A Stationary Weibull-Process and its Applications

Debasis Kundu *

Abstract

In this paper we introduce a discrete time and continuous state space Markov stationary process $\{X_n; n = 1, 2, \ldots\}$, where X_n has a two parameter Weibull distribution, X_n's are dependent and there is a positive probability that $X_n = X_{n+1}$. The motivation came from the gold price data where there are several instances for which $X_n = X_{n+1}$. Hence, the existing methods cannot be used to analyze this data. We derive different properties of the proposed Weibull process. It is observed that the joint cumulative distribution function of X_n and X_{n+1} has a very convenient copula structure. Hence, different dependence properties and dependence measures can be obtained. The maximum likelihood estimators cannot be obtained in explicit forms, we have proposed a simple profile likelihood method to compute these estimators. We have used this model to analyze two synthetic data sets and one gold price data set of the Indian market, and it is observed that the proposed model fits quite well to the data set.

Key Words and Phrases: Weibull distribution, exponential distribution; maximum likelihood estimators; minification process; maximum likelihood predictor.

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

1 Introduction

The aim of this paper is to introduce a new discrete time and continuous state space Markov Weibull process $\{X_n; n = 1, 2, \ldots\}$, which is flexible and it has certain distinct features so

*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.

that it can be used in practice in different areas. In this case the marginals are non-negative two-parameter Weibull distributions and X_ns are dependent. It is a lag-1 process based on minimization approach, and $X_n = X_{n+1}$ with a positive probability. Therefore, if there is a non-negative time series data which are positively skewed and there is a positive probability that the consecutive values might be equal then the proposed model can be used quite effectively to analyze this data set.

The motivation of this work came when we were trying to analyze gold price data of the Indian market during the month of December, 2020 and January 2021. It is observed that the data are coming from a positive valued lag-1 stationary process, there are several instances for which $X_n = X_{n+1}$, and this cannot be ignored. It may be mentioned that there are several positive valued stationary processes available in the literature, for example the exponential process by Tavares [18], Weibull and Gamma processes by Sim [17], the logistic process by Arnold [2], Pareto process by Arnold and Hallet [3], semi-Parero process by Pillai [16], generalized Weibull process by Jayakumar and Girish Babu [9], see also Yeh, Arnold and Robertson [19], Arnold and Robertson [4], Jose, Ristić and Joseph [10] and the reference cited therein. But none of these can be applied in this case, as in all these cases $P(X_n = X_{n+1}) = 0$.

We have provided different properties of the proposed process $\{X_n\}$. The Weibull process $\{X_n\}$ has one shape parameter and two scale parameters. If the shape parameter is one, it becomes a stationary exponential process. The generation from $\{X_n\}$ is quite straight forward, hence, different simulation experiments can be performed quite conveniently. The joint distribution of X_n and X_{n+1} has a very convenient copula structure. Therefore, several dependence properties and the dependence measures can be computed quite conveniently. We provide a characterization of the process. The marginals and the joint PDF can take variety of shapes. The autocovariance and autocorrelation are not in convenient forms.

However, we have provided the necessary expressions in the Appendix for completeness purposes.

The maximum likelihood estimators (MLEs) cannot be obtained in explicit form. Moreover, it cannot be obtained in a routine manner. Based on some re-parameterization and using profile likelihood method the MLEs can be obtained. Parametric bootstrap method can be used to compute the confidence intervals of the unknown parameters. We propose a goodness of fit test based on parametric bootstrap approach. We have analyzed two synthetic data sets and one gold price data set for two months of the Indian market. It is observed that the the proposed model fits the gold price data set quite well and it can be used quite effectively to analyze the gold price of the Indian market. It may be mentioned that the proposed process is a lag one process, but it can be extended to a lag-q process also, and it has been indicated how it can be done.

I think the major difference with the existing literature and the present manuscript is in its construction. The present manuscript allows to have ties, which are not available in the literature. Moreover, in the present manuscript we have provided the detailed inference procedure and showed with real data example how it can be implemented in practice. This seems to be the main contribution of the present manuscript.

The rest of the paper is organized as follows. In Section 2 we have defined the Weibull process and provided its different properties. The maximum likelihood estimators have been discussed in Section 3. In Section 4 one goodness of fit test has been proposed based on parametric bootstrap approach. The analyses of three data sets have been presented in Section 5. Finally we conclude the paper in Section 6.

2 Weibull Process and its Properties

We will use the following notations in this paper. A Weibull random variable with the shape parameter $\alpha > 0$ and the scale parameter $\lambda > 0$ has the following probability density function (PDF);

$$f_{WE}(x;\alpha,\lambda) = \begin{cases} \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}} & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x \le 0, \end{cases} \tag{1}$$

and it will be denoted by WE(α, λ). It has the following cumulative distribution function and hazard function, respectively, for $x > 0$;

$$\begin{aligned} F_{WE}(x;\alpha,\lambda) &= 1 - e^{-\lambda x^{\alpha}}, \\ h_{WE}(x;\alpha,\lambda) &= \alpha\lambda x^{\alpha-1}. \end{aligned}$$

The mean and variance of WE(α, λ) is

$$\frac{1}{\lambda^{1/\alpha}}\Gamma\left(\frac{1}{\alpha}+1\right) \quad \text{and} \quad \frac{1}{\lambda^{2/\alpha}}\left[\Gamma\left(\frac{2}{\alpha}+1\right) - \left(\Gamma\left(\frac{1}{\alpha}+1\right)\right)^2\right], \tag{2}$$

respectively. An uniform random variable on (0,1) will be denoted by $U(0,1)$. Now, we are in a position to define the Weibull process.

Definition: Suppose $U_0, U_1, \ldots$ are independent identically distributed (i.i.d.) $U(0,1)$ random variables, then for $\lambda_0 > 0$, $\lambda_1 > 0$ and $\alpha > 0$, let us define a new sequence of random variables $\{X_n; n = 1, 2, \ldots\}$, where

$$X_n = \min\left\{\left[-\frac{1}{\lambda_0}\ln U_n\right]^{\frac{1}{\alpha}}, \left[-\frac{1}{\lambda_1}\ln U_{n-1}\right]^{\frac{1}{\alpha}}\right\}. \tag{3}$$

Then the sequence of random variables $\{X_n\}$ is called a Weibull process. ∎

A Weibull process as defined in (3) will be denotes by WEP$(\alpha, \lambda_0, \lambda_1)$. Here α is the shape parameter and λ_0 and λ_1 are the scale parameters. It may be mentioned that a location parameter can easily be incorporated in the model, which has not been tried here. The name Weibull process comes from the following results.

Theorem 1: If $\{X_n\}$ is as defined in (3), then

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

(ii) X_n follows $\text{WE}(\alpha, \lambda_0 + \lambda_1)$.

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

$$\begin{aligned} P(X_n > x) &= P\left\{\left[-\frac{1}{\lambda_0}\ln U_n\right]^{\frac{1}{\alpha}} > x, \left[-\frac{1}{\lambda_1}\ln U_{n-1}\right]^{\frac{1}{\alpha}} > x\right\} \\ &= P\left\{U_n < e^{-\lambda_0 x^\alpha}, U_{n-1} < e^{-\lambda_1 x^\alpha}\right\} \\ &= P\left\{U_n < e^{-\lambda_0 x^\alpha}\right\} P\left\{U_{n-1} < e^{-\lambda_1 x^\alpha}\right\} \\ &= e^{-(\lambda_0+\lambda_1)x^\alpha}. \end{aligned}$$

The following result characterizes the Weibull process.

Theorem 2: Let $X_1 \sim \text{WE}(\alpha, \lambda_0+\lambda_1)$, and U_is are i.i.d. random variables with an absolute continuous distribution function $F(\cdot)$ on (0,1). Then the process as defined in (3) is a strictly stationary Markov process if and only if $U_0 \sim U(0,1)$.

Proof: 'If' part is trivial. Now to prove the 'only if' part, we assume $S(x) = 1 - F(x)$, and $F'(x) = -S'(x) = f(x)$, for $x > 0$. Therefore, for $x > 0$,

$$e^{-(\lambda_0+\lambda_1)x^\alpha} = F(e^{-\lambda_0 x^\alpha})F(e^{-\lambda_1 x^\alpha}).$$

If we write $y = e^{-x^\alpha}$, then for $0 < y < 1$,

$$y^{\lambda_0+\lambda_1} = F(y^{\lambda_0})F(y^{\lambda_1}) \Rightarrow \frac{F(y^{\lambda_0})}{y^{\lambda_0}} \times \frac{F(y^{\lambda_1})}{y^{\lambda_1}} = 1 \Rightarrow \frac{F(y)}{y} = 1 \Rightarrow F(y) = y. \qquad \blacksquare$$

Now we present the joint distribution of X_n and X_{n+m}, for $m \geq 1$.

Theorem 3: If $\{X_n\}$ satisfies (3), then the joint survival function of X_n and X_{n+m}, $S_{n,n+m}(x,y) = P(X_n > x, X_{n+m} > y)$ is

$$S_{n,n+m}(x,y) = \begin{cases} e^{-(\lambda_0+\lambda_1)x^\alpha} e^{-(\lambda_0+\lambda_1)y^\alpha} & \text{if} \quad m \geq 2 \\ e^{-\lambda_1 x^\alpha} e^{-\lambda_0 y^\alpha} g(x,y) & \text{if} \quad m = 1, \end{cases} \tag{4}$$

where $g(x,y) = \min\{e^{-\lambda_0 x^\alpha}, e^{-\lambda_1 y^\alpha}\}$.

PROOF: The proof is quite simple and it is avoided. ∎

The above theorem indicates that X_n and X_{n+m} are dependent for $m = 1$, and they are independently distributed if $m > 1$. This makes the process as the lag-1 process. The joint distribution function of X_n and X_{n+1} will help to develop the dependence properties of the Weibull process and we would like to study it in more details. The joint survival function of X_n and X_{n+1} can be written explicitly as

$$S_{n,n+1}(x,y) = \begin{cases} e^{-\lambda_1 x^\alpha} e^{-(\lambda_0+\lambda_1)y^\alpha} & \text{if} \quad \lambda_0 x^\alpha < \lambda_1 y^\alpha \\ e^{-(\lambda_0+\lambda_1)x^\alpha} e^{-\lambda_0 y^\alpha} & \text{if} \quad \lambda_0 x^\alpha > \lambda_1 y^\alpha \\ e^{-z\frac{\lambda_1^2+\lambda_0^2+\lambda_0\lambda_1}{\lambda_0\lambda_1}} & \text{if} \quad \lambda_0 x^\alpha = \lambda_1 y^\alpha = z. \end{cases} \tag{5}$$

Therefore, if $\lambda_0 = \lambda_1 = \lambda$, then

$$S_{n,n+1}(x,y) = \begin{cases} e^{-\lambda x^\alpha} e^{-2\lambda y^\alpha} & \text{if} \quad x < y \\ e^{-2\lambda x^\alpha} e^{-\lambda y^\alpha} & \text{if} \quad x > y \\ e^{-3\lambda z^\alpha} & \text{if} \quad x = y = z. \end{cases} \tag{6}$$

It may be mentioned that (6) is the joint survival function of the Marshall-Olkin bivariate Weibull distribution, and its properties have been well studied in the literature. See for example Kundu and Dey [12] and Kundu and Gupta [14] and the references cited therein. It can be easily seen that the $S_{n,n+1}(x,y)$, for $0 < \delta < 1$, has the following survival copula

$$\widetilde{C}(u,v) = \begin{cases} u^\delta v & \text{if} \quad u^\delta > v^{1-\delta} \\ uv^{1-\delta} & \text{if} \quad u^\delta \le v^{1-\delta}. \end{cases} \tag{7}$$

The corresponding copula density function becomes

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

Based on the copula density function, the Spearman's ρ and Kendall's τ can be obtained as

$$\begin{aligned} \rho &= \frac{3\delta(1-\delta)}{\delta^2 - \delta + 2} \\ \tau &= \frac{\delta(1-\delta)(1-\delta(1-\delta))}{\delta^3 + \delta(1-\delta) + \delta^2(1-\delta^2) + (1-\delta)^3}, \end{aligned}$$

respectively.

We will introduce the following regions, which will be used later.

$$\begin{aligned} S_1 &= \{(x,y); x>0, y>0, \beta x < y\} \\ S_2 &= \{(x,y); x>0, y>0, \beta x > y\} \\ C &= \{(x,y); x>0, y>0, \beta x = y\}. \end{aligned}$$

Here, $\beta = (\lambda_0/\lambda_1)^{1/\alpha}$, and it may be noted that the curve C has the parametric form $(t, \gamma(t))$, for $0 < t < \infty$, where $\gamma(t) = \beta t$. The following results are needed for further development.

THEOREM 4: If $\{X_n\}$ satisfies (3), then the joint survival function of X_n and X_{n+1} can be written as

$$S_{n,n+1}(x,y) = pS_a(x,y) + (1-p)S_s(x,y), \tag{9}$$

here $p = \dfrac{\lambda_0^2 + \lambda_1^2}{\lambda_0^2 + \lambda_1^2 + \lambda_0\lambda_1}$,

$$S_s(x,y) = (g(x,y))^{\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_0\lambda_1}},$$

and $S_a(x,y)$ can be obtained by subtraction, i.e.

$$S_a(x,y) = \begin{cases} \frac{1}{p} e^{-\lambda_1 x^\alpha} e^{-(\lambda_0+\lambda_1)y^\alpha} - \frac{1-p}{p} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_0} y^\alpha} & \text{if} \quad \beta x < y \\ \frac{1}{p} e^{-(\lambda_0+\lambda_1)x^\alpha} e^{-\lambda_0 y^\alpha} - \frac{1-p}{p} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1} x^\alpha} & \text{if} \quad \beta x > y. \end{cases} \tag{10}$$

PROOF: See in the Appendix. ∎

Now we provide the joint probability density function (PDF) of X_n and X_{n+1}, and because of the Markov property, it will be useful to compute the joint PDF of $X_1, \ldots, X_n$. It should be mentioned that since the joint distribution (survival) function is not an absolutely continuous distribution function the joint PDF does not exist in the terms of two dimensional Lebesgue measure dominating. In this case we need to consider the dominating measure in a different way similarly as in Bemis, Bain and Higgins [6]. Here, the dominating measure is the two-dimensional Lebesgue measure on $S_1 \cup S_2$, and one dimensional Lebesgue measure on the curve C. Based on the above dominating measure the joint PDF of X_n and X_{n+1}, for $x >$ and $y > 0$, can be written as follows.

THEOREM 5: If X_n satisfies (3), then the joint PDF of X_n and X_{n+1} is

$$f_{n,n+1}(x,y) = \begin{cases} f_1(x,y) & \text{if} \quad \beta x < y, \\ f_2(x,y) & \text{if} \quad \beta x > y, \\ f_0(x) & \text{if} \quad \beta x = y, \end{cases} \tag{11}$$

where

$$\begin{aligned} f_1(x,y) &= f_{WE}(x;\alpha,\lambda_1) f_{WE}(y;\alpha,\lambda_0+\lambda_1), \\ f_2(x,y) &= f_{WE}(x;\alpha,\lambda_0+\lambda_1) f_{WE}(y;\alpha,\lambda_0), \\ f_0(x) &= \frac{\alpha\lambda_0}{\beta} x^{\alpha-1} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1} x^\alpha}. \end{aligned}$$

PROOF: See in the Appendix. ∎

The following conditional PDF will be useful for prediction purposes. The conditional PDF of X_{n+1} given X_n can be written as follows:

$$f_{X_{n+1}|X_n=x}(y) = \begin{cases} \frac{\lambda_1}{\lambda_0+\lambda_1} e^{\lambda_0 x^\alpha} f_{WE}(y;\alpha,\lambda_0+\lambda_1) & \text{if} \quad \beta x < y \\ f_{WE}(y;\alpha,\lambda_0) & \text{if} \quad \beta x > y \\ \frac{\lambda_0}{\lambda_0+\lambda_1} e^{-\frac{\lambda_0^2}{\lambda_1} x^\alpha} & \text{if} \quad \beta x = y \end{cases} \tag{12}$$

When $\alpha = 1$, it becomes an exponential process. When $\lambda_0 = \lambda_1 = \lambda$, then the joint PDF of X_n and X_{n+1} for an Weibull process is

$$f_{n,n+1}(x,y) = \begin{cases} 2\alpha^2\lambda^2 x^{\alpha-1} e^{-\lambda x^\alpha} y^{\alpha-1} e^{-2\lambda y^\alpha} & \text{if} \quad x < y, \\ 2\alpha^2\lambda^2 x^{\alpha-1} e^{-2\lambda x^\alpha} y^{\alpha-1} e^{-\lambda y^\alpha} & \text{if} \quad x > y, \\ \alpha\lambda x^{\alpha-1} e^{-3\lambda x^\alpha} & \text{if} \quad x = y. \end{cases}$$

The conditional PDF of X_{n+1} given X_n can be written as

$$f_{X_{n+1}|X_n=x}(y) = \begin{cases} \frac{1}{2} e^{\lambda x^\alpha} f_{WE}(y;\alpha,2\lambda) & \text{if} \quad x < y \\ f_{WE}(y;\alpha,\lambda) & \text{if} \quad y > x \\ \frac{1}{2} e^{-\lambda x^\alpha} & \text{if} \quad x = y \end{cases}$$

It follows that both for the Weibull and exponential processes,

$$P(X_n = X_{n+1}) = P(X_n < X_{n+1}) = P(X_n > X_{n+1}) = \frac{1}{3}.$$

The autocovariance and autocorrelation of a Weibull process cannot be obtained in convenient forms. In case of exponential process, however, they can be obtained in explicit forms. We have provided all the necessary expressions in Appendix B, for completeness purposes.

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

$$
\begin{aligned}
f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= f_{X_n|X_{n-1},\ldots,X_1}(x_n) \times \ldots \times f_{X_2|X_1}(x_2) \times f_{X_1}(x_1) && (13) \\
&= f_{X_n|X_{n-1}}(x_n) \times \ldots \times f_{X_2|X_1}(x_2) \times f_{X_1}(x_1) && (14) \\
&= \frac{f_{X_{n-1},X_n}(x_{n-1}, x_n)}{f_{X_{n-1}}(x_{n-1})} \times \ldots \times \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)} \times f_{X_1}(x_1) && (15) \\
&= \frac{\prod_{i=1}^{n-1} f_{X_i,X_{i+1}}(x_i, x_{i+1})}{\prod_{i=2}^{n-1} f_{X_i}(x_i)}. && (16)
\end{aligned}
$$

Note that (13) is obtained by conditioning approach, (14) is obtained by using the Markov property, (15) is obtained by using the conditional density function, and the last step is obtained by simple algebraic calculation.

Now we will be discussing about the stopping time. It may be mentioned that the stopping time has been discussed quite extensively in the time series literature, see for example Christensen [7], Novikov and Shiryaev [15], and see the references cited therein. Let $L > 0$ be a fixed real number, and let us define a new random variable N, such for $k \geq 1$,

$$\{N = k\} \Leftrightarrow \{X_1 > L, \ldots, X_{k-1} > L, X_k \leq L\}.$$

Then clearly, N is a stopping time. Now for $\lambda = \max\{\lambda_0, \lambda_1\}$, first we obtain

$$
\begin{aligned}
P(X_1 > L, \ldots, X_k > L) &= P(U_0 < e^{-\lambda_1 L^\alpha}, U_1 < e^{-\lambda_0 L^\alpha}, \ldots U_{k-1} < e^{-\lambda_1 L^\alpha}, U_k < e^{-\lambda_0 L^\alpha}) \\
&= e^{-(\lambda_0+\lambda_1)L^\alpha} - e^{-(k-1)\lambda L^\alpha}
\end{aligned}
$$

$$
\begin{aligned}
P(N = 1) &= P(X_1 \leq L) = 1 - e^{-(\lambda_0+\lambda_1)L^\alpha} \\
P(N = 2) &= P(X_1 > L, X_2 \leq L) = P(X_1 > L) - P(X_1 > L, X_2 > L)
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-(\lambda_0+\lambda_1)L^\alpha} - P(U_1 < e^{-\lambda_0 L^\alpha}, U_0 < e^{-\lambda_1 L^\alpha}, U_2 < e^{-\lambda_0 L^\alpha}, U_1 < e^{-\lambda_1 L^\alpha}) \\
&= e^{-(\lambda_0+\lambda_1)L^\alpha}(1 - e^{-\lambda L^\alpha}) \\
P(N=3) &= P(X_1 > L, X_2 > L, X_3 \le L) = P(X_1 > L, X_2 > L) - P(X_1 > L, X_2 > L, X_3 > L) \\
&= e^{-(\lambda_0+\lambda_1)L^\alpha} e^{-\lambda L^\alpha}(1 - e^{-\lambda L^\alpha}) \\
P(N=k) &= e^{-(\lambda_0+\lambda_1)L^\alpha} e^{-(k-2)\lambda L^\alpha}(1 - e^{-\lambda L^\alpha}).
\end{aligned}
$$

The probability generating function of N can be obtained as

$$
\begin{aligned}
G(z) &= E(z^N) = \sum_{k=1}^{\infty} P(N=k) z^k \\
&= \frac{z(1 - e^{-(\lambda_0+\lambda_1)L^\alpha}) + z^2(e^{-(\lambda_0+\lambda_1)L^\alpha} - e^{-\lambda L^\alpha})}{(1 - z e^{-\lambda L^\alpha})}.
\end{aligned}
$$

Using the probability generating function, different moments and other properties can be easily derived.

3 Maximum Likelihood Estimators

In this section we consider the maximum likelihood estimators of the unknown parameters of a Weibull process based on a sample of size n, namely $x_1, \ldots, x_n$. We consider two cases separately.

Case 1: $\lambda_0 = \lambda_1$

In this case it is assumed that $\lambda_0 = \lambda_1 = \lambda$. Our problem is to estimate α and λ based on $\mathcal{D} = \{x_1, \ldots, x_n\}$. We use the following notations: $I = \{1, \ldots, n-1\}$

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

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 joint PDF (16), the log-likelihood function can be written as

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

$$= (n_1+n_2+1)\ln\alpha + (n_1+n_2+1)\ln\lambda + (\alpha-1)(\ln x_1 + \sum_{i\in I_1\cup I_2} \ln x_{i+1}) - \lambda g_1(\alpha|\mathcal{D}) \tag{17}$$

where

$$g_1(\alpha|\mathcal{D}) = \sum_{i\in I_1}(x_i^\alpha + 2x_{i+1}^\alpha) + \sum_{i\in I_2}(2x_i^\alpha + x_{i+1}^\alpha) + 3\sum_{i\in I_0} x_i^\alpha - 2\sum_{i=2}^{n-1} x_i^\alpha.$$

It is immediate that for any α, $g_1(\alpha|\mathcal{D}) > 0$. Hence, for a given α, the MLE of λ, say $\widehat{\lambda}(\alpha)$ can be obtained as

$$\widehat{\lambda}(\alpha) = \frac{n_1+n_2+1}{g_1(\alpha|\mathcal{D})}, \tag{18}$$

and the MLE of α, say $\widehat{\alpha}$ can be obtained by maximizing

$$\begin{aligned} h(\alpha) &= (n_1+n_2+1)\ln\alpha + (n_1+n_2+1)(\ln(n_1+n_2+1) - \ln g(\alpha|\mathcal{D})) \\ &\quad + \alpha(\ln x_1 + \sum_{i\in I_1\cup I_2} \ln x_{i+1}). \end{aligned} \tag{19}$$

Once, $\widehat{\alpha}$ is obtained, then the MLE of λ, say $\widehat{\lambda}$ can be obtained as $\widehat{\lambda}(\widehat{\alpha})$. Due to complicated nature of $h(\alpha)$ it is difficult to prove that it is an unimodal function. But in our data analysis it is observed that $h(\alpha)$ is an unimodal function, and it will be explained later. In case of exponential process, the MLE of λ can be obtained as

$$\widehat{\lambda} = \frac{n_1+n_2+1}{g_1(1|\mathcal{D})}. \tag{20}$$

Case 2: $\lambda_0 \neq \lambda_1$

In this section we consider the case when λ_0 and λ_1 are arbitrary. We use the following notations

$$\begin{aligned} I_1(\beta) &= \{i : i \in I, \beta x_i < x_{i+1}\} \\ I_2(\beta) &= \{i : i \in I, \beta x_i > x_{i+1}\} \\ I_0(\beta) &= \{i : i \in I, \beta x_i = x_{i+1}\}, \end{aligned}$$

and $n_0(\beta) = |I_0(\beta)|$, $n_1(\beta) = |I_1(\beta)|$ and $n_2(\beta) = |I_2(\beta)|$. Here, β is same as defined before. Based on the data vector $\mathcal{D}$, the log-likelihood function of λ_0, λ_1 and α becomes

$$\begin{aligned} l(\lambda_0, \lambda_1, \alpha|\mathcal{D}) \;=\; & (n_1(\beta) + n_2(\beta) + 1) \ln \alpha + n_1(\beta) \ln \lambda_1 + (n_2(\beta) + n_0(\beta)) \ln \lambda_0 + \\ & (n_1(\beta) + n_2(\beta) + 2 - n) \ln(\lambda_0 + \lambda_1) + (\alpha - 1) \left\{ \ln x_1 + \sum_{i \in I_1(\beta) \cup I_2(\beta)} \ln x_{i+1} \right\} - \\ & \lambda_1 \sum_{i \in I_1(\beta)} x_i^\alpha - (\lambda_0 + \lambda_1) \sum_{i \in I_2(\beta)} x_i^\alpha - \frac{\lambda_0^2 + \lambda_1^2 + \lambda_0 \lambda_1}{\lambda_1} \sum_{i \in I_0(\beta)} x_i^\alpha - \\ & (\lambda_0 + \lambda_1) \sum_{i \in I_1(\beta)} x_{i+1}^\alpha - \lambda_0 \sum_{i \in I_2(\beta)} x_{i+1}^\alpha - n_0(\beta) \ln \beta + (\lambda_0 + \lambda_1) \sum_{i=2}^{n-1} x_i^\alpha. \end{aligned} \qquad (21)$$

It is not trivial to maximize (21) directly. Hence, we use the following re-parameterization. We use the following transformed parameters, $(\gamma, \lambda_1, \alpha)$, where $\gamma = \frac{\lambda_0}{\lambda_1}$ and based on the transformed parameters, the log-likelihood function can be written as

$$\begin{aligned} l(\gamma, \lambda_1, \alpha) \;=\; & (n_1(\beta) + n_2(\beta) + 1)(\ln \lambda_1 + \ln \alpha) - \lambda_1 g_2(\alpha, \gamma|\mathcal{D}) - (n_0(\beta) - 1) \ln(1 + \gamma) + \\ & \left(\left(n_0(\beta) - \frac{1}{\alpha} \right) + n_2(\beta) \right) \ln(\gamma) + (\alpha - 1) h_2(\mathcal{D}), \end{aligned} \qquad (22)$$

where

$$\begin{aligned} g_2(\alpha, \gamma|\mathcal{D}) \;=\; & \sum_{i \in I_1(\beta)} x_i^\alpha + (1 + \gamma) \left(\sum_{i \in I_1(\beta)} x_{i+1}^\alpha + \sum_{i \in I_2(\beta)} x_i^\alpha \right) + \gamma \sum_{i \in I_2(\beta)} x_{i+1}^\alpha \\ & + (\gamma^2 + \gamma + 1) \sum_{i \in I_0(\beta)} x_i^\alpha - (1 + \gamma) \sum_{i=2}^{n-1} x_i^\alpha \\ h_2(\mathcal{D}) \;=\; & \ln x_1 + \sum_{i \in I_1(\beta) \cup I_2(\beta)} \ln x_{i+1}. \end{aligned}$$

We propose to use profile likelihood method to maximize (22). For fixed γ and α (β is also fixed in that case), first we maximize (22) with respect to λ_1, say $\widehat{\lambda}_1(\gamma, \alpha)$, and it can be obtained in explicit form as

$$\widehat{\lambda}_1(\gamma, \alpha) = \frac{n_1(\beta) + n_2(\beta) + 1}{g_2(\alpha, \gamma|\mathcal{D})}.$$

The MLEs of γ and α, say $\widehat{\gamma}$ and $\widehat{\alpha}$, respectively, can be obtained by maximizing $l(\gamma, \widehat{\lambda}_1(\gamma, \alpha), \alpha)$. Finally, the MLE of λ_1 can be obtained as $\widehat{\lambda}_1(\widehat{\gamma}, \widehat{\alpha})$. We will denote this as $\widehat{\lambda}_1$. Due to complicated nature of the function $g_2(\alpha, \gamma|\mathcal{D})$, it is difficult to prove that it has an unique maximum. But in our data analysis it is observed from the contour plot that $g_2(\alpha, \gamma|\mathcal{D})$ is an unimodal function. In both the cases we have suggested parametric bootstrap method to construct confidence intervals of the unknown parameters. They can be very easily implemented in practice. In case of exponential process, the MLE of λ_1 for a given γ can be obtained as

$$\widehat{\lambda}_1(\gamma) = \frac{n_1(\beta) + n_2(\beta) + 1}{g_2(1, \gamma|\mathcal{D})},$$

and the MLE of γ can be obtained by maximizing $l(\gamma, \widehat{\lambda}_1(\gamma), 1)$. It is an one dimensional optimization problem.

4 Goodness of Fit

In this section we provide a goodness of fit test, so that whether a given data set comes from a Weibull process or not can be tested. Suppose $\{x_1, \ldots, x_n\}$ is a sample from a stationary sequence $\{X_1, \ldots, X_n\}$. We want to test the following null hypothesis

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

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

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

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

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

where $c(\beta)$ is such that

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

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

Algorithm:

Step 1: Obtain $\widehat{\alpha}, \widehat{\lambda}_0, \widehat{\lambda}_1$, the MLEs of $\alpha, \lambda_0, \lambda_1$, respectively, based on $\{x_1, \ldots, x_n\}$.

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

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

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

Step 4: Compute

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

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

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

5 Data Analysis

In this section we have analyzed three data sets; two synthetic data sets and one real gold price data set of the Indian market for two months. The main idea about these data analysis

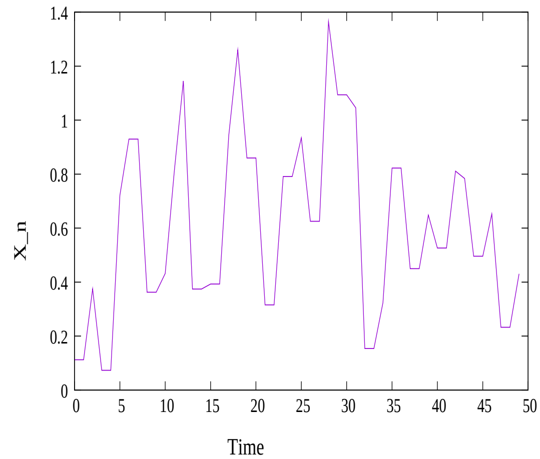

Figure 1: Synthetic data set with $\alpha = 2.0$, $\lambda_0 = \lambda_1 = 1$.

is to see how the proposed MLEs work in practice and also how the proposed model works in real life.

5.1 Synthetic Data Set 1:

In this section we analyze one synthetic data set, and it has been generated using the following model specification: $\alpha = 2.0$, $\lambda_0 = \lambda_1 = 1.0$, $n = 50$. The data set (Data Set 1) has been presented in Figure 1. In this case it is observed $n_0 = 17$, $n_1 = 18$ and $n_2 = 14$. We would like to compute the MLEs of the unknown parameters based on the assumption $\lambda_0 = \lambda_1 = \lambda$. It involves solving a one dimensional optimization problem. The profile log-likelihood function of α has been presented in Figure 2. It is an unimodal function. Based on the profile maximization we obtain the MLEs of α and λ as $\widehat{\alpha} = 1.912$ and $\widehat{\lambda} = 1.068$. The associate 95% confidence intervals are (1.714,2.106) and (0.878,1.245), respectively.

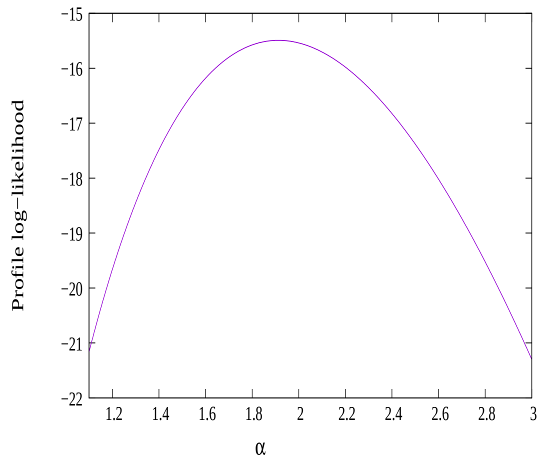

Figure 2: The profile log-likelihood of the Synthetic data set.

5.2 Synthetic Data Set 2:

In this section we analyze one synthetic data set, and it has been generated using the following model specification: $\alpha = 3.0$, $\lambda_0 = 0.15$, $\lambda_1 = 0.04$ and $n = 75$. The data set (Data Set 2) has been presented in Figure 3. Now we would like to compute the MLEs of the unknown parameters. We have adopted the two dimensional grid search method to compute the MLEs of the unknown parameters. The MLEs of α, λ_0 and λ_1 become $\widehat{\alpha} = 3.344$, $\widehat{\lambda}_0 = 0.154$ and $\widehat{\lambda}_1 = 0.029$. The associated 95% bootstrap confidence intervals become (2.876,3.954), (0.137,0.173) and (0.019,0.047), respectively.

5.3 Gold Price Data

In this section we present the analyses of the gold price data in India for two months period from December 01, 2020 to January 31, 2021. The data represents the price of one gram of gold in Indian rupees. It is presented in Figure 4. In this case $n = 62$, the minimum, maximum and median values are 4280, 4580 and 4355, respectively. There are 29, 22 and 10 cases, so that $\{x_i < x_{i+1}\}$, $\{x_i > x_{i+1}\}$ and $\{x_i = x_{i+1}\}$, respectively. We have performed

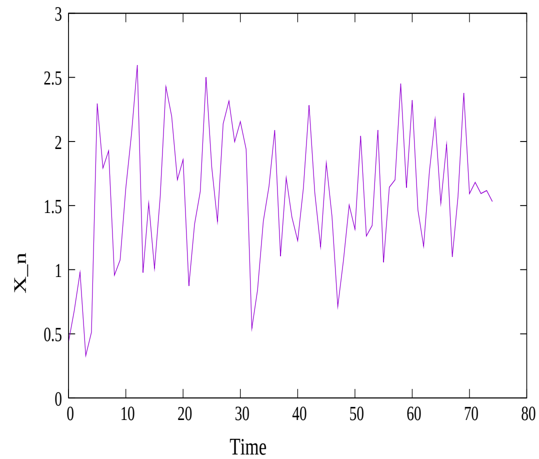

Figure 3: Synthetic data set with $\alpha = 3.0$, $\lambda_0 = 0.15$ and $\lambda_1 = 0.04$.

the run test on the entire data set $\{x_1, \ldots, x_{62}\}$, there are 28 runs, and the associated p value is less than 0.001. Hence, we reject the null hypothesis that they are independently distributed.

We have plotted the autocorrelation function (ACF) and the partial autocorrelation (PACF) of the gold price data in Figures 5 and 6, respectively. It is clear from the ACF and PACF that although x_i and x_{i+k} are correlated, they are uncorrelated given $x_{i+1}, \ldots, x_{i+k-1}$, for $k = 1, 2, \ldots$.

We have performed run tests on two lag-1 series. The number of runs are 16 and 17, respectively. The associated p values are 0.07 and 0.18, respectively. We have performed run tests on three lag-2 series also. The number of runs are 12, 12 and 11, respectively. The associated p values are 0.56, 0.56, 0.25, respectively. Hence, based on the p values we cannot reject the null hypothesis that lag-2 observations are independently distributed. We have fitted Weibull distribution to all the three lag-2 series, the Kolmogorov-Smirnov distances and the associated p values reported in brackets are 0.1453 (0.7920), 0.2163 (0.3066) and 0.2227 (0.2743). Based on the p values we cannot reject the null hypothesis that lag-2 observations are from i.i.d. Weibull distribution.

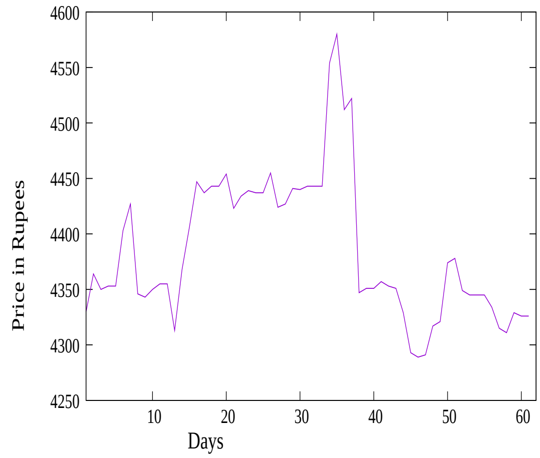

Figure 4: Gold price data in India (Rupees/ gram) from December 01, 2020 to January 31, 2021.

Now we would like to compute the MLEs of the unknown parameters under the assumption $\lambda_1 = \lambda_2 = \lambda$. It may be mentioned that in this case reasonable estimates of α and λ can be obtained in explicit forms without solving any optimization. In case of a Weibull distribution, the approximate MLEs of the unknown parameters can be obtained in explicit forms by expanding the log-likelihood function using the first order Taylor series expansion, see for example Kundu and Gupta [13]. Based on this approach we can obtain estimates of α and λ from the odd sequence as well as from the even sequence of the data. By taking the averages of these two estimates, we obtain estimates of α and λ as 2.9878 and 0.1154, respectively.

Now we would like to obtain the MLEs of α and λ by maximizing the log-likelihood function. The profile log-likelihood function of α has been plotted in Figure 7. By maximizing the profile log-likelihood function, we obtain the MLE of α as 2.6640, the MLE of λ as 0.0729 and the associated log-likelihood value becomes -116.9681. Based on parametric bootstrapping the associated 95% confidence intervals of α and λ are (2.1375,3.1231) and (0.0548,0.0976), respectively.

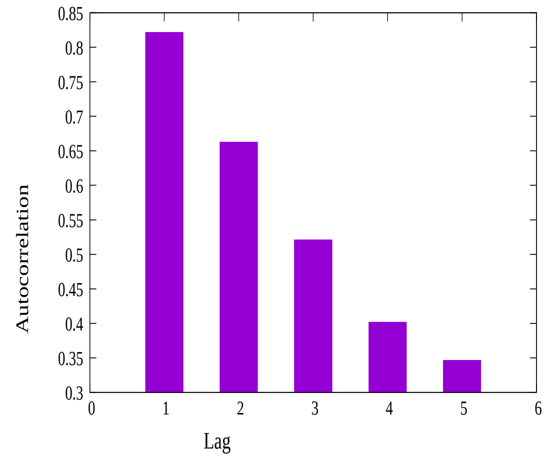

Figure 5: Autocorrelation function of the Gold price data.

Further, we have calculated the MLEs of the unknown parameters when $\lambda_0 \neq \lambda_1$. We have computed the MLEs of the unknown parameters by maximizing the profile log-likelihood of γ and α, i.e. $l(\gamma, \widehat{\lambda}_1(\gamma, \alpha), \alpha)$, with respect to γ and α. It is being performed by using grid search method, and the MLEs are as follows: $\widehat{\alpha} = 3.2238$, $\widehat{\lambda}_0 = 0.0917$,$\widehat{\lambda}_1 = 0.0192$ and the corresponding log-likelihood value becomes -115.9317. The associated 95% confidence intervals of α, λ_0 and λ_1 become (2.6213,3.8231), (0.0529,0.1412), (0.0123,0.0204), respectively.

It is clear that based on BIC model selection criterion we prefer the model WIP$(\alpha, \lambda, \lambda)$ than WIP$(\alpha, \lambda_0, \lambda_1)$. Now we would like to see whether both the models fit the data or not. We have used the bootstrap method proposed in Section 4 with $B = 1000$. The histogram of the generated $\{w^b : 1 \leq b \leq 1000\}$ when $\lambda_0 = \lambda_1 = \lambda$ and when $\lambda_0 \neq \lambda_1$ are provided in Figures 8 and 9, respectively. The test statistic for the model WIP$(\alpha, \lambda, \lambda)$ is 0.1685, and the associated p-value is greater 0.90. It seems it provides a good fit to the data set. The test statistic for the model WIP$(\alpha, \lambda_0, \lambda_1)$ is 0.6291, and the associated p value is less than 0.05. Hence, it does not provide a good fit to the data set.

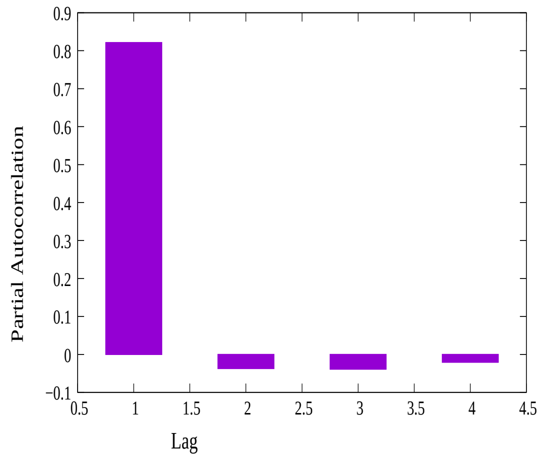

Figure 6: Partial Autocorrelation function of the Gold price data.

6 Conclusions

In this paper we have proposed a new discrete time and continuous state space stochastic process based on the Weibull distributions. The distinct feature of the proposed process is that there is a positive probability that $X_n = X_{n+1}$, for some n. Hence, this model can be used quite effectively when there are ties in the two consecutive time points. We have studied different properties of the proposed process, and also provided the inference procedures of the unknown parameters.

Note that the proposed stochastic process is a lag-1 process, but it can be easily extended to lag-q process as follows: Suppose $U_0, U_1, U_2, \ldots$ are independently and identically distributed (i.i.d.) uniform $U(0,1)$ random variables, and $\alpha > 0, \lambda_0 > 0, \lambda_1, \ldots, \lambda_q > 0$. Then

$$X_n = \min\left\{\left[-\frac{1}{\lambda_0}\ln U_n\right]^{\frac{1}{\alpha}}, \ldots, \left[-\frac{1}{\lambda_q}\ln U_{n-q}\right]^{\frac{1}{\alpha}}\right\}$$

is a lag-q stationary Weibull process. It can be easily checked that there is a positive probability that $X_n = X_{n+1} = \ldots = X_{n+k}$, for some n, and for $k = 1, \ldots, q$. It also has a convenient copula structure. It will be interesting to develop different properties and classical

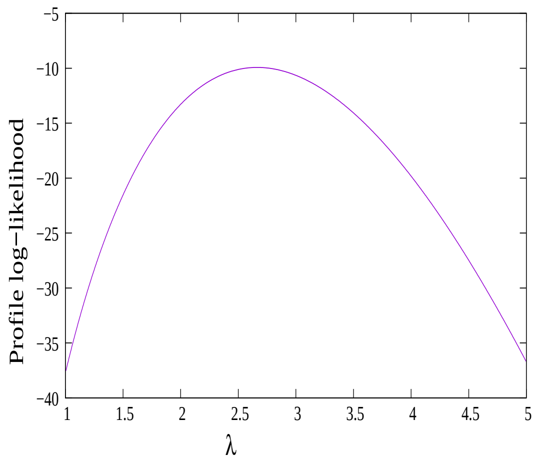

Figure 7: The profile log-likelihood function of α.

inferences of this process. More work is needed in this direction.

Acknowledgements:

The author would like to thank two unknown reviewers for their constructive suggestions, which have helped to improve the paper significantly.

Appendix A: Proofs

Proof of Theorem 4:

Note that p and $S_a(x,y)$ can be obtained from $S_{n,n+1}(x,y)$ as follows:

$$p = \int_0^\infty \int_0^\infty \frac{\partial^2 S_{n,n+1}(x,y)}{\partial x \partial y} dxdy,$$

and

$$pS_a(x,y) = \int_y^\infty \int_x^\infty \frac{\partial^2 S_{n,n+1}(u,v)}{\partial u \partial v} dudv.$$

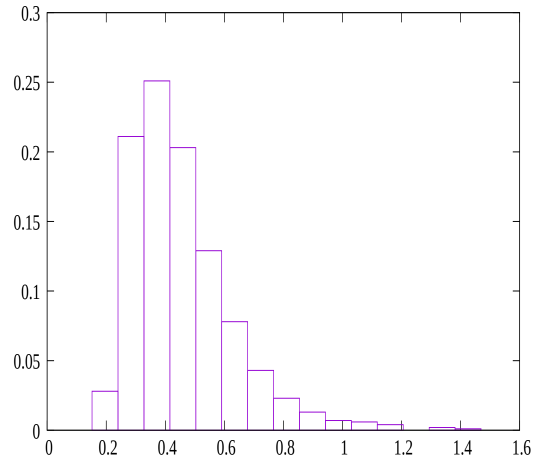

Figure 8: Histogram of the generated test statistics when $\lambda_0 = \lambda_1$

Now, from

$$\frac{\partial^2 S_{n,n+1}(x,y)}{\partial x \partial y} = \begin{cases} f_1(x,y) & \text{if} \quad (x,y) \in S_1 \\ f_2(x,y) & \text{if} \quad (x,y) \in S_2, \end{cases}$$

where

$$\begin{aligned} f_1(x,y) &= f_{WE}(x;\alpha,\lambda_1) f_{WE}(y;\alpha,\lambda_0+\lambda_1) \\ f_2(x,y) &= f_{WE}(x;\alpha,\lambda_0+\lambda_1) f_{WE}(y;\alpha,\lambda_0). \end{aligned}$$

Since

$$\int_0^\infty \int_{\beta x}^\infty f_1(x,y) dy dx = \frac{\lambda_1^2}{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1} \qquad \text{and} \qquad \int_0^\infty \int_{y/\beta}^\infty f_2(x,y) dx dy = \frac{\lambda_0^2}{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1},$$

$$p = \frac{\lambda_0^2+\lambda_1^2}{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}.$$

Using this p, $S_a(x,y)$ can be obtained by simple integration, and after that $S_s(x,y)$ can be obtained by subtraction.

Alternatively, a simple probabilistic arguments also can be given as follows. Suppose A is the following event

$$A = \left\{ \left[-\frac{1}{\lambda_0} \ln U_n \right] < \left[-\frac{1}{\lambda_1} \ln U_{n-1} \right] \right\} \cap \left\{ \left[-\frac{1}{\lambda_1} \ln U_n \right] < \left[-\frac{1}{\lambda_0} \ln U_{n+1} \right] \right\},$$

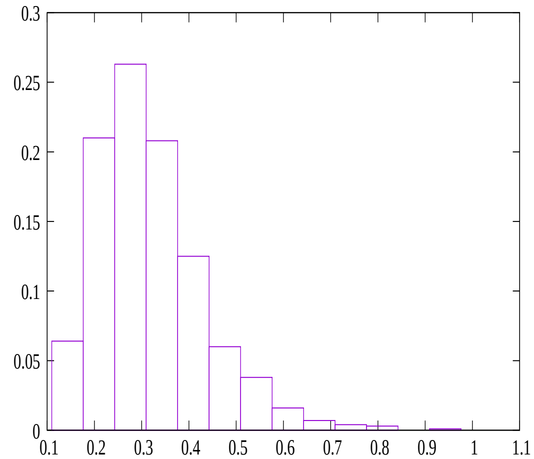

Figure 9: Histogram of the generated test statistics when $\lambda_0 \neq \lambda_1$

then

$$P(A) = P(U_n > U_{n-1}^{\frac{\lambda_0}{\lambda_1}}, U_n > U_{n+1}^{\frac{\lambda_1}{\lambda_0}}) = \int_0^1 u^{\frac{\lambda_0}{\lambda_1}+\frac{\lambda_1}{\lambda_0}} du = \frac{\lambda_0\lambda_1}{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1} = 1-p.$$

Moreover,

$$P(X_n > x, X_{n+1} > y) = P(X_n > x, X_{n+1} > y|A)P(A) + P(X_n > x, X_{n+1} > y|A')P(A'),$$

and

$$P(X_n > x, X_{n+1} > y|A) = P(U_n < e^{-\lambda_0 x^\alpha}, U_n < e^{-\lambda_1 y^\alpha}) = g(x,y).$$

The rest can be obtained by subtraction. ∎

Proof of Theorem 5: We need to show that that for all $0 < x, y < \infty$,

$$S_{n,n+1}(x,y) = \int\int_{B_1} f_1(u,v)dudv + \int\int_{B_2} f_2(u,v)dudv + \int_{h(x,y)}^{\infty} f_0(t)|\gamma'(t)|dt,$$

here for $R(x,y) = \{(u,v); x \leq u < \infty, y \leq v < \infty\}$, $B_1 = R(x,y) \cap S_1$, $B_2 = R(x,y) \cap S_2$, and $h(x,y) = \max\left\{x, \frac{y}{\beta}\right\}$. It has already been shown in Theorem 4 that

$$\int\int_{B_1} f_1(u,v)dudv + \int\int_{B_2} f_2(u,v)dudv = pS_a(x,y),$$

hence, the result is proved if we can show

$$\int_{h(x,y)}^{\infty} f_0(t)|\gamma'(t)|dt = (1-p)S_s(x,y).$$

Since, $|\gamma'(t)| = \beta$ and $(1-p) = \dfrac{\lambda_0\lambda_1}{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}$,

$$\int_{h(x,y)}^{\infty} f_0(t)|\gamma'(t)|dt = \int_{h(x,y)} \alpha\lambda_0 t^{\alpha-1} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1}t^\alpha} dt = (1-p)\int_{v(x,y)} e^{-u} du,$$

where $v(x,y) = \max\left\{\dfrac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1}x^\alpha, \dfrac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_0}y^\alpha\right\}$. Let us remember,

$$S_s(x,y) = \begin{cases} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_0}y^\alpha} & \text{if} \quad y > \beta x \\ e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1}x^\alpha} & \text{if} \quad y < \beta x. \end{cases}$$

Hence, the result follows. ∎

Appendix B: Autocovariance and Autocorrelation Functions

In this section we provide all the expressions of the autocorrelation function of the GE process mainly for completeness purposes. First we will calculate $E(X_{n+1}X_n)$. If I_A denote the indicator function on the set A, then

$$\begin{aligned} E(X_{n+1}X_n) &= E(X_{n+1}X_n \cdot I_{\{\beta X_n < X_{n+1}\}}) + E(X_{n+1}X_n \cdot I_{\{\beta X_n > X_{n+1}\}}) + \\ &\quad E(X_{n+1}X_n \cdot I_{\{\beta X_n = X_{n+1}\}}) \\ &= E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1} \cdot I_{\{\beta X_n < X_{n+1}\}}|X_n) + \\ &\quad E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1} \cdot I_{\{\beta X_n > X_{n+1}\}}|X_n) + \\ &\quad E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1} \cdot I_{\{\beta X_n = X_{n+1}\}}|X_n). \end{aligned}$$

Now

$$\begin{aligned} E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1} \cdot I_{\{\beta X_n = X_{n+1}\}}|X_n) &= \alpha\beta\lambda_0 \int_0^\infty x^{\alpha+1} e^{-\frac{\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1}{\lambda_1}x^\alpha} dx \\ &= \Gamma\left(\frac{2}{\alpha}+1\right) \frac{(\lambda_0\lambda_1)^{1+\frac{1}{\alpha}}}{(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^{\frac{2}{\alpha}+1}} \end{aligned}$$

If we denote $\Gamma(x,a) = \int_x^\infty t^{a-1}e^{-t}dt$ and $\gamma(x,a) = \int_0^x t^{a-1}e^{-t}dt$ as incomplete gamma functions, then

$$\begin{aligned} E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1}\cdot I_{\{\beta X_n<X_{n+1}\}}|X_n) &= \alpha^2\lambda_1 \int_0^\infty x^\alpha e^{-\lambda_1 x^\alpha}\left\{\int_{\beta x}^\infty y^\alpha e^{-(\lambda_0+\lambda_1)y^\alpha}dy\right\}dx \\ &= \frac{\alpha\lambda_1}{(\lambda_0+\lambda_1)^{1/\alpha}}\int_0^\infty x^\alpha e^{-\lambda_1 x^\alpha}\Gamma\left(\frac{\lambda_0(\lambda_0+\lambda_1)x^\alpha}{\lambda_1}, \frac{1}{\alpha}+1\right)dx. \end{aligned}$$

$$\begin{aligned} E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1}\cdot I_{\{\beta X_n>X_{n+1}\}}|X_n) &= \alpha^2\lambda_0(\lambda_0+\lambda_1) \int_0^\infty x^\alpha e^{-(\lambda_0+\lambda_1)x^\alpha}\left\{\int_0^{\beta x} y^\alpha e^{-\lambda_0 y^\alpha}dy\right\}dx \\ &= \frac{\alpha(\lambda_0+\lambda_1)}{\lambda_0^{1/\alpha}}\int_0^\infty x^\alpha e^{-(\lambda_0+\lambda_1)x^\alpha}\gamma\left(\frac{\lambda_0^2 x^\alpha}{\lambda_1}, \frac{1}{\alpha}+1\right)dx. \end{aligned}$$

We have already indicated the mean and variance of a Weibull random variable in (2). Now based on the above expressions, the autocovariance and autocorrelation functions can be obtained. In case of exponential process i.e. when $\alpha = 1$, the above expressions can be obtained in explicit forms. For example

$$E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1}\cdot I_{\{\beta X_n=X_{n+1}\}}|X_n) = \frac{2(\lambda_0\lambda_1)^2}{(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^3}$$

$$\begin{aligned} & E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1}\cdot I_{\{\beta X_n<X_{n+1}\}}|X_n) = \\ & \quad \frac{\lambda_1}{\lambda_0+\lambda_1}\int_0^\infty x e^{-\lambda_1 x}\left(e^{-\frac{\lambda_0(\lambda_0+\lambda_1)}{\lambda_1}x} + \frac{\lambda_0(\lambda_0+\lambda_1)}{\lambda_1}x e^{-\frac{\lambda_0(\lambda_0+\lambda_1)}{\lambda_1}x}\right)dx = \\ & \quad \frac{\lambda_1^3}{(\lambda_0+\lambda_1)(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^2} + \frac{2\lambda_0\lambda_1^3}{(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^3} \\ & E_{X_n}(X_n E_{\{X_{n+1}|X_n\}}(X_{n+1}\cdot I_{\{\beta X_n>X_{n+1}\}}|X_n) = \\ & \quad \frac{\lambda_0+\lambda_1}{\lambda_0}\int_0^\infty x e^{-(\lambda_0+\lambda_1)x}\left(1 - e^{-\frac{\lambda_0^2}{\lambda_1}x} - \frac{\lambda_0^2}{\lambda_1}x e^{-\frac{\lambda_0^2}{\lambda_1}x}\right)dx = \\ & \quad \frac{1}{\lambda_0(\lambda_0+\lambda_1)} - \frac{\lambda_1^2(\lambda_0+\lambda_1)}{\lambda_0(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^2} - \frac{2\lambda_0\lambda_1^2(\lambda_0+\lambda_1)}{(\lambda_0^2+\lambda_1^2+\lambda_0\lambda_1)^3}. \end{aligned}$$

Since,

$$E(X_n) = E(X_{n+1}) = \frac{1}{\lambda_0+\lambda} \quad \text{and} \quad V(X_n) = V(X_{n+1}) = \frac{1}{(\lambda_0+\lambda)^2},$$

the autocovariance and autocorrelation can be obtained in explicit forms.

References

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

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

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

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

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

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

[9] 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.

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

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

[12] 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.

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

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

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

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

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

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

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