A New Reflected Minimax Distribution on a Bounded Domain: Theory and Application

M. Sirajo[1]; J. Y. Falgore[2]; M. S. Najmuddeen; A. A. Umar

Department of Statistics,
Ahmadu Bello University, Zaria, Nigeria.
E-mail: musaddiqsirajo@gmail.com[1]; jyfalgore@gmail.com[2]

Abstract* — In this paper, we introduce a new flexible family of distributions with bounded support, called reflected minimax distribution (RMD), obtained by reflecting Minimax distribution (Jones, 2009) about the y-axis and shifting it $\theta > 0$ units to the right. We proposed standard reflected minimax distribution (SRMD) from this family which includes, as special cases, some distributions on (0,1) support such as the one-parameter minimax distribution, the power distribution, and uniform distribution. This work is an attempt to partially fill a gap regarding the issue of tractability of continuous distributions with bounded support, which appear to be very useful in many real contexts. Some properties of the family, including moments, hazard rate and quantile are investigated. Moreover, the estimators of the parameters are examined using the least squares, maximum likelihood and maximum product of spacings estimation methods, and an application reported.

Keywords: Minimax distribution, Reflected minimax distribution, Failure rate function, Moments, Maximum product of spacings.

I. Introduction

Advances had been made towards the development of standard distributions, with continued renewed interest in more flexible probability distributions. Most of the new proposals in the literature are distributions with unbounded support. According to [1], in the face of the numerous proposals of distributions with unbounded support emerges, undoubtedly, the great scarcity of distributions with bounded support.

There are many real-life situations in which the observations clearly can take values only in a limited range, such as percentages, proportions or fractions. This is often encountered in the study of the finite lifetime of a component or the bounded signals occurring in industrial systems. In this perspective, the models with infinite (unbounded) support can be viewed as an unrealistic approximation of the reality.

Beta and Kumaraswamy distributions, see [2], are the most used laws on bounded support. The latter was referred to as the two-parameter minimax (TPM) distribution in [3], and shares many desirable properties with some tractability advantages over the beta distribution. The major drawback of the beta distribution is the use of special functions in its implementation; the incomplete beta function ratio. Yet, it is still unclear if the tractability advantages of TPM distribution over the beta will be of immense practical significance.

In this paper, we propose a three-parameter reflected minimax distribution (RMD) with the hope of adding to the flexibility of TPM distribution for better application in practical situations.

Other less known two-parameter distributions on bounded support include the standard Two-sided Power distribution by [4], the Log–Lindley distribution by [5, 6] and Log-shifted Gompertz distribution by [7]. Proposals with more parameters include the three-parameter reflected generalized Topp–Leone power series distribution by [8], the four-parameter exponentiated Kumaraswamy-power function distribution by [9] and the five-parameter Kumaraswamy generalized gamma distribution by [10], among others.

.

II. Methodology

A. *Defining the Reflected Minimax Distribution*

The probability density function (pdf) of the TPM distribution is

$$g(x) = \alpha\beta x^{\alpha-1}\left(1-x^{\alpha}\right)^{\beta-1}; \qquad 0 < x < 1.$$

Reflecting the TPM distribution about the y-axis and shifting it $\theta > 0$ units to the right gives a pdf of

$$f\left(x \mid \alpha,\beta,\theta\right) = \alpha\beta\left(\theta - x\right)^{\alpha-1}\left[1-\left(\theta-x\right)^{\alpha}\right]^{\beta-1}, \qquad (1)$$

with corresponding cumulative distribution function (cdf)

$$F(x\mid\alpha,\beta,\theta)=\left[1-(\theta-x)^{\alpha}\right]^{\beta};\quad \theta-1<x<\theta, \tag{2}$$

where $\theta>0$ is the reflecting parameter that will reflect the distribution from positive skewness to negative and vice versa, while $\alpha,\beta>0$ are shape parameters.

B. *Standard Reflected Minimax Distribution*

We define the standardized version of RMD (SRMD) by substituting $\theta=1$ in Eqs. (1) and (2), with the pdf

$$f(x\mid\alpha,\beta)=\alpha\beta(1-x)^{\alpha-1}\left[1-(1-x)^{\alpha}\right]^{\beta-1}, \tag{3}$$

and the cdf

$$F(x\mid\alpha,\beta)=\left[1-(1-x)^{\alpha}\right]^{\beta};\quad 0<x<1. \tag{4}$$

This definition is particularly important because it gives a direct reflection of the TPM distribution studied in [9].

C. *Special Cases of SRMD*

Case 1: When $\beta=1$, SRMD reduces to one-parameter minimax (OPM) distribution with pdf given as

$$f(x\mid\alpha)=\alpha(1-x)^{\alpha-1};\quad 0<x<1,\ \alpha>0. \tag{5}$$

Case 2: When $\alpha=1$, SRMD reduces to Power distribution with pdf given as

$$f(x\mid\beta)=\beta x^{\beta-1};\quad 0<x<1,\ \beta>0. \tag{6}$$

Case 3: When $\alpha=\beta=1$, SRMD reduces to Uniform distribution with probability unity.

D. *Shapes of SRMD*

The SRMD can also be shown to have some basic shape properties as the TPM and beta distributions, which could be constant, increasing, decreasing, unimodal, and uni-antimodal.

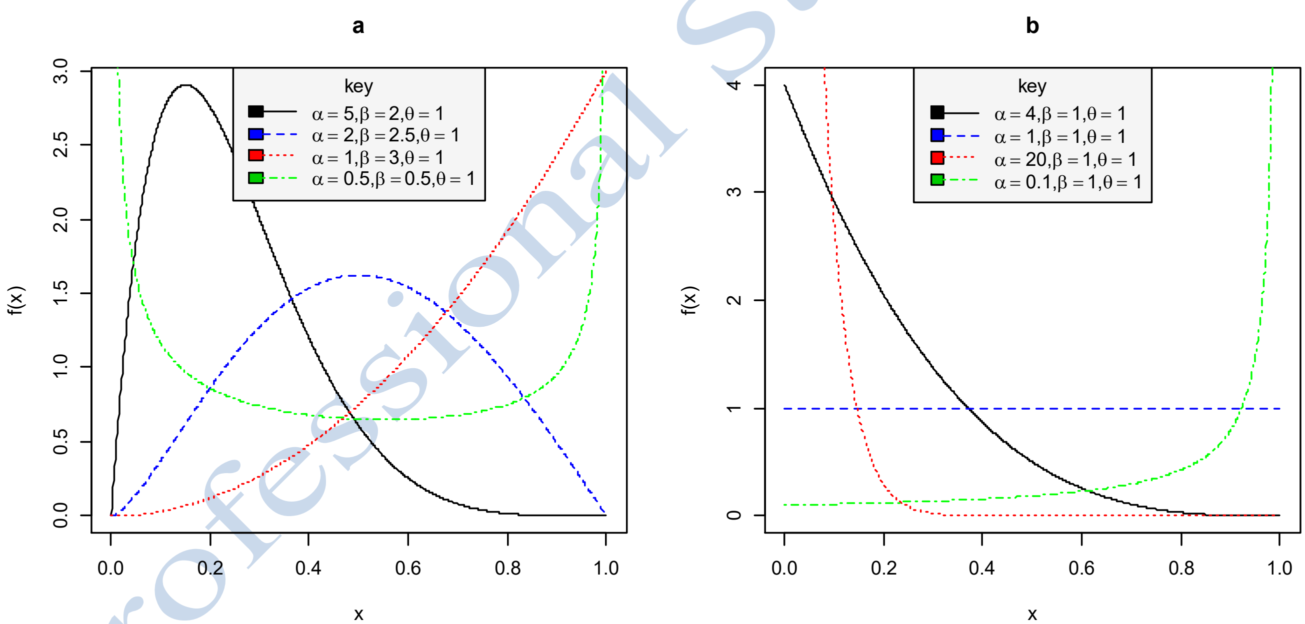

Figure 1: Plots of the SRMD density functions at different paramters values.

III. Prperties of RMD

E. *Failure Rate Function*

From Eqs. (1) and (2) it follows that the failure rate (also called hazard rate) function for an RMD density is:

$$r(x)=\frac{\alpha\beta(\theta-x)^{\alpha-1}\left[1-(\theta-x)^{\alpha}\right]^{\beta-1}}{1-\left[1-(\theta-x)^{\alpha}\right]^{\beta}}.$$

Setting $\beta=1$ gives

$$\Omega(x;\alpha,\theta)=\frac{\alpha}{\theta-x} \qquad (8)$$

which, on setting $\theta=1$, is the failure rate of OPM distribution by [11]. From Eq. (8) one observes that varying values of the parameters α and θ in $\Omega(x;\alpha,\theta)$ would shrink (or stretch) the failure rate of an RMD as compared to an OPM distribution. Thus, θ may be interpreted as a *shrinkage parameter,* i.e., a larger value of θ causes the failure rate to stretch out (to the right) while the smaller value of θ results (continuously) in a drift of the failure rate towards the left-hand side. The failure rate function is generally J-shaped and monotonically increasing.

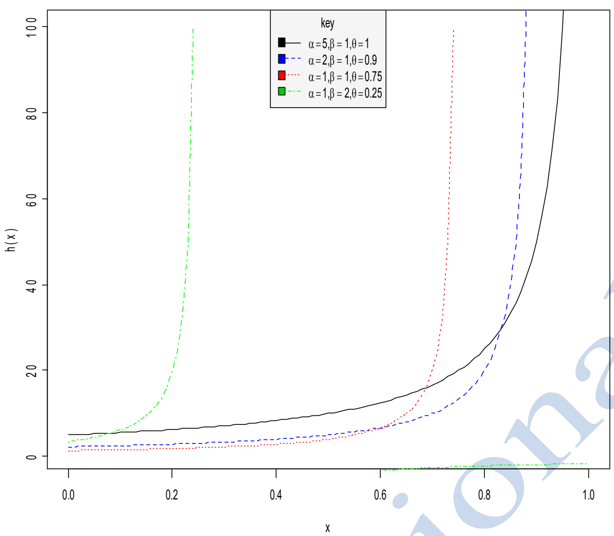

Figure 2: Plots of the RMD failure rate

F. ***Moments***

It would be better to find the moment of the quantity $(\theta-X)^r$ first, and then appropriately manipulate it to obtain the moment of X^r; $r=1,2,\ldots$

$$E(\theta-X)^r=\beta B\left(1+\frac{r}{\alpha},\beta\right). \qquad (9)$$

When $r=1$ in (9), the mean of RMD may be expressed as

$$\mu=E(X)=\theta-\beta B\left(1+\frac{1}{\alpha},\beta\right). \qquad (10)$$

Substituting $\theta=1$ in Eq. (10) yields the mean of SRMD

$$\mu_{\text{SRMD}}=1-\beta B\left(1+\frac{1}{\alpha},\beta\right),$$

where $\beta B\left(1+\frac{1}{\alpha},\beta\right)$ is the mean of a TPM distribution.

By using appropriate moment expressions, the variance of the RMD is obtained as

$$\sigma^2=\beta B\left(1+\frac{2}{\alpha},\beta\right)-\left\{\beta B\left(1+\frac{1}{\alpha},\beta\right)\right\}^2,$$

which is free from the *reflecting parameter* θ, and equivalent to that of the TPM distribution.

G. ***Quantile Function and Median***

The cdf defined in Eq. (2) is readily invertible to obtain the quantile function of RMD as

$$Q(u)=\theta-\left(1-u^{1/\beta}\right)^{1/\alpha}; \quad 0<u<1. \qquad (11)$$

The median of RMD is obtained by setting $u=0.5$ in Eq. (11)

$$\text{Median}=\theta-\left(1-0.5^{1/\beta}\right)^{1/\alpha}.$$

IV. PARAMETER ESTIMATION

In order to estimate the shape parameters (α,β) of the RMD, we consider the methods of least squares, the maximum product of spacings, and maximum likelihood.

A. ***Method of Least-Squares (MLS)***

Suppose $F(X_{i:n})$ denotes the distribution function of the ordered random variables $X_{1:n}<X_{2:n}<\ldots<X_{n:n}$ where $\{X_1,X_2,\ldots,X_n\}$ is a random sample of size n from the distribution function $F(x\,|\,\alpha,\beta)$. The least-squares estimators of α and β can be obtained by minimizing with respect to α and β, the function

$$\sum_{i=1}^{n}\left[F(x_{i:n}\,|\,\alpha,\beta)-\frac{i}{n+1}\right]^2.$$

B. ***Method of Maximum Product of Spacings (MPS)***

Using the same notation as above, we define the uniform spacings of a random sample from the RMD as:

$$D_i(\alpha,\beta)=F(x_{i:n}\,|\,\alpha,\beta)-F(x_{i-1:n}\,|\,\alpha,\beta),\ i=1,2,\ldots,n$$

where $F(x_{0:n}\,|\,\alpha,\beta)=0$, $F(x_{n+1:n}\,|\,\alpha,\beta)=1$ and $\sum_{i=1}^{n+1}D_i(\alpha,\beta)=1$.

following [12] and the reference therein, the maximum product of spacings estimates of α and β are obtained by maximizing, with respect to α and β, the geometric mean of the spacings:

$$G(\alpha,\beta)=\left[\prod_{i=1}^{n+1}D_i(\alpha,\beta)\right]^{\frac{1}{n+1}}.$$

*C. **Method of Maximum Likelihood (MML)***

Let $X_1, X_2, ..., X_n$ be a random sample of size n from a RMD with the parameters α and β unknown, and denote by $x_1, x_2, ..., x_n$ the observed values. The log-likelihood function is obtained as

$$\log L(\alpha,\beta)=n\log\alpha+n\log\beta+(\alpha-1)\sum_{i=1}^{n}\log(\theta-x_i)+(\beta-1)\sum_{i=1}^{n}\log\left[1-(\theta-x_i)^{\alpha}\right].$$

The ML estimates of α and β are the values $\hat{\alpha}$ and $\hat{\beta}$ that maximize the $\log L(\alpha,\beta)$. Thus,

$$\hat{\beta}=-n\bigg/\sum_{i=1}^{n}\log\left[1-(\theta-x_i)^{\alpha}\right],$$

and $\hat{\alpha}$ is obtained by optimizing the score equation (Jones, 2009), given by

$$S(\alpha)=\left\{\frac{1}{\alpha}+T_1(\alpha)+\frac{T_2(\alpha)}{T_3(\alpha)}\right\}=0, \quad \text{where}$$

$$T_1(\alpha)=n^{-1}\sum_{i=1}^{n}\frac{\log(\theta-x_i)}{1-y_i}, \quad T_2(\alpha)=n^{-1}\sum_{i=1}^{n}\frac{y_i\log(\theta-x_i)}{1-y_i},$$

$$T_3(\alpha)=n^{-1}\sum_{i=1}^{n}\log(1-y_i),$$

and $y_i=\theta-x_i^{\alpha}$, $i=1,...,n$.

The estimate $\hat{\alpha}$ cannot be obtained in closed form and therefore, the use of MCMC procedure may be adopted.

V. SIMULATION

We compare the performance of the proposed estimation methods based on the simulation results of 1000 independent replications. Results are summarized in Tables 1-2 for different values of n, α, β and θ.

Table 1: MPS Estimates

	$\alpha=5$		$\beta=2$		$\theta=1$	
	$\hat{\alpha}$	RMSE	$\hat{\beta}$	RMSE	$\hat{\theta}$	RMSE
$n=25$	4.75	1.07	2.25	1.47	0.99	0.06
$n=50$	4.84	0.75	2.15	0.95	0.99	0.03
$n=200$	4.95	0.39	2.05	0.37	1.00	0.01
$n=500$	4.97	0.24	2.02	0.19	1.00	0.01

Table 2: MLS Estimates

	$\alpha=5$		$\beta=2$		$\theta=1$	
	$\hat{\alpha}$	RMSE	$\hat{\beta}$	RMSE	$\hat{\theta}$	RMSE
$n=25$	5.48	0.49	2.18	0.19	0.52	0.48
$n=50$	5.64	0.65	2.09	0.12	0.46	0.54
$n=200$	5.53	0.54	2.08	0.09	0.35	0.65
$n=500$	5.90	0.90	2.03	0.05	0.32	0.68

We observe that MML gives approximately the same estimates as MPS and performed better with less bias and root-mean-square error, than those obtained by MLS. For the sake of saving space, we present only the numerical results obtained using MLS and MPS in the particular case $\theta=1$.

VI. APPLICATION

We have demonstrated the performance of RMD using Antimicrobial Resistance data available in the report (European Centre for Disease Prevention and Control, see [13]). The data represent the annual percentage of antimicrobial resistant isolates in Portugal in the year 2012: 0.01, 0.01, 0.03, 0.05, 0.08, 0.12, 0.14, 0.15, 0.15, 0.16, 0.19, 0.20, 0.20, 0.23, 0.26, 0.30, 0.32, 0.36, 0.39, 0.43, 0.54, 0.58, 0.59. We compared the standardized version of RMD (SRMD) with TPM, beta, Log-shifted Gompertz (LG), Log-Lindley (LL), and Two-sided Power (TP) distributions, and obtained a better fit for the SRMD on the basis of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The results are presented in Table 3.

Table 3: Antimicrobial Resistance Model Fitting Estimates

Distributions	ML Estimates	Log-lik	AIC	BIC
$\text{SRMD}(\alpha,\beta)$ $\alpha\beta(1-x)^{\alpha-1}\left[1-(1-x)^{\alpha}\right]^{\beta-1}$	$\hat{\alpha}=3.538,\ \hat{\beta}=1.116$	11.52	-17.04	-18.36
$\text{TPM}(\alpha,\beta)$ $\alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}$	$\hat{\alpha}=0.8649, \hat{\beta}=2.1116$	7.56	-11.13	-12.45
$\text{beta}(a,b)$ $\frac{1}{\text{B}(a,b)}x^{a-1}(1-x)^{b-1}$	$\hat{a}=0.8562, \hat{b}=2.1635$	7.52	-11.05	-12.37
$\text{LG}(\alpha,\beta)$ $\beta x^{\beta}\left[1+\alpha(1-x^{\beta})\right]e^{-\alpha x^{\beta}}$	$\hat{\alpha}=2.5361, \hat{\beta}=1.0693$	8.41	-12.83	-14.14
$\text{LL}(a,b)$ $a\left[b+a(b-1)\log x\right]x^{a-1}$	$\hat{a}=1.0918, \hat{b}=0.0643$	7.87	-11.74	-13.06
$\text{TP}(\theta,\nu)$ $\begin{cases}\nu\left(\frac{x}{\theta}\right)^{\nu-1}, 0<x\le\theta\\ \nu\left(\frac{1-x}{1-\theta}\right)^{\nu-1}, \theta\le x<1\end{cases}$	$\hat{\theta}=0.010, \hat{\nu}=2.5168$	7.68	-11.37	-12.69

VII. CONCLUSION

A three-parameter probability distribution defined on the bounded domain is derived from the two-parameter minimax distribution. A standardized version of the proposed distribution is also defined on (0, 1), with some special cases on the same support. The new family of distributions has tractable properties. Analytical expressions are provided for the moments, hazard rate and quantile functions. The parameters can be easily estimated by the method of maximum product of spacings (or the alternative method of maximum likelihood), which provided better results than the method of least squares.

Acknowledgment

The authors are grateful to anonymous reviewers for their valuable comments on the original draft of this manuscript

References

[1] Marshall, A. W., Olkin, I. (2007). Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families. New York: Springer.

[2] Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, Vol. 46, 79-88.

[3] Jones M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. Statistical Methodology, 6, 70-81.

[4] Van Dorp, J.R., Kotz, S. (2002). The standard two-sided power distribution and its properties: with applications in financial engineering. Amer. Statist. 56 (2), 90–99.

[5] Gómez-Déniz, E., Sordo, M.A., Calderín-Ojeda, E. (2014). The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance. Insur. Math. Econ. 54, 49–57.

[6] Jodrá, P., Jiménez-Gamero, M.D. (2016). A note on the Log-Lindley distribution. Insur. Math. Econ. 71, 189–194.

[7] Jodrá, P. (2020). A bounded distribution derived from the shifted Gompertz law. Journal of King Saud University – Science 32 (2020), 523–536.

[8] Condino, F., Domma, F., Latorre, G. (2014). Likelihood and Bayesian estimation of P(Y < X) using lower record values from a proportional reversed hazard family. Submitted.

[9] Bursa, N., Ozel, G. (2017). The exponentiated Kumaraswamy-power function distribution. Hacet. J. Math. Stat. 46 (2), 277–292.

[10] Pascoa, M.A.R., Ortega, E.M.M., Cordeiro, G.M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. Stat. Methodol. 8, 411–433.

[11] Oguntunde, P. E., Adejumo, A. O. (2015). A Note on the Minimax Distribution. *Covenant Journal of Physical and Life Sciences. 3(1). 1-8.*

[12] Dey, S., Dey. T., Ali, S., Mulekar, M. S. (2016). Two-parameter Maxwell distribution: Properties and different methods of estimation. J Stat Theory Prac, 10, 291-310.

[13] European Centre for Disease Prevention and Control, 2013. Antimicrobial resistance surveillance in Europe 2012, Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). ECDC, Stockholm. URL https://ecdc.europa.eu