IMA Journal of Management Mathematics (2017) **00**, 1–32
doi: 10.1093/imaman/dpx002

The semi-geometric process and some properties ®

SHAOMIN WU†
Kent Business School, University of Kent, Canterbury, Kent CT2 7FS, UK
†Corresponding author. Email: s.m.wu@kent.ac.uk

AND

GUANJUN WANG
Department of Mathematics, Southeast University, Nanjing, China

[Received on 11 January 2017; accepted on 28 March 2017]

The geometric process has been widely applied in reliability engineering and other areas since its introduction. One of its assumptions is that the times between occurrences of events are independent. This assumption is rather restrictive and can limit its application in the real world. This paper extends the geometric process to a new process, which we call the semi-geometric process, by relaxing this assumption. Some probabilistic properties of the process are derived and parameter estimation is described. A numerical example, based on a real-world data set, is used to illustrate the model and validate the estimation methodology.

Keywords: geometric process; copula; stochastic process; maximum likelihood; maintenance.

1. Introduction

The geometric process (GP) defines a stochastic process: a sequence of independent random variables $\{X_k, k = 1, 2, \dots\}$ is a GP if the cumulative distribution function (cdf) of X_k is given by $F(a^{k-1}x)$ for $k = 1, 2, \dots$ and a is a positive constant. The above definition is given by Lam (1988), though it is likely that this definition was around earlier. For example, in Smith & Leadbetter (1963), it reads '*we consider the situation in which failing components are replaced by new ones with better statistical properties. Specifically, it is assumed that the nth replacement has a lifetime distribution* $F(a^k x)$' and also gives the GP-version renewal function. Nevertheless, most publications typically credit the geometric process to Lam (1988).

The GP has attracted extensive research attention, which includes one monograph (Lam, 2007a) and many research papers, e.g. its probabilistic and statistical properties (Aydoğdu & Altındağ, 2016), and its application in reliability engineering (Jain & Gupta, 2013), warranty cost analysis (Chukova *et al.*, 2005). Meanwhile, some authors either proposed similar definitions to that of the GP (Finkelstein, 1993; Wang & Pham, 1996) or made an attempt to extend the GP (Braun *et al.*, 2005; Wu & Clements-Croome, 2006; Lam, 2007a; Wan & Chan, 2011; Bordes & Mercier, 2013; Wu, 2017).

While the GP is a model that has been widely used to solve problems in various research areas, its scope is still limited and does not fit the purposes of various empirical studies since $\{X_k, k = 1, 2, \dots\}$ in the GP are assumed independent. For example, as mentioned above, the GP has been used to model time-between-failures of multi-component systems, including multi-component systems as shown in

® This paper has been recognized by the editors as meeting the standard of this journal for reproducibility.

Lam (2007a, Section 5.4). However, it may not be realistic to assume that the times-between-failures of a system before repair is statistically independent of that after the repair. This paper aims to extend the GP to a new process, in which times-between-failures are dependent, and then to study the probabilistic and statistical properties of the process.

1.1. *Comments on the geometric process and its extensions*

Numerical examples show that the GP has its merit: in Lam's monograph on the GP (Lam, 2007a), 14 real-world time-between-event data sets are used to validate model performance [in terms of the mean squared error (MSE) or the maximum percentage error]. The results shows that the GP outperforms three other models [i.e. homogeneous Poisson process, non-homogeneous Poisson process (NHPP), following a power law and NHPP following an exponential law] on all data sets. However, the independence assumption of the GP is too restrictive: the GP assumes that $\{X_k, k = 1, 2, ...\}$ is a sequence of independent random variables, but it is true that times between occurrences of events may be statistically dependent in many scenarios in the real world.

To overcome the above limitation, one may relax the assumption that X_{k-1} and X_k are independent. This paper makes such an attempt.

1.2. *Contribution and importance of this work*

This paper proposes a new stochastic process, referred to as the *semi-geometric process* (SGP). One may notice that, in recent years, many authors have devoted considerable effort on developing novel methods to model repair processes, (see, e.g. Jain & Gupta, 2013; Liang & Parlikad, 2015; Sheu *et al.*, 2015; Vu *et al.*, 2015; Wu & Scarf, 2015, 2017; Zhao *et al.*, 2015; Wu, 2017). We consider the current paper and the process we describe as a further important contribution to the literature on the modelling the failure process of a repairable system.

The paper has important managerial implications, because it provides a more flexible model for wider application than the GP. Although this paper uses a case in reliability engineering, the model can also be applied to analyse other recurrent events. Such applications can be found in scientific studies, medical research, marketing research, etc., just as the GP can be used to model recurrent events such as the outbreaks of diseases (Chan *et al.*, 2006) and the electricity price (Chan *et al.*, 2014).

1.3. *Overview*

The rest of the paper is structured as follows. Section 2 introduces the SGP and discusses its probabilistic properties. Section 3 proposes methods of parameter estimation. Section 4 compares the performance [in terms of the Akaike information criterion (AIC)] of the SGP with that of other models based on a data set collected from the real world, respectively. We finish with a conclusion and future work in Section 5.

2. The semi-geometric process and probabilistic properties

2.1. *Definitions*

This section introduces the GP and discusses its limitations in detail. We begin with an important definition on stochastic order.

DEFINITION 2.1 Stochastic order (Ross, 1996, p. 404). Assume that X and Y are two random variables. If for every real number t, the inequality

$$P(X \geq t) \geq P(Y \geq t)$$

holds for all t, then X is stochastically greater than or equal to Y, or $X \geq_{st} Y$. Equivalently, Y is stochastically less than or equal to X, or $Y \leq_{st} X$.

LEMMA 2.1 (Ross, 1996, p. 405) Assume that X and Y are two random variables and their expectations exist, then

$$X \geq_{st} Y \iff \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)],$$

for all increasing functions $u(.)$, where $\iff$ stands for 'if only if'.

The definition of the GP process is introduced in Lam (1988), as shown below.

DEFINITION 2.2 Lam (1988) Given a sequence of non-negative random variables $\{X_k, k = 1, 2, \dots\}$, if they are independent and the cdf of X_k is given by $F(a^{k-1}x)$ for $k = 1, 2, \dots$, where a is a positive constant, then $\{X_k, k = 1, 2, \cdots\}$ is called a GP.

We refer to the random variable X_k as the kth inter-arrival time in what follows.

REMARK 2.1 From Lemma 2.1 and Definition 2.2, we have the following results.

- If $a > 1$, then $\{X_k, k = 1, 2, \cdots\}$ is stochastically decreasing.
- If $a < 1$, then $\{X_k, k = 1, 2, \cdots\}$ is stochastically increasing.
- If $a = 1$, then $\{X_k, k = 1, 2, \cdots\}$ is a renewal process (RP).

X_k and X_{k-1} are assumed to be independent in Definition 2.2. A natural idea to relax the independence assumption is to assume that $P(X_k < x|X_{k-1} = y) \neq P(X_k < x)$. This inspires us to give the following definition.

DEFINITION 2.3 Given a sequence of non-negative random variables $\{X_k, k = 1, 2, \dots\}$, if $P\{X_k < x|X_{k-1} = x_{k-1}, \dots, X_1 = x_1\} = P\{X_k < x|X_{k-1} = x_{k-1}\}$ and the marginal distribution of X_k is given by $P\{X_k < x\} = F_k(x)(\equiv F(a^{k-1}x))$, where a is a positive constant, then $\{X_k, k = 1, 2, \dots\}$ is called an SGP.

In other words, the SGP is a Markovian process having margins distributed as the GP.

Definition 2.3 requires us to find the correlation between X_{k-1} and X_k. An intuitive idea to model the correlation is to use the concept of copulas. Copulas are a tool for constructing multivariate distributions and formalizing the dependence structures between random variables. The notion of copula was first introduced by Abe Sklar in 1959 (Sklar, 1959). It has attracted considerable attention in recent years

in both theoretical and application aspects. Sklar's theorem states that any cumulative distribution function of a random vector can be written in terms of marginal distribution functions and a copula that describes the dependence structure between the variables (Sklar, 1959). That is, given a vector of random variables $(X_1, ..., X_d)$, its cumulative distribution function $H(x_1, ..., x_d)(= P(X_1 \leq x_1, ..., X_d \leq x_d))$ and marginals $F_k(x_k)(= P(X_k \leq x_k)$, where $k = 1, ...d)$, Sklar proved that $H(x_1, ..., x_d)$ can be written as $H(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$ and named $C(.)$ as a *copula* (Sklar, 1959). Copulas are useful in statistical applications, because they allow one to easily model and estimate the distribution of a random vector through estimating the marginals and the copula separately. Copulas have been found applications in a variety of areas, e.g. in reliability engineering problems; Guo *et al.* (2013), Peng *et al.* (2016) and Eryilmaz (2016) use copulas to solve reliability problems.

2.2. *Probabilistic properties*

This section investigates some properties such as the distribution of $X_1 + \cdots + X_n$, which are normally discussed for similar stochastic processes such as the RP and the GP.

Denote the copula between X_{k-1} and X_k as $C_{k-1,k}(u, v; \boldsymbol{\theta})$ and its corresponding density as $c_{k-1,k}(u, v; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is an estimable parameter vector.

LEMMA 2.2 The probability $P(X_k < x|X_{k-1} = y)$ is given by

$$P(X_k < x|X_{k-1} = y) = \frac{\partial C_{k-1,k}(u, v; \boldsymbol{\theta})}{\partial u}|_{(u=F_{k-1}(y), v=F_k(x))},$$

and the hazard function is given by

$$\lambda_{k|k-1}(x|X_{k-1} = y) = \frac{c_{k-1,k}(F_{k-1}(y), F_k(x); \boldsymbol{\theta}) f_k(x)}{1 - \frac{\partial C_{k-1,k}(u,v;\boldsymbol{\theta})}{\partial u}|_{(u=F_{k-1}(y)), v=F_k(x)}}.$$

The proof of each lemma in this paper is given in the Appendix.

Different notions of positive bivariate dependence can be defined, which are invariant by increasing transformations of X_k and X_{k-1}. Barlow & Proschan (1975) give the following Lemma.

LEMMA 2.3 (Barlow & Proschan, 1975, p. 143) X_k is stochastically increasing (SI) in X_{k-1} iff

$$S(x|y) \equiv P[X_k \geq x|X_{k-1} = y] \tag{2.1}$$

is increasing in y, for any $x \in \Re^+$.

Denote $f_{k-1}(y)$ as the density function of X_{k-1}. Based on Lemma 2.3, one can derive the following properties.

LEMMA 2.4 If $f_{k-1}(y) \frac{\partial^2 C_{k-1,k}(u,v;\boldsymbol{\theta})}{\partial u^2}|_{(u=F_{k-1}(y), v=F_k(x))} < 0$, X_k is stochastically increasing in X_{k-1}.

Obtaining the joint distribution of $(X_1, X_2, ..., X_n)$ is important in some applications such as calculating the cumulative failure function. In the two properties below, the bounds of the joint distribution are given.

LEMMA 2.5 The following bounds are given.

- if $a < 1$ then

$$\max\{1 - n + nF_n(x_n), 0\} \leq C_{1,2,\ldots,n}(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \leq F_n(x_n). \tag{2.2}$$

- if $a > 1$ then

$$\max\{1 - n + nF_1(x_1), 0\} \leq C_{1,2,\ldots,n}(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \leq F_1(x_1). \tag{2.3}$$

DEFINITION 2.4 (Joe, 1997, p. 20) Let $Z = (Z_1, Z_2)$ be a bivariate random vector with cdf H_2.

- H_2 is positive quadrant dependent (PQD) if

$$P(Z_1 < z_1, Z_2 < z_2) \geq P(Z_1 < z_1)P(Z_2 < z_2), \tag{2.4}$$

  for $\forall z_1, z_2 \in \Re$.
- H_2 is negative quadrant dependent (NQD) if

$$P(Z_1 < z_1, Z_2 < z_2) \leq P(Z_1 < z_1)P(Z_2 < z_2), \tag{2.5}$$

  for $\forall z_1, z_2 \in \Re$.

LEMMA 2.6 With the definition of the quadrant dependence, one can obtain the following.

- If $P\{X_k < t, X_{k-1} < s\}$ is PQD, then

$$\prod_{k=1}^{n} P(X_k < x_k) \leq P(X_1 < x_1, ..., X_n < x_n) \leq P(X_1 < x_1). \tag{2.6}$$

- If $P\{X_k < t, X_{k-1} < s\}$ is NQD, then

$$\prod_{k=1}^{n} P(X_k < x_k) \geq P(X_1 < x_1, ..., X_n < x_n). \tag{2.7}$$

An interesting question is the expected number of occurrences of the SGP within time interval $[0, t]$. To answer this question, one needs to obtain the probability distribution of the sum $X_1 + X_2 + \cdots + X_n$. Unfortunately, even for two dependent random variables X_1 and X_2, there is no closed form of the probability distribution of the sum $X_1 + X_2$. It is clear that solving this problem is mainly a numerical issue once the joint distribution of $X = (X_1, X_2, \ldots, X_n)$ is completely specified (Bernard & Vanduffel, 2015).

Below we give the bounds of the expected number of occurrences.

LEMMA 2.7 (Cherubini *et al.*, 2011, p. 145) If one denotes with $\mathbf{1}$ the n-dimensional column vector whose components are all equal to 1, and $\boldsymbol{t}$ the n-dimensional column vector, and defines as $T(s)$ the set of vectors such that the sum of their components is equal to 1:

$$T(s) = \left\{ \boldsymbol{t} \in \Re^n : \boldsymbol{t}^{\mathrm{T}} \mathbf{1} = s \right\}, \tag{2.8}$$

then the stochastic upper and lower bounds on $F_S(s)$ are $F_L(s) = \sup_{\boldsymbol{t} \in T(s)} \max \left\{ \sum_{k=1}^{n} F_k(t_k) - n + 1, 0 \right\}$, $F_U(s) = \inf_{\boldsymbol{t} \in T(s)} \min \left\{ \sum_{k=1}^{n} F_k(t_k), 1 \right\}$, respectively.

Denote $H_n(s) = P\{X_1 + X_2 + \cdots + X_n \leq s\}$. Then, it is easy to derive the following lemma.

LEMMA 2.8 The following two bullets hold.

- If $a > 1$, then

$$\inf_{\boldsymbol{t} \in T(s)} \min \left\{ \sum_{k=1}^{n} F_1(t_k) - n + 1, 0 \right\} \leq H_n(s) \leq \sup_{\boldsymbol{t} \in T(s)} \max \left\{ \sum_{k=1}^{n} F_n(t_k) - n + 1, 0 \right\};$$

- If $a < 1$, then

$$\inf_{\boldsymbol{t} \in T(s)} \min \left\{ \sum_{k=1}^{n} F_n(t_k) - n + 1, 0 \right\} \leq H_n(s) \leq \sup_{\boldsymbol{t} \in T(s)} \max \left\{ \sum_{k=1}^{n} F_1(t_k) - n + 1, 0 \right\}.$$

3. Parameter estimation

This section discusses two approaches to estimate the parameters of the SGP: the least squares (LS) method and the maximum likelihood (ML) method. In the LS method, no assumption of the joint distribution of (X_{k-1}, X_k) is made. The ML method, however, requires a known joint distribution or copula of (X_{k-1}, X_k).

3.1. *LS method*

Lam (2007b) uses the following linear regression method to estimate the parameter of the GP.

For the GP $\{X_k, k = 1, 2, \dots\}$ with parameter a, let

$$Y_k = a^{k-1} X_k. \tag{3.1}$$

Then $\{Y_k, k = 1, 2, \dots\}$ is a sequence of i.i.d random variables, so is $\{\ln(Y_k), k = 1, 2, \dots\}$. Taking logarithm on the both sides of equation (3.1) gives

$$\ln Y_k = (k-1) \ln a + \ln X_k. \tag{3.2}$$

On the other hand, for a given n observations $\{x_k, k = 1, 2, \ldots n\}$ of $\{X_k, k = 1, 2, \ldots n\}$, let's set $y_k = a^{k-1}x_k$. Then, one can rewrite

$$\ln y_k = \mu + e_k, \tag{3.3}$$

where μ is the mean of $\ln y_k$, e's are i.i.d. random variables with mean 0 and variance τ^2 (see Lam, 2007b, p. 105). Denote $z_k \equiv \ln(x_k)$ and $\beta \equiv \ln(a)$. Then, equation (3.2) becomes

$$z_k = \mu + (1 - k)\beta + e_k. \tag{3.4}$$

Equation (3.4) is a linear regression equation, and its parameters a and μ can be easily estimated with observations x_k. For more details on the above derivation, the reader is referred to Lam (2007a).

The above linear regression method, however, may not work in our setting because X_k is assumed to depend on X_{k-1}. If one uses equation (3.4), then the residual e_k may be correlated with e_{k-1}.

If e_k is correlated with e_{k-1}, one can assume $e_k = \rho e_{k-1} + \epsilon_k$. Then, ϵ_k is a white noise that follows a normal distribution with mean 0. Substituting $e_k = z_k - \mu - (1 - k)\beta$ and $e_{k-1} = z_{k-1} - \mu - (2 - k)\beta$ into $e_k = \rho e_{k-1} + \epsilon_k$, we obtain

$$z_k - \rho z_{k-1} = \mu' + \beta(1 - k - 2\rho + \rho k) + \epsilon_k, \tag{3.5}$$

where $\mu' = \mu(1 - \rho)$.

REMARK 3.1 Similar to Lemma 2.4 that associates X_k and X_{k-1}, equation (3.5) suggests that z_k and z_{k-1} have a positive correlation if $\rho > 0$ and a negative correlation if $\rho < 0$. Equivalently, X_k and X_{k-1} have a positive correlation if $\rho > 0$, and a negative correlation if $\rho < 0$.

For the given observations x_k of X_k (with $k = 1, 2, ..., n$), one can minimize the following sum of the squares of the errors to estimate the parameters a, b and μ.

$$(\hat{\mu}', \hat{\rho}, \hat{\beta}) = \arg\min_{\mu', \rho, \beta} \sum_{k=2}^{n} \left(z_k - \rho z_{k-1} - \mu' - \beta(1 - k - 2\rho + \rho k)\right)^2. \tag{3.6}$$

Obviously, there is no general closed-form solution for $\hat{\rho}$, $\hat{\mu}'$ and $\hat{\beta}$, one therefore needs to pursue non-linear programming methods to solve the problem. One may pursue a software package such as Matlab® to estimate the parameters.

Alternatively, one may use the Cochrane–Orcutt method (Cochrane & Orcutt, 1949) to solve this non-linear programming problem and estimate the parameters, as shown below.

(a) Obtain μ and β by solving the regression equation (3.4) with LS.

(b) Calculate $e_k = z_k - \mu - (1 - k)\beta$ and regress e_k on e_{k-1} to obtain an estimate of ρ.

(c) Calculate $z_k - \rho z_{k-1}$ and $(1 - k - 2\rho + \rho k)$ and solve equation (3.5) with LS to obtain revised estimates μ and β. Return to (b) and continue until convergence.

3.2. *ML method*

Suppose that the system has failed for n times at time points s_k with $k = 0, 1, \ldots, n$. Let $s_0 = 0$. Then the working times are $s_1 - s_0, s_2 - s_1, \ldots, s_n - s_{n-1}$, respectively. Denote $x_k = s_k - s_{k-1}$ for $k = 1, 2, ..., n$.

Since

$$\begin{aligned} f(x_1, x_2, ..., x_n) &= f_1(x_1) f_{2|1}(x_2|x_1) ... f_{n|n-1}(x_n|x_{n-1}) \\ &= c_{12}(F_1(x_1), F_2(x_2)) ... c_{n-1,n}(F_{n-1}(x_{n-1}), F_n(x_n); \boldsymbol{\theta}) \prod_{k=1}^{n} f_k(x_k) \\ &= \left(\prod_{k=2}^{n} c_{k-1,k}(F_{k-1}(x_{k-1}), F_k(x_k); \boldsymbol{\theta}) \right) \left(\prod_{k=1}^{n} f_k(x_k) \right), \end{aligned} \tag{3.7}$$

and the logarithm likelihood function is given by

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log(f_k(x_k)) + \sum_{k=2}^{n} \log(c_{k-1,k}(F_{k-1}(x_{k-1}), F_k(x_k)); \boldsymbol{\theta}). \tag{3.8}$$

Maximizing the above log-likelihood function by setting the partial derivation of $\ell(\boldsymbol{\theta})$ with respect to each element of $\boldsymbol{\theta}$ to zero, we can obtain $\hat{\boldsymbol{\theta}}$, which are the estimates of the corresponding parameters, respectively. That is

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}). \tag{3.9}$$

3.3. *Confidence intervals*

In this section, estimating confidence intervals is discussed to account for the uncertainty in the estimators of the LS method and the ML method.

3.3.1. *Large sample case* When a large sample of observations are available for estimating the parameters, the confidence intervals for both the LS method and the ML method can be easily obtained. Here, by *large*, it may mean a sample size larger than 50 (e.g. McKnight *et al.*, 2000).

In the case of the LS method, as discussed above, equation (3.6) may be treated as a non-linear programming problem. The confidence intervals of the parameters can then be estimated (see Seber & Wild, 2003, for details).

In the case of the ML method, one may use asymptotic normal approximation to the ML estimates to construct confidence intervals. To use this method, one obtains the ML estimates of $\hat{\boldsymbol{\theta}}$ from equation (3.8). Then the ML estimator $\hat{\boldsymbol{\theta}}$ has a distribution that can be approximated by a multivariate normal distribution $N(\boldsymbol{\theta}, I^{-1}(\boldsymbol{\theta}))$, where $I^{-1}(\boldsymbol{\theta})$ is the Fisher information matrix that can be estimated based on the observed information matrix: $I(\hat{\boldsymbol{\theta}}) = -\mathbb{E}\left(\frac{\partial^2 \log \ell(\boldsymbol{\theta})}{\partial \phi_i \partial \phi_j} \right) |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, which can be used to estimate the asymptotic variance–covariance matrix of $\hat{\boldsymbol{\theta}}$.

3.3.2. *Bootstrap methods* When the sample size is small, the statistical properties of the estimates based on samples of large size may not hold. For example, the coverage probabilities of confidence intervals produced by the normal approximation method may not be close to nominal values. In such cases, variance estimation or confidence intervals may be conveniently approached with the bootstrap (Cook & Lawless, 2007).

In the case of the LS method, to estimate the confidence intervals, one may use the method proposed by McKnight *et al.* (2000), which introduces a double bootstrap method to overcome the limitation of the small sample size ($n < 50$).

In the case of the ML method, one may use the method proposed by Newton & Raftery (1994), which introduces a weighted likelihood bootstrap and use a random weight to the log-likelihood of each observation.

3.4. *Goodness-of-fit test and hypothesis test*

3.4.1. *Detection of the trend* The SGP presents two properties that need testing for real-world data. The two properties are (1) the dependence between X_{k-1} and X_k and (2) the stochastic trend, i.e. either increasing or decreasing, of X_k for $k = 1, 2, \ldots$. The Ljung–Box test (Ljung & Box, 1978) can be carried out to test the dependence between X_{k-1} and X_k and the Mann–Kendall trend test (Mann, 1945; Kendall, 1975) can be conducted to test whether there is a monotonic upwards or downwards trend in $X_k, k = 1, 2, ...\}$.

3.4.2. *Selection of the baseline lifetime distribution and the copula* The AIC, which aims to find the trade-off between the goodness of fit of a model and the complexity of the model, can be used for selection of the baseline lifetime distribution $F(x)$.

The use of the copula approach to depict the dependence between X_{k-1} and X_k has two advantages: one has the freedom to choose the marginal distribution (such as the fat-tailedness of the time series $\{X_k, k = 1, 2, ...\}$), and the copula function separately, and one can also characterize the temporal dependence property such as non-linear, asymmetric dependence, of the time series.

It should be noted that the joint distribution between X_{k-1} and X_k can then be modelled by any $C_{k-1,k}(u, v)$, which include parametric and non-parametric forms. One may use the AIC for selection of the different forms of copulas.

3.5. *Discussion*

3.5.1. *A further extension* Wu (2017) proposes an extension of the GP, which assumes that the distribution of X_k is $F(a^{k-1}x^{h(k)})$ for $k = 1, 2, \ldots$, where a is a positive constant, $h(k)$ is a function of k and the likelihood of the parameters in $h(k)$ has a known closed form, and $h(k) > 0$ for $k \in \mathbb{N}$, and $\{X_k, k = 1, 2, \cdots\}$ is called a doubly geometric process (DGP). It can be seen that the GP is a special case of the DGP.

The SGP merely links the relationship between X_k and its one-step backward counterpart X_{k-1}. A natural extension is to allow X_k to associate with its p historical records $X_{k-p}, X_{k-p+1}, \ldots, X_{k-1}$, which may be applied to both the GP and the DGP. However, it should be noted that on such an extension, the model (i.e. the SGP) becomes more complicated and may cause some problems in its applications, since the sample of time-between-failures is normally not large.

3.5.2. *Application* Consider the following maintenance policy for a system: the system is replaced if the probability of the event that both working times of the two consecutive periods are shorter than a value,

γ, say, is larger than α. In this maintenance policy, the value α is the decision variable. That is, given the cost of failure and cost of replacement, what is the optimal value of α? The real-world application of such a maintenance policy can be found in statistical process control where action may be taken if the working times of two consecutive periods are shorter than a control limit.

Since $P(X_{k-1} < \gamma, X_k < \gamma) = C_{k-1,k}(F_{k-1}(\gamma), F_k(\gamma))$, one can easily obtain value k when the condition $C_{k-1,k}(F_{k-1}(\gamma), F_k(\gamma)) > \alpha$ is satisfied. Often the parameter a is assumed to be larger than 1, as the working times are supposed to be stochastically decreasing. In such a case, for a given γ and large ks, $P(X_{k-1} < \gamma, X_k < \gamma) = 1$. As such, one selects the minimal k that satisfies $C_{k-1,k}(F_{k-1}(\gamma), F_k(\gamma)) > \alpha$.

4. Numerical example

&&&In this section, we fit the SGP, and three other models, the GP, non-homogeneous Poisson process with power law intensity function (NHPP-PL) and generalized renewal process (GRP), to a real-world data set and compare their AIC values with those models on the data set. It is known that the NHPP-PL is often used for minimal repair, and two models for imperfect repair are the GP and the GRP, with $V_k = V_{k-1} + qX_k$ being the virtual age after the kth failure (GRP) being used for imperfect repair.

We then calculate the AIC value from the ML estimation of the model: AIC $= 2p - 2\ln(L)$, where L is the maximized value of the likelihood of each model and p is the number of parameters in a model. In the LS method, the AIC value is calculated with: AIC $= 2p + n\ln(\frac{SME}{n})$, where SME is the squared mean error of the model. For comparison, we calculate SME=$\sum_{k=2}^{n}\left(z_k - \rho z_{k-1} - \mu' \ -\beta(1-k-2\rho+\rho k)\right)^2$ for the SGP and SME=$\sum_{k=2}^{n}(z_k - \mu - (1-k)\beta)^2$ (see equation (3.4)) for the GP. Note that in both SME's k starts from 2. The term $2p$ in the AIC penalizes a model with a large number of parameters. Also note that the SGP with $p \geq 3$ incurs the highest penalty on its AIC value.

In the GP and the GRP, it is assumed that time to first failure has a Weibull distribution $1 - e^{-\left(\frac{t}{\beta}\right)^{\alpha}}$, respectively. In the NHPP-PL model, let $\frac{1}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1}$ be the intensity function of the system. The copulas $C_{k-1,k}(u,v)$ is assumed to be the Clayton copula, i.e. $C_{k-1,k}(u,v) = (u^{-\phi} + v^{-\phi} - 1)^{-1/\phi}$.

A data set published in Musa *et al.* (1987) is used to compare the AIC values of the SGP and the GP. The data set is the times between failures (in CPU seconds, measured in terms of execution time) of a real-time command and control software system and has been often analysed in the reliability literature (see, e.g. Peng *et al.*, 2014). The sample size of the data set is 135, as shown in Fig. 1. There are three observations that are zeroes, we replace them with 0.5. Lam (2007a) built an NHPP, a GP and an RP on this data set and finds that the MSE (squared mean error) of the GP is the smallest.

The Ljung–Box test is carried on the data set and shows that there is strong evidence of non-zero autocorrelations in the series $\{X_k, k = 1, 2, ...\}$ at lag 1, and the Mann–Kendall trend test rejects the null hypothesis that there is no monotonic trend in $\{X_k, k = 1, 2, ...\}$. This suggests that the SGP may be fitted to the data set.

4.1. *The LS method*

One may search for the optimal values of the parameters in the optimization problem shown in equation (3.6) and then find the AIC value on the data set. As can be seen from the results in Table 1, the AIC value of the SGP outperform those of the GP on the data set, which suggests that the SGP fits better than

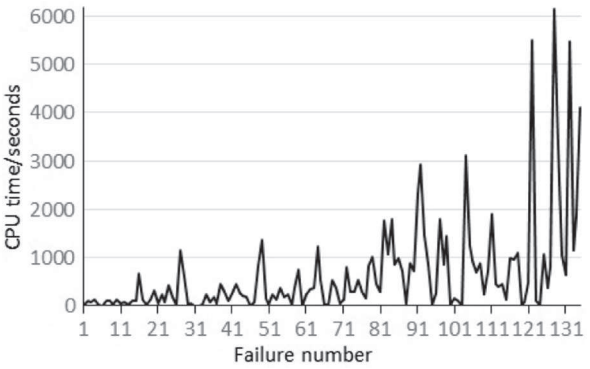

FIG. 1. Time between failure data of a real-time command and control software system.

TABLE 1. *Estimated parameters and AIC values (the LS estimation)*

The GP		The SGP	
Parameters	AIC_{SGP}	Parameters	AIC_{GP}
$\hat{\mu} = 3.857, \hat{a} = 0.978$	139.061	$\hat{\rho} = 0.161, \hat{\mu} = 3.856, \hat{a} = 0.978$	137.557

the GP on the data set. The standard errors of $\hat{\rho}$, $\hat{\mu}'(= (1 - \hat{\rho})\mu)$ and $\hat{\beta}(= \ln(a))$ are 0.0851, 0.435 and 0.00433, respectively. Below interprets the parameters of the SGP.

(a) The parameter $\hat{a} = 0.978$ is smaller than 1, which indicates that the times-between-failures are SI. This agrees with the observations as shown in Fig. 1, which shows X_k is increasing with time t.

Note. Since $\hat{a}(= 0.978)$ is close to 1, one may be curious whether there is a statistically significant difference between the modelling results with $\hat{a} = 0.978$ and those with $\hat{a} = 1$. It should be noted that the cdf of X_k is $F(a^{k-1}x)$ and a^{k-1} is sensitive to its base a. For example, if $a = 0.978$, then $a^{100} = 0.1081149$; if $a = 0.979$, then $a^{100} = 0.1197483$. That is, if a increases 0.001, the result of a^{100} increases 0.01163337, which is approximately 11.6 times larger than 0.001.

(b) Since $\hat{\rho} > 0$, according to equation (3.5) and Remark 3.1, X_k has a positive correlation with X_{k-1}.

4.2. *The ML method*

With the ML method, we obtain the parameters of the models, as shown in Table 2.

From Table 2, the AIC value of the SGP is the smallest among the four models.

When the SGP is built on the data set, the estimated parameters are $\hat{a} = 0.978, \hat{\phi} = 1.125, \hat{\alpha} = 0.854$ and $\hat{\beta} = 96.044$. The parameters of the SGP are interpreted below. The standard errors of $\hat{a}$, $\hat{\phi}$, $\hat{\alpha}$ and $\hat{\beta}$ are 0.00154, 0.0697, 0.0570 and 10.977, respectively.

(a) The parameter $\hat{a} = 0.978$ is smaller than 1, which indicates that the times-between-failures are SI. This agrees with the observations as shown in Fig. 1.

TABLE 2. *Estimated parameters and AIC values (the ML estimation)*

GP		NHPP		GRP		SGP	
Parameters	AIC_{GP}	Parameters	AIC_{NHPP}	Parameters	AIC_{GRP}	Parameters	AIC_{SGP}
$\hat{a} = 0.977, \hat{\alpha} = 0.864$	1925.023	$\hat{\alpha} = 0.495$	1939.969	$\hat{q} = 0.133, \hat{\alpha} = 0.416$	1938.861	$\hat{a} = 0.978, \hat{\phi} = 1.125$	1923.043
$\hat{\beta} = 91.894$		$\hat{\beta} = 2.081$		$\hat{\beta} = 7.576$		$\hat{\alpha} = 0.854, \hat{\beta} = 96.044$	

(b) Since $C_{k-1,k}(u, v) = (u^{-\phi} + v^{-\phi} - 1)^{-1/\phi}$, with Lemma 2.4,

$$\begin{aligned}\frac{\partial S(x|y)}{\partial y} &= f_{k-1}(y)\left[\frac{\partial^2 C_{k-1,k}(u, v; \boldsymbol{\theta})}{\partial u^2}|_{(u=F_{k-1}(y), v=F_k(x))}\right] \\ &= f_{k-1}(y)\left[u^{-\phi-2}(\phi^{-1} - (\phi + 1))(u^{-\phi} + v^{-\phi} - 1)^{-1/\phi}|_{(u=F_{k-1}(y), v=F_k(x))}\right].\end{aligned}$$

If $\phi^{-1} - (\phi + 1) > 0$, i.e. $\phi > \frac{\sqrt{5}-1}{2}$, then $\frac{\partial S(x|y)}{\partial y} > 0$. That implies, $S(x|y)$ is increasing in y. The parameter $\hat{\phi} = 1.125$ in the SGP is greater than $\frac{\sqrt{5}-1}{2}$, which suggests that X_k is SI in X_{k-1}.

This finding agrees with the second finding (b) in the LS method in Section 4.1.

(c) The shape parameter $\hat{\alpha} = 0.854$ in the SGP is less than 1, which suggests that the hazard function X_k is decreasing with respect to time t.

5. Conclusion

The GP is an extension of the RP in a sense that it assumes that the distributions of the times between occurrences of events (gap times) are non-stationary while retaining the assumption that the gap times are statistically independent. This process has been applied in different areas and has attracted extensive attention in maintenance policy optimization.

This paper extends the GP to a process in which times between occurrences of events are dependent. We obtain some probabilistic and statistical properties of the new process and derive two parameter estimation methods (ML estimation and LS estimation). A real-world data set is used to compare the performance (in terms of the AIC value) of the new process with three well-known stochastic processes, the RP, the GP, and the GRP. Our results show that the proposed model outperforms these three models.

Funding

This research was supported by the Economic and Social Research Council of the United Kingdom (Project Ref: ES/L011859/1) and the National Natural Science Foundation of China (Grant No 11671080).

Acknowledgement

The authors are indebted to the two reviewers for their suggestions for improving the clarity of the presentation.

REFERENCES

AYDOĞDU, H., & ALTINDAĞ, Ö. (2016) Computation of the mean value and variance functions in geometric process. *J. Stat. Comput. Simul.*, **86**, 986–995.

BARLOW, R. E., & PROSCHAN, F. (1975) *Statistical Theory of Reliability and Life Testing: Probability Models* (Tech. Rep.). Tallahassee: Florida State University.

BERNARD, C., & VANDUFFEL, S. (2015) A new approach to assessing model risk in high dimensions. *J. Banking Finance*, **58**, 166–178.

BORDES, L., & MERCIER, S. (2013). Extended geometric processes: semiparametric estimation and application to reliability. *J Iran. Stat. Soc.*, **12**, 1–34.

BRAUN, W. J., LI, W., & ZHAO, Y. Q. (2005). Properties of the geometric and related processes. *Nav. Res. Logist.*, **52**, 607–616.

CHAN, J. S., CHOY, S. B., & LAM, C. P. (2014). Modeling electricity price using a threshold conditional autoregressive geometric process jump model. *Commun. Stat. Theory Methods*, **43**, 2505–2515.

CHAN, J. S., YU, P. L., LAM, Y., & HO, A. P. (2006). Modelling sars data using threshold geometric process. *Stat. Med.*, **25**, 1826–1839.

CHERUBINI, U., MULINACCI, S., GOBBI, F., & ROMAGNOLI, S. (2011) *Dynamic Copula Methods in Finance*, Vol. 625. West Essex: John Wiley & Sons.

CHUKOVA, S., HAYAKAWA, Y., & DOHI, T. (2005). Warranty cost analysis: quasi-renewal interrepair times. *Int. J. Qual. Reliab. Manage.*, **22**, 687–698.

COCHRANE, D., & ORCUTT, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *J. Am. Statis. Assoc.*, **44**, 32–61.

COOK, R. J., & LAWLESS, J. F. (2007) *The Statistical Analysis of Recurrent Events.* New York: Springer Science & Business Media.

ERYILMAZ, S. (2016). A reliability model for a three-state degraded system having random degradation rates. *Reliab. Eng. Syst. Saf.*, **156**, 59–63.

FINKELSTEIN, M. (1993). A scale model of general repair. *Microelectron. Reliab.*, **33**, 41–44.

GUO, C., WANG, W., GUO, B., & PENG, R. (2013). Maintenance optimization for systems with dependent competing risks using a copula function. *Eksploatacja i Niezawodność*, **15**, 9–17.

JAIN, M., & GUPTA, R. (2013). Optimal replacement policy for a repairable system with multiple vacations and imperfect fault coverage. *Comput. Ind. Eng.*, **66**, 710–719.

JOE, H. (1997) *Multivariate Models and Multivariate Dependence Concepts*. London: CRC Press.

KENDALL, M. G. (1975) *Rank Correlation Methods*. Londin, England: Charles Griffin.

LAM, Y. (1988). Geometric processes and replacement problem. *Acta Math. Appl. Sin.*, **4**, 366–377.

LAM, Y. (2007a) *The Geometric Process and Its Applications*. Singapore: World Scientific.

LAM, Y. (2007b). A geometric process maintenance model with preventive repair. *Eur. J. Oper. Res.*, **182**, 806–819.

LIANG, Z., & PARLIKAD, A. K. (2015). A tiered modelling approach for condition-based maintenance of industrial assets with load sharing interaction and fault propagation. *IMA J. Manage. Math.*, **26**, 125–144.

LJUNG, G. M., & BOX, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

MANN, H. B. (1945). Nonparametric tests against trend. *Econometrica*, **13**, 245–259.

MCKNIGHT, S. D., MCKEAN, J. W., & HUITEMA, B. E. (2000). A double bootstrap method to analyze linear models with autoregressive error terms. *Psychological Meth.*, **5**, 87–101.

MUSA, J. D., IANNINO, A., & OKUMOTO, K. (1987). *Software Reliability: Measurement, Prediction, Application*. New York: McGraw-Hill, Inc.

NEWTON, M. A., & RAFTERY, A. E. (1994). Approximate bayesian inference with the weighted likelihood bootstrap. *J. R. Statist. Soc. Ser. B (Methodological)*, **56**, 3–48.

PENG, R., LI, Y., ZHANG, W., & HU, Q. (2014). Testing effort dependent software reliability model for imperfect debugging process considering both detection and correction. *Reliab. Eng. Syst. Saf.*, **126**, 37–43.

PENG, W., LI, Y.-F., MI, J., YU, L., & HUANG, H.-Z. (2016). Reliability of complex systems under dynamic conditions: a Bayesian multivariate degradation perspective. *Reliab. Eng. Syst. Saf.*, **153**, 75–87.

Ross, S. M. (1996) *Stochastic Processes*, 2nd edn. New York: John Wiley & Sons.
Seber, G. & Wild, C. (2003) *Nonlinear Regression*. New York: Wiley & Sons.
Sheu, S.-H., Chang, C.-C., Chen, Y.-L., & Zhang, Z. G. (2015). Optimal preventive maintenance and repair policies for multi-state systems. *Reliab. Eng. Syst. Saf.*, **140**, 78–87.
Sklar, A. (1959). Fonctions de répartition án dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231.
Smith, W., & Leadbetter, M. (1963). On the renewal function for the Weibull distribution. *Technometrics*. **5**, 393–396.
Vu, H. C., Do, P., Barros, A., & Bérenguer, C. (2015). Maintenance planning and dynamic grouping for multi-component systems with positive and negative economic dependencies. *IMA J. Manage. Math.*, **26**, 145–170.
Wan, W.-Y., & Chan, J. S.-K. (2011). Bayesian analysis of robust poisson geometric process model using heavy-tailed distributions. *Computat. Statist. Data Anal.*, **55**, 687–702.
Wang, H., & Pham, H. (1996). A quasi renewal process and its applications in imperfect maintenance. *Int. J. Syst. Sci.*, **27**, 1055–1062.
Wu, S. (2017) Doubly geometric processes and applications. *J. Oper. Res. Soc.*, doi:10.1057/s41274-017-0217-4.
Wu, S., & Clements-Croome, D. (2006). A novel repair model for imperfect maintenance. *IMA J. Manage. Math.*, **17**, 235–243.
Wu, S., & Scarf, P. (2015). Decline and repair, and covariate effects. *Eur. J. Oper. Res.*, **244**, 219–226.
Wu, S., & Scarf, P. (2017). Two new stochastic models of the failure process of a series system. *Eur. J. Oper. Res.*, **257**, 763–772.
Zhao, X., Liu, H.-C., & Nakagawa, T. (2015). Where does whichever occurs first hold for preventive maintenance modelings? *Reliab. Eng. Syst. Saf.*, **142**, 203–211.

Appendix

Proof of Lemma 2.2. According to Cherubini *et al.* (2011, p. 40),

$$\begin{aligned} F_{k|k-1}(x|y) &= P(X_k < x|X_{k-1} = y) \\ &= \lim_{h\to 0} P(X_k < x|y \le X_{k-1} \le y+h) \\ &= \lim_{h\to 0} \frac{P(X_k < x, y \le X_{k-1} \le y+h)}{P(y \le X_{k-1} \le y+h)} \\ &= \lim_{h\to 0} \frac{P(X_k < x, X_{k-1} \le y+h) - P(X_k < x, X_{k-1} \le y)}{P(X_{k-1} \le y+h) - P(X_{k-1} \le y)} \\ &= \lim_{h\to 0} \frac{C(F_{k-1}(y+h), F_k(x); \boldsymbol{\theta}) - C(F_{k-1}(y), F_k(x); \boldsymbol{\theta})}{F_{k-1}(y+h) - F_{k-1}(y)}. \end{aligned} \tag{A.1}$$

When $h \to 0$, we may approximate $F_{k-1}(y+h)$ with $F_{k-1}(y) + \Delta h$. Hence, we have

$$\begin{aligned} F_{k|k-1}(x|y) &= \lim_{\Delta h\to 0} \frac{C(F_{k-1}(y) + \Delta h, F_k(x); \boldsymbol{\theta}) - C(F_{k-1}(y), F_k(x); \boldsymbol{\theta})}{\Delta h} \\ &= \frac{\partial C_{k-1,k}(u, v; \boldsymbol{\theta})}{\partial u}|_{(u=F_{k-1}(y), v=F_k(x))}. \end{aligned} \tag{A.2}$$

Denote the density function of $F_{k|k-1}(x|y;\boldsymbol{\theta})$ as $f_{k|k-1}(x|y;\boldsymbol{\theta})$ for each $\boldsymbol{\theta}$ and for $k = 1, 2....$ Then

$$f_{k|k-1}(x|y) = c_{k-1,k}(F_{k-1}(y), F_k(x);\boldsymbol{\theta})f_k(x), \tag{A.3}$$

where $c_{k-1,k}(u, v;\boldsymbol{\theta}) = \frac{\partial^2 C_{k-1,k}(u,v;\boldsymbol{\theta})}{\partial u \partial v}$.

Hence, the hazard function of the item in the kth cycle is given by

$$\lambda_{k|k-1}(x|y) = \frac{f_{k|k-1}(x|y;\boldsymbol{\theta})}{1 - F_{k|k-1}(x|y;\boldsymbol{\theta})} \tag{A.4}$$

$$= \frac{c_{k-1,k}(F_{k-1}(y), F_k(x);\boldsymbol{\theta})f_k(x)}{1 - \frac{\partial C_{k-1,k}(u,v;\boldsymbol{\theta})}{\partial u}|_{(u=F_{k-1}(y),v=F_k(x))}}. \tag{A.5}$$

This establishes the lemma. □

Proof of Lemma 2.4. Recall Lemma 2.2, $P(X_k < x|X_{k-1} = y) = \frac{\partial C_{k-1,k}(u,v;\boldsymbol{\theta})}{\partial u}|_{(u=F_{k-1}(y),v=F_k(x))}$. Hence

$$\frac{\partial S(x|y)}{\partial y} = f_{k-1}(y)\left[\frac{\partial^2 C_{k-1,k}(u, v;\boldsymbol{\theta})}{\partial u^2}|_{(u=F_{k-1}(y),v=F_k(x))}\right].$$

Considering Lemma 2.3, one establishes Lemma 2.4. □

Proof of Lemma 2.5. The Fréchet–Hoeffding theorem states that for any Copula $C : [0, 1]^d \to [0, 1]$ and any $(u_1, \ldots, u_n) \in [0, 1]^n$, the following bounds hold:

$$W(u_1, \ldots, u_n) \le C(u_1, \ldots, u_n) \le M(u_1, \ldots, u_n). \tag{A.6}$$

The functions W and M are called the lower and upper Fréchet–Hoeffding bounds, respectively. They are defined as $W(u_1, \ldots, u_n) = \max\left\{1 - n + \sum_{k=1}^{n} u_k, 0\right\}$ and $M(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}$.

- If $a < 1$, then $a^{k-1} \ge a^k$ and $F(a^{k-1}x) \ge F(a^k x)$. Hence, we have

$$W(u_1, \ldots, u_n)$$
$$= \max\left\{1 - n + \sum_{k=1}^{n} u_k, 0\right\} = \max\left\{1 - n + \sum_{k=1}^{n} F_k(x_k), 0\right\} \ge \max\{1 - n + nF_n(x_n), 0\}$$

and

$$M(u_1, \ldots, u_n) = \min\{F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\} = F_n(x_n).$$

That is,

$$\max\{1 - n + nF_n(x_n), 0\} \le C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \le F_n(x_n). \tag{A.7}$$

- Similarly, for $a > 1$, we have

$$\max\{1 - n + nF_1(x_1), 0\} \le C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \le F_1(x_1). \tag{A.8}$$

This proves the lemma. □

Proof of Lemma 2.6. According to the Chapman–Kolmogorov equation,

$$P(X_1 < x_1, ..., X_n < x_n) = P(X_1 < x_1) \prod_{k=2}^{n} P(X_k < x_k | X_{k-1} < x_{k-1}) \le P(X_1 < x_1). \tag{A.9}$$

Since $P(X_k < x | X_{k-1} < y) = \frac{P(X_{k-1} < y, X_k < x)}{F_{k-1}(y)}$, if $P\{X_k < t, X_{k-1} < s\}$ is PQD, then $P(X_k < x_k, X_{k-1} < x_{k-1}) \ge P(X_k < x_k)P(X_{k-1} < x_{k-1})$. Hence, we can obtain

$$\begin{aligned} P(X_1 < x_1, ..., X_n < x_n) &= P(X_1 < x_1) \prod_{k=2}^{n} \frac{P(X_k < x_k, X_{k-1} < x_{k-1})}{P(X_{k-1} < x_{k-1})} \\ &\ge P(X_1 < x_1) \prod_{k=2}^{n} \frac{P(X_k < x_k) P(X_{k-1} < x_{k-1})}{P(X_{k-1} < x_{k-1})} \\ &= \prod_{k=1}^{n} P(X_k < x_k). \end{aligned} \tag{A.10}$$

This establishes the first part of the lemma.
Similarly, the second part can be established. □

Proof of Lemma 2.8. • If $a > 1$, then $F_1(t_1) \le F_k(t_k) \le F_n(t_n)$ for $n > k$. Then, we have

$$\inf_{t \in T(s)} \min\left\{\sum_{k=1}^{n} F_1(t_k) - n + 1, 0\right\} \le \inf_{t \in T(s)} \min\left\{\sum_{k=1}^{n} F_k(t_k) - n + 1, 0\right\}$$

and

$$\sup_{t \in T(s)} \max\left\{\sum_{k=1}^{n} F_n(t_k) - n + 1, 0\right\} \ge \sup_{t \in T(s)} \max\left\{\sum_{k=1}^{n} F_k(t_k) - n + 1, 0\right\}.$$

- If $a < 1$, then $F_n(t_n) \le F_k(t_k) \le F_1(t_1)$ for $n > k$. Then, we have

$$\inf_{t \in T(s)} \min\left\{\sum_{k=1}^{n} F_n(t_k) - n + 1, 0\right\} \le \inf_{t \in T(s)} \min\left\{\sum_{k=1}^{n} F_k(t_k) - n + 1, 0\right\}$$

and

$$\sup_{\boldsymbol{t}\in T(s)} \max\left\{\sum_{k=1}^{n} F_1(t_k) - n + 1, 0\right\} \geq \sup_{\boldsymbol{t}\in T(s)} \max\left\{\sum_{k=1}^{n} F_k(t_k) - n + 1, 0\right\}.$$

From the above inequalities, it is easy to establish the lemma. □