Sigma J Eng & Nat Sci 37 (3), 2019, 871-882

Sigma Journal of Engineering and Natural Sciences
Sigma Mühendislik ve Fen Bilimleri Dergisi

Research Article

STATISTICAL INFERENCE FOR GEOMETRIC PROCESS WITH THE INVERSE RAYLEIGH DISTRIBUTION

İlhan USTA*[1]

[1]*Eskişehir Technical University, Department of Statistics, ESKISEHIR;* ORCID: 0000-0001-5576-2027

Received: 13.03.2019 Revised: 17.04.2019 Accepted: 03.05.2019

ABSTRACT

This paper deals with the statistical inference for the geometric process (GP), in which the time until the occurrence of the first event is assumed to follow inverse Rayleigh distribution. The maximum likelihood (ML) method is used to derive the estimators of the parameters in GP. Asymptotic distributions of the ML estimators are obtained which help us to construct confidence intervals for the parameters and show the consistency of these estimators. The performances of the ML estimators are also compared with the corresponding non-parametric modified moment estimators in terms of bias, mean squared error and Pitman nearness probability through an extensive simulation study. Finally, a real data set is provided to illustrate the results.

Keywords: Geometric process, inverse Rayleigh distribution, maximum likelihood estimator, modified moment estimator, asymptotic normality.

1. INTRODUCTION

In statistical analysis of a data set which consists of the successive inter-arrival time from a series of events, a counting process is the most commonly used method for analyzing data from this type of study. If the data set is independent and identically distributed and has no trend, a renewal process (RP) or homogeneous Poisson process can be used for modeling the data. If the data set exhibits a trend, an inhomogeneous Poisson process can be applied as a possible approach for modeling this trend (see [1, 2]). However, a more direct approach to model the data with a monotone trend is to apply a stochastically monotone process. The geometric process (GP) as a simple monotone process was first introduced by [3, 4] and defined as follows.

Definition 1. Let X_i denote the inter-arrival time between the $(i-1)$ th and i th events of a counting process $\{N(t), t \geq 0\}$ for $i = 1, 2, \ldots$. A stochastic process $\{X_i, i = 1, 2, \ldots\}$ is said to be a GP, if there exists a real $a > 0$ such that the random variables $Y_i = a^{i-1} X_i$, $i = 1, 2, \ldots$ are independent and identically distributed (iid) with the distribution function F. The number a is called the ratio parameter of the GP.

* Corresponding Author: e-mail: iusta@eskisehir.edu.tr, tel: (222) 335 05 80 / 4684

It can be easil y seen that a GP with ratio a is stochastically decreasing if $a > 1$, and stochastically increasing if $0 < a < 1$. When $a = 1$, the GP will be reduced to a RP [4]. Furthermore, given a GP with ratio a, then it can easily be shown that

$$E(X_i) = \frac{\mu}{a^{i-1}} \text{ and } Var(X_i) = \frac{\sigma^2}{a^{2i-2}} \tag{1}$$

where μ and σ^2 are the expectation and variance of the first inter-arrival time X_1, respectively.

The GP has been applied to the maintenance problem by [4–7]. The basic properties of GP has been studied by several authors, for example: [7–9]. Additionally, the statistical inference results for the GP have been recently presented by assuming that the random variable X_1 follows specific distributions such as lognormal distribution [10], gamma distribution [11], Weibull distribution [12], the inverse Gaussian distribution [13], Power Lindley distribution [14] and Rayleigh distribution [15].

In this study, we assume that the distribution of the first inter-arrival time X_1 in GP follows an inverse Rayleigh (IR) distribution with parameter θ. Therefore; the main purpose of this study is to consider the problem of statistical inference for GP with the IR distribution for X_1, and correspondingly, to obtain the estimators of the parameters a and θ by using the parametric estimation method including maximum likelihood method.

The remainder of this paper is organized as follows: The IR distribution is briefly introduced in Section 2. In the next section, the maximum likelihood (ML) estimators are derived for the unknown parameters along with their asymptotic confidence intervals. In Section 4, the modified moment (MM) estimators are given by using the nonparametric method proposed by [5]. Section 5 presents the results of a Monte Carlo simulation study that compare the performance of the derived ML estimators with the corresponding MM estimators. A real data set is analyzed in Section 6, followed by conclusions in Section 7.

2. INVERSE RAYLEIGH DISTRIBUTION

Let the distribution of first occurrence time X_1 in GP has an IR distribution, which was originally proposed by [16], with the following probability density function (pdf) and the cumulative distribution function (cdf), respectively

$$f(x;\theta) = \frac{2\theta}{x^3}\exp\left(-\frac{\theta}{x^2}\right), \quad x > 0, \ \theta > 0, \tag{2}$$

$$F(x;\alpha) = \exp\left(-\frac{\theta}{x^2}\right), \quad x > 0, \tag{3}$$

where θ is the scale parameter. The mean of the IR distribution is given as

$$\mu = \sqrt{\theta\pi}. \tag{4}$$

The IR distribution has many applications in the area of reliability and the life testing studies including infant mortality, useful life and wear-out periods since the hazard rate function of the IR distribution can be increasing or decreasing depending on θ [17–19]. The IR distribution also has a unimodal pdf and is a member of the exponential family. For further details about the IR distribution, the readers can refer to [19].

3. MAXIMUM LIKELI HOOD ESTIMATION OF PARAMETERS OF GP WITH INVERSE RAYLEIGH DISTRIBUTION

In this section, the ML estimation method is used to obtain the estimators of the parameters of GP, where a and θ are the ratio parameter of GP and the scale parameter of IR distribution, respectively. Besides, the approximate confidence intervals of the parameters based on the asymptotic distributions of the ML estimators are derived.

Let $X=(X_1,X_2,...,X_n)$ denote a random sample of size n drawn from the GP with ratio a and X_1 which follows the IR distribution with the pdf given in Equation (2). For $X_i,\ i=1,2,\cdots$, the likelihood function can be written as

$$L(a,\theta;x)=\prod_{i=1}^{n} f(x_i;a,\theta)=\left(2\theta\right)^n a^{-n(n-1)}\prod_{i=1}^{n} x_i^{-3}\exp\left(-\theta\sum_{i=1}^{n} a^{2-2i}x_i^{-2}\right) \tag{5}$$

The log-likelihood can be obtained by taking the natural logarithm of Equation (5) as

$$l(a,\theta)=\ln L(a,\theta;x)=n\ln 2+n\ln\theta-n(n-1)\ln a-3\sum_{i=1}^{n}\ln x_i-\theta\sum_{i=1}^{n} a^{2-2i}x_i^{-2} \tag{6}$$

After taking the first partial derivatives of $l(a,\theta)$ with respect to a and θ, and equating them to zero, the likelihood equations for the parameters a and θ are obtained as follows

$$\frac{\partial l(a,\theta)}{\partial a}=-\frac{n(n-1)}{a}-\theta\sum_{i=1}^{n}(2-2i)a^{1-2i}x_i^{-2}=0\,, \tag{7}$$

$$\frac{\partial l(a,\theta)}{\partial \theta}=\frac{n}{\theta}-\sum_{i=1}^{n} a^{2-2i}x_i^{-2}=0\,. \tag{8}$$

Then, by solving Equations (7)–(8), the parameter θ is obtained as

$$\theta=\frac{n}{\sum_{i=1}^{n} a^{2-2i}x_i^{-2}}\,. \tag{9}$$

Substituting the Equation (9) into Equation (8), the ML estimator of the parameter a ($\hat{a}_{MLE}$) is obtained by solving the following equation

$$2\sum_{i=1}^{n} i a^{-2i}x_i^{-2}=(n+1)\sum_{i=1}^{n} a^{-2i}x_i^{-2}\,. \tag{10}$$

It is obvious that an explicit form of the solution of Equation (10) does not exist, and hence numerical methods, such as the Newton–Raphson, are required to compute the estimator $\hat{a}_{MLE}$. Then, the ML estimator of the parameter θ ($\hat{\theta}_{MLE}$) is obtained using the numerical solution of $\hat{a}_{MLE}$ as follows

$$\hat{\theta}_{MLE}=\frac{n}{\sum_{i=1}^{n}\hat{a}_{MLE}^{2-2i}X_i^{-2}}\,. \tag{11}$$

Let $\alpha = (a, \theta)^T$ be the unknown parameter vector. Then, according to the large-sample theory, the asymptotic distribution of the ML estimator of α, denoted as $\hat{\alpha}_{MLE} = \left(\hat{a}_{MLE}, \hat{\theta}_{MLE}\right)^T$, is

$$(\hat{\alpha}_{MLE} - \alpha) \to N(0, I^{-1}(\alpha)), \tag{12}$$

where "$\to$" denotes approximately distributed when n is large [20]. Here, $I^{-1}(\alpha)$ is the inverse of the Fisher information matrix $I(\alpha)$ which is defined by

$$I(\alpha) = E\begin{bmatrix} -\dfrac{\partial^2 l(a,\theta)}{\partial a^2} & -\dfrac{\partial^2 l(a,\theta)}{\partial a \partial \theta} \\ -\dfrac{\partial^2 l(a,\theta)}{\partial \theta \partial a} & -\dfrac{\partial^2 l(a,\theta)}{\partial \theta^2} \end{bmatrix}, \tag{13}$$

where

$$-\frac{\partial^2 l(a,\theta)}{\partial a^2} = -\frac{n(n-1)}{a^2} + \theta \sum_{i=1}^{n} (2-2i)(1-2i) a^{-2i} x_i^{-2}, \tag{14}$$

$$-\frac{\partial^2 l(a,\theta)}{\partial a \partial \theta} = -\frac{\partial l(a,\theta)}{\partial \theta \partial a} = \sum_{i=1}^{n} (2-2i) a^{1-2i} x_i^{-2}, \tag{15}$$

$$-\frac{\partial^2 l(a,\theta)}{\partial \theta^2} = \frac{n}{\theta^2}. \tag{16}$$

Let $Y_i = a^{2-2i} X_i^{-2}$, $i = 1, 2, ..., n$, which can be written as $X_i = a^{1-i} Y_i^{-1/2}$, then we have

$$g(y_i) = f(x = a^{1-i} y_i^{-1/2}) \left| \frac{dx_i}{dw_i} \right| = \theta \exp(-\theta y_i), \ y_i > 0.$$

Therefore, each Y_i, $i = 1, 2, ..., n$ has an exponential distribution with the parameter θ and $E(W_i) = 1/\theta$, $i = 1, 2, ..., n$. By using this result, the expected values of the second derivatives given in Equations (14)-(16) are obtained as

$$\begin{aligned} E\left(-\frac{\partial^2 l(a,\theta)}{\partial a^2}\right) &= -\frac{n(n-1)}{a^2} + \theta \sum_{i=1}^{n} (2-2i)(1-2i) E\left(a^{-2i} X_i^{-2}\right) \\ &= -\frac{n(n-1)}{a^2} + \theta \sum_{i=1}^{n} (2-2i)(1-2i) \frac{1}{a^2 \theta} \\ &= -\frac{n(n-1)}{a^2} + \frac{4n^3 - 3n^2 - n}{3a^2} \approx \frac{4n^3}{3a^2}, \end{aligned} \tag{17}$$

$$\begin{aligned} E\left(-\frac{\partial l(a,\theta)}{\partial \theta \partial a}\right) = E\left(-\frac{\partial l(a,\theta)}{\partial a \partial \theta}\right) &= \sum_{i=1}^{n} (2-2i) E\left(a^{1-2i} X_i^{-2}\right) \\ &= \sum_{i=1}^{n} (2-2i) \frac{1}{a\theta} \\ &= \frac{-n^2 + n}{a\theta} \approx -\frac{n^2}{a\theta}, \end{aligned} \tag{18}$$

$$E\left(-\frac{\partial^2 l(a,\theta)}{\partial\theta^2}\right)=E\left(\frac{n}{\theta^2}\right)=\frac{n}{\theta^2}\,. \tag{19}$$

Substituting Equations (17)-(19) into Equation (13), the inverse of the Fisher information matrix, $I^{-1}(\alpha)$, is derived as follows

$$I^{-1}(\alpha)=\begin{bmatrix}\frac{4n^3}{3a^2} & -\frac{n^2}{a\theta}\\ -\frac{n^2}{a\theta} & \frac{n}{\theta^2}\end{bmatrix}^{-1}=\begin{bmatrix}\frac{3a^2}{n^3} & \frac{3a\theta}{n^2}\\ \frac{3a\theta}{n^2} & \frac{4\theta^2}{n}\end{bmatrix}. \tag{20}$$

By using the asymptotic normality of $\hat{\alpha}_{MLE}$, and the estimated variances, thus, the approximate $(1-\gamma)100\%$ confidence intervals for a and θ are derived, respectively, as

$$\hat{a}_{MLE}\pm z_{\gamma/2}\sqrt{\frac{3\left(\hat{a}_{MLE}\right)^2}{n^3}}\ \text{ and }\ \hat{\theta}_{MLE}\pm z_{\gamma/2}\sqrt{\frac{4\left(\hat{\theta}_{MLE}\right)^2}{n}}\,, \tag{21}$$

where $z_{\gamma/2}$ is the percentile of the standard normal distribution with right-tail $\gamma/2$.

Furthermore, it is to be noted that both the estimators $\hat{a}_{MLE}$ and $\hat{\theta}_{MLE}$ are asymptotically unbiased and consistent, since the asymptotic variance of each of $\hat{a}_{MLE}$ and $\hat{\theta}_{MLE}$ converges to zero as n goes to infinity.

4. MODIFIED MOMENT ESTIMATION OF PARAMETERS OF GP WITH INVERSE RAYLEIGH DISTRIBUTION

In this section, the MM estimator of the parameter θ, $\hat{\theta}_{MME}$, is obtained by using the nonparametric estimate of the ratio parameter a, a commonly used method for GP. Lam [5] derived a nonparametric estimator for the ratio parameter a as the following

$$\hat{a}_{NP}=\exp\left(\frac{6}{(n-1)n(n+1)}\sum_{i=1}^{n}(n-2i+1)\ln X_i\right)\,. \tag{22}$$

The estimator $\hat{a}_{NP}$ is both an unbiased and a consistent estimator of the ratio parameter a. It is also asymptotically normally distributed. Furthermore, Lam [5] proposed a nonparametric estimator for Y_i, $i=1,2,\cdots$ based on $\hat{a}_{NP}$ as

$$\hat{Y}_i=\hat{a}_{NP}^{i-1}X_i\,,\ i=1,2,\cdots \tag{23}$$

Thus, for the sample $X_1,X_2,...,X_n$ drawn from the GP with ratio a and X_1 which follows the IR distribution, the first sample moment, say m_1, is computed as follows

$$m_1=\frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i=\frac{1}{n}\sum_{i=1}^{n}\hat{a}_{NP}^{i-1}X_i\,. \tag{24}$$

In addition to these, the first population moment of the IR distribution given in Equation (4) can be expressed as

$$\mu_1=\sqrt{\theta\pi}\,. \tag{25}$$

Then, equating the first population moment given by Equation (25) with the corresponding sample moment given in Equation (24), the MM estimator of the parameter θ , $\hat{\theta}_{MME}$, is obtained as follows

$$\hat{\theta}_{MME} = \frac{1}{\pi n^2}\left(\sum_{i=1}^{n}\hat{a}_{NP}^{i-1}X_i\right)^2 . \tag{26}$$

5. SIMULATION STUDY

In this section, a Monte Carlo simulation study is conducted to compare the performances of the ML and the MM estimators for the parameters a and θ . The simulation is carried out for different sample sizes $n = 25, 50, 100$. For each case, the parameter θ is taken as $0.5, 1, 2$ and also the ratio parameter a is chosen as $0.9, 0.95, 1.05$ and 1.1. For 5000 repetitions, the performances of the ML and MM estimators are measured with different criteria such as bias, mean square error (MSE) and Pitman nearness (PN) probability (see [21, 22]), given as follows:

$$Bias = \frac{1}{5000}\sum_{i=1}^{5000}(\hat{\alpha}_i - \alpha) , \tag{27}$$

$$MSE = \frac{1}{5000}\sum_{i=1}^{5000}(\hat{\alpha}_i - \alpha)^2 , \tag{28}$$

$$PN = P\left(\left|\hat{\alpha}_i - \alpha\right| < \left|\hat{\alpha}_j - \alpha\right|\right) = \frac{1}{5000}\#\left\{\left|\hat{\alpha}_i - \alpha\right| < \left|\hat{\alpha}_j - \alpha\right|\right\} , \tag{29}$$

where $\hat{\alpha}_i (= \hat{a}_i, \hat{\theta}_i)$ is the estimate of $\alpha (= a, \theta)$ for the i^{th} simulated sample. It can be said that if $PN > 50$ (%) , $\hat{\alpha}_i$ outperforms $\hat{\alpha}_j$ with regard to PN criteria.

The simulation results are reported in Tables 1–3, show the mean, bias and MSE values for the ML and MM estimators of a and θ . Tables 1–3 also exhibit the PN probability values (as percentages) of the considered estimators relative to each other.

Table 1. The Means, Biases, MSEs and PN probability values (%) of the estimators of a and θ when $\theta = 0.5$

			$\hat{a}$				$\hat{\theta}$			
$\boldsymbol{a}$	$\boldsymbol{n}$	**Method**	**Mean**	**Bias**	**MSE**	**PN**	**Mean**	**Bias**	**MSE**	**PN**
0.9	25	MME	0.89930	-0.000699	0.000253	38.1	0.56685	0.06685	0.41882	34.2
		MLE	0.89951	-0.000494	0.000166	61.9	0.55521	0.05521	0.05845	65.8
	50	MME	0.89976	-0.000241	0.000034	36.5	0.54083	0.04083	0.15608	31.7
		MLE	0.89977	-0.000230	0.000020	63.5	0.52659	0.02659	0.02413	68.3
	100	MME	0.90008	0.000084	0.000004	37.7	0.52771	0.02771	0.14864	31.6
		MLE	0.90009	0.000086	0.000002	62.3	0.51829	0.01829	0.01082	68.4
0.95	25	MME	0.95005	0.000053	0.000266	39.6	0.57451	0.07451	0.50296	32.3
		MLE	0.94993	-0.000069	0.000177	60.4	0.56046	0.06046	0.06259	67.7
	50	MME	0.95005	0.000046	0.000035	39	0.54995	0.04995	0.39360	30.1
		MLE	0.94995	-0.000050	0.000022	61	0.52962	0.02962	0.02425	69.9
	100	MME	0.94994	-0.000058	0.000004	39.1	0.52191	0.02191	0.06525	31.2
		MLE	0.94995	-0.000054	0.000003	60.9	0.51364	0.01364	0.01074	68.8
1.05	25	MME	1.05055	0.000547	0.000341	38.3	0.58226	0.08226	0.60377	34.2
		MLE	1.05098	0.000976	0.000221	61.7	0.57871	0.07871	0.06810	65.8
	50	MME	1.05017	0.000168	0.000042	37.5	0.52858	0.02858	0.10166	34.3
		MLE	1.05015	0.000152	0.000028	62.5	0.53764	0.03764	0.02910	65.7
	100	MME	1.05001	0.000007	0.000005	37.5	0.51258	0.01258	0.04412	31.8
		MLE	1.05005	0.000055	0.000003	62.5	0.51707	0.01707	0.01017	68.2
1.1	25	MME	1.10115	0.001153	0.000392	37.2	0.56654	0.06654	0.31695	33.7
		MLE	1.10114	0.001138	0.000231	62.8	0.57251	0.07251	0.06648	66.3
	50	MME	1.09969	-0.000312	0.000045	40.7	0.53768	0.03768	0.16650	30.9
		MLE	1.09966	-0.000339	0.000029	59.3	0.52055	0.02055	0.02260	69.1
	100	MME	1.10005	0.000050	0.000006	36.3	0.53445	0.03445	0.14462	32.5
		MLE	1.10003	0.000031	0.000004	63.7	0.51818	0.01818	0.01126	67.5

Table 2. The Means, Biases, MSEs and PN probability values (%) of the estimators of a and θ when $\theta = 1$

			$\hat{a}$				$\hat{\theta}$			
$\boldsymbol{a}$	$\boldsymbol{n}$	**Method**	**Mean**	**Bias**	**MSE**	**PN**	**Mean**	**Bias**	**MSE**	**PN**
0.9	25	MME	0.90007	0.000188	0.000244	40.5	1.06817	0.06817	0.50546	33.4
		MLE	0.90012	0.000119	0.000169	59.5	1.13926	0.13926	0.28231	66.6
	50	MME	0.90019	0.000074	0.000031	36.8	1.03470	0.03470	0.29616	32.6
		MLE	0.90008	0.000081	0.000019	63.2	1.05209	0.05209	0.09832	67.4
	100	MME	0.89997	-0.000031	0.000004	39.2	1.02428	0.02428	0.16582	32.7
		MLE	0.90002	0.000018	0.000003	60.8	1.03465	0.03465	0.05092	67.3
0.95	25	MME	0.95034	0.000344	0.000293	39.8	1.07654	0.07654	0.53803	31.2
		MLE	0.94996	-0.000037	0.000194	60.2	1.12849	0.12849	0.26840	68.8
	50	MME	0.95025	0.000248	0.000033	42.5	1.05607	0.05607	0.34766	33
		MLE	0.95018	0.000078	0.000024	57.5	1.08136	0.08136	0.12356	67
	100	MME	0.95004	0.000043	0.000005	37.6	1.02694	0.02694	0.18809	31.1
		MLE	0.95005	0.000049	0.000003	62.4	1.03508	0.03508	0.04650	68.9
1.05	25	MME	1.05014	0.000239	0.000339	37.5	1.03673	0.03673	0.49214	34.3
		MLE	1.04972	-0.000276	0.000216	62.5	1.11628	0.11628	0.24426	65.7
	50	MME	1.05024	0.000139	0.000044	39.3	1.03458	0.03458	0.29624	35.2
		MLE	1.05027	0.000272	0.000029	60.7	1.07438	0.07438	0.10682	64.8
	100	MME	1.04994	-0.000061	0.000005	41.2	1.00889	0.00889	0.15609	31
		MLE	1.04998	-0.000016	0.000003	58.8	1.03265	0.03265	0.04454	69
1.1	25	MME	1.09922	-0.000779	0.000396	39.5	1.03600	0.03600	0.47973	34.3
		MLE	1.09959	-0.000423	0.000277	60.5	1.11383	0.11383	0.26804	65.7
	50	MME	1.09954	-0.000465	0.000048	36.3	1.01263	0.01263	0.23140	33.9
		MLE	1.09958	-0.000410	0.000031	63.7	1.04994	0.04994	0.10501	66.1
	100	MME	1.10003	0.000026	0.000006	39.8	1.01557	0.01157	0.16768	28.1
		MLE	1.09999	-0.000014	0.000004	60.2	1.02811	0.02811	0.04590	71.9

Table 3. The Means, Biases, MSEs and PN probability values (%) of the estimators of a and θ when $\theta = 1.5$

			$\hat{a}$				$\hat{\theta}$			
a	n	**Method**	**Mean**	**Bias**	**MSE**	**PN**	**Mean**	**Bias**	**MSE**	**PN**
0.9	25	MME	0.89941	-0.000588	0.000259	38.8	1.55692	0.05692	1.11640	34.7
		MLE	0.89955	-0.000452	0.000174	61.2	1.67207	0.17207	0.55901	65.3
	50	MME	0.89981	-0.000187	0.000032	37.9	1.55547	0.05547	0.67215	30.7
		MLE	0.89976	-0.000242	0.000020	62.1	1.56610	0.06610	0.20384	69.3
	100	MME	0.90003	0.000027	0.000004	38.4	1.53229	0.03229	0.39064	29.2
		MLE	0.90002	0.000024	0.000003	61.6	1.54648	0.04648	0.10606	70.8
0.95	25	MME	0.94916	-0.000842	0.000261	38.7	1.54171	0.07215	0.96315	34.2
		MLE	0.94961	-0.000387	0.000176	61.3	1.67757	0.17757	0.53421	65.8
	50	MME	0.95001	-0.000092	0.000037	38.4	1.57215	0.04171	0.74136	33.4
		MLE	0.94981	-0.000189	0.000023	61.6	1.58312	0.08312	0.23115	66.6
	100	MME	0.94991	0.000009	0.000004	38.5	1.53622	0.03622	0.44856	31
		MLE	0.94991	-0.000087	0.000003	61.5	1.53408	0.03408	0.09914	69
1.05	25	MME	1.05058	0.000582	0.000289	39.4	1.58607	0.08607	1.04389	32.1
		MLE	1.05041	0.000406	0.000208	60.6	1.71391	0.21391	0.58859	67.9
	50	MME	1.05010	0.000098	0.000046	38.3	1.54697	0.04697	0.63773	30.7
		MLE	1.04991	-0.000113	0.000028	61.7	1.56918	0.06918	0.20557	69.3
	100	MME	1.04988	-0.000092	0.000006	36	1.55786	0.05786	0.52408	30.2
		MLE	1.04989	-0.000095	0.000003	64	1.53035	0.03035	0.09654	69.8
1.1	25	MME	1.10001	0.000115	0.000371	40.1	1.53964	0.09706	1.03460	33.5
		MLE	1.10055	0.000555	0.000256	59.9	1.69172	0.19172	0.52160	66.5
	50	MME	1.10011	0.000112	0.000049	38	1.59706	0.06145	0.89679	31.5
		MLE	1.09996	-0.000042	0.000031	62	1.59264	0.09264	0.24326	68.5
	100	MME	1.10012	0.000009	0.000006	38.3	1.56145	0.03964	0.40444	32
		MLE	1.10013	0.000127	0.000004	61.7	1.56092	0.06092	0.10523	68

Some of the points are quite clear from Tables 1-3. First of all, as the sample size n increases, both the bias and MSE values decrease for all estimators of a and θ. In fact, we were expecting to see this result because the estimators are asymptotically unbiased and consistent. It is also observed that the ML estimators of a and θ outperform the corresponding MM estimators in terms of MSE for all the cases. On the other hand, according to PN probability, the ML estimators show better performance than the MM estimators in all the cases.

6. APPLICATION TO A REAL DATA SET

In order to verify how the estimators considered in this study work in a real-life context, a real data set is investigated in this section. The dataset consists of 29 software error times which are taken from Xu [23]. It was also studied by [24].

The software data set is here analyzed from the point of view of the GP with ratio a and X_1 which follows the IR distribution. Therefore, we first investigate whether the underlying distribution of the data set is the IR distribution. In order to test whether the IR distribution is consistent with the data set, it follows from definition 1 that $Y_i = a^{i-1}X_i$, $i = 1, 2, \cdots$., where Y_i 's follow the IR distribution. Then, it can be expressed as $\ln Y_i = (i-1)\ln a + \ln X_i$, $i = 1, 2, \cdots$, by taking the natural logarithm of Y_i. Here, it is known that $\ln Y_i$'s are iid random variables with the log-IR distribution. Thus, a simple linear regression model can be written as

$$\ln X_i = \beta - (i-1)\ln a + \varepsilon_i, \quad i = 1, 2, \cdots, \tag{30}$$

where $\beta = E(\ln Y_i)$ and $\exp(\varepsilon_i)$ follows the IR distribution. If the distribution of the exponential error terms is IR, then it is said that the underlying distribution of the data set is IR. In this case, the error terms can be estimated as the following

$$\hat{\varepsilon}_i = \ln X_i - \hat{\beta} + (i-1)\ln \hat{a}_{NP}, \quad i = 1, 2, \cdots, \tag{31}$$

where $\hat{\beta} = \dfrac{2}{n(n+1)} \sum_{i=1}^{n} (2n - 3i + 2)\ln X_i$ and $\hat{a}_{NP}$ is given in (22).

In order to analyze whether the exponential residuals, $\exp(\hat{\varepsilon}_i)$, follow the IR distribution, a QQ-plot is created first. This plot is depicted in Figure 1. It is seen from Figure 1 that the data points do not diverge much from a straight line. Hence, it can be concluded that the underlying distribution of this software data is consistent with the IR distribution. This result is also supported by using the Kolmogorov–Smirnov test ($KS = 0.2248$ and the corresponding $p-$value $= 0.09$), a well-known goodness of fit test.

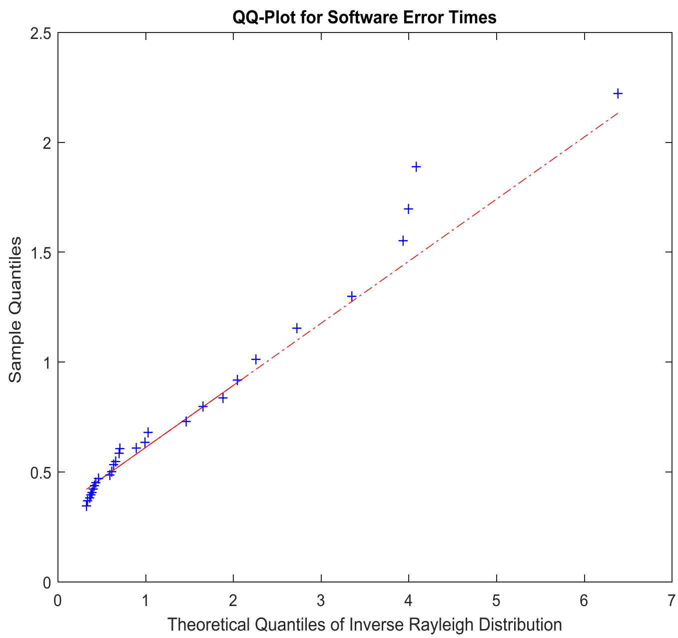

Figure 1. Q-Q plot for the software data set.

When the software data set is modeled by using a GP with the IR distribution, the estimates of the parameters a and θ obtained by using the ML and MM estimators are reported in Table 4. The standard errors of the estimators are given in parentheses in Table 4.

Table 4. Estimation of parameters for the software data set.

Method	$\hat{a}$	$\hat{\theta}$
MME	0.9654 (0.0198)	1.3772 (0.9663)
MLE	0,9672 (0.0107)	1,0212 (0.3793)

Based on the simulation results presented in Table 2, it can be concluded that for the software data set, the ML estimates of the parameters are preferable to the MM estimates in terms of the MSE and PN probability criteria, since the ML estimators show better performance than the MM estimators according to MSE and PN probability when the ratio parameter a is less than 1.

8. CONCLUSIONS

In this paper, we study the statistical inference problem for the geometric process (GP) when the distribution of the first occurrence time is assumed to be an inverse Rayleigh (IR). The maximum likelihood (ML) estimators are derived for the shape parameter θ of IR distribution and the ratio parameter a of GP along with their asymptotic confidence intervals. Additionally, the modified moment (MM) estimator for θ is obtained by using a nonparametric estimator of a , derived by Lam [5]. The performances of the ML estimators are compared with the corresponding MM estimators through an extensive simulation study. From the simulation results, it is observed that the ML estimators outperform the MM estimators in terms of the MSE and Pitman nearness probability criteria for all of the cases. Finally, a real data set is provided to illustrate the results.

REFERENCES

[1] Cox D.R., Lewis P.A.W., (1966) The Statistical Analysis of Series of Events. *Mathuen*, London.

[2] Ascher H., Feingold H., (1984) Repairable Systems Reliability. *Marcel Dekker*, New York.

[3] Lam Y., (1988) A Note on the optimal replacement problem, *Advances in Applied Probability* 20, 479–482.

[4] Lam Y., (1988) Geometric process and replacement problem, *Acta Mathematicae Applicatae Sinica* 4, 366–377.

[5] Lam Y., (1992) Nonparametric inference for geometric process, *Communications in Statistics – Theory and Methods* 21, 2083–2105.

[6] Lam Y., Zhu L.X., Chan J.S.K., Liu Q., (2004) Analysis of data from a series of events by a geometric process model, *Acta Mathematicae Applicatae Sinica* 20, 263–282.

[7] Lam Y., (2007) The Geometric Process and Its Applications. *World Scientific*, Singapore.

[8] Lam Y., Zheng Y.H., Zhang Y.L., (2003) Some limit theorems in geometric process, *Acta Mathematicae Applicatae Sinica* 19, 405–416.

[9] Braun W.J., Li W., Zhao Y.P., (2005) Properties of the geometric and related processes, *Naval Research Logistic* 52, 607–616.

[10] Lam Y., Chan S.K., (1998) Statistical inference for geometric processes with lognormal distribution, *Computational Statistics and Data Analysis* 27, 99–112.

[11] Chan S.K., Lam Y., Leung Y.P., (2004) Statistical inference for geometric processes with gamma distribution, *Computational Statistics and Data Analysis* 47, 565–581.

[12] Aydoğdu H., ¸Senoğlu B., Kara M., (2010) Parameter estimation in geometric process with Weibull distribution, *Applied Mathematics and Computation* 217, 2657–2665.

[13] Kara M., Aydoğdu H., Türkşen Ö. (2015) Statistical inference for geometric process with the inverse Gaussian distribution, *Journal of Statistical Computation and Simulation* 85, 3206–3215.

[14] Biçer C. (2018) Statistical inference for geometric process with the power Lindley distribution, *Entropy* 20, 723–743

[15] Biçer C., Biçer H.D., Kara M., Aydoğdu H. (2019) Statistical inference for geometric process with the Rayleigh distribution, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics* 68, 149–160.

[16] Trayer V.N., (1964) Doklady Acad. Nauk, Belorus, U.S.S.R.

[17] Voda, R. G., (1972). On the inverse Rayleigh variable, *Reports of Statistical Application Research of the Union of Japanese Scientists and Engineers*. 19, 15–21.

[18] Dey S. (2012) Bayesian estimation of the parameter and reliability function of an inverse Rayleigh distribution, *Malaysian Journal of Mathematical Sciences* 6,113–124.

[19] Maleki Jebely F., Zare K., Deiri E., (2018) Efficient estimation of the pdf and the cdf of the inverse Rayleigh distribution, *Journal of Statistical Computation and Simulation* 88, 75–88.

[20] Ferguson, T.S., (1996), A Course in Large-Sample Theory. *Chapman and Hall*, London.

[21] Usta. I., (2013). Different estimation methods for the parameters of the extended Burr XII distribution, *Journal of Applied Statistics* 40, 397–414.

[22] Muslehv, R.M., Helu, A., (2014) Estimation of the inverse Weibull distribution based on progressively censored data: Comparative study, *Reliability Engineering & System Safety* 131, 216–27.

[23] Xu, R. Z. (1985) An empirical investigation: software errors and their influence upon development of WPADT, *Proceedings Ninth International Computer Software & Applications Conferenc*e, 9–10 October 1985, Chicago, USA.

[24] Wang Z., Wang J., Liang X. (2007) Non-parametric estimation for NHPP software reliability models, *Journal of Applied Statistics* 34,107 –119.