International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:9, No:7, 2015

A New Distribution and Application on the Lifetime Data

Gamze Ozel, Selen Cakmakyapan

***Abstract*—**We introduce a new model called the Marshall-Olkin Rayleigh distribution which extends the Rayleigh distribution using Marshall-Olkin transformation and has increasing and decreasing shapes for the hazard rate function. Various structural properties of the new distribution are derived including explicit expressions for the moments, generating and quantile function, some entropy measures, and order statistics are presented. The model parameters are estimated by the method of maximum likelihood and the observed information matrix is determined. The potentiality of the new model is illustrated by means of a simulation study.

***Keywords*—**Marshall-Olkin distribution, Rayleigh distribution, estimation, maximum likelihood.

I. Introduction

THE Rayleigh distribution is a special case of the Weibull distribution and useful for several areas including engineering, statistics, life testing and reliability which age with time as its hazard rate is a linear function of time. It is a popular distribution for the lifetime of components that age rapidly with time since its probability density function (pdf) has a linearly increasing failure rate. The Rayleigh distribution was originally introduced by Rayleigh [1] in the field of acoustics; since its introduction, many researchers have used the distribution in different fields of science and technology.

The pdf of a Rayleigh random variable X, with scale parameter $\sigma > 0$ is given by

$$f(x) = \frac{x}{\sigma^2}\exp\left(\frac{-x^2}{2\sigma^2}\right),\ x > 0 \tag{1}$$

and the corresponding cumulative distribution function (cdf) of X is

$$F(x) = 1 - \exp\left(\frac{-x^2}{2\sigma^2}\right),\ x > 0\ . \tag{2}$$

Several extensions of the Rayleigh distribution proposed in the literature and these extensions provided great flexibility in modelling data in practice. Vodà [2] introduced the generalized Rayleigh distribution, its mathematical properties, and left-truncated form of the distribution. Leao et al. [3] derived the beta inverse Rayleigh distribution. Recently, Ahmad et al. [4] have proposed the transmuted inverse Rayleigh distribution and some of its properties.

The parameter(s) induction to the baseline distribution has received increased attention in recent years to explore properties and for efficient estimation of the parameters. Marshall & Olkin [5] introduced a new family of distributions by adding a new parameter. The resulting distribution, known as Marshall-Olkin distribution, includes the baseline distributions are also known as the proportional odds family or family with parameter.

Let $\bar{F}(x) = 1 - F(x)$ be the survival function of any distribution, then, the survival function of Marshall-Olkin (MO) family of distributions is given by

$$\bar{G}(x) = \frac{\gamma\bar{F}(x)}{1-(1-\gamma)\bar{F}(x)} \tag{3}$$

where γ is an additional positive parameter and Marshall & Olkin [5] have called it the tilt parameter. For $\gamma = 1$, we have $G = F$. The cdf and pdf for the new distribution are given by

$$G(x) = \frac{F(x)}{\left[1-(1-\gamma)\bar{F}(x)\right]^2}, \tag{4}$$

$$g(x) = \frac{\gamma f(x)}{\left[1-(1-\gamma)\bar{F}(x)\right]^2}, \tag{5}$$

respectively.

The main aim of this paper is to provide another extension of the Rayleigh distribution using the MO transformation. We propose the new Marshall-Olkin Rayleigh ("MOR" for short) distribution by adding extra parameter to the Rayleigh model. The objectives of the research are to study some structural properties of the proposed distribution.

II. Properties of Marshall-Olkin Rayleigh Distribution

A. The Probability Density and Cumulative Distribution Functions

We obtain the cdf of the MOR distribution by inserting (2) in (4) as

$$G(x) = \frac{1-\exp\left(-\frac{x^2}{2\sigma^2}\right)}{1-(1-\gamma)\exp\left(-\frac{x^2}{2\sigma^2}\right)}, \tag{6}$$

Gamze Ozel is with the Department of Statistics, Hacettepe University, 06800, Ankara, Turkey (corresponding author to provide phone: +90312-299-20-16/139; fax: +90312-297-79-13; e-mail: gamzeozl@ hacettepe.edu.tr).

Selen Cakmakyapan is with the Department of Statistics, Hacettepe University, 06800, Ankara, Turkey (e-mail: cakmakyapan@hacettepe.edu.tr).

where $x, \sigma, \gamma > 0$. Then, the pdf of corresponding to (6) is given by

$$g(x) = \frac{\gamma x \exp\left(-\frac{x^2}{2\sigma^2}\right)}{\sigma^2\left[1-(1-\gamma)\exp\left(-\frac{x^2}{2\sigma^2}\right)\right]^2}, \tag{7}$$

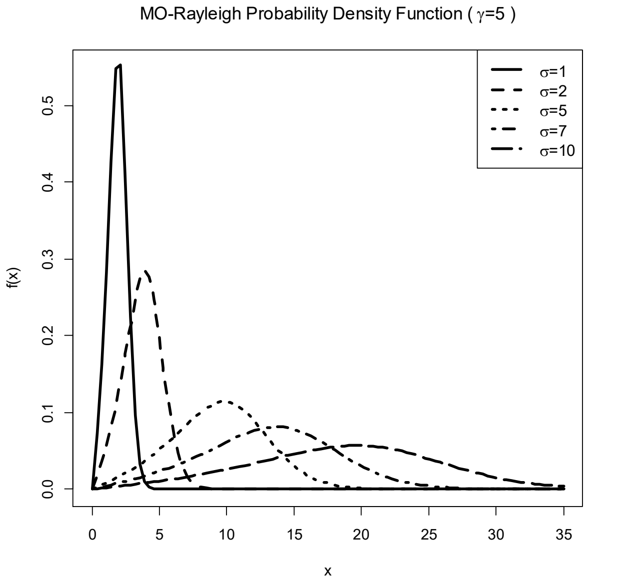

Fig. 1 The probability density function of the MOR distribution for several parameter values (for fixed $\gamma = 5$)

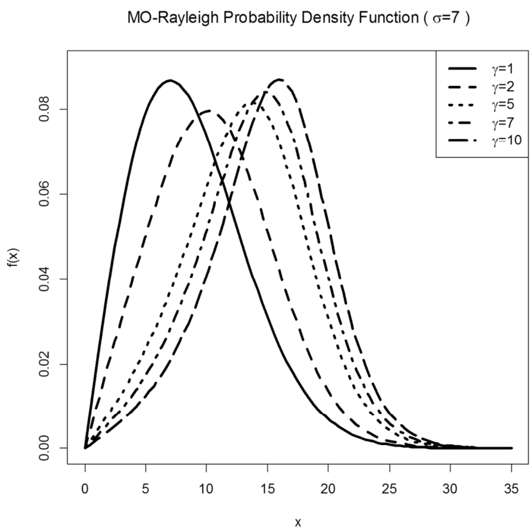

Fig. 2 The probability density function of the MOR distribution for several parameter values (for fixed $\sigma = 7$)

Fig. 1 shows the plot for the probability density function of the MOR distribution for several parameter values. As seen from Figs. 1 and 2, the density function can take various forms depending on the parameter values. Its density function is symmetrical, left-skewed, and right-skewed. It is evident that the MOR distribution is much more flexible than the Rayleigh distribution, i.e. the additional parameter γ allow for a high degree of flexibility of the MOR distribution.

B. Survival Function

The survival function $\bar{G}(x)$ for the MOR distribution are given by

$$\bar{G}(x) = \frac{\gamma \exp\left(-\frac{x^2}{2\sigma^2}\right)}{1-(1-\gamma)\exp\left(-\frac{x^2}{2\sigma^2}\right)} \tag{8}$$

The MOR distribution can be applied in survival analysis, hydrology, economics, among others, as the Rayleigh distribution and can be used to model reliability problems.

The other characteristic of the random variable is the hrf which is an important quantity characterizing life phenomenon. It can be loosely interpreted as the conditional probability of failure, given it has survived to time t. Then, the hrf of X is given by

$$h(x) = \frac{x}{\sigma^2\left[1-(1-\gamma)\exp\left(-\frac{x^2}{2\sigma^2}\right)\right]} \tag{9}$$

Figs. 3 and 4 illustrate some of the possible shapes of the hazard functions of the MOR distribution, respectively, for several parameter values. We can verify that this distribution can have an increasing hrf depending on the values of its parameters.

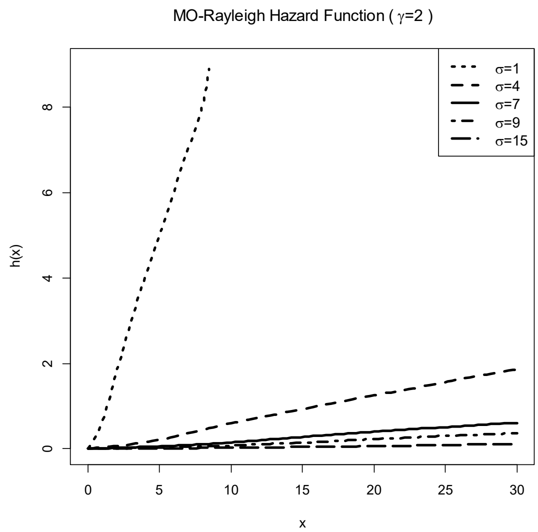

Fig. 3 The hazard functions of the MOR distribution for several parameter values (for fixed $\gamma = 2$)

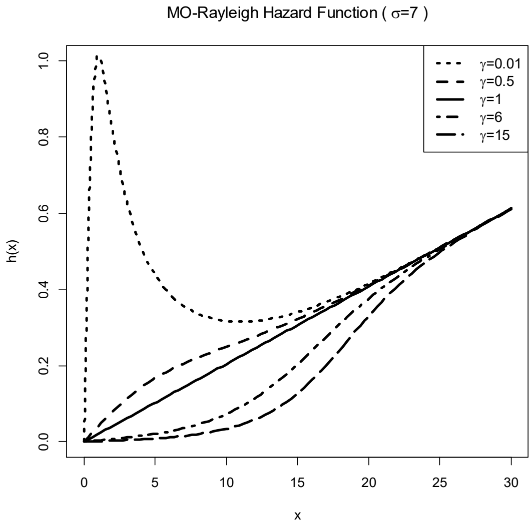

Fig. 4 The hazard functions of the MOR distribution for several parameter values (for fixed $\sigma = 7$)

C. Quantile Function

Quantile functions are in widespread use in general statistics and often find representations in terms of lookup tables for key percentiles. Inverting $F(x) = u$, the quantile function of the Rayleigh distribution is given by

$$F^{-1}(u) = \left[-2\sigma^2 \log\left(1-u\right)\right]^{1/2}, \quad 0 < u < 1.$$

Let $G(x)$ be the cdf of MOR and let $G^{-1}(u)$ be the quantile function of the distribution for $0 < u < 1$ and obtained as

$$G^{-1}(u) = \left\{\left[-2\sigma^2\left[\ln(1-u) - \ln(1-u+u\gamma)\right]\right]\right\}^{1/2} \tag{10}$$

D. Entropy Function

The entropy of a random variable X with density function $f(x)$ is a measure of variation of the uncertainty [6], [7]. Here, we derive expressions for the Rényi entropies of the MOR distribution. The Rényi entropy of a random variable with $f(x)$ is defined as

$$I_R(\gamma) = \frac{\delta}{1-\delta}\log\int_0^{\infty} g^{\delta}(x)\,dx \tag{11}$$

for $\delta > 0$ which implies that

$$g^{\delta}(x) = \frac{\gamma^{\delta} f^{\delta}(x)}{\Gamma(2\delta)}\sum_{j=0}^{\infty}(1-\gamma)^j\Gamma(2\delta+j)\frac{\left[1-F(x)\right]^j}{j!}.$$

Then, the Rényi entropy for $\gamma \in (0,1)$ is given as

$$I_R(\gamma) = \frac{\delta}{1-\delta}\log\left[\sum_{j=0}^{\infty} e_j \frac{\sigma^{1-\delta}}{(\delta+j)^{\frac{\delta-1}{2}+1}} 2^{\frac{\delta-1}{2}}\Gamma\left(1+\frac{\delta-1}{2}\right)\right] \tag{12}$$

where

$$e_j = \frac{\gamma^{\delta}(1-\gamma)^j\Gamma(2\delta+j)}{\Gamma(2\delta)j!}.$$

For $\delta > 1$, we have

$$g^{\delta}(x) = \frac{f^{\delta}(x)}{\gamma^{\delta}\Gamma(2\delta)}\sum_{j=0}^{\infty}(1-\gamma)^j\Gamma(2\delta+j)\frac{\left[F(x)\right]^j}{j!}.$$

Then, the Rényi entropy for $\gamma > 1$ is given as

$$I_R(\gamma) = \frac{\delta}{1-\delta}\log\left(\sum_{j=0}^{\infty} h_j \int_0^{\infty} f^{\delta}(x)F^j(x)\,dx\right)$$

$$I_R(\gamma) = \frac{\delta}{1-\delta}\log\left[\sum_{j=0}^{\infty}\sum_{i=0}^{j} h_j(-1)^i\binom{j}{i}\frac{\sigma^{1-\delta}}{(\delta+i)^{\frac{\delta-1}{2}+1}} 2^{\frac{\delta-1}{2}}\Gamma\left(1+\frac{\delta-1}{2}\right)\right]$$

where

$$h_j = \frac{(\gamma-1)^j}{\gamma^{\delta}\Gamma(2\delta)j!}.$$

III. Maximum Likelihood Estimation

Several approaches for parameter estimation have been proposed in the literature but the maximum likelihood method is the most commonly employed. Here, we consider estimation of the unknown parameters of the MOR distribution by the method of maximum likelihood. Let $x_1, x_2, ..., x_n$ be observed values from the MOR distribution with parameters γ and σ. The likelihood function for (γ, σ) is given by

$$L = \prod_{i=1}^{n}\left\{\frac{\gamma x_i \exp(-\frac{x_i^2}{2\sigma^2})}{\sigma^2\left[1-(1-\gamma)\exp(-\frac{x_i^2}{2\sigma^2})\right]^2}\right\} \tag{13}$$

and the corresponding log-likelihood function is given by

$$\log L = n\log\gamma + \sum_{i=1}^{n}\log x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n}\log\left[1-(1-\gamma)\exp(-\frac{x_i^2}{2\sigma^2})\right]. \tag{14}$$

The first derivatives of the log-likelihood function with respect to the parameters γ and σ are

$$\frac{\partial \log L}{\partial \gamma} = \frac{n}{\gamma} - \sum_{i=1}^{n} \frac{2\exp(-\frac{x_i^2}{2\sigma^2})}{\left[1-(1-\gamma)\exp(-\frac{x_i^2}{2\sigma^2})\right]},$$

$$\frac{\partial \log L}{\partial \sigma} = \frac{1}{\sigma^3}\sum_{i=1}^{n} x_i^2 - \frac{2(1-\gamma)}{\sigma^3}\sum_{i=1}^{n} \frac{x_i^2\exp(-\frac{x_i^2}{2\sigma^2})}{\left[1-(1-\gamma)\exp(-\frac{x_i^2}{2\sigma^2})\right]}.$$

The MLEs of (γ,σ), say $(\hat{\gamma},\hat{\sigma})$, are the simultaneous solutions of the equations $\frac{\partial \log L}{\partial \gamma}=0$ and $\frac{\partial \log L}{\partial \sigma}=0$. Maximization of (14) can be performed by using *nlm* or *optimize* in R statistical package. For interval estimation of (γ,σ) and tests of hypothesis, we require the Fisher information matrix.

IV. Simulation Study

We conduct Monte Carlo simulation studies to assess on the finite sample behavior of the Mean Squared Errors (MSEs) of γ and σ. All results were obtained from 1000 Monte Carlo replications and the simulations were carried out using the statistical software package R. In each replication, a random sample of size n is drawn from $X \sim MO-Rayleigh(\gamma,\sigma)$ distribution and the BFGS method has been used by the authors for maximizing the total log-likelihood function log L. The MOR random number generation was performed using the inversion method. The true parameter values used in the data generating processes are $\gamma = 0.5, \sigma = 2$ and $\gamma = 3, \sigma = 5$. Table I lists the means of the MSEs of the four parameters that index the MOR distribution along with the respective biases for sample sizes $n = 50$, $n = 100$, $n = 150$, $n = 200$. Table I indicates that the MSEs of (γ,σ) decay as the sample size increases, as expected.

TABLE I
MSEs of (γ,σ)

	Real Parameters			
	$\gamma = 0.5, \sigma = 2$		$\gamma = 3, \sigma = 5$	
n	Estimated Parameters (Mean Squared Error)			
50	0,6383807	1,9695970	3,6452330	4,9614980
	(0,152058)	(0,132218)	(4,033918)	(0,266293)
10	0,5644994	1,9857230	3,2980440	4,9804960
0	(0,053041)	(0,067304)	(1,464633)	(0,132438)
15	0,5436154	1,9877790	3,1850690	4,9897920
0	(0,031194)	(0,043940)	(0,838079)	(0,087561)
20	0,5306474	1,9931910	3,1377940	4,9898750
0	(0,021708)	(0,033041)	(0,597327)	(0,066841)

V. Conclusion

In this study, we propose a two parameter lifetime distribution, so-called the Marshall-Olkin Rayleigh distribution which is an extension of the Rayleigh distribution. Our proposed model has increasing and decreasing hazard rate functions. We provide a mathematical treatment of this distribution including reliability and entropy. We examine a maximum likelihood estimation of the parameters. A simulation study is given to demonstrate that it can be used quite effectively to provide better fits than other available models. We hope that this generalization may attract wider application in statistics.

References

[1] L. Rayleigh, "On the stability or instability of certain fluid motions", in *Proc. of the London Mathematical Society,* vol. 11, 1880, pp. 57-70.
[2] V.G. Vodà, "Note on the truncated Rayleigh variate", *Revista Colombiana de Matemáticas,* vol. 9, 1975, pp. 1–7.
[3] J. Leao, H. Saulo, M. Bourguignon, R. Cintra, L. Rego, and G. Cordeiro, "On some properties of the beta inverse Rayleigh distribution", *Chilean Journal of Statistics,* vol. 4, 2013, pp. 111–131.
[4] A. Ahmad, S.P. Ahmad, and A. Ahmed, "Transmuted inverse Rayleigh distribution: A generalization of the inverse Rayleigh distribution", *Mathematical Theory and Modeling,* vol. 4, 2014, pp. 90-98.
[5] A.W. Marshall and I. Olkin, "A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families", *Biometrika*, vol. 84, 1997, pp. 641-652.
[6] A. Rényi, "On measures of entropy and information". in *Proc. of the 4th Berkeley Symposium*, 1961.
[7] C.E. Shannon, "Prediction and entropy of printed English". *Bell System Technical Journal*, vol. 30, 1951, pp. 50-64.

G. Ozel is an associate professor in the Department of Statistics at Hacettepe University. She received her Ph.D., Master of Science, and Bachelor of Science degrees in the Department of Statistics from Hacettepe University. Her research interests include stochastic processes, probability theory, survival analysis, and applied statistics.

S. Çakmakyapan was born in Ankara, Turkey in 1985. She is a Ph.D. student in the Department of Statistics at Hacettepe University. She received her B.Sc. and M.Sc. degrees from the Department of Statistics at Mugla University. Her research interests include stochastic processes, probability theory, Monte Carlo simulations, survival analysis, and applied statistics.