

Transmuted Rayleigh Distribution

Faton Merovci

Department of Mathematics, University of Prishtina, Kosovo

Abstract: In this article, we generalize the Rayleigh distribution using the quadratic rank transmutation map studied by Shaw et al. (2009) to develop a transmuted Rayleigh distribution. We provide a comprehensive description of the mathematical properties of the subject distribution along with its reliability behavior. The usefulness of the transmuted Rayleigh distribution for modeling data is illustrated using real data.

Zusammenfassung: Wir verallgemeinern die Rayleigh-Verteilung, indem wir die in Shaw et al. (2009) untersuchte quadratische Rangumwandlungsabbildung verwenden, um eine transmuted Rayleigh-Verteilung herzuleiten. Wir stellen eine umfassende Beschreibung der Eigenschaften dieser Verteilung bereit und besprechen ihr Verhalten in der Zuverlässigkeitstheorie. Der Nutzen der transmuted Rayleigh-Verteilung bei der Modellierung von wird anhand realer Daten illustriert.

Keywords: Rayleigh Distribution, Hazard Rate Function, Reliability Function, Parameter Estimation.

1 Introduction

In many applied sciences such as medicine, engineering and finance, amongst others, modeling and analyzing lifetime data is crucial. Several lifetime distributions have been used to model such kind of data. The quality of the procedures used in a statistical analysis depends heavily on the assumed probability model or distribution. Because of this, considerable effort has been expended in the development of large classes of standard probability distributions along with relevant statistical methodologies. However, there still remain many important problems where the real data does not follow any of the classical or standard probability models.

In this article we present a new generalization of the Rayleigh distribution called the transmuted Rayleigh distribution.

Definition 1 A random variable X is said to have a transmuted distribution if its cumulative distribution function(cdf) is given by

$$G(x) = (1 + \lambda)F(x) - \lambda F^2(x), \quad |\lambda| \leq 1,$$

where $F(x)$ is the cdf of the base distribution.

Observe that at $\lambda = 0$ we have the base distribution of the random variable X . Aryal et al. (2009) studied the transmuted Gumbel distribution and it has been observed that the transmuted Gumbel distribution can be used to model climate data. In the present study we will provide mathematical formulations of the transmuted Rayleigh distribution and also some of its properties.

2 Transmuted Rayleigh Distribution

Definition 2 The density function (pdf) of a Rayleigh distribution is

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

and the respective cdf is

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

The transmuted cdf is

$$G(x, \sigma, \lambda) = \left(1 - \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \left(1 + \lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right),$$

with transmuted pdf

$$g(x, \sigma, \lambda) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + 2\lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right).$$

Note that the transmuted Rayleigh distribution is an extended model to analyze more complex data. The Rayleigh distribution is clearly a special case for $\lambda = 0$. Figure 1 illustrates some of the possible shapes of the pdf of a transmuted Rayleigh distribution for selected values of the parameters λ and σ .

3 Moments

Theorem 1 The r th moment $E(X^r)$ of a transmuted Rayleigh distributed random variable X is given as

$$E(X^r) = \frac{1}{2} \sigma^r r \Gamma\left(\frac{r}{2}\right) (\lambda + 2^{\frac{r}{2}}(1 - \lambda)).$$

Especially we have

$$\begin{aligned} E(X) &= \frac{1}{2} \sigma \sqrt{\pi} (\lambda + \sqrt{2}(1 - \lambda)), \\ \text{var}(X) &= E(X^2) - E^2(X) = \sigma^2 \left(2 - \lambda - \frac{\pi}{4} (\lambda + \sqrt{2}(1 - \lambda))^2\right). \end{aligned}$$

Proof.

$$\begin{aligned} E(X^r) &= \int_0^\infty x^r f(x) dx \\ &= \int_0^\infty \frac{x^{r+1}}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + 2\lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) dx \\ &= \frac{(1 - \lambda)}{\sigma^2} \int_0^\infty x^{r+1} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx + \frac{2\lambda}{\sigma^2} \int_0^\infty x^{r+1} \exp\left(-\frac{x^2}{\sigma^2}\right) dx \\ &= 2^{\frac{r-2}{2}} (1 - \lambda) \sigma^r r \Gamma\left(\frac{r}{2}\right) + \frac{\lambda r}{2} \sigma^r \Gamma\left(\frac{r}{2}\right) \\ &= \frac{1}{2} \sigma^r r \Gamma\left(\frac{r}{2}\right) (\lambda + 2^{\frac{r}{2}}(1 - \lambda)). \end{aligned}$$

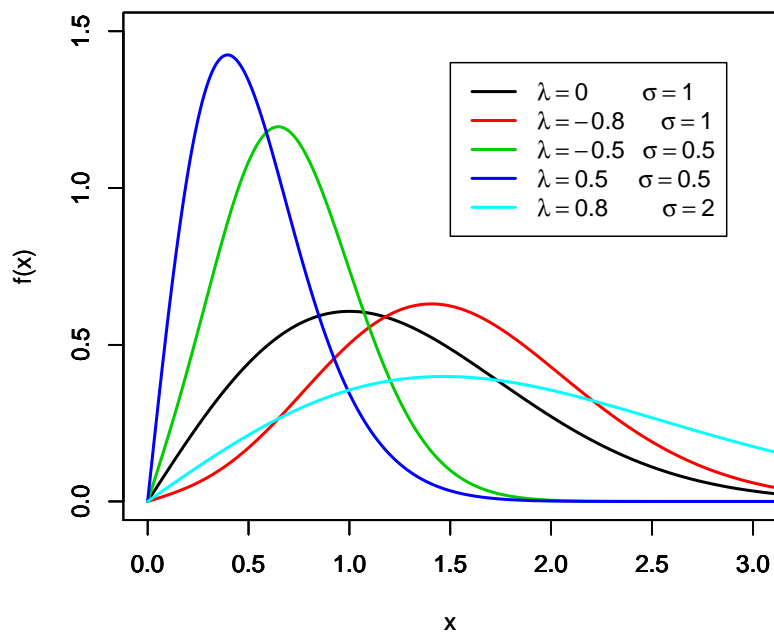


Figure 1: The pdf's of various transmutated Rayleigh distributions.

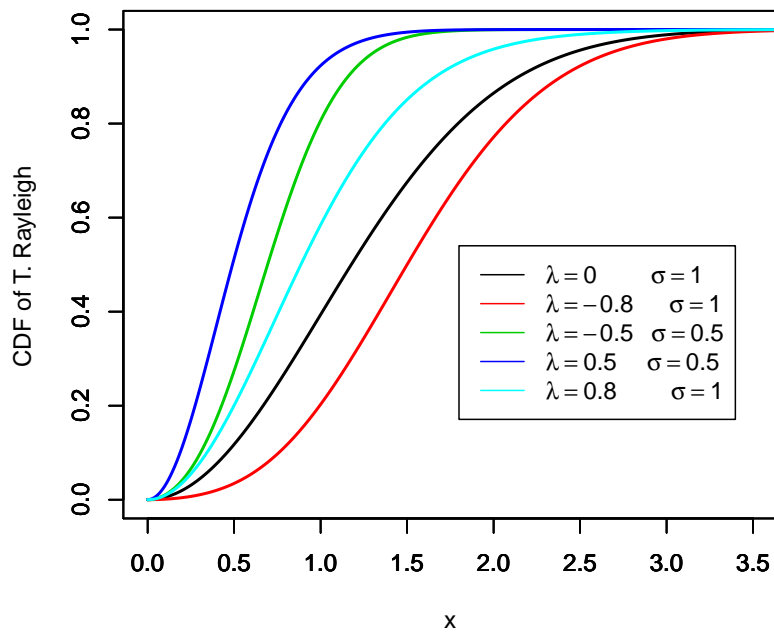


Figure 2: The cdf's of various transmutated Rayleigh distributions.

Here, we used

$$\int_0^{\infty} x^{\nu-1} \exp(-\mu x^p) dx = \frac{1}{p} \mu^{-\frac{\nu}{p}} \Gamma\left(\frac{\nu}{p}\right),$$

for $p, \nu, \mu > 0$ (see Gradshtein and Ryzhnik(2000), Sec. 3.478). \square

Theorem 2 Let X have a transmuted Rayleigh distribution. Then the moment generating function of X , say $M_X(t)$, is

$$M_X(t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} \frac{1}{2} \sigma^i i \Gamma\left(\frac{i}{2}\right) \left(\lambda + 2^{\frac{i}{2}}(1-\lambda)\right).$$

Proof.

$$\begin{aligned} M_X(t) &= \mathbf{E}(e^{tX}) = \int_0^{\infty} \exp(tx) f(x) dx \\ &= \int_0^{\infty} \left(1 + tx + \frac{t^2 x^2}{2!} + \cdots + \frac{t^n x^n}{n!} + \cdots\right) f(x) dx \\ &= \sum_{i=0}^{\infty} \frac{t^i \mathbf{E}(X^i)}{i!} \\ &= \sum_{i=0}^{\infty} \frac{t^i}{i!} \frac{1}{2} \sigma^i i \Gamma\left(\frac{i}{2}\right) \left(\lambda + 2^{\frac{i}{2}}(1-\lambda)\right). \end{aligned}$$

\square

4 Parameter Estimators

The maximum likelihood estimator (MLE) of the parameters that is inherent within the transmuted Rayleigh pdf is discussed next. The sample likelihood function under this model is

$$\begin{aligned} L(\sigma, \lambda; x) &= \prod_{i=1}^n g(x_i, \sigma, \lambda) \\ &= \frac{\prod_{i=1}^n x_i}{\sigma^{2n}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2\right) \prod_{i=1}^n \left(1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right), \quad (1) \end{aligned}$$

with respective sample log-likelihood function

$$\begin{aligned} \ell(\sigma, \lambda; x) &= \log L(\sigma, \lambda; x) \\ &= \sum_{i=1}^n \log(x_i) - n \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 + \sum_{i=1}^n \log\left(1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right). \end{aligned}$$

For ease of notation, we will denote the first partial derivatives of any function $f(x, y)$ by f_x and f_y , and its second partial derivatives by f_{xx} , f_{yy} , f_{xy} , and f_{yx} .

Now setting

$$\ell_\sigma = 0 \quad \text{and} \quad \ell_\lambda = 0,$$

we have

$$-\frac{2n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n x_i^2 + \frac{2\lambda}{\sigma^3} \sum_{i=1}^n \frac{x_i^2 \exp\left(-\frac{x_i^2}{2\sigma^2}\right)}{1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)} = 0, \quad (2)$$

and

$$\sum_{i=1}^n \frac{2 \exp\left(-\frac{x_i^2}{2\sigma^2}\right) - 1}{1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)} = 0. \quad (3)$$

The MLE $\hat{\theta} = (\hat{\sigma}, \hat{\lambda})$ of $\theta = (\sigma, \lambda)$ is obtained by solving this nonlinear system of equations. It is usually more convenient to use nonlinear optimization algorithms such as the quasi-Newton algorithm to numerically maximize the sample likelihood function given in (1). Applying the usual large sample approximation, the MLE $\hat{\theta}$ can be treated as being approximately bivariate normal with mean θ and variance-covariance matrix equal to the inverse of the expected information matrix, i.e.

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0, nI^{-1}(\theta)),$$

where $I^{-1}(\theta)$ is the limiting variance-covariance matrix of $\hat{\theta}$. The elements of the 2×2 matrix $I(\theta)$ can be estimated by $I_{ij}(\hat{\theta}) = -\ell_{\theta_i \theta_j} \big|_{\theta=\hat{\theta}}$, $i, j \in \{1, 2\}$.

From (2) and (3) the second partial derivatives of the log-likelihood function are found to be

$$\begin{aligned} \ell_{\sigma\sigma} &= \frac{2n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^n x_i^2 - \frac{6}{\sigma^4} \sum_{i=1}^n \frac{x_i^2 \exp\left(-\frac{x_i^2}{2\sigma^2}\right)}{1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)} \\ &\quad + \frac{2\lambda(1-\lambda)}{\sigma^6} \sum_{i=1}^n \frac{x_i^4 \exp\left(-\frac{x_i^2}{2\sigma^2}\right)}{\left(1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right)^2}, \\ \ell_{\sigma\lambda} &= \frac{2}{\sigma^3} \sum_{i=1}^n \frac{x_i^2 \exp\left(-\frac{x_i^2}{2\sigma^2}\right)}{\left(1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right)^2} \\ \ell_{\lambda\lambda} &= - \sum_{i=1}^n \left(\frac{2 \exp\left(-\frac{x_i^2}{2\sigma^2}\right) - 1}{1 - \lambda + 2\lambda \exp\left(-\frac{x_i^2}{2\sigma^2}\right)} \right)^2. \end{aligned}$$

Approximate two sided $100(1-\alpha)\%$ confidence intervals for σ and for λ are, respectively, given by

$$\hat{\sigma} \pm z_{\alpha/2} \sqrt{I_{11}^{-1}(\hat{\theta})} \quad \text{and} \quad \hat{\lambda} \pm z_{\alpha/2} \sqrt{I_{22}^{-1}(\hat{\theta})},$$

where z_α is the upper α th quantile of the standard normal distribution. Using R we can easily compute the Hessian matrix and its inverse and hence the standard errors and asymptotic confidence intervals.

We can compute the maximized unrestricted and restricted log-likelihood functions to construct the likelihood ratio (LR) test statistic for testing on some transmuted Rayleigh sub-models. For example, we can use the LR test statistic to check whether the a transmuted Rayleigh distribution for a given data set is statistically *superior* to the Rayleigh distribution. In any case, hypothesis tests of the type $H_0 : \theta = \theta_0$ versus $H_0 : \theta \neq \theta_0$ can be performed using a LR test. In this case, the LR test statistic for testing H_0 versus H_1 is $\omega = 2(\ell(\hat{\theta}; x) - \ell(\hat{\theta}_0; x))$, where $\hat{\theta}$ and $\hat{\theta}_0$ are the MLEs under H_1 and H_0 , respectively. The statistic ω is asymptotically (as $n \rightarrow \infty$) distributed as χ_k^2 , where k is the length of the parameter vector θ of interest. The LR test rejects H_0 if $\omega > \chi_{k;\gamma}^2$, where $\chi_{k;\gamma}^2$ denotes the upper $100\gamma\%$ quantile of the χ_k^2 distribution.

5 Reliability Analysis

The reliability function $R(t)$, which is the probability of an item not failing prior to some time t , is defined by $R(t) = 1 - F(t)$. The reliability function of a transmuted Rayleigh distribution is given by

$$R(t, \sigma, \lambda) = \exp\left(-\frac{t^2}{2\sigma^2}\right) \left(1 - \lambda + \lambda \exp\left(-\frac{t^2}{2\sigma^2}\right)\right).$$

The other characteristic of interest of a random variable is the hazard rate function defined by

$$h(t) = \frac{f(t)}{1 - F(t)},$$

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 . The hazard rate function for a transmuted Rayleigh random variable is given by

$$h(t, \sigma, \lambda) = \frac{t \left(1 - \lambda + 2\lambda \exp\left(-\frac{t^2}{2\sigma^2}\right)\right)}{\sigma^2 \left(1 - \lambda + \lambda \exp\left(-\frac{t^2}{2\sigma^2}\right)\right)}.$$

Figure 3 illustrates the reliability function of a transmuted Rayleigh distribution for different values of the parameters σ and λ .

6 Order Statistics

The k th order statistic of a sample is its k th smallest value. For a sample of size n , the n th order statistic (or largest order statistic) is the maximum, that is,

$$X_{(n)} = \max\{X_1, \dots, X_n\}.$$

The sample range is the difference between the maximum and minimum. It is clearly a function of the order statistics:

$$\text{range}\{X_1, \dots, X_n\} = X_{(n)} - X_{(1)}.$$

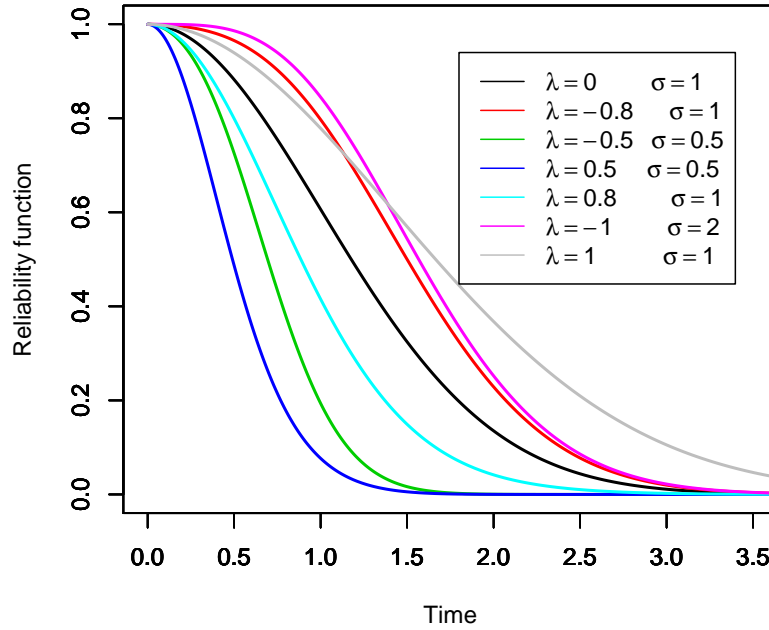


Figure 3: The reliability function of a transmuted Rayleigh distribution.

We know that if $X_{(1)} \leq \dots \leq X_{(n)}$ denotes the order statistic of a random sample X_1, \dots, X_n from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$ then the pdf of $X_{(j)}$ is given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) (F_X(x))^{j-1} (1 - F_X(x))^{n-j} ,$$

for $j = 1, \dots, n$. The pdf of the j th order statistic for a transmuted Rayleigh distribution is given by

$$\begin{aligned} f_{X_{(j)}}(x) &= \frac{n!}{(j-1)!(n-j)!} \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + 2\lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \\ &\cdot \left(1 - \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \left(1 + \lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right)^{j-1} \\ &\cdot \left(\exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + \lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right)^{n-j} . \end{aligned}$$

Therefore, the pdf of the largest order statistic $X_{(n)}$ is

$$\begin{aligned} f_{X_{(n)}}(x) &= \frac{nx}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + 2\lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \\ &\cdot \left(1 - \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \left(1 + \lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right)^{n-1} , \end{aligned}$$

and the pdf of the smallest order statistic $X_{(1)}$ is

$$f_{X_{(1)}}(x) = \frac{nx}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + 2\lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \cdot \left(\exp\left(-\frac{x^2}{2\sigma^2}\right) \left(1 - \lambda + \lambda \exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\right)^{n-1}.$$

7 Application

Now we use a real data set to show that the transmuted Rayleigh distribution can be a better model than the Rayleigh distribution.

We work with nicotine measurements made from several brands of cigarettes in 1998. The data have been collected by the Federal Trade Commission which is an independent agency of the US government, whose main mission is the promotion of consumer protection.

The report entitled tar, nicotine, and carbon monoxide of the smoke of 1206 varieties of domestic cigarettes for the year of 1998 available at <http://www.ftc.gov/reports/tobacco> and consists of the data sets and some information about the source of the data, smoker's behaviour and beliefs about nicotine, tar and carbon monoxide contents in cigarettes. The free form data set can be found at <http://pw1.netcom.com/rdavis2/smoke.html>.

The site <http://home.att.net/rdavis2/cigra.html> contains $n = 384$ observations. We analysed data on nicotine, measured in milligrams per cigarette, from several cigarette brands. Some summary statistics for the nicotine data are as follows: mean = 0.852, median = 0.9, minimum = 0.1 and maximum = 2.

The variance covariance matrix of the MLEs under the transmuted Rayleigh distribution is computed as

$$I(\hat{\theta})^{-1} = \begin{pmatrix} 0.183 \times 10^{-3} & 0.470 \times 10^{-3} \\ 0.470 \times 10^{-3} & 0.529 \times 10^{-2} \end{pmatrix}.$$

Thus, the variances of the MLE of σ and λ is $\text{var}(\hat{\sigma}) = 0.183 \times 10^{-3}$ and $\text{var}(\hat{\lambda}) = 0.529 \times 10^{-2}$. Therefore, 95% confidence intervals for σ and λ are $[0.528, 0.582]$ and $[-0.914, -0.629]$.

Table 1: Estimated parameters of the Rayleigh and Transmuted Rayleigh distribution for the nicotine measurements data.

Model	Parameter Estimate	Standard Error	$-\ell(\cdot; x)$
Transmuted	$\hat{\sigma} = 0.5555$	0.0135	121.224
Rayleigh	$\hat{\lambda} = -0.7718$	0.0728	
Rayleigh	$\hat{\sigma} = 0.6475$	0.0175	142.3572

The LR test statistic to test the hypotheses $H_0 : \lambda = 0$ versus $H_1 : \lambda \neq 0$ is $\omega = 42.2664 > 3.841 = \chi_{1,0.05}^2$, so we reject the null hypothesis.

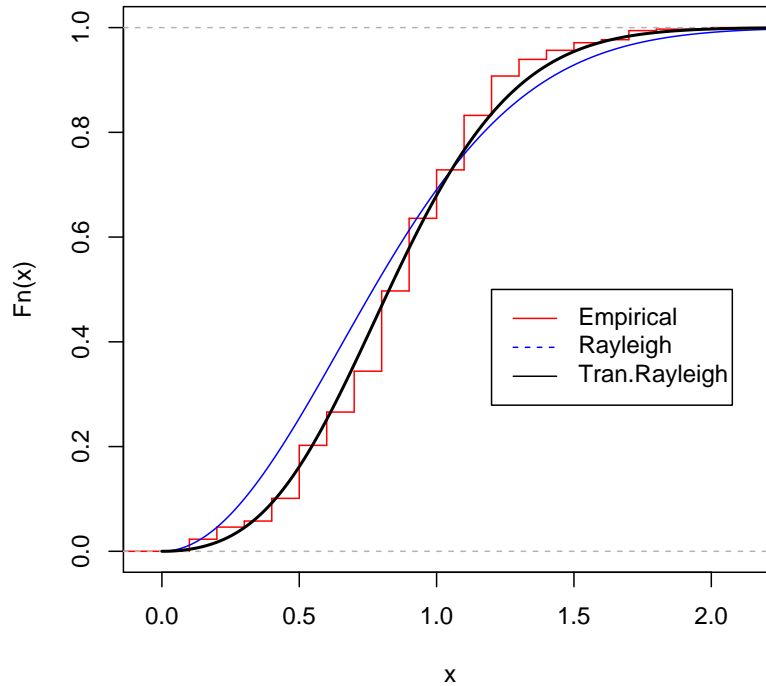


Figure 4: Empirical, fitted Rayleigh, and transmuted Rayleigh cdf of the nicotine measurements data.

Table 2: Criteria for comparison.

Model	KS	-2ℓ	AIC	AICC	BIC
Rayleigh	0.184	284.714	286.714	285.714	296.407
Transmuted Rayleigh	0.124	242.448	246.448	243.445	265.833

In order to compare the two distribution models, we consider criteria like KS (Kolmogorow Smirnow), -2ℓ , AIC (Akaike information criterion), AICC (corrected Akaike information criterion), and BIC (Bayesian information criterion) for the data set. The better distribution corresponds to smaller KS, -2ℓ , AIC, AICC, and BIC values:

$$\text{AIC} = 2k - 2\ell, \quad \text{AICC} = \text{AIC} + \frac{2k(k+1)}{n-k-1},$$

and

$$\text{BIC} = k \log(n) - 2\ell,$$

where k is the number of parameters in the statistical model, n the sample size and ℓ is the maximized value of the log-likelihood function under the considered model. Also, here for calculating the values of KS we use the sample estimates of λ and σ . Table 1 shows the MLEs under both distributions, Table 2 shows the values of KS, -2ℓ , AIC, AICC, and BIC values. The values in Table 2 indicate that the transmuted Rayleigh distribution leads to a better fit than the Rayleigh distribution.

8 Conclusion

Here we propose a new model, the so-called the transmuted Rayleigh distribution which extends the Rayleigh distribution in the analysis of data with real support. An obvious reason for generalizing a standard distribution is because the generalized form provides larger flexibility in modeling real data. We derive expansions for the mean, variance, moments and for the moment generating function. The estimation of parameters is approached by the method of maximum likelihood, also the information matrix is derived. We consider the likelihood ratio statistic to compare the model with its baseline model. An application of the transmuted Rayleigh distribution to real data show that the new distribution can be used quite effectively to provide better fits than the Rayleigh distribution.

Acknowledgements

The author would like to thank the Editor Herwig Friedl and the referee for carefully reading the manuscript and for their comments which greatly improved the presentation.

References

- Aryal, G. R., and Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods and Applications*, 71, 1401-1407.
- Aryal, G. R., and Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 4, 89-102.
- Gradshteyn, I. S., and Ryzhik, I. M. (2000). *Table of Integrals, Series, and Products*. San Diego: Academic Press.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. New York: Wiley.
- Khan, M. S., and King, R. (2013). Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 6, 66-88.
- Kundu, D., and Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimation. *Computational Statistics and Data Analysis*, 49, 187-200.
- Lee, E. T., and Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*. New York: Wiley.
- Mahdi, S. (2006). Improved parameter estimation in Rayleigh model. *Metodoloski zvezki*, 3, 63-74.
- Manesh, S. F., and Khaledi, B. (2008). On the likelihood ratio order for convolutions of independent generalized Rayleigh random variables. *Statistics & Probability Letters*, 78, 3139-3144.
- Merovci, F. (2013). Transmuted Lindley distribution. *International Journal of Open Problems in Computer Science and Mathematics*, 6, (to appear).

- Shaw, W. T., and Buckley, I. R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint*, arXiv:0901.0434.
- Van Der Vart, A. W. (1998). *Asymptotic Statistics*. New York: Cambridge University Press.
- Vodă, V. G. (1972). On the inverse Rayleigh distributed random variable. *Reports of Statistical Application Research*, 19, 13-21.
- Vodă, V. G. (1975). Note on the truncated Rayleigh variate. *Revista Colombiana de Matematicas*, 9, 1-7.
- Vodă, V. G. (1976). Inferential procedures on a generalized Rayleigh variate I. *Applied Mathematics*, 21, 395-412.

Authors address:

Faton Merovci
Department of Mathematics
University of Prishtina
Prishtinë, 10000
Kosovo

E-mail: fmerovci@yahoo.com