Proposed Methods for Estimating Parameters of the Generalized Raylieh Distribution in the Presence of One Outlier

Dhwyia S. Hassan¹, Faten F. Albadri², Nathier A. Ibrahim³, Hani Aziz Ameen⁴*

¹Head Department of Industrial Management, College of Administration & Economic, Baghdad University, Iraq
²Department of Statistics, College of Administration & Economic, Baghdad University, Iraq
³Department of Banks Economics, College of Business Economics, Al-Nahrain University, Iraq
⁴Dies and Tools Eng. Dept., Technical College Baghdad – Iraq

Abstract In this paper, the probability distribution (Generalized Rayleigh) with two parameters \((\theta, \sigma^2)\), in case of outlier, is developed, where the probability density function \((p.d.f)\) is defined, and its moment generating function is derived, to help us in finding the moments, also its cumulative distribution function is found to be used, in obtaining the least squares estimator of the parameters \(\sigma^2\) and \(\theta\). The parameters are estimated also by method of moments and method of least square, and also mixture of the estimators are derived, and explained, the estimators of maximum likelihood for \(\theta, \sigma^2\) are also obtained.

Keywords Generalized two Parameter Rayleigh Distribution, Uniform Distribution Outlier, MLE, Moments and Least Square and Mixture Estimators

1. Introduction

The probability Rayleigh distribution naturally arises in cases when the wind speed data is analysis into two-orthogonal dimensional vector components, where the magnitude of components is independent and normally distributed with equal variances. Also this distribution arises in the case of random complex numbers whose real and imaginary components are iid as normal.

The two parameters parameters, Burr Type X, which introduced by [7], is called Generalized Rayleigh distribution. In this paper, we introduce a new two-parameter Generalized Rayleigh in presence of one outlier generated from another distribution, after definition of proposed distribution, its Moment-Generating function is derived, to help us in finding the first and second moments for this distribution, these moments are used to obtain the mixture estimator of parameters, as well as the Maximum Likelihood estimators. The paper is organized as follows: In section (2) we present the GR\((\theta, \sigma^2)\) and its properties, and the provides its \(m.g.f\). Then section (3), discuss finding moment generating function and the methods of estimating parameters which are Maximum Likelihood and Least Squares and Mixture of estimators are derived. These estimators are compared using (MSE), through simulation programs, prepared for this purpose.

In this paper, we introduce a new family of continuous distribution, called a new – two parameter generalized Rayleigh GR\((\theta, \sigma^2)\) in presence of outlier generated distribution from another distribution i.e the distribution depend on mixing the distribution of \((x_1, x_2, \ldots, x_{n-1})\) random variables, distributed as Rayleigh with scale parameter \((\sigma^2 = 1)\), and shape parameter \(\theta\) while the \((x_n)\) random variable which represents the outlier is one random variable that is uniformly distributed \(x_n \sim \text{uniform}(0, \theta \cdot \sigma^2)\). So the aim of research is how to find the marginal \(p.d.f\) of this type of distribution, in precence of outlier, and also how to derive its cumulative distribution function, and it's moment generating function to help in obtaining the moments after all required derivation three methods are introduced include moment estimators and least squares estimators, and maximum likelihood estimators.

2. Definition of Distribution

Let the random variables \((x_1, x_2, \ldots, x_n)\) are such that \((n - 1)\) of them are distributed as Rayleigh with scale parameter \((\sigma^2 = 1)\), and shape parameter \((\theta)\), i.e.

\[
f(x, \theta) = \theta x e^{-\frac{x^2}{2}} (1 - e^{-\frac{x^2}{2}})^{\theta - 1}, \quad x > 0
\] (1)

With cumulative \(c.d.f\) is:
The second term integral can be evaluated as follows:

\[ f(x, \theta, \sigma^2) = \frac{1}{\eta \sigma^2} \cdot I(\theta, \sigma^2) \]

\[ + \left( \frac{n-1}{n} \right) \theta x e^{-\frac{x^2}{2}} \left( 1 - e^{-\frac{x^2}{2}} \right)^{\theta-1} \]

\[ x > 0, \quad \theta > 0, \sigma^2 > 0 \]

Therefore the marginal distribution of \( X \) is:

\[ F(x, \theta, \sigma^2) = \frac{x}{n \theta \sigma^2} + \left( \frac{n-1}{n} \right) \left( 1 - e^{-\frac{x^2}{2}} \right)^{\theta} \]

### 3. Moment Generating Function

The \( r \)th moments of \( x \) may be determined directly or using Moment generating function technique. Now we derive \( M_x(t) \) for the distribution defined in (1):

\[ M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x, \theta, \sigma^2) dx \]

\[ = \int_0^\infty e^{tx} \frac{1}{\eta \sigma^2} dx + \int_0^\infty e^{tx} \frac{n-1}{n} \theta x e^{-\frac{x^2}{2}} \left( 1 - e^{-\frac{x^2}{2}} \right)^{\theta-1} dx \]

\[ = \frac{1}{n \theta \sigma^2} \left( e^{t \sigma^2} - 1 \right) + \left( \frac{n-1}{n} \theta \right) \int_0^\infty e^{tx} x e^{-\frac{x^2}{2}} \left( 1 - e^{-\frac{x^2}{2}} \right)^{\theta-1} dx \]

\[ = I_2 \]

The second term integral can be evaluated as follows:

Let \( u = 1 - e^{-\frac{x^2}{2}} \Rightarrow du = e^{-\frac{x^2}{2}} x dx \)

\[ \Rightarrow e^{-\frac{x^2}{2}} = 1 - u \Rightarrow x = \left( \ln(1 - u)^2 \right)^{1/2} \]

\[ \therefore I_2 = \left( \frac{n-1}{n} \right) \theta \int_0^1 e^{t \ln(1-u)^2} u^{\theta-1} du \]

after simplification of (7), it can be written as:

Let \( y = 1 - u \Rightarrow u = 1 - y \Rightarrow du = -dy \)

\[ e^{t \ln(1-y)^2} \left( 1 - y \right)^{\theta-1} dy \]

Therefore equation (7) can be written as (8):

\[ \left( \frac{n-1}{n} \right) \theta \int_0^1 (y)^{\theta-1} \left( 1 - y \right)^{\theta-1} dy \]

According to Beta formula:

\[ \left( \frac{n-1}{n} \right) \theta \beta(\theta, 1-2t) = \left( \frac{n-1}{n} \right) \theta \frac{\Gamma(\theta) \Gamma(1-2t)}{\Gamma(\theta+1-2t)} \]

Therefore \( M_x(t) \):

\[ M_x(t) = \frac{1}{n \theta \sigma^2} \left( \frac{\theta \sigma^2 - 1}{t} \right) + \left( \frac{n-1}{n} \right) \theta \beta(\theta, 1-2t) \]

\[ M_x(t) = \frac{1}{n \theta \sigma^2} \left( \frac{\theta \sigma^2 - 1}{t} \right) + \left( \frac{n-1}{n} \right) \theta \frac{\Gamma(\theta) \Gamma(1-2t)}{\Gamma(\theta+1-2t)} \]

Differentiating \( M_x(t) \) and evaluating at \( t = 0 \) we get \( E(x) \) and \( E(x^2) \) as:

\[ M_x'(t) = \frac{1}{n \theta \sigma^2} \left( \frac{\theta \sigma^2 - 1}{t^2} \right) + \left( \frac{n-1}{n} \right) \Gamma(\theta + 1) \left[ \frac{\Gamma(\theta + 1 - 2t) \Gamma'(\theta + 1 - 2t)}{\Gamma^2(\theta + 1 - 2t)} \right] - \left[ \frac{\Gamma(1-2t) \Gamma'(\theta + 1 - 2t)}{\Gamma^2(\theta + 1 - 2t)} \right] \]

\[ M_x'(t) = \frac{1}{n \theta \sigma^2} \left( \frac{\theta \sigma^2 - 1}{t^2} \right) + \left( \frac{\theta \sigma^2 - 1}{t^2} \right) \left[ \frac{2(n-1)}{n} \right] \Gamma(\theta + 1) \frac{\Gamma(1-2t) \Gamma'(\theta + 1 - 2t)}{\Gamma^2(\theta + 1 - 2t)} - \left[ \frac{\Gamma'(1-2t)}{\Gamma(\theta + 1 - 2t)} \right] \]
\[
M'_c(0) = \frac{\theta \sigma^2}{2n} + \left[\frac{2(n - 1)}{n}\right] \Gamma(\theta + 1) \frac{\Gamma'(\theta + 1)}{\Gamma^2(\theta + 1)} \frac{\Gamma''(\theta + 1)}{\Gamma'(\theta + 1)} - \frac{2(n - 1)}{n} \Gamma(\theta + 1)
\]

Taking the second derivative \(M''_c(t)\), we have \(E(x^2)\) as:

\[
M''_c(t) = \frac{1}{n\theta \sigma^2}\left(\frac{t e^{\theta \sigma^2} (e^{\theta \sigma^2})^2 - \theta^2 e^{\theta \sigma^2} (e^{\theta \sigma^2})^2 - (e^{\theta \sigma^2} - 1) (2t)}{t^2 (e^{\theta \sigma^2} - 1) (2t - 1)}\right) + \frac{2(n - 1)}{n} \Gamma(\theta + 1)
\]

Then:

\[
M'_c(0) = \frac{(\theta \sigma^2)^2}{3n} + \frac{2(n - 1)}{n} \Gamma(\theta + 1)\left[\frac{2\Gamma''(\theta + 1) - 2\Gamma''(\theta + 1)\Gamma' + 2\Gamma'(\theta + 1)}{2\Gamma' + 2\Gamma'(\theta + 1)} + \frac{2\Gamma'(\theta + 1)}{\Gamma^2(\theta + 1)}\right] + \frac{4\Gamma^{-1}(\theta + 1)}{\Gamma^2(\theta + 1)} + \frac{2(n - 1)}{n} \Gamma(\theta + 1)\left[\frac{\Gamma'(\theta + 1)\Gamma' + 1}{\Gamma^2(\theta + 1)} - \frac{\Gamma'(\theta + 1)}{\Gamma^2(\theta + 1)}\right]
\]

Therefore:

\[
E(x^2) = \frac{(\theta \sigma^2)^2}{3n} + \frac{2(n - 1)}{n} \Gamma(\theta + 1)\left[\frac{2\Gamma''(\theta + 1) - 2\Gamma''(\theta + 1)\Gamma' + 2\Gamma'(\theta + 1)}{2\Gamma' + 2\Gamma'(\theta + 1)} + \frac{2\Gamma'(\theta + 1)}{\Gamma^2(\theta + 1)}\right] + \frac{4\Gamma^{-1}(\theta + 1)}{\Gamma^2(\theta + 1)} + \frac{2(n - 1)}{n} \Gamma(\theta + 1)\left[\frac{\Gamma'(\theta + 1)\Gamma' + 1}{\Gamma^2(\theta + 1)} - \frac{\Gamma'(\theta + 1)}{\Gamma^2(\theta + 1)}\right]
\]

To estimate \(\theta\) we can solve equation (3) by Newton Raphson method. The solution of equation (15) is:

\[
\theta_{i+1} = \theta_i - \frac{g(\theta_i)}{g'(\theta_i)}
\]
If \( \hat{\sigma}^2 \) known then:
\[
\hat{\theta} = \frac{2n(\bar{x} - \mu)}{\hat{\sigma}^2} \tag{17}
\]
This value can be considered as initial value for solving equation (16) by Newton Raphson method.
Also to proceed in finding the moment estimator for the parameter \( \sigma^2 \), the following equation is applied as follows:
\[
\hat{\theta} \hat{\sigma}^2 = 2n(m_1 - A) \tag{18}
\]
\[
m_2 - B = \frac{3n}{m_2 - B} = \frac{4n^2(m_1 - A)^2}{3} \tag{18}
\]
Also the estimate \( \hat{\sigma}^2 \) can be obtained by solving equation (18) by Newton Raphson method:
\[
\sigma_{i+1}^2 = \sigma_i^2 - \frac{g(\sigma_i^2)}{\frac{\partial g(\sigma_i^2)}{\partial \sigma_i^2}} \tag{19}
\]
where:
\[
g(\sigma_i^2) = m_2 - B - (\sigma_i^2)^2 - \frac{4n}{3}(m_1 - A(\sigma_i^2)^2) + \frac{8n}{3}(m_1 - A(\sigma_i^2))A(\sigma_i^2) \tag{19}
\]
\[
B = \frac{-4(n - 1) \Gamma(\theta + 1) - \frac{n}{\Gamma(\theta + 1)} + \frac{8(n - 1) \Gamma(\theta + 1)}{\Gamma(\theta + 1)} \Gamma(\theta + 1)}{\frac{n}{\Gamma(\theta + 1)}} \tag{19}
\]
\[
\Gamma''(1) + \frac{8(n - 1) \Gamma(\theta + 1)}{\Gamma(\theta + 1)} \Gamma'(1) + \frac{4(n - 1)}{\Gamma(\theta + 1)} \Gamma'(1) \tag{19}
\]
\[
\frac{n}{\Gamma(\theta + 1)} \Gamma'(1) + \frac{8(n - 1) \Gamma(\theta + 1)}{\Gamma(\theta + 1)} \Gamma'(1) - \frac{8(n - 1) \Gamma'\Gamma'(\theta + 1)}{\Gamma(\theta + 1)} \tag{19}
\]
\[
\frac{4(n - 1)}{\Gamma(\theta + 1)} \Gamma(\theta + 1) - \frac{n}{\Gamma(\theta + 1)} \Gamma(\theta + 1) \tag{19}
\]

4. Least Squares Estimators

In this section we provide the regression based method estimators of unknown parameters, which was originally suggested by [16] to estimate the parameters of Beta distribution.
Suppose \( Y_{(i)}; i = 1, 2, \ldots, n \) denotes the ordered sample, where:
\[
E[G(Y_{(i)})] = \frac{i}{n + 1} \quad \text{var}[G(Y_{(i)})] = \frac{i(n - i + 1)}{(n + 1)^2(n + 2)}
\]
The \( LS \) estimators for the unknown parameters \( (\theta, \sigma^2) \) of the proposed \( p.d.f \) of Generalized Rayleigh with two parameters \( (\theta, \sigma^2) \) in case of outlier (i.e equation 3), and its c.d.f (equation 5), can be obtained by minimizing the quantity \( T \), defined in equation (19):
\[
T = \sum_{i=1}^{n} \left( F(x_i) - \frac{i}{n + 1} \right)^2 \tag{19}
\]
with respect to \( \theta \) and \( \sigma^2 \):
\[
\hat{T} = \sum_{i=1}^{n} \left( x_i - \frac{n-1}{n} \right) \left( 1 - e^{-\frac{x_i^2}{\sigma^2}} \right) - i \left( \frac{n}{n + 1} \right)
\]
\[
\frac{\partial T}{\partial \theta} = 2 \sum_{i=1}^{n} \frac{x_i}{\theta \sigma^2} + \frac{n - 1}{n} \left( 1 - e^{-\frac{x_i^2}{\sigma^2}} \right) - \frac{i}{n + 1}
\]
\[
\frac{\partial T}{\partial \sigma^2} = 2 \sum_{i=1}^{n} \frac{n - 1}{\theta \sigma^2} + \frac{n - 1}{n} \left( 1 - e^{-\frac{x_i^2}{\sigma^2}} \right) \tag{20}
\]
Also we find \( \frac{\partial \theta}{\sigma^2} \):
\[
\frac{\partial \theta}{\sigma^2} = 2 \sum_{i=1}^{n} \frac{x_i}{\theta \sigma^2} + \frac{n - 1}{n} \left( 1 - e^{-\frac{x_i^2}{\sigma^2}} \right) \tag{20}
\]
\[
\sum_{i=1}^{n} \frac{x_i}{\theta \sigma^2} - \frac{i}{n + 1} \sum_{i=1}^{n} \frac{x_i}{\theta \sigma^2} \tag{20}
\]
\[
\sum_{i=1}^{n} \frac{x_i^2}{\theta \sigma^2} = \hat{\sigma}^2 \left( \frac{n - 1}{n} \right) \left( \frac{1}{n \theta} \right) \tag{20}
\]
\[
\sum_{i=1}^{n} \frac{x_i}{\theta \sigma^2} = \hat{\sigma}^2 \left( \frac{n - 1}{n} \right) \left( \frac{1}{n \theta} \right) \tag{20}
\]
\[
\sum_{i=1}^{n} \frac{x_i^2}{\theta \sigma^2} = \hat{\sigma}^2 \left( \frac{n - 1}{n} \right) \left( \frac{1}{n \theta} \right) \tag{20}
\]
\[
\hat{\sigma}^2_{LS} = \frac{\Sigma_{i=1}^{n} \frac{x_i^2}{\theta \sigma^2}}{\Sigma_{i=1}^{n} x_i (1 - e^{-\frac{x_i^2}{\theta \sigma^2}}) - \Sigma_{i=1}^{n} \frac{x_i}{\theta \sigma^2}} \tag{21}
\]
Where \( \hat{\theta}_{LS} \) is estimated from equation (20) which is non linear equation can be solved using Newton Raphson method, or we can use Moment estimator of \( \theta \) (equation 17) to obtain least square estimator of parameter \( (\sigma^2) \), and then obtain the mixed estimator.

5. Maximum Likelihood Estimators

Let the random variables \( (x_1, x_2, \ldots, x_n) \) are such that \( (n - 1) \) of them are distributed with \( p.d.f \) \( f(x, \theta) \) and
one variable is distributed with p.d.f. \( f(x, \theta \sigma^2) \), so the joint distribution of \( (x_1, x_2, ..., x_n) \) are:

\[
\begin{align*}
&f(x_1, x_2, ..., x_n, \theta, \sigma^2) = \\
&\frac{(n-1)!}{n!} \prod_{i=1}^{n} f(x_i, \theta) \sum_{A_1} = 1 g(x_{A_1}) \\
&f(x, \theta) = 8x e^{-\frac{x^2}{2}} (1 - e^{-\frac{x^2}{2}})^{\theta-1} \\
g(x, \theta \sigma^2) = \frac{1}{\theta \sigma^2} \\
f(x_1, x_2, ..., x_n, \theta, \sigma^2) = \frac{(n-1)!}{n!} \frac{x^n e^{-\frac{x^2}{2}}}{(1 - e^{-\frac{x^2}{2}})^{\theta-1}} \prod_{i=1}^{n} \frac{x_i}{\theta \sigma^2} (1 - e^{-\frac{x_i^2}{2}}) ^{\theta-1}
\end{align*}
\tag{21}
\]

6. Since the c.d.f. was found in closed from equation (5), it can be used to find the LS estimator, through minimizing equation (19), the estimators of LS are obtained from equations (20, 21).

7. The Likelihood function is defined in equations (21), (22) and (23), and then the Maximum Likelihood estimator's are explained in equations (24) and (25).

8. Any estimators of Moments, LS can be used in MLE to obtain the Mixture estimators.

6. Conclusions

This paper offers a new family of distributions, the two parameter Rayleigh distribution in the presence of one outlier, which is important for analysis lifetime data. The distribution has two parameters (scale parameter \( \sigma^2 \), and shape parameter \( \theta \)), and consist of mixing the distribution of \( (x_1, x_2, ..., x_{n-1}) \) random variables with the distribution of random variable \( (x_n) \), i.e. the distribution of \( x_1, x_2, ..., x_{n-1} \) is Rayleigh distribution with \( (\sigma^2 = 1) \) and \( x_n \) uniformly with \( (0, \theta \sigma^2) \), and the replicate each experiment \( (R=1000) \), and to use mean square error (MSE) or integrated mean square error (IMSE) for comparison.

We have studied various method for estimating the parameters, (Least Squares, Moments, Mixture), and derived the Moment Generating function, which is used to obtain the first and second Moment of this GR(\( \theta, \sigma^2 \)), and then used as a possible alternative method for estimating Parameters.

This work will be done in another suggestion in future, to apply this method on another distribution.

REFERENCES

[1] Aslam, M., (2008). "Economic reliability acceptance sampling plan for generalized Rayleigh distribution", Journal of Statistics, 15(1), 26-35.

[2] Aslam, M., C.-H. Jun, Y.L. Lio, and M. Ahmad, (2011) “Group acceptance sampling plans for the generalized Rayleigh distribution", International Journal of Intelligent Technologies & Applied Statistics, 4(3), 355-365, September 2011.

[3] A. Soliman, Essam A. Amin, and Alaa A. Abd-El Aziz,(2010) , "Estimation and Prediction from Inverse Rayleigh Distribution Based on Lower Record Values", Applied Mathematical Sciences, Vol. 4, 2010, no. 62, 3057 – 3066.

[4] Hryniewicz, Q., (2008)," Statistics with fuzzy data in statistical quality control", Soft Company. 12:229 – 234.

[5] Kishore K. Das and Tanusree Deb Roy, (2011)," Applicability of Length Biased Weighted Generalized Rayleigh Distribution" Pelagia Research Library Advances in Applied Science Research, 2011, 2 (4) 320 - 327.

[6] Mohamed Saber Fallah, Seyed Taghi Akhavan, and Mohamed Hossein Abooe, (2011) " A new Acceptance sampling plan based on cumulative sums of conforming Run
References

[7] Muhammad Aslam, (2008), "Economic Reliability Acceptance Sampling Plan for Generalized Rayleigh Distribution" ISSN 1684-8403 Journal of Statistics Volume 15, 2008, pp. 26-35.

[8] Muhammad Aslam1, Chi-Hyuck Jun, Y. L. Lio and Munir Ahmad,(2011)," Group Acceptance Sampling Plans for the Generalized Rayleigh Distribution" International Journal of Intelligent Technologies and Applied Statistics Vol.4, No.3 (2011) pp.355-365.

[9] Raqab, M. Z. (1998); "Order Statistics from the Burr Type X Mode"; Computer Mathematics and Applications, Vol. 36, 111-120.

[10] Raqab, M. Z. and Kundu, D. (2006); "Burr Type X Distribution: Revisited"; Journal of Probability and Statistical Science, Vol. 4, 179-193.

[11] Saima Altaf, and Muhamed Aslam, (2011), " Bayesian analysis of the Davidson model for paired comparision with order effect using non informative priors", Pak. J. Stats. Vol. 27(2), 171 – 185.

[12] Sanka Dey, (2008), "Minimax estimation of the parameter of the Rayleigh distribution under quadratic loss function", Data Science Journal, Vol. 7,23.

[13] Surles, J. G. and Padgett, W. J. (1998); " inference for \( p(y < x) \) in the Burr type X model", Journal of applied statistical sciences, 7, 225-238.

[14] Surles, J. G. and Padgett, W. J. (2001); "inference For Reliability and Stress-Strength For a Scaled Burr Type X Distribution"; Lifetime Data Analysis, Vol. 7, 187-200.

[15] Surles, J. G. and Padgett, W. J. (2005); "Some Properties of a Scaled Burr Type X Distribution"; Journal of Statistical Planning and Inference, Vol. 128, 271-280.

[16] Swain, J.J., Venkatraman S., Wilson, J.R. (1988), "Distribution Selection and Validation", Journal of Statistical Computation and Simulation, Vol. 29, pp. 271-297.

[17] Tang, K. (1988)," Economic design of product specifications for a complete inspection plan", International Journal of production research, Vol. 26. No. 2, 203 – 219.