Extended Geometric Processes for Application to Reliability

Sophie MERCIER[a*], Laurent BORDES[a]
[a]Laboratoire de Mathématiques et de leurs Applications (UMR CNRS 5142), Pau, France

Abstract: Renewal processes have been widely used in reliability, to describe successive failure times of systems submitted to perfect and instantaneous maintenance actions. In case of imperfect maintenance, different models have been developed to take this feature into account, among which geometric processes introduced by [Y. Lam, *The Geometric Process and its Applications.* World Scientific, 2007]. In such a model, successive lifetimes are independent and identically distributed up to a multiplicative scale parameter $a > 0$, in a geometric fashion. A drawback in Lam's setting is the fast increase or decrease of the successive periods, induced by the geometric progression. We here envision a more flexible progression, where the multiplicative scaling factor is not necessarily a geometric progression any more. The corresponding counting process is here named Extended Geometric Process (EGP).

As a first step in the study of an EGP, we consider its semiparametric estimation based on the observation of the n first gap times. We start with the estimation of the Euclidean parameter a following the regression method proposed by Lam. We next proceed to the estimation of the unknown distribution of the underlying renewal process. Several consistency results, including convergence rates, are obtained.

We next turn to applications of EGPs to reliability, where successive arrival times stand for failure (and instantaneous maintenance) times. A first quantity of interest is the pseudo-renewal function associated to an EGP, which is proved to fulfill a pseudo-renewal equation. When the system is deteriorating (case $a < 1$), a preventive renewal policy is proposed: as soon as a lifetime is observed to be too short, under a predefined threshold, the system is considered as too deteriorated and replaced by a new one. This renewal policy is assessed through a cost function, on an infinite horizon time. Numerical experiments illustrate the study.

Keywords: Imperfect maintenance, Renewal processes, Semiparametric estimation.

1. INTRODUCTION

From several years, many attention has been paid to the modeling of recurrent event data. Application fields are various and include reliability, medicine, insurance, etc., see [7] for an overview of models and their applications. In reliability, the events of interest typically are successive failures of a system submitted to instantaneous repair. In case of perfect repairs (As Good As New repairs), the underlying process describing the system evolution is a renewal process, which has been widely used in reliability, see [2]. In case of imperfect repair, the successive inter-failure times may however become shorter and shorter, leading to some (stochastically) decreasing sequence of lifetimes. In case of improving systems such as software releases e.g., inter-failure times may also be increasing. This has lead to the development of different models taking into account such features, among which geometric processes introduced by [14]. In such a model, successive lifetimes $X_1, X_2, \ldots, X_n, \ldots$ are independent with identical distributions up to a multiplicative scale parameter: $X_n = a^{n-1} Y_n$ where $(Y_n)_{n\geq 1}$ is a sequence of independent and identically distributed random variables (the interarrival times of a renewal process). According to $a \geq 1$ or $0 < a < 1$, the sequence $(X_n)_{n\geq 1}$ may be (stochastically) non-decreasing or non-increasing. A drawback in Lam's setting is the fast increase or decrease of the successive periods, induced by a geometric progression. We here envision a more general scaling factor, where X_n is of the shape $X_n = a^{b_n} Y_n$ and $(b_n)_{n\geq 1}$ stands for a non decreasing sequence. This allows for more flexibility in the progression of the X_n's. The corresponding counting process is named Extended Geometric Process (EGP) in the sequel.

As a first step in the study of an EGP, we consider its semiparametric estimation based on the observation of the n first gap times. The sequence $(b_n)_{n\geq 1}$ is assumed to be known and we start with the estimation of the Euclidean parameter a. Following the regression method proposed by [12], several consistency results are obtained for the estimate $\hat{a}$, including convergence rates. We next proceed to the estimation of the unknown distribution of the underlying renewal process. The estimation method relies on a pseudo version $(\tilde{Y}_n)_{n\geq 1}$ of the points $(Y_n)_{n\geq 1}$ of the underlying renewal process, which is obtained by setting $\tilde{Y}_n = \hat{a}^{-b_n} X_n$. Again, several convergence results are obtained, such as strong uniform consistency.

We next turn to applications of EGPs to reliability, with the previous interpretation of arrivals of an EGP as successive failure times. A first quantity of interest then is the mean number of instantaneous repairs on some time interval $[0, t]$, which corresponds to the pseudo-renewal function associated to an EGP, seen as some pseudo-renewal process. The pseudo-renewal function is proved to fulfill a pseudo-renewal equation, and tools are provided for its numerical solving. In case $a < 1$, the system is aging and requires some action to prevent successive lifetimes to become shorter and shorter. In that case, a replacement policy is proposed: as soon as a lifetime is observed to be too short - bellow a predefined threshold -, the system is considered as too degraded and it is replaced by a new one. In case $a \geq 1$, the system is improving at each corrective action and no replacement policy is required. In case $a < 1$, the replacement policy is assessed through a cost function, which is provided in full form. The replacement policy proposed here is an alternative to the one considered by [14], where the failure times are modelled by a geometric process and where the system is replaced by a new one once it has been repaired N times (with N fixed). Non negligible repair times are also considered by [14] (modelled by another geometric process), which we do not envision here.

This paper is organized as follows: Section 2 is devoted to the semiparametric estimation of an EGP and applications to reliability are developed in Section 3. Both of these sections include numerical experiments, to illustrate the results. Concluding remarks end this paper in Section 4.

2. ESTIMATION OF EXTENDED GEOMETRIC PROCESSES

2.1. The model

Let $(T_n)_{n\geq 0}$ stand for the successive failure times of a system, with $0 = T_0 < T_1 < \cdots < T_n < \cdots$ We set $X_n = T_n - T_{n-1}$ for $n \geq 1$ and we assume that $(X_n)_{n\geq 1}$ satisfies $X_n = a^{b_n} Y_n$, where $(Y_n)_{n\geq 1}$ are the interarrival times of a renewal process (RP), $a \in (0, +\infty)$ and $(b_n)_{n\geq 1}$ is a non decreasing sequence of non negative real numbers such that $b_1 = 0$ and b_n tends to infinity when n goes to infinity. To prevent trivialities, Y_1 is assumed to be non zero: $\mathbb{P}(Y_1 > 0) > 0$. This setting enlarges classical geometric processes, for which $b_n = n - 1$ and $X_n = a^{n-1} Y_n$ (all $n \geq 1$), see [14].

Though parameter b_n could be only known up to an euclidean parameter, the sequence $(b_n)_{n\geq 1}$ is here assumed to be fully known. Unknown parameters hence are $a \in (0, +\infty)$ and F the cumulative distribution function (c.d.f.) of the underlying RP. As a consequence, it is a semiparametric model.

2.2. The estimation method

Assuming that $T_1 < \cdots < T_n$ are observed, we consider the problem of estimating a and F. As for the euclidian parameter a, we follow the same estimation method considered by Lam in a series of papers, see [12–14], which is based on a classical regression: writing $Z_n = \log X_n$ for $n \geq 1$, we have $Z_n = b_n \beta + \mu + e_n$ where $\beta = \log a$, $\mu = \mathbb{E}[\log Y_1]$ and $e_n = \log Y_n - \mu$ are independent and identically distributed (i.i.d.) centered errors. Parameters μ and β are next estimated by a least square method:

$$(\hat{\mu}_n, \hat{\beta}_n) = \arg\min_{\mu,\beta} \sum_{k=1}^{n} (Z_k - \beta b_k - \mu)^2 ,$$

which provides

$$\hat{\beta}_n = \frac{n^{-1} \sum_{k=1}^n b_k Z_k - n^{-2} \sum_{k=1}^n Z_k \sum_{i=1}^n b_i}{n^{-1} \sum_{k=1}^n b_k^2 - \left(n^{-1} \sum_{k=1}^n b_k\right)^2} \quad \text{and} \quad \hat{\mu}_n = \bar{Z}_n - \hat{\beta}_n \bar{b}_n,$$

where $\bar{b}_n = (b_1 + \cdots + b_n)/n$ and $\bar{Z}_n = (Z_1 + \cdots + Z_n)/n$. The euclidian parameter a is then estimated by $\hat{a}_n = \exp(\hat{\beta}_n)$.

Starting from $\hat{a}_n$, we next construct a pseudo version $(\tilde{Y}_n)_{n\geq 1}$ of the inter-arrival times $(Y_n)_{n\geq 1}$ by setting $\tilde{Y}_n = \hat{a}_n^{-b_n} X_n$. Then, we can expect to estimate F by the empirical c.d.f. $\hat{F}_n$ defined by

$$\hat{F}_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{\tilde{Y}_k \leq x\}} \tag{1}$$

for all $x \in \mathbb{R}^+$, where $\mathbf{1}_{\{\cdot\}}$ denotes the set indicator function.

Assuming that $\mathbb{E}(\log^2(Y_n))$ exists, let us define $\mathbb{V}\mathrm{ar}(e_n) = \sigma^2$. Easy computations then provide $\mathbb{E}(\hat{\beta}_n) = \beta$ and $\mathbb{V}\mathrm{ar}(\hat{\beta}_n) = \dfrac{\sigma^2}{n\alpha_n^2}$, where $\alpha_n^2 = \frac{1}{n}\sum_{k=1}^n b_k^2 - \left(\frac{1}{n}\sum_{k=1}^n b_k\right)^2$. If a central limit theorem holds, its formulation can consequently only be $\theta_n(\hat{\beta}_n - \beta) \xrightarrow{d} \mathcal{N}(0,\sigma^2)$, where $\xrightarrow{d}$ stands for the convergence in distribution and $\theta_n = \alpha_n\sqrt{n}$. The only possible order for the convergence rate of $\hat{\beta}_n$ towards β hence is θ_n.

2.3. Asymptotics

We first provide asymptotic results on the euclidian estimates $\hat{\beta}_n$ and $\hat{a}_n$ when $n \to +\infty$. Such results are based on strong laws of large numbers for weighted sums of i.i.d. random variables, as provided by [8, 3, 4], and on the δ-method theorem (e.g. [19]). Because of the reduced size of the present paper, details are omitted, which may be found in [5], as for all proofs, later on.

Proposition 1 *Suppose that $\mathbb{E}(Z_1^2) < +\infty$. Then:*

1. *We have $\alpha_n(\hat{\beta}_n - \beta) \xrightarrow{a.s.} 0$ (**Strong consistency**).*
2. *We have $\limsup_{n\to+\infty} \frac{\sqrt{n}\alpha_n^2}{b_n\sqrt{\log n}}|\hat{\beta}_n - \beta| \leq 2\sqrt{2}\sigma \quad$ a.s. (**Law of Iterated Logarithm**),*
3. *If, addingly, $\theta_n/b_n \to +\infty$ (where $\theta_n = \alpha_n\sqrt{n}$), then $\theta_n(\hat{\beta}_n - \beta) \xrightarrow{d} \mathcal{N}(0,\sigma^2)$ and $\theta_n(\hat{a}_n - a) \xrightarrow{d} \mathcal{N}(0, a^2\sigma^2)$ (**Central Limit Theorem**).*

It is straightforward to verify that $\alpha_{n+1}^2 = \alpha_n^2 + \frac{n}{n+1}\left(b_{n+1} - \bar{b}_n\right)^2$, which implies that $(\alpha_n)_{n\geq 1}$ is a non decreasing sequence. This monotonicity plus the previous consistency result imply that $\hat{\beta}_n \xrightarrow{a.s.} \beta$.
It is known from standard results on linear regression that

$$\hat{\sigma}_n^2 = \frac{1}{n-2}\sum_{k=1}^{n}\left(Z_k - \hat{\beta}_n b_k - \hat{\mu}_n\right)^2$$

is an unbiased consistent estimator of σ^2. Then, the asymptotic variance of $\theta_n(\hat{a}_n - a)$ is consistently estimated by $\hat{a}_n^2\hat{\sigma}_n^2$.

Example 2 *If $b_n = (n-1)^\alpha$ with $\alpha > 0$, we have*

$$\theta_n \overset{+\infty}{\sim} \frac{\alpha n^{\alpha+1/2}}{(\alpha+1)\sqrt{2\alpha+1}}$$

and the condition $\theta_n/b_n \to +\infty$ is true.

In case $b_n = n-1$, we get $n^{3/2}(\hat{\beta}_n - \beta) \xrightarrow{d} \mathcal{N}(0, 12\sigma^2)$, which is consistent with the central limit result from [13].
If $b_n = \log n$, we have $\theta_n \overset{+\infty}{\sim} \sqrt{n}$ and the condition $\theta_n/b_n \to +\infty$ is true.

The c.d.f. F is now estimated by the empirical c.d.f., as defined by (1). The following uniform consistency result is based on the Glivenko-Cantelli theorem.

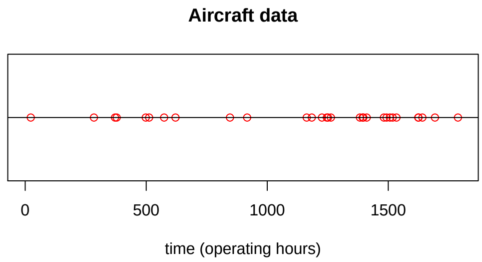


Figure 1: Cumulated times between successive failures of the air-conditioning equipment.

Proposition 3 (Uniform Strong Consistency) *Assume that $Z_1 = \ln(X_1)$ admits a bounded density g with respect to Lebesgue measure, that Z_1 has a second order moment and that*

$$\limsup_{n\to+\infty} \frac{b_n^2\sqrt{\log n}}{\sqrt{n}\alpha_n^2} = 0. \tag{2}$$

Then $\|\hat{F}_n - F\|_\infty$ converges to 0 almost surely as n tends to infinity.

The boundedness condition on g is satisfied whenever f belongs to the usual parametric families (Weibull, Gamma, log-normal, etc.). Condition (2) on the sequence $\left(\frac{b_n^2\sqrt{\log n}}{\sqrt{n}\alpha_n^2}\right)_{n\geq 1}$ is satisfied for many non decreasing sequences $(b_n)_{n\geq 1}$ tending to infinity. For example:

- Condition (2) is true as soon as $b_n^2\sqrt{\log n}/\sqrt{n} \to 0$, due to decreasingness of $\left(\alpha_n^2\right)_{n\in\mathbb{N}}$. As a special case, Condition (2) is true for $b_n = (\log n)^\alpha$ with $\alpha > 0$.
- if $b_n = (n-1)^\alpha$ with $\alpha > 0$ then

$$\frac{b_n^2\sqrt{\log n}}{\sqrt{n}\alpha_n^2} \overset{+\infty}{\sim} \frac{(\alpha+1)^2(2\alpha+1)\sqrt{\log n}}{\alpha^2\sqrt{n}} \to 0$$

(see Example 2) and Condition (2) is satisfied.

2.4. Numerical experiments

2.4.1. Illustrative example

We here consider a data set of size $n = 29$, for successive failures of an air-conditioning equipment of a Boeing 720 aircraft. This data set is taken among 13 ones corresponding to 13 different aircrafts, that were studied in [18] and are available in [15]. Figure 1 shows the cumulated times (operating hours) between the successive failures. Table 1 summarizes the results obtained for the estimation of parameter a for various sequences $(b_n)_{n\geq 1}$. The estimation $\hat{a}$ of a is given with a 95% asymptotic confidence interval $[\hat{a}_{\min}, \hat{a}_{\max}]$ which is computed using point 3 of Proposition 1. We can see that whatever the choice for b_n, the times between successive failures seem to be stochastically increasing. Finally, Figure 2 shows the empirical c.d.f.s for the different choices of the b_n's.

2.4.2. Monte Carlo study of the estimators

Figure 3 shows three boxplots obtained from estimates of $a \in \{0.85, 0.9, 0.95\}$ for various sequences $(b_n)_{n\geq 1}$ based on 1000 simulated samples of size $n = 50$. Here, the underlying renewal process is generated using independent inter-arrival times that follow a Weibull distribution with shape parameter 2 and scale parameter 10. These figures show that the convergence rate of $\hat{a}_n$ heavily depends on b_n. This is coherent with the results of Example 2, from where we know that for $b_n = n-1$, $\sqrt{n-1}$ or $\log n$, the convergence rate θ_n of $\hat{a}_n$ is proportional to $n^{3/2}$, n or $\sqrt{n}$, respectively. Figure 4 next illustrates the uniform consistency result obtained in Proposition 3, where $\hat{F}_n$ of F is based on the empirical distribution function obtained from the n first observations of the pseudo renewal process $(\tilde{Y}_n)_{n\geq 1}$ defined by $\tilde{Y}_n = \hat{a}^{-b_n}X_n$: the c.d.f. F (black line) is compared with several estimates $\hat{F}_n$ for $n \in \{50, 100, 200, 300, 400, 500\}$.

b_n	$n-1$	$\sqrt{n-1}$	$\log n$
$\hat{a}$	1.078 [1.052,1.104]	1.740 [1.555,1926]	2.590 [2.233,2.947]

Table 1: Estimates of a with 95% asymptotic confidence intervals (within brackets) based on different b_n's.

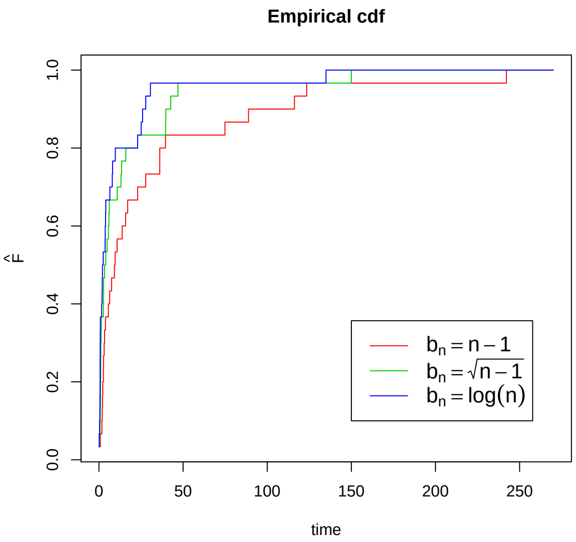


Figure 2: Empirical cumulative distribution function for the air-conditioning equipment.

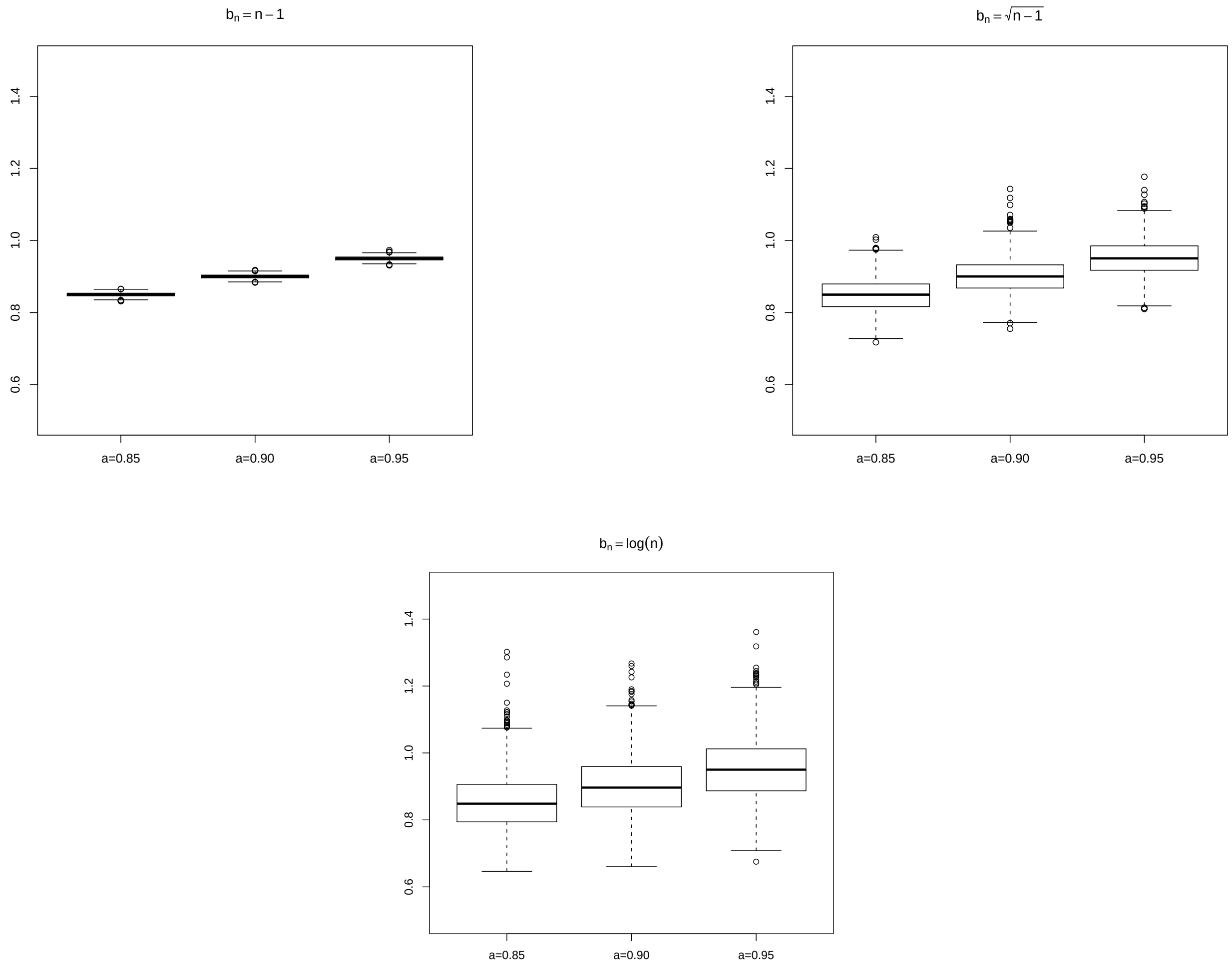


Figure 3: Comparison of boxplots of 1000 estimates of $a \in \{0.85, 0.9, 0.95\}$ obtained from samples of size 50 for $b_n = n-1$, $\sqrt{n-1}$and $\log n$.

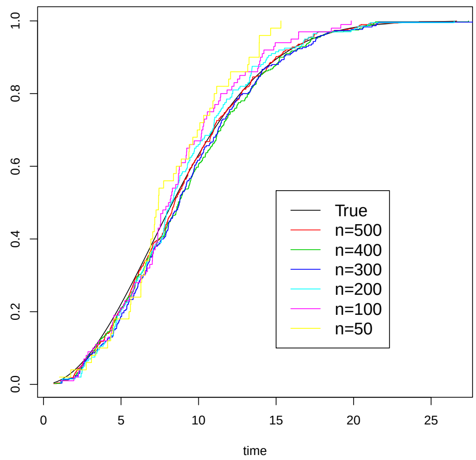


Figure 4: Comparison of $\hat{F}_n$ and F for various values of n.

3. APPLICATION TO RELIABILITY

A repairable system is now considered, with instantaneous repairs at failures and successive life-times modelled by an EGP. Once the process has been statistically estimated, it may be used for prediction purposes and/or optimization of replacement policies. As for prediction purpose, a typical quantity of interest is the mean number of failures on some time interval $[0, t]$, namely the pseudo-renewal function associated to an EGP seen as a counting process.

3.1. The pseudo-renewal function

We first provide conditions for the pseudo-renewal function to be finite. Using results from [6] (in the more general set up of Markov renewal functions), a necessary condition is known to be $\lim_{n\to+\infty} T_n = \infty$ a.s. Using a strong law of large numbers for independent but non identically distributed random variables from [17], we may here prove that this condition is equivalent to $\sum_{i=1}^{+\infty} a^{b_i} = +\infty$, which is hence assumed in all the following.

Example 4 *In case $b_n = n^\alpha (\log n)^\beta$ with $\alpha, \beta \geq 0$ and $a \in (0,1)$, one can prove that $\sum_{i=1}^{+\infty} a^{b_i} = +\infty$ if and only if $\alpha = 0$ and either $\beta < 1$ or ($\beta = 1$ and $a \geq 1/e$), using Raabe's rule for the last condition.*

To write a "pseudo-renewal" equation fullfilled by the "pseudo-renewal" function, we need to envision the case where the first interarrival time X_1 of the EGP is distributed as $X_k = a^{b_k} Y_k$, with $k \geq 1$. Under this condition, the system has already been repaired $k-1$ times at time $T_0 = 0$ and the successive interarrival times are distributed as $X_k, X_{k+1}, \ldots$ This situation is denoted by $Z_0 = k$. More generally, we also set $Z_{T_n} = k$ in case X_{n+1} is distributed as $a^{b_k} Y_k$ (all $k \geq n+1$) and $Z_t = Z_{T_n}$ for $T_n \leq t < T_{n+1}$. The process $(Z_t)_{t\geq 0}$ then appears as a semi-Markov process with semi-Markov kernel provided by $q(i, j, dx) = \mathbf{1}_{\{j=i+1\}} dF_i(x)$.

For $k \geq 1$, we next set $\mathbb{P}_k$ to be the conditional probability measure given that $Z_0 = k$, with $k \geq 1$ and $\mathbb{E}_k$ the associated conditional expectation (with $\mathbb{P}_1 = \mathbb{P}$ and $\mathbb{E}_1 = \mathbb{E}$). For all $t \geq 0$, we also set $N(t)$ to be the number of arrivals of the EGP on $[0, t]$. Given that $Z_0 = k$, the "pseudo-renewal" function associated to the EGP then is

$$n_k(t) = \mathbb{E}_k(N(t)) = \textstyle\sum_{n=1}^{+\infty} \mathbb{P}_k(T_n \leq t) \qquad (\text{with } n_1(t) = n(t))$$

and $n_k(t)$ stands for the mean number of failures on $[0, t]$. The function $n_k(t)$ also appears as the Markov renewal function associated to the semi-Markov process $(Z_t)_{t\geq 0}$.

A sufficient condition for $n_k(t)$ to be finite for all $t \geq 0$ is provided by the following proposition, which may be proved using Raabe's rule again. The pseudo-renewal function $n_k(t)$ is also proved to fullfill a Markov renewal equation, using classical renewal arguments ([6] e.g.).

Proposition 5 *Assume that* $\lim_{n\to+\infty} na^{b_n} > 1/\mathbb{E}(Y_1)$ *(and hence* $\sum_{i=1}^{+\infty} a^{b_i} = +\infty$*).*

1. *We then have* $n_k(t) < +\infty$ *for all* $t \geq 0$ *and all* $k \geq 1$.

2. *The function* n_k *fulfills the following Markov renewal equation:*

$$n_k = F_k + f_k * n_{k+1} \tag{3}$$

for all $k \geq 1$*, where* F_k *(resp.* f_k*) stands for the c.d.f. (resp. p.d.f.) of* X_k.

Example 6 *For* $b_n = (\log n)^{\beta}$ *with* $\beta \geq 0$ *and* $a \in (0,1)$*, we derive that* $n_k(t)$ *is finite for all* $t \geq 0$ *and all* $k \geq 1$ *as soon as one of the following condition is fulfilled: 1.* $\beta < 1$*, 2.* $\beta = 1$ *and* $a > 1/e$*, 3.* $\beta = 1$*,* $a = 1/e$*, and* $\mathbb{E}(Y_1) > 1$.

Using bounding arguments from [10], we now provide practical tools for the numerical assessment of $n_k(t)$ in case $a \geq 1$.

Corollary 7 *Assume* $a \geq 1$*. Setting* $u_n(t) = \mathbb{P}(T_n \leq t)$ *for all* $n \geq 1$*, we have:*

$$0 \leq \frac{n(t) - \sum_{n=1}^{N} u_n(t)}{n(t)} \leq u_N(t) \tag{4}$$

for all $N \geq 1$*. Also,* $(u_n(t))_{n\geq 1}$ *may be computed recursively using*

$$\begin{aligned} u_1(t) &= F(t), \\ u_{n+1}(t) &= (f_{n+1} * u_n)(t) = \frac{1}{a^{b_{n+1}}} \int_0^t u_n(u)\, f\left(\frac{t-u}{a^{b_{n+1}}}\right) du \end{aligned} \tag{5}$$

for all $n \geq 1$*, where* F *(resp.* f*) stands for the c.d.f. (resp. p.d.f.) of* Y_1.

Remark 8 *From a numerical point of view, the* $u_i(t)$*'s are computed using discrete convolutions in (5), which induces numerical errors. Such errors might be quantified using similar methods as in [16].*

In case $a < 1$, the previous result is not valid because $n_N(t) \geq n(t)$. In that case, Monte-Carlo simulations may be used to compute the pseudo-renewal function. A lower bound $n^c(t)$ may also be given, which converges to $n(t)$ when c goes to zero and hence provides a lower approximation for $n(t)$. This approximation is constructed and computed via the following proposition.

Proposition 9 *For* $c > 0$ *and* $t \geq 0$*, let*

$$\tau^c = \inf(n \geq 1 : X_n < c) \quad \text{and} \quad n^c(t) = \mathbb{E}\left(\sum_{n=1}^{\tau^c - 1} \mathbf{1}_{\{T_n \leq t\}}\right) \tag{6}$$

(0 in case of an empty sum).

1. *We have:* $n^c(t) \leq n(t)$ *and* $\lim_{c\to 0^+} n^c(t) = n(t)$.

2. *Setting* $u_n^c(t) = \mathbb{P}(T_n \leq t, X_1 \geq c, \ldots, X_n \geq c)$ *for all* $n \geq 1$*, we have:*

$$n^c(t) = \sum_{n=1}^{\lfloor \frac{t}{c} \rfloor} u_n^c(t), \tag{7}$$

where $\lfloor\cdot\rfloor$ *stands for the floor function. Also,* $(u_n^c(t))_{n\geq 1}$ *may be computed recursively using*

$$u_1^c(t) = (F(t) - F(c))^+ \quad \text{and} \quad u_{n+1}^c(t) = \frac{1}{a^{b_{n+1}}} \int_0^{(t-c)^+} u_n^c(u)\, f\left(\frac{t-u}{a^{b_{n+1}}}\right) du \quad \text{for all } n \geq 1. \tag{8}$$

3.2. A replacement policy

In case of non increasing lifetimes ($a < 1$), a preventive replacement policy is studied, where the system is instantaneously replaced at some cost c_R as soon as a lifetime X_i is observed to be shorter than a predefined threshold s ($s > 0$). Between replacements, the cost of an instantaneous repair after failure is denoted by c_F, with $c_R \geq c_F$. We set $C(s)$ to be the asymptotic unitary cost per unit time. The next proposition use classical results from renewal theory to derive the existence of $C(s)$, and its expression.

Proposition 10 *Assume $a \in (0,1)$. Setting $C(s;[0,t])$ to be the cumulated cost on $[0,t]$ for the threshold s, the asymptotic unitary cost per unit time exists a.s. and is provided by*

$$C(s) = \lim_{t \to +\infty} \frac{C(s;[0,t])}{t} = \frac{c_R + c_F \mathbb{E}(\tau^s - 1)}{\mathbb{E}(T_{\tau^s})} \quad a.s. \tag{9}$$

where τ^s is defined in (6). Besides:

$$\mathbb{E}(\tau^s - 1) = \sum_{k=1}^{+\infty} v_k^s \quad \text{and} \quad \mathbb{E}(T_{\tau^s}) = \mathbb{E}(Y_1)\left(1 + \sum_{k=1}^{+\infty} a^{b_{k+1}}\ v_k^s\right)$$

with

$$v_k^s = \prod_{i=1}^{k} \bar{F}\left(\frac{s}{a^{b_i}}\right) \text{ for all } k \geq 1 \text{ and } \bar{F} = 1 - F. \tag{10}$$

We next provide tools for the numerical assessment of $C(s)$.

Proposition 11 *Assume $a \in (0,1)$. We have:*

$$\left| C(s) - \frac{m_C(s) + M_C(s)}{2} \right| \leq \Delta C_{\max}(s) := \frac{M_C(s) - m_C(s)}{2},$$

where

$$m_C(s) = \frac{c_R + c_F\ S_1^N(s)}{\mathbb{E}(Y_1)\left(1 + S_2^N(s) + a^{b_{N+2}}\ v_{N+1}^s / F\left(\frac{s}{a^{b_{N+2}}}\right)\right)},$$

$$M_C(s) = \frac{c_R + c_F\ \left(S_1^N(s) + v_{N+1}^s / F\left(\frac{s}{a^{b_{N+2}}}\right)\right)}{\mathbb{E}(Y_1)\left(1 + S_2^N(s)\right)},$$

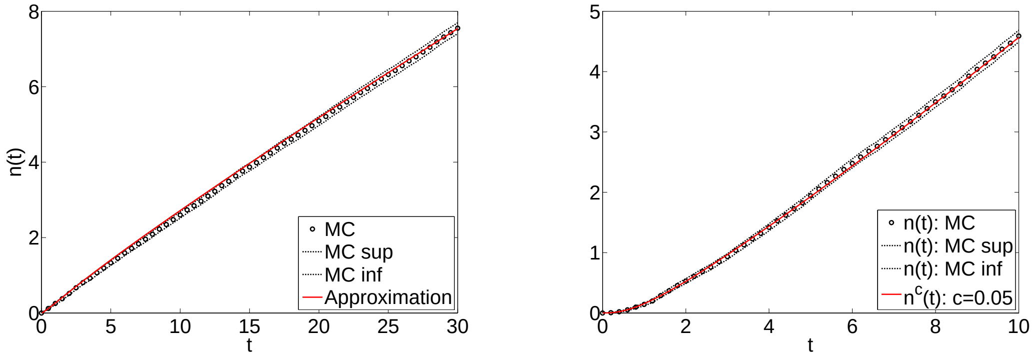
and $S_1^N(s) = \sum_{k=1}^{N} v_k^s$, $S_2^N(s) = \sum_{k=1}^{N} a^{b_{k+1}} v_k^s$ (with v_k^s defined by (10)).

3.3. Numerical experiments

3.3.1. Computation of the pseudo-renewal function

We first consider the case where $a \geq 1$: the random variable Y_1 follows $\Gamma(1.2, 2.5)$ with $\mathbb{E}(Y_1) = 3$, $\mathbb{V}\mathrm{ar}(Y_1) = 7.5$, $b_n = n^{0.3}$ and $a = 1.2$. The approximation of $n(t)$ provided by Corollary 7 is plotted against t in Figure 5a for $N = 20$. The maximal relative error provided by the approximation is about 4.2×10^{-6}. We also plot $n(t)$ computed by Monte-Carlo (MC) simulations and the 95% confidence band for 10^3 stories in the same Figure. The results are quite similar.

We next consider the case where $a < 1$ (and $\lim_{n \to +\infty} n a^{b_n} > 1/\mathbb{E}(Y_1)$): the random variable Y_1 follows $\Gamma(2.5, 1)$ with $\mathbb{E}(Y_1) = \mathbb{V}\mathrm{ar}(Y_1) = 2.5$, $b_n = (\log n)^{0.7}$ and $a = 0.8$. The approximating lower bound $n^c(t)$ for $n(t)$ is computed via the results of Proposition 9 for $c = 0.05$. It is compared in Figure 5b to the results for $n(t)$ by MC simulations with 95% confidence band for 10^3 stories. As expected, we observe that $n^c(t)$ is a good approximation of $n(t)$ for small c.



(a) Case $a > 1$ (b) Case $a < 1$

Figure 5: Computation of $n(t)$ by MC simulations and by the paper approximations

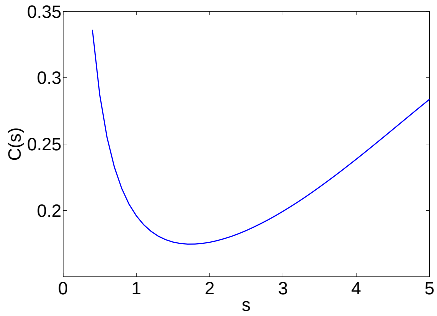


Figure 6: The cost function $C(s)$with respect of s.

3.3.2. The replacement policy

The random variable Y_1 follows $\Gamma(2.5, 1)$ with $\mathbb{E}(Y_1) = \mathbb{V}\text{ar}(Y_1) = 2.5$, $b_n = (\log n)^{0.7}$, $a = 0.8$, $c_R = 1$ and $c_F = 0.5$. For $N = 100$, the maximal absolute error $\Delta C_{\max}(s)$ decreases very quickly as s increases ($\Delta C_{\max}(0.4) \simeq 8 \times 10^{-5}$, $\Delta C_{\max}(0.7) \simeq 3 \times 10^{-12}$, beyond the machine precision for $s \geq 0.9$). The cost function $C(s)$ is plotted against s in Figure 6. The cost function reaches its minimum at $s^{opt} \simeq 1.70$, with $\min_{s>0} C(s) = C\left(s^{opt}\right) \simeq 0.17$.

4. CONCLUDING REMARKS AND PROSPECTS

Contrary to classical renewal processes and just as the geometric processes proposed by [14], extended geometric processes allow to account for some stochastic monotony property for the successive inter-arrival times of a counting process, with some more flexibility in the modelling than geometric processes however. Extended geometric processes may hence be a simple alternative to the virtual age models proposed by [9] and [11] for the modeling of imperfect maintenance actions e.g.

From the estimation point of view, we saw that the convergence rate of the estimator of the Euclidean parameter strongly depends on the sequence $(b_n)_{n\geq 1}$. A miss-specification of the sequence $(b_n)_{n\geq 1}$ will naturally lead to biased estimates. To make the model more flexible, it would be interesting to consider a parametrized version of the sequence $(b_n)_{n\geq 1}$ setting for example $b_n = g(n, \theta)$, where θ would be an additional Euclidean parameter, to be estimated. We can also mention the lack of a central limit theorem for the estimator $\hat{F}$ of the underlying c.d.f. F. Indeed, standard methods cannot be used here, because of the deterministic nature of the b_n's. This problem hence requires some more investigation. Such a result would however be useful for testing the hypothesis that the underlying c.d.f. F belongs to some parametric family. Another possible issue would be to include covariates in this model in order to describe (e.g.) the effect of the environment on the monotonicity of the EGP.

In case $a < 1$, a lower bound has been provided for the pseudo-renewal function, which is easy to compute using Lemma 9. However, we haven't been able to provide a computable upper bound, though necessary for the numerical assessment of the results precision. Indeed, usual tools such as those used in case $a \geq 1$ are inappropriate here, and new tools should be developed. As for the replacement policy, because of the random character of the successive lifetimes, an alternate policy based on a predefined number m of consecutive lifetimes under a threshold s, might be better adapted than the present policy, based on a replacement at the first observation of a single lifetime bellow s.

REFERENCES

[1] Asmussen S. *Applied probability and queues.* Applications of Mathematics **51**. Springer-Verlag, New York, 2003 (sec. ed.).

[2] Barlow R. E. and Proschan F. *Mathematical theory of reliability.* Classics in Applied Mathematics **17**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1965.

[3] Bai Z. D., Cheng P. E., Zhang C. H. An extension of Hardy-Littlewood strong law. *Statistica Sinica*: 923–928, 1997.

[4] Bai Z. D., Cheng P. E. Marcinkiewicz strong laws for linear statistics. *Statistics & Probability Letters* **46**: 105–112, 2000.

[5] Bordes L., Mercier S. Extended geometric processes: semiparametric estimation and application to reliability, *submitted*.

[6] Çinlar E. *Introduction to Stochastic Processes.* Prentice-Hall. Englewood Cliffs, NJ, 1975.

[7] Cook, R.J. and Lawless J.F. *The Statistical Analysis of Recurrent Events.* Springer, 2007.

[8] Cuzick, J. A strong law for large for weighted sums of I.I.D. random variables. *Journal of Theoretical Probability* **8**(3): 625–641, 1995.

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

[10] Feller W. *An introduction to probability theory and its applications. Vol. I*, John Wiley & Sons Inc., New York, 1968 (third ed.).

[11] Kijima M. Generalized Renewal Processes and General Repair Models in *Stochastic models in reliability and maintenance*, Osaki, Shunji ed., 145–164, Springer, Berlin, 2002.

[12] Lam Y. Nonparametric inference for geometric processes. *Communications in Statistics-Theory and Methods* **21**: 2083–2105, 1992.

[13] Lam Y., Zhu, L. X., Chan, S. K. and Liu, Q. Analysis of data from a series of events by a geometric process model. *Acta Mathematicae Applicatae Sinica* **20**: 263–282, 2004.

[14] Lam Y. *The Geometric Process and its Applications.* World Scientific, 2007.

[15] Lindsey J. K. *Statistical Analysis of Stochastic Processes in Time*, Cambridge University Press, Cambridge, 2004.

[16] Mercier S. Numerical bounds for semi-markovian quantities and application to reliability. *Methodology and Computing in Applied Probability* **10**(2): 179–198, 2008.

[17] Petrov V. V. *Limit Theorems of Probability Theory: Sequences of Independent Random Variables.* Oxford University Press, 1995.

[18] Proschan F. Theoretical explanation of observed decreasing failure rate. *Technometrics* **5**: 375–383, 1963.

[19] van der Vaart A. W. *Asymptotic Statistics*, Cambridge University Press, Cambridge, 1998.