

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/331806715>

Statistical inference for geometric process with the Rayleigh distribution

Article in *Communications* · April 2018

DOI: 10.31801/cfsuasmas.443690

CITATIONS

3

READS

78

4 authors, including:



Cenker Biçer

Kirikkale University

16 PUBLICATIONS 22 CITATIONS

[SEE PROFILE](#)



Mahmut hüdai Kara

Istanbul University

32 PUBLICATIONS 247 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Alternative matrix decompositions for Latent Semantic Analysis [View project](#)



STATISTICAL INFERENCE FOR GEOMETRIC PROCESS WITH THE RAYLEIGH DISTRIBUTION

CENKER BIÇER, HAYRINISA DEMIRCI BIÇER, MAHMUT KARA,
AND HALİL AYDOĞDU

ABSTRACT. The aim of this study is to investigate the solution of the statistical inference problem for the geometric process (GP) when the distribution of first occurrence time is assumed to be Rayleigh. Maximum likelihood (ML) estimators for the parameters of GP, where a and λ are the ratio parameter of GP and scale parameter of Rayleigh distribution, respectively, are obtained. In addition, we derive some important asymptotic properties of these estimators such as normality and consistency. Then we run some simulation studies by different parameter values to compare the estimation performances of the obtained ML estimators with the non-parametric modified moment (MM) estimators. The results of the simulation studies show that the obtained estimators are more efficient than the MM estimators.

1. INTRODUCTION

Counting process is quite suitable and widely used method for the statistical analysis of the occurrence times of successive events. Let we consider a set of data with successive arrival times. Renewal process (RP) can be used for analyzing of this data, if successive arrival times are independent and identically distributed (iid). Although this approach seems theoretically convenient, the data set often contains a monotone trend in real life problems due to the ageing effect and the accumulated wear [6], i.e., the successive arrival times may be independent but not identically distributed. There are more possible approaches in the literature for the analysis of set of successive arrival times with trend, such as non-homogeneous Poisson process and GP [2,7,17].

GP was firstly introduced by Lam [11,12] as a generalization of a renewal process and he applied to replacement problems. To understand GP, see the following definition, [9].

Received by the editors: June 21, 2017, Accepted: November 07, 2017.

2010 *Mathematics Subject Classification.* 60K05, 62F10, 62F12.

Key words and phrases. Parameter estimation; geometric process ; maximum likelihood estimators; asymptotic distribution.

©2018 Ankara University.
Communications Faculty of Sciences University of Ankara-Series A1 Mathematics and Statistics.
Communications de la Faculté des Sciences de l'Université d'Ankara-Séries A1 Mathématiques et Statistiques.

Definition 1. Let X_i be the interarrival time the $(i - 1)$ th and i th events of a counting process $\{N(t), t \geq 0\}$ for $i = 1, 2, \dots$. The counting process $\{N(t), t \geq 0\}$ is said to be a GP with parameter a if there exists a real number $a > 0$ such that $Y_i = a^{i-1}X_i, i = 1, 2, \dots$, are iid random variables with the distribution function F .

a is the ratio parameter of GP. Obviously, GP is a simple monotonic stochastic process. The monotonicity of GP according to the ratio parameter a is given in Table 1.

TABLE 1. Behavior of GP according to values of the ratio parameter a

Parameter Value	Behavior of X_i random variables
$a > 1$	X_i 's are stochastically decreasing
$a < 1$	X_i 's are stochastically increasing
$a = 1$	X_i 's are iid and GP reduces to RP

In the literature, there is a wide range of study on GP. Lam [13], Lam [14], Lam et al. [15] and Braun et al. [5] investigated some of the basic properties of GP by their studies. Until now, the problem of parameter estimation for GP has been solved by assuming that the distribution of the first occurrence time is the Gamma [6], Weibull [3], log-normal [14] and inverse Gaussian [9] distribution.

Estimation of the mean and variance of the first occurrence time X_1 and also ratio parameter a are very important for GP. Because of the fact that they are completely determine the mean and variance of $X_i, i = 1, 2, \dots$. Let $E(X_1) = \beta$ and $Var(X_1) = \theta^2$ for a GP with the ratio parameter a . The mean and variance of X_i 's are as below.

$$E(X_i) = \frac{\beta}{a^{i-1}} \quad i = 1, 2, \dots \quad (1.1)$$

$$Var(X_i) = \frac{\theta^2}{a^{2(i-1)}}, \quad i = 1, 2, \dots \quad (1.2)$$

The main objective of this study is to estimate the parameters in GP when the distribution of first occurrence time X_1 is Rayleigh with parameter λ . In fact, the Rayleigh distribution with parameter λ is a special case of the Weibull distribution with the shape parameter 2 and the scale parameter $\lambda\sqrt{2}$. The problem of statistical inference for GP with the Weibull distribution has been investigated by Aydogdu et al. [3] within the framework of the modified maximum likelihood method (MML). But, it is known that the ML method works better than the MML method in the small sample sizes. As a result of this, evaluating the statistical inference problem for GP with the Weibull distribution within the ML methodology is quite important. However, ML estimators for parameters of GP with the Weibull distribution cannot be obtain explicitly, because of the fact that the first derivatives of the likelihood function involve power functions of the ratio parameter a and shape parameter

of Weibull distribution. Due to divergence problems, they cannot also be solved by numerical methods. Thus, the statistical inference for GP with the Rayleigh distribution within the framework of the ML methodology is of quite importance. The main contribution of this paper is obtain the ML estimators for the parameters of GP with the Rayleigh distribution.

The rest of this paper is organized as follow: Section 2 presents basic information on the Rayleigh distribution. In Section 3, in accordance with the purpose of this study, by using the ML method, the estimators of the parameters a and λ in GP are obtained. Furthermore, asymptotic distributions and consistency properties of ML estimators of the parameters a and λ are investigated. The numerical simulation for comparing the efficiencies of the obtained ML estimators with the MM estimators is given in Section 4. The conclusions of this study are discussed in Section 5.

2. OVERVIEW TO RAYLEIGH DISTRIBUTION

The Rayleigh distribution is frequently used distribution for modelling of positive data from different areas such as communication, health, engineering and reliability etc.. Let X is a Rayleigh distributed random variable with the parameter λ , from now on, will be indicated as $X \sim R(\lambda)$ for brevity. X has the probability density function (pdf)

$$f(x; \lambda) = \frac{x}{\lambda^2} e^{-x^2/2\lambda^2}, \quad x > 0, \quad (2.1)$$

and cumulative distribution function (cdf)

$$F(x, \lambda) = 1 - e^{-x^2/2\lambda^2}, \quad x > 0 \quad (2.2)$$

where λ is the positive and real valued scale parameter of the distribution [10]. If $\lambda = 1$, then distribution is called the standart Rayleigh distribution. The pdf of Rayleigh distribution is unimodal and skewed to the right. The expected value and variance for the Rayleigh distributed random variable X are $E(X) = \lambda\sqrt{\frac{\pi}{2}}$ and $Var(X) = \frac{4-\pi}{2}\lambda^2$. Also, the skewness and kurtosis values of X are $\frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}}$ and $-\frac{6\pi^2-24\pi+16}{(4-\pi)^2}$, respectively.

Let us assume that $X \sim R(\lambda)$. It can be shown that for a constant $c > 0$

$$X \sim R(\lambda) \Rightarrow cX \sim R(c\lambda). \quad (2.3)$$

For more information on the Rayleigh distribution, we refer the readers to [8] and [10].

3. INFERENCE FOR GP

Let X_1, X_2, \dots, X_n be a random sample from a GP with ratio a and $X_1 \sim R(\lambda)$ with the pdf (2.1). From Equation (2.3), X_i has the distribution $R(\frac{\lambda}{a^{i-1}})$ for all $i = 1, 2, \dots, n$ is

$$L(a, \lambda) = \frac{a^{n(n-1)}}{\lambda^{2n}} \prod_{i=1}^n x_i e^{-(a^{i-1}x_i)^2/2\lambda^2}. \quad (3.1)$$

We can write the natural logarithm of the likelihood function given in Equation (3.1) as shown below.

$$\ln L(a, \lambda) = n(n-1) \ln a - 2n \ln \lambda + \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \frac{(a^{i-1}x_i)^2}{2\lambda^2}. \quad (3.2)$$

If the first derivatives of Equation (3.2) according to a and λ are taken, we reach to the following likelihood equations.

$$\frac{\partial \ln L(a, \lambda)}{\partial a} = \frac{n(n-1)}{a} - \frac{1}{a\lambda^2} \sum_{i=1}^n (a^{i-1}x_i)^2 (i-1) = 0 \quad (3.3)$$

$$\frac{\partial \ln L(a, \lambda)}{\partial \lambda} = -\frac{2n}{\lambda} + \frac{1}{\lambda^3} \sum_{i=1}^n (a^{i-1}x_i)^2 = 0 \quad (3.4)$$

Then, from the solution of Equations (3.3)-(3.4), the parameter λ is obtained as

$$\lambda = \left(\frac{1}{2n} \sum_{i=1}^n (a^{i-1}x_i)^2 \right)^{1/2}. \quad (3.5)$$

By substituting the solution of λ into Equation (3.3), we have

$$\left(\frac{n(n-1)}{a} \right) - \left(2n \sum_{i=1}^n (i-1) x_i^2 a^{2i-3} \right) \left(\sum_{i=1}^n (a^{i-1}x_i)^2 \right)^{-1} = 0 \quad (3.6)$$

Let us denote that the ML estimators of a and λ are \hat{a}_L and $\hat{\lambda}_L$, respectively. The \hat{a}_L cannot be obtained analytically from solution of Equation (3.6), because of the power functions of the parameter a .

Equation (3.6) can be solved by using a numerical method such as the Newton Raphson method. The Newton-Raphson iterative formula for the solution of (3.6) is given as

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)} \quad (3.7)$$

where f is considered as an objective function given in Equation (3.6). If we substitute the numerical solution of \hat{a}_L into Equation (3.5), the ML estimator of λ is obtained as below.

$$\hat{\lambda}_{ML} = \left(\frac{1}{2n} \sum_{i=1}^n (\hat{a}_{ML}^{i-1} X_i)^2 \right)^{1/2}, \quad (3.8)$$

The joint distribution of \hat{a}_{ML} and $\hat{\lambda}_{ML}$ estimators is asymptotically normal with mean vector (a, λ) and covariance matrix I^{-1} , (see [4]), that is,

$$\begin{pmatrix} \hat{a}_{ML} \\ \hat{\lambda}_{ML} \end{pmatrix} \sim AN \left(\begin{pmatrix} a \\ \lambda \end{pmatrix}, I^{-1} \right), \quad (3.9)$$

where I^{-1} is the inverse of the Fisher information matrix I , given as

$$I^{-1} = \begin{bmatrix} \frac{3a^2}{n^3} & \frac{3a\lambda}{2n^2} \\ \frac{3a\lambda}{2n^2} & \frac{\lambda^2}{n} \end{bmatrix}. \quad (3.10)$$

See appendix for the derivation of I^{-1} .

(3.9) yields the marginal distribution of \hat{a}_L and $\hat{\lambda}_{ML}$ estimators as

$$\hat{a}_{ML} \sim AN \left(a, \frac{3a^2}{n^3} \right) \quad (3.11)$$

and

$$\hat{\lambda}_{ML} \sim AN \left(\lambda, \frac{\lambda^2}{n} \right),$$

respectively.

Hence, both \hat{a}_{ML} and $\hat{\lambda}_{ML}$ are asymptotically unbiased estimators and they are also consistent, because the asymptotic variance of each of \hat{a}_{ML} and $\hat{\lambda}_{ML}$ converges to zero as $n \rightarrow \infty$.

Also, by considering (3.11), the following hypothesis

$$H_0 : a = 1 \text{ vs. } H_1 : a \neq 1 \quad (3.12)$$

can be tested by using the statistic

$$U = \frac{n^{3/2} (\hat{a}_{ML} - 1)}{\sqrt{3\hat{a}_{ML}^2}}. \quad (3.13)$$

Here, \hat{a}_{ML} is the ML estimate of the parameter a which is obtained using the iterative method given by (3.7). Under hypothesis H_0 given by (3.12), by Slutsky theorem, from (3.11) and consistency of \hat{a}_{ML} , the statistic U is asymptotically normally (AN) distributed with mean zero and variance 1, in other words $U \sim AN(0, 1)$. Thus, by using the statistic U , it can be decided whether GP is suitable or not for given a data set.

4. MONTE CARLO SIMULATION STUDY

In this section, a simulation study was performed to evaluate the estimation performance of the ML estimators obtained in previous section and to compare the efficiencies of the obtained estimators and MM estimators given by [6],[16]:

$$\hat{a}_{MM} = \exp \left(\frac{6}{(n-1)n(n+1)} \sum_{i=1}^n (n-2i+1) \ln X_i \right) \quad (4.1)$$

and

$$\hat{\lambda}_{MM} = \sqrt{\frac{2}{\pi}} \frac{1}{n} \sum_{i=1}^n \hat{Y}_i, \quad (4.2)$$

where $\hat{Y}_i = \hat{a}_{MM}^{i-1} X_i$. Throughout the simulation study, the parameter λ was chosen as 0.5, 1, 1.5, 2, 4. The means, biases and $n \times$ MSEs for the ML and MM estimators were computed for different sample sizes $n = 30, 50, 100$ and the ratio parameters $a = 0.90, 0.95, 1.05, 1.10$. The study results based on $[100,000/n]$ Monte Carlo simulations are given in Table 2-6.

TABLE 2. The simulated means, Biases and $n \times$ MSEs for the ML and MM estimators of the parameters a and λ , when $\lambda = 0.5$

a	n	Method	\hat{a}			$\hat{\lambda}$		
			Mean	Bias	$n \times$ MSE	Mean	Bias	$n \times$ MSE
0.9	30	ML	0.90012	0.00012	0.00284	1.99992	0.00008	3.84934
		MM	0.90000	0.00000	0.00441	2.02659	0.02659	5.80704
	50	ML	0.90001	0.00001	0.00104	2.00149	0.00149	4.06576
		MM	0.90002	0.00002	0.00161	2.02129	0.02129	6.03156
	100	ML	0.89999	0.00001	0.00025	1.99862	0.00138	3.93192
		MM	0.89999	0.00001	0.00040	2.00846	0.00846	6.02902
0.95	30	ML	0.95001	0.00001	0.00323	1.99776	0.00224	3.92955
		MM	0.95011	0.00011	0.00507	2.03201	0.03201	5.97559
	50	ML	0.94998	0.00002	0.00113	1.99656	0.00344	3.98855
		MM	0.94996	0.00004	0.00182	2.01547	0.01547	6.03692
	100	ML	0.95001	0.00001	0.00027	2.00186	0.00186	3.93329
		MM	0.95001	0.00001	0.00044	2.01218	0.01218	5.98921
1.05	30	ML	1.04994	0.00006	0.00380	1.99304	0.00696	3.82719
		MM	1.05003	0.00003	0.00587	2.02445	0.02445	5.63419
	50	ML	1.05010	0.00010	0.00139	2.00333	0.00333	3.95386
		MM	1.05017	0.00017	0.00220	2.02500	0.02500	6.02510
	100	ML	1.04999	0.00001	0.00033	1.99805	0.00195	3.86163
		MM	1.05000	0.00000	0.00054	2.00818	0.00818	5.85188
1.1	30	ML	1.10006	0.00006	0.00438	1.99817	0.00183	4.00071
		MM	1.10025	0.00025	0.00685	2.03509	0.03509	6.16833
	50	ML	1.10007	0.00007	0.00151	2.00380	0.00380	4.00610
		MM	1.10004	0.00004	0.00237	2.02164	0.02164	5.93794
	100	ML	1.09999	0.00001	0.00037	1.99788	0.00212	3.96683
		MM	1.09999	0.00001	0.00060	2.00856	0.00856	6.00147

As can be clearly seen from Table 2, when the number of observations n increases, both bias and $n \times$ MSE values decrease for all the estimators of a and λ . This is an expected result owing to these estimators are both asymptotically unbiased and consistent. Also, according to the results given in Table 2-6, ML estimators have smaller MSE values than MM estimators for all cases. Therefore, we can say that their estimation performance is better than MM estimators. The diagonal elements in I^{-1} given in Equation (3.10) are also known as the minimum variance bounds (MVBs) for estimating a and λ . The simulated variances of the ML estimators and the corresponding MVB values with $a = 1.10$ and $\lambda = 2$ are presented in Table 7. From Table 7, the simulated variances of the ML estimators and the corresponding

TABLE 3. The simulated means, Biases and $n \times$ MSEs for the ML and MM estimators of the parameters a and λ , when $\lambda = 1$

a	n	Method	\hat{a}			$\hat{\lambda}$		
			Mean	Bias	$n \times$ MSE	Mean	Bias	$n \times$ MSE
0.9	30	ML	0.90022	0.00022	0.00287	2.00567	0.00567	3.94513
		MM	0.90021	0.00021	0.00448	2.03610	0.03610	5.91755
	50	ML	0.90005	0.00005	0.00101	2.00182	0.00182	3.91768
		MM	0.89999	0.00001	0.00162	2.01929	0.01929	5.94595
	100	ML	0.90001	0.00001	0.00025	1.99901	0.00099	4.01014
		MM	0.90002	0.00002	0.00040	2.01007	0.01007	5.98562
0.95	30	ML	0.95009	0.00009	0.00323	2.00009	0.00009	3.98117
		MM	0.95017	0.00017	0.00497	2.03391	0.03391	5.96713
	50	ML	0.95009	0.00009	0.00113	2.00353	0.00353	3.96231
		MM	0.95010	0.00010	0.00178	2.02297	0.02297	5.93633
	100	ML	0.95003	0.00003	0.00028	2.00233	0.00233	3.99716
		MM	0.95002	0.00002	0.00045	2.01124	0.01124	6.10600
1.05	30	ML	1.05019	0.00019	0.00385	2.00108	0.00108	3.92159
		MM	1.05024	0.00024	0.00606	2.03296	0.03296	5.91685
	50	ML	1.04996	0.00004	0.00140	1.99976	0.00024	3.99969
		MM	1.04999	0.00001	0.00223	2.02019	0.02019	6.03730
	100	ML	1.05001	0.00001	0.00033	2.00064	0.00064	3.97149
		MM	1.05001	0.00001	0.00055	2.01167	0.01167	6.01511
1.1	30	ML	1.09994	0.00006	0.00424	1.99638	0.00362	3.90510
		MM	1.09999	0.00001	0.00661	2.02830	0.02830	5.88341
	50	ML	1.10001	0.00001	0.00150	2.00082	0.00082	3.92722
		MM	1.10002	0.00002	0.00236	2.02124	0.02124	5.95966
	100	ML	1.09999	0.00001	0.00037	1.99870	0.00130	4.01233
		MM	1.09998	0.00002	0.00059	2.00749	0.00749	5.86874

TABLE 4. The simulated means, Biases and $n \times$ MSEs for the ML and MM estimators of the parameters a and λ , when $\lambda = 1.5$

a	n	Method	\hat{a}			$\hat{\lambda}$		
			Mean	Bias	$n \times$ MSE	Mean	Bias	$n \times$ MSE
0.9	30	ML	0.90005	0.00005	0.00284	2.00119	0.00119	3.90264
		MM	0.90016	0.00016	0.00439	2.03493	0.03493	5.93560
	50	ML	0.89995	0.00005	0.00103	1.99929	0.00071	3.99919
		MM	0.89987	0.00013	0.00159	2.01367	0.01367	5.92713
	100	ML	0.90000	0.00000	0.00025	2.00045	0.00045	3.89088
		MM	0.90000	0.00000	0.00040	2.01035	0.01035	5.89549
0.95	30	ML	0.95001	0.00001	0.00313	1.99717	0.00283	3.82946
		MM	0.95013	0.00013	0.00494	2.03151	0.03151	5.84001
	50	ML	0.95006	0.00006	0.00111	2.00218	0.00218	3.88205
		MM	0.95008	0.00008	0.00180	2.02400	0.02400	5.95549
	100	ML	0.95002	0.00002	0.00027	2.00174	0.00174	3.98022
		MM	0.95003	0.00003	0.00045	2.01338	0.01338	6.07587
1.05	30	ML	1.04978	0.00022	0.00384	1.99469	0.00531	3.90250
		MM	1.04988	0.00012	0.00598	2.02831	0.02831	5.88681
	50	ML	1.04999	0.00001	0.00136	1.99626	0.00374	3.88286
		MM	1.05000	0.00000	0.00216	2.01569	0.01569	5.84545
	100	ML	1.04998	0.00002	0.00034	1.99825	0.00175	3.94494
		MM	1.04998	0.00002	0.00055	2.00876	0.00876	5.97694
1.1	30	ML	1.10012	0.00012	0.00431	1.99943	0.00057	3.99093
		MM	1.10024	0.00024	0.00659	2.03171	0.03171	5.88111
	50	ML	1.10007	0.00007	0.00151	2.00158	0.00158	3.95243
		MM	1.10008	0.00008	0.00240	2.02120	0.02120	5.91306
	100	ML	1.09997	0.00003	0.00038	1.99582	0.00418	4.09140
		MM	1.09997	0.00003	0.00061	2.00666	0.00666	6.08929

TABLE 5. The simulated means, Biases and $n \times \text{MSEs}$ for the ML and MM estimators of the parameters a and λ , when $\lambda = 2$

a	n	Method	\hat{a}			$\hat{\lambda}$		
			Mean	Bias	$n \times \text{MSE}$	Mean	Bias	$n \times \text{MSE}$
0.9	30	ML	0.89992	0.00008	0.00284	1.99636	0.00364	3.82012
		MM	0.89986	0.00014	0.00442	2.02506	0.02506	5.75547
	50	ML	0.90001	0.00001	0.00101	2.00008	0.00008	3.96042
		MM	0.89997	0.00003	0.00162	2.01595	0.01595	5.96905
	100	ML	0.90001	0.00001	0.00025	2.00088	0.00088	3.95471
		MM	0.90001	0.00001	0.00040	2.01110	0.01110	5.98187
0.95	30	ML	0.95016	0.00016	0.00321	2.00324	0.00324	3.92126
		MM	0.95024	0.00024	0.00507	2.03673	0.03673	5.98280
	50	ML	0.94995	0.00005	0.00109	1.99438	0.00562	3.91541
		MM	0.94996	0.00004	0.00175	2.01461	0.01461	5.83645
	100	ML	0.95000	0.00000	0.00028	2.00031	0.00031	4.07335
		MM	0.95002	0.00002	0.00045	2.01258	0.01258	6.15760
1.05	30	ML	1.05016	0.00016	0.00388	1.99936	0.00064	3.88975
		MM	1.05022	0.00022	0.00599	2.03052	0.03052	5.82419
	50	ML	1.05005	0.00005	0.00138	2.00267	0.00267	3.98249
		MM	1.05010	0.00010	0.00215	2.02408	0.02408	5.92146
	100	ML	1.04997	0.00003	0.00033	1.99719	0.00281	3.87698
		MM	1.04998	0.00002	0.00054	2.00810	0.00810	5.90144
1.1	30	ML	1.10002	0.00002	0.00422	1.99815	0.00185	3.91876
		MM	1.09992	0.00008	0.00653	2.02601	0.02601	5.88347
	50	ML	1.10008	0.00008	0.00150	2.00039	0.00039	3.95590
		MM	1.10008	0.00008	0.00244	2.01933	0.01933	6.06236
	100	ML	1.09998	0.00002	0.00037	1.99797	0.00203	3.92220
		MM	1.09998	0.00002	0.00060	2.00836	0.00836	5.99295

TABLE 6. The simulated means, Biases and $n \times \text{MSEs}$ for the ML and MM estimators of the parameters a and λ , when $\lambda = 1$

a	n	Method	\hat{a}			$\hat{\lambda}$		
			Mean	Bias	$n \times \text{MSE}$	Mean	Bias	$n \times \text{MSE}$
0.9	30	ML	0.90009	0.00009	0.00277	2.00069	0.00069	3.82511
		MM	0.90010	0.00010	0.00437	2.03149	0.03149	5.82083
	50	ML	0.90005	0.00005	0.00101	2.00240	0.00240	3.98809
		MM	0.89998	0.00002	0.00160	2.01773	0.01773	5.97136
	100	ML	0.89999	0.00001	0.00025	1.99803	0.00197	3.94618
		MM	0.90001	0.00001	0.00040	2.00993	0.00993	5.93718
0.95	30	ML	0.94997	0.00003	0.00324	1.99879	0.00121	3.91719
		MM	0.94997	0.00003	0.00503	2.02922	0.02922	5.79434
	50	ML	0.94996	0.00004	0.00112	1.99900	0.00100	3.93871
		MM	0.94997	0.00003	0.00179	2.01798	0.01798	5.92703
	100	ML	0.94999	0.00001	0.00027	1.99956	0.00044	3.91735
		MM	0.94997	0.00003	0.00044	2.00791	0.00791	5.95418
1.05	30	ML	1.05019	0.00019	0.00391	2.00269	0.00269	3.87857
		MM	1.05021	0.00021	0.00596	2.03344	0.03344	5.80297
	50	ML	1.05005	0.00005	0.00139	1.99822	0.00178	4.02246
		MM	1.05012	0.00012	0.00219	2.01988	0.01988	5.94967
	100	ML	1.04999	0.00001	0.00034	1.99839	0.00161	3.99821
		MM	1.04999	0.00001	0.00054	2.00857	0.00857	5.92792
1.1	30	ML	1.10026	0.00026	0.00426	2.00355	0.00355	3.91609
		MM	1.10020	0.00020	0.00661	2.03346	0.03346	5.93627
	50	ML	1.10001	0.00001	0.00147	1.99938	0.00062	3.87188
		MM	1.10000	0.00000	0.00240	2.01862	0.01862	5.91393
	100	ML	1.10002	0.00002	0.00037	2.00168	0.00168	3.99584
		MM	1.10000	0.00000	0.00060	2.01039	0.01039	6.00551

TABLE 7. The simulated variances of the ML estimators and the corresponding MVB values.

n	Simulated variances		MVB	
	a	λ	a	λ
30	1.42603E-04	0.1413	1.34444E-04	0.1333
50	2.99881E-05	0.0819	2.90400E-05	0.0800
100	3.77741E-06	0.0406	3.63000E-06	0.0400

MVB values become close as n increases. It is clear to say that the ML estimators are highly efficient estimators.

5. APPLICATION

In this section, in order to illustrate the data analysis, a real data set is analysed by using the ML and MM estimators. This data set is about the coal mining disaster.

Coal mining disaster data

The coal-mining disaster data set has 190 observations showing that the intervals in days between successive disasters in Great Britain [1]. To test whether the data set $\{X_1, X_2, \dots, X_n\}$ is consistent with the Rayleigh distribution, let's write $Y_i = a^{i-1}X_i, i = 1, 2, \dots, n$. By taking the logarithm of Y_i , we obtain $\ln Y_i = (i - 1) \ln a + \ln X_i, i = 1, 2, \dots, n$. It is known that $\ln Y_i$'s are iid random variables with extreme value distribution $EV(\delta, \eta)$ by the pdf $f(x) = \frac{1}{\eta} \exp(\frac{x-\delta}{\eta}) \exp(-\exp(\frac{x-\delta}{\eta})), x \in \mathbb{R}; \eta > 0, \delta \in \mathbb{R}$, where $\delta = \ln(\sqrt{2}\lambda)$ and $\eta = 0.5$. Then, a simple linear regression model is given by $\ln X_i = \mu - (i - 1) \ln a + \varepsilon_i, i = 1, 2, \dots, n$ where $\mu = E(\ln Y_i)$ and $\varepsilon_i \sim EV(\delta, 0.5)$. For this data set, it is obtained $\varepsilon_i \sim EV(0.2886, 0.5)$, where $\hat{\varepsilon}_i = \ln x_i - \hat{\mu} + (i - 1) \ln \hat{a}_{MM}, i = 1, 2, \dots, n$ and $\hat{\mu} = \frac{2}{n(n+1)} \sum_{i=1}^n (2n - 3i + 2) \ln x_i$.

Thus, to obtain an idea whether the underlying distribution of data set is the Rayleigh, a Q-Q plot can be constructed by plotting the ordered residuals $\hat{\varepsilon}_i$ against the quantiles of the $EV(0.2886, 0.5)$ distribution, see Figure 1.

It is clear from Figure 1 that the data points fall approximately on the straight line, thus it can be concluded that the Rayleigh is an appropriate distribution for the coal mining disaster data. This is also supported by the Z^* test statistic proposed by Tiku [18] ($Z^*=1.0049$ and $p\text{-value}=0.8938$). Moreover, for this data set, the value of statistic U given in Equation 3.13 and respective p -value are calculated as $U = -12.8417$ and $p\text{-value} = 9.5745e - 038$, respectively. According to result of this test, the data follow a GP with $a \neq 1$. This data was also studied by Lam et. al (2004) who showed that the data come from a GP and the ratio parameter a is less than 1. Thus, we can say that the data set can be modeled by a GP with the Rayleigh distribution.

The estimates of the parameters a and λ when the coal mining disaster data set is modeled by a GP with Rayleigh distribution are given in Table 8. Values given in parantheses in Table 8 are the standart errors (SE) of the estimators.

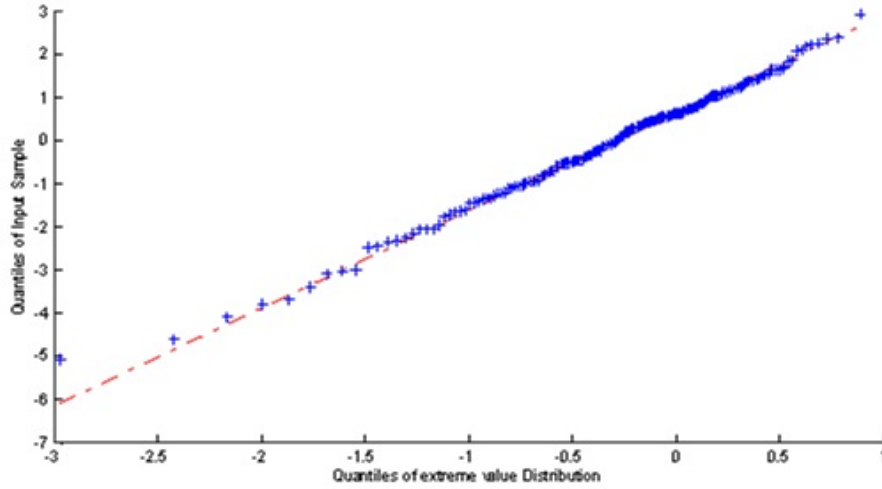


FIGURE 1. EV Q-Q plot of the coal mining data.

TABLE 8. Estimation of parameters for the coal mining disaster data

Method	\hat{a}	$\hat{\lambda}$
ML	0.9916 (6.5425x10-5)	91.2931 (4.6812)
MM	0.9909 (0.0018)	62.2422 (12.1883)

6. CONCLUSION

In this paper, we consider the parameter estimation problem in the GP by assuming that distribution of the first occurrence time is Rayleigh with the scale parameter λ . ML estimators for both the ratio parameter a of GP and scale parameter λ of Rayleigh distribution are also obtained and it is proved that these estimators are asymptotically normal distributed and consistent estimators. In addition, the ML estimators are compared to MM estimators with a simulation study which evaluates the means, biases and $n \times \text{MSE}$ for estimators. According to simulated results, ML estimators are more efficient than MM estimators and they have smaller $n \times \text{MSE}$ values.

7. APPENDIX. THE DERIVATION OF I^{-1}

The second derivatives of the logarithmic likelihood function given in Equation (3.2) are

$$\begin{aligned}\frac{\partial^2 \ln L}{\partial a^2} &= -\frac{n(n-1)}{a^2} - \frac{1}{a^2 \lambda^2} \sum_{i=1}^n (a^{i-1} x_i)^2 (i-1)(2i-3) \\ \frac{\partial^2 \ln L}{\partial \lambda^2} &= \frac{2n}{\lambda^2} - \frac{3}{\lambda^4} \sum_{i=1}^n (a^{i-1} x_i)^2 \\ \frac{\partial^2 \ln L}{\partial \lambda \partial a} &= \frac{2}{a \lambda^3} \sum_{i=1}^n (a^{i-1} x_i)^2 (i-1)\end{aligned}$$

Furthermore, since $E(a^{i-1} X_i) = \lambda \sqrt{\frac{\pi}{2}}$ and $E[(a^{i-1} X_i)^2] = 2\lambda^2$ the expected values of the second derivatives are obtained as

$$\begin{aligned}E\left(-\frac{\partial^2 \ln L}{\partial a^2}\right) &= \frac{n(n-1)}{a^2} + \frac{1}{a^2 \lambda^2} \sum_{i=1}^n E[(a^{i-1} X_i)^2] (i-1)(2i-3) \\ &= \frac{n(n-1)}{a^2} + \frac{1}{a^2 \lambda^2} \sum_{i=1}^n 2\lambda^2 (i-1)(2i-3) \\ &= \frac{1}{a^2} \left(\frac{4}{3} n^3 - 3n^2 + \frac{5}{3} n \right) - \frac{1}{a^2} (n^2 - n) \approx \frac{4}{3a^2} n^3\end{aligned}$$

$$\begin{aligned}E\left(-\frac{\partial^2 \ln L}{\partial \lambda^2}\right) &= -\frac{2n}{\lambda^2} + \frac{3}{\lambda^4} \sum_{i=1}^n E[(a^{i-1} X_i)^2] \\ &= -\frac{2n}{\lambda^2} + \frac{3}{\lambda^4} \sum_{i=1}^n 2\lambda^2 \\ &= \frac{4n}{\lambda^2}\end{aligned}$$

$$\begin{aligned}E\left(-\frac{\partial^2 \ln L}{\partial \lambda \partial a}\right) &= -\frac{2}{a \lambda^3} \sum_{i=1}^n E[(a^{i-1} X_i)^2] (i-1) \\ &= -\frac{2}{a \lambda^3} \sum_{i=1}^n 2\lambda^2 (i-1) \\ &= -\frac{2}{a \lambda^3} (n^2 \lambda^2 - n \lambda^2) \approx -\frac{2n^2}{a \lambda}\end{aligned}$$

where the symbol \approx stands for ‘asymptotically equivalent’. These are the components of the Fisher information matrix I and its inverse is given Equation (3.10).

REFERENCES

- [1] Andrews D. F., Herzberg. A. M., Data. New York: Springer, 1985.
 - [2] Ascher H, Feingold H., Repairable systems reliability, New York: Marcel Dekker; 1984.
 - [3] Aydoğdu H., Şenoğlu B., Kara M., Parameter estimation in geometric process with Weibull distribution, Appl Math Comput. (2010), 217, 2657–2665.
 - [4] Barndorff-Nielsen O. E., Cox D. R., Inference and asymptotics, London: Chapman & Hall, 1994.
 - [5] Braun W. J. , Li W., Zhao Y. P., Properties of the geometric and related processes, Nav Res Log., (2005), 52, 607–616.
 - [6] Chan S. K., Lam Y., Leung Y.P., Statistical inference for geometric processes with gamma distribution, Comput. Stat. Data Anal. (2004), 47, 565–581..
 - [7] Cox D.R., Lewis P.A.W., The statistical analysis of series of events, London: Mathuen, 1966.
 - [8] Forbes C., Evans M., Hastings N., Peacock B., Statistical distributions, New Jersey: John Wiley & Sons; 2011.
 - [9] Kara, M., Aydoğdu, H., & Türkşen, Ö. Statistical inference for geometric process with the inverse Gaussian distribution, Journal of Statistical Computation and Simulation, (2015), 85(16), 3206-3215.
 - [10] Kececioglu, Dimitri. Reliability engineering handbook, Prentice-Hall Inc., 1991.
 - [11] Lam Y., A note on the optimal replacement problem, Adv Appl Probab. (1988), 20, 479–482.
 - [12] Lam Y., Geometric process and replacement problem, Acta Math Appl Sin. (1988), 366–377.
 - [13] Lam Y., Nonparametric inference for geometric processes. Commun Stat Theor M. (1992), 21, 2083–2105.
 - [14] Lam Y, Chan S. K., Statistical inference for geometric processes with lognormal distribution, Comput Stat Data Anal. (1998), 27, 99–112.
 - [15] Lam Y, Zheng Y. H, Zhang Y. L., Some limit theorems in geometric process, Acta Math Appl Sin. (2003),19(3), 405– 416.
 - [16] Lam Y, Zhu L.X., Chan JSK, Liu Q. Analysis of data from a series of events by a geometric process model, Acta Math Appl Sin. (2004),20(2), 263–282.
 - [17] Lam Y., The geometric process and its applications, Singapore: World Scientific; 2007.
 - [18] Tiku M. L., Goodness-of-fit statistics based on the spacings of complete or censored samples, Austral. J. Statist. (1980) 22, 260–275.
- Current address:* Cenker Biçer: Department of Statistics, Faculty of Science and Arts, Kirikkale University, Kirikkale, Turkey
E-mail address: cbicer@kku.edu.tr
 ORCID Address: <https://orcid.org/0000-0003-2222-3208>
- Current address:* Hayrinisa Demirci Biçer: Department of Statistics, Faculty of Science and Arts, Kirikkale University, Kirikkale, Turkey
E-mail address: hdbicer@hotmail.com
 ORCID Address: <https://orcid.org/0000-0002-1520-5004>
- Current address:* Mahmut Kara: Department of Statistics, Faculty of Science and Arts, Yüzüncü Yıl University, Van, Turkey
E-mail address: mkara2581@gmail.com
 ORCID Address: <https://orcid.org/0000-0001-7678-8824>
- Current address:* Halil Aydoğdu: Department of Statistics, Faculty of Science , Ankara University, Ankara Turkey
E-mail address: aydogdu@ankara.edu.tr
 ORCID Address: <https://orcid.org/0000-0001-5337-5277>