DOI: https://doi.org/10.21123/bsj.2023.7962

Estimating the Parameters of Exponential-Rayleigh Distribution under Type-I Censored Data

Rihaam N. Shatti * ***Iden H. Al-Kinani***

Department of Mathematics, College of Science for women, University of Baghdad, Baghdad, Iraq
*Corresponding author: riham.najm1203a@csw.uobaghdad.edu.iq
E-mail address: iden_alkanani@yahoo.com

Received 16/10/2022, Revised 17/12/2022, Accepted 18/12/2022, Published Online First 20/5/2023, Published 01/1/2024

Abstract:

This paper discusses estimating the two scale parameters of Exponential-Rayleigh distribution for singly type one censored data which is one of the most important Rights censored data, using the maximum likelihood estimation method (MLEM) which is one of the most popular and widely used classic methods, based on an iterative procedure such as the Newton-Raphson to find estimated values for these two scale parameters by using real data for COVID-19 was taken from the Iraqi Ministry of Health and Environment, AL-Karkh General Hospital. The duration of the study was in the interval 4/5/2020 until 31/8/2020 equivalent to 120 days, where the number of patients who entered the (study) hospital with sample size is (n=785). The number of patients who died during the period of study was (m=88). And the number of patients who survived during the study period was (n-m=697), then utilized one of the most important non-parametric tests which is the Chi-square test to determine if the sample (data) corresponded with the Exponential-Rayleigh distribution (ER). then, after estimating the parameters of ER distribution for singly type-I censoring data, compute the survival function, hazard function, and probability density function.

Keywords: COVID-19, Exponential-Rayleigh distribution (ER), Hazard function, Singly type one censored data, Survival function.

Introduction:

In many life testing and reliability studies, the experimenters and researchers may not get complete information on failure times for experimental units, the Data obtained from experiments are called censoring data [1]. The censored sample is divided into three types: Right-censored sample, Left-censored sample, and Interval-censored sample.

The right-censored sample is also divided into three branches, singly type one censoring sample, singly type two censoring sample, and progressively censoring sample. This paper depends on the most important Rights censored data which type one censored data.

In 2013, Hussain I, et al, estimated the unknown parameters of the generalized Rayleigh distribution for singly type one censored sample by using one of the most important classical estimation methods which maximum likelihood estimation method [2].

In 2016, Makhdoom I, et al, estimated the parameters of an Exponential distribution using approximation forms of Lindley, and approximation of Tierney and Kadane under type two censoring samples. And to compare the effectiveness of the various methods, a Monte Carlo simulation is used [1].

In 2020, Abadi QS, et al, used the (MLE) method to estimate three parameters of a new mixture distribution consisting of one shape parameter and two scale parameters for type-I censored data and progressively censored data [3, 4].

In 2021, Heidari KF, et al, generalized the estimated parameters of Rayleigh distribution by using the Bayesian method for type two censoring data under the squared error loss function [5].

In 2021, Mazaal AR, et al, based on a singly type II censored sample for Weibull Stress- Strength with considering the stress-strength Reliability estimation. And considered the Bayesian analysis using the loss function under (Extension of Jeffery and Gamma) which are prior functions [6].

Baghdad Science Journal
2024, 21(1): 146-150
P-ISSN: 2078-8665
E-ISSN: 2411-7986

The properties of Exponential-Rayleigh distribution

This paper depends on the formula that was proven by Hussein LK, Hussein IH, and Rasheed HA in 2021, by using the mix between the cumulative of each distribution as follows [7] :

Let $T = \max(v, w)$, where v and w are two independent random variables, then:

$F(t;\theta,\beta) = p_r(T \le t)$

$F(t;\theta,\beta) = 1 - p_r(T > t)$

$F(t;\theta,\beta) = 1 - p_r(\max(v,w) > t)$

$F(t;\theta,\beta) = 1 - [\, p_r(v > t).\, p_r(w > t)]$

$F(t;\theta,\beta) = 1 - [(\int_t^\infty f(v;\theta)dv).(\int_t^\infty f(w;\beta)dw)]$

$F(t;\theta,\beta) = 1 - [(\int_t^\infty \theta e^{-\theta v}dv).\left(\int_t^\infty \beta w e^{-\frac{\beta}{2}w^2}dw\right)$

$$F(t;\theta,\beta) = 1 - e^{-(\theta t+\frac{\beta}{2}t^2)} \quad 1$$

- The probability density function is given by:

$$f(t;\theta,\beta) = (\theta+\beta t)e^{-(\theta t+\frac{\beta}{2}t^2)} \quad t \ge 0 \;;\; \theta,\beta > 0 \quad 2$$

And zero for otherwise, where θ and β are scale parameters.

- The survival function is obtained by:

$$S(t;\theta,\beta) = e^{-(\theta t+\frac{\beta}{2}t^2)} \quad t \ge 0 \;;\; \theta,\beta > 0 \quad 3$$

- The Hazard rate function can be expressed as:

$$h(t;\theta,\beta) = (\theta+\beta t) \quad t \ge 0 \;;\; \theta,\beta > 0 \quad 4$$

Parameters Estimation

This section shows the derivative and estimates the unknown parameters of (ER) distribution, using the maximum likelihood estimation method for type-one censoring sample.

Type-One Censoring Data [3]

It is one of the most common right-censored types, the time of experiments denoted by t, which is fixed, while the number of failures time denoted by m observed is random. However, n represents the individuals or items that are placed in the study, and the units or individuals that did not fail or survive are denoted by (n-m).

Maximum Likelihood Estimation Method for Type-One Censoring Data

The maximum likelihood estimation method is one of the most popular and widely used classic methods. This method was initially used in 1912 by R.A. Fisher in his first statistical papers. Maximum likelihood describes a technique for estimating the unknown parameters of any distribution [8].

$$L = \frac{n!}{(n-m)!}[s(t_m;\theta,\beta)]^{n-m}\prod_{i=1}^{m}[f(t_i;\theta,\beta)] \quad 5$$

let $\frac{n!}{(n-m)!} = k$, then:

$$L = k[e^{-\left(\theta t_m+\frac{\beta}{2}t_m^2\right)}]^{n-m}\prod_{i=1}^{m}[(\theta+\beta t_i)e^{-\left(\theta t_i+\frac{\beta}{2}t_i^2\right)}] \quad 6$$

Taking the natural logarithm for both sides of the equation:

$$\ln L = \ln k - (n-m)\theta t_m - \frac{(n-m)}{2}\beta t_m^2 + \sum_{i=1}^{m}\ln(\theta+\beta t_i) - \sum_{i=1}^{m}(\theta t_i + \frac{\beta}{2}t_i^2) \quad 7$$

Deriving Eq.7 partially with respect to α and β respectively, and setting it equal to zero:

$$\frac{\partial \ln l}{\partial\theta} = -(n-m)t_m + \sum_{i=1}^{m}\frac{1}{(\theta+\beta t_i)} - \sum_{i=1}^{m}t_i = 0 \quad 8$$

$$\frac{\partial \ln l}{\partial\beta} = -\frac{(n-m)}{2}t_m^2 + \sum_{i=1}^{m}\frac{t_i}{(\theta+\beta t_i)} - \frac{1}{2}\sum_{i=1}^{m}t_i^2 = 0 \quad 9$$

Now, putting $\frac{\partial lnl}{\partial\theta}$ as a function f(θ) and put $\frac{\partial lnl}{\partial\beta}$ as a function f(β)

$$f(\acute{\theta}_k) = -(n-m)t_m + \sum_{i=1}^{m}\frac{1}{(\theta+\beta t_i)} - \sum_{i=1}^{m}t_i \quad 10$$

$$f(\acute{\beta}_k) = -\frac{(n-m)}{2}t_m^2 + \sum_{i=1}^{m}\frac{t_i}{(\theta+\beta t_i)} - \frac{1}{2}\sum_{i=1}^{m}t_i^2 \quad 11$$

Note that Eq.10 and Eq.11 are so difficult to solve, then employing them by an iterative method such as Newton Raphson procedure [9] to find the value of $\hat{\theta}$ and $\hat{\beta}$ as:

$$\theta_{k+1} = \theta_k - \frac{f(\theta_k)}{\acute{f}(\theta_k)} \quad 12$$

Where $\acute{f}(\theta_k) = -\sum_{i=1}^{m}\frac{1}{(\theta+\beta t_i)^2} \quad 13$

And $\beta_{k+1} = \beta_k - \frac{f(\beta_k)}{\acute{f}(\beta_k)} \quad 14$

Where $\acute{f}(\beta_k) = -\sum_{i=1}^{m}\frac{t_i^2}{(\theta+\beta t_i)^2} \quad 15$

The error term denoted by ϵ, which is very small and assumed value, is the value for the difference between the new values of θ and β in the new iterative with the previous values of θ and β in the past iterative.

The error term is formulated as:

$\epsilon_{k+1}(\theta) = \theta_{k+1} - \theta_k$

$\epsilon_{k+1}(\beta) = \beta_{k+1} - \beta_k$

Where θ_k and β_k are initial values that are assumed.

Results and Discussion:

This study is based on a sample and real data taken from the Iraqi Ministry of Health and Environment, AL-Karkh General Hospital about COVID-19. The duration of the study was the

interval from 4/5/2020 until 31/8/2020 equivalent to 120 days. The period of study was (May, June, July, and August).
The number of patients who entered the (study) hospital which is the sample size was (n=785).
The number of patients who died during the period of study was (m=88).
The number of patients who survived during the study period was (n-m=697).
In this study, the Chi-square test which is one of the non-parametric tests is used to determine if the sample (data) correspond with the Exponential-Rayleigh distribution (ER).
The null and alternative hypotheses for Chi-square test are:
H_0 : The data are distributed as (ER) distribution.
H_1 : The data are not distributed as (ER) distribution.
The formula of this test is given by:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{O_i}$$

Where:
O_i refers to the observed in class i.
E_i refers to the expected frequency in class i.
k refers to the number of class.

The calculated value (15.32008) is less than the tabulated value (21.67) with the degree of freedom (9) and level of significance (0.01). That means accepting the null hypothesis H_0 and the data is distributed as (ER) distribution.

After that, employing MATLAB programming (version 2021) to estimate values of the following parameters: $\hat{\theta} = 0.0100$, $\hat{\beta} = 0.0183$. When the initial values are $\theta_0 = 0.0120$, $\beta_0 = 0.0001$.
Substitute the estimated values of parameters and the values of the lifetime in Eq.1 , Eq.2 , Eq.3. And Eq.4 Then get the values of $\hat{f}(t)$, $\hat{F}(t)$, $\hat{S}(t)$, and $\hat{h}(t)$ in the following Table.

Table 1. The values of $\hat{f}$(t), $\hat{F}$(t), $\hat{S}$(t), and $\hat{h}$(t) by (MLE) method for singly type one censored data.

patients	Life time t	$\hat{f}(t)$	$\hat{F}(t)$	$\hat{S}(t)$	$\hat{h}(t)$
1	1	0.02776	0.018968	0.981032	0.0283
2	1	0.02776	0.018968	0.981032	0.0283
3	1	0.02776	0.018968	0.981032	0.0283
4	1	0.02776	0.018968	0.981032	0.0283
5	1	0.02776	0.018968	0.981032	0.0283
6	1	0.02776	0.018968	0.981032	0.0283
7	1	0.02776	0.018968	0.981032	0.0283
8	1	0.02776	0.018968	0.981032	0.0283
9	1	0.02776	0.018968	0.981032	0.0283
10	1	0.02776	0.018968	0.981032	0.0283
11	1	0.02776	0.018968	0.981032	0.0283
12	1	0.02776	0.018968	0.981032	0.0283
13	1	0.02776	0.018968	0.981032	0.0283
14	1	0.02776	0.018968	0.981032	0.0283
15	1	0.02776	0.018968	0.981032	0.0283
16	2	0.04404	0.05503	0.94497	0.04660
17	2	0.04404	0.05503	0.94497	0.04660
18	2	0.04404	0.05503	0.94497	0.04660
19	2	0.04404	0.05503	0.94497	0.04660
20	2	0.04404	0.05503	0.94497	0.04660
21	2	0.04404	0.05503	0.94497	0.04660
22	2	0.04404	0.05503	0.94497	0.04660
23	3	0.05800	0.10627	0.89373	0.06490
24	3	0.05800	0.10627	0.89373	0.06490
25	3	0.05800	0.10627	0.89373	0.06490
26	3	0.05800	0.10627	0.89373	0.06490
27	3	0.05800	0.10627	0.89373	0.06490
28	4	0.06905	0.17006	0.82994	0.08320
29	4	0.06905	0.17006	0.82994	0.08320
30	4	0.06905	0.17006	0.82994	0.08320
31	4	0.06905	0.17006	0.82994	0.08320
32	4	0.06905	0.17006	0.82994	0.08320
33	4	0.06905	0.17006	0.82994	0.08320
34	4	0.06905	0.17006	0.82994	0.08320
35	4	0.06905	0.17006	0.82994	0.08320

36	4	0.06905	0.17006	0.82994	0.08320
37	5	0.07681	0.24327	0.75673	0.10150
38	5	0.07681	0.24327	0.75673	0.10150
39	5	0.07681	0.24327	0.75673	0.10150
40	5	0.07681	0.24327	0.75673	0.10150
41	6	0.08116	0.32254	0.67746	0.11980
42	6	0.08116	0.32254	0.67746	0.11980
43	6	0.08116	0.32254	0.67746	0.11980
44	6	0.08116	0.32254	0.67746	0.11980
45	6	0.08116	0.32254	0.67746	0.11980
46	6	0.08116	0.32254	0.67746	0.11980
47	6	0.08116	0.32254	0.67746	0.11980
48	6	0.08116	0.32254	0.67746	0.11980
49	6	0.08116	0.32254	0.67746	0.11980
50	6	0.08116	0.32254	0.67746	0.11980
51	6	0.08116	0.32254	0.67746	0.11980
52	6	0.08116	0.32254	0.67746	0.11980
53	6	0.08116	0.32254	0.67746	0.11980
54	6	0.08116	0.32254	0.67746	0.11980
55	6	0.08116	0.32254	0.67746	0.11980
56	6	0.08116	0.32254	0.67746	0.11980
57	6	0.08116	0.32254	0.67746	0.11980
58	7	0.08224	0.4045	0.59550	0.13810
59	7	0.08224	0.4045	0.59550	0.13810
60	7	0.08224	0.4045	0.59550	0.13810
61	7	0.08224	0.4045	0.59550	0.13810
62	7	0.08224	0.4045	0.59550	0.13810
63	8	0.08038	0.48603	0.51397	0.15639
64	8	0.08038	0.48603	0.51397	0.15639
65	8	0.08038	0.48603	0.51397	0.15639
66	8	0.08038	0.48603	0.51397	0.15639
67	9	0.07609	0.56445	0.43555	0.17470
68	9	0.07609	0.56445	0.43555	0.17470
69	9	0.07609	0.56445	0.43555	0.17470
70	10	0.06994	0.6376	0.36240	0.19299
71	10	0.06994	0.6376	0.36240	0.19299
72	11	0.06256	0.70393	0.29607	0.21130
73	12	0.05453	0.7625	0.23750	0.2296
74	12	0.05453	0.7625	0.23750	0.2296
75	12	0.05453	0.7625	0.23750	0.2296
76	12	0.05453	0.7625	0.23750	0.2296
77	12	0.05453	0.7625	0.23750	0.2296
78	12	0.05453	0.7625	0.23750	0.2296
79	13	0.04637	0.81294	0.18706	0.24789
80	15	0.03125	0.89016	0.10984	0.28450
81	15	0.03125	0.89016	0.10984	0.28450
82	15	0.03125	0.89016	0.10984	0.28450
83	16	0.02480	0.91811	0.08189	0.30285
84	17	0.01925	0.94006	0.05994	0.32115
85	18	0.01462	0.95692	0.004308	0.33937
86	18	0.01462	0.95692	0.004308	0.33937
87	18	0.01462	0.95692	0.004308	0.33937
88	18	0.01462	0.95692	0.004308	0.33937

1. The values of $\hat{f}(t)$ are increasing until t=7 and it is decreasing when $8 \leq t \leq 18$.
2. The values of $\hat{F}(t)$ are increasing with the increase in failure times.
3. The values of $\hat{S}(t)$ are decreasing with the increase in failure times.
4. The values of $\hat{h}(t)$ are increasing with the increase in failure times.

Conclusions:

In this paper, the parameters of Exponential-Rayleigh distribution can be derived for singly type one censored sample, the estimator values of these scale parameters can be obtained by using the maximum likelihood estimation method (MLE) based on an iterative procedure such as Newton-Raphson. Based on real data on Coronavirus. Note that there is a direct relationship between the hazard rate function and the failure times, $\hat{h}(t)$ is increasing with the increase in failure times. Observing that there is an opposite relationship between the survival function and the failure time s, $\hat{S}(t)$ are decreasing with the increase in failure times.

Authors' Declaration:

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are mine.
- Ethical Clearance: The project was approved by the local ethical committee in University of Baghdad.

Authors' Contributions Statement:

I. H. A. proposed the idea of the research, review, suggest the modifications and give advice how to write the paper. R. N. S. calculated the results, write the research and programming the estimated values of the parameters.

References:

1. Makhdoom I, Nasiri P, Pak A. Estimating the Parameter of Exponential Distribution under Type II Censoring From Fuzzy Data. J Mod Appl Stat Methods. 2016; 15(2): 495–509.
2. Hussain I, Kadhum H. Survival estimation for singly type one censored sample based on generalized Rayleigh distribution. Baghdad Sci J. 2014; 11(2): 193–201. https://doi.org/10.21123/bsj.2014.11.2.193-201
3. Abadi QS, Al-kanani IH. Estimating of Survival Function under Type One Censoring Sample for Mixture Distribution. IHJPAS 2020; 33(4): 102–9.
4. Abadi QS, AL-Kanani IH. Progressively censoring data to estimate the survival function in the new mixture distribution. J Crit Rev. 2020; 7(1): 527–30.
5. Heidari KF, Deiri E, Jamkhaneh EB. E-Bayesian and Hierarchical Bayesian Estimation of Rayleigh Distribution Parameter with Type-II Censoring from Imprecise Data. J Indian Soc Probab Stat. 2022; 23(1): 63–76. https://doi.org/10.1007/s41096-021-00112-3
6. Mazaal AR, Karam NS, Karam GS. Comparing weibull stress-strength reliability Bayesian estimation for singly type II censored data under different loss functions. Baghdad Sci J. 2021; 18(2): 306–14. https://doi.org/10.21123/bsj.2021.18.2.0306
7. Hussein LK, Hussein IH, Rasheed HA. An Estimation of Survival and Hazard Rate Functions of Exponential Rayleigh Distribution. Ibn AL- Haitham J Pure Appl Sci. 2021; 34(4): 93–107.
8. Rosenblad A. The Concise Encyclopedia of Statistics. J App Stat. 2011: 867–868.
9. Hameed HH, Al-Saedi HM. The majorant function modelling to solve nonlinear algebraic system. J Interdiscip Math. 2021; 24(7): 1825–9.

تقدير معلمات توزيع رايلي الأسي للنوع الأول من البيانات الخاضعة للرقابة

ريهام نجم شاطي ايدن حسن الكناني

قسم الرياضيات، كلية العلوم للبنات، جامعة بغداد، بغداد، العراق.

الخلاصة:

يهتم هذا البحث في مناقشة تقدير معلمات توزيع رايلي_الأسي للنوع الاول للبيانات الخاضعة للرقابة و ذلك بأستخدام طريقة الامكان الاعظم وهي واحدة من أهم وأشهر طرق التقدير الغير بيزية لتقدير معلمات التوزيع، بعد ذلك تم استخدام احدى الطرق التكرارية وهي طريقة نيوتن رافسن لإيجاد قيم المعلمات المقدرة. في هذه الدراسة تم الاعتماد على بيانات حقيقية خاصة بمرضى بفايروس كورونا تم الحصول عليها من وزارة الصحة العراقية/ مستشفى الكرخ العام وكانت مدة الدراسة من 2020/5/4 الى 2020/8/31 بما يكافئ 120 يوم وبلغ العدد الكلي للمصابين الداخلين خلال فترة الدراسة 785 مريض. بلغ عدد المرضى المتوفين خلال فترة الدراسة 88 مريض وبلغ عدد المرضى المتعافين خلال فترة الدراسة 697 شخص. تم استخدام اختبار كاي_سكوير لأختبار العينة المستخدمة في الدراسة تتوزع توزيع رايلي_الاسي..في النهاية تم حساب الدالة التجميعية ,دالة الكثافة الاحتمالية , دالة البقاء و دالة المخاطرة لتوزيع رايلي-الأسي.

الكلمات المفتاحية: كوفيد-19، توزيع رايلي الأسي، دالة المخاطرة، النوع الأول من البيانات الخاضعة للرقابة، دالة البقاء.