Statistics and Applications {ISSN 2454-7395 (online)}
Special Issue in Honour of Prof. V K Gupta
Volume 24, No. 2, 2026 (New Series), pp 1–22
http://www.ssca.org.in/journal

A Proportional Hazard Class Frailty Process

Debasis Kundu[1]

[1]*Department of Mathematics and Statistics*
Indian Institute of Technology Kanpur, Kanpur-208016, India

Received: dd Month yyyy; Revised: dd Month yyyy; Accepted: dd Month yyyy

Abstract

In this paper we have proposed a new discrete time continuous state space stationary process. The important feature of the proposed process is that it has positive probability of ties of two consecutive values, although the marginal distributions are absolutely continuous. The motivation of this work came when we were trying to analyze the gold price data of the Indian market. It has been observed that there are significant amount of time when there are ties between the two consecutive values, although the gold price data at any particular time point can be treated as a continuous random variable. Further, it is observed from the autocorrelation function that the gold price data in Indian market possesses persistent dependence. In this paper we have tried to take care both these aspects. We have proposed a stationary process based on proportional hazard class distributions in presence of frailty. We have considered different baseline distributions and developed several properties of the proposed process. The maximum likelihood estimators cannot be obtained in explicit form, and we have proposed EM algorithm for this purpose. Extensive simulations have been performed to show the effectiveness of the proposed EM algorithm. We have provided an empirical goodness of test also. The analyses of the gold price data set has been performed and it has been observed that the proposed model fits the data set quite well.

Key words: Marshall-Olkin bivariate exponential distribution; Marshall-Olkin bivariate Weibull distribution; bivariate singular distribution; bivariate copula; positive dependence, maximum likelihood estimators; EM algorithm.

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

1. Introduction

In this paper we have proposed a new discrete time continuous state space stochastic process based on proportional hazard class distributions in presence of frailty. It is a very flexible process as one can choose different baseline distribution functions for the proportional hazard class. The important feature of the process is that it has the property $P(X_n = X_{n+1}) > 0$, although X_n's are absolutely continuous random variables. The motivation of

Corresponding Author: Debasis Kundu
Email: kundu@iitk.ac.in

this work came when we were trying to analyze the gold price data of the Indian market. In this case it has been observed that although X_n's are positive valued continuous random variables, there is a significant number of n values for which $X_n = X_{n+1}$, which cannot be ignored. Further, from the autocorrelation function of the gold price data it is observed that it possesses persistent dependence. Hence, it is important to develop a process $\{X_n\}$, where X_n's have absolutely continuous distributions with $P(X_n = X_{n+1}) > 0$ and it also possesses persistent dependence. This is an attempt towards that direction.

There are several discrete time and continuous state space stochastic processes available in the literature for more than four decades. Among the early ones, Tavares (1980) proposed a Markovian lag-1 stationary process where all the marginals follow exponential distribution. He has developed several properties of the proposed exponential process. Sim (1986) introduced two autoregressive processes; one is gamma process based on stochastic difference equation and the other one is Weibull process based on minimization approach. In the first case the marginals follow gamma distribution and in the second case they follow Weibull distribution. The author has provided efficient simulation techniques to generate from these processes. The autoregressive Pareto process was proposed by Yeh *et al.* (1988). It is also based on minimization approach and here the marginals have Pareto distribution. They have developed several properties of the proposed process and also discussed the classical inference procedure, see also Arnold and Hallet (1989) in this respect. Following similar procedure as in Yeh *et al.* (1988), Pillai (1991) proposed autoregressive semi-Pareto process and provided its characterization.

Jose *et al.* (2010) introduced Marshall-Olkin q-Weibull distribution, and based on the minification process they have developed autoregressive lag-1 process which has the above marginal distribution. Jayakumar and Girish Babu (2015) proposed a two-sided modified Weibull distribution which has support on the whole real line, and based on minimization approach they have developed modified Weibull process. It is a lag-1 stationary process and they have analyzed Bombay stock exchange data and showed that the proposed process provided a good fit. Arnold *et al.* (2026) recently proposed a power function distribution process based on proportional reversed hazard class distribution and using similar procedures as in Kundu (2024). But similar to Kundu (2024), they are all lag-1 processes. Arnold *et al.* (2026) have derived different properties of the proposed processes and provided simple method of moment estimators of the unknown parameters, which can be obtained quite conveniently. In all the above processes the marginals have absolutely continuous probability density functions, but $P(X_n = X_{n+1}) = 0$.

It may be recalled that Marshall and Olkin (1967) introduced a bivariate exponential random variable (X, Y), where both X and Y follow exponential distributions and $P(X = Y) > 0$. From now on we call this random variable as the Marshall-Olkin bivariate exponential (MOBE) random variable. The construction of the MOBE random variable is as follows. Suppose U_1, U_2 and U_3 are three independent exponential random variables, then (X, Y), where $X = \min\{U_1, U_3\}$ and $Y = \min\{U_2, U_3\}$, is said to be MOBE random variable. Note that although U_1, U_2 and U_3 are independent, X and Y are dependent due to the presence of U_3. Part of the main idea of the proposed process mainly comes from the above construction. Here we have used more general proportional hazard class models and also we have used frailty as introduced by Vaupel et al. (1979), to incorporate some hidden dependence between X and Y. In this way we have created a stochastic process

which possesses persistent dependence.

First we have introduced the process based on general proportional hazard distributions and using frailty. We call this process as the proportional hazard class frailty (PHCF) process. It is a persistent dependence stationary process whose marginals have absolutely continuous distribution function and it has the required property $P(X_n = X_{n+1}) > 0$. We have considered some of the special cases of the baseline distribution namely; exponential, Rayleigh, Weibull and linear failure rate and discussed several properties of the proposed PHCF process. The maximum likelihood estimators of the unknown parameters cannot be obtained in explicit forms and it involves solving a higher dimensional optimization problem. We have proposed to use the EM algorithm, which significantly reduces the computational burden. Extensive simulations have been performed to show the effectiveness of the proposed EM algorithm. We have proposed some empirical goodness of fit test. Finally, we have analyzed the gold price data set, and it is observed that the proposed model provides a good fit to the data set, when the baseline distribution is Weibull.

The rest of the paper is organized as follows. In Section 2 we have provided some preliminary results. The PHCF process has been introduced and several of its properties have been discussed in Section 3. In Section 4 we have presented the EM algorithm and simulation results have been provided in Section 5. The empirical goodness of fit test has been discussed in Section 6. In section 7 the analysis of the gold price data has been presented. Finally we have concluded the paper in Section 8.

2. Some Preliminary Results

In this manuscript we have assumed that all the univariate random variables are positive valued random variables and they are absolutely continuous. We will be using the following notations for the rest of the paper. A random variable X is said to be an exponential random variable with the scale parameter $\lambda > 0$, denoted by $\mathrm{EX}(\lambda)$, if X has the following probability density function (PDF)

$$f_{EX}(t;\lambda) = \lambda e^{-\lambda t}; \quad \text{for} \quad t > 0, \tag{1}$$

and 0, otherwise. It has the cumulative distribution function (CDF), survival function (SF) and hazard function (HF), respectively, for $t > 0$, as

$$S_{EX}(t;\lambda) = e^{-\lambda t}, \quad F_{EX}(t;\lambda) = 1 - e^{-\lambda t}, \quad h_{EX}(t;\lambda) = \lambda. \tag{2}$$

A random variable X is said to be a Weibull random variable with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$, denoted by $\mathrm{WE}(\alpha, \lambda)$, if it has the following PDF

$$f_{WE}(t;\alpha,\lambda) = \alpha\lambda t^{\alpha-1} e^{-\lambda t^{\alpha}}; \quad \text{for} \quad t > 0, \tag{3}$$

and 0, otherwise. It has the cumulative distribution function (CDF), survival function (SF) and hazard function (HF), respectively, for $t > 0$, as

$$S_{WE}(t;\alpha,\lambda) = e^{-\lambda t^{\alpha}}, \quad F_{WE}(t;\alpha,\lambda) = 1 - e^{-\lambda t^{\alpha}}, \quad h_{WE}(t;\alpha,\lambda) = \alpha\lambda t^{\alpha-1}. \tag{4}$$

A random variable X is said to have a linear failure rate (LFR) distribution with parameters $\alpha > 0$ and $\beta > 0$, denoted by $\mathrm{LFR}(\alpha, \beta)$, if it has the following PDF

$$f_{LFR}(t;\alpha,\beta) = (\alpha + 2\beta t) e^{-(\alpha t + \beta t^2)}; \quad \text{for} \quad t > 0, \tag{5}$$

and 0, otherwise. It has the cumulative distribution function (CDF), survival function (SF) and hazard function (HF), respectively, for $t > 0$, as

$$S_{LFR}(t;\alpha,\beta) = e^{-(\alpha t+\beta t^2)}, \quad F_{LFR}(t;\alpha,\beta) = 1 - e^{-(\alpha t+\beta t^2)}, \quad h_{LFR}(t;\alpha,\beta) = \alpha + 2\beta t. \tag{6}$$

A random variable X is said to have a gamma distribution with the shape parameter $\alpha > 0$ and the scale parameter $\lambda > 0$, denoted by $\text{GA}(\alpha,\lambda)$, if it has the PDF

$$f_{GA}(t;\alpha,\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}; \quad \text{for} \quad t > 0, \tag{7}$$

and 0, otherwise.

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 $\lambda > 0$, which has the following survival function for $t > 0$;

$$S(t;\lambda) = (S_0(t))^\lambda. \tag{8}$$

Clearly, $F(t;\lambda) = 1 - S(t;\lambda)$ is an absolutely continuous distribution function. It will be denoted by $\text{PHC}(\lambda, S_0)$. This is known as the PHC of distributions because if $h_0(t)$ is the hazard function of $F_0(t)$, then $\lambda h_0(t)$ is the hazard function of $F(t;\lambda)$. We denote $H_0(t) = \int_0^t h_0(u)du$, as the cumulative hazard function of $h_0(t)$. Note that if $f_0(t)$ denotes the PDF of $F_0(t)$, then the PDF of $S(t;\lambda)$ is

$$f_{PHC}(t;\lambda,S_0) = \lambda f_0(t)(S_0(t))^{\lambda-1} = \lambda h_0(t) e^{-\lambda H_0(t)}. \tag{9}$$

This class of distribution functions was originally introduced by Cox (1972). Note that 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 (1984) in this respect.

Following similar approach as Marshall and Olkin (1967) defined the MOBE distribution we define bivariate proportional hazard class (BPHC) distribution. Suppose U_1 follows ($\sim$) $\text{PHC}(\lambda_1, S_0)$, $U_2 \sim \text{PHC}(\lambda_2, S_0)$ and $U_3 \sim \text{PHC}(\lambda_3, S_0)$ and they are independently distributed, then (X, Y), where $X = \min\{U_1, U_3\}$ and $Y = \min\{U_2, U_3\}$ is said to be BPHC random variable. The joint survival function of (X, Y) can be written as

$$P(X > x, Y > y) = S_{BPHC}(x,y) = \begin{cases} (S_0(x))^{\lambda_1}(S_0(y))^{\lambda_2+\lambda_3} & \text{if} \quad x < y \\ (S_0(x))^{\lambda_1+\lambda_3}(S_0(y))^{\lambda_2} & \text{if} \quad x \geq y \end{cases} \tag{10}$$

The joint PDF of (X, Y) can be written as

$$f_{BPHC}(x,y) = \begin{cases} f_{PHC}(x;\lambda_1,S_0) f_{PHC}(y;\lambda_2+\lambda_3,S_0) & \text{if} \quad x < y \\ f_{PHC}(x;\lambda_1+\lambda_3,S_0) f_{PHC}(y;\lambda_2,S_0) & \text{if} \quad x > y \\ \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f_{PHC}(z;\lambda_1+\lambda_2+\lambda_3,S_0) & \text{if} \quad x = y = z \end{cases} \tag{11}$$

From now on it will be denoted by $\text{BPHC}(\lambda_1, \lambda_2, \lambda_3, S_0)$. It can be easily seen that

$$P(X < Y) = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3}, \quad P(Y < X) = \frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_3}, \quad P(X = Y) = \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3}.$$

The correlation coefficient of X and Y varies from zero to one. X and Y are independent if $\lambda_3 = 0$ and the correlation tends to one as $\lambda_3 \to \infty$.

Now we will introduce the proportional hazard (gamma) frailty (PHF) model. Suppose $V \sim \text{GA}(\beta, \beta)$, $F_0(t) = 1 - S_0(t)$ is a survival function with the corresponding PDF $f_0(t)$ and $\lambda > 0$. Let X be a random variable such that for $t > 0$,

$$P(X > t|V = v) = (S_0(t))^{\lambda v}. \tag{12}$$

Then the survival function of X becomes

$$S_X(t; \lambda, \beta, S_0) = P(X > t) = \frac{\beta^\beta}{\Gamma(\beta)} \int_0^\infty (S_0(t))^{\lambda v} v^{\beta-1} e^{-v\beta} dv = \frac{\beta^\beta}{(\beta - \lambda \ln S_0(t))^\beta}. \tag{13}$$

Hence, the corresponding PDF becomes

$$f_X(t; \lambda, \beta, S_0) = \frac{\lambda \beta^{\beta+1} h_0(t)}{(\beta - \lambda \ln S_0(t))^{\beta+1}} \tag{14}$$

From now on it will be called PHF model and it will be denoted by $\text{PHF}(\lambda, \beta, S_0)$.

Now we will introduce bivariate proportional hazard (gamma) frailty (BPHF) model based on BPHC model. Suppose V is same as defined above and (X, Y) is a bivariate random variable such that $\{(X, Y)|V = v\} \sim \text{BPHC}(v\lambda_1, v\lambda_2, v\lambda_3, S_0)$, i.e. for $x > 0$ and $y > 0$,

$$P(X > x, Y > y|V = v) = \begin{cases} (S_0(x))^{v\lambda_1} (S_0(y))^{v(\lambda_2+\lambda_3)} & \text{if} \quad x < y \\ (S_0(x))^{v(\lambda_1+\lambda_3)} (S_0(y))^{v\lambda_2} & \text{if} \quad x \geq y. \end{cases} \tag{15}$$

Hence, the joint survival function of (X, Y) becomes

$$S_{BPHF}(x, y) = P(X > x, Y > y) = \begin{cases} \frac{\beta^\beta}{(\beta - \lambda_1 \ln S_0(x) - (\lambda_2+\lambda_3) \ln S_0(y))^\beta} & \text{if} \quad x < y \\ \frac{\beta^\beta}{(\beta - (\lambda_1+\lambda_3) \ln S_0(x) - \lambda_2 \ln S_0(y))^\beta} & \text{if} \quad x \geq y. \end{cases} \tag{16}$$

The joint PDF of (X, Y) becomes

$$f_{BPHF}(x, y) = \begin{cases} \frac{\beta^{\beta+1}(\beta+1)\lambda_1(\lambda_2+\lambda_3) h_0(x) h_0(y)}{(\beta - \lambda_1 \ln S_0(x) - (\lambda_2+\lambda_3) \ln S_0(y))^{\beta+2}} & \text{if} \quad x < y \\ \frac{\beta^{\beta+1}(\beta+1)(\lambda_1+\lambda_3)\lambda_2 h_0(x) h_0(y)}{(\beta - (\lambda_1+\lambda_3) \ln S_0(x) - \lambda_2 \ln S_0(y))^{\beta+2}} & \text{if} \quad x > y \\ \frac{\beta^{\beta+1} \lambda_3 h_0(z)}{(\beta - (\lambda_1+\lambda_2+\lambda_3) \ln S_0(z))^{\beta+1}} & \text{if} \quad x = y = z \end{cases} \tag{17}$$

The BPHF has the following survival copula

$$\overline{C}(u, v) = \begin{cases} \left(\frac{\lambda_1}{\lambda_1+\lambda_3}(u^{-1/\beta} - 1) + v^{-1/\beta}\right)^{-\beta} & \text{if} \quad (\lambda_2 + \lambda_3)(u^{-1/\beta} - 1) < (\lambda_1 + \lambda_3)(v^{-1/\beta} - 1) \\ \left(u^{-1/\beta} + \frac{\lambda_2}{\lambda_2+\lambda_3}(v^{-1/\beta} - 1)\right)^{-\beta} & \text{if} \quad (\lambda_2 + \lambda_3)(u^{-1/\beta} - 1) > (\lambda_1 + \lambda_3)(v^{-1/\beta} - 1) \end{cases} \tag{18}$$

If $\lambda_1 = \lambda_2$ and if we denote $\delta = \dfrac{\lambda_1}{\lambda_1 + \lambda_3} = \dfrac{\lambda_2}{\lambda_2 + \lambda_3}$, then (18) becomes

$$\overline{C}(u, v) = \begin{cases} \left(\delta(u^{-1/\beta} - 1) + v^{-1/\beta}\right)^{-\beta} & \text{if} \quad u > v \\ \left(u^{-1/\beta} + \delta(v^{-1/\beta} - 1)\right)^{-\beta} & \text{if} \quad u < v \end{cases} \tag{19}$$

Hence, different dependency properties of BPHF may be explored through copula function.

3. Proportional Hazard Class Frailty Process

In this section we define the proportional hazard class frailty process and derive its different properties. Suppose $\{U_n; n \geq 1\}$ is a sequence of identically distributed random variables, similarly, $\{W_n; n \geq 1\}$ is also a sequence of identically distributed random variables. Let V be a gamma random variable with the shape and scale parameters both are β. It is assumed $\{U_n|V = v\} \sim \text{PHC}(\lambda_1 v, S_0)$ for $n = 1, 2, \ldots$, and they are independent and identically distributed (i.i.d.). Similarly, $\{W_n|V = v\} \sim \text{PHC}(\lambda_2 v, S_0)$ for $n = 1, 2, \ldots$, and they are also i.i.d. Further, $\{U_n|V = v; n = 1, 2, \ldots\}$ and $\{W_n|V = v; n = 1, 2, \ldots\}$ are independently distributed. . We define $\{X_n; n \geq 1\}$, where

$$X_n = \min\{U_n, U_{n+1}, W_n\}; \quad n = 1, 2, \ldots. \tag{20}$$

Then $\{X_n; n \geq 1\}$ is called as the proportional hazard (class) frailty process (PHFP) with parameters λ_1, λ_2, β and the baseline survival function $S_0(t)$. It will be denoted by $\text{PHFP}(\lambda_1, \lambda_2, \beta, S_0)$.

Theorem 1: Suppose $\{X_n\}$ is a $\text{PHFP}(\lambda_1, \lambda_2, \beta, S_0)$, then

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

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

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

$$\begin{aligned} P(X_n > t) &= \int_0^\infty P(X_n > t|V = v) f_{GA}(v) dv \\ &= \int_0^\infty P(U_n > t|V = v) P(U_{n+1} > t|V = v) P(W_n > t|V = v) f_{GA}(v) dv \\ &= \int_0^\infty (S_0(t))^{v(2\lambda_1+\lambda_2)} f_{GA}(v) dv \\ &= \frac{\beta^\beta}{(\beta - (2\lambda_1 + \lambda_2) \ln S_0(t))^\beta}. \end{aligned}$$

Theorem 2: Suppose $\{X_n\}$ is a $\text{PHFP}(\lambda_1, \lambda_2, \beta, S_0)$, then

(i) The joint survival function of $X_1, \ldots, X_n$, for $x_1 > 0, \ldots, x_n > 0$ is

$$S_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \beta^\beta \left(\beta + \lambda_1 \left(S_0(x_1) + \sum_{i=1}^{n-1} S_0(z_i) + S_0(x_n) \right) + \lambda_2 \left(\sum_{i=1}^{n} S_0(x_i) \right) \right)^{-\beta},$$

here $z_i = \max\{x_i, x_{i+1}\}$, for $i = 1, \ldots, n-1$.

(ii) $\min\{X_1, \ldots, X_n\} \sim \text{PHF}((n+1)\lambda_1 + n\lambda_2, \beta, S_0)$.

Proof: To prove (i) note that

$$\begin{aligned}
&P(X_1 > x_1, \ldots, X_n > x_n) = S_{X_1,\ldots,X_n}(x_1, \ldots, x_n) \\
&= P\left(U_1 > x_1, \prod_{i=1}^{n-1} \{U_{i+1} > z_i\}, U_{n+1} > x_n, \prod_{i=1}^{n} \{W_i > x_i\}\right) \\
&= \int_0^\infty P\left(U_1 > x_1, \prod_{i=1}^{n-1} \{U_{i+1} > z_i\}, U_{n+1} > x_n, \prod_{i=1}^{n} \{W_i > x_i\} | V = v\right) f_{GA}(v) dv \\
&= \beta^\beta \left(\beta + \lambda_1 \left(S_0(x_1) + \sum_{i=1}^{n-1} S_0(z_i) + S_0(x_n)\right) + \lambda_2 \left(\sum_{i=1}^{n} S_0(x_i)\right)\right)^{-\beta}.
\end{aligned}$$

(ii) It mainly follows from (i) by using $x_1 = \ldots = x_n = x$.

The following results will be useful in deriving the joint PDF of $X_1, \ldots, X_n$ in a convenient form. It might have some independent interest also.

Theorem 3: Suppose $\{X_n\}$ is a PHFP$(\lambda_1, \lambda_2, \beta, S_0)$, then

(i) $\{X_n|V = v\} \sim$ PHC$(S_0, v(2\lambda_1 + \lambda_2))$.

(ii) $\{(X_n, X_{n+1})|V = v\} \sim$ BPHC$(v(\lambda_1 + \lambda_2), v(\lambda_1 + \lambda_2), v\lambda_1)$.

(iii) $\{X_n|V = v\}$ and $\{X_{n+m}|V = v\}$ are independently distributed for $n \geq 1$ and $m > 1$.

(iv) For any $x_1 > 0, \ldots, x_n > 0, v > 0$,

$$P(X_n \geq x_n|X_{n-1} = x_{n-1}, \ldots, X_1 = x_1, V = v) = P(X_n \geq x_n|X_{n-1} = x_{n-1}, V = v).$$

Proof: (i) mainly follows from (ii) of Theorem 1. To prove (ii) note that for $x > 0$, $y > 0$ and $z = \max\{x, y\}$,

$$P(X_n > x, X_{n+1} > y|V = v) = [S_0(x)]^{v(\lambda_1+\lambda_2)}[S_0(y)]^{v(\lambda_1+\lambda_2)}[S_0(z)]^{v\lambda_1}.$$

Hence, the result follows. (iii) and (iv) follow from the definition.

Now we will obtain the joint PDF of $X_1, \ldots, X_{n+1}$, and it will be used to develop the likelihood based inference of the unknown parameters. The joint PDF of $\{(X_k, X_{k+1})|V = v\}$ can be written as

$$f_{X_k,X_{k+1}|V=v}(x_k, x_{k+1}) = \begin{cases} f_{PHC}(x_k; v(\lambda_1 + \lambda_2), S_0) f_{PHC}(x_{k+1}; v(2\lambda_1 + \lambda_2), S_0) & \text{if} \quad x_k < x_{k+1} \\ f_{PHC}(x_k; v(2\lambda_1 + \lambda_2), S_0) f_{PHC}(x_{k+1}; v(\lambda_1 + \lambda_2), S_0) & \text{if} \quad x_k > x_{k+1} \\ \frac{\lambda_1}{(3\lambda_1+2\lambda_2)} f_{PHC}(z_k; v(3\lambda_1 + 2\lambda_2), S_0) & \text{if} \quad x_k = x_{k+1} = z_k \end{cases}$$

Hence, using (iv) of Theorem 3, the joint PDF of $X_1, \ldots, X_{n+1}$ can be written as

$$\begin{aligned}
f_{X_1,\ldots,X_{n+1}|V=v}(x_1, \ldots, x_{n+1}) &= f_{X_1|V=v}(x_1) f_{X_2|X_1,V=v}(x_2) \ldots f_{X_{n+1}|X_n,V=v}(x_{n+1}) \quad (21) \\
&= \frac{f_{X_1|V=v}(x_1) \prod_{k=1}^n f_{X_k,X_{k+1}|V=v}(x_k, x_{k+1})}{\prod_{j=1}^n f_{X_j|V=v}(x_j)} \quad (22) \\
&= \frac{\prod_{k=1}^n f_{X_k,X_{k+1}|V=v}(x_k, x_{k+1})}{\prod_{j=2}^n f_{X_j|V=v}(x_j)} \quad (23)
\end{aligned}$$

for $x_1 > 0, \ldots, x_{n+1} > 0, v > 0$. Let us use the following notations: $I = \{1, \ldots, n\}$, $I_1 = \{i : i \in I, x_i < x_{i+1}\}$, $I_2 = \{i : i \in I, x_i > x_{i+1}\}$, $I_3 = \{i : i \in I, x_i = x_{i+1}\}$, and n_1, n_2, n_3 denote the number of points in I_1, I_2 and I_3, respectively. Based on the above notations,

$$\begin{aligned}\prod_{k=1}^{n} f_{X_k,X_{k+1}|V=v}(x_k, x_{k+1}) &= v^{2n_1+2n_2+n_3}\lambda_1^{n_3}(\lambda_1+\lambda_2)^{n_1+n_2}(2\lambda_1+\lambda_2)^{n_1+n_2} \times \\ &\prod_{k\in I_1}[S_0(x_k)]^{v(\lambda_1+\lambda_2)}[S_0(x_{k+1})]^{v(2\lambda_1+\lambda_2)} \times \\ &\prod_{k\in I_2}[S_0(x_k)]^{v(2\lambda_1+\lambda_2)}[S_0(x_{k+1})]^{v(\lambda_1+\lambda_2)} \times \\ &\prod_{k\in I_3}[S_0(x_k)]^{v(3\lambda_1+2\lambda_2)} \times \prod_{k\in I} h_0(x_k) \times \prod_{k\in I_1\cup I_2} h_0(x_{k+1})\end{aligned}$$

and

$$\prod_{j=2}^{n} f_{X_j|V=v}(x_j) = v^{n-1}(2\lambda_1+\lambda_2)^{n-1}\prod_{j=2}^{n}\{[S_0(x_j)]^{v(2\lambda_1+\lambda_2)}h_0(x_j)\}.$$

Since, $H_0(x) = -\ln(S_0(x))$ and if we denote $\mathcal{D} = \{x_1, \ldots, x_{n+1}\}$,

$$\begin{aligned}A_1(S_0, D) &= 2\sum_{k\in I_1} H_0(x_{k+1}) + \sum_{k\in I_2} H_0(x_{k+1}) + \sum_{k\in I_3} H_0(x_k) + 2H_0(x_1) - \sum_{k\in I_1} H_0(x_k) \\ A_2(S_0, D) &= \sum_{k\in I_1\cup I_2} H_0(x_{k+1}) + H_0(x_1) + \sum_{k\in I_3} H_0(x_k),\end{aligned}$$

then

$$\begin{aligned}f_{X_1,\ldots,X_{n+1}|V=v}(x_1, \ldots, x_{n+1}) &= \frac{v^{n_1+n_2-1}\lambda_1^{n_3}(\lambda_1+\lambda_2)^{n_1+n_2}}{(2\lambda_1+\lambda_2)^{n_3-1}} e^{-v\lambda_1 A_1(S_0,\mathcal{D})-v\lambda_2 A_2(S_0,\mathcal{D})} \times \\ &\qquad h_0(x_1)\prod_{k\in I_1\cup I_2} h_0(x_{k+1}). \qquad (24)\end{aligned}$$

Therefore,

$$\begin{aligned}f_{X_1,\ldots,X_{n+1}}(x_1, \ldots, x_{n+1}) &= \int_o^\infty f_{X_1,\ldots,X_{n+1}|V=v}(x_1, \ldots, x_{n+1}) f_{GA}(v; \beta) dv \qquad (25) \\ &= \frac{\beta^\beta}{\Gamma(\beta)} \frac{\lambda_1^{n_3}(\lambda_1+\lambda_2)^{n_1+n_2}}{(2\lambda_1+\lambda_2)^{n_3-1}} \times h_0(x_1)\prod_{k\in I_1\cup I_2} h_0(x_{k+1}) \times \\ &\qquad \int_0^\infty v^{n_1+n_2+\beta-2} e^{-v(\beta+\lambda_1 A_1(S_0,\mathcal{D})+\lambda_2 A_2(S_0,\mathcal{D}))} dv \\ &= \frac{\lambda_1^{n_3}(\lambda_1+\lambda_2)^{n_1+n_2}}{(2\lambda_1+\lambda_2)^{n_3-1}} \times \frac{\Gamma(n_1+n_2+\beta-1)}{\Gamma(\beta)} \times h_0(x_1)\prod_{k\in I_1\cup I_2} h_0(x_{k+1}) \times \\ &\qquad \frac{\beta^\beta}{(\beta+\lambda_1 A_1(S_0,\mathcal{D})+\lambda_2 A_2(S_0,\mathcal{D}))^{n_1+n_2+\beta-1}}. \qquad (26)\end{aligned}$$

It can be easily seen that

$$V|\{X_1, \ldots, X_{n+1}\} \sim \text{Gamma}(n_1+n_2+\beta-1, \beta+\lambda_1 A_1(S_0,\mathcal{D})+\lambda_2 A_2(S_0,\mathcal{D})),$$

therefore

$$E(V|X_1,\ldots,X_{n+1}) = \frac{n_1+n_2+\beta-1}{\beta+\lambda_1 A_1(S_0,\mathcal{D})+\lambda_2 A_2(S_0,\mathcal{D})}.$$

The above result will be used for implementing the EM algorithm in the next section.

Stopping time plays an important role in any stationary time series. It has been discussed in the literature by several authors. It has important applications in different areas including quality control, finance, sequential analysis, change-point detection etc. Christensen (2012) discussed the optimal stopping time for autoregressive processes. Kundu (2024) provided the optimal stopping rule for a large class of stationary processes. Novikov and Shiryaev (2007) obtained the optimal stopping rule for stationary processes with independent increments. Surya (2007) discussed the optimal stopping problems driven by Levy processes. For some of the related works see the references cited therein.

Suppose $\{X_n\}$ is a PHFP$(\lambda_1,\lambda_2,\beta,S_0)$. Let use define the stopping time N for $\{X_n\}$, for a given L as follows:

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

Since, the event $\{N=n\}$ is independent of $X_{n+1},X_{n+2},\ldots$, for all $n\ge 1$, therefore, N is a stopping time for the above PHFP. Let us compute the probability mass function of N, for $n=1,2,\ldots$.

$$\begin{aligned}
P(N=n) &= P(X_1>L,\ldots,X_{n-1}>L,X_n\le L)\\
&= P(X_1>L,\ldots,X_{n-1}>L)-P(X_1>L,\ldots,X_{n-1}>L,X_n>L)\\
&= \left(1+\frac{\lambda_1}{\beta}nS_0(L)+\frac{\lambda_2}{\beta}(n-1)S_0(L)\right)^{-\beta}-\left(1+\frac{\lambda_1}{\beta}(n+1)S_0(L)+\frac{\lambda_2}{\beta}nS_0(L)\right)^{-\beta}\\
&= (1+An+B(n-1))^{-\beta}-(1+A(n+1)+Bn)^{-\beta},
\end{aligned}$$

where $A=\frac{\lambda_1}{\beta}S_0(L)$, $B=\frac{\lambda_2}{\beta}S_0(L)$, and it is defined that $P(N=0)=1-(1+A)^{-\beta}$. The k-th moment of N exists for $\beta>k+1$, and it is

$$\begin{aligned}
E(N^k) &= \sum_{n=1}^{\infty}P(N=n)n^k\\
&= \sum_{n=1}^{\infty}n^k\left(1+An+B(n-1)\right)^{-\beta}-\sum_{n=1}^{\infty}n^k\left(1+A(n+1)+Bn\right)^{-\beta}.
\end{aligned}$$

Using the Hurwitz zeta function $\zeta(s,q)=\sum_{m=0}^{\infty}(m+q)^{-s}$, and if we use $C=\frac{1-B}{A+B}$ and $D=\frac{1+A}{A+B}$, then it can be seen that for $\beta>2$, $E(N)$ exits and for $\beta>3$, $E(N^2)$ exists. They are as follows:

$$\begin{aligned}
E(N) &= (A+B)^{-\beta}[\zeta(\beta-1,D)-C\zeta(\beta,D)-\zeta(\beta-1,1+D)+D\zeta(\beta,1+D)],\\
E(N^2) &= (A+B)^{-\beta}[\zeta(\beta-2,D)-2C\zeta(\beta-1,D)+C^2\zeta(\beta,D)-\\
&\qquad \zeta(\beta-2,1+D)+2D\zeta(\beta-1,1+D)-D^2\zeta(\beta,1+D)].
\end{aligned}$$

4. Inference

In this section we discuss about the maximum likelihood estimators (MLEs) of the unknown parameters based on the random sample $\mathcal{D} = \{x_1, \ldots, x_{n+1}\}$. Note that the baseline survival function can be completely known, say for example in case of exponential, where $S_0(t) = e^{-t}$, or in case of Rayleigh, where $S_0(t) = e^{-t^2}$. Otherwise, the baseline survival function also contains some unknown parameters, say for example in case of Weibull distribution where $S_0(t) = e^{-t^\theta}$ and in case of linear failure rate distribution $S_0(t) = e^{-(1+\theta t^2)}$. Therefore, if the baseline survival function is completely known then the unknown parameters are λ_1, λ_2 and β. When the baseline survival function also contains unknown parameter(s), say θ, where θ can be vector valued also, then the unknown parameters are λ_1, λ_2, β and θ. Let us assume that θ is of dimension $p \geq 1$. We consider both the cases, when the baseline survival function is completely known and also when it has also some unknown parameter(s).

4.1. Baseline is Completely Known

In this case it is assumed that the baseline survival function $S_0(\cdot)$ is completely known. Taking logarithm of the joint PDF of $X_1, \ldots, X_{n+1}$ as given in (26) based on the data $\mathcal{D}$, the log-likelihood function without the additive constant can be written as

$$\begin{aligned} l_1(\lambda_1, \lambda_2, \beta|\mathcal{D}) &= n_3 \ln \lambda_1 + (n_1 + n_2) \ln(\lambda_1 + \lambda_2) - (n_3 - 1) \ln(2\lambda_1 + \lambda_2) + \\ &\quad \ln \Gamma(n_1 + n_2 + \beta - 1) - \ln \Gamma(\beta) + \beta \ln \beta - \\ &\quad (n_1 + n_2 + \beta - 1) \ln(\beta + \lambda_1 A_1(S_0, \mathcal{D}) + \lambda_2 A_2(S_0, \mathcal{D})). \end{aligned} \tag{27}$$

The MLEs of λ_1, λ_2 and β can be obtained by maximizing (27) with respect to the unknown parameters. It involves solving a three dimensional optimization problem. To avoid that we treat this problem as a missing value problem, and we use the EM algorithm. It is observed that it avoids solving three dimensional optimization problem. One needs to solve only one dimensional optimization problem. The basic idea is to use profile likelihood method to compute the MLEs of λ_1, λ_2 and β. For a given β, first we can obtain the MLEs of λ_1 and λ_2 using EM algorithm. Then we maximize the profile log-likelihood function of β to obtain the MLE of β. The exact procedure is as follows.

The complete data set in this case is $\mathcal{D}_c = (\mathcal{D}, v)$. First we assume that β is known. Now using (24) and the PDF of the gamma distribution, based on the complete data set, the log-likelihood function without the additive constant becomes

$$\begin{aligned} l_{1c}(\lambda_1, \lambda_2|\mathcal{D}_c, \beta) &= (n_1 + n_2 - 1) \ln v + n_3 \ln \lambda_1 + (n_1 + n_2) \ln(\lambda_1 + \lambda_2) - \\ &\quad (n_3 - 1) \ln(2\lambda_1 + \lambda_2) - v(\beta + \lambda_1 A_1(S_0, \mathcal{D}) + \lambda_2 A_2(S_0, \mathcal{D})). \end{aligned} \tag{28}$$

Now we re-parameterize $\lambda_1 = \gamma_1$, $\lambda_2 = \gamma_1 \times \gamma_2$ and let us denote $R(S_0, \mathcal{D}) = \dfrac{A_1(S_0\mathcal{D})}{A_2(S_0, \mathcal{D})}$. Suppose $\widetilde{\lambda}_1(v)$ and $\widetilde{\lambda}_2(v)$ maximize (28) and we denote $\widetilde{\gamma}_1(v) = \widetilde{\lambda}_1(v)$ and $\widetilde{\gamma}_2(v) = \dfrac{\widetilde{\lambda}_2(v)}{\widetilde{\lambda}_1(v)}$. Then it can be shown (see Appendix A) that $\widetilde{\gamma}_2(v)$ can be obtained as the positive root of the following quadratic equation

$$-n_3\gamma^2 + \gamma[R(S_0, \mathcal{D})(n + 1 - 2n_3) - n - 2] + R(S_0, \mathcal{D})(2n - 3n_3 + 1) - 2(n_1 + n_2 + 1) = 0. \tag{29}$$

and

$$\widetilde{\lambda}_1(v) = \widetilde{\gamma}_1(v) = \frac{n_1 + n_2 + 1}{v(A_1(S_0, \mathcal{D}) + \widetilde{\gamma}_2(v) A_2(S_0, \mathcal{D}))}. \tag{30}$$

Based on the above observation, we propose to use the following procedure to compute the MLEs of λ_1, λ_2 and β as follows. For fixed β, we use the EM algorithm, to compute the MLEs of λ_1 and λ_2, say $\widehat{\lambda}_1(\beta)$ and $\widehat{\lambda}_2(\beta)$, respectively using Algorithm 1. Suppose at the k-th stage of the EM algorithm, the estimates of λ_1 and λ_2 are $\lambda_1^{(k)}$ and $\lambda_2^{(k)}$, respectively.

Algorithm 1:

Step 1: Compute $\gamma_2^{(k)}$ as the positive root of (29). Note that it does not depend on k. Hence, we just need to compute it once only.

Step 2: Compute

$$a^{(k)} = \frac{n_1 + n_2 + \beta - 1}{\beta + \lambda_1^{(k)} A_1(S_0, \mathcal{D}) + \lambda_2^{(k)} A_2(S_0, \mathcal{D})}.$$

Step 3: Obtain

$$\lambda_1^{(k+1)} = \frac{n_1 + n_2 + 1}{a^{(k)}(A_1(S_0, \mathcal{D}) + \gamma_2^{(k+1)} A_2(S_0, \mathcal{D}))} \quad \text{and} \quad \lambda_2^{(k+1)} = \lambda_1^{(k+1)} \times \gamma_2^{(k+1)}.$$

Step 4: Go back to Step 2, and the process continues until convergence.

Once we get $\widehat{\lambda}_1(\beta)$ and $\widehat{\lambda}_2(\beta)$, we maximize $l_1(\widehat{\lambda}_1(\beta), \widehat{\lambda}_2(\beta), \beta|\mathcal{D})$ as in (27) with respect to β, to obtain the MLE of β, say $\widehat{\beta}$.

Exponential Baseline Distribution: In this case $h_0(x) = 1$ and $H_0(x) = x$, for $x > 0$. Hence,

$$\begin{aligned} A_1(S_0, \mathcal{D}) &= 2 \sum_{k \in I_1} x_{k+1} + \sum_{k \in I_2} x_{k+1} + \sum_{k \in I_3} x_k + 2x_1 - \sum_{k \in I_1} x_k \\ A_2(S_0, \mathcal{D}) &= \sum_{k \in I_1 \cup I_2} x_{k+1} + x_1 + \sum_{k \in I_3} x_k. \end{aligned}$$

Rayleigh Baseline Distribution: In this case $h_0(x) = 2x$ and $H_0(x) = x^2$, for $x > 0$. Hence,

$$\begin{aligned} A_1(S_0, \mathcal{D}) &= 2 \sum_{k \in I_1} x_{k+1}^2 + \sum_{k \in I_2} x_{k+1}^2 + \sum_{k \in I_3} x_k^2 + 2x_1^2 - \sum_{k \in I_1} x_k^2 \\ A_2(S_0, \mathcal{D}) &= \sum_{k \in I_1 \cup I_2} x_{k+1}^2 + x_1^2 + \sum_{k \in I_3} x_k^2. \end{aligned}$$

4.2. Baseline has Unknown Parameters

In this case it is assumed that the baseline survival function $S_0(\cdot)$ has some unknown parameter(s), say for example Weibull, Pareto or Lomax distributions, where baseline distributions have unknown parameters. In this section, we denote the baseline survival function, hazard function, cumulative hazard function and PDF as $S_0(t;\theta)$, $h_0(t;\theta)$, $H_0(t;\theta)$ and $f_0(t;\theta)$, respectively. Similarly, $A_1(S_0,\mathcal{D})$, $A_2(S_0,\mathcal{D})$ $R(S_0,\mathcal{D})$ as $A_1(S_0(\theta),\mathcal{D})$, $A_2(S_0(\theta),\mathcal{D})$ $R(S_0(\theta),\mathcal{D})$, respectively. Note that θ can be vector valued also. Based on the data $\mathcal{D}$, the log-likelihood function can be written as

$$\begin{aligned} l_2(\lambda_1,\lambda_2,\theta,\beta|\mathcal{D}) &= n_3\ln\lambda_1+(n_1+n_2)\ln(\lambda_1+\lambda_2)-(n_1+n_3-1)\ln(2\lambda_1+\lambda_2)+ \\ &\quad \ln\Gamma(n_1+n_2+\beta-1)-\ln\Gamma(\beta)+\beta\ln\beta- \\ &\quad (n_1+n_2+\beta-1)\ln(\beta+\lambda_1A_1(S_0(\theta),\mathcal{D})+\lambda_2A_2(S_0(\theta),\mathcal{D})). \end{aligned} \tag{31}$$

We modify the EM algorithm as follows. As before, first we fixed β. For fixed β, we use EM algorithm to compute the MLEs of λ_1, λ_2 and θ. To implement EM algorithm for fixed β, we re-write eqn. (28) as follows;

$$\begin{aligned} l_{2c}(\lambda_1,\lambda_2,\theta|\mathcal{D},v,\beta) &= (n_1+n_2-1)\ln v+n_3\ln\lambda_1+(n_1+n_2)\ln(\lambda_1+\lambda_2)- \\ &\quad (n_3-1)\ln(2\lambda_1+\lambda_2)-v(\beta+\lambda_1A_1(S_0(\theta),\mathcal{D})+\lambda_2A_2(S_0(\theta),\mathcal{D})). \end{aligned} \tag{32}$$

Now the EM algorithm proceeds as follows. Suppose at the k-th stage of the EM algorithm, the estimates of λ_1, λ_2 and θ are $\lambda_1^{(k)}$, $\lambda_2^{(k)}$ and $\theta^{(k)}$, respectively.

Algorithm 2:

Step 1: Compute $\gamma_2^{(k+1)}$ as the positive root of (29) replacing S_0 by $S_0(\theta^{(k)})$.

Step 2: Compute

$$a^{(k+1)}=\frac{n_1+n_2+\beta-1}{\beta+\lambda_1^{(k)}A_1(S_0(\theta^{(k)}),\mathcal{D})+\lambda^{(k)}A_2(S_0(\theta^{(k)}),\mathcal{D})}.$$

Step 3: Obtain $\theta^{(k+1)}$, by the argument maximum of $g(\theta)$, where

$$\begin{aligned} g(\theta) &= n_3\ln\lambda_1^{(k+1)}(\theta)+(n_1+n_2)\ln(\lambda_1^{(k+1)}(\theta)+\lambda_2^{(k+1)}(\theta))- \\ &\quad (n_3-1)\ln(2\lambda_1^{(k+1)}(\theta)+\lambda_2^{(k+1)}(\theta))- \\ &\quad a^{(k+1)}(\beta+\lambda_1^{(k+1)}(\theta)A_1(S_0(\theta),\mathcal{D})+\lambda_2^{(k+1)}(\theta)A_2(S_0(\theta),\mathcal{D})). \end{aligned} \tag{33}$$

Here

$$\lambda_1^{(k+1)}(\theta)=\frac{n_1+n_2+1}{a^{(k+1)}(A_1(S_0(\theta),\mathcal{D})+\gamma_2^{(k+1)}A_2(S_0(\theta),\mathcal{D}))} \quad\text{and}\quad \lambda_2^{(k+1)}(\theta)=\lambda_1^{(k+1)}(\theta)\times\gamma_2^{(k+1)}(\theta).$$

Note that (33) has been obtained from (32) by replacing v with its expectation at the k-th stage.

Step 4: Once $\theta^{(k+1)}$ is obtained obtain $\lambda_1^{(k+1)} = \lambda_1^{(k+1)}(\theta^{(k+1)})$ and $\lambda_2^{(k+1)} = \lambda_2^{(k+1)}(\theta^{(k+1)})$.

Step 5: Go back to Step 1, and the process continues until convergence.

Once we get $\widehat{\lambda}_1(\beta)$, $\widehat{\lambda}_2(\beta)$ and $\widehat{\theta}(\beta)$, we maximize $l_2(\widehat{\lambda}_1(\beta), \widehat{\lambda}_2(\beta), \widehat{\theta}(\beta), \beta|\mathcal{D})$ as in (31) with respect to β, to obtain the MLE of β, say $\widehat{\beta}$.

Weibull Baseline Distribution: In this case for $\theta = \alpha > 0$, $h_0(x) = \alpha x^{\alpha-1}$ and $H_0(x) = x^\alpha$, for $x > 0$. Hence,

$$\begin{aligned} A_1(S_0(\alpha), \mathcal{D}) &= 2\sum_{k\in I_1} x_{k+1}^\alpha + \sum_{k\in I_2} x_{k+1}^\alpha + \sum_{k\in I_3} x_k^\alpha + 2x_1^\alpha - \sum_{k\in I_1} x_k^\alpha \\ A_2(S_0(\alpha), \mathcal{D}) &= \sum_{k\in I_1\cup I_2} x_{k+1}^\alpha + \sum_{k\in I_3} x_k^\alpha + x_1^\alpha. \end{aligned}$$

Linear Hazard Rate Baseline Distribution: In this case for $\theta = \alpha$, where $\alpha > 0$ and $h_0(x) = 1 + 2\alpha x$ and $H_0(x) = x + \alpha x^2$, for $x > 0$. Hence,

$$\begin{aligned} A_1(S_0(\alpha), \mathcal{D}) &= 2\sum_{k\in I_1}(x_{k+1} + \alpha x_{k+1}^2) + \sum_{k\in I_2}(x_{k+1} + \alpha x_{k+1}^2) + \sum_{k\in I_3}(x_k + \alpha x_k^2) + 2(x_1 + \alpha x_1^2) - \\ &\quad \sum_{k\in I_1}(x_k + \alpha x_k^2) \\ A_2(S_0(\alpha), \mathcal{D}) &= \sum_{k\in I_1\cup I_2}(x_{k+1} + \alpha x_{k+1}^2) + \sum_{k\in I_3}(x_k + \alpha x_k^2) + (x_1 + \alpha x_1^2). \end{aligned}$$

5. Simulation Results

We have performed some simulation experiments to see how the proposed method works in practice for different models and for different parameter values and for different sample sizes. We have considered four different baseline distribution functions; (i) exponential, (ii) Rayleigh, (iii) Weibull and (iv) linear failure rate. Note that for (i) and (ii), the baseline distribution functions are completely known, and for (iii) and (iv), the baseline distribution functions have one unknown parameter. In all the cases we have taken four different sample sizes. In case of exponential and Rayleigh we have taken the following two sets of parameter values. Set 1: $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$ and Set 2: $\lambda_1 = 0.4$, $\lambda_2 = 0.2$, $\beta = 1.0$. In case of Weibull and linear failure rate distributions we have taken the following two sets of parameters. Set 3: $\alpha = 1.25$, $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$ and Set 4: $\alpha = 1.5$, $\lambda_1 = 0.4$, $\lambda_2 = 0.2$, $\beta = 1.0$.

In all the cases we have computed the MLEs of the unknown parameters based on EM algorithm. We have started the EM algorithm with all the initial guesses as 1 for all the parameters. We stop the iterations whenever the relative difference between two consecutive log-likelihood values is less than 10^{-6}. The results are reported in Tables 1 to 8. In all the cases the iterations stop within 15 iterations. Even if we start with other initial guesses, it has been observed that the algorithm converges to the same point, the number of iterations may be different. Some of the points are clear from the simulation results. The implementation of the proposed EM algorithm is quite simple in practice and it works quite well for both the cases. As the sample size increases in all the cases the biases and the mean squared errors (MSEs) decrease. It indicates the consistency properties of the MLEs of the unknown parameters.

Table 1: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is exponential. Here $\lambda_1 = 0.1$ and $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	λ_1	λ_2	β
25	0.1014	0.2098	1.2121
	(0.00015)	(0.00031)	(0.2312)
50	0.1010	0.2102	1.1889
	(0.00008)	(0.00022)	(0.2011)
75	0.1011	0.2047	1.1453
	(0.00006)	(0.00017)	(0.1587)
100	0.1009	0.2031	1.0298
	(0.00003)	(0.00011)	(0.1015)

Table 2: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is exponential. Here $\lambda_1 = 0.4$ and $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	λ_1	λ_2	β
25	0.4112	0.2047	1.2257
	(0.00051)	(0.00035)	(0.2417)
50	0.4089	0.2098	1.1778
	(0.00025)	(0.00027)	(0.2034)
75	0.4032	0.2055	1.1461
	(0.00017)	(0.00016)	(0.1452)
100	0.4008	0.2022	1.0115
	(0.00013)	(0.00014)	(0.1023)

6. Goodness of Fit Test

In this section we have proposed an empirical goodness of test of the proposed PHFP. The idea is similar to the idea proposed by the author in Kundu (2023). The problem can be mathematically formulated as follows. Suppose $\{X_1, \ldots, X_n\}$ is a sample from the PHFP$(\lambda_1, \lambda_2, \beta, S_0)$, we want to test the following hypothesis:

$$H_0 : \{X_1, \ldots, X_n\} \sim \text{PHFP}(\lambda_1, \lambda_2, \beta, S_0).$$

We denote $X_{1:n} < \ldots < X_{n:n}$ as the ordered $\{X_1, \ldots, X_n\}$ and the corresponding sample version of $\{x_1, \ldots, x_n\}$ as $\{x_{1:n} < \ldots < x_{n:n}\}$. We also denote $a_1 = E_{H_0}(X_{1:n}) \ldots, a_n = E_{H_0}(X_{n:n})$ as their ordered expected values under H_0. Although, $a_1, \ldots, a_n$ depend on θ, β, λ_1 and λ_2, we do not make it explicit for brevity. We use the following statistic for the goodness of fit test:

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

Table 3: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is Rayleigh. Here $\lambda_1 = 0.1$ and $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	λ_1	λ_2	β
25	0.1017	0.2110	1.2341
	(0.00018)	(0.00034)	(0.2298)
50	0.1015	0.2111	1.1799
	(0.00010)	(0.00023)	(0.2101)
75	0.1013	0.2057	1.1541
	(0.00008)	(0.00014)	(0.1498)
100	0.1004	0.2041	1.0119
	(0.00006)	(0.00012)	(0.1017)

Table 4: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is Rayleigh. Here $\lambda_1 = 0.4$ and $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	λ_1	λ_2	β
25	0.4222	0.2056	1.2311
	(0.00058)	(0.00041)	(0.2398)
50	0.4101	0.2078	1.1889
	(0.00031)	(0.00026)	(0.2067)
75	0.4057	0.2045	1.1367
	(0.00021)	(0.00017)	(0.1387)
100	0.4008	0.2018	1.0110
	(0.00015)	(0.00018)	(0.1016)

It is expected that W_n should be small under H_0, and large otherwise. We use the following test criterion for a given level of significance $0 < \gamma < 1$

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

where $c_n(\gamma)$ is such that

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

Clearly, $c_n(\gamma)$ also depends on the true parameter values, but we do not make it explicit here also. It is difficult to compute $c_n(\gamma)$ even asymptotically. Note that if we want to compute $c_n(\gamma)$, for a given (known) λ_1, λ_2, β and S_0, we can use the following simulation technique for this purpose. It is a two stage process, first we will compute $a_i = E_{H_0}(X_{i:n})$, for $i = 1, \ldots, n$, and then we will compute $c_n(\gamma)$. In the first stage we generate $\{(x_1^j, \ldots, x_n^j); j = 1, \ldots, N\}$ i.i.d. samples from PHFP$(\lambda_1, \lambda_2, \beta, S_0)$. Here, N is very large. We order them as $\{(x_{1:n}^j, \ldots, x_{n:n}^j); j = 1, \ldots, N\}$. Now we approximate a_i's as $\widetilde{a}_i = \dfrac{1}{N}\sum_{j=1}^{N} x_{i:n}^j$, for $i = 1, \ldots, n$.

Table 5: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is Weibull. Here $\alpha = 1.25$, $\lambda_1 = 0.1$, $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	α	λ_1	λ_2	β
25	1.2378	0.1074	0.2217	1.3551
	(0.0356)	(0.00032)	(0.00178)	(0.3287)
50	1.2454	0.1031	0.2165	1.2487
	(0.0256)	(0.00019)	(0.00113)	(0.2619)
75	1.2478	0.1014	0.2119	1.1921
	(0.0178)	(0.00013)	(0.00098)	(0.1998)
100	1.2501	0.1011	0.2087	1.1101
	(0.0093)	(0.00009)	(0.00051)	(0.1315

Table 6: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is Weibull. Here $\alpha = 1.5$, $\lambda_1 = 0.4$, $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	α	λ_1	λ_2	β
25	1.4218	0.4123	0.2278	1.3217
	(0.0512)	(0.00056)	(0.00047)	(0.3671)
50	1.4587	0.4101	0.2167	1.2541
	(0.0468)	(0.00035)	(0.00029)	(0.2589)
75	1.4781	0.4019	0.2141	1.1317
	(0.0328)	(0.00023)	(0.00018)	(0.1991)
100	1.4981	0.4004	0.2015	1.1127
	(0.0167)	(0.00015)	(0.00011)	(0.1418)

At the second stage we can approximate $c_n(\gamma)$ as follows. We generate again $\{(y_1^j, \ldots, y_n^j); j = 1, \ldots, M\}$ i.i.d. samples from PHFP$(\lambda_1, \lambda_2, \beta, S_0)$. Here, M is very large. We order them as $\{(y_{1:n}^j, \ldots, y_{n:n}^j); j = 1, \ldots, M\}$. We compute $\{(w_n^j = \max_{1 \le i \le n} |y_{i:n}^j - \tilde{a}_i|); j = 1, \ldots, M\}$. We order $\{w_n^j; j = 1, \ldots, M\}$. The upper γ-th percentile point will be an approximation of $c_n(\gamma)$. As $M \to \infty$ and $N \to \infty$, both the approximate values will converge to the true values.

In practice we do not have the true values of λ_1, λ_2, θ and β. Hence, we suggest the following simulation technique to approximate $c_n(\gamma)$ from the given sample $\{x_1, \ldots, x_n\}$.

Algorithm:

Step 1: Obtain $\widehat{\beta}$, $\widehat{\lambda}_1$, $\widehat{\lambda}_2$ and $\widehat{\theta}$, the MLEs of β, λ_1, λ_2 and θ, respectively, based on the sample $\{x_1, \ldots, x_n\}$.

Table 7: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is linear failure rate. Here $\alpha = 1.25$, $\lambda_1 = 0.1$, $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	α	λ_1	λ_2	β
25	1.2392	0.1101	0.2234	1.3481
	(0.0348)	(0.00029)	(0.00181)	(0.3198)
50	1.2448	0.1029	0.2157	1.2501
	(0.0268)	(0.00020)	(0.00117)	(0.2598)
75	1.2498	0.1016	0.2121	1.2018
	(0.0169)	(0.00013)	(0.00101)	(0.1889)
100	1.2505	0.1014	0.2101	1.1008
	(0.0101)	(0.00011)	(0.00071)	(0.1298)

Table 8: Average estimates and the associated mean squared errors based on 1000 replications. The baseline distribution is linear failure rate. Here $\alpha = 1.5$, $\lambda_1 = 0.4$, $\lambda_2 = 0.2$ and $\beta = 1.0$.

n	α	λ_1	λ_2	β
25	1.4198	0.4117	0.2175	1.3210
	(0.0499)	(0.00078)	(0.00052)	(0.3598)
50	1.4489	0.4098	0.2158	1.2482
	(0.0387)	(0.00056)	(0.00041)	(0.2437)
75	1.4844	0.4023	0.2137	1.1254
	(0.0316)	(0.00033)	(0.00025)	(0.1783)
100	1.4991	0.4011	0.2023	1.1015
	(0.0178)	(0.00018)	(0.00015)	(0.1366)

- Step 2: Generate a sample of size n from a PHFP($\widehat{\beta}$,$\widehat{\lambda}_1$, $\widehat{\lambda}_2$,$S_0(\widehat{\theta})$) and order them. Let us denote them as $(x^1_1, \ldots, x^1_{n:n})$. Repeat this procedure $B_1 + B_2$ times, and obtain $\{(x^b_{1:n}, \ldots, x^b_{n:n}); b = 1, \ldots, B_1 + B_2\}$.

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

$$\widehat{a}_i = \frac{1}{B_2} \sum_{b=B_1+1}^{B_1+B_2} x^b_{i:n}; \quad i = 1, \ldots, n.$$

- Step 4: Compute

$$w^b = \max_{1 \le i \le n} \{|x^b_{i:n} - \widehat{a}_i|\}; \quad b = 1, \ldots, B_1.$$

- Step 5: Order $\{w_1, \ldots, w_{B_1}\}$ as $w^{(1)} < \ldots < w^{(B_1)}$, then $\widehat{c}_n(\gamma) = w^{([B_1 100(1-\gamma)])}$ is an estimate of $c_n(\gamma)$.

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

We have performed some small simulation experiments to see how the proposed method works in practice. We have taken the following two case: Case 1: H_0: $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$ and the baseline is exponential. We have considered different H_1: (i) $\lambda_1 = 0.1$, $\lambda_1 = 0.2$, $\beta = 1.0$ and the baseline is Rayleigh, (ii) $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$, and the baseline distribution is Weibull with $\alpha = 1.25$, (iii) $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$, and the baseline distribution is linear failure rate with $\alpha = 1.25$. Case 2: H_0: $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$, and the baseline distribution is Weibull with $\alpha = 1.25$, We have considered different H_1: (i) $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$ and the baseline is exponential. (ii) $\lambda_1 = 0.1$, $\lambda_1 = 0.2$, $\beta = 1.0$ and the baseline is Rayleigh, (iii) $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$, and the baseline distribution is linear failure rate with $\alpha = 1.25$. For both the cases, we have taken $\gamma = 0.05$. We have computed the size and the power of the proposed test, reported below in brackets in each box, for different sample sizes based on 1000 replications. We have used $M = N = 500$ to compute a_i's and $c_n(\gamma)$, respectively. The results are reported in Table 9 and Table 10, respectively.

Table 9: The size and power of the goodness of fit test for Case 1. Both under H_0 and H_1 in all the cases $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$. H_0: $S_0(x) = e^{-x}$ and H_1: (i) $S_0(x) = e^{-x^2}$, (ii) $S_0(x) = e^{-x^{1.25}}$, (iii) $S_0(x) = e^{-x-1.25x^2}$.

n	Rayleigh	Weibull	LFR
25	0.059	0.059	0.059
	(0.478)	(0.515)	(0.495)
50	0.057	0.057	0.057
	(0.531)	(0.567)	(0.549)
75	0.053	0.053	0.053
	(0.721)	(0.849)	(0.757)
100	0.052	0.052	0.052
	(0.811)	(0.838)	(0.822)

It is observed that the proposed test works works well in practice. As the sample size increase it is observed that it maintains the level of the test and the power also increases for all the different alternatives.

7. Data Analysis

This data set represents the price of 1gm of 24 Carat (Kt) gold in Indian Rupees from Dec. 01, 2020 till Jan 31, 2021 of total 62 days. The gold price data has been plotted in Figure 1. 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. Some of the basic statistics are as follows: The minimum and maximum values are 4230.02 and 4515.21, respectively. The 25th percentile, median and 75th percentile values are 4433.57, 4486.16 and 4515.21, respectively. The augmented Dickey-Fuller test suggests that it is from a stationary time series. The auto-correlation function (ACF) and partial auto-correlation function (PACF) have been plotted in Figure 2. It indicates that it possesses

Table 10: The size and power of the goodness of fit test for Case 1. Both under H_0 and H_1 in all the cases $\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\beta = 1.0$. H_0: $S_0(x) = e^{-x^{1.25}}$ and H_1: (i) $S_0(x) = e^{-x}$, (ii) $S_0(x) = e^{-x^2}$, , (iii) $S_0(x) = e^{-x-1.25x^2}$.

n	Exponential	Rayleigh	LFR
25	0.061	0.061	0.061
	(0.611)	(0.423)	(0.472)
50	0.056	0.056	0.056
	(0.721)	(0.522)	(0.538)
75	0.052	0.052	0.052
	(0.831)	(0.613)	(0.698)
100	0.049	0.049	0.049
	(0.911)	(0.745)	(0.811)

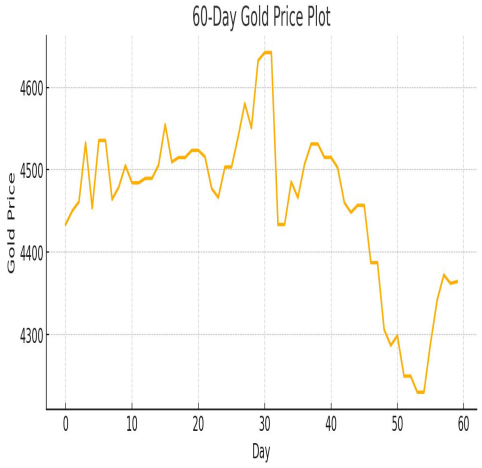


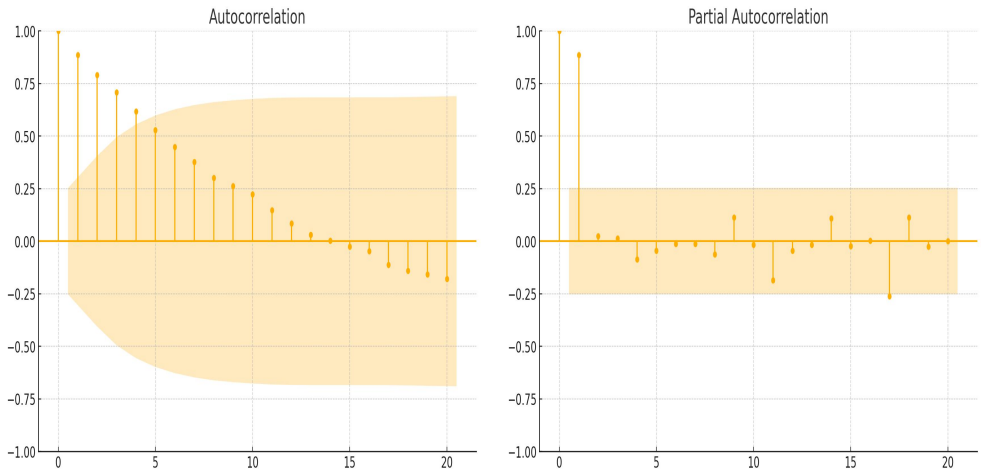
Figure 1: (a) The plot of the gold price data.

persistent dependence.

We want to fit all the four models to this data set. First we have subtracted 4225 from all the data points. We have computed the MLEs of the unknown parameters and the associated 95% confidence intervals, for each model based on EM algorithm as it has been suggested in Section 4. We have also performed the goodness of fit test based on parametric bootstrap of 500 replications, as suggested in Section 6 and the obtained the associated p-values. The results are reported in Table 11. From the table values it is clear that the proposed PHFP model with the baseline distribution as the Weibull provides the best fit to the gold price data.

8. Conclusions

In this paper we have proposed a new discrete time continuous state space stationary process with $P(X_n = X_{n+1}) > 0$. It is possessing persistent dependence and the marginals



(a) (b)

Figure 2: (a) The autocorrelation function and (b) the partial autocorrelation function of the gold price data.

can be very flexible depending on the baseline distribution. We have studied different properties of the proposed process, and we have considered in details four different baseline distributions. We have developed an efficient EM algorithm to compute the MLEs of the unknown parameters. We have provided an empirical goodness of fit test also. Indian gold price data has been analyzed using this model, and it is observed that the proposed model can be used quite effectively to analyze the gold price data of the Indian market.

Appendix A

In this Appendix, we will show how (29) can be obtained. Note that using the notation as in Section 4, (28) can be written as

$$\begin{aligned} l_{1c}(\gamma_1,\gamma_2|\mathcal{D}_c,\beta) &= (n_1+n_2-1)\ln v + n_3\ln\gamma_1 + (n_1+n_2)\ln(\gamma_1+\gamma_1\gamma_2) - \\ &\quad (n_3-1)\ln(2\gamma_1+\gamma_1\gamma_2) - v\beta - v(\gamma_1 A_1(S_0,\mathcal{D}) + \gamma_1\gamma_2 A_2(S_0,\mathcal{D})). \\ &= (n_1+n_2-1)\ln v + n_3\ln\gamma_1 + (n_1+n_2)\ln\gamma_1 + (n_1+n_2)\ln(1+\gamma_2) - \\ &\quad (n_3-1)\ln\gamma_1 - (n_3-1)\ln(2+\gamma_2) - v\beta - v\gamma_1 A_2(S_0,\mathcal{D})(R_0(S_0,\mathcal{D})+\gamma_2) \\ &= (n_1+n_2-1)\ln v - v\beta + (n_1+n_2+1)\ln\gamma_1 - v\gamma_1 A_2(S_0,\mathcal{D})(R_0(S_0,\mathcal{D})+\gamma_2) + \\ &\quad (n_1+n_2)\ln(1+\gamma_2) - (n_3-1)\ln(2+\gamma_2). \end{aligned} \tag{34}$$

Hence, the argument maximum of (34) with respect to γ_1 and γ_2 is same as the argument maximum of

$$\begin{aligned} l_0(\gamma_1,\gamma_2) &= (n_1+n_2+1)\ln\gamma_1 - v\gamma_1 A_2(S_0,\mathcal{D})(R_0(S_0,\mathcal{D})+\gamma_2) + (n_1+n_2)\ln(1+\gamma_2) - \\ &\quad (n_3-1)\ln(2+\gamma_2). \end{aligned} \tag{35}$$

Note that the maximization of (35) can be obtained in two steps. Observe that for fixed γ_2, if $\widetilde{\gamma}_1(\gamma_2)$ maximizes (35), then

$$\widetilde{\gamma}_1(\gamma_2) = \frac{n_1+n_2+1}{vA_2(S_0,\mathcal{D})(R_0(S_0,\mathcal{D})+\gamma_2)}.$$

Table 11: The MLEs and the associated confidence intervals of the unknown parameters for different models in case of gold price data. The p-values of the corresponding goodness of fit test have been reported.

Model	λ_1	λ_2	β	θ	log-lik	p-value
Exponential	0.2314×10^{-2}	0.7891×10^{-2}	0.9671		-401.121	0.005
	(0.2134×10^{-3})	(0.4541×10^{-3})	(0.2019)			
Rayleigh	0.9867×10^{-2}	1.9878×10^{-2}	1.3878		-382.453	0.071
	(0.4325×10^{-3})	(0.7531×10^{-3})	(0.2341)			
Weibull	1.6376×10^{-2}	3.0394×10^{-2}	1.6956	2.1518	-324.7195	0.344
	(0.9167×10^{-3})	(1.1165×10^{-3})	(0.4562)	(0.7545)		
LFR	3.6751×10^{-2}	4.2541×10^{-2}	2.1165	1.9876	-378.556	0.154
	(1.2134×10^{-3})	(1.7341×10^{-3})	(0.7334)	(0.6874)		

At the second step we obtain γ_2 that maximizes $l_0(\tilde{\gamma}_1(\gamma_2), \gamma_2)$. Now,

$$\begin{aligned} l_0(\tilde{\gamma}_1(\gamma_2), \gamma_2) &= (n_1+n_2+1)\ln\tilde{\gamma}_1(\gamma_2) - v\tilde{\gamma}_1(\gamma_2)A_2(S_0,\mathcal{D})(R_0(S_0,\mathcal{D})+\gamma_2) + \\ &\quad (n_1+n_2)\ln(1+\gamma_2) - (n_3-1)\ln(2+\gamma_2) \\ &= c - (n_1+n_2+1)\ln(R_0(S_0,\mathcal{D})+\gamma_2) + (n_1+n_2)\ln(1+\gamma_2) - (n_3-1)\ln(2+\gamma_2). \end{aligned}$$

Here c denotes the quantity which is independent of γ_2. To find the value of γ_2 which maximizes $l_0(\tilde{\gamma}_1(\gamma_2), \gamma_2)$, we differentiate it with respect to γ_2 and equate it to zero. It leads to solving

$$-\frac{n_1+n_2+1}{R_0(S_0,\mathcal{D})+\gamma_2} + \frac{n_1+n_2}{1+\gamma_2} - \frac{n_3-1}{2+\gamma_2} = 0. \tag{36}$$

To find $\gamma_2 > 0$, which satisfies (36), is equivalent to find the positive root of the quadratic equation (29).

Acknowledgement

The author would like to thank the unknown reviewer for his/her constructive suggestions which have helped to improve the manuscript significantly.

Conflict of interest

The author does not have any financial or non-financial conflict of interest to declare for the research work included in this article.

References

Arnold, B. and Hallet, T. (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*, **8**, 377–380.

Arnold, B., Sachdev, S., and Manjunath, B. (2026). Some power function distribution process. *Metrika*, **doi:10.1007/s00184-026-01019-4**.

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

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.

Cox, D. and Oakes, D. (1984). *Analysis of Survival Data*. Chapman & Hall.

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

Jose, K., Naik, S., and Ristić, M. (2010). Marshall-Olkin q-Weibull distribution and max-min processes. *Statistical Papers*, **51**, 837–851.

Kundu, D. (2023). Bivariate distributions with singular components. *Springer Handbook of Engineering Statistics, Editor: Hoang Pham*, , 733–761.

Kundu, D. (2024). A stationary proportional hazard class process and its applications. *Methodology and Computing in Applied Probability*, **26**, Article no. 45.

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

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

Pillai, R. (1991). Semi-pareto processes. *Journal of Applied Probability*, **28**, 461–465.

Sim, C. (1986). Simulation of weibull and gamma autoregressive stationary processes. *Communications in Statistics - Simulation and Computation*, **15**, 1141–1146.

Surya, B. (2007). An approach for solving perpetual optimal stopping problems driven by lévy processes. *Stochastics*, **79**, 337–361.

Tavares, L. (1980). An exponential markovian stationary process. *Journal of Applied Probability*, **17**, 1117–1120.

Yeh, H., Arnold, B., and Robertson, C. (1988). Pareto process. *Journal of Applied Probability*, **25**, 291–301.