A Stationary Proportional Hazard Class Process and its Applications

Debasis Kundu *

Abstract

The motivation of this work came when we were trying to analyze gold price data of the Indian market and the exchange rate data between Indian Rupees and US Dollars. It is observed that in both the cases there is a significant amount of time when $X_n = X_{n+1}$, hence they cannot be ignored. In this paper we have introduced a very flexible discrete time and continuous state space stationary stochastic process $\{X_n\}$, where X_n has a proportional hazard class of distribution and there is a positive probability that $X_n = X_{n+1}$. We have assumed a very flexible piecewise constant hazard function of the base line distribution of the proportional hazard class. Various properties of the proposed class has been obtained. Various dependency properties have been established. Estimating the cut points of the piecewise constant hazard function is an important problem and it has been addressed here. The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in closed form, and we have proposed to use profile likelihood method to compute the estimators. The gold price data set and the exchange rate data set have been analyzed and the results are quite satisfactory.

Key Words and Phrases: Weibull distribution, exponential distribution; maximum likelihood estimators; piecewise constant hazard function; copula.

AMS Subject Classifications: 62F10, 62F03, 62H12.

*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

The motivation of this work came when we were trying to analyze the gold price data of the Indian market during the month of December 2020 and January 2021, and also the exchange rate data of the Indian Rupees and US dollars during January 2023 and July 2023. It is observed in both the cases that the data are coming from a positive valued lag-1 stationary process and there are several instances for which $X_n = X_{n+1}$, which cannot be ignored. It may be mentioned that there are several positive valued lag-1 stationary processes available in the literature; for example the exponential process by Tavares [27], Weibull and gamma process by Sim [26], Pareto process by Yeh et al. [28], semi-Pareto process by Pillai [24], Weibull process by Jayakumar and Girish Babu [14], and see the references cited there in. Unfortunately, none of these processes can be applied in this case as in all these cases $P(X_n = X_{n+1}) = 0$. It may be mentioned that the Pareto process introduced by Arnold and Hallet [3] can take a positive probability on $X_n = X_{n+1}$, but they did not develop any inference procedures for that process. Moreover, the proposed process is more general in certain aspects than the process of Arnold and Hallet [3]. Kundu [18] also introduced a stationary process which can take a positive probability on $X_n = X_{n+1}$, but the proposed model is easier to implement in practice and it can be more flexible than the existing models.

The main aim of this paper is to introduce a discrete time and continuous state space lag-1 stationary process $\{X_n\}$, for which $P(X_n = X_{n+1}) > 0$. We want $X_n > 0$, and it should be flexible enough to be applicable in different cases. In this case the marginals belong to the proportional hazard class (PHC) of distribution as introduced by Cox [9]. Due to this reason it will be called as the PHC-process. The PHC of distributions is a very flexible family of distribution functions and they have been used quite extensively in lifetime data analysis. The exponential and Weibull distributions belong to the PHC of distributions. In this paper we have not assumed any specific parametric family of the base line distribution function. It is assumed that the base line distribution has a piecewise constant hazard function, which makes the proposed PHC-process to be a very flexible process. Hence, if there is a non-negative time series data which are skewed and there is a positive probability that the consecutive values might be equal, then the proposed PHC-process can be used quite effectively to analyze the data set.

We have provided various properties of the proposed PHC-process. Since exponential and Weibull distributions are members of the PHC of distributions, the exponential process and Weibull process can be obtained as special cases of the proposed process. The joint distribution function of $\{X_n\}$ has been obtained, and it is observed that X_n is independent of X_{n+k}, for $k \geq 2$. The joint distribution of X_n and X_{n+1} has a very convenient copula structure, and based on this copula structure several dependency properties can be established. Generation of the random sample from this PHC-process is quite straight forward, hence bootstrapping or simulation experiments can be performed quite conveniently.

The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in explicit forms. It involves multidimensional optimization problem. We have suggested to use profile likelihood technique and it reduces the computational burden significantly. In this profile likelihood method, one needs to solve only one two dimensional optimization problem to compute the MLEs of the unknown parameters. The parametric bootstrap confidence intervals can be used to construct confidence intervals of the unknown parameters.

Another important point which has been addressed is the estimation of the change point of the base line hazard function. It is an important problem, and it has not been addressed in the literature properly. In this paper it is assumed that the position of the cut points has the same distribution as the arrival times of a non-homogeneous Poisson process. Based on this assumption we have obtained the maximum likelihood estimators of the position of the cut points. Two real life data sets have been analyzed to see how the proposed model works in practice. The results are quite satisfactory.

The rest of the paper is organized as follows. In Section 2, we provide the PHC of distributions and the piece wise constant hazard function. The PHC process and various properties are provided in Section 3. The two data sets and some preliminary data analysis have been performed in Section 4. In Section 5 we present the MLEs of the unknown parameters for different models. The data analyses are presented in Section 7 and we conclude the paper in Section 8.

2 Preliminaries and Some New Results

2.1 Proportional Hazard Class

Suppose $F_0(t)$ is an absolutely continuous distribution function with the support on the positive real axis, and $S_0(t) = 1 - F_0(t)$ is the corresponding survival function. Define a class of distribution function parameterized by $\alpha > 0$, which has the survival function

$$S(t;\alpha) = (S_0(t))^{\alpha}; \qquad t \geq 0. \tag{1}$$

Clearly, $F(t;\alpha) = 1 - (t;\alpha)$ is an absolutely continuous distribution function. It will be denoted by PHC(S_0, α). This class of distribution functions has been introduced by Cox [9]. This is known as the PHC of distributions because if $h_0(t)$ is the hazard function of $F_0(t)$, then $\alpha h_0(t)$ is the hazard function of $F(t;\alpha)$. Note that if $f_0(t)$ denotes the probability density function (PDF) of $F_0(t)$, then the PDF of $S(t;\alpha)$ is of the form

$$f_{PHC}(t;\alpha, S_0) = \alpha f_0(t)(S_0(t))^{\alpha-1}. \tag{2}$$

Throughout this article it is assumed that $S_0(t) = 1$ for all $t \leq 0$, although all the results can be extended for general S_0 also. Most of the results are valid when $S_0(t)$ has bounded support also. Several well known distribution functions belong to the PHC of distributions, for example exponential, Weibull, Rayleigh, generalized Pareto, Lomax, Gompertz etc. An extensive amount of work has been done in developing several properties of the PHC of distributions, see for example Cox and Oakes [10] in this respect.

2.2 Piecewise Constant Hazard Function

It is assumed that the base line hazard function $h_0(t)$ is piecewise constant, and $M-1$ number of cut points $0 = \tau_0 < \tau_1 < \ldots \tau_{M-1} < \tau_M = \infty$ to be used. At the beginning both M and the cut points are assumed to be known. Later we will discuss how to estimate these constants. Further, it is assumed that $h_0(t)$ has the following form;

$$h_0(t) = \sum_{j=1}^{M} \theta_j I_{[\tau_{j-1}\tau_j)}(t), \tag{3}$$

here θ_j's are unknown positive constants and $I_{[\tau_{j-1}\tau_j)}(t)$ is the indicator function, i.e.

$$I_{[\tau_{j-1}\tau_j)}(t) = \begin{cases} 1 & t \in [\tau_{j-1}, \tau_j) \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that under this set up $\theta_1, \ldots, \theta_M$ are identifiable only up to a multiplicative constant. Hence, without loss of generality, it is assumed that $\theta_M = 1$. It may be observed that when $M = 1$, $h_0(t)$ becomes an exponential hazard. Note that the distribution function having the piecewise constant hazard function as defined in (3) is a very flexible class of distribution functions. The hazard function can take variety of shapes, like increasing, decreasing, bathtub, unimodal shaped etc. One does not need to assume apriori the shape of the hazard function. It is more like a data driven technique, the data will decide about the shape of the fitted distribution. On the other hand it is not a purely non-parametric approach, where one needs a very large sample size to estimate the hazard as well as the survival functions. Therefore, it is more like a semi-parametric approach which has been used effectively to estimate the hazard and survival functions with a reasonable sample size.

Assumption of the piecewise constant hazard function for analyzing lifetime data is not very uncommon in the statistical literature, see for example Murphy [22], Goodman et al. [13], Balakrishnan et al. [6], Kundu [17] and references cited therein.

Now let us compute $S_0(t)$ and $f_0(t)$ which will be be needed later. Note that, the cumulative baseline hazard function $H_0(t)$ can be written as

$$H_0(t) = \sum_{j=1}^{M} \theta_j(\min\{t, \tau_j\} - \tau_{j-1}) I_{[\tau_{j-1}, \tau_M)}(t).$$

From now on we will use the notation $A_j = [\tau_{j-1}, \tau_j)$, for $j = 1, \ldots, M$. The cumulative hazard function

$$H_0(t) = \begin{cases} \theta_1 t & \text{if} \quad t \in A_1 \\ \theta_1\tau_1 + \theta_2(t - \tau_1) & \text{if} \quad t \in A_2 \\ \vdots & \vdots \quad \vdots \\ \sum_{j=1}^{k-1} \theta_j(\tau_j - \tau_{j-1}) + \theta_k(t - \tau_{k-1}) & \text{if} \quad t \in A_k \\ \vdots & \vdots \quad \vdots \\ \sum_{j=1}^{M-1} \theta_j(\tau_j - \tau_{j-1}) + \theta_M(t - \tau_{M-1}) & \text{if} \quad t \in A_M. \end{cases} \tag{4}$$

Once $H_0(t)$ is obtained, $S_0(t)$ and $f_0(t)$ can be obtained as

$$S_0(t) = e^{-H_0(t)} \quad \text{and} \quad f_0(t) = h_0(t)e^{-H_0(t)}. \tag{5}$$

The mean and variance can be obtained by simple integration. We just provide the mean residual time $M(a) = E(T|T > a)$ which are quite useful in lifetime data analysis.

$$M(a) = E(T|T > a) = e^{H_0(a)} \left[\int_a^{\tau_k} t h_0(t) e^{-H_0(t)} dt + \int_{\tau_k}^{\infty} t h_0(t) e^{-H_0(t)} dt \right] = A + B(\text{say}),$$

where

$$\begin{aligned} B &= e^{H_0(a)} \sum_{m=k+1}^{M} c_m \int_{\tau_{m-1}}^{\tau_m} t e^{-\sum_{j=1}^{M} c_j(\min\{t,\tau_j\} - \tau_{j-1}) I_{[\tau_{j-1},\tau_M)}(t)} dt \\ &= e^{H_0(a)} \sum_{m=k+1}^{M} c_m \int_{\tau_{m-1}}^{\tau_m} t e^{-\sum_{j=1}^{m} c_j(\min\{t,\tau_j\} - \tau_{j-1})} dt \\ &= e^{H_0(a)} \sum_{m=k+1}^{M} c_m e^{-\sum_{j=1}^{m-1} c_j(\tau_j - \tau_{j-1})} \int_{\tau_{m-1}}^{\tau_m} t e^{-c_m(t - \tau_{m-1})} dt \end{aligned}$$

$$
\begin{aligned}
&= e^{H_0(a)} \sum_{m=k+1}^{M} e^{-\sum_{j=1}^{m-1} c_j(\tau_j-\tau_{j-1})} \left[\tau_{m-1} + \frac{1}{c_m} - e^{-c_m(\tau_m-\tau_{m-1})}(\frac{1}{c_m} + \tau_m)\right]. \\
A &= e^{H_0(a)} c_k \int_a^{\tau_k} t e^{-\sum_{j=1}^{M} c_j(\min\{t,\tau_j\}-\tau_{j-1}) I_{[\tau_{j-1},\tau_M)}(t)} dt \\
&= e^{H_0(a)} c_k \int_a^{\tau_k} t e^{-\sum_{j=1}^{k} c_j(\min\{t,\tau_j\}-\tau_{j-1})} dt \\
&= e^{H_0(a)} c_k e^{-\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1})} \int_a^{\tau_k} t e^{-c_k(t-\tau_{k-1})} dt \\
&= e^{H_0(a)} e^{-\sum_{j=1}^{k-1} c_j(\tau_j-\tau_{j-1})} \left[a + \frac{1}{c_k} - e^{-c_k(\tau_k-a)}(\frac{1}{c_k} + \tau_k)\right] \\
&= e^{c_k(\tau_k-a)} \left[a + \frac{1}{c_k} - e^{-c_k(\tau_k-a)}(\frac{1}{c_k} + \tau_k)\right].
\end{aligned}
$$

3 Proportional Hazard Class Process

In this section we will define the PHC process (PHCP) and derive its properties.

Definition 1: Suppose $\{Y_n; n \geq 1\}$ is a sequence of independent and identically distributed (i.i.d.) random variables, Y_n follows($\sim$) PHC(S_0, λ_1) and $\{Z_n; n \geq 1\}$ is a sequence of i.i.d random variables independent of $\{Y_n; n \geq 1\}$ also and $Z_n \sim$ PHC(S_0, λ_2), for $n = 1, 2, \ldots$. Let us define $X_n = \min\{Y_n, Y_{n+1}, Z_n\}$, for $n = 1, 2, \ldots$. Then the sequence of random variables $\{X_n\}$ is called a PHCP. It will be denoted by PHCP$(S_0, \lambda_1, \lambda_2)$.

If $S_0(t) = e^{-t}$ for $t \geq 0$, then it becomes an exponential process and similarly if $S_0(t) = e^{-t^\alpha}$, then it becomes a Weibull process. Similarly, by choosing different $S_0(t)$, different processes can be obtained. From now let us use the following notations. $F_0(t) = 1 - S_0(t)$, $f_0(t)$, $h_0(t)$ and $H_0(t)$, the base line cumulative distribution function (CDF), the base line probability density function (PDF), the base line hazard function (HF) and the base line cumulative hazard function (CHF), respectively. The process $\{X_n\}$ is called the PHCP due to the following theorem.

Theorem 1: If $\{X_n\}$ is a PHCP$(S_0, \lambda_1, \lambda_2)$, then

(i) $\{X_n\}$ is a stationary process.

(ii) $X_n \sim \text{PHC}(S_0, 2\lambda_1 + \lambda_2)$

Proof: Part (i) follows from the definition. To prove part (ii), note that

$$P(X_n > t) = P(Y_n > t, Y_{n+1} > t, Z_n > t) = P(Y_n > t)P(Y_{n+1} > t)P(Z_n > t) = (S_0(t))^{2\lambda_1+\lambda_2}. \blacksquare$$

Theorem 2: If $\{X_n\}$ is a PHCP$(S_0, \lambda_1, \lambda_2)$, then the joint survival function of X_n and X_{n+m}, $S_{n,n+m}(x, y) = P(X_n > x, X_{n+m} > y)$, for $z = \max\{x, y\}$ is

$$S_{n,n+m}(x,y) = \begin{cases} (S_0(x))^{2\lambda_1+\lambda_2}(S_0(y))^{2\lambda_1+\lambda_2} & \text{if} \quad m \geq 2 \\ (S_0(x))^{\lambda_1+\lambda_2}(S_0(y))^{\lambda_1+\lambda_2}(S_0(z))^{\lambda_1} & \text{if} \quad m = 1. \end{cases}$$

Proof: Simply writing X_n in terms of Y_n,Y_{n+1},Z_n and similarly, X_{n+m}, the result follows. $\blacksquare$.

The above result indicates that X_n and X_{n+m} are independent if $m > 1$, and they are dependent if $m = 1$. This makes the process to be a lag-1 process. The joint distribution function of X_n and X_{n+1} will provide the dependence structure of the proposed PHCP. We will explore more later about the dependence properties of the PHCP. The joint survival function of X_n and X_{n+1} can be written explicitly as follows:

$$S_{n,n+m}(x,y) = \begin{cases} (S_0(x))^{\lambda_1+\lambda_2}(S_0(y))^{2\lambda_1+\lambda_2} & \text{if} \quad x < y \\ (S_0(x))^{2\lambda_1+\lambda_2}(S_0(y))^{\lambda_1+\lambda_2} & \text{if} \quad x \geq y. \end{cases} \tag{6}$$

It may be seen that (6) becomes the joint survival function of the Marshall-Olkin bivariate exponential (MOBE) distribution if we take $S_0(x) = e^{-x}$ and similarly, (6) becomes the joint survival function of the Marshall-Olkin bivariate Weibull (MOBW) distribution if we take $S_0(x) = e^{-x^\alpha}$, see for example Kundu and Dey [19] or Kundu and Gupta [21]. From Theorem 2, it follows that $\{X_n\}$ has a Markov property, i.e. the distribution of $X_n|\{X_{n-1}, \ldots, X_1\}$ is same as the distribution of $X_n|X_{n-1}$, since X_n is independent of $X_{n-2}, \ldots, X_1$.

It can be easily verified that X_n and X_{n+1} has the following Marshall-Olkin survival copula for $0 < \delta < 1$,

$$\widetilde{C}(u,v) = \begin{cases} u^{\delta}v & \text{if} \quad u \leq v \\ uv^{\delta} & \text{if} \quad u > v. \end{cases} \tag{7}$$

The corresponding copula density function becomes

$$c(u,v) = \frac{\partial^2}{\partial u \partial v}\widetilde{C}(u,v) = \begin{cases} \delta u^{\delta-1} & \text{if} \quad u \leq v \\ \delta v^{\delta-1} & \text{if} \quad u > v. \end{cases} \tag{8}$$

Several dependence measures and dependence properties can be obtained from this. For example the Kendal's τ and Spearman's ρ become

$$\tau = \frac{1-\delta}{1+\delta} \quad \text{and} \quad \rho = \frac{12}{(\delta+1)(\delta+3)} - 3.$$

Now we will provide the joint PDF of X_n and X_{n+1}, and because of the Markov property it will be helpful to compute the joint PDF of $X_1, \ldots, X_n$. It may be noted that the joint distribution function is not an absolutely continuous distribution function, hence the joint PDF does not exist in the usual sense, i.e. in terms of two dimensional Lebesgue measure, similar to MOBE or MOBW distributions. In this case the dominating measure is a two-dimensional Lebesgue measure on $x \neq y$ and one dimensional Lebesgue measure on $x = y$, see for example Bemis et al. [7]. Based on the above dominating measure the joint PDF of X_n and X_{n+1} can be written as follows.

$$f_{X_n,X_{n+1}}(x,y) = \begin{cases} f_1(x,y) & \text{if} \quad x < y \\ f_2(x,y) & \text{if} \quad x > y \\ f_0(x,y) & \text{if} \quad x = y, \end{cases} \tag{9}$$

where

$$\begin{aligned} f_1(x,y) &= f_{PHC}(x;\lambda_1+\lambda_2,S_0) f_{PHC}(y;2\lambda_1+\lambda_2,S_0) \\ f_2(x,y) &= f_{PHC}(x;2\lambda_1+\lambda_2,S_0) f_{PHC}(y;\lambda_1+\lambda_2,S_0) \\ f_0(x,x) &= \frac{\lambda_1}{3\lambda_1+2\lambda_2} f_{PHC}(x;3\lambda_1+2\lambda_2,S_0), \end{aligned}$$

and $f_0(x,y) = 0$ if $x \neq y$. If we denote

$$f_{ac}(x,y) = \frac{3\lambda_1+2\lambda_2}{2(\lambda_1+\lambda_2)} \begin{cases} f_{PHC}(x;\lambda_1+\lambda_2,S_0) f_{PHC}(x;2\lambda_1+\lambda_2,S_0) & \text{if} \quad 0 < x < y \\ f_{PHC}(x;2\lambda_1+\lambda_2,S_0) f_{PHC}(x;\lambda_1+\lambda_2,S_0) & \text{if} \quad 0 < y < x, \end{cases}$$

then $f_{ac}(x, y)$ is a proper bivariate PDF. The joint PDF of X_n and X_{n+1} can be written as

$$f_{X_n,X_{n+1}}(x, y) = \frac{2\lambda_1 + 2\lambda_2}{3\lambda_1 + 2\lambda_2} f_{ac}(x, y) + \frac{\lambda_1}{3\lambda_1 + 2\lambda_2} f_0(x, y).$$

Hence, $P(X_n = X_{n+1}) = \dfrac{\lambda_1}{3\lambda_1 + 2\lambda_2}$, and it varies between $(0, 1/3)$. Therefore, if the number of cases $X_n = X_{n+1}$ is more than one third of the total sample size, this model should not be used. Further, $P(X_n \neq X_{n+1}) = \dfrac{2(\lambda_1 + \lambda_2)}{3\lambda_1 + 2\lambda_2}$ and because of symmetry, $P(X_n < X_{n+1}) = P(X_n > X_{n+1}) = \dfrac{\lambda_1 + \lambda_2}{3\lambda_1 + 2\lambda_2}$. These probabilities are useful in estimating λ_1 and λ_2 also.

The following quantity might be of interest for some practical purposes

$$P(X_{n+1} > a | X_n = a) = \frac{P(X_{n+1} > a, X_n = a)}{P(X_n = a)} = \frac{(\lambda_1 + \lambda_2)}{(2\lambda_1 + \lambda_2)} (S_0(a))^{\lambda_1+\lambda_2}.$$

The following conditional density function is needed to compute the joint PDF of $X_1, \ldots, X_n$.

$$f_{X_{k+1}|X_k=x}(y) = \frac{f_{k,k+1}(x, y)}{f_{X_k}(x)} = \begin{cases} (\lambda_1 + \lambda_2)(S_0(x))^{-\lambda_1} h_0(y)(S_0(y))^{2\lambda_1+\lambda_2} & \text{if} \quad x < y \\ (\lambda_1 + \lambda_2) h_0(y)(S_0(y))^{\lambda_1+\lambda_2} & \text{if} \quad x > y \\ \frac{\lambda_1}{2\lambda_1+\lambda_2} h_0(x)(S_0(x))^{\lambda_1+\lambda_2} & \text{if} \quad x = y. \end{cases}$$

The joint PDF of $\{X_1, \ldots, X_n\}$ can be obtained as

$$\begin{aligned} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= f_{X_n|X_{n-1},\ldots,X_1}(x_n) \times \ldots \times f_{X_2|X_1}(x_2) \times f_{X_1}(x_1) & (10) \\ &= f_{X_n|X_{n-1}}(x_n) \times \ldots \times f_{X_2|X_1}(x_2) \times f_{X_1}(x_1) & (11) \\ &= \frac{f_{X_{n-1},X_n}(x_{n-1}, x_n)}{f_{X_{n-1}}(x_{n-1})} \times \ldots \times \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)} \times f_{X_1}(x_1) & (12) \\ &= \frac{\prod_{i=1}^{n-1} f_{X_i,X_{i+1}}(x_i, x_{i+1})}{\prod_{i=2}^{n-1} f_{X_i}(x_i)}. & (13) \end{aligned}$$

In this case (10) is obtained by the conditioning approach, (11) is obtained by using the Markov property, (12) is obtained due to the conditional density function and finally the last step by simple algebraic calculation.

Another important concept in any stationary time series is 'stopping time'. It has been discussed quite extensively in the time series literature. Christensen [8] discussed about optimal stopping time for autoregressive processes and Novikov and Shiryaev [23] also discussed about the optimal stopping time for processes with independent increment. Let us define the stopping time N for the PHCP $\{X_n\}$, for a given $L > 0$ as follows;

$$\{N = n\} \Leftrightarrow \{X_1 > L, \ldots, X_{n-1} > L, X_n < L\}.$$

Since the event $\{N = n\}$ is independent of $X_{n+1}, X_{n+2}, \ldots$, for all $n \geq 1$, therefore, N is a stopping time. To compute the probability mass function (PMF) of N, first let us compute the following

$$\begin{aligned} P(X_1 > L, \ldots, X_n > L) &= P(Y_1 > L, Y_2 > L, Z_1 > L, \ldots, Y_n > L, Y_{n+1} > L, Z_n > L) \\ &= (S_0(L))^{(n+1)\lambda_1 + n\lambda_2}. \end{aligned}$$

Hence,

$$\begin{aligned} P(N = 1) &= P(X_1 \leq L) = 1 - P(X_1 > L) = 1 - (S_0(L))^{2\lambda_1 + \lambda_2} \\ P(N = k) &= P(X_1 > L, X_2 > L, \ldots, X_k \leq L) \\ &= P(X_1 > L, \ldots, X_{k-1} > L) - P(X_1 > L, \ldots, X_k > L) \\ &= (S_0(L))^{k\lambda_1 + (k-1)\lambda_2} - (S_0(L))^{(k+1)\lambda_1 + k\lambda_2} = (S_0(L))^{k\lambda_1 + (k-1)\lambda_2}(1 - (S_0(L))^{\lambda_1 + \lambda_2}). \end{aligned}$$

If we denote $p = S_0(L)$, the probability generating function of N can be obtained as

$$G(z) = E(Z^N) = \sum_{k=1}^{\infty} P(N = k) z^k = \frac{zp^{\lambda_1}[1 - p^{\lambda_1 + \lambda_2}]}{1 - zp^{\lambda_1 + \lambda_2}}.$$

Hence

$$E(N) = \frac{p^{\lambda_1}}{1 - p^{\lambda_1 + \lambda_2}} \quad \text{and} \quad V(N) = \frac{p^{\lambda_1}(1 - p^{2(\lambda_1 + \lambda_2)} - p^{\lambda_1}(1 - p^{\lambda_1 + \lambda_2}))}{(1 - p^{\lambda_1 + \lambda_2})^3}.$$

4 Data Description

4.1 Data Set 1: Gold Price Data

This data set represents the price of 1 gm. of 24 Carat (Kt) gold in Indian Rupees From Dec. 01, 2020 till Jan. 31, 2021 of total 62 days. The price has been plotted in Figure 1 and it is presented in Appendix B. The minimum, maximum and the median value of the price

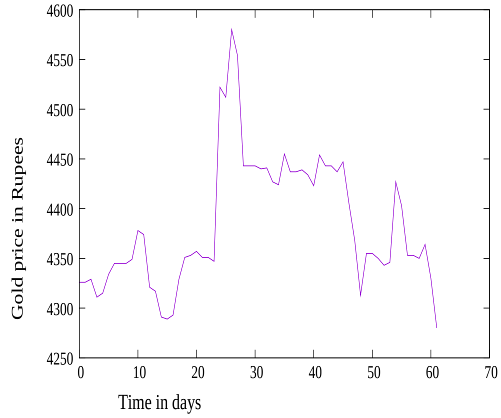
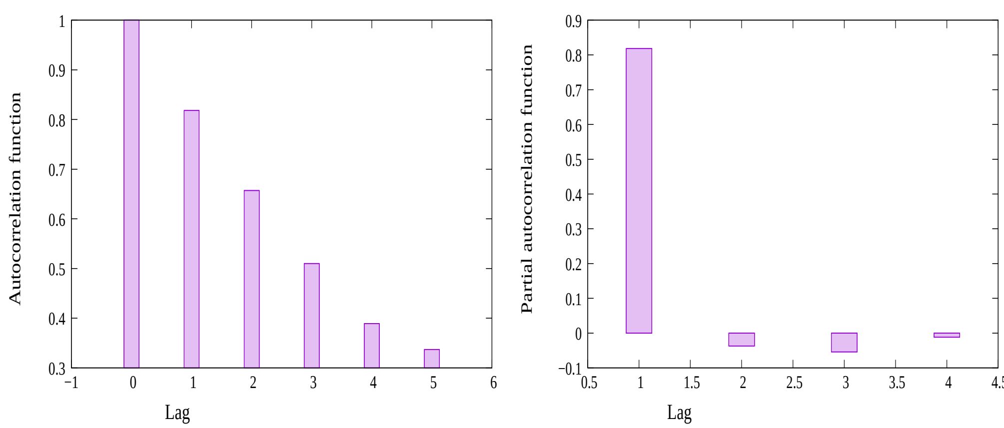


Figure 1: The price of 1 gm of 24 Carat gold in Indian Rupees from December 01, 2020 till January 31, 2021.

during this period is Rupees 4280, 4580 and 4355, respectively. There are 22, 29 and 10 cases such that $X_n < X_{n+1}$, $X_n > X_{n+1}$ and $X_n = X_{n+1}$, respectively. The autocorrelation (ACF) and the partial autocorrelation (PACF) functions of the data have been plotted in Figures 2. Now to check whether the observations of the lag-1 difference or lag-2 difference



(a) (b)

Figure 2: The ACF and PACF of the gold price data have been plotted in (a) and (b), respectively.

are independent distributed or not, the run test have been performed. The results for the

No.	Series	No. of runs	No. of '+' sign	No. of '-' sign	Test Statistic	p value
1	$\{X_1, X_3, \ldots,\}$	17	15	15	0.3716	0.7102
2	$\{X_2, X_4, \ldots,\}$	16	13	17	0.1095	0.9130
3	$\{X_1, X_4, \ldots,\}$	10	11	8	-0.1276	0.8984
4	$\{X_2, X_5, \ldots,\}$	12	7	12	1.0994	0.2716
5	$\{X_3, X_6, \ldots,\}$	10	10	9	-0.2243	0.8226

Table 1: Run test statistic and the associated p values for different series for the gold price data.

different series are presented in Table 1. It is clear from the PACF and also from the results in Table 1 that the observations of lag-1 series as well as lag-2 series are independently distributed.

4.2 Data Set 2: Exchange Rate Data

This data set represents the exchange rate of US one dollar against Indian rupees during April 05, 2021 to Sept 30, 2021. The exchange rate of US dollar one against Indian rupees has been plotted in Figure 3 and it is presented in Appendix B. The minimum, maximum and the

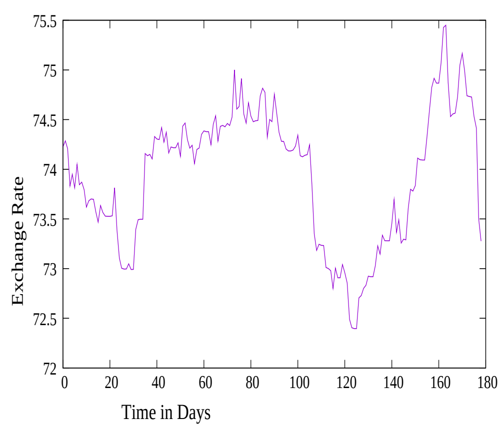


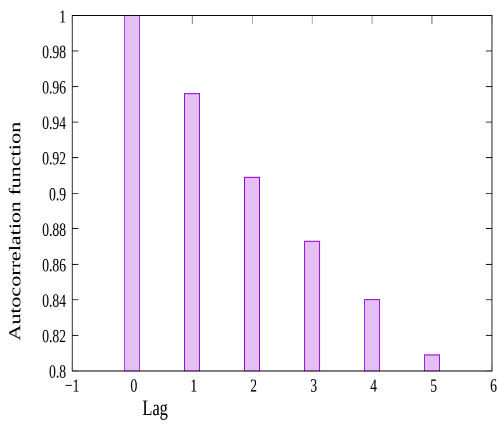
Figure 3: The exchange rate of US dollar one against Indian Rupees during April 05, 2021 to Sept 30, 2021.

median exchange rate during this period is Rupees 72.396, 75.450 and 74.139, respectively.

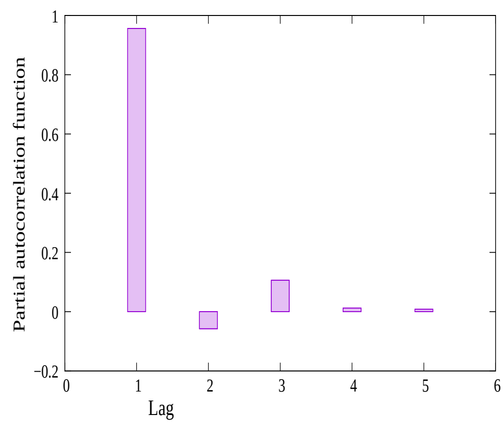
No.	Series	No. of runs	No. of '+' sign	No. of '-' sign	Test Statistic	p value
1	$\{X_1, X_3, \ldots, \}$	52	44	45	1.3872	0.1654
2	$\{X_2, X_4, \ldots, \}$	42	43	45	-0.6388	0.5230
3	$\{X_1, X_4, \ldots, \}$	30	35	24	0.1431	0.8862
4	$\{X_2, X_5, \ldots, \}$	30	28	31	-0.1116	0.9112
5	$\{X_3, X_6, \ldots, \}$	24	26	32	-1.5240	0.1276

Table 2: Run test statistic and the associated p values for different series for the exchange rate data.

There are 82, 85 and 11 cases so that $X_n < X_{n+1}$, $X_n > X_{n+1}$ and $X_n = X_{n+1}$, respectively. The autocorrelation (ACF) and the partial autocorrelation (PACF) functions of the data have been plotted in Figures 2. We have performed the run tests on the two lag-1 series and



(a)



(b)

Figure 4: The ACF and PACF of the exchange rate data have been plotted in (a) and (b), respectively.

three lag-2 series, and the results are presented in Table 2. In this case also it is clear that the observations of lag-1 series as well as lag-2 series are independently distributed.

5 Maximum Likelihood Estimators

5.1 Likelihood Function: General

In this section we provide the likelihood function of the observed random sample $\mathcal{D} = \{x_1, \ldots, x_n\}$ from the PHCP$(S_0, \lambda_1, \lambda_2)$. We consider three different forms of the base line distribution, namely (a) exponential, i.e. $S_0(t) = e^{-t}$, (b) Weibull, i.e. $S_0(t) = e^{-t^\alpha}$, and (c) piecewise constant hazard function as defined in (5). First we will provide the general form of the likelihood function, and then consider the three different cases explicitly. We use the following notations in all these cases. $I = \{1, \ldots, n-1\}$, $I_{-1} = \{2, \ldots, n-1\}$,

$$I_1 = \{i : i \in I, x_i < x_{i+1}\}, \quad I_2 = \{i : i \in I, x_i > x_{i+1}\}, \quad I_0 = \{i : i \in I, x_i = x_{i+1}\}.$$

The number of elements in I_0, I_1 and I_2 are denoted by n_0, n_1 and n_2,respectively. Based on the observed sample the likelihood function can be written as

$$L(\lambda_1, \lambda_2, \Theta|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}}{(2\lambda_1+\lambda_2)^{n_0-1}} h_0(x_1;\Theta) e^{-\lambda_1 E_1(\mathcal{D},\Theta) - \lambda_2 E_2(\mathcal{D},\Theta)} \prod_{i \in I_1 \cup I_2} h_0(x_{i+1};\Theta), \quad (14)$$

where

$$\begin{aligned} E_1(\mathcal{D},\Theta) &= 2\sum_{i\in I_1} H_0(x_{i+1}) + \sum_{i\in I_2} H_0(x_{i+1}) + \sum_{i\in I_0} H_0(x_i) - \sum_{i\in I_1} H_0(x_i) + 2H_0(x_1) \\ E_2(\mathcal{D},\Theta) &= \sum_{i\in I_1\cup I_2} H_0(x_{i+1}) + \sum_{i\in I_0} H_0(x_i) + H_0(x_1). \end{aligned}$$

5.2 Likelihood Function: Exponential and Weibull

In case of exponential distribution when $S_0(x) = e^{-x}$, the likelihood function can be written as

$$L_{EX}(\lambda_1, \lambda_2|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}}{(2\lambda_1+\lambda_2)^{n_0-1}} e^{-\lambda_1 E_1(\mathcal{D})} e^{-\lambda_2 E_2(\mathcal{D})}, \quad (15)$$

where

$$\begin{aligned} E_1(\mathcal{D}) &= 2\sum_{i\in I_1} x_{i+1} + \sum_{i\in I_2} x_{i+1} + \sum_{i\in I_0} x_i - \sum_{i\in I_1} x_i + 2x_1 \\ E_2(\mathcal{D}) &= \sum_{i\in I_1\cup I_2} x_{i+1} + \sum_{i\in I_0} x_i + x_1. \end{aligned}$$

In case of Weibull distribution, when $S_0(x) = e^{-x^\alpha}$, the likelihood function can be written as

$$L_{WE}(\lambda_1, \lambda_2, \alpha|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}\alpha^{n_1+n_2+1}x_1^{\alpha-1}}{(2\lambda_1+\lambda_2)^{n_0-1}} e^{-\lambda_1 E_1(\mathcal{D};\alpha)-\lambda_2 E_2(\mathcal{D};\alpha)} \prod_{i\in I_1\cup I_2} x_{i+1}^{\alpha-1}, \quad (16)$$

where

$$\begin{aligned} E_1(\mathcal{D};\alpha) &= 2\sum_{i\in I_1} x_{i+1}^\alpha + \sum_{i\in I_2} x_{i+1}^\alpha + \sum_{i\in I_0} x_i^\alpha - \sum_{i\in I_1} x_i^\alpha + 2x_1^\alpha \\ E_2(\mathcal{D};\alpha) &= \sum_{i\in I_1\cup I_2} x_{i+1}^\alpha + \sum_{i\in I_0} x_i^\alpha + x_1^\alpha. \end{aligned}$$

5.3 Likelihood Function: PCH Function

First observe that in presence of λ_1 and λ_2, $\theta_1, \ldots, \theta_M$ are identifiable up to a multiplicative constant. Without loss of generality, it is assumed that $\theta_1 + \ldots + \theta_M = 1$. Further, the following notations might be useful to write the likelihood function in a compact manner. For any $i \in I_1 \cup I_2$, the corresponding (x_i, x_{i+1}) will be denoted as follows: x_i will be denoted by a_{uvw}, here $u = j$, if $i \in I_j$, for j = 1,2, the index v indicates $x_i \in A_v$, for $v = 1, \ldots, M$ and $w \in \{1, \ldots, n_{1uv}\}$, here n_{1uv} denotes the number of x_i's in A_v, for $i \in I_u$. Similarly, y_i will be denoted by b_{uvw}, here $u = j$, if $i \in I_j$, for j = 1,2, the index v indicates $x_{i+1} \in A_v$, for $v = 1, \ldots, M$ and $w \in \{1, \ldots, n_{2uv}\}$, here n_{2uv} denotes the number of x_{i+1}'s in A_v, for $i \in I_u$. For any $i \in I_0$, the corresponding x_i will be denoted by c_{vw}, here the index v indicates $x_i \in A_v$, for $v = 1, \ldots, M$ and $w \in \{1, \ldots, n_{0v}\}$, here n_{0v} denotes the number of x_i's in A_v, for $i \in I_0$. Finally we will denote $n_{3m} = \#\{i : i \in I_{-1}, x_i \in A_m\}$, $m = 1, \ldots, M$. It is assumed that $x_1 \in A_k$. Based on the observed sample the likelihood function in this case can be written as

$$L(\lambda_1, \lambda_2, \Theta|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}\theta_k}{(2\lambda_1+\lambda_2)^{n_0-1}} \left\{ \prod_{m=1}^{M} \theta_m^{n_{21m}+n_{22m}+n_{0m}} \right\} e^{\left\{-\lambda_1 \sum_{j=1}^{M} C_j\theta_j - \lambda_2 \sum_{j=1}^{M} D_j\theta_j\right\}}, \tag{17}$$

where for $j < r$,

$$\begin{aligned} C_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (2n_{21m} + n_{22m} + n_{0m} - n_{11m} + 2) \\ &\quad +2\sum_{k=1}^{n_{21j}}(b_{1jk} - \tau_{j-1}) + \sum_{k=1}^{n_{22j}}(b_{2jk} - \tau_{j-1}) + \sum_{k=1}^{n_{0j}}(c_{jk} - \tau_{j-1}) - \sum_{k=1}^{n_{11j}}(a_{1jk} - \tau_{j-1}) \\ D_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (n_{21m} + n_{22m} + n_{0m} + 1) \\ &\quad +\sum_{k=1}^{n_{21j}}(b_{1jk} - \tau_{j-1}) + \sum_{k=1}^{n_{22j}}(b_{2jk} - \tau_{j-1}) + \sum_{k=1}^{n_{0j}}(c_{jk} - \tau_{j-1}) \end{aligned}$$

for $j = r$,

$$\begin{aligned} C_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (2n_{21m} + n_{22m} + n_{0m} - n_{11m} + 2) \\ &\quad +2\sum_{k=1}^{n_{21j}}(b_{1jk} - \tau_{j-1}) + \sum_{k=1}^{n_{22j}}(b_{2jk} - \tau_{j-1}) + \sum_{k=1}^{n_{0j}}(c_{jk} - \tau_{j-1}) - \sum_{k=1}^{n_{11j}}(a_{1jk} - \tau_{j-1}) + 2(x_1 - \tau_{j-1}) \\ D_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (n_{21m} + n_{22m} + n_{0m} + 1) \\ &\quad +\sum_{k=1}^{n_{21j}}(b_{1jk} - \tau_{j-1}) + \sum_{k=1}^{n_{22j}}(b_{2jk} - \tau_{j-1}) + \sum_{k=1}^{n_{0j}}(c_{jk} - \tau_{j-1}) + (x_1 - \tau_{j-1}) \end{aligned}$$

and for $j > r$,

$$\begin{aligned} C_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (2n_{21m} + n_{22m} + n_{0m} - n_{11m}) \\ &\quad +2\sum_{k=1}^{n_{21j}}(b_{1jk} - \tau_{j-1}) + \sum_{k=1}^{n_{22j}}(b_{2jk} - \tau_{j-1}) + \sum_{k=1}^{n_{0j}}(c_{jk} - \tau_{j-1}) - \sum_{k=1}^{n_{11j}}(a_{1jk} - \tau_{j-1}) \\ D_j &= (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} (n_{21m} + n_{22m} + n_{0m}) \end{aligned}$$

$$+\sum_{k=1}^{n_{21j}}(b_{1jk}-\tau_{j-1})+\sum_{k=1}^{n_{22j}}(b_{2jk}-\tau_{j-1})+\sum_{k=1}^{n_{0j}}(c_{jk}-\tau_{j-1}).$$

See Appendix A for details.

5.4 MLEs: Exponential and Weibull

In this section we will discuss about the maximum likelihood estimators of the unknown parameters for the exponential and Weibull processes. Based on the likelihood function (15), the log-likelihood function for the exponential process can be written as

$$l_{EX}(\lambda_1,\lambda_2|\mathcal{D}) = n_0\ln\lambda_1+(n_1+n_2)\ln(\lambda_1+\lambda_2)-(n_0-1)\ln(2\lambda_1+\lambda_2)-\lambda_1E_1(\mathcal{D})-\lambda_2E_2(\mathcal{D}) \quad (18)$$

where $E_1(\mathcal{D})$ and $E_2(\mathcal{D})$ are same as defined in (15). We make the following reparametrization $\beta=\lambda_1$ and $\beta\delta=\lambda_2$. Then the log-likelihood function can be written as

$$\begin{aligned} l_{EX}(\beta,\delta|\mathcal{D}) &= (n_1+n_2+1)\ln\beta+(n_1+n_2)\ln(1+\delta)-(n_0-1)\ln(2+\delta) \\ &\quad -\beta\left(E_1(\mathcal{D})+\delta E_2(\mathcal{D})\right). \end{aligned} \quad (19)$$

Hence, for given δ, the MLE of β, say $\widehat{\beta}(\delta)$ can be obtained as

$$\widehat{\beta}(\delta)=\frac{n_1+n_2+1}{E_1(\mathcal{D})+\delta E_2(\mathcal{D})}, \quad (20)$$

and the MLE of δ can be obtained by maximizing the profile log-likelihood function (without an additive constant) of δ, namely

$$l_{EX}(\widehat{\beta}(\delta),\delta|\mathcal{D}) = -(n_1+n_2+1)\ln(E_1(\mathcal{D})+\delta E_2(\mathcal{D}))+(n_1+n_2)\ln(1+\delta)-(n_0-1)\ln(2+\delta). \quad (21)$$

It can be easily seen that the $\widehat{\delta}$ which maximizes (21) can be obtained as the positive root of the following quadratic equation for $R=E_1(\mathcal{D})/E_2(\mathcal{D})$,

$$p_{EX}(\delta) = -n_0\delta^2+\delta(R(n_1+n_2+1-n_0)+(n_0-2))+R(2(n_1+n_2)-(n_0-1))-2(n_1+n+2+1)=0.$$

Once, the MLEs of β and δ are obtained, the MLEs of λ_1 and λ_2 can be obtained by transformation.

The log-likelihood function for the Weibull process without some additive constant, based on (16) can be written as

$$\begin{aligned} l_{WEB}(\lambda_1, \lambda_2, \alpha|\mathcal{D}) &= n_0 \ln \lambda_1 + (n_1+n_2)\ln(\lambda_1+\lambda_2) - (n_0-1)\ln(2\lambda_1+\lambda_2) + (n_1+n_2+1)\ln\alpha \\ &+\alpha\left(\sum_{i\in I_1\cup I_2} \ln x_{i+1} + \ln x_1\right) - \lambda_1 E_1(\mathcal{D};\alpha) - \lambda_2 E_2(\mathcal{D};\alpha). \end{aligned} \quad (22)$$

In this case also we make the same reparametrization as the exponential case: $\beta = \lambda_1$ and $\beta\delta = \lambda_2$. Then the log-likelihood function (22) can be written as

$$\begin{aligned} l_{WEB}(\beta, \delta, \alpha|\mathcal{D}) &= (n_1+n_2+1)\ln\beta + (n_1+n_2)\ln(1+\delta) - (n_0-1)\ln(2+\delta) \\ &+(n_1+n_2+1)\ln\alpha + \alpha\left(\sum_{i\in I_1\cup I_2} \ln x_{i+1} + \ln x_1\right) - \beta\left(E_1(\mathcal{D};\alpha) + \delta E_2(\mathcal{D};\alpha)\right). \end{aligned}$$

Hence, for given δ and α, the MLE of β, say $\widehat{\beta}(\delta,\alpha)$ can be obtained as

$$\widehat{\beta}(\delta,\alpha) = \frac{n_1+n_2+1}{E_1(\mathcal{D};\alpha) + \delta E_2(\mathcal{D};\alpha)}. \quad (23)$$

For given α, $\widehat{\delta}(\alpha)$, the MLE of δ, can be obtained as the unique positive root of

$$\begin{aligned} p_{WEB}(\delta|\alpha) &= -n_0\delta^2 + \delta(R(\alpha)(n_1+n_2+1-n_0) + (n_0-2)) \\ &+R(\alpha)(2(n_1+n_2) - (n_0-1)) - 2(n_1+n+2+1) = 0. \end{aligned}$$

Here $R(\alpha) = E_1(\mathcal{D};\alpha)/E_2(\mathcal{D};\alpha)$. The MLE of α, $\widehat{\alpha}$, is obtained by maximizing $g(\alpha)$, the profile log-likelihood of α without the additive constant, where

$$\begin{aligned} g(\alpha) &= -(n_1+n_2+1)\ln(E_1(\mathcal{D},\alpha) + \widehat{\delta}(\alpha)E_2(\mathcal{D},\alpha)) + (n_1+n_2)\ln(1+\widehat{\delta}(\alpha)) \\ &-(n_0-1)\ln(2+\widehat{\delta}(\alpha)) + (n_1+n_2+1)\ln\alpha + \alpha\left(\sum_{i\in I_1\cup I_2} \ln x_{i+1} + \ln x_1\right). \end{aligned}$$

5.5 MLEs: PCH Function

In this section first we will we discuss about the MLEs of the unknown parameters of the PCH function when M and $\tau_1, \ldots, \tau_M$ are known, and then we will discuss when they are unknown. The log-likelihood function of the PCH function process can be written as

$$\begin{aligned} l_{PCH}(\lambda_1, \lambda_2, \Theta|\mathcal{D}) &= (n_1+n_2)\ln(\lambda_1+\lambda_2) + n_0 \ln \lambda_1 - (n_0-1)\ln(2\lambda_1+\lambda_2) + \sum_{m=1}^{M} d_m \ln \theta_m \\ &\quad -\lambda_1 \sum_{j=1}^{M} C_j \theta_j - \lambda_2 \sum_{j=1}^{M} D_j \theta_j, \end{aligned}$$

where $d_m = n_{21m} + n_{22m} + n_{0m}$, for $m = 1, \ldots, M$, $m \neq k$, $d_k = n_{21k} + n_{22k} + n_{0k} + 1$, C_j and D_j are same as defined in (17). For given λ_1 and λ_2, the MLEs of θ_m, can be obtained as

$$\widehat{\theta}_m(\lambda_1, \lambda_2) = \frac{d_m}{\lambda_1 C_m + \lambda_2 C_m}; \quad m = 1, \ldots, M-1. \tag{24}$$

Hence, the MLEs of λ_1 and λ_2 can be obtained by maximizing the profile log-likelihood function $p_{PCH}(\lambda_1, \lambda_2) = l_{PCH}(\lambda_1, \lambda_2, \widehat{\Theta}(\lambda_1, \lambda_2)|\mathcal{D})$.

5.6 Estimation of Cut Points

Now we discuss about the estimation of the unknown parameters of the PCH function when M and $\tau_1, \ldots, \tau_M$ are unknown. Let us denote the maximum time point as τ. It is assumed that the position of the cut points are modeled as the arrival times of the events in $[0, \tau]$, from a non-homogeneous Poisson process $\{N(t); t \geq 0\}$ with intensity function $u(t) = \delta t^{\alpha-1}$, for $\delta > 0$. But the arrival times of the events from $N(t)$ are not observed, hence they are being treated as missing values. For the given number of cut points $M-1$ in $[0, \tau]$, the likelihood function of the observation is treated as a function of the unknown parameters $\theta_1, \ldots, \theta_M, \delta$, when the arrival times of the events from $N(t)$, which are missing, are replaced by their corresponding expected values. For a given M we compute the maximum likelihood estimators of the unknown parameters, and then based on some information theoretic criteria, like Akike Information Criterion (AIC) or Bayes Information Criterion (BIC), we estimate M also.

Now we provide the joint PDF of the arrival times of a non-homogeneous Poisson process. Suppose $\{N(t); t \geq 0\}$ is a non-homogeneous Poisson process with intensity function $u(t) = \delta\lambda t^{\delta-1}$, for $\delta > 0$. Let $T_1, T_2, \ldots$ denote the arrival times of the first, second etc. arrival times of the events from $N(t)$. Then the joint PDF of $T_1, \ldots, T_n$, given that $N(\tau) = n$, can be written as

$$f_{T_1,\ldots,T_n|N(\tau)=n}(t_1, \ldots, t_n) = \frac{n!\delta^n}{\tau^{n\delta}} t_1^{\delta-1} \ldots t_n^{\delta-1}; \qquad 0 < t_1 < t_2 \ldots < t_n < \tau, \tag{25}$$

and zero, otherwise. Hence, for $j = 1, \ldots, n$,

$$\begin{aligned}
E(T_j|N(\tau) = n) &= \frac{n!\delta^n}{\tau^{n\delta}} \int_0^{t_2} \int_0^{t_3} \cdots \int_0^{t_n} \int_0^{\tau} t_1^{\delta-1} \ldots t_{j-1}^{\delta-1} t_j^{\delta} t_{j+1}^{\delta-1} \ldots t_n^{\delta-1} dt_1 dt_2 \ldots dt_{n-1} dt_n \\
&= \frac{n!\delta^n}{\tau^{n\delta}} \times \frac{\tau^{n\delta+1}}{\alpha(2\delta) \ldots ((j-1)\delta)(j\delta+1)((j+1)\delta+1) \ldots (n\delta+1)} \\
&= \frac{n!}{(j-1)!} \frac{\delta^{n-j+1}\tau}{(j\delta+1)((j+1)\delta+1) \ldots (n\delta+1)} \\
&= \frac{\tau}{(1+\frac{1}{j\delta}) \ldots + (1+\frac{1}{n\delta})}
\end{aligned}$$

Therefore, for homogeneous Poisson process i.e. when $\delta = 1$,

$$E(T_j|N(\tau) = n) = \frac{j\tau}{n+1}.$$

In this case the cut points are equidistant.

Hence, in practice we can implement the above method as follows. For fixed δ and M, we can write the the likelihood function in terms of the unknown parameters. Then for a given δ and M we can obtain the maximum likelihood estimators of the other parameters. We treat this as profile likelihood of δ and then obtain the maximum likelihood estimator of δ also for a given M. Then we can find M by using AIC or BIC.

6 Goodness of Fit

It is important to provide a goodness of test to verify whether the proposed model can be fitted to a given data set or not. We have considered a similar goodness of fit test as it has been proposed in Kundu [18]. Suppose $\{x_1, \ldots, x_n\}$ is a sample from a stationary sequence $\{X_1, \ldots, X_n\}$. We want to test the following null hypothesis

$$H_0: \quad \{X_1, \ldots, X_n\} \sim \text{PHCP}(S_0, \lambda_0, \lambda_1).$$

Here, S_0 is a specific known survival function, λ_0 and λ_1 are also known. Let us use the following notations. We denote $X_{1:n} < \ldots < X_{n:n}$ as the ordered $\{X_1, \ldots, X_n\}$, similarly, $x_{1:n} < \ldots < x_{n:n}$ as the ordered $\{x_1, \ldots, x_n\}$, and $a_1 = E_{H_0}(X_{1:n}), \ldots, a_n = E_{H_0}(X_{n:n})$ as their ordered expected values under H_0. Here, $a_1, \ldots, a_n$ depend on $S_0, \lambda_0, \lambda_1$, but we do not make it explicit. We use the following statistic for goodness of fit test.

$$W_n = \max_{1 \le i \le n} |X_{i:n} - a_i|.$$

It is expected that if H_0 is true, then W_n should be small. Hence, we use the following test criterion for a given level of significance $0 < \beta < 1$

$$\text{Reject } H_0 \text{ if } W_n > c_n(\beta),$$

where $c_n(\beta)$ is such that

$$P_{H_0}(W_n > c_n(\beta)) = \beta.$$

Note that $c_n(\beta)$ also depends on $S_0, \lambda_0, \lambda_1$, but we do not make it explicit for brevity. It is difficult to obtain $c_n(\beta)$ theoretically even for large n. Hence, we propose to use parametric bootstrap technique to approximate $c_n(\beta)$ from a given observed sample $\{x_1, \ldots, x_n\}$

Algorithm:

Step 1: Obtain $\widehat{S}_0, \widehat{\lambda}_0, \widehat{\lambda}_1$, the MLEs of $S_0, \lambda_0, \lambda_1$, respectively, based on $\{x_1, \ldots, x_n\}$. Note that in case of exponential $S_0(x) = \widehat{S}_0 = e^{-x}$. In case of Weibull or piecewise constant hazard function $\widehat{S}_0$ is obtained by replacing the parameters associated with the baseline distribution with their estimates.

Step 2: Generate a sample of size n from a WEP$(\widehat{S}_0, \widehat{\lambda}_0, \widehat{\lambda}_1)$, order them. Let us denote them as $(x^1_{1:n}, \ldots, x^1_{n:n})$. Repeat this procedure B times, and obtain $\{(x^b_{1:n}, \ldots, x^b_{n:n}); b = 1, \ldots, B\}$.

Step 3: Obtain estimates of $a_1, \ldots, a_n$ as

$$\widehat{a}_i = \frac{1}{B} \sum_{b=1}^{B} x^b_{i:n}; \qquad i = 1, \ldots, n.$$

Step 4: Compute

$$w^b = \max_{1 \le i \le n} \{|x^b_{i:n} - \widehat{a}_i|\}; \qquad b = 1, \ldots, B.$$

Step 5: Order $\{w^1, \ldots, w^B\}$ as $w^{(1)} < \ldots < w^{(B)}$, then $\widehat{c}_n(\beta) = w^{[(100(1-\beta))]}$ is an estimate of $c_n(\beta)$.

Hence, if $w_n = \max\limits_{1 \le i \le n} |x_{i:n} - \widehat{a}_i| > \widehat{c}_n(\beta)$, then we reject the null hypothesis with $\beta\%$ level of significance, otherwise we accept the null hypothesis.

7 Data Analysis

Data Set 1: Gold Price Data

In this section we have fitted the exponential model, Weibull model and PCH model with different cut points to the gold price data set. It has been observed by several authors, see for example Balakrishnan et al. [6], Goodman et al. [13], Samanta and Kundu [25], Ganguly et al. [12] that in case of PCH model, usually $M = 3$ or 4 are sufficient. Hence, here we have used PCH model with $M =$ 2,3 and 4. Note that when $M = 1$, it becomes an exponential distribution. We have computed the maximum likelihood estimators in each case. In case of the PCH model we have computed the cut points also as mentioned in the previous section.

Note that MLEs of the exponential parameters can be obtained in explicit forms. In case of Weibull distribution they can be obtained by solving a one-dimensional optimization problem. For the PCH model, for a given α, the MLEs of the unknown parameters can be obtained by solving a two-dimensional optimization problem, and then the estimate of α, hence the cut points, can be obtained by grid search method. The MLEs, the corresponding log-likelihood values, and the associated Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values are also reported in Table 3.

Process $\longrightarrow$ Parameters $\downarrow$	Exponential	Weibull	PCH-1CP	PCH-2CP	PCH-3CP
λ_1	15.1059	13.4754	75.5229	83.1960	87.0856
λ_2	24.1695	21.5607	120.8367	133.1136	139.3370
α	—-	2.9891	—-	—-	—-
θ_1	—-	—-	0.1393	0.0498	0.0208
θ_2	—-	—-	—-	0.3638	0.1624
θ_3	—-	—-	—-	—-	0.3316
τ_1	—-	—-	4425	4333	4299
τ_2	—-	—-	—	4427	4341
τ_3	—-	—-	—-	—-	4425
δ	——	——	1.345	0.579	0.482
log-lik	-346.419	-326.532	-339.231	- 321.754	- 317.452
AIC	696.838	659.064	684.462	651.508	644.904
BIC	701.092	665.445	690.843	660.016	655.539

Table 3: The estimates of the unknown parameters and the associated 95% confidence intervals, estimation of the cut points for PCH model and the corresponding log-likelihood (log-lik) values in case of Data Set 1.

Now based on the AIC and BIC values we conclude that among these the PCH process with three cut points is the best fitted model. Now for the given cut points, based on bootstrap method, we have obtained 95% confidence intervals of λ_1, λ_2, θ_1, θ_2, θ_3 and δ as follows: (75.2345,96.3498), (125.6923,154.9876), (0.0156,0.0269), (0.1487,0.1854), (0.2786,0.3944), (0.397,0.592), respectively. One natural question is whether the above model fits the data or not. For that we have used the test statistic as proposed in Section 6. We

$w_n = 0.0298$, and based on simulation the associated p value lies in $(0.80, 0.85)$. Hence, it indicates that the proposed model provides a good fit to the data.

Data Set 2: Exchange Range Data

In this case also as before we have fitted same set of distributions to this data set also, and the results are presented in Table 4. Similarly, as before the AIC, BIC values, and the KS distances and the associated p values of the odd and even series data to the fitted distributions have been presented in Table ??. In this case also it has been concluded that the PCH model with three cut points fit the exchange rate data quite well. Further, we have obtained the 95% confidence intervals of λ_1, λ_2, θ_1, θ_2, θ_3 and δ as follows: (24.3451,38.6751), (212.876,230.542), (0.0123,0.0169), (0.0823,0.1256), (0.1331,0.1999), (0.411,0.482), respectively. In this case w_n = 0.0423 and the associated p value lies in (0.35,0.40). Hence, in this case also the proposed model provides a good fit to the data.

Process $\longrightarrow$ Parameters $\downarrow$	Exponential	Weibull	PCH-1CP	PCH-2CP	PCH-3CP
λ_1	6.0081	5.3695	27.5188	29.4074	31.2742
λ_2	42.3299	37.8310	195.1336	208.5253	221.7627
α	—-	3.0926	—-	—-	—-
θ_1	—-	—-	0.1091	0.0523	0.0147
θ_2	—-	—-	—-	0.1658	0.1006
θ_3	—-	—-	—-	—-	0.1678
τ_1	—-	—-	74.0555	73.0451	72.6261
τ_2	—-	—-	—	74.0812	73.1917
τ_3	—-	—-	—-	—-	74.1149
δ	——	——	1.327	0.524	0.441
log-lik	-302.1141	-288.676	-293.342	-282.251	-279.222
AIC	608.228	583.352	592.684	572.502	568.444
BIC	614.711	593.077	602.409	585.468	584.653

Table 4: The estimates of the unknown parameters and the associated 95% confidence intervals, estimation of the cut points for PCH model and the corresponding log-likelihood (log-lik) values in case of Data Set 2.

8 Conclusions

In this paper we have proposed a new stationary process whose marginals follow proportional hazard class model. There are two important features of this process. First of all we have kept the base line distribution to be a quite flexible one. It have been assumed that the base line distribution has piece wise constant hazard function, which makes the process to be a very flexible one. Estimating the cut points of the piece wise constant hazard function is an important issue, and it has not been addressed in the literature properly. We have provided the method of estimating the cut points based on non-homogeneous Poisson process approach. Another important feature of the proposed process is that there is a positive probability that two consecutive values might be equal. We have provided different properties of the proposed process, and discussed the MLEs. Two real life data sets have been analyzed, and the proposed model fits both the data sets quite well. It may be mentioned that although we have developed this model for lag-1 process, but it can be generalized to a general lag-p process also, more work is needed in that direction.

Acknowledgements

The author would like to thank the reviewer and the Associate editor for many constructive suggestions which have helped to improve the earlier version of the paper significantly.

Funding Declaration: No funding has been received to prepare this manuscript.

Competing Interest: There is no competing interest to prepare this manuscript.

Availability of Data: The data are available in Appendix B.

Appendix A: Log-Likelihood Function

$$l(\lambda_1, \lambda_2, \mathbf{\Theta}|\mathcal{D}) = \sum_{i \in I_0 \cup I_1 \cup I_2} \ln f_{X_i, X_{i+1}}(x_i, x_{i+1}) - \sum_{i=2}^{n-1} \ln f_{X_i}(x_i).$$

Now

$$\begin{aligned}
\sum_{i \in I_1} \ln f_{X_i, X_{i+1}}(x_i, x_{i+1}) &= n_1 \ln(\lambda_1 + \lambda_2) + n_1 \ln(2\lambda_1 + \lambda_2) + \sum_{m=1}^{M} (n_{11m} + n_{21m}) \ln \theta_m - \\
&\quad (\lambda_1 + \lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{11m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{j=1}^{n_{11m}} (a_{1mj} - \tau_{m-1}) \right\} \right] - \\
&\quad (2\lambda_1 + \lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{21m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{j=1}^{n_{21m}} (b_{1mj} - \tau_{m-1}) \right\} \right]
\end{aligned}$$

$$\begin{aligned}
\sum_{i \in I_2} \ln f_{X_i, X_{i+1}}(x_i, x_{i+1}) &= n_2 \ln(\lambda_1 + \lambda_2) + n_2 \ln(2\lambda_1 + \lambda_2) + \sum_{m=1}^{M} (n_{12m} + n_{22m}) \ln \theta_m - \\
&\quad (2\lambda_1 + \lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{12m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{j=1}^{n_{12m}} (a_{2mj} - \tau_{m-1}) \right\} \right] - \\
&\quad (\lambda_1 + \lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{22m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{j=1}^{n_{22m}} (b_{2mj} - \tau_{m-1}) \right\} \right]
\end{aligned}$$

$$\begin{aligned}
\sum_{i \in I_0} \ln f_{X_i, X_{i+1}}(x_i, x_{i+1}) &= n_0 \ln \lambda_2 - n_0 \ln(3\lambda_1 + 2\lambda_2) + \sum_{m=1}^{M} n_{0m} \ln \theta_m - \\
&\quad (3\lambda_1 + 2\lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{0m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{j=1}^{n_{0m}} (c_{mj} - \tau_{m-1}) \right\} \right].
\end{aligned}$$

$$\begin{aligned}
\sum_{i \in I_{-1}} \ln f_{X_i}(x_i) &= (n-2) \ln(2\lambda_1 + \lambda_2) + \sum_{m=1}^{M} n_{3m} \ln \theta_m - \\
&\quad (2\lambda_1 + \lambda_2) \left[\sum_{m=1}^{M} \left\{ n_{3m} \sum_{j=1}^{m-1} \theta_j (\tau_j - \tau_{j-1}) + \theta_m \sum_{i: x_i \in I_{-1} \cap A_m} (x_i - \tau_{m-1}) \right\} \right].
\end{aligned}$$

Let us denote for $j = 1, \ldots, M$, when $\displaystyle\sum_{m=M+1}^{M} (\cdot) = 0$.

$$C_j = (\tau_j - \tau_{j-1}) \left(\sum_{m=j+1}^{M} (n_{11m} + 2n_{21m} + 2n_{12m} + n_{22m} + 3n_{0m} - 2n_{3m}) \right)$$

$$+\sum_{k=1}^{n_{11j}}(a_{1jk}-\tau_{j-1})+2\sum_{k=1}^{n_{21j}}(b_{1jk}-\tau_{j-1})+2\sum_{k=1}^{n_{12j}}(a_{2jk}-\tau_{j-1})+\sum_{k=1}^{n_{22j}}(b_{2jk}-\tau_{j-1})$$
$$+3\sum_{k=1}^{n_{0j}}(c_{jk}-\tau_{j-1})-2\sum_{k:x_k\in I\cap A_j}(x_k-\tau_{j-1}) \tag{26}$$

$$\begin{aligned} D_j \;=\; & (\tau_j-\tau_{j-1})\left(\sum_{m=j+1}^{M}(n_{11m}+n_{21m}+n_{12m}+n_{22m}+2n_{0m}-n_{3m})\right) \\ & +\sum_{k=1}^{n_{11j}}(a_{1jk}-\tau_{j-1})+\sum_{k=1}^{n_{21j}}(b_{1jk}-\tau_{j-1})+\sum_{k=1}^{n_{12j}}(a_{2jk}-\tau_{j-1})+\sum_{k=1}^{n_{22j}}(b_{2jk}-\tau_{j-1}) \\ & +2\sum_{k=1}^{n_{0j}}(c_{jk}-\tau_{j-1})-\sum_{k:k\in I_{-1}\cap A_j}(x_k-\tau_{j-1}). \end{aligned} \tag{27}$$

Appendix B: DATA

Daily gold price of 1 gm. in Indian Rupees during December 01, 2020 to January 31, 2021

4326 4326 4329 4311 4315 4334 4345 4345 4345 4349 4378 4374 4321 4317 4291 4289 4293 4329 4351 4353 4357 4351 4351 4347 4522 4512 4580 4554 4443 4443 4443 4440 4441 4427 4424 4455 4437 4437 4439 4434 4423 4454 4443 4443 4437 4447 4406 4368 4313 4355 4355 4350 4343 4346 4427 4403 4353 4353 4350 4364 4330 4280

Exchange rate between Indian Rupees and US Dollar during September 30, 2021 to April 05, 2021

74.224 74.286 74.207 73.833 73.950 73.813 74.049 73.844 73.870 73.795 73.620 73.683 73.701 73.698 73.569 73.468 73.635 73.565 73.529 73.526 73.526 73.532 73.814 73.388 73.107 73.004 72.995 72.997 73.049 72.991 72.992 73.395 73.493 73.497 73.497 74.158 74.136 74.149 74.104 74.330 74.305 74.299 74.420 74.273 74.373 74.164 74.224 74.217 74.217 74.266 74.130 74.434 74.465 74.299 74.213 74.242 74.057 74.199 74.214 74.351 74.386 74.379 74.379 74.247 74.452

74.538 74.284 74.432 74.441 74.427 74.461 74.440 74.525 75.001 74.606 74.631 74.913 74.556 74.464 74.669 74.536 74.478 74.488 74.492 74.738 74.816 74.774 74.323 74.501 74.480 74.755 74.562 74.372 74.282 74.280 74.204 74.184 74.184 74.194 74.234 74.344 74.136 74.126 74.141 74.147 74.246 73.840 73.350 73.183 73.245 73.234 73.232 73.012 73.001 72.979 72.801 73.005 72.908 72.908 73.041 72.960 72.857 72.490 72.405 72.396 72.396 72.706 72.729 72.801 72.833 72.924 72.918 72.918 73.029 73.230 73.146 73.336 73.281 73.280 73.280 73.454 73.696 73.365 73.492 73.257 73.296 73.290 73.608 73.799 73.780 73.836 74.113 74.096 74.093 74.093 74.333 74.580 74.824 74.916 74.867 74.867 75.070 75.427 75.450 74.874 74.531 74.557 74.565 74.730 75.048 75.167 74.990 74.740 74.732 74.727 74.533 74.414 73.501 73.277

References

[1] Al-Hussaini, E.K. (1999), "Predicting observable from a general class of distributions", *Journal of Statistical Planning and Inference*, vol. 79, pp. 79 – 81.

[2] Arnold, B.C. (1993), "Logistic process involving Markovian minimization", *Communications in Statistics - Theory and Methods*, vol. 22, 1699 – 1707.

[3] Arnold, B.C. and Hallet, T.J. (1989), "A characterization of the Pareto process among stationary processes of the form $X_n = c\min(X_{n-1}, Y_n)$", *Statistics and Probability Letters*, vol. 8, 377 – 380.

[4] Arnold, B.C., Robertson, C.A. (1989), "Autoregressive logistic processes", *Journal of Applied Probability*, vol. 26, 524–531.

[5] Balakrishna, N. and Jayakumar, K. (1997), "Bivariate semi-Pareto distributions and processes", *Statistical Papers*, vol. 38, 149–165.

[6] Balakrishnan, N., Koutras, M.V., Milenos, F.S., Pal, S. (2016),“Piecewise linear approximation for cure rate models and associated inferential issues”, *Methodology and Computing in Applied Probability*, vol. 18, 937 – 966.

[7] Bemis,B., Bain,L.J. and Higgins,J.J. (1972), “Estimation and hypothesis testing for the parameters of a bivariate exponential distribution”, *Journal of the American Statistical Association*, vol. 67, 927-929.

[8] Christensen, S. (2012), “Phase-type distributions and optimal stopping for autoregressive processes”, *Journal of Applied Probability*, vol. 49, 22 – 39.

[9] Cox, D.R. (1972), “Regression Models and Life-Tables”, *Journal of the Royal Statistical Society*, Series B. vol. 34, 187–220.

[10] Cox, D. R. and Oakes, D. (1984), *Analysis of Survival Data*, Chapman & Hall, New York.

[11] Dellaportas, P. and Wright, D. (1991), “Numerical prediction for the two-parameter Weibull distribution”, *The Statistician*, vol. 40, pp. 365 – 372.

[12] Ganguly, G., Mitra, D., Balakrishnan, N. and Kundu, D. (2023), “A flexible model based on piecewise linear approximation for the analysis of left truncated right censored data with covariates, and applications to Worcester Heart Attack study data and Channing House data”, *Statistics in Medicine*, to appear.

[13] Goodman, M.S., Li, Y., Tiwari, R.C. (2011), “Detecting multiple change points in piecewise constant hazard functions”, *Journal of Applied Statistics*, vol. 38, 2523 – 2532.

[14] Jayakumar, K. and Girish Babu, M. (2015), "Some generalizations of Weibull distribution and related processes", *Journal of Statistical Theory and Applications*, vol. 14, 425–434.

[15] Jose, K.K., Ristić, M.M. and Joseph, A. (2011), "Marshall-Olkin bivariate Weibull distributions and processes", *Statistical Papers*, vol. 52, 789–798.

[16] Kaminsky, K.S. and Rhodin, L.S. (1985), "Maximum likelihood prediction", *Annals of the Institute of Statistical Mathematics*, vol. 37, 507 – 517.

[17] Kundu, D. (2022), "Bivariate semi-parametric singular family of distributions and its applications", *Sankhya, Ser B*, vol. 84, Part 2, 846 – 872.

[18] Kundu, D. (2023), "A Stationary Weibull-process and its applications", *Journal of Applied Statistics*, vol. 50, no. 13, 2681 – 2700.

[19] Kundu, D. and Dey, A.K. (2009), "Estimating the parameters of the Marshall Olkin bivariate Weibull distribution by EM Algorithm", *Computational Statistics and Data Analysis*, vol. 53, no. 4, 956 - 965.

[20] Kundu, D. and Gupta, R.D. (2006), "Estimation of $P(Y < X)$ for Weibull distribution", *IEEE Transactions on Reliability*, vol. 55, no. 2, 270 – 280.

[21] Kundu, D. and Gupta, A. (2013), "Bayes estimation for the Marshall-Olkin bivariate Weibull distribution", *Computational Statistics and Data Analysis*, vol. 57, 271 - 281.

[22] Murphy, A. (1996), "A piecewise-constant hazard-rate model for the duration of unemployment in single-interview samples of the stock of unemployed", *Economics Letters*, vol. 51, 177–183.

[23] Novikov, A. and Shiryaev, A. (2007), "On solution of the optimal stopping problem for processes with independent increments", *Stochastics 79*, vol. , 393–406.

[24] Pillai R.N. (1991), "Semi-Pareto processes", *Journal of Applied Probability*, vol. 28, 461–465.

[25] Samanta, D. and Kundu, D. (2023), "Bivariate semi-parametric model: Bayesian inference", *Methodology and Computing in Applied Probability*, (to appear).

[26] Sim C.H. (1986), "Simulation of Weibull and gamma autoregressive stationary processes", *Communications in Statistics - Simulation and Computation*, vol. 15, 1141–1146.

[27] Tavares, L.V. (1980) "An exponential Markovian stationary process", *Journal of Appl Probab*, vol. 17, 1117–1120.

[28] Yeh, H.C., Arnold, B.C., Robertson, C.A. (1988), "Pareto process", *Jouranl of Applied Probability*, vol. 25, 291–301.