International Journal of Electronics Communication and Computer Engineering
Volume 8, Issue 3, ISSN (Online): 2249–071X

Bivariate Rayleigh Distribution with Application

M. K. ABD ELAAL and L.A. Baharith*
*Corresponding author email id: lbaharith@kau.edu.sa

Date of publication (dd/mm/yyyy): 09/05/2017

***Abstract* — The Rayleigh distribution has received special attention from many researchers due to its close relations with important distributions which makes it useful in analyzing lifetime applications. In this study, a new bivariate Rayleigh distribution based on the Gaussian copula and mixture is suggested. The bivariate Rayleigh distribution is constructed with univariate Rayleigh distribution as marginals. Bayesian estimation is used to estimate the unknown parameters of the proposed distribution. Simulation study is conducted to study the effectiveness of the proposed bivariate Rayleigh distribution. One real data application is analyzed for illustrative purpose.**

***Keywords* — Bivariate Rayleigh Distribution, M Mixture Representation, Gaussian Copula, Bayesian Estimation.**

I. Introduction

The Rayleigh distribution was first introduced by Lord Rayleigh [17] as a problem in the field of acoustics. It has relations with important distributions such as the Weibull, Chi-square, and extreme value distributions. The decreasing hazard function of time is another important characteristics of Rayleigh distribution in life time applications. Therefore, it becomes more useful in physics studies such as signal processing and the study of various types of radiation. Moreover, one of its applications is a model of wind speed variation. This type of analysis is used for estimating the energy recovery from a wind turbine. For more details see [1], [7], [8], [13].

The probability density function (Pdf) and the cumulative function of Rayleigh distribution are, respectively, given by

$$f(\mathrm{t}) = 2\alpha t exp(-\alpha t^2), \quad (1)$$

$$F(\mathrm{t}) = 1 - exp(-\alpha t^2), \quad (2)$$

where α>0 is the scale parameter. Interestingly, although much work has been done on the univariate Rayleigh distribution, not much work has been done to the bivariate Rayleigh distribution. Various methods of constructing multivariate or bivariate distributions have been studied in the literatures. For a complete review of these methods, see for example, [5], [14], [16], [18].

Adham and Walker [3] combined the mixture and copula ideas to introduce the bivariate Gompertz distribution which could be extended naturally to the multivariate distribution. They concluded that the resulting bivariate distribution is easy to analyze and has a full dependence structures. AL-Dayian et al. [4] have adopted this method to propose the bivariate half- logistic distribution. In addition, the bivariate half- logistic-type distributions is introduced by Adham et al. [2]. Recently, Elaal et al. [9] introduced the bivariate Burr X distribution.

According to [3], [18], the mixture representation for a Pdf of a random variable T on [0,∞) can be written in the following form:

$$f(t) = \int_\Omega f(t|u) f(u)\, du, for all u \in \Omega, \quad (3)$$

where u is non-negative latent variable that follows a gamma distribution with shape parameter 2 and scale parameter 1. That is

$$f(u) = u exp(-u) \quad (4)$$

Then, the mixture representation for any lifetime distribution can be written as

$$f(t|u) = \frac{h(t)}{u}, u > H(t), \quad (5)$$

where $h(\mathrm{t})$ and $H(\mathrm{t})$ are the hazard rate function and the cumulative hazard rate function of T, respectively.

The paper is outlined as follows: Section 2 presents the construction and estimation for the unknown parameters of the univariate Rayleigh distribution. Section 3 presents the construction of the proposed bivariate Rayleigh distribution. Bayesian estimation of the parameters of the bivariate Rayleigh is discussed in Section 4. Simulation study is carried out in Section 5 to illustrate the performance of the proposed bivariate Rayleigh distribution. Real dataset application is analyzed to illustrate the results in Section 6. Concluding remarks appear in Section 7.

II. The Univariate Rayleigh Distribution

The hazard rate function (HRF) and the cumulative hazard rate function (CHRF) of a continues random variable T that follows a Rayleigh distribution are given, respectively, by

$h(\mathrm{t}) = 2\alpha\mathrm{t}$,(6)

and

$$H(\mathrm{t}) = \alpha t^2, \quad (7)$$

The mixture representation for the Rayleigh is obtained by substituting (4) and (5) in (3) as follows

$$f(\mathrm{t}|u) = \frac{2\alpha\mathrm{t}}{u}, \mathrm{u} > ln(\alpha t^2) \quad (8)$$

Therefore, the Pdf of Rayleigh can be written as

$$f(\mathrm{t}) = 2\alpha\mathrm{t}\, exp(-u),\ \mathrm{u} > \alpha t^2 \quad (9)$$

A. Estimation of the Parameters of Univariate Rayleigh Distribution

The Bayesian method is used to obtain estimates of the Rayleigh parameter via Gibbs sampler algorithm. The Gibbs sampler will be used to obtain random variables from posterior distributions of the Rayleigh distribution. See [3], [11], [12].

Suppose that $T = t_1, \dots, t_n$ is a random sample from Rayleigh distribution, and $U = u_1, \dots, u_n$ is a random sample from gamma(2,1) distribution given in (4). Then, the likelihood function can be written as

$$L(\alpha|T,U) = (2\alpha)^n exp(\sum_{i=1}^{n} ln(t_i)) exp(-\sum_{i-1}^{n} u_i) \quad (10)$$

A non-informative priors for the parameter α is given by

$$\pi(\alpha) = \frac{1}{\alpha}\ . \quad (11)$$

Then, the posterior distribution is

$$f(\alpha|T,U) = \alpha^{n-1}exp(\sum_{i=1}^{n} ln(t_i))exp(-\sum_{i-1}^{n} u_i) \quad (12)$$

where $\alpha < min\left(\frac{u_i}{t_i^2}\right)$. It is clear that the conditional distribution of the parameters α is not in closed form. In order to apply Monte Carlo integration, the full conditional distribution is sampled as follows:

First: we sample the marginal posterior of u_i, i=1,…,n which is given by

$$f(u_i|\alpha,T) = exp(-u_i), \quad u > \alpha t_i^2. \quad (13)$$

This is exponential restricted to $(\alpha t^2, \infty)$.

Second: we sample the marginal posterior of α, which is given by

$$f(\alpha|T,U) = \alpha^{n-1}exp(\sum_{i=1}^{n} ln(t_i)), \alpha < (u_i/t_i^2) \quad (14)$$

It can be seen that the above marginal posteriors of α is not in closed form. Therefore, Algorithm 1in Appendix is applied to sample the posterior distribution of the parameter α.

III. Bivariate Rayleigh Distribution

The idea of constructing univariate Rayleigh distribution illustrated in section 2 could be extended to bivariate Rayleigh based on copula and M mixture representation, where M denotes the set of densities for a random variable T on [0,∞). This can be obtained by considering bivariate gamma distribution of the latent variables $\underline{U} = (U_1, U_2)$,with two marginal gamma distribution given in (4) using copula with unknown correlation parameter ρ. Then obtaining the joint bivariate distribution of $\underline{U}$ and $\underline{T} = (T_1, T_2)$ and assuming that T_1, T_2 are conditionally independent given U_1, U_2.

Then, the joint Pdf of the bivariate distribution based on M mixture representation with Gaussian copula can be written as

$$f(T_1,T_2) = \int_{G_1}^{\infty}\int_{G_2}^{\infty}\prod_{j=1}^{2} f(t_j|u_j)f(u_j)\, C_G'\, du_1 du_2,$$

where $G_j = H(t_j)$, $f(t_j|u_j)$is given by (5), $f(u_j)$ is given by (4), and C_G' is the Pdf of Gaussian copula given by

$$C_G^{`} = \frac{\exp\left\{\frac{-1}{2(1-\rho^2)}(y_1^2 - 2\rho y_1 y_2 + y_2^2)\right\}}{2\pi\sqrt{1-\rho^2}}, \quad (15)$$

where y_1, y_2 have standard normal distribution.

Therefore, the joint pdf of bivariate Rayleigh distribution based on mixture with Gaussian copula can be written as

$$f(T_1,T_2) = \int_{G_1}^{\infty}\int_{G_2}^{\infty} 4\alpha_1 t_1 \alpha_2 t_2 \exp(u_1)\exp(u_2) C_G'\, du_1 du_2. \quad (16)$$

IV. Estimation of the Parameters of Bivariate Rayleigh Distribution

Bayesian method is used to estimate the vector of parameters of the bivariate Rayleigh, $(\alpha_1, \alpha_2, \rho)$. The Gibbs sampler procedure is used to obtain random variables from the posterior distribution associated with the bivariate Rayleigh distribution. Then assuming that the prior distribution of the parameter α_j as

$$\pi(\alpha_j) \propto \frac{1}{\alpha_j}, j = 1,2. \quad (17)$$

In addition, assuming that the correlation parameter ρ has a uniform prior distribution defined on the interval (-1, 1).

Then, the joint posterior distribution of the parameters given a random sample of size n of the bivariate Rayleigh is given by

$$f\left(\underline{\alpha},\rho,\underline{U}\middle|\underline{T}\right) = \prod_{j=1}^{2}\pi(\alpha_j)\pi(\rho)\prod_{i=1}^{n} f\left(t_{ji}\middle|u_{ji}\right)f(u_{ji})\, C_G'$$

where $u_{ji} > \alpha_j t_j^2$, j=1,2. Then, we sample the following:

First: Sample the marginal posterior of u_{ji} from $f\left(u_{ji}|\alpha, u_{-ji}, \underline{T}\right) \propto exp\left\{-u_{ji} - \frac{x_i}{2}\right\}$

where

$$x_i = \frac{y^2{}_{ji} - 2\rho y_{ji} y_{-ji} + y^2{}_{-ji}}{1-\rho^2}.$$

Algorithm 2 in Appendix is applied to sample the above conditional distribution.

Second: Sample α_j from

$$f\left(\alpha_j\middle|\rho,\underline{T},\underline{U}\right) = \alpha_j^{n-1}exp\left(\sum_{i=1}^{n} ln(t_{ji})\right), \alpha_j < \left(u_{ji}/t_{ji}^2\right).$$

Algorithm 1 in Appendix is used to sample this conditional distribution.

Finally, sample ρ from its posterior distribution

$$f\left(\rho|\underline{\alpha},\underline{T},\underline{U}\right) \propto (1-\rho^2)^{\frac{-n}{2}} exp\left\{\sum_{i=1}^{n}\frac{y^2{}_{1i} - 2\rho y_{1i} y_{2i} + y^2{}_{2i}}{-2(1-\rho^2)}\right\}$$

This full conditional distribution can be sampled using metropolis Hasting Algorithm.

V. Simulation Study

In this Section, simulation study is carried out to examine the performance of the Bayesian estimation for different sample sizes and different parameter values for the constructed bivariate Rayleigh distribution. The performances of the Bayesian estimates are studied mainly with respect to their mean squared error (MSE) over 10000 iterations. A random samples of sizes (n=15,30,50) observations are generated from the bivariate Rayleigh distribution with marginal distributions Rayleigh($\alpha_1 = 0.9$) and Rayleigh($\alpha_2 = 1.8$) with ρ=0.7. In addition, random samples of sizes (n=15, 30, 50) observations are also generated from the bivariate Rayleigh distribution with marginal distributions Rayleigh($\alpha_1 = 3$) and Rayleigh($\alpha_2 = 4$) with ρ=0.7. The Gibbs sampler is used to estimate the parameters of the bivariate Rayleigh distribution in which the Markov chain ran for 10000 times.

The results are illustrated in Tables I and II for different sample sizes (n=15, 30, 50) and given values of the parameter.

It can be seen from Tables I and II that parameter estimates improved with increases in sample size. In addition, for all selected values of α_1, α_2, and ρ the MSE of the estimates $\hat{\alpha}_1$,$\hat{\alpha}_2$, and $\hat{\rho}$become smaller as the sample size increases. For α_1 smaller than one, the behavior of the estimates α_1 and α_2 get better based on MSE.

Table I: Mean estimates and the corresponding Bias and MSE of the Bayesian estimators for $\alpha_1 = 0.9, \alpha_2 = 1.8, \rho = 0.7$.

n	Parameter estimate		Bias	MSE
15	$\hat{\alpha}_1$	0.9046	0.0046	0.0045
	$\widehat{\alpha_2}$	1.8099	0.0099	0.0179
	$\hat{\rho}$	0.6129	0.0871	0.0240
30	$\hat{\alpha}_1$	0.9013	0.0013	0.0010
	$\widehat{\alpha_2}$	1.8026	0.0026	0.0041
	$\hat{\rho}$	0.6579	0.0421	0.0094
50	$\hat{\alpha}_1$	0.9011	0.0011	0.0004
	$\widehat{\alpha_2}$	1.8023	0.0023	0.0016
	$\hat{\rho}$	0.6763	0.0236	0.0046

Table II: Mean estimates and the corresponding Bias and MSE of the Bayesian estimators for $\alpha_1 = 3, \alpha_2 = 4, \rho = 0.7$.

n	Parameter estimate		Bias	MSE
15	$\hat{\alpha}_1$	3.0156	0.0156	0.0410
	$\hat{\alpha}_2$	4.0221	0.0221	0.0886
	$\hat{\rho}$	0.6129	0.0871	0.0240
30	$\hat{\alpha}_1$	3.0045	0.0045	0.0114
	$\hat{\alpha}_2$	4.0059	0.0059	0.0202
	$\hat{\rho}$	0.6579	0.0421	0.0094
50	$\hat{\alpha}_1$	3.0038	0.0038	0.0044
	$\hat{\alpha}_2$	4.0051	0.0051	0.0078
	$\hat{\rho}$	0.6763	0.0237	0.0046

The predicted observations can be obtained by applying the inverse of the conditional distribution function of the Rayleigh distribution as

$$t_{jk} = \sqrt{\frac{u_{jk}v_{jk}}{\alpha_{jk}}}, \; j = 1,2,$$

where v~uniform(0,1), $\alpha_{jk,}is$ the sampled values of α_j at iteration k of the MCMC run and u_{jk} is the jth element of a new bivariate vector of the observations generated from the bivariate gamma mixing distribution with correlation parameter sampled at the kth iteration of the MCMC.

VI. Data Analysis

The American football league data obtained from the matches played in 1986 with two variables where;

T_1: the game time to the first points scored by kicking the ball between goalposts

T_2 : the game time to the first points scored by moving the ball into the end zone, See [6].

These times are important to know how long one has to wait to watch a touchdown or to an observer who is interested only at the beginning stages of a game. T_1 and T_2 are positively correlated. Therefore, we use bivariate Rayleigh distribution to fit this data

As mentioned by [15] that the marginals are tested using Kolmogorov-Smirnov (K-S) separately. The K-S test values and the associated p values (reported in brackets) for T_1 and T_2 are 0.1419 (0.3666) and 0.1525 (0.2557), which indicate that the bivariate Rayleigh distribution provides appropriate fit for this bivariate data.

In addition, goodness of fit test is conducted to examine the appropriateness of Gaussian copula for this dataset based on *Sn* statistics which showed that it is suitable with (p-value > 0.05). The copula parameter is estimated through inversion of Kendall's tau to be 0.88 which can be used as initial value when fitting bivariate Rayleigh distribution. For details, see [10].

Bayesian estimates and associated credible intervals of the bivariate Rayleigh parameters are reported in Table III.

Table III: Bayesian estimates and the corresponding95% credible interval of the bivariate Rayleigh parameters.

Parameter estimate		95% Credible Interval	
		2.5%	97.5%
$\hat{\alpha}_1$	5.8499	6.3849	6.0015
$\hat{\alpha}_2$	7.8001	8.5132	8.0021
$\hat{\rho}$	0.6858	0.6208	0.7431

VII. Concluding Remarks

In this paper, we proposed a new bivariate Rayleigh distribution based on the Gaussian copula with univariate Rayleigh distribution as marginals. Bayesian estimation is performed to estimate the unknown parameters of the proposed distribution. Simulation study and real data application showed the flexibility of the proposed bivariate Rayleigh distribution. It is also provides a suitable fit for the data.

Appendix

***Algorithm* 1**

Note that for the univariate case, j=1.

Given a random sample of $T_{ji} = (T_{1i}, T_{2i})$, i = 1,…,n from Rayleigh distribution, and a random sample of $U_{ji} = (U_{1i}, U_{2i})$, i = 1,…,n from gamma(2,1) distribution.

1. Introduce a non-negative latent variablev_j. The joint pdf of v_j and α_j is given by

$$f(\alpha_j, v) \propto \alpha^{n-2}, \alpha_j > A_j, v < \frac{d_j}{\alpha_j},$$

where

$$A_j = \frac{u_{ji}}{t_{ji}^2}, \qquad d_j = exp\left(\sum_{i=1}^{n} \ln(t_{ji}\right)$$

2. Give a value of the parameter α, v is sampled from $\text{Uniform}(0, \alpha_j d_j)$.

3. Finally, use the distribution function inverse method to sample α_j

$$f(\alpha_j|v_j) \propto \alpha_j^{n-2}, (B_{ji} < \alpha_j < A_{ji}), \quad B_j = \frac{v_{ji}}{d_j}.$$

Thus,

$$\alpha_j = \left(\left[A_j^{n-1} - B_j^{n-1}\right]\delta + B_j^{n-1}\right)^{\frac{1}{n-1}},$$

where δ is Uniform$(0\,,1)$.

Algorithm 2

1. Introduce a non-negative latent variableτ, such that

$$f\left(u_{ji}\,,\tau\right) \propto e^{-u_{ji}}\, I\left(\tau < exp\left(\frac{-x_i}{2}\right)\right), u_{ji} > \alpha_j t_{ji}.$$

2. Choose the initial values of u_{ji} to be

$$u_{ji} = \ln\left(\alpha_j t_{ji}\right) + 1\,, j = 1\,,2.$$

3. Sample τ from Uniform $\left(0\,, \exp\left(\frac{-x_i}{2}\right)\right)$.

4. Sample $u_{ji}$from $f\left(u_{ji}|\tau\right)$

$$f\left(u_{ji}|\tau\right) \propto e^{-u_{ji}}\,, \left(A_i < u_{ji} < B_i\right),$$

where

$$A_i = \max\left[\alpha_j t_{ji}\,, F_{u_j}^{-1}\{\Phi\,(\delta_{1i})\}\right],$$

$$B_i = F_{u_j}^{-1}\{\Phi\,(\delta_{2i})\},$$

$$\delta_{1i} = \frac{y_{-ji}}{\rho} - q_i\,, \delta_{2i} = \frac{y_{-ji}}{\rho} + q$$

$$q_i = \sqrt{-2\,(1 - \rho^2)\left\{\ln(\tau) + \frac{y_{-ji}^2}{2\rho^4}\right\}}$$

$f\left(u_{ji}|\tau\right)$ is a double truncated distribution which easily sampled using inverse transform sampling method. Then, for $\nu \sim$ Uniform $(0\,,1)$

$$u_{ji} = -\ln\left(e^{-B_i} + \nu\left[e^{-A_i} - e^{-B_i}\right]\right), \text{see[9]}$$

References

[1] A. M. Abd Elfattah, A. S. Hassan, and D. M. Ziedan, “Efficiency of maximum likelihood estimators under different censored sampling schemes for Rayleigh distribution,” Interstat, 2006.

[2] M. K. Adham, S. A., AL-Dayian, G. R., El Beltagy, S. H., and Abd Elaal, “Bivariate half- logistic-type distribution,” Acad. Bus. Journal, AL-Azhar Univ., vol. 2, 2009,pp. 92–107.

[3] S. A. Adham and S. G. Walker, “A multivariate Gompertz-type distribution,” J. Appl. Stat., vol. 28, no. 8, 2001, pp. 1051–1065.

[4] G. R. AL- Dayian, S. A. Adham, S. H. El Beltagy, and A. Elaal, “Bivariate Half-Logistic Distributions Based on Mixtures and Copula,” Acad. Bus. J., vol. 2, 2008,pp. 92–107.

[5] E. K. Al-Hussaini and S. F. Ateya, “Bayes estimations under a mixture of truncated type I generalized logistic components model,” J. Stat. Theory Appl, vol. 4, no. 2, 2005,pp. 183–208.

[6] S. Csörgő and A. H. Welsh, “Testing for exponential and Marshall--Olkin distributions,” J. Stat. Plan. Inference, vol. 23, no. 3, 1989,pp. 287–300.

[7] S. Dey, “Comparison of Bayes estimators of the parameter and reliability function for Rayleigh distribution under different loss functions,” Malaysian J. Math. Sci., vol. 3, 2009,pp. 247–264.

[8] S. Dey, T. Dey, and D. Kundu, “Two-parameter Rayleigh distribution: different methods of estimation,” Am. J. Math. Manag. Sci., vol. 33, no. 1, 2014,pp. 55–74.

[9] M. K. A. Elaal, M. R. Mahmoud, M. M. EL-Gohary, and L. A. Baharith, “Univariate And Bivariate Burr Type X Distributions Based On Mixtures And Copula,” Int. J. Math. Stat., vol. 17, no. 1, 2016, pp. 113–127.

[10] C. Genest, B. Rémillard, and D. Beaudoin, “Goodness-of-fit tests for copulas: A review and a power study,” Insur. Math. Econ., vol. 44, no. 2, 2009,pp. 199–213.

[11] W. R. Gilks and P. Wild, “Adaptive rejection sampling for Gibbs sampling,” Appl. Stat., 1992,pp. 337–348.

[12] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, “Markov Chain Monte Carlo in Practice Chapman and HallCRC.” London, 1996.

[13] N. L. Johnson, S. Kotz, and N. Balakrishnan, “Continuous univariate distributions, vol. 2 of wiley series in probability and mathematical statistics: applied probability and statistics.” Wiley, New York, 1995.

[14] N. L. Johnson, S. Kotz, and N. Balakrishnan, Continuous Multivariate Distributions, Volume 1, Models and Applications, vol. 59. New York: John Wiley & Sons, 2002.

[15] D. Kundu and R. D. Gupta, “Absolute continuous bivariate generalized exponential distribution,” AStA Adv. Stat. Anal., vol. 95, no. 2, 2011,pp. 169–185.

[16] 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, no. 3, 1997,pp. 641–652.

[17] Lord Rayleigh, “Xii. on the resultant of a large number of vibrations of the same pitch and of arbitrary phase,” London, Edinburgh, Dublin Philos. Mag. J. Sci., vol. 10, no. 60, 1880,pp. 73–78.

[18] S. G. Walker and D. A. Stephens, “Miscellanea. A multivariate family of distributions on (0,∞) p,” Biometrika, vol. 86, no. 3, 1999,pp. 703–709.

Authors' Profiles

MERVAT K. ABD ELAAL[1,2],

[1]Assistant professor of statistics in Department of Statistics, Faculty of Science, Main Campus, King Abdulaziz University (KAU), Jeddah, Saudi Arabia.

[2] Department of Statistics Faculty of Commerce Al-Azhar University Girls Branch Cairo, Egypt

She has been worked in KAU since 2007. She taught many graduate and undergraduate courses in Statistics.

Dr. ABD ELAALis a member of graduate committee in Department of Statistics.

Email : mkabdelaal@kau.edu.sa

Lamya A. Baharith

Assistant professor of statistics in Department of Statistics, Faculty of Science, Main Campus, King Abdulaziz University (KAU), Jeddah, Saudi Arabia.

She has been worked in KAU since 1999. She taught many graduate and undergraduate courses in Statistics. Additionally, she presented workshops in Analyzing datausing different statistical packages.

Currently, she is the Vice Dean of Deanship of e-Learning and Distance Education at KAU. She was a head of statistics department during 2006-2010. She attended international conferences and published many articles.

Dr. Baharith is a member of several academic committees in KAU.

Email: lbaharith@kau.edu.sa