Generalized Method of Moments for an Extended Gamma Process

Z. Al Masry, S. Mercier & G. Verdier
Laboratoire de Mathématiques et de leurs Applications - Pau (UMR CNRS 5142),
Université de Pau et des Pays de l'Adour, Pau, France

ABSTRACT: In reliability theory, a widely used process to model the phenomena of the cumulative deterioration of a system over time is the standard Gamma process. Based on several restrictions, such as a constant variance-to-mean ratio, this process is not always a suitable choice to describe the deterioration. A way to overcome these restrictions is to use an extended version of the Gamma process introduced by Cinlar (1980). In this paper, the aim is to propose statistical methods which enable us to estimate the unknown parameters of an EGP from a parametric form of the shape and scale functions. We here develop a generalized method of moments, which was introduced by Hansen (1982), based on the moments or the Laplace transform to estimate the parameters. Asymptotic properties are provided and the performance of the proposed estimation methods is illustrated on simulated data.

1 INTRODUCTION

Standard Gamma process is a widely used process to model the cumulative deterioration of a system over time (see van Noortwijk (2009) for an overview). However, this process is not a suitable choice when the variance-to-mean ratio of a system is not constant over time. A way to overcome this restriction is to use an extended Gamma process (see Çinlar (1980)), defined as a stochastic integral with respect to a standard Gamma process. It is a non decreasing process with independent increments and it is characterized by a shape function and a scale function.

In this paper, the aim is to propose statistical methods which enable us to estimate the unknown parameters of an extended Gamma process (EGP) from a parametric form of the shape and scale functions. Parameter estimation is an important task for a practical use of an EGP in an industrial reliability context. Maximum likelihood estimation (MLE) is the most classically applied estimation technique which however requires the complete specification of the probability density function (pdf). Based on the fact that an explicit expression is not available for the pdf of an EGP, standard maximum likelihood estimation is here not possible. One could also think of the empirical likelihood method of Qin & Lawless (1994). However, this method does not seem to be adapted to the present case of an EGP, since it requires estimating too many parameters. The moments and Laplace transform are available in full form for an EGP. We here develop generalized methods of moments (GMM) based on either one of these quantities. Recall that GMM was introduced by Hansen (1982). It does not require a full knowledge of the pdf and it relies on a set of population moment conditions upon which estimation is based. GMM based on the empirical characteristic function has been the subject of many papers in the literature (Feuerverger & McDunnough 1981, Carrasco & Florens 2002).

The remaining of the paper is organized as follows. Section 2 briefly introduces the general approach of GMM. Then, we elaborate about the two approaches of GMM for an EGP in Section 3 and provide asymptotic properties. In Section 4, the focus is on a parametric form of the shape and scale functions. Illustrations are presented in Section 5 and we finally conclude in Section 6.

2 GENERAL SPECIFICATION OF GMM

Let $\boldsymbol{W}$ be a random vector of dimension d and $\{\boldsymbol{W}_n, n = 1, \ldots, N\}$ a set of independent and identically distributed (i.i.d.) random vectors sharing the same distribution with $\boldsymbol{W}$. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ be a parameter vector that indexes the distribution of $\boldsymbol{W}$ and $\Theta \subseteq \mathbb{R}^p$ the parameter space. Let also $\boldsymbol{f} : \mathbb{R}^d \times \Theta \to \mathbb{R}^q$ ($q \geq p$) be a function such that

$$\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{f}^{(1)}(w^{(1)}, \boldsymbol{\theta}) \\ \vdots \\ \boldsymbol{f}^{(d)}(w^{(d)}, \boldsymbol{\theta}) \end{pmatrix}, \tag{1}$$

where $\boldsymbol{w} = \left(w^{(1)}, \ldots, w^{(d)}\right)$ and $\boldsymbol{f}^{(i)}(w^{(i)}, \boldsymbol{\theta})$, $i = 1, \ldots, d$ is a column vector with dimension k.
The gradient of $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ can be given by

$$\frac{\partial \boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial \boldsymbol{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ \vdots \\ \frac{\partial \boldsymbol{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \end{pmatrix} = \begin{pmatrix} \frac{\partial \boldsymbol{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \boldsymbol{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & & \\ \frac{\partial \boldsymbol{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \boldsymbol{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix} \quad (2)$$

if it exists, for all $\boldsymbol{w} \in \mathbb{R}^d, \boldsymbol{\theta} \in \Theta$.

The first step of GMM begins by defining the population moment condition and the sample moment condition.

Definition 2.1. (Hall 2005, Definition 1.1, p.14) Let $\boldsymbol{\theta}_0$ be the true unknown parameter vector to be estimated. Then the population moment condition takes the form

$$\mathbb{E}[\boldsymbol{f}(\boldsymbol{W}, \boldsymbol{\theta}_0)] = 0. \quad (3)$$

The population moment condition should provide sufficient information to identify the unknown parameters. It represents a set of $q = kd$ equations for p unknowns which are exactly solved by $\boldsymbol{\theta}_0$. The corresponding sample moment condition for an arbitrary $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{g}}_N(\boldsymbol{\theta}) = 0 \quad (4)$$

where $\hat{\boldsymbol{g}}_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{f}(\boldsymbol{W}_n, \boldsymbol{\theta})$.

Next, the GMM estimator is defined as follows:

Definition 2.2. (Hall 2005, Definition 1.2, p.14) Let $(\boldsymbol{P}_N)$ be a sequence of positive semi-definite weighting matrices that converges in probability to a constant positive definite matrix $\boldsymbol{P}$. Then, the GMM estimator based on the population moment condition (3) is given by

$$\hat{\boldsymbol{\theta}}_N = \underset{\boldsymbol{\theta} \in \Theta}{\arg\min} \hat{\boldsymbol{g}}_N(\boldsymbol{\theta})^T \boldsymbol{P}_N \, \hat{\boldsymbol{g}}_N(\boldsymbol{\theta}). \quad (5)$$

Under some technical assumptions, GMM estimator is consistent and asymptotically normal (Hansen 1982, Hall 2005, Newey & McFadden 1994).

3 GMM FOR AN EGP

Section 2 introduced the general framework of GMM. The focus here is on GMM for a particular model, which is the EGP.

3.1 *Definition of an EGP*

Throughout we let $\mathbb{R}_+ = [0, +\infty)$ and $\mathbb{R}_+^* = (0, +\infty)$. Let $A : [0, T] \times \Theta \to \mathbb{R}_+$ be a measurable, increasing and right-continuous function with $A(0, \boldsymbol{\theta}) = 0$ for all $\boldsymbol{\theta} \in \Theta$ and let $b : (0, T] \times \Theta \to \mathbb{R}_+^*$ be a measurable positive function such that, for $t \in (0, T], \boldsymbol{\theta} \in \Theta$:

$$\int_{(0,t]} \frac{a(s, \boldsymbol{\theta}) ds}{b(s, \boldsymbol{\theta})} < \infty \quad (6)$$

where $a(t, \boldsymbol{\theta})$ is the derivative of $A(t, \boldsymbol{\theta})$ with respect to t on a given compact set $[0, T]$.
Following (Çinlar 1980, Dykstra & Laud 1981), the process $\boldsymbol{X} = (X_t)_{t \in [0,T]}$ is said to be an EGP with shape function $A(t, \boldsymbol{\theta})$ and scale function $b(t, \boldsymbol{\theta})$ (written $\boldsymbol{X} \sim \Gamma(A(t, \boldsymbol{\theta}), b(t, \boldsymbol{\theta}))$) if it can be represented as a stochastic integral with respect to a standard Gamma process $(Y_t)_{t \in [0,T]} \sim \Gamma(A(t, \boldsymbol{\theta}), 1)$:

$$X_t = \int_{(0,t]} \frac{dY_s}{b(s, \boldsymbol{\theta})}, \quad \forall t \in (0, T], \boldsymbol{\theta} \in \Theta \quad (7)$$

and $X_0 = 0$.

An EGP has independent increments and an explicit formula is available for the Laplace transform of an increment, with

$$\mathcal{L}_{X_{t+h} - X_t}(\lambda, \boldsymbol{\theta})$$

$$:= \mathbb{E}\left(e^{-\lambda(X_{t+h} - X_t)}\right) \quad (8)$$

$$= \exp\left(-\int_{(t,t+h]} \log\left(1 + \frac{\lambda}{b(s, \boldsymbol{\theta})}\right) a(s, \boldsymbol{\theta}) ds\right), \quad (9)$$

for all $t \in [0, T]; \boldsymbol{\theta} \in \Theta; \lambda \geq 0$ and $h > 0$.

The mean and variance of an EGP are given by

$$\mathbb{E}(X_t) = \int_{(0,t]} \frac{a(s, \boldsymbol{\theta}) ds}{b(s, \boldsymbol{\theta})}, \quad (10)$$

$$\mathbb{V}(X_t) = \int_{(0,t]} \frac{a(s, \boldsymbol{\theta}) ds}{b(s, \boldsymbol{\theta})^2}. \quad (11)$$

Let us now consider $t_0 = 0 < t_1 < \cdots < t_d = T$. We assume that we observe N trajectories $\boldsymbol{X}_n, n = 1, \ldots, N$ at the same t_i's, $i = 1, \ldots, d$ which are issued from $\boldsymbol{\theta}_0$. For a generic trajectory $\boldsymbol{X}$, we set

$$W^{(i)} = X_{t_i} - X_{t_{i-1}}, i = 1, \ldots, d \quad (12)$$

to be the i-th increment. This leads to a sequence of i.i.d. random vectors $\boldsymbol{W}_n, n = 1, \ldots, N$.

Now we discuss the choice of the population moment condition for each of the following approaches: GMM based on the moments and GMM based on the Laplace transform.

3.2 *Approaches*

3.2.1 *GMM based on the moments*

As in the classical method of moments, population moment condition for GMM based on the moments is defined by matching the theoretical moments of an EGP with the appropriate empirical ones. This matching is done for each increment. Note that, in the case of GMM, there are more equations than unknowns. Define

$$\boldsymbol{f}(\boldsymbol{w},\boldsymbol{\theta}) = \left(\boldsymbol{f}^{(i)}(w^{(i)},\boldsymbol{\theta})\right)_{1\le i\le d}$$

$$= \begin{pmatrix} w^{(i)} - m^{(i)}(\boldsymbol{\theta}) \\ \left(w^{(i)} - m^{(i)}(\boldsymbol{\theta})\right)^2 - v^{(i)}(\boldsymbol{\theta}) \end{pmatrix}_{1\le i\le d} \tag{13}$$

where

$$m^{(i)}(\boldsymbol{\theta}) = \int_{t_{i-1}}^{t_i} \frac{a(s,\boldsymbol{\theta})ds}{b(s,\boldsymbol{\theta})};$$

$$v^{(i)}(\boldsymbol{\theta}) = \int_{t_{i-1}}^{t_i} \frac{a(s,\boldsymbol{\theta})ds}{b(s,\boldsymbol{\theta})^2}.$$

Obviously, $\boldsymbol{f}$ satisfies $\mathbb{E}[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)] = 0$. We have $\mathbb{E}[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta})] < \infty$ for all $\boldsymbol{\theta} \in \Theta$ (from (6)) and then

$$\hat{\boldsymbol{g}}_N(\boldsymbol{\theta})$$

$$= \begin{pmatrix} \frac{1}{N}\sum_{n=1}^{N}\left[W_n^{(i)} - m^{(i)}(\boldsymbol{\theta})\right] \\ \frac{1}{N}\sum_{n=1}^{N}\left[\left(W_n^{(i)} - m^{(i)}(\boldsymbol{\theta})\right)^2 - v^{(i)}(\boldsymbol{\theta})\right] \end{pmatrix}_{1\le i\le d}$$

$$= \begin{pmatrix} \hat{m}^{(i)} - m^{(i)}(\boldsymbol{\theta}) \\ \hat{v}^{(i)}(\boldsymbol{\theta}) - v^{(i)}(\boldsymbol{\theta}) \end{pmatrix}_{1\le i\le d} \tag{14}$$

with

$$\hat{m}^{(i)} = \frac{1}{N}\sum_{n=1}^{N} W_n^{(i)};$$

$$\hat{v}^{(i)}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{n=1}^{N}\left(W_n^{(i)} - m^{(i)}(\boldsymbol{\theta})\right)^2.$$

3.2.2 *GMM based on the Laplace transform*

A similar procedure as for the previous method is followed. Instead of using moments, we rely on the Laplace transform at distinct points $\lambda_l, 1 \le l \le o$ and we match it with the empirical Laplace transform at the same points λ_l. Let

$$\left(\boldsymbol{f}^{(i)}(w^{(i)},\boldsymbol{\theta})\right)_{1\le i\le d}$$

$$= \left(\exp(-\lambda_l w^{(i)}) - \mathcal{L}^{(i)}(\lambda_l,\boldsymbol{\theta})\right)_{\substack{1\le i\le d\\ 1\le l\le o}} \tag{15}$$

where

$$\mathcal{L}^{(i)}(\lambda_l,\boldsymbol{\theta}) = \exp\left(-\int_{t_{i-1}}^{t_i} \ln\left(1 + \frac{\lambda_l}{b(s,\boldsymbol{\theta})}\right) a(s,\boldsymbol{\theta})ds\right).$$

If the arbitrary $\boldsymbol{\theta}$ is replaced by the true value, we have $\mathbb{E}[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)] = 0$. Also, $\mathbb{E}[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta})] < \infty$ for all $\boldsymbol{\theta} \in \Theta$ and the sample moment is given by

$$\hat{\boldsymbol{g}}_N(\boldsymbol{\theta}) = \left(\hat{\mathcal{L}}^{(i)}(\lambda_l) - \mathcal{L}^{(i)}(\lambda_l,\boldsymbol{\theta})\right)_{\substack{1\le i\le d\\ 1\le l\le o}} \tag{16}$$

with

$$\hat{\mathcal{L}}^{(i)}(\lambda_l) = \frac{1}{N}\sum_{n=1}^{N}\exp(-\lambda_l W_n^{(i)}).$$

3.3 *Asymptotic properties*

In this subsection, general asymptotic results for GMM for an EGP are presented. For any parametric form of the shape and scale functions, one need to verify the assumptions given below to show consistency and asymptotic normality of GMM for an EGP estimator. The results are derived from (Newey & McFadden 1994, Theorem 2.6, p.2132), (Hansen 1982, Theorem 2.1, p.1035), (Newey & McFadden 1994, Theorem 3.2, p.2145) and (Hall 2005, Theorem 3.4, p.88).

3.3.1 *GMM based on the moments*

A consistency result for GMM based on the moments is now formulated:

Theorem 3.1 (Consistency)**.** *If*

- (H_1) $(\boldsymbol{P}_N)$ *converges in probability (respectively, almost surely) to a constant positive definite matrix* $\boldsymbol{P}$*,*
- (H_2) $\mathbb{E}[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta})] = 0$ *if and only if* $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ *(identification condition),*
- (H_3) Θ *is compact,*
- (H_4) *There exists* $I_1 : [0,T] \to \mathbb{R}_+$ *such that* $\left|\frac{a(s,\boldsymbol{\theta})}{b(s,\boldsymbol{\theta})}\right| \le I_1(s)$ *for all* $s \in [0,T], \boldsymbol{\theta} \in \Theta$ *with* $\int_0^T I_1(s)ds < \infty$*,*
- (H_5) *There exists* $I_2 : [0,T] \to \mathbb{R}_+$ *such that* $\left|\frac{a(s,\boldsymbol{\theta})}{b(s,\boldsymbol{\theta})^2}\right| \le I_2(s)$ *for all* $s \in [0,T], \boldsymbol{\theta} \in \Theta$ *with* $\int_0^T I_2(s)ds < \infty$*.*

Then $\hat{\boldsymbol{\theta}}_N$ *converges in probability (respectively, almost surely) to* $\boldsymbol{\theta}_0$.

In the following, the convergence in probability is denoted by "Prob", the almost sure convergence by "a.s." and the convergence in distribution by "$\mathcal{D}$".

In addition to assumptions $(H_1 - H_5)$ given in Theorem 3.1, more conditions are needed for the asymptotic normality:

Theorem 3.2 (Asymptotic Normality). *If $(H_1 - H_5)$ and the following conditions hold*

(H_6) *$\boldsymbol{\theta}_0$ is an interior point in Θ,*

(H_7) *$\boldsymbol{D}_0^T \boldsymbol{P} \boldsymbol{D}_0$ is non-singular with*

$$\boldsymbol{D}_0 = \begin{pmatrix} -\frac{\partial m^{(i)}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \\ -\frac{\partial v^{(i)}(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \end{pmatrix}_{1 \le i \le d},$$

(H_8) *There exists $J_1 : [0,T] \to \mathbb{R}_+$ such that $\left| \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{a(s,\boldsymbol{\theta})}{b(s,\boldsymbol{\theta})} \right) \right| \le J_1(s)$ for all $s \in [0,T], \boldsymbol{\theta} \in \Theta$ with $\int_0^T J_1(s) ds < \infty$,*

(H_9) *There exists $J_2 : [0,T] \to \mathbb{R}_+$ such that $\left| \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{a(s,\boldsymbol{\theta})}{b(s,\boldsymbol{\theta})^2} \right) \right| \le J_2(s)$ for all $s \in [0,T], \boldsymbol{\theta} \in \Theta$ with $\int_0^T J_2(s) ds < \infty$,*

(H_{10}) *$\int_0^T \frac{a(s,\boldsymbol{\theta}) ds}{b(s,\boldsymbol{\theta})^3} < \infty$, $\int_0^T \frac{a(s,\boldsymbol{\theta}) ds}{b(s,\boldsymbol{\theta})^4} < \infty$ for all $\boldsymbol{\theta} \in \Theta$.*

Then $\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)$ is asymptotically normal with mean 0 and asymptotic variance $\boldsymbol{V} = \boldsymbol{H} \boldsymbol{S} \boldsymbol{H}^T$ where

$$\boldsymbol{H} = (\boldsymbol{D}_0^T \boldsymbol{P} \boldsymbol{D}_0)^{-1} \boldsymbol{D}_0^T \boldsymbol{P}.$$

Theorem 3.2 indicates that the weighting matrix affects the asymptotic properties of the estimator via the covariance matrix. The optimal weighting matrix, in the sense of minimizing the asymptotic covariance matrix, is given by the following theorem.

Theorem 3.3 (Optimal weighting matrix). *If assumptions $(H_1 - H_{10})$ hold, then the minimal asymptotic variance of $\hat{\boldsymbol{\theta}}_N$ is $\left(\boldsymbol{D}_0^T \boldsymbol{S}^{-1} \boldsymbol{D}_0\right)^{-1}$ and this can be obtained by setting $\boldsymbol{P} = \boldsymbol{S}^{-1}$ with $\boldsymbol{S} = \mathbb{E}\left[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)^T\right]$.*

3.3.2 *GMM based on the Laplace transform*

The following theorems state the asymptotic results for GMM based on the Laplace transform.

Theorem 3.4 (Consistency). *Assumptions $(H_1 - H_4)$ imply*

$$\hat{\boldsymbol{\theta}}_N \xrightarrow[N\to\infty]{Prob} \boldsymbol{\theta}_0.$$

Theorem 3.5 (Asymptotic Normality). *Assumptions $(H_1 - H_4)$ and $(H_6 - H_8)$ imply*

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \boldsymbol{V})$$

with

$\boldsymbol{V} = \boldsymbol{H} \boldsymbol{S} \boldsymbol{H}^T$;

$\boldsymbol{H} = (\boldsymbol{D}_0^T \boldsymbol{P} \boldsymbol{D}_0)^{-1} \boldsymbol{D}_0^T \boldsymbol{P}$;

$\boldsymbol{D}_0 = \left(-\frac{\partial \mathcal{L}^{(i)}(\lambda_l,\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right)_{\substack{1 \le i \le d \\ 1 \le l \le o}}$.

Theorem 3.6 (Optimal weighting matrix). *If $(H_1 - H_4)$ and $(H_6 - H_8)$ hold, then the minimal asymptotic variance of $\hat{\boldsymbol{\theta}}_N$ is $\left(\boldsymbol{D}_0^T \boldsymbol{S}^{-1} \boldsymbol{D}_0\right)^{-1}$ and this can be obtained by setting $\boldsymbol{P} = \boldsymbol{S}^{-1}$ with $\boldsymbol{S} = \mathbb{E}\left[\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)\boldsymbol{f}(\boldsymbol{W},\boldsymbol{\theta}_0)^T\right]$.*

Notice that less assumptions are required to show the asymptotic properties of GMM based on the Laplace transform than GMM based on the moments.

3.4 *Estimation*

In practice, the optimal weighting matrix $\boldsymbol{S}^{-1}$ should be estimated. As $\boldsymbol{\theta}_0$ is unknown, a consistent GMM estimator is required in order to estimate $\boldsymbol{S}^{-1}$.

Let $\boldsymbol{S}_N = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{f}(\boldsymbol{W}_n,\boldsymbol{\theta}_0)\boldsymbol{f}(\boldsymbol{W}_n,\boldsymbol{\theta}_0)^T$ be an empirical version of $\boldsymbol{S}$. An easy way to proceed is to adopt a two-step procedure (see Hansen (1982)):

1. Set $\boldsymbol{P}_N = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix, and compute
$$\hat{\boldsymbol{\theta}}_N^{(1)} = \arg\min_{\boldsymbol{\theta}\in\Theta} \hat{\boldsymbol{g}}_N(\boldsymbol{\theta})^T \hat{\boldsymbol{g}}_N(\boldsymbol{\theta}). \quad (17)$$
2. Construct an estimator of $\boldsymbol{S}$ based on the initial GMM estimator $\hat{\boldsymbol{\theta}}_N^{(1)}$
$$\hat{\boldsymbol{S}}_N = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{f}(\boldsymbol{W}_n, \hat{\boldsymbol{\theta}}_N^{(1)}) \boldsymbol{f}(\boldsymbol{W}_n, \hat{\boldsymbol{\theta}}_N^{(1)})^T. \quad (18)$$

For GMM based on the moments, $\hat{\boldsymbol{S}}_N$ can be shown to be almost surely non singular. Also, under the condition that λ_i's, $i = 2, \ldots, o$ are multiples of λ_1, $\hat{\boldsymbol{S}}_N$ can be shown to be almost surely non singular for GMM based on the Laplace transform. Then the optimal two-step GMM estimator of $\boldsymbol{\theta}_0$ is given by

$$\hat{\boldsymbol{\theta}}_N = \arg\min_{\boldsymbol{\theta}\in\Theta} \hat{\boldsymbol{g}}_N(\boldsymbol{\theta})^T \hat{\boldsymbol{S}}_N^{-1} \hat{\boldsymbol{g}}_N(\boldsymbol{\theta}). \quad (19)$$

Furthermore, for both approaches, it can be shown that

$$\hat{\boldsymbol{S}}_N \xrightarrow[N\to\infty]{a.s.} \boldsymbol{S}.$$

Let $\hat{\boldsymbol{V}}_N = \left(\boldsymbol{D}_0^T \hat{\boldsymbol{S}}_N^{-1} \boldsymbol{D}_0\right)^{-1}$ denote the estimator of the asymptotic variance $\boldsymbol{V}$. Theorems 3.2, 3.5 imply that an approximate $100 \times (1-\epsilon)\%$ confidence interval for $\boldsymbol{\theta}_0(i)$, $i = 1, \ldots, p$ is given by

$$\hat{\boldsymbol{\theta}}_N(i) \pm q_{\epsilon/2}\sqrt{\hat{\boldsymbol{V}}_N(i,i)/N}] \tag{20}$$

where $q_{\epsilon/2}$ is the $(1-\epsilon/2)$ quantile of the standard normal distribution.

4 A PARAMETRIC FORM FOR THE SCALE AND SHAPE FUNCTIONS

Let $\boldsymbol{\theta} = (a, \alpha, b, \beta, c)$, $a(t, \boldsymbol{\theta}) = at^{\alpha}$ and $b(t, \boldsymbol{\theta}) = b(t+c)^{\beta}$. The parameter space is given by

$$\Theta = \left(\mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+\right)$$
$$\cup \{(a, \alpha, b, \beta, c) \in \mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \{0\}$$
$$\text{such that } \alpha > 2\beta - 1\}. \tag{21}$$

Based on Section 3, we have the following results for this parametric form using both approaches of GMM for an EGP. Firstly, we address the issue of identification (see (H_2)) which is the most technical point to verify. Secondly, we provide conditions under which GMM estimator using this parametric form is consistent and asymptotically normal.

4.1 *GMM based on the moments*

Theorem 4.1. *Consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (13). If at least 5 increments are observed ($d \geq 5$), then (H_2) is satisfied, namely the model is identifiable.*

From Theorem 3.1, it follows

Corollary 4.2. *Let us consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (13) with $d \geq 5$ and $\boldsymbol{\theta}_0 \in \overline{\Theta} \subset \Theta$ where $\overline{\Theta}$ is a compact set. Assume that assumption (H_1) holds, then*

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{Prob} \boldsymbol{\theta}_0.$$

Now, we define

$$\Theta^{\mathcal{N}} = \left(\mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+\right)$$
$$\cup \{(a, \alpha, b, \beta, c) \in \mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \{0\}$$
$$\text{such that } \alpha > 4\beta - 1\}. \tag{22}$$

Corollary 4.3. *Consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (13) with $d \geq 5$ and a compact set $\overline{\Theta}^{\mathcal{N}} \subset \Theta^{\mathcal{N}}$. If $\boldsymbol{\theta}_0$ is an interior point in $\overline{\Theta}^{\mathcal{N}}$ and (H_1) and (H_7) are satisfied for $\boldsymbol{P} = \boldsymbol{S}^{-1}$, then*

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\boldsymbol{D}_0^T \boldsymbol{S}^{-1} \boldsymbol{D}_0)^{-1}).$$

4.2 *GMM based on the Laplace transform*

Theorem 4.4. *Consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (15). If $d \geq 4$ and $o \geq 3$, then (H_2) is satisfied.*

Corollary 4.5. *Let us consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (15) with $d \geq 4, o \geq 3$ and $\boldsymbol{\theta}_0 \in \overline{\Theta} \subset \Theta$ where $\overline{\Theta}$ is a compact set. Under assumption (H_1) it follows that*

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{Prob} \boldsymbol{\theta}_0.$$

Corollary 4.6. *Consider $\boldsymbol{f}(\boldsymbol{w}, \boldsymbol{\theta})$ as defined in (15) with $d \geq 4, o \geq 3$ and a compact set $\overline{\Theta} \subset \Theta$. If $\boldsymbol{\theta}_0$ is an interior point in $\overline{\Theta}$ and if we assume that (H_1) and (H_7) hold for $\boldsymbol{P} = \boldsymbol{S}^{-1}$, then*

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\boldsymbol{D}_0^T \boldsymbol{S}^{-1} \boldsymbol{D}_0)^{-1}).$$

5 NUMERICAL EXPERIMENTS

Here, we test GMM based on the moments (GMM_{MM}) and Laplace transform (GMM_{Lap}) on simulated data.

We first generate N sample paths on $[0, T]$ using a simulation technique from Al Masry et al. (2015). By taking $\boldsymbol{P} = \boldsymbol{I}$, an initial GMM estimator is obtained (see step 1, Subsection 3.4). Then, using this estimator, we construct $\hat{\boldsymbol{S}}_N$ and compute $\hat{\boldsymbol{\theta}}_N$ using (18) and (19) respectively. In order to study the behavior of the estimator, we consider r sets of N sample paths. For each set, we compute $\hat{\boldsymbol{\theta}}_N$ and then we report the mean, the standard deviation and the quantiles based on these r estimations of $\hat{\boldsymbol{\theta}}_N$.

Moreover, by the asymptotic normality of GMM estimators, we give the coverage probabilities (CP) of hyper-ellipsoid confidence region. Firstly, we evaluate the squared distance from $\hat{\boldsymbol{\theta}}_N$ to $\boldsymbol{\theta}_0$ given by $N(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\hat{\boldsymbol{V}}^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)^T$. Secondly, we compare it to the chi-squared value $\chi^2_{0.95,p}$ and compute

$$CP = \frac{\#\{N(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\hat{\boldsymbol{V}}^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)^T \leq \chi^2_{0.95,p}\}}{r}. \tag{23}$$

If the squared distance from $\hat{\boldsymbol{\theta}}_N$ to $\boldsymbol{\theta}_0$ is less than the chi-squared value then the true value $\boldsymbol{\theta}_0$ lies inside the region. The CP should be close to the nominal value which is here 95%.

As a first step, the methods are compared for estimating two unknown parameters. Let $d = 10$, $T = 10$, $N = 400$, $r = 500$ and let us consider an EGP $\boldsymbol{X}$ with $A(t, \boldsymbol{\theta}) = t^{\alpha}$ and $b(t, \boldsymbol{\theta}) = t^{\beta}$. The results are presented in Tables 1 and 2. As we can observe, GMM_{MM} and GMM_{Lap} provide similar results. The CP, for all two approaches, is approximately

Table 1: Mean, standard deviation and quantiles for $A(t,\boldsymbol{\theta}) = t^{\alpha}$, $b(t,\boldsymbol{\theta}) = t^{\beta}$, $d = 10$, $T = 10$, $N = 400$, $r = 500$.

	$\hat{\alpha}$	$\hat{\beta}$
True value	2	0.5
	Mean (std)	Mean (std)
GMM_{MM}	1.9954 (0.0363)	0.4942 (0.0457)
GMM_{Lap}	1.9994 (0.0320)	0.4991 (0.0413)
	$[Q_{.025}, Q_{.975}]$	$[Q_{.025}, Q_{.975}]$
GMM_{MM}	$[1.9235, 2.0655]$	$[0.4027, 0.5839]$
GMM_{Lap}	$[1.9398, 2.0610]$	$[0.4228, 0.5793]$

Table 2: Coverage probability of hyper-ellipsoid confidence region for $A(t,\boldsymbol{\theta}) = t^{\alpha}$, $b(t,\boldsymbol{\theta}) = t^{\beta}$, $d = 10$, $T = 10$, $N = 400$, $r = 500$.

	CP
GMM_{MM}	94.8%
GMM_{Lap}	94.4%

equal to 95%. Also, the computing time of $\hat{\boldsymbol{\theta}}_N^{(2)}$ using GMM_{MM} or GMM_{Lap} is around 1s. Thus, both methods can be used to estimate two parameters.

We next present the estimation results for four unknowns. Let us consider $A(t,\boldsymbol{\theta}) = at^{\alpha}, b(t,\boldsymbol{\theta}) = bt^{\beta}$ and generate 500 sets of 500 samples paths. Then we compute the four estimates using both GMM approaches. From Table 3, it is clear that GMM_{Lap} shows better performance against GMM_{MM}. The time needed for computing the estimators using GMM_{MM} is around 1s, while it is around 7s for GMM_{Lap}. In addition, in Table 4, we see that the CP for GMM_{Lap} becomes closer to 95% by increasing N while for GMM_{MM} it tends to be less than the nominal value. GMM based on the moments seems less accurate than GMM based on the Laplace transform when we increase the number of parameters to be estimated.

6 CONCLUSION

In this work, we have dealt with parametric estimation techniques to estimate the shape and scale functions of an EGP. Two approaches dedicated for an EGP have been proposed: GMM based on the moments and GMM based on the Laplace transform. General asymptotic results for GMM for an EGP were presented and the results were illustrated for one parametric form of the shape and scale functions. Note that the results of Section 3 could be used for studying other parametric forms (we have studied another parametric form, where $A(t,\boldsymbol{\theta}) = a(1 - \exp(-\alpha t))$ and $b(t,\boldsymbol{\theta}) = b(1 - \exp(-\beta t))$, but is not presented here due to the reduced size of the paper).

As for the numerical assessment, GMM based on the Laplace transform is more performing than GMM based on the moments and particularly when we need to estimate more than two parameters.

In regard to inspections, the proposed methods are applicable when the interval time is common to all trajectories. This is not always the case in an industrial context. In this purpose, we have tested an approximate likelihood method. From Al Masry et al. (2015), an EGP can be approximated by another EGP with a piecewise constant scale function. This provides an approximation of the pdf of an increment and the likelihood function can be computed using this approximation. However, the computational time seems to be very long and we do not have any theoretical results. It could be interesting to study other parameter estimation techniques which take into consideration this problem.

Table 3: Mean, standard deviation and quantiles for $A(t,\boldsymbol{\theta}) = at^{\alpha}$, $b(t,\boldsymbol{\theta}) = bt^{\beta}$, $d = 10$, $T = 10$, $N = 500$, $r = 500$.

	True value	Method	Mean (std)	$[Q_{.025}, Q_{.975}]$
$\hat{a}$	1	GMM_{MM}	1.0402 (0.0635)	$[0.9138, 1.1725]$
		GMM_{Lap}	1.0053 (0.0501)	$[0.9159, 1.1081]$
$\hat{\alpha}$	2	GMM_{MM}	1.9897 (0.0282)	$[1.9369, 2.0478]$
		GMM_{Lap}	2.0030 (0.0247)	$[1.9534, 2.0513]$
$\hat{b}$	1	GMM_{MM}	1.0398 (0.0538)	$[0.9387, 1.1457]$
		GMM_{Lap}	1.0083 (0.0440)	$[0.9301, 1.1025]$
$\hat{\beta}$	0.5	GMM_{MM}	0.4873 (0.0299)	$[0.4286, 0.5487]$
		GMM_{Lap}	0.5023 (0.0270)	$[0.4490, 0.5582]$

Table 4: Coverage probability of hyper-ellipsoid confidence region for $A(t,\boldsymbol{\theta}) = at^{\alpha}$, $b(t,\boldsymbol{\theta}) = bt^{\beta}$, $d = 10$, $T = 10$, $r = 500$.

	CP (N=500)	CP (N=800)
GMM_{MM}	84.8%	87.6%
GMM_{Lap}	89.8%	93.4%

REFERENCES

Al Masry, Z., S. Mercier, & G. Verdier (2015). Approximate simulation techniques and distribution of an extended gamma process. *Methodol. Comput. Appl. Probab. available online*.

Carrasco, M. & J. Florens (2002). Efficient GMM estimation using the empirical characteristic function. *Working Paper, Department of Economics. University of Rochester*.

Çinlar, E. (1980). On a generalization of gamma processes. *J. Appl. Probab. 17*, 467–480.

Dykstra, R. L. & P. Laud (1981). A Bayesian nonparametric approach to reliability. *Ann. Stat. 9*(2), 356–367.

Feuerverger, A. & P. McDunnough (1981). On some Fourier methods for inference. *J. Am. Stat. Assoc. 76*(374), 379–387.

Hall, A. R. (2005). *Generalized method of moments.* Oxford University Press, Oxford, UK.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), 1029–1054.

Newey, W. K. & D. L. McFadden (1994). *Handbook of econometrics, Large sample estimation and hypothesis testing.*, Volume 4, pp. 2113–247. Elsevier Science Publishers, Amsterdam, The Netherlands.

Qin, J. & J. Lawless (1994). Empirical likelihood and general estimating equations. *Ann. Stat. 22*, 300–325.

van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliab. Eng. Syst. Safe. 94*, 2–21.