Bayesian Inference of a Stationary Proportional Hazard Class Process

Debasis Kundu *

Abstract

In this paper we consider one discrete time and continuous state space stationary stochastic process $\{X_n; n \geq 1\}$, where X_n's are positive valued random variables. Further, the marginal distribution of X_n belongs to the proportional hazard class of distributions, and there is a positive probability that $X_n = X_{n+1}$. We provide different properties of the proposed process. We consider different distributions of X_n, namely (a) exponential, (b) Weibull and (c) piecewise constant hazard function. All these distributions belong to the proportional hazard class of distributions. We provide the Bayesian inference of the unknown parameters in all these cases under a very flexible set of priors. The Bayes estimators and the associated credible intervals cannot be obtained in closed forms. We propose to use very convenient importance sampling technique to compute the Bayes estimators and the associated credible intervals. We consider one gold price data set of the Indian market and one exchange rate data set between Indian Rupees and US dollars and in both these data sets there is a significant number of n, for which $X_n = X_{n+1}$, hence they cannot be ignored. Since, the proposed model is capable of handing this issue, we use the proposed model to analyze both the data sets. The results are quite satisfactory.

Key Words and Phrases: Weibull distribution, exponential distribution; piecewise constant hazard function distribution; Bayes estimates; posterior distribution; stationary process.

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 Motivation & Introduction

The proportional hazard class (PHC) of distributions is a very flexible class of distribution functions. Several well known distribution functions like exponential, Weibull, Rayleigh etc. belong to the PHC class of distribution functions. Several interesting properties of this class of distribution functions can be found in Cox and Oakes [10]. In this paper we have considered one discrete time and continuous state space stationary stochastic process $\{X_n; n \geq 1\}$, where X_n's are non-negative random variables and the distribution function of X_n belongs to the PHC of distribution functions. Due to this reason we call this process as the PHC process. The proposed PHC process is a stationary Markov process and one important characteristic of this process is that there is a positive probability that $X_n = X_{n+1}$. The motivation of this work came when we were trying to analyze the gold price data in the Indian market and also the exchange rate data of the Indian rupees and US dollars. It is observed in both the cases that there is a significant proportion where $X_n = X_{n+1}$, hence they cannot be ignored. We have used the proposed PHC process to analyze these two data sets, and the results are quite satisfactory.

It may be mentioned that there are several positive valued discrete time and continuous state space processes available in the literature. Among these processes, the important ones are the exponential process by Tavares [30]. A more general stationary exponential family processes has been introduced by Dinwoodie [11]. The Weibull process by Lee and Lee [20] and Jayakumar and Girish Babu [14]. The Weibull and gamma autoregressive processes by Sim [29]. The Pareto process by Yeh et al. [32], see also Arnold and Hallet [2], the semi-Pareto process by Pillai [28], bivariate semi-Pareto process by Balakrishna and Jayakumar[4], Marshall-Olkin bivariate Weibull process by Jose et al. [15], the logistic process and the autoregressive logistic process by Arnold [1] and Arnold and Robertson [3], respectively and see the references cited therein. But in all these cases $P(X_n = X_{n+1}) = 0$, hence they cannot

be used to analyze these two data sets.

Let us recall that if $F_0(t)$ denotes an absolutely continuous distribution function with the support on the positive real axis, and $S_0(t) = 1 - F_0(t)$, then the class of distribution function parameterized by $\lambda > 0$ which has the survival function

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

is known as the PHC of distributions. The PHC of distribution was originally proposed by Cox [9] and later it has been used quite extensively in the statistical literature. From now on the class of distribution function characterized by (1) will be denoted by PHC(S_0, λ). In this case $F_0(t)(S_0(t))$ is known as the base line distribution (survival) function. If $f_0(t)$ denotes the base line probability density function (PDF), then the PDF of $S(t;\lambda)$ has the following form;

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

In this paper we have considered the proposed PHC process under different base line distributions namely (a) exponential, (b) Weibull, and (c) piecewise constant hazard function. It is well known that exponential distribution has constant hazard function, where as the Weibull has monotone hazard function only. The piecewise constant hazard function (PCHF) distribution is a very flexible class of distribution functions. The shape of the hazard function can be of any type. It can be monotone, unimodal, bathtub shape or some other non-conventional form. It can be determined based on the data, and one does not need to decide before hand. Hence, it is essentially a data driven approach. Therefore, it can be a very flexible class of distribution functions. In this approach the base line hazard function $h_0(t)$ is assumed to be of the following form:

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

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

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

and $0 = \tau_0 < \tau_1 < \ldots \tau_{M-1} < \tau_M = \infty$. It is assumed that M and τ_i's are known, where as θ_j's are unknown. Observe that in presence of λ in (2), $\theta_1, \ldots, \theta_M$ are identifiable only up to a multiplicative constant. Hence, in this paper without loss of generality it is assumed $\theta_1 + \ldots + \theta_M = 1$. The PCHF distribution was first considered in details by Loubert [22]. Since then several other researchers have considered this distribution in lifetime data analysis, see for example Goodman et al. [13], Balakrishnan et al. [5], Murphy [24], Kundu [16, 17], Ganguly et al. [12], Pal et al. [26] and the references cited therein.

We have provided several properties of the general PHC process. Since, exponential, Weibull and PCHF distribution are special cases of this PHC of distributions, exponential process, Weibull process and PCHF process can be obtained as special cases of this general PHC process. It is observed that the proposed process is a lag-1 process and the joint distribution of X_n and X_{n+1} has a very convenient copula structure. It can be used to derive several dependency properties also. We have assumed very flexible set of priors in all the cases, and provided the Bayesian inference of the unknown parameters. The Bayes estimators cannot be obtained in closed forms, and we have proposed to use very convenient importance sampling techniques to compute the Bayes estimators and the associated highest posterior density (HPD) credible intervals.

Another important point which has been addressed in this paper is that we have provided an efficient method of estimating M and the cut points τ_j's. So far it has not been addressed satisfactorily in the literature. In all the cases it has been assumed that they are known. In this paper it has been assumed that the given the number of cut points, the position of the cut points follows the arrival times of a non-homogeneous Poisson process. Based on this approach we estimate the position of the cut points, and following the Akaike Information

Criterion (AIC) or Bayes Information Criterion (BIC) we also estimate the number of cut points also. We analyze the two data sets using the approach, and the results are quite satisfactory.

The rest of the paper is organized as follows. In Section 2 we have provided the description of the two data sets. In Section 3 we define the PHC process and the likelihood functions for the different cases are provided in Section 4. The priors, the posterior analysis and the goodness of fit test have been presented in Section 5 and Section 6, respectively. In Section 7 two illustrative examples have been provided, and finally in Section 8 we have concluded the paper.

2 Two Data Sets:

In this section we present the two data sets; (a) Gold Price Data and (b) Exchange Rate Data, and provide some basic analyses. The original data are presented in the Appendix A. The Gold Price Data set has been extracted from https://www.bankbazaar.com/gold-rate/gold-rate-trend-in-india.html. The Exchange Rate data set has been extracted form https://in.investing.com/currencies/usd-inr-historical-data. Both the data sets are freely available.

2.1 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. The minimum, maximum and the median values of the price 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 auto-

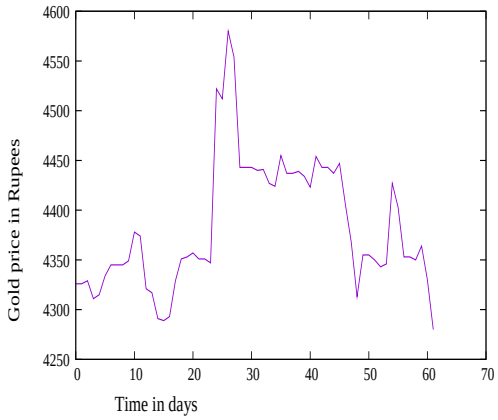
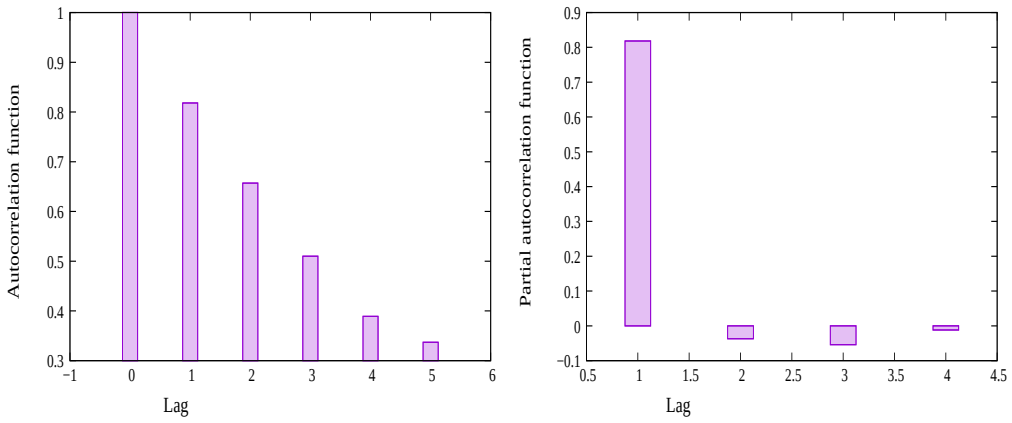


Figure 1: The price of 1 gm of 24 Carat gold in Indian Rupees from December 01, 2020 till January 31, 2021.

correlation (PACF) functions of the data have been plotted in Figure 2. To check whether



(a) (b)

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

the observations of the lag-1 difference or lag-2 difference are independently distributed or not, the run test were been performed. The results for the different series are presented in Table 1. It is clear from the PACF and from the results in Table 1 that the observations of lag-1 series as well as lag-2 series are independently distributed.

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.

2.2 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 $1 against Indian rupees has been plotted in Figure 3. The minimum, maximum and the median exchange rate during this

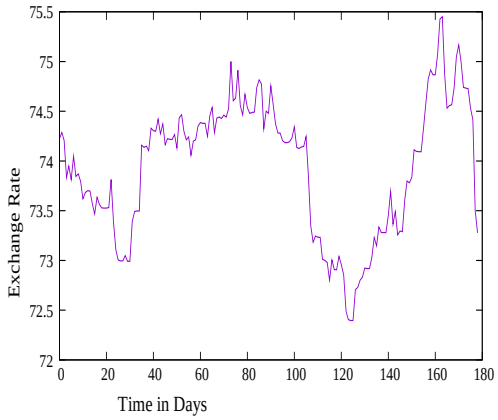
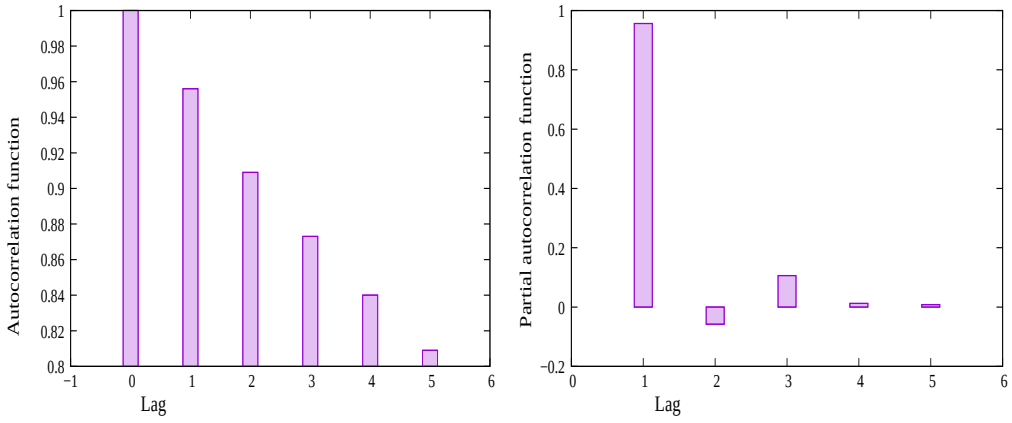


Figure 3: The exchange rate of US dollar one against Indian Rupees during April 05, 2021 to Sept 30, 2021.

period are Rupees 72.396, 75.450 and 74.139, respectively. 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 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



(a) (b)

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

well as lag-2 series are independently distributed.

3 Proportional Hazard Class Process

In this section first we will define the PHC process (PHCP) in general and derive its properties. Later we will consider the specific cases also in details.

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

$$X_n = \min\{Y_n, Y_{n+1}, Z_n\}, \qquad n = 1, 2, \ldots. \tag{4}$$

Then the sequence of random variables $\{X_n\}$ is called a PHCP. It will be denoted by PHCP$(S_0, \lambda_1, \lambda_2)$.

Because of the structure of (4) it may be called as a minification process. Note that the minification process has received a considerable amount of attention in the stochastic process literature, see for example Lewis and McKenzie [21] or Thomas and Jose [31] and the references cited therein. But in all these cases the construction is different from the present one. In the literature all the existing processes are based on the minimization of two independent random variables, but here it is based on three independent random variables. Due to this reason there is a positive probability that $X_n = X_{n+1}$. The maxification process also can be defined. The following results are easy to derive

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

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

(ii) $X_n \sim$ PHC$(S_0, 2\lambda_1 + \lambda_2)$.

(iii) 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

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

Note that because of (i) and (ii) it is called a PHCP, and (iii) will be used to derive the likelihood function of the data. Moreover, (iii) can be used to provide several dependency properties of the PHCP. First let us observe 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

survival function of X_n and X_{n+1} can be written explicitly as follows:

$$S_{n,n+m}(x,y) = \begin{cases} (S_0(x))^{\lambda_1+\lambda_2}(S_0(y))^{2\lambda_1+\lambda_2} & \text{if} \quad x < y \\ (S_0(x))^{2\lambda_1+\lambda_2}(S_0(y))^{\lambda_1+\lambda_2} & \text{if} \quad x \geq y. \end{cases} \tag{5}$$

From (5) it can be seen that (X_n, X_{n+1}) follows a Marshall-Olkin bivariate exponential (MOBE) distribution, see Marshall and Olkin [23], if we take $S_0(x) = e^{-x}$. Similarly it follows a Marshall-Olkin bivariate Weibull (MOBW) distribution if we take $S_0(x) = e^{-x^\alpha}$, see for example Kundu and Dey [18] or Kundu and Gupta [19]. If we take $\alpha = 2$, it becomes Marshall-Olkin bivariate Rayleigh (MOBR). 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 $X_{n-2}, \ldots, X_1$. Hence, $\{X_n\}$ has a Markov property. Further, X_n and X_{n+1} has the following Marshall-Olkin survival copula for $0 < \delta = \dfrac{\lambda_1+\lambda_2}{2\lambda_1+\lambda_2} < 1$,

$$\widetilde{C}(u,v) = \begin{cases} u^\delta v & \text{if} \quad u \leq v \\ uv^\delta & \text{if} \quad u > v. \end{cases} \tag{6}$$

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

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

Now we will provide the joint PDF of X_n and X_{n+1}, and because of the Markov property it will be easy 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. [6]. Based on the above dominating measure the joint PDF of X_n and X_{n+1} can be written as follows.

$$f_{X_n,X_{n+1}}(x,y) = \begin{cases} f_1(x,y) & \text{if} \quad x < y \\ f_2(x,y) & \text{if} \quad x > y \\ f_0(x,y) & \text{if} \quad x = y, \end{cases} \tag{7}$$

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

$$\begin{aligned} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) &= f_{X_n|X_{n-1},\ldots,X_1}(x_n)\times\ldots\times f_{X_2|X_1}(x_2)\times f_{X_1}(x_1) && (8) \\ &= f_{X_n|X_{n-1}}(x_n)\times\ldots\times f_{X_2|X_1}(x_2)\times f_{X_1}(x_1) && (9) \\ &= \frac{f_{X_{n-1},X_n}(x_{n-1},x_n)}{f_{X_{n-1}}(x_{n-1})}\times\ldots\times\frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}\times f_{X_1}(x_1) && (10) \\ &= \frac{\prod_{i=1}^{n-1}f_{X_i,X_{i+1}}(x_i,x_{i+1})}{\prod_{i=2}^{n-1}f_{X_i}(x_i)} && (11) \end{aligned}$$

In this case (8) is obtained by the conditioning approach, (9) is obtained by using the Markov property, (10) 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 [7] discussed about optimal stopping time for autoregressive processes and Novikov and Shiryaev [25] 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

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

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 Likelihood Function

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 four different forms of the base line distribution; (a) exponential, (b) Weibull, (c) Rayleigh and (d) PCH function. 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

$$\begin{aligned} L(\lambda_1, \lambda_2, \Theta|\mathcal{D}) &= \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}}{(2\lambda_1+\lambda_2)^{n_0-1}} \left\{\prod_{i\in I_1\cup I_2} h_0(x_{i+1};\Theta)\right\}\left\{\prod_{i\in I_0} h_0(x_i;\Theta)\right\} h_0(x_1;\Theta) \times \\ & \quad e^{-\lambda_1 E_1(\mathcal{D},\Theta)} e^{-\lambda_2 E_2(\mathcal{D},\Theta)}, \end{aligned} \tag{12}$$

where Θ is the parameter which is associated with the base line distribution function and

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

4.1 Likelihood Function: Exponential, Weibull and Rayleigh

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

$$L_{EX}(\lambda_1, \lambda_2|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}}{(2\lambda_1+\lambda_2)^{n_0-1}} e^{-\lambda_1 E_1(\mathcal{D})} e^{-\lambda_2 E_2(\mathcal{D})}, \tag{13}$$

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

$$L_{WE}(\lambda_1, \lambda_2, \alpha|\mathcal{D}) = \frac{(\lambda_1+\lambda_2)^{n_1+n_2}\lambda_1^{n_0}\alpha^n}{(2\lambda_1+\lambda_2)^{n_0-1}} \left\{\prod_{i\in I_1\cup I_2} x_{i+1}^{\alpha-1}\right\}\left\{\prod_{i\in I_0} x_i^{\alpha-1}\right\} x_1^{\alpha-1} e^{-\lambda_1 E_1(\mathcal{D};\alpha)} e^{-\lambda_2 E_2(\mathcal{D};\alpha)}, \tag{14}$$

where

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

In case of Rayleigh distribution, the likelihood function can be obtained by putting $\alpha = 2$, in (14).

4.2 Likelihood Function: PCH Function

In case of base line distribution having PCH function it is assumed that the base line hazard function $h_0(t)$ has the form defined in (3). We will use the notation $A_j = [\tau_{j-1}, \tau_j)$, for $j = 1, \ldots, M$. The base line cumulative hazard function $H_0(t) = \int_0^t h_0(u)du$, becomes

$$H_0(t) = \begin{cases} \theta_1 t & \text{if} \quad t \in A_1 \\ \theta_1\tau_1 + \theta_2(t-\tau_1) & \text{if} \quad t \in A_2 \\ \vdots & \vdots \quad \vdots \\ \sum_{j=1}^{k-1}\theta_j(\tau_j - \tau_{j-1}) + \theta_k(t-\tau_{k-1}) & \text{if} \quad t \in A_k \\ \vdots & \vdots \quad \vdots \\ \sum_{j=1}^{M-1}\theta_j(\tau_j - \tau_{j-1}) + \theta_M(t-\tau_{M-1}) & \text{if} \quad t \in A_M. \end{cases} \tag{15}$$

Once $H_0(t)$ is obtained, the base line survival function $S_0(t)$ and the base line PDF $f_0(t)$ can be obtained as

$$S_0(t) = e^{-H_0(t)} \quad \text{and} \quad f_0(t) = h_0(t)e^{-H_0(t)}. \tag{16}$$

First observe that in presence of λ_1 and λ_2, $\theta_1, \ldots, \theta_M$ are identifiable upto 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 $j = r$,

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

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

and for $j > r$,

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

See Appendix B for details.

5 Priors and Posterior Analysis

5.1 Priors: Exponential and Weibull

In case of exponential and Rayleigh distributions, we have assumed the following flexible Beta-Gamma (BG) prior on λ_1 and λ_2, as suggested by Pena and Gupta [27]. The joint prior on λ_1 and λ_2 can be written as

$$
\pi_1(\lambda_1, \lambda_2; a, b, a_1, a_2) = \frac{\Gamma(a_1 + a_2)}{\Gamma(a)} (b(\lambda_1 + \lambda_2))^{a_1 + a_2 - a} \times \prod_{i=1}^{2} \frac{b^{a_i}}{\Gamma(a_i)} \lambda_i^{a_i - 1} e^{-b\lambda_i}. \tag{18}
$$

Here $a > 0$, $b > 0$, $a_1 > 0$, $a_2 > 0$ are hyperparameters, and the above prior will be denoted by $\mathrm{BG}(a_1, a_2, a, b)$. It is a very flexible conjugate prior. Apriori λ_1 and λ_2 can be positively correlated, negatively correlated or independent depending on whether $a_1 + a_2 > a$, $a_1 + a_2 < a$ or $a_1 + a_2 = a$, respectively.

In case of Weibull, it is assumed that jointly λ_1 and λ_2 has the same BG(a_1, a_2, a, b) prior, and the prior on α, $\pi_2(\alpha|c_0, d_0)$, is gamma with the shape parameter c_0 and scale parameter d_0, denoted by GA(c_0, d_0), has the following probability density function;

$$f_{GA}(x; c_0, d_0) = \frac{d_0^{c_0}}{\Gamma(c_0)} x^{c_0-1} e^{-d_0 x}, \quad x > 0, \tag{19}$$

and it is independent of (λ_1, λ_2).

5.2 Prior: PCH Function

In this case also the (λ_1, λ_2) has the same BG(a_1, a_2, a, b) prior, and $\theta_1, \ldots, \theta_M$ have the Dirichlet prior as follows:

$$\pi_3(\theta_1, \ldots, \theta_M|\boldsymbol{d}) = \frac{1}{B(d)} \prod_{i=1}^{M} \theta_i^{d_i-1}, \tag{20}$$

for $0 < \theta_1, \ldots, \theta_M < 1$, and $\sum_{i=1}^{M} \theta_i = 1$. Here $\boldsymbol{d} = (d_1, \ldots, d_M)^\top$, $d_1 > 0, \ldots, d_M > 0$ and $B(\boldsymbol{d}) = \left\{\prod_{i=1}^{M} \Gamma(d_i)\right\} \left\{\Gamma\left(\sum_{i=1}^{M} d_i\right)\right\}^{-1}$ is the normalizing constant. It will be denoted by DI$(\boldsymbol{d})$, and the PDF (20) will be denoted by $f_{DI}(\cdot|\boldsymbol{d})$ Further, $\theta_1, \ldots, \theta_M$ are independent of (λ_1, λ_2).

5.3 Posterior Analysis: Exponential

In this case the posterior distributed of λ_1 and λ_2 given the data $\mathcal{D}$ can be written

$$\begin{aligned} \pi_{EX}(\lambda_1, \lambda_2|\mathcal{D}) &= K \times h(\lambda_1, \lambda_2) \times f_{GA}(\lambda_1; n_0 + a_1 - 1, E_1(D) + b) \times \\ &\qquad f_{GA}(\lambda_2; n_1 + n_2 - n_0 + a_1 + a_2 - a + 1, E_2(D) + b), \end{aligned} \tag{21}$$

where K is the normalizing constant and

$$h(\lambda_1, \lambda_2) = \left(\frac{\lambda_1}{\lambda_2} + 1\right)^{n_1+n_2+a_1+a_2-a} \left(\frac{2\lambda_1}{\lambda_2} + 1\right)^{1-n_0}. \tag{22}$$

The following algorithm can be used to compute the Bayes estimate of $g(\lambda_1, \lambda_2)$, any function of λ_1 and λ_2 with respect to squared error loss function, and the associated $100(1-\gamma)\%$ highest posterior density (HPD) credible interval.

Algorithm 1:

Step 1: Generate $\lambda_1 \sim \text{GA}(n_0 + a_1 - 1, E_1(D) + b)$ and $\lambda_2 \sim$ Gamma $(n_1 + n_2 - n_0 + a_1 + a_2 - a + 1, E_2(D) + b)$.

Step 2: Repeat this procedure to obtain $\{(\lambda_{1i}, \lambda_{2i}); i = 1, \ldots, N\}$.

Step 3: The Bayes estimate of $g(\lambda_1, \lambda_2)$ with respect to the squared error loss function can be approximated as

$$\widehat{g}(\lambda_1, \lambda_2) = \frac{\sum_{i=1}^{N} g_i h(\lambda_{1i}, \lambda_{2i})}{\sum_{i=1}^{N} h(\lambda_{1i}, \lambda_{2i})},$$

here $g_i = g(\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \theta_{1i}, \ldots, \theta_{Mi})$.

Step 4: Now to construct 100(1-γ)% HPD credible interval of $g(\lambda_1, \lambda_2)$, first compute

$$w_i = \frac{h(\lambda_{1i}, \lambda_{2i})}{\sum_{j=1}^{N} h(\lambda_{1j}, \lambda_{2j})}.$$

Rearrange $\{(g_1, w_1), \ldots, (g_M, w_M)\}$ as $\{(g_{(1)}, w_{[1]}), \ldots, (g_{(M)}, w_{[M]})\}$,where $g_{(1)} < \ldots < g_{(M)}$, and $w_{[i]}$'s are not ordered, they are just associated with $g_{(i)}$'s. For any $0 < p < 1$, let us define the integer N_p, such that

$$\sum_{i=1}^{N_p} w_{[i]} \leq p \leq \sum_{i=1}^{N_p+1} w_{[i]} \quad \text{and} \quad \widehat{g}_p = g_{(N_p)}.$$

A 100(1-γ)% credible interval of $g(\lambda_1, \lambda_2)$ can be obtained as $(\widehat{g}_\delta, \widehat{g}_{\delta+1-\gamma})$, for $\delta = w_{[1]}, w_{[1]} + w_{[2]}, \ldots, \sum_{i=1}^{N_\alpha} w_{[i]}$. Hence, a HPD 100(1-$\gamma$)% credible interval of $g(\lambda_1, \lambda_2)$ can be obtained as $(\widehat{g}_{\delta^*}, \widehat{g}_{\delta^*+1-\gamma})$, where

$$\widehat{g}_{\delta^*+1-\gamma} - \widehat{g}_{\delta^*} \leq \widehat{g}_{\delta+1-\gamma} - \widehat{g}_\delta, \quad \text{for all } \delta.$$

5.4 Posterior Analysis: Weibull

In this case the posterior distribution of α, λ_1 and λ_2 given the data $\mathcal{D}$ can be written as

$$\begin{aligned}\pi_{WE}(\alpha,\lambda_1,\lambda_2|\mathcal{D}) &= K\times h(\lambda_1,\lambda_2)\times f_{GA}(\lambda_1;n_0+a_1-1,E_1(D;\alpha)+b)\\ &\quad\times f_{GA}(\lambda_2;n_1+n_2-n_0+a_1+a_2-a+1,E_2(D;\alpha)+b)\\ &\quad\times f_{GA}\left(\alpha;n+c_0-1,d_0-\sum_{i\in I_1\cup I_2}\ln x_{i+1}-\sum_{i\in I_0}\ln x_i-\ln x_1\right)\end{aligned}\tag{23}$$

here K is a normalizing constant, $h(\lambda_1,\lambda_2)$ is same as defined in (22). The following algorithm can be used to compute Bayes estimate of $g(\alpha,\lambda_1,\lambda_2)$, any function of α, λ_1 and λ_2 and to construct the associate 100(1-γ)% HPD credible interval.

Algorithm 2:

Step 1: Generate α from

$$f_{GA}\left(\alpha;n+c_0-1,d_0-\sum_{i\in I_1\cup I_2}\ln x_{i+1}-\sum_{i\in I_0}\ln x_i-\ln x_1\right)$$

Step 2: For a given α, generate λ_1 from $\lambda_1\sim$ GA$(n_0+a_1-1,E_1(D)+b)$ and λ_2 from $\lambda_2\sim$ Gamma $(n_1+n_2-n_0+a_1+a_2-a+1,E_2(D)+b)$.

Step 3: Repeat Step 1 and Step 2 to generate $\{(\alpha_i,\lambda_{1i},\lambda_{2i});i=1,\ldots,N\}$.

Now we can obtain the Bayes estimate of $g(\alpha,\lambda_1,\lambda_2)$ and the associated HPD credible interval following Step 3 and Step 4 of Algorithm 1.

5.5 Posterior Analysis: PCH Function

The posterior distribution of λ_1,λ_2, $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_M)^\top$ given the data $\mathcal{D}$ can be written as

$$\pi_{PCH}(\boldsymbol{\theta},\lambda_1,\lambda_2|\mathcal{D}) = K\times h(\lambda_1,\lambda_2)\times f_{GA}(\lambda_1;n_0+a_1-1,\sum_{j=1}^{M}C_j\theta_j+b)\times$$
$$f_{GA}(\lambda_2;n_1+n_2-n_0+a_1+a_2-a+1,\sum_{j=1}^{M}D_j\theta_j+b)\times f_{DI}(\boldsymbol{\theta};\widetilde{\boldsymbol{d}}), \tag{24}$$

here $\widetilde{\boldsymbol{d}}=(\widetilde{d}_1,\ldots,\widetilde{d}_M)^\top$, and $\widetilde{d}_m=n_{21m}+n_{22m}+n_{0m}+d_m$, for $m=1,\ldots,M$, $m\neq k$, and $\widetilde{d}_k=n_{21k}+n_{22k}+n_{0k}+d_k+1$.

The following algorithm can be used to compute Bayes estimate of $g(\boldsymbol{\theta},\lambda_1,\lambda_2)$, any function of $\boldsymbol{\theta}$, λ_1 and λ_2 and to construct the associate 100(1-γ)% HPD credible interval.

Algorithm 3:

Step 1: Generate $\boldsymbol{\theta}$ from $f_{DI}(\boldsymbol{\theta};\widetilde{\boldsymbol{d}})$.

Step 2: Given $\boldsymbol{\theta}$ generate λ_1 from $f_{GA}(\lambda_1;n_0+a_1-1,\sum_{j=1}^{M}C_j\theta_j+b)$ and λ_2 from $f_{GA}(\lambda_2;n_1+n_2-n_0+a_1+a_2-a+1,\sum_{j=1}^{M}D_j\theta_j+b)$.

Step 3: Repeat Step 1 and Step 2 to generate $\{(\theta_{1i},\ldots,\theta_{Mi},\lambda_{1i},\lambda_{2i});i=1,\ldots,N\}$.

Now we can obtain the Bayes estimate of $g(\boldsymbol{\theta},\lambda_1,\lambda_2)$ and the associated HPD credible interval following Step 3 and Step 4 of Algorithm 1.

5.6 Choice of Cut Points:

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 $u(t)=\beta t^{\beta-1}$, for $\beta>0$. 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 $\theta_1, \ldots, \theta_M, \beta$, 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 Akaike 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 $u(t) = \alpha\lambda t^{\alpha-1}$, for $\alpha > 0$. 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

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

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

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

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

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

In this case the cut points are equidistant.

Hence, in practice we can implement the above method as follows. For fixed β and M, we can write the the likelihood function in terms of the unknown parameters. Then for given β and M we can obtain the maximum likelihood estimators of the other parameters. We

treat this as profile likelihood of β and then obtain the maximum likelihood estimator of α also for a given M. Then we can find M by using AIC or BIC.

6 Goodness of Fit

In this section we provide a goodness of fit test based on bootstrap technique to test whether a given data set comes from a given PHCP or not. It means, suppose we have a sample of size n, say $\{x_1, \ldots, x_n\}$ from a stationary sequence $\{X_1, \ldots, X_n\}$, we want to test the following null hypothesis;

$$H_0 : \quad \{X_1, \ldots, X_n\} \sim \text{PHCP}(S_0, \lambda_1, \lambda_2),$$

here S_0, λ_1 and λ_2 are known. We use the following notations: $\{X_{1:n} < \ldots < X_{1:n}\}$ as the ordered $\{X_1, \ldots, X_n\}$, similarly, $\{x_{1:n} < \ldots < x_{n:n}\}$ are observed values of $\{X_{1:n} < \ldots < X_{1:n}\}$. Further, $a_1 = E_{H_0}(X_{1:n}), \ldots, a_n = E_{H_0}(X_{n:n})$. We use the following statistic for the goodness of fit test:

$$W_n = \max_{1 \le i \le n} |X_{i:n} - a_i|. \tag{26}$$

It is expected under H_0, W_n should be small compared to when H_0 is not true. Hence, we use the following test criterion for a given level of significance $0 < \gamma < 1$;

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

We propose to use parametric bootstrap technique to obtain approximate value of $c_n(\gamma)$, from a given sample $\{x_1, \ldots, x_n\}$, as it is difficult to compute theoretically $c_n(\gamma)$, even for a large value of n. The following algorithm can be used for that purpose.

Algorithm

Step 1: Obtain $\widehat{S}_0$, $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$, from the observed sample.

Step 2: Generate a sample of size n from the PHCP$(\widehat{S}_0, \widehat{\lambda}_1, \widehat{\lambda}_2)$. We denote the ordered sample as $\{x^1_{1:n}, \ldots, x^1_{n:n}\}$. Repeat this procedure B times, and we 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

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

Step 4: Compute

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

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

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

7 Two Illustrative Examples

We analyze two data sets previously described, based on different baseline distributions.

7.1 Gold Price Data

We have already seen in Section 2 that the elements of the odd sequences are independently distributed, and the same is true for the elements of the even sequence also. For computational purposes we have made the following data transformation to all the values $(X_t - 4200)/1000$. It is not going to make any difference in the inference procedure. In Figure 5 we have plotted the empirical survival function (ESF) of the odd and even sequneces of the gold price data. The two ESFs are very close to each other. The Kolmogorov-Smirnov

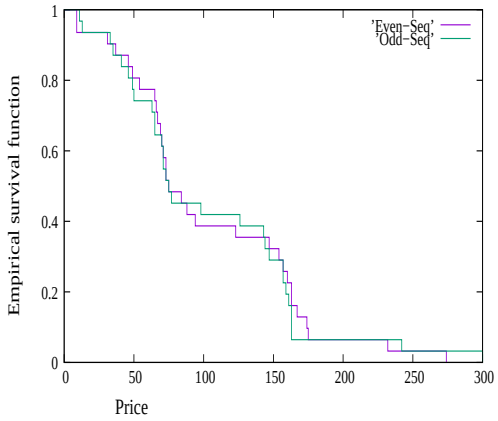


Figure 5: The empirical survival functions of the elements of the odd sequence and the even sequence of the gold price data.

distance between the two is 0.0967. Since the two sequences are dependent it is very difficult to compute the associated p value. Ignoring the dependence, the associate p value becomes 0.9976. We have fitted exponential (EXP), Weibull (WE), PCH model with one cut point (PCH1), two cut points (PCH2) and three cut points (PCH3) to both the series. The KS distances and the associated p values are presented in Table 3. From Table 3 it is clear that

Process $\longrightarrow$ Parameters $\downarrow$	Exponential	Weibull	PCH1	PCH2	PCH3
$\{X_1, X_3, \ldots\}$	0.3939	0.1872	0.1607	0.1493	0.1801
	(0.0001)	(0.2275)	(0.3998)	(0.4941)	(0.2673)
$\{X_2, X_4, \ldots\}$	0.3893	0.1716	0.1668	0.1714	0.1984
	(0.0002)	(0.3208)	(0.3544)	(0.3222)	(0.1743)

Table 3: Kolmogorov-Smirnov distances and the associated p-values for the different fitted models to the even and odd series in case of Data Set 1.

the exponential does not fit the two series, where as Weibull, PCH1 PCH2 models fit the two series quite well. We obtain the Bayes estimates and the associated HPD credible intervals for Weibull, PCH1 and PCH2 models.

Exponential and Weibull Models

In case of exponential and Weibull models we have taken the following hyperparameters: $a = b = a_1 = a_2 = 0.0001$, and $a = b = a_1 = a_2 = c_0 = d_0 = 0.0001$, respectively. They behave like almost non-informative priors, see Congdon [8], although they are proper priors. Based on the above priors we have used Algorithm 1 in case of exponential and Algorithm 2 in case of Weibull 2, with $N = 10000$. In case of exponential the Bayes estimates (posterior mean) and the associated 95% HPD credible intervals, within brackets are as follows: $\widehat{\lambda}_1 = 1.5101$ (0.9923,2.0123) and $\widehat{\lambda}_2 = 2.4189$ (1.8923,2.9965). In case of Weibull, the Bayes estimates and the associated HPD credible intervals are as follows: $\widehat{\alpha} = 2.9884$(2.0114,3.8975), $\widehat{\lambda}_1 = 31.1701$(28.4397,34.0012), $\widehat{\lambda}_2 = 49.8859$(45.4587,53.1120).

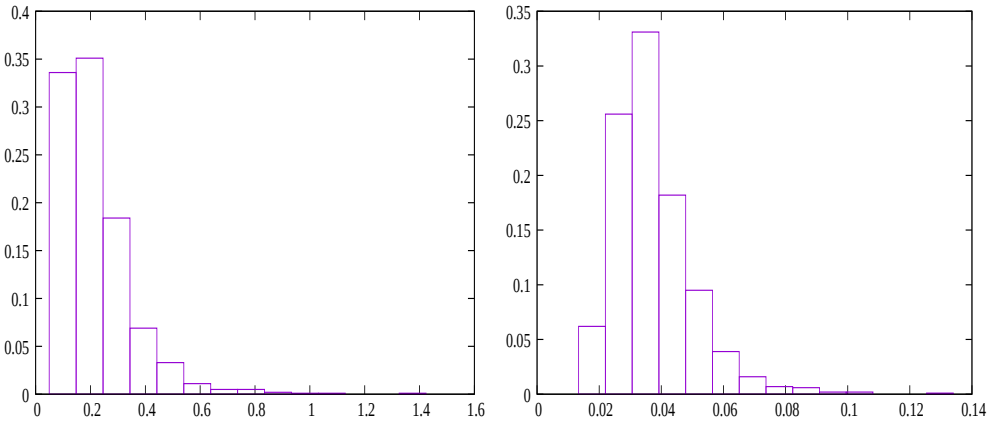


Figure 6: Histogram of the generated W_n when the base line is exponential (left) and Weibull (right).

Now based on the goodness of fit test as described in Section 6 we want to test whether exponential process and Weibull process can be fitted to the gold price data or not. We have obtained the histograms based on 10000 replications of W_n in both the cases and they are plotted in Figure 6. In case of exponential process, the test statistic $w_n = 0.4852$, and from the simulated W_n, we have obtained that the corresponding $0.01 < p < 0.05$. In case of Weibull $w_n = 0.0368$, and the associate $0.40 < p < 0.45$. Hence, exponential process cannot

be used in this case, where as Weibull process provides a good fit. The same conclusions we have drawn based on the the odd and even sequences also.

Piecewise Constant Hazard Models

We have used PCH1 and PCH2 models. We have used almost non-informative priors as before, i.e. in case of PCH1 model $a = b = a_1 = a_2 = d_1 = 0.0001$ and for PCH2 model $a = b = a_1 = a_2 = d_1 = d_2 = 0.0001$. For PCH1 model, first we obtain $\tau = 0.0177$, and the Bayes estimates (posterior means) of the parameters and the associated 95% HPD credible intervals are $\widehat{\theta}_1 = 0.0614$ (0.0518,0721), $\widehat{\theta}_2 = 0.9386$ (0.8612,0.9925), $\widehat{\lambda}_1 = 3.8081$ (3.0123,4.3125), $\widehat{\lambda}_2 = 6.0923$ (5.1123,7.1211). For PCH2 model, we obtain $\tau_1 = 0.1008$, $\tau_2 = 0.1691$, and the Bayes estimates are $\widehat{\theta}_1 = 0.0190$ (0.0101,0211), $\widehat{\theta}_2 = 0.2941$ (0.2113,0.3443), $\widehat{\theta}_3 = 0.6869$ (0.5843,0.7884), $\widehat{\lambda}_1 = 6.9027$ (5.1123,8.2314), $\widehat{\lambda}_2 = 11.2408$ (9.5612,13.0102).

For the goodness of fit test, we have obtained the histograms based on 10000 replications of W_n in both the cases and they are plotted in Figure 7. In case of PCH1 process, the test statistic $w_n = 0.0938$, and from the simulated W_n, we have obtained that the corresponding $0.35 < p < 0.40$. In case of PCH2 $w_n = 0.0313$, and the associate $0.75 < p < 0.80$. Hence, from the goodness of fit, it is observed that any one of the three namely Weibull, PCH1 and PCH2 can be used to fit this data. Now between the three based on Akaike Information Criteion, PCH1 is the preferred model.

7.2 Exchange Rate Data

We have seen in Section 2 that the entries in the odd and even sequences are independently distributed. Now to check whether or not they can be considered as identically distributed or not, we have plotted the empirical survival functions of the two sequences in Figure 8. The KS distance is 0.0604. It is clear that they are very close, ignoring the dependence

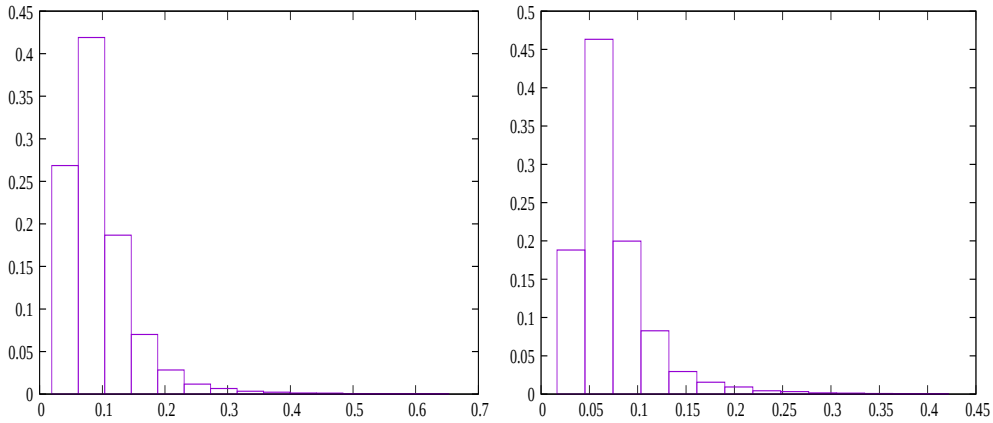


Figure 7: Histogram of the generated W_n when the base line is PCH1 (left) and PCH2 (right).

the associated p values becomes 0.9956. We have made the transformation $(X_t - 72)/10$. Similarly as the previous case, we have fitted EXP, WE, PCH1, PCH2 and PCH3 to the odd and even sequences of the exchange rate data. The results are presented in Table 4. Based on the marginals, we can say that EXP, WE and PCH1 do not fit both the marginals well.

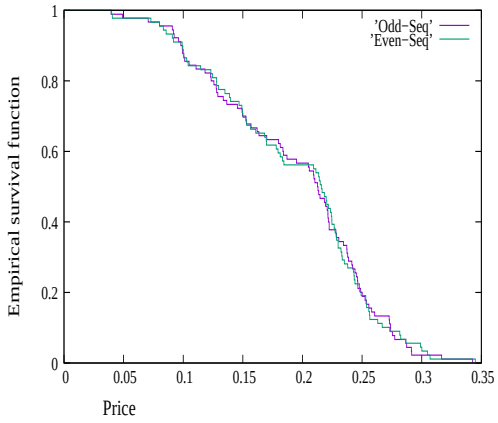


Figure 8: The empirical survival functions of the elements of the odd sequence and the even sequence of the gold price data.

We have obatined the Bayes estimates of the unknown parameters for PCH2 and PCH3 processes. We have taken the same set of almost non-informatice priors as gold-price data set.

Process $\longrightarrow$ Parameters $\downarrow$	Exponential	Weibull	PCH1	PCH2	PCH3
$\{X_1, X_3, \ldots\}$	0.3292	0.1342	0.1673	0.0853	0.0775
	(< 0.0001)	(0.0781)	(0.0129)	(0.5287)	(0.6513)
$\{X_2, X_4, \ldots\}$	0.3102	0.1554	0.1556	0.0808	0.0515
	(< 0.0001)	(0.0270)	(0.0269)	(0.6059)	(0.9721)

Table 4: Kolmogorov-Smirnov distances and the associated p-values for the different fitted models to the even and odd series in case of Data Set 2.

In case of PCH2, $\tau_1 = 0.1004$ and $\tau_2 = 0.2079$, and the Bayes estimates and the associated 95% HPD credible intervals are: $\widehat{\theta}_1 = 0.0389$ (0.0278,0488), $\widehat{\theta}_2 = 0.1367$ (0.1223,0.1487), $\widehat{\theta}_3 = 0.8244$ (0.6897,0.9654), $\widehat{\lambda}_1 = 9.9095$ (7.8978,11.1098), $\widehat{\lambda}_2 = 14.4134$ (12.3421,16.2113). In case of PCH3, $\tau_1 = 0.0637$, $\tau_2 = 0.1193$, $\tau_3 = 0.2114$, and the Bayes estimates and the associated 95% HPD credible intervals are: $\widehat{\theta}_1 = 0.0097$ (0.0078,0110), $\widehat{\theta}_2 = 0.0785$ (0.0523,0.0997), $\widehat{\theta}_3 = 0.1310$ (0.1167,0.1543), $\widehat{\theta}_4 = 0.7808$ (0.5876,0.9451) $\widehat{\lambda}_1 = 10.1158$ (8.2124,12.3211), $\widehat{\lambda}_2 = 16.1832$ (13.1245,19.1167).

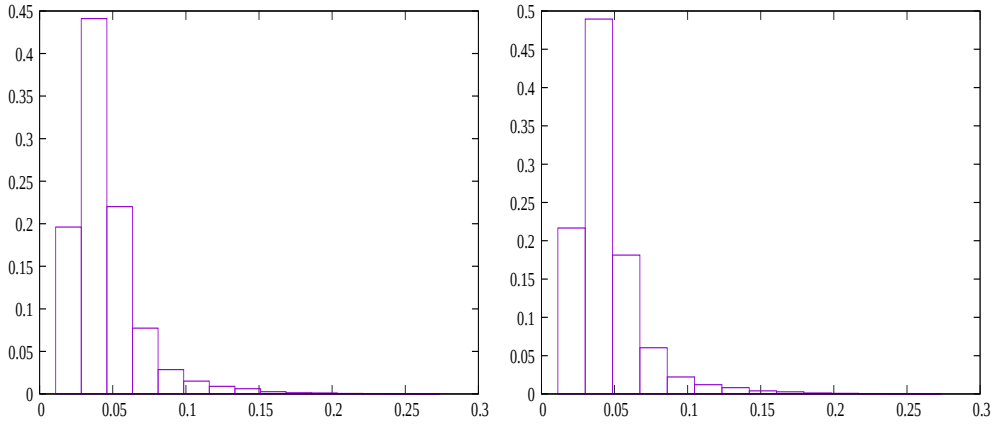


Figure 9: Histogram of the generated W_n when the base line is PCH2 (left) and PCH3 (right).

In this case also in Figure 9 we have plotted the histogram of W_n based on 10000 replications for PCH2 and PCH3 models. In case of PCH2 model $w_n = 0.0445$ and the associated $0.38 < p < 0.42$ and for PCH3 model $w_n = 0.0469$ and $0.31 < p < 0.35$. Hence, both PCH2

and PCH3 models fit the exchange rate data well. Based on AIC we preferred to choose PCH2 model in this case.

8 Conclusions

The essential feature for our process it is allowed to have $P(X_n = X_{n+1}) > 0$. We have developed the Bayesian inference of the unknown parameters under a very flexible set of priors under different base line distributions. We have developed a goodness of test procedure also. Two data sets have been analyzed to demonstrate that the proposed process can be used quite successfully in certain situations. Although we have developed the model based on a minification process, similar procedure may be devloped for a maxification process also.

Acknowledgements:

The author would like to thank the reviewers for their constructive comments which have helped to improve the manuscript significantly.

Appendix A: 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

Appendix B: PHC Likelihood Function

In this Appendix we provide the PHC likelihood function. Observe that

$$\sum_{i\in I_1} H_0(x_i) = \sum_{k=1}^{M}\sum_{w=1}^{n_{11k}} H_0(a_{1kw}) = \sum_{j=1}^{M} C_{1j}\theta_j,$$

where

$$C_{1j} = (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} n_{11m} + \sum_{k=1}^{n_{11j}} (a_{1jk} - \tau_{j-1}),$$

$$\sum_{i\in I_1} H_0(x_{i+1}) = \sum_{k=1}^{M}\sum_{w=1}^{n_{21k}} H_0(b_{1kw}) = \sum_{j=1}^{M} C_{1j}^{+}\theta_j,$$

where

$$C_{1j}^{+} = (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} n_{21m} + \sum_{k=1}^{n_{21j}} (b_{1jk} - \tau_{j-1}),$$

$$\sum_{i \in I_2} H_0(x_{i+1}) = \sum_{k=1}^{M} \sum_{w=1}^{n_{22k}} H_0(b_{2kw}) = \sum_{j=1}^{M} C_{2j}^{+} \theta_j,$$

where

$$C_{2j}^{+} = (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} n_{22m} + \sum_{k=1}^{n_{22j}} (b_{2jk} - \tau_{j-1}),$$

$$\sum_{i \in I_0} H_0(x_i) = \sum_{k=1}^{M} \sum_{w=1}^{n_{0k}} H_0(c_{kw}) = \sum_{j=1}^{M} C_{0j} \theta_j,$$

where

$$C_{0j} = (\tau_j - \tau_{j-1}) \sum_{m=j+1}^{M} n_{0m} + \sum_{k=1}^{n_{0j}} (c_{jk} - \tau_{j-1}).$$

Finally for $x_1 \in A_r$,

$$H_0(x_1) = \theta_1 \tau_1 + \theta_2(\tau_2 - \tau_1) + \ldots + \theta_{r-1}(\tau_{r-1} - \tau_{r-2}) + \theta_r(x_1 - \tau_{r-1}).$$

Combining the above results, the PHC likelihood function can be obtained.

References

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

[2] 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.

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

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

[5] 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.

[6] 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.

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

[8] Congdon, P. (2003), *Applied Bayesian Modeling*, John Wiley and Sons, New York.

[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] Dinwoodie, I.H. (1995), "Stationary exponential families", *The Annals of Statistics*, vol. 23, no. 1, 327 - 337.

[12] Ganguly, A., Mitra, D., Balakrishnan, N. and Kundu, D. (2024), "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*, vol. 43, no. 2, 233–255.

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

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

[18] 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.

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

[20] Lee, L. and Lee, S.K. (1978), "Some results on inference for the Weibull process", *Technometrics*, vol. 20, no. 1, 41 - 45.

[21] Lewis, P.A.W. and McKenzie, E. (1991), "Minification processes and their transfomations", *Journal of Applied Probability*, vol. 28, 45 – 57.

[22] Loubert, S.K. (1986), "Inference procedures for the piecewise exponential model when the data are arbitrarily censored", *Ph.D. Thesis*, Iowa State University, USA, ProQuest LLC, Ann Arbor, MI, 1986. 188 pp.

[23] Marshall, A.W. and Olkin, I. (1967), "A multivariate exponential distribution", *Journal of the American Statistical Association*, vol. 62, 30–44.

[24] 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.

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

[26] Pal, A., Samanta, D. & Kundu, D. (2025), "A semiparametric approach for simple step-stress model", TEST, vol. 34, 91–124.

[27] Pena, A. and Gupta, A.K. (1990), "Bayes estimation for the Marshall-Olkin exponential distribution", *Journal of the Royal Statistical Society*, Ser B, vol. 52, 379–389.

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

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

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

[31] Thomas, A. and Jose, K.K. (2004), "Bivariate semi-Pareto minification processes", *Metrika*, vol. 59, 305 – 313.

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