

Kent Academic Repository

Full text document (pdf)

Citation for published version

Wu, Shaomin (2018) Doubly geometric processes and applications. *Journal of the Operational Research Society*, 69 (1). pp. 66-77. ISSN 0160-5682.

DOI

<https://doi.org/10.1057/s41274-017-0217-4>

Link to record in KAR

<http://kar.kent.ac.uk/60730/>

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

Doubly geometric processes and applications

Shaomin Wu*

Kent Business School, University of Kent, Canterbury, Kent CT2 7PE, UK

Abstract

The geometric process has attracted extensive research attention from authors in reliability mathematics since its introduction. However, it possesses some limitations, which include that: (1) it can merely model stochastically increasing or decreasing inter-arrival times of recurrent event processes, and (2) it cannot model recurrent event processes where the inter-arrival time distributions have varying shape parameters. Those limitations may prevent it from a wider application in the real world.

In this paper, we extend the geometric process to a new process, the *doubly geometric process*, which overcomes the above two limitations. Probability properties are derived and two methods of parameter estimation are given. Application of the proposed model is presented: one is on fitting warranty claim data and the other is to compare the performance of the doubly geometric process with the performance of other widely used models in fitting real world datasets, based on the corrected Akaike information criterion.

Keywords: failure process modelling, geometric processes, recurrent events, Poisson processes, maintenance.

1 Introduction

1.1 Motivation

Since its introduction by Lam (1988), the geometric process (GP) has attracted extensive research attention. A considerable bulk of research on the GP, including more than 200 papers and one monograph (Lam, 2007), has been published. For example, the GP has been applied in system

*E-mail: s.m.wu@kent.ac.uk. This paper is accepted by *Journal of the Operational Research Society* and is in press.

23 reliability analysis (Yuan and Meng, 2011; Jain and Gupta, 2013), maintenance policy optimisation
 24 (Zhang, et.al, 2002; Liu and Huang, 2010; Wang, 2011; Zhang, et.al, 2013), warranty cost analysis
 25 (Chukova, et. al., 2005), modelling of the outbreak of an epidemic disease (Chan, et.al., 2006),
 26 and modelling of electricity prices (Chan, et al., 2014). In the meantime, some authors propose
 27 extended models to overcome the limitations of the GP (Finkelstein, 1993; Wang and Pham, 1996;
 28 Braun, et. al., 2005; Chan, et.al., 2006; Wu and Clements-Croome, 2006).

29 The GP is a stochastic process that is defined as (Lam, 1988): a sequence of random variables
 30 $\{X_k, k = 1, 2, \dots\}$ is a GP if the cdf (cumulative distribution function) of X_k is given by $F(a^{k-1}t)$
 31 for $k = 1, 2, \dots$ and a is a positive constant.

32 As can be seen, the distinction between the GP and the renewal process lies in the fact that
 33 the inter-arrival times of the renewal process have the same distribution $F(t)$ over k 's and the
 34 inter-arrival times of the GP have a cdf $F(a^{k-1}t)$, which changes over k 's. In some scenarios such
 35 as reliability mathematics, this distinction makes the GP more attractive in application as it can
 36 model the failure process of ageing or deteriorating systems, which may have decreasing working
 37 times between failures.

38 While the GP is an important model and has been widely used in solving problems in various
 39 research areas, its scope is still limited and does not fit the purposes of various empirical studies.
 40 First, this model is not suitable for a stochastic process in which the inter-arrival times may need
 41 to be modelled by distributions with varying shape parameters. Second, it can merely describe
 42 stochastically increasing or decreasing stochastic processes. This paper aims to propose a new
 43 process that can overcome those two limitations and to study its probabilistic properties.

44 1.2 The geometric process and related work

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

47 **Definition 1** *Stochastic order (p. 404 in Ross (1996)). Assume that X and Y are two random*
 48 *variables. If for every real number r , the inequality*

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

49 *holds, then X is stochastically greater than or equal to Y , or $X \geq_{st} Y$. Equivalently, Y is*
 50 *stochastically less than or equal to X , or $Y \leq_{st} X$.*

51 From Definition 1, one can define the monotonicity of a stochastic process: Given a stochastic
 52 process $\{X_k, k = 1, 2, \dots\}$, if $X_k \leq_{st} X_{k+1}$ ($X_k \geq_{st} X_{k+1}$) for $k = 1, 2, \dots$, then $\{X_k, k = 1, 2, \dots\}$ is

53 said stochastically to be increasing (decreasing).

54 **Lemma 1** (*p. 405 in Ross (1996)*) Assume that X and Y are two random variables, then

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

55 for all increasing functions $u(\cdot)$.

56 Lam proposes the definition of the GP, as shown below (Lam, 1988).

57 **Definition 2** (*Lam, 1988*) Given a sequence of non-negative random variables $\{X_k, k = 1, 2, \dots\}$,
58 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
59 constant, then $\{X_k, k = 1, 2, \dots\}$ is called a geometric process (GP).

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

61 **Remark 1** From Definition 1 and Lemma 2, we have the following results.

- 62 • If $a > 1$, then $\{X_k, k = 1, 2, \dots\}$ is stochastically decreasing.
- 63 • If $a < 1$, then $\{X_k, k = 1, 2, \dots\}$ is stochastically increasing.
- 64 • If $a = 1$, then $\{X_k, k = 1, 2, \dots\}$ is a renewal process (RP).
- 65 • If $\{X_k, k = 1, 2, \dots\}$ is a GP and X_1 follows the Weibull distribution, then the shape param-
66 eter of X_k for $k = 2, 3, \dots$ remains the same as that of X_1 . This observation is not specific
67 to the Weibull distribution and holds for many other distributions with a scale and shape
68 parameter such as the Gamma distribution.

69 The GP offers an alternative process to model recurrent event processes. For example, in reliability
70 mathematics, the renewal process (RP) and the non-homogeneous Poisson process (NHPP) are
71 two widely used stochastic processes. The RP is normally used to model working times of a system
72 if the system is renewed (or replaced with new and identical items upon failures) and the NHPP
73 is used to model working times of a system where a repair restores the system to the status just
74 before the failure happened, i.e., the repair is a minimal repair. Those assumptions of the RP and
75 the NHPP may be too stringent in real applications. On the other hand, repairing a given item
76 may have a limited number of methods, which implies that repair effect on the item is not random
77 (Kijima, 1989). Meanwhile, the reliability of the item may decrease over time. Considering those
78 facts, time between failures may therefore become shorter and shorter. The GP can model time
79 between failures of such items.

80 Meanwhile, some authors either proposed similar definitions to that of the GP (Finkelstein,
 81 1993; Wang and Pham, 1996) or made an attempt to extend the GP (Braun, et. al., 2005; Wu
 82 and Clements-Croome, 2006; Lam, 2007). Those different versions can be unified as: they replace
 83 a^{k-1} with $g(k)$, where $g(k)$ is a function of k and is defined differently by different authors, as
 84 discussed below.

85 For a sequence of non-negative random variables $\{X_k, k = 1, 2, \dots\}$, different consideration
 86 has been laid on the distribution of X_k , as illustrated in the following (in chronological order).

87 (i) Finkelstein (1993) proposes a process, named the *general deteriorating renewal process*, in
 88 which the distribution of X_k is $F_k(x)$, where $F_{k+1}(x) \leq F_k(x)$. A more specific model is
 89 defined such that $F_k(x) = F(a_k x)$ where $1 = a_1 \leq a_2 \leq a_3 \leq \dots$ and a_k are parameters. In
 90 this model, $g(k) = a_k$.

91 (ii) Wang and Pham (1996) defines a quasi-renewal process, which assumes $X_1 = W_1, X_2 = aW_2,$
 92 $X_3 = a^2W_3, \dots$, and the W_k are independently and identically distributed and $a > 0$ is
 93 constant. Here, $g(k) = a^{1-k}$.

94 (iii) Braun, et. al. (2005) proposes a variant, which assumes that the distribution of X_k is
 95 $F_k(x) = F(k^{-a}x)$, or $g(k) = k^{-a}$. The authors proved that the expected number of event
 96 counts before a given time, or analogously, the Mean Cumulative Function (MCF) (or, the
 97 renewal function), tends to infinite for the decreasing GP. As such, they propose the process
 98 as a complement.

99 (iv) Wu and Clements-Croome (2006) set $g(k) = \alpha a^{k-1} + \beta b^{k-1}$, where α, β, a and b are param-
 100 eters. Their intention is to extend the GP to model more complicated failure patterns such
 101 as the bathtub shaped failure patterns.

102 (v) Chan, et.al. (2006) extends the GP to the threshold GP: A stochastic process $\{Z_n, n =$
 103 $1, 2, \dots\}$ is said to be a threshold geometric process (threshold GP), if there exists real numbers
 104 $a_i > 0, i = 1, 2, \dots$, and integers $\{1 = M_1 < M_2 < \dots\}$ such that for each $i = 1, 2, \dots$,
 105 $\{a_i^{n-M_i} Z_n, M_i \leq n < M_{i+1}\}$ forms a renewal process.

106 Apparently, the model proposed in Finkelstein (1993) has a limitation in common: there is a
 107 need to estimate a large number of parameters, which may be problematic in real applications as
 108 a large number of failure data are needed to estimate the parameters. It should be noted that it
 109 is notoriously difficult to collect a large number of failure data in practice.

1.3 Comments on the geometric process and its extensions

While the GP is an important model and widely used, its scope is still limited and does not fit the purposes of various empirical studies due to the following two limitations.

- *Invariance of the shape parameter.* Suppose the cdf $F_k(x)$ of X_k in the GP have a scale parameter and a shape parameter. Then, all of the above discussed GP-like variants and extensions implicitly make an assumption: the processes merely change the scale parameter of $F_k(x)$, but keep their shape parameter constant over k 's. In other words, none of the existing GP-like processes can model a recurrent event process whose shape parameter of $F_k(x)$ changes over k . To elaborate, let us take the Weibull distribution as an example. Assume that the cdf of X_1 is $F(x) = 1 - e^{-\left(\frac{x}{\theta_1}\right)^{\theta_2}}$. Then according to the GP-like processes, the cdf of X_k is $F(g(k)x) = 1 - \exp\left\{-\left(\frac{x}{\theta_1 g^{-1}(k)}\right)^{\theta_2}\right\}$. That is, the scale parameter $\theta_1 g^{-1}(k)$ is a function of k and it changes over k 's, but the shape parameter θ_2 is independent of k and remains constant over different k 's. This assumption may be too stringent and should be relaxed for a wider application. To this end, one may assume a natural extension of the GP, in which X_k has a cdf $F(g(k)x^{h(k)})$, where $h(k)$ is a function of k and the parameters in $h(k)$ are estimable. As a result, in the Weibull distribution case, for example, the inter-arrival times, X_k 's, may be fitted with cdf $F(g(k)x^{h(k)}) = 1 - \exp\left\{-\left(\frac{x}{(\theta_1 g^{-1}(k))^{1/h(k)}}\right)^{\theta_2 h(k)}\right\}$.

A similar description of the above paragraph is the invariance of the CV (coefficient of variation). Assume that $\{X_1, X_2, \dots\}$ follows the GP. Denote $\lambda_{11} = \mathbb{E}[X_1]$ and $\lambda_{21} = \mathbb{E}[X_1^2] - \lambda_{11}^2$. Then it is easy to obtain the expected value and the variance of X_k : $\mathbb{E}[X_k] = a^{(1-k)}\lambda_{11}$ and $\mathbb{V}[X_k] = a^{(2-2k)}\lambda_{21}$, respectively. The coefficient of variation (CV) of X_k is therefore given by $\gamma_k = \frac{\sqrt{\mathbb{V}[X_k]}}{\mathbb{E}[X_k]} = \sqrt{\lambda_{21}}/\lambda_{11}$, which suggests that the CVs are independent of k and keep constant over k 's.

An example of such a process with varying shape parameters in $F_k(x)$ can be found in Chan, et.al. (2006), in which X_k are the number of daily infected cases of an epidemic disease (i.e., the severe acute respiratory syndrome) in Hong Kong in 2003 are assumed to be independent and follow the threshold geometric process, in which $F_k(x)$ have different shape parameters for $k = 1, 2, \dots$.

- *Monotonicity of the GP.* From Remark 1, the GP $\{X_k, k = 1, 2, \dots\}$ change monotonously. That is, it can merely model the processes with increasing or decreasing inter-arrival times, or renewal processes. It is known, however, that the inter-arrival time processes of some real-world systems may exhibit non-monotonous failure patterns. For those systems, using the GP to model their failure processes is apparently inappropriate.

143 1.4 Contribution and importance of this work

144 This paper proposes a new stochastic process, the *doubly geometric process (DGP)*, which makes
145 contribution to the literature in the following aspects.

- 146 • First, the DGP can model recurrent event processes where $F_k(x)$'s have different shape
147 parameters over k 's, which can be done by neither the GP-like models nor other repair
148 models such as reduction of age models discussed in Doyen and Gaudoin (2004). One
149 may note that the DGP differs from the research that treats the parameters in a lifetime
150 distribution as functions of time (Zuo et al., 1999).
- 151 • Second, the DGP can model not only monotonously increasing or decreasing stochastic
152 processes, but also processes with complicated failure intensity functions such as the bathtub
153 shaped curves and the upside-down bathtub shaped curves, as can be seen from examples
154 shown in Fig. 1. Noteworthy, although the models proposed by Wu and Clements-Croome
155 (2006) and Chan, et.al. (2006) can also model complicated failure intensity functions, they
156 assume that $F_k(x)$'s have constant shape parameters over k 's and they need more parameters
157 than the DGP (i.e., the DGP needs 2 parameters whereas the models proposed by Wu and
158 Clements-Croome (2006) and Chan, et.al. (2006) need at least 3 parameters).
- 159 • Third, as Braun, et. al. (2005) points out, the GP has a limitation that *it only allows for*
160 *logarithmic growth or explosive growth*. The DGP can overcome this limitation.

161 One may also notice that, in recent years, many authors have devoted considerable effort on
162 developing novel methods to model repair processes, see Wu and Scarf (2015), for example. The
163 current paper can of course be regarded as a new contribution to the literature of modelling repair
164 processes.

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

171 1.5 Overview

172 The rest of the paper is structured as follows. Section 2 introduces the DGP and discusses its
173 probabilistic properties. Section 3 proposes methods of parameter estimation. Section 4 compares

174 the performance of the DGP with that of other models based on datasets collected from the
 175 real-world. We finish with a conclusion and future work in Section 5.

176 2 A doubly geometric process and its probabilistic prop- 177 erties

178 In this section, we propose the following definition and then discuss its statistical properties.

179 **Definition 3** *Given a sequence of non-negative random variables $\{X_k, k = 1, 2, \dots\}$, if they are
 180 independent and the cdf of X_k is given by $F(a^{k-1}x^{h(k)})$ for $k = 1, 2, \dots$, where a is a positive
 181 constant, $h(k)$ is a function of k and the likelihood of the parameters in $h(k)$ has a known closed
 182 form, and $h(k) > 0$ for $k \in \mathbb{N}$, then $\{X_k, k = 1, 2, \dots\}$ is called a doubly geometric process (DGP).*

184 In the above definition, for the sake of simplicity, we call the process as *doubly geometric process*
 185 since the process can include two geometric processes: $\{a^{k-1}, k = 1, 2, \dots\}$ is a geometric series
 186 and $\{h(k), k = 1, 2, \dots\}$ can be a geometric series.

187 We refer to a^{k-1} as the scale impact factor and $h(k)$ as the shape impact factor. It should be
 188 noted that the cdf of X_1 is $F(x)$.

189 **Remark 2** *Similar to the definition of the quasi-renewal process given by Wang and Pham (1996),
 190 one may give an alternative definition of Definition 3 as: assume $X_1 = W_1$, $X_2 = (a^{-1}W_2)^{1/h(1)}$,
 191 \dots , $X_k = (a^{1-k}W_2)^{1/h(k)}$, \dots and the W_k are i.i.d., then the process $\{X_k, k = 1, 2, \dots\}$ is called a
 192 doubly geometric process.*

193 Although the extension from the GP to the DGP seems quite natural, it may create difficul-
 194 ties in mathematical derivation. For example, deriving some probability properties of the DGP
 195 becomes much more complicating than that of the GP, it is difficult to derive a closed-form of the
 196 MCF for the DGP whereas an explicit iteration equation of the MCF for the GP can be derived.

197 **Remark 3** *From Definition 3, it follows the results below.*

198 (i) *If $h(k) = 1$, then $\{X_k, k = 1, 2, \dots\}$ reduces to the geometric process.*

199 (ii) *Denote $\lambda_{1k} = \mathbb{E}[X_1^{h^{-1}(k)}] = \int_0^\infty x^{h^{-1}(k)} f(x) dx$ and $\lambda_{2k} = \mathbb{E}[X_1^{2h^{-1}(k)}] = \int_0^\infty x^{2h^{-1}(k)} f(x) dx$,
 200 where $f(x) = \partial F(x)/\partial x$ exists and $h^{-1}(k) = \frac{1}{h(k)}$. Assume that $\mathbb{E}[X_1^{h^{-1}(k)}] < \infty$ and
 201 $\mathbb{E}[X_1^{2h^{-1}(k)}] < \infty$. Then it is easy to obtain the expected value and the variance of X_k :
 202 $\mathbb{E}[X_k] = a^{(1-k)h^{-1}(k)} \lambda_{1k}$ and $\mathbb{V}[X_k] = a^{(2-2k)h^{-1}(k)} \lambda_{2k} - \lambda_{1k}^2$ for $k = 1, 2, \dots$.*

203 (iii) If X_1 follows the exponential distribution and

204 (a) if $\{X_k, k = 1, 2, \dots\}$ follows the GP, then X_k (for $k = 2, 3, \dots$) follows the exponential
205 distribution with different rate parameters from that of X_1 ,

206 (b) if $\{X_k, k = 1, 2, \dots\}$ follows the DGP, then X_k (for $k = 2, 3, \dots$) follows the Weibull
207 distribution,

208 (iv) If $\{X_k, k = 1, 2, \dots\}$ follows the DGP and X_1 follows the Weibull distribution, then X_k (for
209 $k > 1$) follows the Weibull distribution with different shape and scale parameters from those
210 of X_1 .

211 If we assume that $\{X_1, X_2, \dots\}$ follows the DGP. Then from (ii) in Remark 3, the coefficient
212 of variation (CV) of X_k is $\gamma_k = \frac{\sqrt{\text{V}[X_k]}}{\mathbb{E}[X_k]} = \frac{\sqrt{a^{(2-2k)h^{-1}(k)}\lambda_{2k} - \lambda_{1k}^2}}{a^{(1-k)h^{-1}(k)}\lambda_{1k}}$, which implies that the CVs change
213 over k 's. Hence, we can make the following conclusion.

214 **Lemma 2** Suppose that $\{X_k, k = 1, 2, \dots\}$ is a GP, then the coefficient of variation (CV) of X_k
215 changes over k 's.

216 Now a question arisen is the selection of the forms of $h(k)$. In what follows, we investigate the
217 DGP with the $h(k)$ defined below:

$$h(k) = (1 + \log(k))^b, \quad (1)$$

218 where \log is the logarithm with base 10 and b is a parameter.

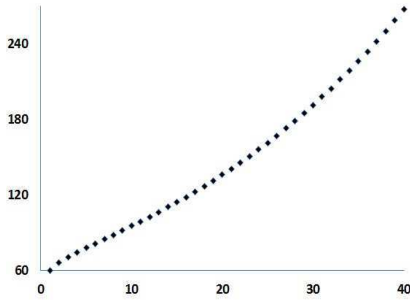
219 **2.1 Probabilistic properties of the DGP with $h(k) = (1 + \log(k))^b$**

220 In this entire section, i.e., Section 2.1, we assume $h(k) = (1 + \log(k))^b$.

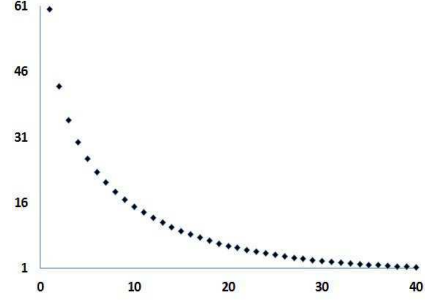
221 The reason that we select $h(k) = (1 + \log(k))^b$ is: we have fit the DGP with different $h(k)$,
222 which are b^{k-1} , $b^{\log(k)}$, and $1 + b \log(k)$, on ten real-world datasets (see Section 4) and found that
223 the DGP with $h(k) = (1 + \log(k))^b$ outperforms the processes with the other three $h(k)$'s. In real
224 applications, it is suggested that other form of $h(k)$ may also be investigated and selected once a
225 comparison on the performance of difference $h(k)$ has been made.

226 In selecting $h(k)$, one may set some conditions, for example, $h(1) = 1$ and $h(k) > 0$ for
227 $k = 1, 2, \dots$.

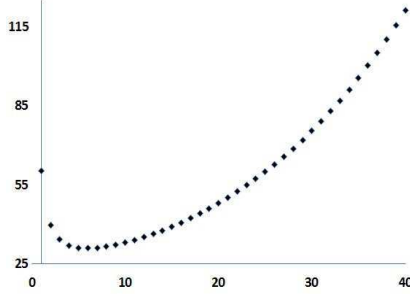
228 Unlike the GP that can only be either stochastically increasing or stochastically decreasing,
229 the DGP can model more flexible processes, as shown in the four examples in Figure 1.



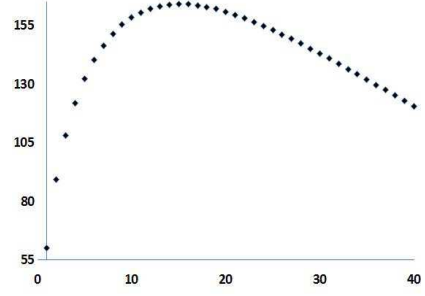
(a) $a = 0.97, b = -0.05, \theta_1 = 40$ and $\theta_2 = 0.6$.



(b) $a = 1.1, b = 0.2, \theta_1 = 40$ and $\theta_2 = 0.6$.



(c) $a = 0.92, b = 0.4, \theta_1 = 40$ and $\theta_2 = 0.6$.



(d) $a = 1.02, b = -0.3, \theta_1 = 40$ and $\theta_2 = 0.6$.

Figure 1: DGPs with different parameter settings.

230 **Proposition 1** Given a DGP $\{X_k, k = 1, 2, \dots\}$,

231 (i) if $0 < a < 1$, $P(X_1 > 1) = 1$, and $b < 0$, then $\{X_k, k = 1, 2, \dots\}$ is stochastically increasing.

232 (ii) if $a > 1$, $P(0 < X_1 < 1) = 1$, and $b < 0$, then $\{X_k, k = 1, 2, \dots\}$ is stochastically decreasing.

233 (iii) if $0 < a < 1$, $P(0 < X_1 < 1) = 1$, and $0 < b < 4.898226$, then $\{X_k, k = 1, 2, \dots\}$ is
234 stochastically increasing.

235 (iv) if $a > 1$, $P(X_1 > 1) = 1$, and $0 < b < 4.898226$, then $\{X_k, k = 1, 2, \dots\}$ is stochastically
236 decreasing.

237 **Proposition 2** Given a DGP $\{X_k, k = 1, 2, \dots\}$ with $h(k) = (1 + \log(k))^b$, if $(1 + \log(k +$
238 $1))^{-b}(\log(y) - k \log(a)) + (1 + \log(k))^{-b}((k - 1) \log(a) - \log(y))$ varies between negative and positive
239 values, then the DGP is not stochastically monotonous over k 's, where y represents all the possible
240 values on X_k (for $k = 1, 2, \dots$).

241 Stochastic ageing properties are widely discussed in the reliability literature. For example,
242 $F(t)$ is IFR (Increasing Failure Rate) if $\frac{f(t)}{F(t)}$ is increasing in t for all $t \geq 0$, where $f(t) = \frac{dF(t)}{dt}$

243 and $\bar{F}(t) = 1 - F(t)$. With regard to the stochastic ageing properties of the DGP, we have the
 244 following proposition.

245 **Proposition 3** Suppose $\{X_k, k = 1, 2, \dots\}$ follows the DGP. If $b > 0$ and $F(t)$ is IFR, then the
 246 cdf $F_k(t)$ of X_k is IFR.

247 Suppose $\{X_k, k = 1, 2, \dots\}$ follows the DGP, denote $S_n \equiv \sum_{k=1}^n X_k$ with $S_0 \equiv 0$. Then the
 248 distribution of S_n is

$$\begin{aligned}
 P(S_n \leq t) &= P(S_{n-1} + X_n \leq t) \\
 &= \int_0^t F^{(n-1)}(t-u) dF_n(u) \\
 &= \int_0^t F^{(n-1)}(t-u) \left(a^{n-1} (1 + \log(n))^b u^{(1+\log(n))^b-1} f(a^{n-1} u^{(1+\log(n))^b}) \right) du \\
 &= \int_0^{a^{n-1} t^{(1+\log(n))^b}} F^{(n-1)}(t - a^{(1-n)(1+\log(n))^{-b}} v^{(1+\log(n))^{-b}}) f(v) dv, \tag{2}
 \end{aligned}$$

249 where $F^{(0)}(t) = 1$ and $F^{(n)}(t) \equiv P(S_n \leq t)$. Let $N(t) = \max\{n : S_n \leq t\}$, then the MCF, $m(t)$, is
 250 given by

$$m(t) = \mathbb{E}[N(t)] = \sum_{n=1}^{\infty} P(S_n \leq t). \tag{3}$$

251 Denote

$$m_1(t) = \sum_{n=1}^{\infty} P\left(\sum_{k=1}^n Y_k \leq t\right), \tag{4}$$

252 where $\{Y_k : k \geq 1\}$ is a renewal process with $Y_k > 0$ and the cdf of the inter-arrival times
 253 is $F(x)$ (which has the same as the cdf of X_1). Then, equivalently, $m_1(t)$ is the MCF of the
 254 ordinary renewal process $\{N_1(t) : t \geq 0\}$ with $N_1(t) \equiv \max\{n : \sum_{k=1}^n Y_k \leq t\}$. For $\{Y_k : k \geq 1\}$,
 255 $m_1(t) = F(t) + \int_0^t m_1(t-y) dF(y)$, as can be seen in many textbooks of stochastic processes (for
 256 example, see Ross (1996)).

257 Unlike the MCF, $m_1(t)$, for the ordinary renewal process where an iteration equation can be
 258 given, deriving an iteration equation for $m(t)$ defined in Eq. (3) seems not an easy task. In real
 259 applications, numerical analysis may be sought. For example, on the four examples used in Figure
 260 1, we run the Monte Carlo simulation for 2000 times and estimate the values of the MCF for each
 261 example. Figure 2 shows the values of the MCF of the four examples with the parameter settings
 262 shown in Figure 1.

263 Below, the lower bounds or the upper bounds are given for two scenarios.

264 **Proposition 4** (i) Given that $m_1(t)$ and $m(t)$ are defined in Eq. (3) and Eq. (4), respectively,

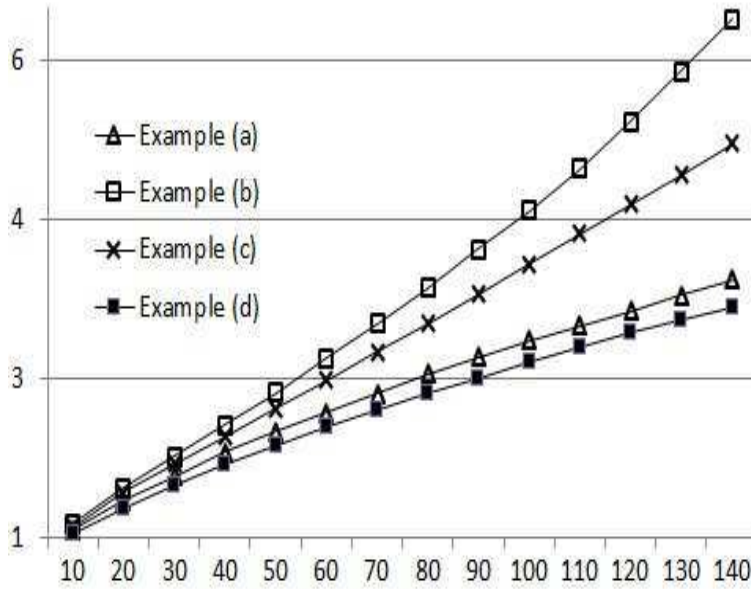


Figure 2: The MCF, $m(t)$, of the four examples shown in Figure 1.

265 if $\{X_k, k = 1, 2, \dots\}$ is stochastically non-decreasing, then

$$m(t) \leq m_1(t). \quad (5)$$

266 (ii) Suppose that $\{X_k, k = 1, 2, \dots\}$ follows the DGP and $P(X_k < c) = 1$ for $k = 1, 2, \dots$ and
 267 c is a positive real number. Denote $\Lambda_n = \sum_{k=1}^n \mathbb{E}[X_k]$ and $\sigma^2 = \frac{1}{n} \sum_{k=1}^n \mathbb{V}[X_k]$. Assume that
 268 $\{X_k, k = 1, 2, \dots\}$ is stochastically non-increasing and $t > \lim_{n \rightarrow \infty} \Lambda_n (< +\infty)$, then

$$m(t) \geq \max \left\{ m_1(t), \sum_{n=1}^{\infty} \left[1 - \exp \left(-\frac{n\sigma^2}{c^2} H \left(\frac{ct - c\Lambda_n}{n\sigma^2} \right) \right) \right] \right\}. \quad (6)$$

269 The following proposition compares the MCFs of the GP and the DGP.

270 **Proposition 5** Suppose that $\{X_k^g, k = 1, 2, \dots\}$ is a GP with $X_k^g \sim F(a^{k-1}x)$ and $\{X_k^d, k =$
 271 $1, 2, \dots\}$ is a DGP with $X_k^d \sim F(a^{k-1}x^{(1+\log(k))^b})$. Denote $m^g(t) = \sum_{n=1}^{\infty} P(\sum_{k=1}^n X_k^g \leq t)$ and
 272 $m^d(t) = \sum_{n=1}^{\infty} P(\sum_{k=1}^n X_k^d \leq t)$. Then,

273 (i) $m^g(t) > m^d(t)$ if $0 < a < 1, b < 0$ and $P(X_1 > 1) = 1$, or if $a > 1, b > 0$ and $P(0 < X_1 <$
 274 $1) = 1$.

275 (ii) $m^g(t) < m^d(t)$ if $0 < a < 1, b > 0$ and $P(X_1 > 1) = 1$, or if $a > 1, b < 0$ and $P(0 < X_1 <$
 276 $1) = 1$.

277 The following proposition compares the MCFs of two DGPs.

278 **Proposition 6** Suppose that $\{X_k^{d_1}, k = 1, 2, \dots\}$ with $X_k^{d_1} \sim F(a_1^{k-1}x^{(1+\log(k))^{b_1}})$ is a DGP and
 279 $\{X_k^{d_2}, k = 1, 2, \dots\}$ with $X_k^{d_2} \sim F(a_2^{k-1}x^{(1+\log(k))^{b_2}})$ is a DGP. Denote $m^{d_1}(t)$
 280 $= \sum_{n=1}^{\infty} P(\sum_{k=1}^n X_k^{d_1} \leq t)$ and $m^{d_2}(t) = \sum_{n=1}^{\infty} P(\sum_{k=1}^n X_k^{d_2} \leq t)$.

281 (i) If $a_1 = a_2$ and $b_1 > b_2$,

- 282 • $m^{d_1}(t) < m^{d_2}(t)$ if $a > 1$ and $P(0 < X_1 < 1) = 1$,
- 283 • $m^{d_1}(t) > m^{d_2}(t)$ if $0 < a < 1$ and $P(X_1 > 1) = 1$.

284 (ii) $m^{d_1}(t) < m^{d_2}(t)$ if $b_1 = b_2$ and $a_1 > a_2$.

285 (iii) $m^{d_1}(t) > m^{d_2}(t)$ if $a_2 > a_1 > 1$, $b_1 > b_2$, and $P(X_1 > 1) = 1$.

286 (iv) $m^{d_1}(t) < m^{d_2}(t)$ if $0 < a_1 < a_2 < 1$, $b_1 > b_2$, and $P(X_1 < 1) = 1$.

287 Proposition 1 shows the monotonicity property of the DGP, but it has not shown the conver-
 288 gence of the DGP in probability. The following property addresses this issue.

289 **Proposition 7** Given a DGP $\{X_k, k = 1, 2, \dots\}$,

- 290 (i) if $0 < a < 1$, then X_k converges to infinity in probability as $k \rightarrow \infty$,
- 291 (ii) if $a > 1$, then X_k converges to zero in probability as $k \rightarrow \infty$.

292 2.2 Discussion

293 We make the following discussion.

- 294 • *On the scale impact factor $g(k)$ and the shape impact factor $h(k)$.* Although we only discussed
 295 the DGP in which the scale impact factor is set to $g(k) = a^{k-1}$, $g(k)$ may also be replaced with
 296 other forms of functions such as those proposed in Finkelstein (1993); Braun, et. al. (2005);
 297 Wu and Clements-Croome (2006); Chan, et.al. (2006). The function $h(k) = (1 + \log(k))^b$
 298 in Eq. (1) can be replaced with any other functions of k , for example, $h(k) = b^{k-1}$, or
 299 $h(k) = b^{\log(k)}$ etc. However, the propositions of DGPs with different $g(k)$ and $h(k)$ are
 300 discussed in the following bullet.
- 301 • *On the propositions.* Among the propositions discussed in Section 2.1, Proportion 4 holds
 302 for any $g(k)$ and $h(k) > 0$ as both $g(k)$ and $h(k) > 0$ are not involved in the proof process
 303 of Proposition 4. But the other propositions are discussed for the case where $g(k) = a^{k-1}$
 304 and $h(k) = (1 + \log(k))^b$.

3 Estimation of the parameters in the DGP

In this section, we discuss two methods of estimation of the parameters in the DGP.

3.1 Least squares method

For the geometric process, Lam (1992) develops a method, which is a least squares method, to estimate the parameters in the GP. With a similar method, we estimate the parameters in the DGP in this section.

Suppose that a process $\{X_k, k = 1, 2, \dots\}$ follows the DGP with $X_k \sim F(a^{k-1}x^{(1+\log(k))^b})$. Let

$$Z_k = a^{k-1}X_k^{(1+\log(k))^b}. \quad (7)$$

Then $\{Z_k, k = 1, 2, \dots\}$ follows an ordinary renewal process. Given observations x_k of X_k (for $k = 1, 2, \dots$), from Eq. (7), we can have

$$\mu = a^{k-1}x_k^{(1+\log(k))^b} + e_k \quad (8)$$

where $\mu = \mathbb{E}[Z_k]$ and e_k are i.i.d. random variables each having mean 0 and a constant variance.

When $b \neq 0$, it is not possible to linearise model (8) by means of a suitable transformation, that is, model (8) is intrinsically nonlinear.

For given observations x_k of X_k (with $k = 1, 2, \dots, N_0$), one can minimise the following sum of the squares of the errors to estimate the parameters a , b and μ .

$$(\hat{\mu}, \hat{a}, \hat{b}) = \arg \min_{\mu, a, b} \sum_{k=1}^{N_0} \left(x_k - (\mu a^{1-k})^{(1+\log(k))^{-b}} \right)^2. \quad (9)$$

Obviously, there is no general closed-form solution for $\hat{\mu}$, \hat{a} , and \hat{b} , one needs therefore pursue nonlinear programming methods to solve the problem.

The reader is referred to Theorem 2.1 in page 24 in the book by Seber and Wild (2003) for obtaining the asymptotic distributions of $(\hat{\mu}, \hat{a}, \hat{b})$.

3.2 Maximum likelihood method

Suppose that one observes N systems starting from time 0 until time T . Assume that system j ($j = 1, 2, \dots, N$) has failed for N_j times at time points $s_{j,k}$ with $k = 0, 1, \dots, N_j$. Let $s_{j,0} = 0$. Then the working times of system j are $s_{j,1} - s_{j,0}$, $s_{j,2} - s_{j,1}$, \dots , $s_{j,N_j} - s_{j,N_j-1}$, and $T - s_{j,N_j}$,

327 respectively. Denote $x_{j,i} = s_{j,i} - s_{j,i-1}$ for $i = 1, 2, \dots, N_j$ and $x_{j,N_j+1} = T - s_{j,N_j}$.

328 Then, for the DGP with $h(k) = (1 + \log(k))^b$, the likelihood function is given by

$$\begin{aligned}
 L(a, b, \boldsymbol{\theta}) &= \prod_{j=1}^N \left\{ \left[1 - F(a^{N_j} (x_{j,N_j})^{(1+\log(N_j+1))^b}) \right] \prod_{k=1}^{N_j} f_k(x_{j,i}) \right\} \\
 &= \prod_{j=1}^N \left\{ \left[1 - F(a^{N_j} (x_{j,N_j})^{(1+\log(N_j+1))^b}) \right] \right. \\
 &\quad \left. \times \prod_{k=1}^{N_j} \left[a^{k-1} (1 + \log(k))^b (x_{j,i})^{(1+\log(k))^b-1} f(a^{k-1} (x_{j,i})^{(1+\log(k))^b}) \right] \right\}, \quad (10)
 \end{aligned}$$

329 where $\prod_{k=1}^{N_j} \bullet = 1$ for $N_j = 0$, $\boldsymbol{\theta}$ is the vector of the parameters of distribution $F(x)$.

330 Maximising the above likelihood function, we can obtain \hat{a} , \hat{b} , and $\hat{\boldsymbol{\theta}}$, which are the estimates
 331 of the corresponding parameters, respectively. That is

$$(\hat{a}, \hat{b}, \hat{\boldsymbol{\theta}}) = \arg \max_{a, b, \boldsymbol{\theta}} L(a, b, \boldsymbol{\theta}). \quad (11)$$

332 Denote $\boldsymbol{\vartheta} = (a, b, \boldsymbol{\theta})$, where $\vartheta_1 = a$, $\vartheta_2 = b$. The Fisher information matrix $I_{N_0}(\hat{a}, \hat{b}, \hat{\boldsymbol{\theta}})$ can
 333 then be calculated by $I_{N_0}(\hat{a}, \hat{b}, \hat{\boldsymbol{\theta}}) = -\mathbb{E} \left(\frac{\partial^2 \log L(a, b, \boldsymbol{\theta})}{\partial \vartheta_i \partial \vartheta_j} \right) |_{\boldsymbol{\vartheta}=(\hat{a}, \hat{b}, \hat{\boldsymbol{\theta}})}$, which can be used to estimate the
 334 asymptotic variance-covariance matrix of $(\hat{a}, \hat{b}, \hat{\boldsymbol{\theta}})$. In this paper, the Fisher information matrix
 335 will be used to calculate the standard deviations of the estimated parameters.

336 Obviously, there is no general closed-form solution in Eq. (10) for the MLE of \hat{a} , \hat{b} , and $\hat{\boldsymbol{\theta}}$.

337 4 Applications of the DGP

338 In Section 4.1 and Section 4.2, two case studies based on real-world datasets are conducted to
 339 compare the performance of the DGP with $h(k) = (1 + \log(k))^b$, in terms of the corrected Akaike
 340 information criterion, or AICc for short.

- 341 • For the least squares method, model performance is measured by the root mean squared
 342 error (RMSE) = $\sqrt{\frac{1}{N_0} \sum_{k=1}^{N_0} (x_k - \hat{x}_k)^2}$, where \hat{x}_k is the estimate of the x_k .
- 343 • For the maximum likelihood method, model performance is measured with the AICc value,
 344 $N_0 \ln(L) + 2p + \frac{2p(p+1)}{n-p+1}$, where p is the number of parameters in the model and L is the
 345 maximised likelihood. The reader is referred to Burnham and Anderson (2004) for more

346 discussion on the AICc. The value $2p + \frac{2p(p+1)}{n-p+1}$ in the AICc value is a penalty term that is
 347 proportional to the number p of parameters in a model.

348 4.1 Estimating the number of warranty claims

349 Table 1 shows warranty claim data that were collected from a networking card manufacturer.
 350 The manufacturer ships a certain number of items to its retailers on a month basis and then the
 351 warranty agency manages warranty claims. The exact number of the items sold in a shipment
 352 is unknown to the warranty agency. It includes the number of warranty claims in consecutive
 353 12 months on 20 shipments. For example, the underlined number 8 in month 2 and shipment 3
 354 means that 8 2-month-old items that were claimed were from shipment 3 (or they were shipped
 355 in month 3). The last column shows the CV of the warranty claims in each month.

356 Figure 3 illustrates the coefficient of variation (CV) on the warranty claims over the 12 months.
 357 As can be seen, the CV values show an increasing trend. Following Lemma 2, it is more appropriate
 358 to use the DGP to fit the data than the GP.

359 We fit the data with the nonparametric method by solving the problem for the DGP:

$$(\hat{\mu}, \hat{a}, \hat{b}) = \arg \min_{\mu, a, b} \sum_{i=1}^{20} \sum_{k=1}^{12} \left(x_{k,i} - (\mu a^{1-k})^{(1+\log(k))^{-b}} \right)^2 \quad (12)$$

360 where $x_{k,i}$ is the number of warranty claims of k -month-old items that are shipped in month
 361 i . Similarly, the parameters of the GP are estimated. For the DGP model, $\hat{\mu} = 9.19(3.495)$,
 362 $\hat{a} = 1.00232(0.114)$ and $\hat{b} = 0.250(0.739)$ (the values in the brackets are the estimate errors of
 363 the corresponding estimates). The AICc values are $\text{AICc}_{\text{DGP}} = 630.090$ and $\text{AICc}_{\text{GP}} = 630.242$,
 364 which suggests that the DGP outperforms the GP.

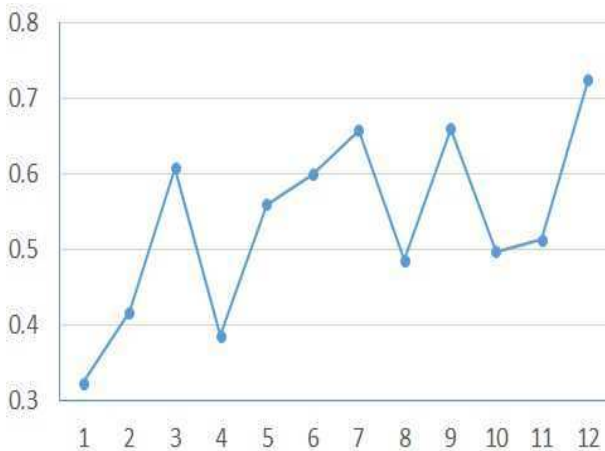


Figure 3: Change of the CVs over 12 months.

Table 1: Time between warranty claims of 22 identical items (unit: day).

Shipments Months	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	CV
1	10	8	13	7	8	16	9	6	7	15	11	9	13	7	9	6	13	10	9	5	0.323
2	7	4	8	6	9	6	1	8	8	9	11	10	10	9	7	8	1	3	9	12	0.417
3	11	7	15	3	4	3	3	13	9	13	6	4	3	5	5	6	3	2	8	5	0.607
4	8	3	12	6	7	6	11	9	9	7	10	7	8	11	6	5	8	5	6	17	0.385
5	4	3	4	2	8	6	7	15	7	9	10	5	2	6	4	14	3	7	10	13	0.559
6	11	8	5	10	4	5	7	8	1	6	11	1	3	4	3	9	4	5	16	13	0.599
7	7	7	22	3	5	14	12	5	4	7	9	4	4	6	17	4	13	3	6	5	0.658
8	11	8	4	5	4	12	6	10	3	4	8	3	5	12	9	10	3	11	4	4	0.486
8	4	3	16	7	1	8	3	6	1	5	6	4	4	12	5	2	4	5	5	6	0.660
10	2	5	9	4	3	10	11	8	1	12	8	6	10	7	2	3	9	10	6	9	0.497
11	5	4	8	4	7	12	1	9	5	8	4	7	3	2	3	5	13	8	7	6	0.513
12	4	5	2	6	1	7	6	10	4	3	12	2	2	17	4	13	6	1	9	5	0.724

365 4.2 Modelling time-between-failure data

366 4.2.1 The datasets

367 Two datasets published in Kumar and Klefsjö (1992); Ascher and Feingold (1984) are used in this
 368 section. Both datasets are collected from the real world and are time-between-failures. The names
 369 and the sample sizes of the datasets are shown in Table 2, where N_0 is the sample size. Kumar and
 370 Klefsjö (1992) develop a power-law-based non-homogeneous Poisson process (PL-NHPP) model
 371 on dataset 1 and Lam (2007) develops geometric process models and PL-NHPP models on dataset
 372 2, which allow us to compare the performance of the DGP with their results.

Table 2: The datasets, including TBF(Time between failures).

No.	Dataset	N_0	References
1	Hydraulic system (LHD3)	25	Kumar and Klefsjö (1992)
2	Propulsion diesel engine failure data	71	Ascher and Feingold (1984)

373 In the following, we compare the performance of the models that are estimated with the least
 374 squares and the maximum likelihood estimation methods, respectively.

375 4.2.2 Model comparison

376 Definition 3 assumes that $\{X_k, k = 1, 2, \dots\}$ in the DGP are independent. We therefore use the
 377 Box-Ljung test to check the hypothesis that a given series of data is independent (Ljung and
 378 Box, 1978). Applying the Box-Ljung test on datasets 1 and 2, the result fails to reject the null
 379 hypothesis that observations in datasets 1 and 2 are independent at the 5% level of significance.

380 On the two datasets listed in Table 2, we use both the least squares method and the maximum
 381 likelihood method to estimate the parameters and then compare the performance of the DGP
 382 with the GP.

383 With the least squares method, both the DGP and the GP are estimated and their RMSE
 384 values are denoted by RMSE_{DGP} and RMSE_{GP} , respectively. The estimated parameters and their
 385 standard deviations (which are shown in brackets under the estimated parameters), and the RMSE
 386 values of both the DGP and the GP are shown in Table 3. As can be seen, the RMSE values
 387 of the DGP on each dataset is smaller than the RMSE values of the GP, based on which one can
 388 conclude the DGP outperforms the GP on both datasets.

Table 3: Comparison of the performance of the GP and the DGP based on the least squares method.

No.	Parameters of the DGP			Parameters of the GP		RMSE_{DGP}	RMSE_{GP}
	\hat{a}	\hat{b}	$\hat{\mu}$	\hat{a}	$\hat{\mu}$		
1	0.944 (0.0559)	0.499 (0.174)	531.406 (109.390)	1.0382 (0.0315)	209.841 (67.652)	<u>111.729</u>	144.431
2	0.909 (0.0607)	0.488 (0.280)	147.624 (62.664)	0.972 (0.0181)	56.702 (20.486)	<u>65.670</u>	69.810

389 Suppose $F(t) = 1 - e^{-(\frac{t}{\theta_1})^{\theta_2}}$. With the maximum likelihood method, we use the DGP, the GP,
 390 the PL-NHPP to fit the two datasets, and denote their corresponding AICc values as AICc_{DGP} ,
 391 AICc_{GP} , and AICc_{PL} , respectively. The number of the parameters (i.e., a, b, θ_1, θ_2) in the DGP
 392 and the number of the parameters (i.e., a, θ_1, θ_2) in the GP are 4 and 3, respectively, i.e., $p = 4$ for
 393 the DGP and $p = 3$ for the GP. The number of the parameters in the PL-NHPP is 2 (i.e., $p = 2$).
 394 The results are shown in Table 4. The estimated parameters and their standard deviations (which
 395 are shown in brackets under the estimated parameters) of the DGP are also given in the table.
 396 On the rest comparison, the AICc values of the DGP are the smallest.

397 In addition to the independence test conducted before, to test the assumption that the DGP
 398 can model datasets 1 and 2, we use the Cramér-von-Mises test to test the null hypotheses that
 399 $\{\hat{a}^{k-1} X_k^{(1+\log(k))^{\hat{b}}}, k = 1, \dots, N_0\}$ on datasets 1 and 2 follow the Weibull distribution, respectively.
 400 We conduct the hypothesis testing with a R-package *EWGoF* (Krit, 2014). The results fail to
 401 reject the null hypotheses at the 5% level of significance.

402 4.3 Comparison between different forms of $h(k)$

403 In the preceding sections, we set $h(k) = (1 + \log(k))^b$ in Definition 3. By setting other forms of
 404 $h(k)$ such as $h(k) = b^{k-1}$, $h(k) = b^{\log(k)}$, or $h(k) = 1 + b \log(k)$, one can define other forms of

Table 4: Comparison of the performance of the GP and the DGP based on the maximum likelihood method.

No.	Estimated Parameters of the DGP				Estimated Parameters of the GP			AIC _C DGP	AIC _C GP	AIC _C PL
	\hat{a}	\hat{b}	$\hat{\theta}_1$	$\hat{\theta}_2$	\hat{a}	$\hat{\theta}_1$	$\hat{\theta}_2$			
1	0.884 (0.0938)	0.638 (0.352)	449.165 (337.92)	0.789 (0.227)	1.0147 (0.0230)	168.807 (58.139)	1.0287 (0.159)	<u>301.376</u>	304.182	311.851
2	0.899 (0.0714)	0.502 (0.349)	147.636 (103.569)	0.964 (0.281)	0.983 (0.0151)	73.070 (19.461)	1.295 (0.182)	<u>318.030</u>	319.445	323.094

405 the DGP. To differentiate them, we refer to the processes with $h(k) = (1 + \log(k))^b$, $h(k) = b^{k-1}$,
406 $h(k) = b^{\log(k)}$ and $h(k) = 1 + b \log(k)$ as DGP_{log1}, DGP_{exp}, DGP_{log2}, and DGP_{log3}, respectively.
407 Similarly, one can estimate parameters a and b of the DGP_{exp}, DGP_{log2}, and DGP_{log3} with either
408 the least squares or the maximum likelihood estimation method. We have compared the AICc
409 values of the DGP_{log1} with the AICc values of the rest three models on the ten datasets and found
410 that the AICc value of the DGP_{log} on each dataset is smaller than those of the other three models,
411 respectively, which implies that the DGP with $h(k) = (1 + \log(k))^b$ outperforms. That is the
412 reason that we investigated the GDP with $h(k) = (1 + \log(k))^b$ in this paper.

413 5 Conclusion and future work

414 This paper proposed a new stochastic process, the doubly geometric process (DGP), which extends
415 the geometric process (GP). The DGP can overcome three limitations inherent in the GP. The
416 paper discussed probabilistic properties of the DGP with $h(k) = (1 + \log(k))^b$, compared the
417 mean cumulative functions between the DGP and other processes, and then proposed methods of
418 estimation of the parameters in the DGP.

419 The paper also applied the DGP to fit two inter-arrival time datasets collected from the real
420 world and then compared its performance with the performance of other models. It is found that
421 the DGP outperforms the other models on those datasets. This has practical implications for
422 lifecycle costing, for example.

423 As the DGP is a new model, there are plenty of questions waiting for answers. Those questions
424 include, for example, what are the differences between the DGP and the other models in terms
425 of the application of the DGP in reliability mathematics? Before we fit a given dataset with the
426 DGP, how can we test if the dataset agrees with the DGP? To answer those questions will be our
427 future work.

428 Acknowledgements

429 The authors are indebted to the reviewers and the editor for their comments.

430 Appendix

431 Proof of Proposition 1.

432 Let $u(x)$ denote a non-decreasing function. With Lemma 1, X_k is stochastically increasing if
 433 $\mathbb{E}[u(X_{k+1})] - \mathbb{E}[u(X_k)] > 0$ and let $y = a^{k-1}x^{(1+\log(k))^b}$, then we can obtain

$$\begin{aligned}
 \mathbb{E}[u(X_{k+1})] - \mathbb{E}[u(X_k)] &= \int_0^{+\infty} u(x) dF(a^k x^{(1+\log(k+1))^b}) - \int_0^{+\infty} u(x) dF(a^{k-1} x^{(1+\log(k))^b}) \\
 &= \int_0^{+\infty} \left(u(a^{-k(1+\log(k+1))^{-b}} y^{(1+\log(k+1))^{-b}}) \right. \\
 &\quad \left. - u(a^{(1-k)(1+\log(k))^{-b}} y^{(1+\log(k))^{-b}}) \right) dF(y) \\
 &> 0.
 \end{aligned} \tag{13}$$

434 Let y represent all the possible values on X_k (for $k = 1, 2, \dots$). Hence, $\mathbb{E}[u(X_{k+1})] - \mathbb{E}[u(X_k)] >$
 435 0 if $u(a^{-k(1+\log(k+1))^{-b}} y^{(1+\log(k+1))^{-b}}) - u(a^{(1-k)(1+\log(k))^{-b}} y^{(1+\log(k))^{-b}}) > 0$. As $u(\cdot)$ is a non-decreasing
 436 function, $u(a^{-k(1+\log(k+1))^{-b}} y^{(1+\log(k+1))^{-b}}) - u(a^{(1-k)(1+\log(k))^{-b}} y^{(1+\log(k))^{-b}}) > 0$ iff

$$\frac{a^{-k(1+\log(k+1))^{-b}} y^{(1+\log(k+1))^{-b}}}{a^{(1-k)(1+\log(k))^{-b}} y^{(1+\log(k))^{-b}}} = a^{-k(1+\log(k+1))^{-b} + (k-1)(1+\log(k))^{-b}} y^{(1+\log(k+1))^{-b} - (1+\log(k))^{-b}} > 1. \tag{14}$$

437 From equality (14), we have

438 • if $b < 0$, then $-k(1+\log(k+1))^{-b} + (k-1)(1+\log(k))^{-b} < 0$ and $(1+\log(k+1))^{-b} - (1+$
 439 $\log(k))^{-b} > 0$. That implies,

440 (i) if $0 < a < 1$, $P(X_1 > 1) = 1$, and $b < 0$, the inequality (14) holds. Then $\{X_k, k =$
 441 $2, 3, \dots\}$ is stochastically increasing, and

442 (ii) if $a > 1$, $P(0 < X_1 < 1) = 1$, and $b < 0$, the greater-than sign in the inequality (14)
 443 should be changed to the smaller-than sign. Then $\{X_k, k = 2, 3, \dots\}$ is stochastically
 444 decreasing.

445 • On the other hand, if $b > 0$, $(1+\log(k+1))^{-b} - (1+\log(k))^{-b} < 0$. But if $b > 0$,
 446 $-k(1+\log(k+1))^{-b} + (k-1)(1+\log(k))^{-b}$ may be positive or negative, which can be

447 equivalently expressed as

$$\frac{k-1}{k} < \left(\frac{1+\log(k)}{1+\log(k+1)} \right)^b \quad (15)$$

448 may hold and

$$\frac{k-1}{k} > \left(\frac{1+\log(k)}{1+\log(k+1)} \right)^b \quad (16)$$

449 may hold as well.

450 If b is small ($b = 1$, for example), then inequality (15) holds. If b is large, then inequality
 451 (16) holds (this is because $\left(\frac{1+\log(k)}{1+\log(k+1)}\right)^b \rightarrow 0$ for $b \rightarrow \infty$). Since $\left(\frac{1+\log(k)}{1+\log(k+1)}\right)^b$ is a decreasing
 452 function of b , we can find a value of b , denoted as b_0 , which satisfies: if $0 < b < b_0$, then
 453 inequality (15) always holds for any k . Taking the logarithm on both sides of inequality (15)
 454 and then dividing both sides by $\log(1+\log(k)) - \log(1+\log(k+1))$, then inequality (15)
 455 becomes

$$\frac{\log(k-1) - \log(k)}{\log(1+\log(k)) - \log(1+\log(k+1))} < b. \quad (17)$$

456 Let $b_0 = \min_k \left\{ \frac{\log(k-1) - \log(k)}{\log(1+\log(k)) - \log(1+\log(k+1))}, k = 2, 3, \dots \right\}$. One can obtain $b_0 = 4.898226$. If
 457 $0 < b < b_0$, then $-k(1+\log(k+1))^{-b} + (k-1)(1+\log(k))^{-b} < 0$ and $(1+\log(k+1))^{-b} -$
 458 $(1+\log(k))^{-b} < 0$, the inequality (14) holds. That implies

459 (iii) if $0 < a < 1$, $0 < b < b_0$, and $P(0 < X_1 < 1) = 1$, then $\{X_k, k = 2, 3, \dots\}$ is
 460 stochastically increasing, and

461 (iv) if $a > 1$, $0 < b < b_0$, and $P(X_1 > 1) = 1$, then $\{X_k, k = 2, 3, \dots\}$ is stochastically
 462 decreasing.

463 This completes the proof. □

464 **Proof of Proposition 2.** Denote

$$U = a^{-k(1+\log(k+1))^{-b} + (k-1)(1+\log(k))^{-b}} y^{(1+\log(k+1))^{-b} - (1+\log(k))^{-b}}. \quad (18)$$

465 Similar to the proof of Proposition 1, if $\log(U) = (1+\log(k+1))^{-b}(\log(y) - k\log(a)) + (1+$
 466 $\log(k))^{-b}((k-1)\log(a) - \log(y))$ varies between negative and positive values, the left hand side
 467 of Eq. (14) changes between $(0, 1)$ and $(1, +\infty)$. That is, the process $\{X_k, k = 1, 2, \dots\}$ is
 468 stochastically non-monotonous.

469 This completes the proof. □

470 **Proof of Proposition 3.**

$$f_k(t) = a^{k-1}(1 + \log(k))^b t^{(1+\log(k))^b-1} f(a^{k-1}t^{(1+\log(k))^b}). \quad (19)$$

471 Denote $r(t) = \frac{f(t)}{\bar{F}(t)}$. We have,

$$\begin{aligned} r_k(t) &= \frac{f_k(t)}{\bar{F}_k(t)} \\ &= \frac{a^{k-1}(1 + \log(k))^b t^{(1+\log(k))^b-1} f(a^{k-1}t^{(1+\log(k))^b})}{\bar{F}(a^{k-1}t^{(1+\log(k))^b})} \\ &= a^{k-1}(1 + \log(k))^b t^{(1+\log(k))^b-1} r(a^{k-1}t^{(1+\log(k))^b}), \end{aligned} \quad (20)$$

472 If $b > 0$, then $t^{(1+\log(k))^b}$ is increasing in t . Since $r(t)$ is an increasing function in t , $r_k(t)$ is increasing
473 in t . □

474 **Proof of Proposition 4.**

475 If $\{X_k, k = 1, 2, \dots\}$ is stochastically non-decreasing, for every real numbers r_0 and r_1 , we have
476 $P(X_k > r_0) \geq P(X_1 > r_0)$, or $P(X_k < r_0) \leq P(X_1 < r_0)$. Then we have $P(\sum_{i=1}^n X_k < r_1) \leq$
477 $P(\sum_{i=1}^n Y_i < r_1)$, which implies that inequality $m(t) \leq m_1(t)$ holds.

478 Similarly, we can prove that $m(t) \geq m_1(t)$ if $\{X_k, k = 1, 2, \dots\}$ is stochastically non-decreasing.

479 According to Bennett's inequality (Bennett, 1962) below,

$$P\left(\sum_{k=1}^n (X_k - \mathbb{E}[X_k]) > t\right) \leq \exp\left(-\frac{n\sigma^2}{c^2} H\left(\frac{ct}{n\sigma^2}\right)\right), \quad (21)$$

480 where $H(u) = (1 + u) \ln(1 + u) - u$, we can obtain

$$P(S_n < t) \geq 1 - \exp\left(-\frac{n\sigma^2}{c^2} H\left(\frac{ct - c\Lambda_n}{n\sigma^2}\right)\right). \quad (22)$$

481 Hence,

$$m(t) \geq \sum_{n=1}^{\infty} \left[1 - \exp\left(-\frac{n\sigma^2}{c^2} H\left(\frac{ct - c\Lambda_n}{n\sigma^2}\right)\right)\right]. \quad (23)$$

482 This completes the proof. □

483 **Proof of Proposition 5.**

484 In the following, we prove (i).

485 According to Definition 1, if $X_k^g <_{st} X_k^d$, we have $m^g(t) > m^d(t)$. For a given non-decreasing

486 function $u(x)$, with Lemma 1, $X_k^g <_{st} X_k^d$ if $\mathbb{E}[u(X_k^g)] < \mathbb{E}[u(X_k^d)]$. Since

$$\begin{aligned} \mathbb{E}[u(X_k^g)] - \mathbb{E}[u(X_k^d)] &= \int_0^{+\infty} u(x) dF(a^{k-1}x) - \int_0^{+\infty} u(x) dF(a^{k-1}x^{(1+\log(k))^b}) \\ &= \int_0^{+\infty} \left(u(a^{(1-k)}y) - u(a^{(1-k)(1+\log(k))^{-b}}y^{(1+\log(k))^{-b}}) \right) dF(y), \end{aligned} \quad (24)$$

487 $\mathbb{E}[u(X_k^g)] < \mathbb{E}[u(X_k^d)]$ if $u(a^{(1-k)}y) < u(a^{(1-k)(1+\log(k))^{-b}}y^{(1+\log(k))^{-b}})$. As $u(\cdot)$ is a non-decreasing
488 function, $u(a^{(1-k)}y) < u(a^{(1-k)(1+\log(k))^{-b}}y^{(1+\log(k))^{-b}})$ holds if

$$\frac{a^{(1-k)}y}{a^{(1-k)(1+\log(k))^{-b}}y^{(1+\log(k))^{-b}}} = a^{(1-k)(1-(1+\log(k))^{-b})}y^{1-(1+\log(k))^{-b}} < 1. \quad (25)$$

489 Inequality (25) holds if either of the following conditions is true,

- 490 • if $0 < a < 1$, $b < 0$ and $P(X_1 > 1) = 1$,
- 491 • if $a > 1$, $b > 0$ and $P(0 < X_1 < 1) = 1$.

492 Similarly, the other bullet (ii) can be established.

493 This completes the proof. □

494 **Proof of Proposition 6.**

495 Similar to the proof for Proposition 5, Proposition 6 can be established. □

496 **Proof of Proposition 7.**

- 497 • For any given $M > 0$,

$$\lim_{k \rightarrow \infty} P(|X_k| < M) = \lim_{k \rightarrow \infty} P(0 < X_k < M) = \lim_{k \rightarrow \infty} P(X_1 < a^{k-1}M^{(1+\log(k))^b}). \quad (26)$$

498 If $0 < a < 1$, then $\lim_{k \rightarrow \infty} a^{k-1}M^{(1+\log(k))^b} = 0$. Since X_1 is non-negative, $\lim_{k \rightarrow \infty} P(X_1 <$
499 $a^{k-1}M^{(1+\log(k))^b}) = 0$, or $\lim_{k \rightarrow \infty} P(|X_k| < M) = 0$. That is, X_k converges to infinity in
500 probability as $k \rightarrow \infty$.

- 501 • For any given $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} P(|X_k| > \varepsilon) = \lim_{k \rightarrow \infty} P(X_k > \varepsilon) = 1 - \lim_{k \rightarrow \infty} P(X_1 \leq \varepsilon) = 1 - \lim_{k \rightarrow \infty} P(X_k \leq a^{k-1}\varepsilon^{(1+\log(k))^b}). \quad (27)$$

502 If $a > 1$, then $\lim_{k \rightarrow \infty} a^{k-1}\varepsilon^{(1+\log(k))^b} = \infty$. That implies $\lim_{k \rightarrow \infty} P(X_1 \leq a^{k-1}\varepsilon^{(1+\log(k))^b}) = 1$, or
503 $\lim_{k \rightarrow \infty} P(|X_k| > \varepsilon) = 0$. That is, X_k converges to zero in probability as $k \rightarrow \infty$.

504 This completes the proof. □

505 Acknowledgements

506 The author is indebted to the reviewers and the editor for their comments.

507 References

508 Ascher H and Feingold H (1984). *Repairable systems reliability: Modelling, inference, misconcep-*
509 *tions and their causes*. Marcel Dekker: New York.

510 Bennett G (1962). Probability inequalities for the sum of independent random variables. *Journal*
511 *of the American Statistical Association* **57** (297): 33–45.

512 Braun W J, Li W and Zhao Y Q (2005). Properties of the geometric and related processes. *Naval*
513 *Research Logistics* **52** (7): 607–616.

514 Chan J S, Choy S B and Lam C P (2014). Modeling electricity price using a threshold condi-
515 tional autoregressive geometric process jump model. *Communications in Statistics-Theory and*
516 *Methods* **43** (10-12): 2505–2515.

517 Chan J S, Yu P L, Lam Y and Ho AP (2006). Modelling sars data using threshold geometric
518 process. *Statistics in Medicine* **25** (11): 1826–1839.

519 Chukova S, Hayakawa Y and Dohi T (2005). Warranty cost analysis: quasi-renewal inter-repair
520 times. *International Journal of Quality & Reliability Management* **22** (7): 687–698.

521 Burnham K P and Anderson D R (2004). Multimodel inference understanding AIC and BIC in
522 model selection. *Sociological methods & research* **33** (2): 261-304.

523 Doyen L and Gaudoin O (2004). Classes of imperfect repair models based on reduction of failure
524 intensity or virtual age. *Reliability Engineering & System Safety* **84** (1): 45–56.

525 Finkelstein M (1993). A scale model of general repair. *Microelectronics Reliability* **33** (1): 41–44.

526 Jain M and Gupta R (2013). Optimal replacement policy for a repairable system with multiple
527 vacations and imperfect fault coverage. *Computers & Industrial Engineering* **66** (4): 710–719.

528 Kijima M (1989). Some results for repairable systems with general repair. *Journal of Applied*
529 *Probability* **26** (1): 89–102.

530 Krit M (2014). *Goodness-of-fit tests for the Exponential and two-parameter Weibull distributions*
531 *(Package: EWGoF)*, R Foundation for Statistical Computing, Vienna, Austria.

- 532 Kumar U and Klefsjö B (1992). Reliability analysis of hydraulic systems of LHD machines using
533 the power law process model. *Reliability Engineering & System Safety* **35** (3): 217–224.
- 534 Lam Y (1988). Geometric processes and replacement problem. *Acta Mathematicae Applicatae*
535 *Sinica* **4** (4): 366–377.
- 536 Lam Y (1992). Nonparametric inference for geometric processes. *Communications in Statistics-*
537 *Theory and Methods* **21** (7): 2083–2105.
- 538 Lam Y (2007). *The geometric process and its applications*, World Scientific: Singapore.
- 539 Liu Y and Huang H Z (2010). Optimal replacement policy for multi-state system under imperfect
540 maintenance. *IEEE Transactions on Reliability* **59** (3): 483–495.
- 541 Ljung G M and Box G E (1978). On a measure of lack of fit in time series models. *Biometrika*
542 **65** (2): 297–303.
- 543 Ross S M (1996). *Stochastic processes*. John Wiley & Sons: New York.
- 544 Seber G and Wild C (2003) *Nonlinear regression*. Wiley & Sons: New York.
- 545 Wang H and Pham H (1996). A quasi renewal process and its applications in imperfect mainte-
546 nance. *International Journal of Systems Science* **27** (10): 1055–1062.
- 547 Wang W (2011). Maintenance models based on the np control charts with respect to the sampling
548 interval. *Journal of the Operational Research Society* **62**(1): 124-133.
- 549 Wu S and Clements-Croome D (2006). A novel repair model for imperfect maintenance. *IMA*
550 *Journal of Management Mathematics* **17** (3): 235–243.
- 551 Wu S and Scarf P (2015). Decline and repair, and covariate effects. *European Journal of Opera-*
552 *tional Research* **244** (1): 219–226.
- 553 Yuan L and Meng X Y (2011). Reliability analysis of a warm standby repairable system with
554 priority in use. *Applied Mathematical Modelling* **35** (9): 4295–4303.
- 555 Zhang Y L, Yam R C M and Zuo M J (2002). Optimal replacement policy for a multistate
556 repairable system. *Journal of the Operational Research Society* **53**(3): 336-341.
- 557 Zhang M, Xie M and Gaudoin O (2013). A bivariate maintenance policy for multi-state repairable
558 systems with monotone process. *IEEE Transactions on Reliability* **62** (4): 876–886.

559 Zuo M J, Jiang R and Yam R (1999). Approaches for reliability modeling of continuous-state
560 devices. *IEEE Transactions on Reliability* **48** (1): 9–18.