Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 14(3), 1280-1290, 2024
Journal of the Institute of Science and Technology, 14(3), 1280-1290, 2024
ISSN: 2146-0574, eISSN: 2536-4618
Mathematics **DOI: 10.21597/jist.1478445**
Research Article
Received: 04.05.2024 **Accepted: 29.05.2024**

To Cite: Altındağ, Ö., Kara, M. & Aydoğdu, H. (2024). Estimation in α-Series Processes with Exponential Inter-Arrival Times under Censored Data. *Journal of the Institute of Science and Technology*, 14(3), 1280-1290.

Estimation in α-Series Processes with Exponential Inter-Arrival Times under Censored Data

Ömer ALTINDAĞ[1*], Mahmut KARA[2], Halil AYDOĞDU[3]

Highlights:
- Maximum likelihood estimation of model parameters under censored data
- Asymptotic properties of the estimators
- Efficiencies of the estimators

Keywords:
- α-series process
- Maximum likelihood method
- Exponential distribution
- Multi-sample
- Censored data

ABSTRACT:

The α-series process is an important counting process commonly used to model data sets having monotonic trend. It is especially utilized in reliability analysis of deteriorating systems and warranty analysis of repairable systems. When a data set is compatible with the α-series process, it is important to make inference for model parameters of the process. All the studies in the literature only consider single realization of the process which only has complete samples. However, multi-sample of the process may be observed. In this situation, the data set includes both complete and censored samples. In this study, estimation problem for an α-series process under censored data is studied by assuming inter-arrival times of the process have exponential distribution and all samples are homogeneous. Maximum likelihood estimators of the model parameters are obtained and their asymptotic properties such as asymptotic normality and consistency are proved. Also, their small sample performances have been investigated by a simulation study.

[1] Ömer ALTINDAĞ (**Orcid ID:** 0000-0002-7035-9612), Bilecik Şeyh Edebali University, Faculty of Science, Department of Statistics and Computer Sciences, Bilecik, Türkiye
[2] Mahmut KARA (**Orcid ID:** 0000-0001-7678-8824), Van Yüzüncü Yıl University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Van, Türkiye
[3] Halil Aydoğdu (**Orcid ID:** 0000-0001-5337-5277), Ankara University, Faculty of Science, Department of Statistics, Ankara, Türkiye
***Corresponding Author:** Ömer ALTINDAĞ, e-mail: omer.altindag@bilecik.edu.tr

INTRODUCTION

Counting processes are basic examples of stochastic processes modeling the number of events randomly occurring in a specified period. One of the well-known counting processes is homogeneous Poisson process (HPP) in which the inter-arrival times of randomly occurred consecutive events are independent and identical exponentially distributed. For most data sets representing the inter-arrival times of certain events, the independency condition may hold but the exponential distribution assumption may be restrictive. So, it can be assumed that they have a general distribution. In such a case, the counting process turns out to a renewal process (RP) in which the inter-arrival times are independent and identically distributed with a general distribution function F. If F is chosen as exponential distribution, the RP reduces to an HPP. Therefore, the RP is a generalization of HPP. The RP has been widely used in the fields of applied probability such as reliability analysis, warranty analysis, risk analysis etc. since its introduction in the 1950s. For basics and recent applications of RP, see (Barlow and Proschan, 1996; Chukova and Hayakawa, 2004; Blischke and Murthy, 2011; Fleming and Harrington, 2013; Jiang, 2020; Altındağ and Aydoğdu, 2021). If a data set representing the inter-arrival times of consecutive events doesn't hold identically distributed feature, the non-homogenous Poisson process (NHPP), in which the inter-arrival times are neither independent nor identically distributed, may be used. The non-identical property of inter-arrival times of NHPP allows us to model data sets having trend properly. However, the dependency between inter-arrival times of consecutive events may result in some difficulties in modeling the data set. To overcome the difficulty of dependency, (Lam, 1988) introduced a monotonic counting process model in which the inter-arrival times are assumed to be independent but may be stochastically monotone rather than identically distributed. This monotonic counting process model proposed by (Lam, 1988) is called as geometric process (GP). The formal definition of GP is given below.

Let $X_1, X_2, \ldots$ be non-negative random variables representing the inter-arrival times of consecutive events, then the process $\{X_k,\ k = 1,2,\ldots\}$ said to be a GP with trend parameter $a > 0$, if the random variables $a^{k-1}X_k,\ k = 1,2,\ldots$ are independent and identically distributed with a general distribution function F. The GP reduces to RP when $a = 1$. Therefore, it may be considered as a generalization of RP allowing the inter-arrival times may not to be identically distributed. The GP is stochastically increasing when $a < 1$ and, stochastically decreasing when $a > 1$. This feature allows the GP to model data sets having monotonic trend in time. The GP has been utilized for many fields of applied probability such as reliability analysis, warranty analysis, medicine applications etc. For a comprehensive consideration of GP and its recent applications, we refer to (Lam, 2007; Aydoğdu and Altındağ, 2016; Pekalp and Aydoğdu, 2021).

Although the GP is easily applicable for data sets having monotonic trend, it has some disadvantages. Let $N(t) = \sup\{n: X_1 + X_2 + \cdots X_n \leq t\}$. Then, $E[N(t)]$, which gives the expected number of events occurring in $(0, t]$, is not defined when the GP is stochastically increasing, i.e. $a < 1$. Further, the monotonic trend exhibited by GP is either logarithmically slow or exponentially fast. To overcome these disadvantages of GP, (Braun et al., 2005) introduced a monotonic counting process model, called as α-series process (ASP), as an alternative to GP. The ASP is defined as follows.

Let $X_1, X_2, \ldots$ be non-negative random variables representing the inter-arrival times of consecutive events, then the process $\{X_k,\ k = 1,2,\ldots\}$ said to be an ASP with trend parameter $\alpha \in \mathbb{R}$ if the random variables $k^{\alpha}X_k,\ k = 1,2,\ldots$ are independent and identically distributed with a general distribution function F. It is obvious that, the ASP reduces to RP when $\alpha = 0$. So, the ASP is another generalization of RP like the GP. Note that, the ASP is stochastically increasing when $\alpha < 0$ and, stochastically decreasing when $\alpha > 0$. But, unlike the GP, the $E[N(t)]$ is defined for an ASP either the ASP is stochastically increasing or stochastically decreasing. Furthermore, the monotonic trend exhibited by ASP is moderate compared to GP, see for details (Braun et al., 2005; 2008).

Let $\{X_k,\ k = 1,2,\ldots\}$ be an ASP with trend parameter α and $E(X_1) = \mu$, $Var(X_1) = \sigma^2$. Then, $E(X_k) = \mu k^{-\alpha}$ and $Var(X_k) = \sigma^2 k^{-2\alpha}$ for $k = 1,2,\ldots$. From a statistical point of view, it is important to estimate the parameters α, μ and σ^2 when there exists a data set compatible with ASP. The estimation problem of these parameters is well studied in the literature. (Aydoğdu and Kara, 2012) considered non-parametric estimation of the parameters by utilizing the linear regression method. (Kara et al., 2017a) studied statistical inference for ASP with gamma distributed inter-arrival times. (Kara et al., 2017b) considered statistical inference for ASP with inverse-Gaussian distributed inter-arrival times. (Kara et al., 2019) studied parameter estimation for ASP with log-normal distributed inter-arrival times. In these studies, the data sets are assumed to consist of only complete observations coming from a single realization of the process. However, the data may come from multiple processes which yields that some inter-arrival times may be observed as censored. The data structure with censored observations is illustrated in Figure 1 below.

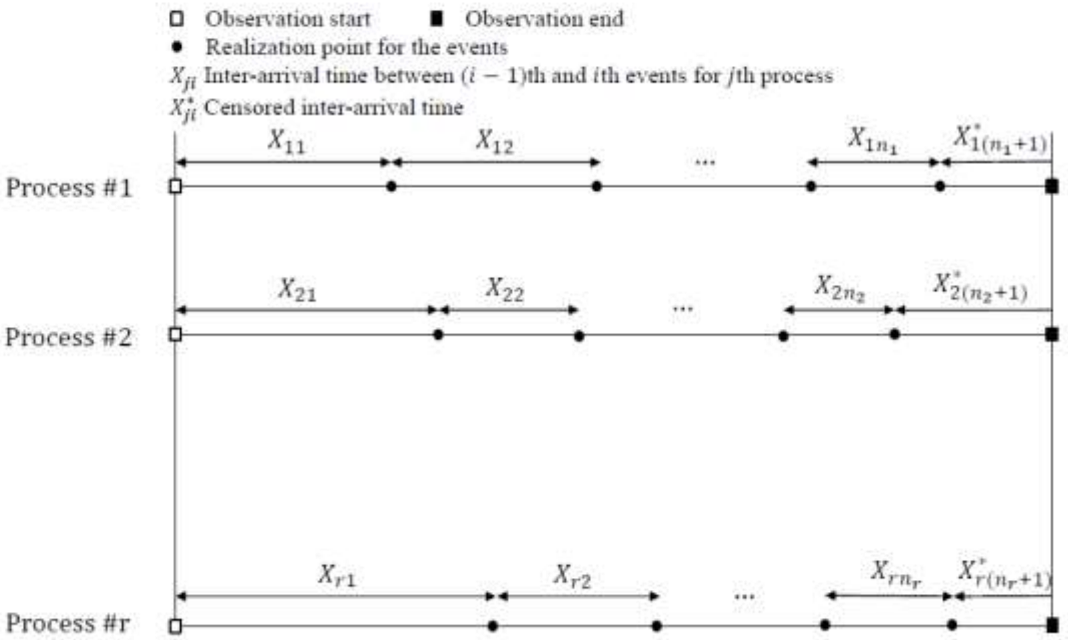

Figure 1. Data structure for ASP with both complete and censored observations

It is important to estimate the parameters α, μ and σ^2 based on both complete observations X_{ji}, $i = 1,\ldots,n_j$, $j = 1,\ldots,r$ and censored observations X^*_{ji}, $i = n_j + 1$, $j = 1,\ldots,r$. There is no study in the literature considering this type of data structure for ASP. For this reason, we consider the censored data to estimate the model parameters of ASP by assuming the inter-arrival times are distributed as exponential and all the processes are homogeneous.

MATERIALS AND METHODS

Let r homogeneous ASPs with common trend parameter α are observed until a pre-determined time, say T, and $\{X_{ji}, i = 1,..,n_j + 1\}$, $j = 1,\ldots,r$ be the inter-arrival times of jth process, X_{j1} has distribution function $F(x) = 1 - e^{-\lambda x}, x \geq 0; \lambda > 0$, $\mu \coloneqq E(X_{j1}) = 1/\lambda$, $\sigma^2 \coloneqq Var(X_{j1}) = 1/\lambda^2$ for $j = 1,\ldots,r$. Then, distribution function and probability density function of X_{ji} are given as $F_i(x) = 1 - e^{-i^\alpha \lambda x}, x \geq 0; \lambda > 0$, $f_i(x) = i^\alpha \lambda e^{-i^\alpha \lambda x}, x \geq 0; \lambda > 0$, $i = 1,..,n_j + 1$ for $j = 1,\ldots,r$. Therefore, mean and variance of X_{ji} are $\mu_i \coloneqq E(X_{ji}) = \mu/i^\alpha$, $\sigma_i^2 \coloneqq Var(X_{ji}) = \sigma^2/i^{2\alpha}$, $i = 1,..,n_j + 1$ for $j = 1,\dots,r$. Note that, the inter-arrival times $\{X_{ji}, i = 1,..,n_j\}$ for $j = 1,\dots,r$ are complete while $\{X_{j(n_j+1)}\}$ for $j = 1,\dots,r$ are right censored as demonstrated in Figure 1.

To estimate the parameters α, μ and σ^2 based on the observations $\{X_{ji}, i = 1,..,n_j + 1,\ j = 1,\dots,r\}$, we will use maximum likelihood method due to its easy implementation and asymptotically well-behaviour.

Maximum Likelihood Estimators

Let's denote the complete inter-arrival times as $\boldsymbol{X}_{com} = \{X_{ji}, i = 1,\dots,n_j, j = 1,\dots,r\}$, censored inter-arrival times as $\boldsymbol{X}_{cens} = \{X_{j(n_j+1)}, j = 1,\dots,r\}$ and total data as $\boldsymbol{X}_t = (\boldsymbol{X}_{com}, \boldsymbol{X}_{cens})$. Let $\boldsymbol{x}_t$ be the sample points of $\boldsymbol{X}_t$. Then, the likelihood function based on the sample $\boldsymbol{x}_t$ is obtained as

$$\begin{aligned} L(\alpha,\lambda;\boldsymbol{x}_t) &= \prod_{j=1}^{r}\left[\prod_{i=1}^{n_j} f_i(x_{ji})\left[1 - F_{n_j+1}\left(x_{j(n_j+1)}\right)\right]\right] \\ &= \prod_{j=1}^{r}\left[\prod_{i=1}^{n_j} i^\alpha \lambda e^{-i^\alpha \lambda x_{ji}}\left[e^{-i^\alpha \lambda x_{j(n_j+1)}}\right]\right]. \end{aligned} \tag{1}$$

Therefore, the log-likelihood function is

$$\ln L(\alpha,\lambda;\boldsymbol{x}_t) = \alpha\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i) + \lambda\sum_{j=1}^{r} n_j - \lambda\sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha x_{ji}\,. \tag{2}$$

By taking partial derivatives of the log-likelihood function with respect to α and λ and equating them to zero, we obtain the following equations:

$$\frac{1}{\lambda}\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i) - \sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha \ln(i) x_{ji} = 0 \tag{3}$$

$$\sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha x_{ji} - \frac{1}{\lambda}\sum_{j=1}^{r} n_j = 0 \tag{4}$$

The common solution of these equations gives maximum likelihood estimators of the parameters α and λ. If we take λ as $(\sum_{j=1}^{r} n_j)/(\sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha x_{ji})$ in Equation (3), the following non-linear equation is obtained for α:

$$\sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha x_{ji}\left[\frac{\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i)}{\sum_{j=1}^{r} n_j} - \ln(i)\right] = 0 \tag{5}$$

Solution of Equation (5) gives the maximum likelihood estimator of α. It is clear that, this equation can't be solved analytically. So, it must be solved numerically. To solve this non-linear equation, the Newton-Raphson method can be applied. Let

$$g(\alpha) = \sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^\alpha x_{ji}\left[\frac{\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i)}{\sum_{j=1}^{r} n_j} - \ln(i)\right]. \tag{6}$$

Then, the first derivate of $g(\alpha)$ is obtained as

$$g'(\alpha) = \sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^{\alpha}\ln(i)x_{ji}\left[\frac{\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i)}{\sum_{j=1}^{r}n_j} - \ln(i)\right]. \tag{7}$$

Let $\alpha(1) = 0$ and $\alpha(k+1) = \alpha(k) - \frac{g(\alpha(k))}{g'(\alpha(k))}$ for $= 1,2,, \ldots$. Then, numerical solution of Equation (5) is obtained by repeating the iterative steps until the condition $|\alpha(k+1) - \alpha(k)| < \varepsilon$ holds where $\varepsilon > 0$ is a pre-defined tolerance level. Once the condition $|\alpha(k+1) - \alpha(k)| < \varepsilon$ holds, the maximum likelihood estimator of α is obtained as

$$\hat{\alpha} = \alpha(k+1). \tag{8}$$

Hence, the maximum likelihood estimator of λ is

$$\hat{\lambda} = \frac{\sum_{j=1}^{r} n_j}{\sum_{j=1}^{r}\sum_{i=1}^{n_j+1} i^{\hat{\alpha}}x_{ji}}. \tag{9}$$

Further, maximum likelihood estimators of the parameters μ and σ^2 are obtained as

$$\hat{\mu} = \frac{1}{\hat{\lambda}}, \tag{10}$$

$$\hat{\sigma}^2 = \frac{1}{\hat{\lambda}^2}. \tag{11}$$

The asymptotic properties of the estimators are given below.

Theorem 1. Let $\hat{\alpha}$ and $\hat{\lambda}$ be maximum likelihood estimators of the parameters α and λ based on the data $\boldsymbol{X}_t$. Then,

$$\begin{bmatrix}\hat{\alpha}\\ \hat{\lambda}\end{bmatrix} \sim AN\left(\begin{bmatrix}\hat{\alpha}\\ \hat{\lambda}\end{bmatrix}, I^{-1}(\alpha,\lambda;\boldsymbol{X}_t)\right), \tag{12}$$

as $T \to \infty$, where AN stands for asymptotically normal and $I^{-1}(\alpha,\lambda;\boldsymbol{X}_t)$ is inverse of Fisher information such that

$$I(\alpha,\lambda;\boldsymbol{X}_t) = \begin{bmatrix} \sum_{j=1}^{r}\left[\sum_{i=1}^{n_j}\ln^2(i) + \ln^2(n_j+1)F_{n_j+1}(t_j)\right] & \frac{1}{\lambda}\sum_{j=1}^{r}\left[\sum_{i=1}^{n_j}\ln(i) + \ln(n_j+1)F_{n_j+1}(t_j)\right] \\ \frac{1}{\lambda}\sum_{j=1}^{r}\left[\sum_{i=1}^{n_j}\ln(i) + \ln(n_j+1)F_{n_j+1}(t_j)\right] & \frac{1}{\lambda^2}\sum_{j=1}^{r}\left[n_j + F_{n_j+1}(t_j)\right] \end{bmatrix}. \tag{13}$$

Here, t_j is the censoring time of $X_{j(n_j+1)}$ such that $t_j := x_{j(n_j+1)} = T - \sum_{i=1}^{n_j} x_{ji}$ given $X_{ji} = x_{ji}$ and $F_{n_j+1}(t_j) = 1 - e^{-(n_j+1)^{\alpha}\lambda t_j}$ for $j = 1, \ldots, r$.

Proof of Theorem 1. Let $I(\alpha,\lambda;\boldsymbol{X}_{com})$ be the Fisher information based on $\boldsymbol{X}_{com}$ and $I(\alpha,\lambda;\boldsymbol{X}_{cens})$ be the Fisher information based on $\boldsymbol{X}_{cens}$. Then, $I(\alpha,\lambda;\boldsymbol{X}_t) = I(\alpha,\lambda;\boldsymbol{X}_{com}) + I(\alpha,\lambda;\boldsymbol{X}_{cens})$. So, we need to calculate $I(\alpha,\lambda;\boldsymbol{X}_{com})$ and $I(\alpha,\lambda;\boldsymbol{X}_{cens})$ separately. Let $L(\alpha,\lambda;\boldsymbol{X}_{com})$ be the likelihood function based on $\boldsymbol{X}_{com}$. Then, it is obtained that,

$$L(\alpha,\lambda;\boldsymbol{X}_{com}) = \prod_{j=1}^{r}\left[\prod_{i=1}^{n_j} i^{\alpha}\lambda e^{-i^{\alpha}\lambda X_{ji}}\right], \tag{14}$$

and

$$\ln L(\alpha,\lambda;\boldsymbol{X}_{com}) = \alpha\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i) + \lambda\sum_{j=1}^{r}n_j - \lambda\sum_{j=1}^{r}\sum_{i=1}^{n_j} i^{\alpha}X_{ji}. \tag{15}$$

If we take partial derivatives of $\ln L(\alpha,\lambda;\boldsymbol{X}_{com})$, it is obtained that,

$$\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\alpha^2} = -\lambda\sum_{j=1}^{r}\sum_{i=1}^{n_j} i^{\alpha}\ln^2(i)\,X_{ji}\,, \tag{16}$$

$$\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\lambda^2} = -\frac{1}{\lambda^2}\sum_{j=1}^{r} n_j, \tag{17}$$

$$\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\alpha\partial\lambda} = -\sum_{j=1}^{r}\sum_{i=1}^{n_j} i^{\alpha}\ln(i)\,X_{ji}\,. \tag{18}$$

Therefore, negative expectations of partial derivates are obtained as

$$E\left[-\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\alpha^2}\right] = \sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln^2(i)\,, \tag{19}$$

$$E\left[-\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\lambda^2}\right] = \frac{1}{\lambda^2}\sum_{j=1}^{r} n_j, \tag{20}$$

$$E\left[-\frac{\partial^2 L(\alpha,\lambda;\boldsymbol{X}_{com})}{\partial\alpha\partial\lambda}\right] = \frac{1}{\lambda}\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i)\,, \tag{21}$$

since $E\left[i^{\alpha}X_{ji}\right] = 1/\lambda$. Hence,

$$I(\alpha,\lambda;\boldsymbol{X}_{com}) = \begin{bmatrix} \sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln^2(i) & \frac{1}{\lambda}\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i) \\ \frac{1}{\lambda}\sum_{j=1}^{r}\sum_{i=1}^{n_j}\ln(i) & \frac{1}{\lambda^2}\sum_{j=1}^{r} n_j \end{bmatrix}. \tag{22}$$

As for the $I(\alpha,\lambda;\boldsymbol{X}_{cens})$, we first introduce the Fisher information for a non-negative right censored random variable. Let $h_{n_j+1}(x)$ be the hazard function and t_j be the censoring time of censored variable $X_{j(n_j+1)}$ for $j = 1,\dots,r$. Then,

$$I\left(\alpha,\lambda;X_{j(n_j+1)}\right)_{11} = \int_0^{t_j}\left(\frac{\partial}{\partial\alpha}\ln\left(h_{n_j+1}(x)\right)\right)^2 f_{n_j+1}(x)dx = \ln^2\left(n_j+1\right)F_{n_j+1}\left(t_j\right), \tag{23}$$

$$I\left(\alpha,\lambda;X_{j(n_j+1)}\right)_{22} = \int_0^{t_j}\left(\frac{\partial}{\partial\lambda}\ln\left(h_{n_j+1}(x)\right)\right)^2 f_{n_j+1}(x)dx = \frac{1}{\lambda^2}F_{n_j+1}\left(t_j\right), \tag{24}$$

$$\begin{aligned} I\left(\alpha,\lambda;X_{j(n_j+1)}\right)_{12} &= \int_0^{t_j}\left(\frac{\partial}{\partial\alpha}\ln\left(h_{n_j+1}(x)\right)\right)\left(\frac{\partial}{\partial\lambda}\ln\left(h_{n_j+1}(x)\right)\right)f_{n_j+1}(x)dx \\ &= \frac{1}{\lambda}\ln\left(n_j+1\right)F_{n_j+1}\left(t_j\right), \end{aligned} \tag{25}$$

since $\ln\left(h_{n_j+1}(x)\right) = \alpha\ln\left(n_j+1\right) + \ln(\lambda)$. For the Fisher information of right censored non-negative random variables, see (Zheng and Gastwirth, 2001; Park et al., 2008.) Therefore, it is obtained that,

$$I(\alpha,\lambda;\boldsymbol{X}_{cens}) = \begin{bmatrix} \sum_{j=1}^{r}\ln^2(n_j+1)\,F_{n_j+1}(t_j) & \frac{1}{\lambda}\sum_{j=1}^{r}\ln(n_j+1)\,F_{n_j+1}(t_j) \\ \frac{1}{\lambda}\sum_{j=1}^{r}\ln(n_j+1)\,F_{n_j+1}(t_j) & \frac{1}{\lambda^2}\sum_{j=1}^{r}F_{n_j+1}(t_j) \end{bmatrix}. \tag{26}$$

Consequently, the result is clear, and the proof is completed.

Corollary 1. Let $\hat{\alpha}$ and $\hat{\lambda}$ be maximum likelihood estimators of the parameters α and λ based on the data $\boldsymbol{X}_t$. Then, as $T\to\infty$,

$$\hat{\alpha}\sim AN(\alpha,\vartheta_{11}), \tag{27}$$

$$\hat{\lambda}\sim AN(\lambda,\vartheta_{22}), \tag{28}$$

where

$$\vartheta_{11} = \frac{D}{AD-B^2}, \tag{29}$$

$$\vartheta_{22} = \lambda^2\frac{A}{AD-B^2}, \tag{30}$$

$$A = \sum_{j=1}^{r}\left[\sum_{i=1}^{n_j}\ln^2(i) + \ln^2(n_j+1)\,F_{n_j+1}(t_j)\right], \tag{31}$$

$$B = \sum_{j=1}^{r}\left[\sum_{i=1}^{n_j}\ln(i) + \ln(n_j+1)\,F_{n_j+1}(t_j)\right], \tag{32}$$

$$D = \sum_{j=1}^{r}\left[n_j + F_{n_j+1}(t_j)\right]. \tag{33}$$

Proof of Corollary 1. The result is easily obtained by inverting the Fisher information matrix $I(\alpha,\lambda;\boldsymbol{X}_t)$.

Corollary 2. Let $\hat{\alpha}$ and $\hat{\lambda}$ be maximum likelihood estimators of the parameters α and λ based on the data $\boldsymbol{X}_t$. Then,

$$\hat{\alpha}\xrightarrow{P}\alpha, \tag{34}$$

$$\hat{\lambda}\xrightarrow{P}\lambda, \tag{35}$$

as $T\to\infty$, where $\xrightarrow{P}$ denotes convergence in probability, that is, the estimators $\hat{\alpha}$ and $\hat{\lambda}$ are consistent.

Proof of Corollary 2. To prove consistencies of the estimators, it is sufficient to show that the asymptotic variances ϑ_{11} and ϑ_{22} converge to zero as $T\to\infty$. It is clear that, the number of completely observed inter-arrival times n_j for each process increases as the observation ending time T increases. That is, $n_j\to\infty$ for $j=1,\dots,r$ as $T\to\infty$. Then, $F_{n_j+1}(t_j)$ converges to 1 if $\alpha>0$, or it converges to 0 if $\alpha<0$ as $T\to\infty$, regardless of the value of censoring time t_j, for $j=1,\dots,r$. It is obvious that, $n_j\simeq n_j+1$, where " $\simeq$ " denotes asymptotic equivalence. (Kara et al., 2019) showed that,

$$\sum_{i=1}^{n}\ln(i)\simeq n(\ln(n)-1), \tag{36}$$

$$\sum_{i=1}^{n}\ln^2(i)\simeq n\left(2+\ln^2(n)-2\ln(n)\right). \tag{37}$$

Furthermore, all n_j's are asymptotically equal as $T \to \infty$. Let n^* denotes an asymptotical equivalent of n_j's. Then,

$$\vartheta_{11} \simeq \frac{rn^*}{\left[r\sum_{i=1}^{n^*} \ln^2(i)\right]rn^* - \left[r\sum_{i=1}^{n^*} \ln(i)\right]^2}. \tag{38}$$

$$\vartheta_{22} \simeq \lambda^2 \frac{r\sum_{i=1}^{n^*} \ln^2(i)}{\left[r\sum_{i=1}^{n^*} \ln^2(i)\right]rn^* - \left[r\sum_{i=1}^{n^*} \ln(i)\right]^2}. \tag{39}$$

By considering the asymptotic equivalences given in Equation (36) and (37), $\vartheta_{11} \to 0$ and $\vartheta_{22} \to 0$ as $T \to \infty$, since the denominators of both terms are of higher order than the numerators. So, the proof is concluded.

Asymptotic distributions of the maximum likelihood estimators of μ and σ^2 are given below.

Corollary 3. Let $\hat{\mu}$ and $\hat{\sigma}^2$ be maximum likelihood estimators of the parameters μ and σ^2 based on the data $\boldsymbol{X}_t$. Then, as $T \to \infty$,

$$\hat{\mu} \sim AN(\mu, \mu^4 \vartheta_{22}), \tag{40}$$

$$\hat{\sigma}^2 \sim AN(\sigma^2, 4(\sigma^2)^3 \vartheta_{22}). \tag{41}$$

Proof of Corollary 3. It is known that, $\hat{\mu} = 1/\hat{\lambda}$ and $\hat{\sigma}^2 = 1/\hat{\lambda}^2$. Let $g_1(x) = 1/x$ and $g_2(x) = 1/x^2$. Then, $\hat{\mu} = g_1(\hat{\lambda})$ and $\hat{\sigma}^2 = g_2(\hat{\lambda})$. First derivatives of the functions $g_1(x)$ and $g_2(x)$ are obtained as $g_1'(x) = -1/x^2$ and $g_2'(x) = -2/x^3$. The result is clear via the well-known delta method since $[g_1'(\lambda)]^2 = 1/\lambda^4$ and $[g_2'(\lambda)]^2 = 4/\lambda^6$, $\mu = 1/\lambda$ and $\sigma^2 = 1/\lambda^2$.

Corollary 4. Let $\hat{\mu}$ and $\hat{\sigma}^2$ be maximum likelihood estimators of the parameters μ and σ^2 based on the data $\boldsymbol{X}_t$. Then, as $T \to \infty$,

$$\hat{\mu} \xrightarrow{P} \mu, \tag{42}$$

$$\hat{\sigma}^2 \xrightarrow{P} \sigma^2, \tag{43}$$

that is, the estimators $\hat{\mu}$ and $\hat{\sigma}^2$ are consistent.

Proof of Corollary 4. The result is obvious by continuous mapping theorem since the estimators $\hat{\mu}$ and $\hat{\sigma}^2$ are functions of $\hat{\lambda}$, which is proved to be consistent in Corollary 2.

It should be noted that, the maximum likelihood estimators of the parameters are derived under the ASP model. However, the goodness-of-fit of the ASP model for the data $\boldsymbol{X}_t$ must be tested. For this purpose, the hypothesis $H_0\colon \alpha = 0$ against $H_1\colon \alpha \neq 0$ may be tested with the following test statistic $S = \hat{\alpha}/\hat{\vartheta}_{11}$ to distinguish the ASP from its non-monotonic counterpart RP. Here, $\hat{\vartheta}_{11}$ is obtained by replacing $F_{n_j+1}(t_j)$ with its estimation $\hat{F}_{n_j+1}(t_j) = 1 - e^{-(n_j+1)^{\hat{\alpha}}\hat{\lambda}t_j}$. From Slutsky and continuous mapping theorem, $S \sim AN(0,1)$ under H_0. Therefore, the hypothesis $H_0\colon \alpha = 0$ is rejected at significance level α^* if $|S| > z_{\alpha^*/2}$, where $z_{\alpha^*/2}$ denotes upper $\alpha^*/2$ tail of the standard normal distribution. Then, it is concluded that the data $\boldsymbol{X}_t$ has a trend and it can be modelled by the ASP.

Simulation Study

In this section, we carry out a Monte Carlo simulation to observe small sample performances of the maximum likelihood estimators $\hat{\alpha}$, $\hat{\mu}$ and $\hat{\sigma}^2$. All the results are given in Table 1 below.

In simulations, the number of replications is chosen as $N = 1000$. The simulation has been conducted under different parameter settings but, for sake of simplicity, some of them are summarized in Table 1 since the results are similar. The number of independent samples is chosen as $r = 2, 3, 4$ and the observation ending time is chosen as $T = 20, 30, 50$. In the numerical computation of $\hat{\alpha}$ while utilizing the Newton-Raphson algorithm, the tolerance level is chosen as $\varepsilon = 1/1000$.

Table 1. Results for the maximum likelihood estimators under different settings of parameters

$\alpha = -0.8,\ \lambda = 5,\ \mu = 0.2,\ \sigma^2 = 0.04$					
	n^*	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\sigma}^2$
$r = 2, T = 20$	18	-0.7964	5.4748	0.2317	0.1063
		0.0448	4.4986	0.0527	0.5252
$r = 3, T = 20$		-0.7993	5.3398	0.2153	0.0647
		0.0287	2.9050	0.0184	0.0806
$r = 4, T = 20$		-0.8056	5.2779	0.2073	0.0480
		0.0189	2.2302	0.0050	0.0019
$r = 2, T = 30$	22	-0.8008	5.4757	0.2351	0.1449
		0.0372	4.0617	0.0898	1.4882
$r = 3, T = 30$		-0.8095	5.3567	0.2073	0.0496
		0.0178	2.6159	0.0067	0.0047
$r = 4, T = 30$		-0.8029	5.2474	0.2073	0.0474
		0.0145	2.0981	0.0044	0.0014
$r = 2, T = 50$	30	-0.8007	5.3340	0.2153	0.0587
		0.0191	3.0708	0.0124	0.0216
$r = 3, T = 50$		-0.7997	5.2027	0.2096	0.0485
		0.0123	2.1538	0.0046	0.0014
$r = 4, T = 50$		-0.8031	5.1916	0.2059	0.0457
		0.0096	1.6487	0.0033	0.0009
$\alpha = 0.2,\ \lambda = 1,\ \mu = 1,\ \sigma^2 = 1$					
	n^*	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\sigma}^2$
$r = 2, T = 20$	34	0.1504	1.1485	0.9203	0.8970
		0.0082	0.0804	0.0555	0.2332
$r = 3, T = 20$		0.1668	1.1179	0.9748	1.0475
		0.0090	0.1040	0.1080	0.5880
$r = 4, T = 20$		0.1845	0.9774	1.0365	1.0885
		0.0025	0.0135	0.0158	0.0697
$r = 2, T = 30$	56	0.1929	1.0855	1.0333	1.2119
		0.0115	0.1274	0.1602	0.9905
$r = 3, T = 30$		0.1949	0.9996	1.0316	1.1005
		0.0022	0.0305	0.0402	0.1984
$r = 4, T = 30$		0.2029	1.0533	0.9749	0.9748
		0.0027	0.0338	0.0272	0.1052
$r = 2, T = 50$	102	0.2065	1.0152	1.0707	1.2457
		0.0077	0.0925	0.1102	0.6089
$r = 3, T = 50$		0.1941	1.0432	1.0242	1.1153
		0.0050	0.0801	0.0736	0.3123
$r = 4, T = 50$		0.1977	1.0299	1.0117	1.0667
		0.0033	0.0467	0.0480	0.2144

It should be noted that, the number of inter-arrival times for each sample is random due to randomness of inter-arrival times although the observation ending time is pre-defined. That is, the sample sizes in each simulation are random. So, we give mean number of inter-arrival times in each sample as n^*. It is obvious that, the number of inter-arrival times in each sample increases as the observation ending time T increases. In Table 1, the first rows give simulation mean of the estimators and the second rows give simulation variance of the estimators.

RESULTS AND DISCUSSION

The ASP is an important monotonic stochastic model commonly used in applied probability fields. It is more convenient than its some counterparts due to its moderate trend behaviour. When a data set having monotonic trend is analysed, it is important to estimate the parameters of model. In the literature, estimation problem for an ASP is well studied. All existing studies consider single realization of the process, that is, there is only one sample of data. However, there may occur multi-sample of the process. There isn't any study in the literature dealing with this situation. It should be noted, all the existing studies only consider complete sample case though the multi-sample may include both complete and censored samples. For this purpose, the multi-sample case for an ASP has been considered statistically.

When the results given in Table 1 are analysed, it is seen that the maximum likelihood estimators $\hat{\alpha}$, $\hat{\mu}$ and $\hat{\sigma}^2$ perform well regardless of the different parameter settings. All the estimators have small biases even if the observation ending time T is relatively short. However, they seem to be asymptotically unbiased as the biases decrease as observation ending time T or number of independent samples r increases. Further, their variances decrease as T or r increases. Because the number of observed inter-arrival times increases as T or r increases. These results support the consistencies of the estimators which is theoretically proved above.

It has been observed that, the model parameters of an ASP can be effectively estimated based on the multiple homogeneous samples which may include both complete and censored inter-arrival times. The maximum likelihood estimators perform well with different parameter settings.

CONCLUSION

In this study, the ASP has been analysed statistically by assuming that the inter-arrival times have exponential distribution and that the data available consists of multiple homogeneous samples which have both complete and censored samples. The maximum likelihood estimators for the parameters of ASP have been obtained and their asymptotic properties have been established. Asymptotic distributions have been derived and consistencies of the estimators have been proved. Besides the asymptotic properties of the estimators, their small sample behaviours have been investigated. Also, a test statistic to distinguish ASP from a RP has been introduced. It has been exhibited that, statistical estimation for an ASP based on multi-sample is quite efficient. It should be noted that, the multiple samples of an ASP have been assumed to be homogeneous. However, if there is not enough evidence to assume homogeneity, it must be tested statistically whether the samples are homogeneous. Further, the inter-arrival times are assumed to have exponential distribution. To widen the present study, some general distributions such as gamma distribution, Weibull distribution, log-normal distribution, etc. may be taken as the distribution of inter-arrival times. These cases should be considered as a future study.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their careful reading and contribution.

Conflict of Interest

The article authors declare that there is no conflict of interest between them.

Author's Contributions

The authors declare that they have contributed equally to the article.

REFERENCES

Altındağ, Ö., & Aydoğdu, H. (2021). Estimation of renewal function under progressively censored data and its applications. *Reliability Engineering & System Safety, 216,* 107988.

Aydoğdu, H., & Altındağ, Ö. (2016). Computation of the mean value and variance functions in geometric process. *Journal of Statistical Computation and Simulation, 86*(5), 986-995.

Aydoğdu, H., & Kara, M. (2012). Nonparametric estimation in α-series processes. *Computational statistics & data analysis, 56*(1), 190-201.

Barlow, R. E., & Proschan, F. (1996). *Mathematical theory of reliability*. Society for Industrial and Applied Mathematics.

Blischke, W. R., & Murthy, D. P. (2011). *Reliability: modeling, prediction, and optimization*. John Wiley & Sons.

Braun, W. J., Li, W., & Zhao, Y. Q. (2005). Properties of the geometric and related processes. *Naval Research Logistics (NRL)*, *52*(7), 607-616.

Braun, W. J., Li, W., & Zhao, Y. Q. (2008). Some theoretical properties of the geometric and α-series processes. *Communications in Statistics—Theory and Methods*, *37*(9), 1483-1496.

Chukova, S., & Hayakawa, Y. (2004). Warranty cost analysis: Non-zero repair time. *Applied Stochastic Models in Business and Industry*, *20*(1), 59-71.

Fleming, T. R., & Harrington, D. P. (2013). *Counting processes and survival analysis* (Vol. 625). John Wiley & Sons.

Jiang, R. (2020). A novel two-fold sectional approximation of renewal function and its applications. *Reliability Engineering & System Safety*, *193*, 106624

Kara, M., Aydoğdu, H., & Şenoğlu, B. (2017a). Statistical inference for α-series process with gamma distribution. *Communications in Statistics-Theory and Methods*, *46*(13), 6727-6736.

Kara, M., Türkşen, Ö., & Aydoğdu, H. (2017b). Statistical inference for α-series process with the inverse Gaussian distribution. *Communications in Statistics-Simulation and Computation*, *46*(6), 4938-4950.

Kara, M., Altındağ, Ö., Pekalp, M. H., & Aydoğdu, H. (2019). Parameter estimation in α-series process with lognormal distribution. *Communications in Statistics-Theory and Methods*, *48*(20), 4976-4998.

Lam, Y. (1988). Geometric processes and replacement problem. *Acta Mathematicae Applicatae Sinica*, *4*, 366-377.

Lam, Y. (2007). *The geometric process and its applications*. World Scientific.

Park, S., Balakrishnan, N., & Zheng, G. (2008). Fisher information in hybrid censored data. *Statistics & probability letters*, *78*(16), 2781-2786.

Pekalp, M. H., & Aydoğdu, H. (2021). Power series expansions for the probability distribution, mean value and variance functions of a geometric process with gamma interarrival times. *Journal of Computational and Applied Mathematics*, *388*, 113287.

Zheng, G., & Gastwirth, J. L. (2001). On the Fisher information in randomly censored data. *Statistics & probability letters*, *52*(4), 421-426.