table 3 and Table 4 are similar those obtained from Table 1 and Table 2. In Table 3 and Table 4, the sample size is increased to 200 therefore all estimators gain efficiency. For example, biases, MSEs and therefore Def values decrease as expected. Results show that MPS is more preferable to its rivals in most of the cases. ML is the second best and LS presents relatively good results in some cases.

Overall, we recommend to use MPS method used to estimate the parameters of FEP distribution since it has the smallest bias, MSE and Def values in most of the cases.

Table 1. Simulated biases, MSEs and Def values of the ML, LS, WLS, CVM and MPS estimators for $\mu = 2$, $\sigma = 1$ and $n = 100$.

Estimation Method	μ		σ		p		Def
	Bias	MSE	Bias	MSE	Bias	MSE	
$p = 0.8$							
ML	0.0035	0.0108	-0.0017	0.0225	0.0549	0.0578	0.0911
LS	-0.0084	0.0108	0.0281	0.0240	0.0713	0.0865	0.1214
WLS	-0.0061	0.0119	0.0219	0.0247	0.0670	0.8715	0.9081
CVM	-0.0030	0.0107	0.0072	0.0223	0.0865	0.0979	0.1310
MPS	-0.0041	0.0114	0.0147	0.0232	-0.0198	0.0394	0.0739
$p = 1$							
ML	-0.0049	0.0125	0.0009	0.0184	0.0774	0.1034	0.1342
LS	-0.0170	0.0125	0.0242	0.0192	0.0670	0.1301	0.1618
WLS	-0.0135	0.0131	0.0184	0.0193	0.0685	1.0731	1.1055
CVM	-0.0122	0.0123	0.0047	0.0179	0.0930	0.1484	0.1786
MPS	-0.0130	0.0129	0.0157	0.0190	-0.0303	0.0676	0.0995
$p = 1.5$							
ML	0.0003	0.0137	0.0138	0.0178	0.2179	0.6035	0.6351
LS	-0.0098	0.0139	0.0223	0.0173	0.1027	0.5064	0.5376
WLS	-0.0060	0.0134	0.0165	0.0173	0.1188	1.6329	1.6636
CVM	-0.0072	0.0136	0.0072	0.0166	0.1641	0.6004	0.6307
MPS	-0.0072	0.0142	0.0176	0.0173	-0.0047	0.3261	0.3576
$p = 2$							
ML	-0.0044	0.0121	0.0252	0.0172	0.3314	0.9214	0.9508
LS	-0.0095	0.0126	0.0232	0.0150	0.1414	0.9621	0.9897
WLS	-0.0071	0.0118	0.0186	0.0146	0.1702	2.1992	2.2256
CVM	-0.0082	0.0124	0.0117	0.0147	0.2408	1.1537	1.1808
MPS	-0.0095	0.0125	0.0174	0.0146	-0.0045	0.5180	0.5452
$p = 2.5$							
ML	-0.0059	0.0092	0.0178	0.0171	0.3494	1.1601	1.1865
LS	-0.0060	0.0105	0.0094	0.0138	0.2253	2.4578	2.4821
WLS	-0.0055	0.0094	0.0054	0.0124	0.1919	2.7287	2.7505
CVM	-0.0054	0.0103	0.0003	0.0139	0.3618	2.7212	2.7454
MPS	-0.0077	0.0095	0.0015	0.0123	-0.0586	0.9356	0.9575

Table 2. Simulated biases, MSEs and Def values of the ML, LS, WLS, CVM and MPS estimators for $\mu = 5$, $\sigma = 1$ and $n = 100$.

Estimation Method	μ		σ		p		Def
	Bias	MSE	Bias	MSE	Bias	MSE	
$p = 0.8$							
ML	-0.0762	0.0099	-0.0092	0.0204	0.0131	0.0338	0.0641
LS	0.0055	0.0125	0.0229	0.0205	0.0146	0.0485	0.0815
WLS	0.0064	0.0136	0.0147	0.0204	0.0217	0.8222	0.8562
CVM	0.0064	0.0126	0.0059	0.0196	0.0431	0.0547	0.0869
MPS	0.0060	0.0102	0.0018	0.0190	-0.0674	0.0286	0.0577
$p = 1$							
ML	-0.0810	0.0107	-0.0106	0.0183	0.0105	0.0502	0.0792
LS	0.0024	0.0122	0.0142	0.0177	0.0278	0.0802	0.1101
WLS	0.0029	0.0127	0.0072	0.0177	0.0334	1.0345	1.0650
CVM	0.0027	0.0122	-0.0003	0.0174	0.0652	0.0920	0.1216
MPS	0.0015	0.0117	-0.0049	0.0171	-0.0835	0.0451	0.0738
$p = 1.5$							
ML	-0.0901	0.0124	0.0059	0.0140	0.1186	0.2061	0.2325
LS	0.0035	0.0117	0.0075	0.0151	0.0568	0.2770	0.3037
WLS	0.0028	0.0116	0.0028	0.0145	0.0609	1.5646	1.5908
CVM	0.0035	0.0117	-0.0037	0.0151	0.1205	0.3272	0.3540
MPS	0.0043	0.0125	-0.0091	0.0136	-0.1139	0.1162	0.1424
$p = 2$							
ML	-0.0838	0.0111	0.0105	0.0148	0.2480	0.7436	0.7695
LS	-0.0005	0.0117	0.0040	0.0137	0.0943	0.8095	0.8349
WLS	0.0004	0.0112	0.0006	0.0128	0.0932	2.1019	2.1259
CVM	-0.0005	0.0117	-0.0053	0.0139	0.1936	0.9926	1.0183
MPS	0.0003	0.0114	-0.0083	0.0119	-0.1309	0.3034	0.3268
$p = 2.5$							
ML	-0.0733	0.0085	0.0139	0.0152	0.3781	1.8727	1.8965
LS	0.0008	0.0095	0.0028	0.0155	0.1895	2.4575	2.4825
WLS	0.0008	0.0085	-0.0006	0.0139	0.1587	2.6839	2.7063
CVM	0.0007	0.0095	-0.0051	0.0158	0.3405	3.2004	3.2257
MPS	0.0006	0.0083	-0.0111	0.0124	-0.1708	0.6494	0.6701

Table 3. Simulated biases, MSEs and Def values of the ML, LS, WLS, CVM and MPS estimators for $\mu = 2, \sigma = 1$ and $n = 200$.

Estimation Method	μ		σ		p		Def
	Bias	MSE	Bias	MSE	Bias	MSE	
$p = 0.8$							
ML	-0.0017	0.0052	-0.0009	0.0125	0.0120	0.0187	0.0364
LS	-0.0077	0.0056	0.0139	0.0129	0.0234	0.0306	0.0491
WLS	-0.0065	0.0061	0.0100	0.0129	0.0181	0.8185	0.8375
CVM	-0.0050	0.0056	0.0033	0.0125	0.0292	0.0322	0.0503
MPS	-0.0050	0.0053	0.0075	0.0127	-0.0292	0.0166	0.0345
$p = 1$							
ML	-0.0014	0.0065	-0.0010	0.0092	0.0429	0.0355	0.0512
LS	-0.0057	0.0064	0.0104	0.0094	0.0530	0.0639	0.0797
WLS	-0.0042	0.0068	0.0071	0.0096	0.0440	1.0459	1.0622
CVM	-0.0034	0.0064	0.0007	0.0091	0.0652	0.0687	0.0842
MPS	-0.0064	0.0066	0.0064	0.0094	-0.0187	0.0283	0.0443
$p = 1.5$							
ML	0.0031	0.0059	0.0012	0.0069	0.0802	0.1021	0.1150
LS	-0.0007	0.0059	0.0067	0.0070	0.0393	0.1513	0.1643
WLS	0.0010	0.0059	0.0040	0.0070	0.0522	1.5549	1.5677
CVM	0.0004	0.0059	-0.0007	0.0069	0.0675	0.1643	0.1771
MPS	-0.0006	0.0060	0.0019	0.0068	-0.0376	0.0774	0.0902
$p = 2$							
ML	0.0009	0.0052	0.0100	0.0065	0.1262	0.2496	0.2613
LS	-0.0009	0.0058	0.0081	0.0067	0.0219	0.3589	0.3715
WLS	0.0001	0.0054	0.0070	0.0064	0.0541	2.0570	2.0688
CVM	-0.0004	0.0058	0.0025	0.0067	0.0661	0.3900	0.4025
MPS	-0.0014	0.0053	0.0042	0.0060	-0.0591	0.1720	0.1833
$p = 2.5$							
ML	0.0008	0.0042	0.0052	0.0064	0.1546	0.4281	0.4387
LS	0.0020	0.0050	0.0028	0.0072	0.0674	0.6629	0.6751
WLS	0.0018	0.0045	0.0011	0.0066	0.0739	2.5793	2.5904
CVM	0.0022	0.0050	-0.0015	0.0072	0.1292	0.7253	0.7375
MPS	0.0000	0.0043	-0.0045	0.0059	-0.0949	0.2915	0.3017

Table 4. Simulated biases, MSEs and Def values of the ML, LS, WLS, CVM and MPS estimators for $\mu = 5, \sigma = 1$ and $n = 200$.

Estimation Method	μ		σ		p		Def
	Bias	MSE	Bias	MSE	Bias	MSE	
$p = 0.8$							
ML	-0.0538	0.0048	-0.0068	0.0090	0.0054	0.0146	0.0284
LS	-0.0017	0.0062	0.0090	0.0091	0.0019	0.0211	0.0364
WLS	-0.0013	0.0068	0.0055	0.0091	0.0096	0.8097	0.8255
CVM	-0.0013	0.0062	0.0006	0.0089	0.0152	0.0222	0.0373
MPS	0.0008	0.0047	-0.0022	0.0088	-0.0439	0.0144	0.0279
$p = 1$							
ML	-0.0631	0.0062	-0.0057	0.0086	0.0032	0.0253	0.0401
LS	0.0009	0.0068	0.0049	0.0084	0.0063	0.0385	0.0537
WLS	0.0008	0.0072	0.0019	0.0083	0.0140	1.0142	1.0296
CVM	0.0010	0.0068	-0.0024	0.0083	0.0236	0.0409	0.0560
MPS	0.0020	0.0063	-0.0058	0.0082	-0.0545	0.0245	0.0389
$p = 1.5$							
ML	-0.0643	0.0065	0.0010	0.0072	0.0509	0.0580	0.0718
LS	-0.0003	0.0062	0.0052	0.0078	0.0320	0.1262	0.1402
WLS	-0.0004	0.0061	0.0032	0.0074	0.0381	1.5396	1.5531
CVM	-0.0003	0.0062	-0.0005	0.0078	0.0613	0.1368	0.1507
MPS	-0.0004	0.0062	-0.0057	0.0069	-0.0740	0.0519	0.0651
$p = 2$							
ML	-0.0584	0.0055	0.0048	0.0059	0.0984	0.1578	0.1692
LS	-0.0007	0.0058	0.0039	0.0072	0.0435	0.3222	0.3352
WLS	-0.0013	0.0054	0.0028	0.0066	0.0498	2.0523	2.0643
CVM	-0.0007	0.0058	-0.0008	0.0073	0.0869	0.3524	0.3655
MPS	-0.0018	0.0053	-0.0068	0.0060	-0.1061	0.1184	0.1297
$p = 2.5$							
ML	-0.0510	0.0041	-0.0005	0.0059	0.1202	0.2914	0.3015
LS	0.0008	0.0051	-0.0054	0.0070	0.0309	0.5913	0.6034
WLS	0.0006	0.0046	-0.0063	0.0063	0.0368	2.5381	2.5490
CVM	0.0008	0.0051	-0.0095	0.0071	0.0900	0.6484	0.6606
MPS	0.0009	0.0043	-0.0156	0.0058	-0.1685	0.2196	0.2297

4. Applications

In this section two real life data set taken from the literature is analyzed. It should be mentioned that we use ML and MPS methods in this section since they are more preferable in terms of efficiency, see Table 1 - Table 4. In applications, we also conduct Kolmogorov-Simornov tests. Therefore, test statistic (KS) and corresponding p −values obtained from Kolmogrov-Smirnov test are also provided. The method having smaller (larger) KS (p −) value is more desirable.

4.1. Lead wires data

This application is taken from Leone et al. [5] in which FN distribution is used. The data consists of camber of 497 lead wires which are used in manufacture of miniature radio tubes. The full data set and more explanation are available in Leone et al. [5]. See also Cooray et al. [8] in which FL distribution is considered for modelling purposes. Different from these studies, we here use FEP distribution to model lead wires data. Parameter estimates based on FEP, FN and FL distributions and corresponding KS test results are given in Table 5. It should be noted that parameter estimates are obtained using the ML and MPS methods for FEP distibution. On the other hand, results for FN and FL distributions are based on the ML method. It is clear from Table 5 that the ML is more desirable according to KS and p − values for FEP distribution. It should be recognized that the fitting performances of the ML and MPS methods are similar, see Figure 2 in which histogram of data is given along with fitted densities based on ML and MPS estimates. Cooray et al. [8] show that FL distribution is more appropriate than FN distribution for modelling this data. Results given in Table 5 provide that the FEP distribution can also be considered as an alternative to FL distribution.

Table 5. Parameter estimates, KS and p −values for lead wires data.

Distribution	Estimation Method	$\hat{\mu}$	$\hat{\sigma}$	$\hat{p}$	KS	p −value
FEP	ML	13.1291	7.7628	1.4488	0.0612	0.0482
	MPS	13.0694	7.7759	1.3951	0.0632	0.0375
FN	ML	13.6120	8.4626	-	0.0709	0.0135
FL	ML	13.2860	4.8400	-	0.0611	0.0487

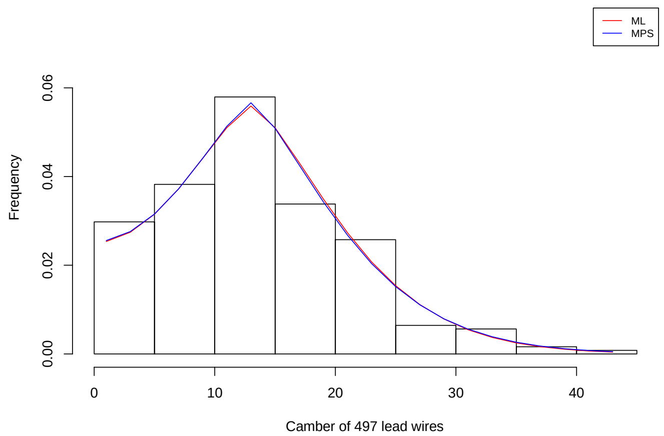


Figure 2. The histogram and fitted densities for lead wires data.

4.2. Log-Norwegian fire claim data

Brazauskas and Kleefeld [2] use this data in the context of insurance claims distributional modeling. The full data set is available at the following web site with name NORWEGIANFIRE.TXT: http://lstat.kuleuven.be/Wiley/. It should be noted that log-Norwegian fire claim data is obtained by taking logarithm of original data divided by 500 and is for the year 1988.

Nadarajah and Bakar [1] consider different folded distributions, including FEP, to model log-Norwegian fire claim data. They show that FEP is the best distribution modeling log-Norwegian fire claim by using ML estimates. In addition to ML, we use MPS method to estimate the parameters of FEP distribution in the current study. Results are given in Table 6. It is clear that MPS method is more preferable for this data set. See also Figure 3 in which histogram and fitted densities are provided.

Table 6. Parameter estimates, KS and p −values for log-Norwegian fire claim data.

Dsitribution	Estimation Method	$\hat{\mu}$	$\hat{\sigma}$	$\hat{p}$	KS	p −value
FEP	ML	0.6109	0.9322	1.1282	0.0213	0.8475
	MPS	0.6109	0.9259	1.0981	0.0211	0.8566

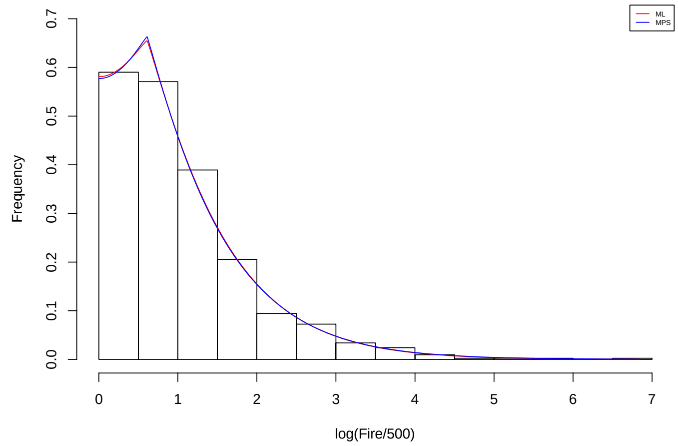


Figure 3. The histogram and fitted densities log-Norwegian fire claim data.

5. Conclusion

In this study, we consider the ML, LS, WLS, CVM and MPS methods to estimate the unknown parameters of FEP distribution. As far as we know, LS, WLS, CVM and MPS methods have not been used for FEP distribution previously. We aim to investigate the most efficient estimator in terms of some criteria such as bias, MSE and Def. Results of the simulation study show that MPS is slightly better than the other estimators. It should be noticed that the ML estimator has also a remarkable performance since it is the second best estimator. Two real life examples taken from literature are considered in the application, i.e. Lead wires and log Norwegian data. They are modelled using FEP distribution and the corresponding parameter estimates are obtained based on ML and MPS methods. KS test results show that these methods are desirable for modelling the data sets.

References

[1] S. Nadarajah, S.A.A. Bakar, 2015, New folded models for the log-transformed Norwegian fire claim data, *Communications in Statistics-Theory and Methods*, 44(20), 4408-4440.

[2] V. Brazauskas, A. Kleefeld, 2011, Folded and log-folded-t distributions as models for insurance loss data, *Scandinavian Actuarial Journal*, 2011(1), 59-74.

[3] M. Chen, F. Kianifard, 2003, Estimation of treatment difference and standard deviation with blinded data in clinical trials, *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 45(2), 135-142.

[4] S. Jung, M. Foskey, J. S. Marron, 2011, Principal arc analysis on direct product manifolds, *The Annals of Applied Statistics*, 5(1), 578-603.

[5] F. C. Leone, L. S. Nelson, R. B. Nottingham, 1961, The folded normal distribution, *Technometrics*, 3, 543-550.

[6] M. Y. Liao, 2010, Economic tolerance design for folded normal data, *International Journal of Production Research*, 48(14), 4123-4137.

[7] S. Psarakis, J. Panaretoes, 1990, The folded t distribution, *Communications in Statistics-Theory and Methods*, 19, 2717-2734.

[8] K. Cooray, S. Gunasekera, M. M. A. Ananda, 2006, The folded logistic distribution, *Communications in Statistics-Theory and Methods*, 35, 385-393.

[9] M. T. Subbotin, 1923, On the law of frequency of errors, *Matematicheskii Sbornik*, 31, 296-301.

[10] K. Cooray, 2010, Generalized Gumbel distribution, *Journal of Applied Statistics*, 37, 171-179.

[11] W. Gui, 2013, Statistical inferences and applications of the half exponential power distribution, *Journal of Quality and Reliability Engineering*, Article ID 219473, 9 pages.

[12] F. G. Akgul, 2018, Comparison of the estimation methods for the parameters of exponentiated reduced kies distribution, *Süleyman Demirel University Journal of Natural and Applied Sciences*, 22(3), 1209-1216.

[13] Y. M. Kantar, B. Senoglu, 2008, A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter, *Computers & Geosciences*, 34(12), 1900-1909.

[14] F. Louzada, P. L. Ramos, G. S. Perdona, 2016, Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data, *Computational and mathematical methods in medicine*, Article ID 8727951, 12 pages.

[15] J. J. Swain, S. Venkatraman, J. R. Wilson, 1988, Least-squares estimation of distribution functions in Johnson’s translation system, *Journal of Statistical Computation and Simulation*, 29(4), 271-297.

[16] R. C. H. Cheng, N. A. K. Amin, 1983, Estimating parameters in continuous univariate distributions with a shifted origin, *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(3), 394-403.

[17] B. Ranneby, 1984, The maximum spacing method. An estimation method related to the maximum likelihood method, *Scandinavian Journal of Statistics*, 11(2), 93-112.

[18] Statisticat, LLC., 2018, LaplacesDemon: Complete Environment for Bayesian Inference. Bayesian-Inference.com. R package version 16.1.1. https://web.archive.org/web/20150206004624/http://www.bayesian-inference.com/software.