2

Gamma Processes

As already seen in Chapter 1, the Wiener process models a continuous deterioration with an increasing trend. However, it has non-monotonous trajectories and this is one major drawback of this model for the analysis of degradation measurements. Consequently, the gamma process is sometimes preferred to the Wiener process. As we will see, it is a pure jump process with an infinite number of jumps over any finite time interval. Hence, the gamma process is well suited to model some non-decreasing deterioration that accumulates over time through many tiny increments.

When not further specified, all random variables and stochastic processes are defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

2.1. Gamma distribution

We recall here a few basic facts on the gamma distribution.

DEFINITION 2.1.– *Let $a, b > 0$. A random variable X is said to be gamma distributed with shape parameter a and rate parameter b (parameters (a, b) for short) if its distribution admits the following probability density function with respect to Lebesgue measure:*

$$f_{a,b}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \mathbf{1}_{\mathbb{R}_+}(x), \text{ for all } x \in \mathbb{R},$$

where

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} \, \mathrm{d}x$$

is the gamma function.

Degradation Processes in Reliability, First Edition. Waltraud Kahle, Sophie Mercier and Christian Paroissin.

The gamma distribution with parameters (a, b) is denoted by $\mathcal{G}(a, b)$ in the following, and its p.d.f., c.d.f. and survival function by $f_{a,b}$, $F_{a,b}$ and $\bar{F}_{a,b}$, respectively. Note that another standard parameterization in the literature is (a, β) with $\beta = \frac{1}{b}$. In that case, parameter β is called *scale* parameter.

In the following, we will allow the shape parameter a to be zero and the gamma distribution $\mathcal{G}(0, b)$ will refer to the Dirac mass at point 0, namely to the distribution of the constant r.v. equal to 0.

A random variable X with gamma distribution is non-negative and its distribution is characterized by its Laplace transform $\mathcal{L}_X$, where we recall that, for any non-negative r.v. X with p.d.f. f_X, we have:

$$\mathcal{L}_X(s) = \mathbb{E}\left(e^{-sX}\right) = \int_0^{+\infty} e^{-sx} f_X(x)\, \mathrm{d}x, \text{ for all } s \geq 0.$$

PROPOSITION 2.2.– Let X be gamma distributed $\mathcal{G}(a, b)$, with $a, b > 0$. Then, its Laplace transform is:

$$\mathcal{L}_X(s) = \left(\frac{b}{b+s}\right)^a, \text{ for all } s \geq 0.$$

Also:

$$\mathbb{E}(X) = \frac{a}{b} \quad \text{and} \quad \mathrm{Var}(X) = \frac{a}{b^2}.$$

Note that the previous results are also valid for $a = 0$.

PROOF.– When $a > 0$, we have:

$$\begin{aligned}\mathcal{L}_X(s) &= \int_0^{+\infty} e^{-sx} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}\, \mathrm{d}x \\ &= \frac{b^a}{(b+s)^a} \int_0^{+\infty} f_{a,b+s}(x)\, \mathrm{d}x = \left(\frac{b}{b+s}\right)^a,\end{aligned}$$

where $f_{a,b+s}$ is the p.d.f. of $\mathcal{G}(a, b+s)$. We easily derive $\mathbb{E}(X)$ and $\mathrm{Var}(X)$ using that $\mathbb{E}(X) = -\mathcal{L}'_X(0^+)$ and $\mathbb{E}\left(X^2\right) = \mathcal{L}''_X(0^+)$. □

COROLLARY 2.3.– Let X_1, ..., X_n be independent gamma distributed r.v.s with parameters (a_1, b), ..., (a_n, b), respectively (same rate parameter). Then, $\sum_{i=1}^n X_i$ is gamma distributed $\mathcal{G}\left(\sum_{i=1}^n a_i, b\right)$.

PROOF.– Based on the independence, we have:

$$\mathcal{L}_{\sum_{i=1}^n X_i}(s) = \mathbb{E}\left(e^{-s\sum_{i=1}^n X_i}\right) = \prod_{i=1}^n \mathbb{E}\left(e^{-sX_i}\right)$$
$$= \prod_{i=1}^n \left(\frac{b}{b+s}\right)^{a_i} = \left(\frac{b}{b+s}\right)^{\sum_{i=1}^n a_i}$$

where the last expression is the Laplace transform of $\mathcal{G}\left(\sum_{i=1}^n a_i, b\right)$. □

As the gamma distribution $\mathcal{G}(1, \lambda)$ coincides with the exponential distribution $\mathcal{E}(\lambda)$ with mean $\frac{1}{\lambda}$, we directly get the following corollary.

COROLLARY 2.4.– Let X_1, ..., X_n be independent and identically exponentially distributed r.v.s with mean $\frac{1}{\lambda}$. Then, $\sum_{i=1}^n X_i$ is gamma distributed $\mathcal{G}(n, \lambda)$ with p.d.f.

$$f(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x).$$

We can also derive another basic fact from proposition 2.2.

COROLLARY 2.5.– Let X be gamma distributed $\mathcal{G}(a, b)$ and let $c > 0$. Then, cX is gamma distributed $\mathcal{G}\left(a, \frac{b}{c}\right)$.

PROOF.– For all $s \geq 0$:

$$\mathcal{L}_{cX}(s) = \mathbb{E}\left(e^{-scX}\right) = \left(\frac{b}{b+sc}\right)^a = \left(\frac{b/c}{b/c+s}\right)^a$$

which provides the result. □

Recall that the beta distribution $\mathcal{B}(a_1, a_2)$ with parameters $a_1, a_2 > 0$ admits the following p.d.f. with respect of Lebesgue measure:

$$f(x) = \frac{x^{a_1}(1-x)^{a_2}}{B(a_1, a_2)} \mathbf{1}_{[0,1]}(x),$$

where

$$B(a_1, a_2) = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\,\Gamma(a_2)}$$

is the beta function. We then have the following result.

PROPOSITION 2.6.– Let X and Y be two random variables and let $a_1, a_2, b > 0$. Then, the two following points are equivalent:

1) $X + Y$ and $\frac{X}{X+Y}$ are independent, with respective distributions $\mathcal{G}(a_1 + a_2, b)$ and $B(a_1, a_2)$;

2) X and Y are independent with respective distributions $\mathcal{G}(a_1, b)$ and $\mathcal{G}(a_2, b)$.

PROOF.– Let $(Z, U) = \left(X + Y, \frac{X}{X+Y}\right)$, so that $(X, Y) = (ZU, Z(1-U))$. We set φ to be any measurable non-negative function. Assume point 1 to be true. Then:

$$\begin{aligned}
&\mathbb{E}(\varphi(X, Y)) = \mathbb{E}(\varphi(ZU, Z(1-U))) \\
&= cst \times \int_0^{+\infty} \int_0^1 \varphi(zu, z(1-u))\, z^{a_1+a_2-1} e^{-bz} u^{a_1-1} (1-u)^{a_2-1}\, \mathrm{d}u\, \mathrm{d}z \\
&= cst \times \int_0^{+\infty} \int_0^{+\infty} \varphi(x, y)\, (x+y)^{a_1+a_2-2} e^{-b(x+y)} \\
&\qquad \times \left(\frac{x}{x+y}\right)^{a_1-1} \left(\frac{y}{x+y}\right)^{a_2-1} \mathrm{d}x\, \mathrm{d}y \\
&= cst \times \int_0^{+\infty} \int_0^{+\infty} \varphi(x, y)\, e^{-b(x+y)} x^{a_1-1} y^{a_2-1} \mathrm{d}x\, \mathrm{d}y
\end{aligned}$$

where cst stands for a constant and where we set $x = zu$ and $y = z(1-u)$ (with $\mathrm{d}u\, \mathrm{d}z = \frac{1}{x+y}\, \mathrm{d}x\, \mathrm{d}y$). This shows that the joint p.d.f. of (X, Y) is of the form

$$cst \times e^{-b(x+y)} x^{a_1-1} y^{a_2-1} = cst \times e^{-bx} x^{a_1-1} \times e^{-by} y^{a_2-1}$$

and point 2 is true.

Conversely, assume point 2 to be true. Then:

$$\begin{aligned}\mathbb{E}\left(\varphi\left(Z,U\right)\right) &= \mathbb{E}\left(\varphi\left(X+Y,\frac{X}{X+Y}\right)\right)\\ &= cst\times\int_0^{+\infty}\int_0^{+\infty}\varphi\left(x+y,\frac{x}{x+y}\right)e^{-b(x+y)}x^{a_1-1}y^{a_2-1}\mathrm{d}x\,\mathrm{d}y\\ &= cst\times\int_0^{+\infty}\int_0^{1}\varphi\left(z,u\right)e^{-bz}z^{a_1+a_2-1}u^{a_1-1}\left(1-u\right)^{a_2-1}\mathrm{d}u\,\mathrm{d}z\end{aligned}$$

where $(z,u)=\left(x+y,\frac{x}{x+y}\right)$ (with $\mathrm{d}x\,\mathrm{d}y=z\,\mathrm{d}u\,\mathrm{d}z$), which shows the result. □

We now provide some stochastic comparison results between gamma distributed r.v.s and we first recall the definition of several stochastic orders.

DEFINITION 2.7.– *Let X and Y be two random variables with c.d.f. F_X and F_Y and survival functions $\bar{F}_X$ and $\bar{F}_Y$, respectively.*

The r.v. X is said to be smaller than the r.v. Y with respect to the usual stochastic order ($X\prec_{sto}Y$) if $\bar{F}_X\leq\bar{F}_Y$.

The r.v. X is said to be smaller than the r.v. Y with respect to the hazard rate order ($X\prec_{hr}Y$) if

$$\bar{F}_X\left(y\right)\bar{F}_Y\left(x\right)\leq\bar{F}_X\left(x\right)\bar{F}_Y\left(y\right)$$

for all $y\geq x$.

The r.v. X is said to be smaller than the r.v. Y with respect to the reversed hazard rate order ($X\prec_{rhr}Y$) if

$$F_X\left(y\right)F_Y\left(x\right)\leq F_X\left(x\right)F_Y\left(y\right)$$

for all $y\geq x$.

Assume that the distributions of X and Y are absolutely continuous with respect to Lebesgue measure with p.d.f. f_X and f_Y, respectively. Then, the r.v. X is said to be smaller than the r.v. Y with respect to the likelihood ratio order if

$$f_X\left(y\right)f_Y\left(x\right)\leq f_X\left(x\right)f_Y\left(y\right)$$

for all $y\geq x$.

We recall that the likelihood ratio order implies both hazard and reversed hazard rate orders, which themselves imply the usual stochastic order, see [SHA 07] for more details on these notions.

PROPOSITION 2.8.– Let X and Y be gamma distributed r.v.s with distributions $\mathcal{G}(a,b)$ and $\mathcal{G}(a',b)$ (same rate parameter), respectively, where $a \leq a'$. Then, X is smaller than Y with respect to the likelihood ratio order, and consequently also with respect to both hazard and reversed hazard rate orders and to the usual stochastic order.

PROOF.– Let $x \leq y$. We have:

$$f_X(x) f_Y(y) - f_X(y) f_Y(x) = \frac{b^{a'}}{\Gamma(a')} \frac{b^a}{\Gamma(a)} e^{-b(x+y)} (xy)^{a-1} \left(y^{a'-a} - x^{a'-a} \right) \geq 0$$

because $a \leq a'$. □

2.2. Poisson random measures

For sake of completeness, we here introduce Poisson random measures, which are used in section 2.4 to construct series representations of gamma processes. Note that we do not intend to make a general presentation of (Poisson) random measures and we mostly restrict ourselves to the specific setting of purely atomic random measures on $\left(\mathbb{R}_+^d, \mathcal{B}\left(\mathbb{R}_+^d\right)\right)$, which is sufficient for our applicative purposes, where we only need purely atomic random measures on $\mathbb{R}_+^2$ and $\mathbb{R}_+^3$. The interested reader may for example refer to [ÇIN 11, JAC 87, NEV 77] for a more detailed account. *This section can be skipped in the first reading.*

We first recall some elementary vocabulary. Let $\lambda > 0$. A random variable X is said to be Poisson distributed with parameter λ (denoted by $X \sim \mathcal{P}(\lambda)$) if

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \text{ for all } k \in \mathbb{N}.$$

The moments of a Poisson distributed r.v. X with parameter λ are $\mathbb{E}(X) = \text{Var}(X) = \lambda$ and its Laplace transform is

$$\mathcal{L}_X(s) = \mathbb{E}\left(e^{-sX}\right) = \sum_{k=0}^{+\infty} e^{-sk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda\left(1-e^{-s}\right)}, \text{ for all } s \geq 0. \quad [2.1]$$

the random measure $M = \sum_{n\geq 1} \delta_{(U_n, \mathbf{V}_n)}$ is Poisson with intensity measure $(\alpha \otimes \mu) = (\lambda \otimes \nu)$ based on corollary 2.22. The points of M that belong to $[0,T] \times \mathbb{R}_+^d$ are now easy to simulate based on algorithm 2.10. This provides the following algorithm (assuming $A(t)$ to be one-to-one).

ALGORITHM 2.23.– Simulation of an observation $(u_i, \mathbf{v}_i)_{1\leq i\leq n}$ of the points of the Poisson random measure M on $[0,T] \times \mathbb{R}_+^d$:

– generate an observation n of N_T according to $\mathcal{P}(A(T))$;

– generate n i.i.d. observations $x_1, \cdots, x_n$ according to $\mathcal{U}([0,1])$ (independently of n);

– set $u_i = A^{-1}(A(t)\, x_i)$ for all $1 \leq i \leq n$;

– generate n i.i.d. observations $\mathbf{v}_1, \cdots, \mathbf{v}_n$ with distribution $\mu(\mathrm{d}\mathbf{v})$ (independently of n and of the u_i's).

If needed, we can also add the sorting step of the points $x_1, \cdots, x_n$ of the Poisson process $(N_t)_{t\geq 0}$ as in algorithm 2.10.

We are now ready to define and study gamma processes.

2.3. Definition and basic properties of a gamma process

DEFINITION 2.24.– *Let $A : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ be a non-decreasing and right-continuous function such that $A(0) = 0$ and let $b > 0$. Let $(X_t)_{t\geq 0}$ be a right-continuous stochastic process with left-side limits (càdlag process). Then, $(X_t)_{t\geq 0}$ is called a (non-homogeneous) gamma process with shape function $A(t)$ and rate parameter b ($\mathcal{G}(A(t), b)$ for short) if:*

– *$X_0 = 0$ almost surely;*

– *$(X_t)_{t\geq 0}$ is a process with independent increments: for all $n \in \mathbb{N}^*$, all $(t_1, \cdots, t_{n+1}) \in \mathbb{R}^{n+1}$ such that $0 \leq t_1 < t_2 < \cdots < t_{n+1}$, the increments $X_{t_{i+1}} - X_{t_i}$ with $1 \leq i \leq n$ are independent;*

– *for all $0 < s < t$, the r.v. $X_t - X_s$ is gamma distributed $\mathcal{G}(A(t) - A(s), b)$.*

In case $A(t) = at$ for all $t \geq 0$ with $a > 0$, the increment $X_t - X_s$ is gamma distributed $\mathcal{G}(a(t-s), b)$ and the process is called a homogeneous gamma process with parameters (a, b).

Let us first remark that the third point of the definition implies that, for all $0 < s < t$, the r.v. $X_t - X_s$ is non-negative. As a result, a gamma process is a *non-decreasing* process.

Also, considering $0 \leq s < t < u$, the gamma preservation property from corollary 2.3 implies that the random variable

$$X_u - X_s = (X_u - X_t) + (X_t - X_s)$$

is gamma distributed with shape parameter

$$(A(u) - A(t)) + (A(t) - A(s)) = A(u) - A(s),$$

based on the independence between $X_t - X_s$ and $X_u - X_t$ and on the fact that their respective distributions are $\mathcal{G}(A(t) - A(s), b)$ and $\mathcal{G}(A(u) - A(t), b)$. The definition of a gamma process hence is coherent with this gamma preservation property.

Finally, as will be seen in the following, a gamma process is stochastically continuous as soon as the shape function $A(t)$ is continuous. In that case, a gamma process is an additive process (see definition I.1 in the Introduction). The distribution of X_t is hence known to be infinitely divisible [SAT 99]. In the specific case of a *homogeneous* gamma process (with $A(t) = at$), the distribution of $X_t - X_s$ only depends on $t - s$ for all $0 \leq s \leq t$ and the increments are stationary. A homogeneous gamma process is hence a Lévy process. The infinitely divisibility property is here clear, writing

$$X_t = \sum_{i=1}^{n} (X_{\frac{i}{n}\,t} - X_{\frac{i-1}{n}\,t})$$

and noting that the $X_{\frac{i}{n}\,t} - X_{\frac{i-1}{n}\,t}$'s are i.i.d. gamma distributed $\mathcal{G}(a\frac{t}{n}, b)$.

Based on corollary 2.5, we get directly the following scaling result, which often allows us to concentrate on the special case $b = 1$, with no restriction.

PROPOSITION 2.25.–

1) Let $(X_t)_{t\geq 0}$ be a gamma process $\mathcal{G}(A(t), b)$ and let $c > 0$. Then, $(c\,X_t)_{t\geq 0}$ is a gamma process $\mathcal{G}\left(A(t), \frac{b}{c}\right)$.

2) Let $b > 0$ and $Y_t = bX_t$, for all $t \geq 0$. Then, the process $(X_t)_{t\geq 0}$ is a gamma process $\mathcal{G}(A(t), b)$ if and only if $(Y_t)_{t\geq 0}$ is a gamma process $\mathcal{G}(A(t), 1)$.

Based on proposition 2.8, we directly get the following comparison result, which may be, e.g. useful for the theoretical study of preventive maintenance policies.

PROPOSITION 2.26.– Let $(X_t)_{t\geq 0}$ be a gamma process $\mathcal{G}(A(t), b)$. Then, if $0 \leq t_1 \leq t_2$, the r.v. X_{t_1} is smaller than X_{t_2} in the sense of the likelihood ratio order and, consequently, in the sense of the (reversed) hazard rate and usual stochastic orders too.

The remaining of this section can be skipped in the first reading.

In the definition, we have assumed that $A : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is a non-decreasing and right-continuous function such that $A(0) = 0$. The function A is hence continuous on $\mathbb{R}^+$, except on a set $\mathcal{I} \subset \mathbb{R}^+$ which is at most countable. Following [ÇIN 80], we can set A^c and A^d to be the continuous and discontinuous parts of A, with

$$A^d(t) = \sum_{s\in\mathcal{I}, s\leq t} \Delta A(s),$$

$$A^c(t) = A(t) - A^d(t),$$

where $\Delta A(s) = A(s) - A(s^-)$.

Introducing $(X_t^c)_{t\geq 0}$ and $\left(X_t^d\right)_{t\geq 0}$ as two independent gamma processes with respective parameters $\left(A^c(t), b\right)$ and $\left(A^d(t), b\right)$ and using again corollary 2.3, the random variable X_t with distribution $\mathcal{G}\left(A^c(t) + A^d(t), b\right)$ is identically distributed as the sum $X_t^c + X_t^d$:

$$X_t \stackrel{d}{=} X_t^c + X_t^d. \qquad [2.13]$$

Let us now look at the fixed discontinuities of $\left(X_t^d\right)_{t\geq 0}$ and $(X_t^c)_{t\geq 0}$, where we recall that a right-continuous process $(Y_t)_{t\geq 0}$ is said to admit a fixed discontinuity at point $t > 0$ as soon as

$$\mathbb{P}\left(\{\omega \in \Omega : \Delta Y_t(\omega) \neq 0\}\right) > 0,$$

where $\Delta Y_t = Y_t - Y_{t^-}$.

Based on these tables, we can see that the RWA method behaves well when tested at points of its generation grid and even better than Bondesson's method in the case $A(t) = t^{1.5}$ (look at the computation times). However, when tested at points out of its generation grid, RWA method requires a longer computation time than Bondesson's method to get a similar accuracy.

As a conclusion to this section, it seems that the simple RWA method is well adapted when the point is to sample trajectories of a gamma process at fixed points known in advance. In the other case, Bondesson's (or the Rejection) method performs better.

2.6. Hitting time (time-to-failure) and overshoot

2.6.1. *Introduction*

In this section, $(X_t)_{t\geq 0}$ is a gamma process $\mathcal{G}(A(t), 1)$, where $A : \mathbb{R}_+ \to \mathbb{R}_+$ is assumed to be continuous, non-decreasing and such that $A(0) = 0$ and $a(\mathrm{d}t)$ is the only measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ such that $A(t) = a([0,t])$, for all $t \geq 0$. As a non-decreasing function, $A(t)$ is almost surely differentiable (everywhere differentiable except on a set which is at most countable, see [RUD 87]). We set $a(s) = A'(s)$ a.s. so that $a(\mathrm{d}s) = a(s)\,\mathrm{d}s$. Finally, the trajectories of $(X_t)_{t\geq 0}$ are almost surely continuous so that

$$\mathbb{P}(X_t < x) = \mathbb{P}(X_t \leq x)$$

for all $t > 0$ and all $x \geq 0$.

In case X_t represents the deterioration accumulated by a system on the time interval $[0,t]$, the system is typically considered as failed as soon the deterioration level is beyond a given failure threshold (say h). The time-to-failure of the system hence is the hitting time τ_h of the Borel set $[h, +\infty)$ by the process $(X_t)_{t\geq 0}$, with

$$\tau_h = \inf(t \geq 0 : X_t \geq h).$$

Apart from their interpretation as failure times, a lot of preventive maintenance policies are also based on hitting times, the reaching of a preventive threshold m (with $m < h$) triggering a preventive maintenance action, see, e.g. [VAN 09] for a lot of references on the subject. Hitting times

τ_h (or τ_m) are hence of major importance. As a gamma process is a pure jump process, a threshold h is crossed by a jump and the after-jump level X_{τ_h} is greater than h almost surely ($X_{\tau_h} > h$ a.s.). The exceedance $X_{\tau_h} - h$ is known as the *overshoot* of the gamma process. Note that in an applicative context, the failure may be all the more severe as the overshoot is higher. Also, even if $m < h$, both levels m and h may be crossed at the same time τ_m in case of a big jump at time τ_m, so that we may have $X_{\tau_m} = X_{\tau_h}$. Hence the interest of studying the after-jump level X_{τ_h} jointly with the jump time τ_h. We begin with providing the joint distribution of $\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)$, as well as a few marginal distributions of $\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)$, where $X_{\tau_h^-}$ stands for the left-side limit of $(X_t)_{t\geq 0}$ at time τ_h, namely for the level just before the jump. We next study aging properties of the hitting time τ_h and we end with comparing the different methods at our disposal for simulating τ_h.

Note that if $(Y_t)_{t\geq 0}$ is a gamma process $\mathcal{G}(A(t), b)$ with $b > 0$, then

$$\tilde{\tau}_h = \inf(t \geq 0 : Y_t \geq h) = \inf(t \geq 0 : X_t \geq bh)$$

where $(X_t)_{t\geq 0} = (bY_t)_{\geq 0}$ is a gamma process $\mathcal{G}(A(t), 1)$, see proposition 2.25. Hence $\tilde{\tau}_h = \tau_{bh}$ and it is enough to consider the case where $b = 1$.

2.6.2. *Theoretical results*

Let us first recall that the p.d.f., c.d.f. and survival function of $\mathcal{G}(\alpha, \beta)$ are denoted by $f_{\alpha,\beta}$, $F_{\alpha,\beta}$ and $\bar{F}_{\alpha,\beta}$, respectively.

THEOREM 2.37.– The joint p.d.f. of $\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)$ is given as

$$f_{\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)}(s, x, z) = \mathbf{1}_{\mathbb{R}_+}(s)\, \mathbf{1}_{\{0<x<h\leq z\}} f_{A(s),1}(x)\, \frac{e^{-(z-x)}}{z-x} a(s)\, \mathrm{d}s\, \mathrm{d}x\, \mathrm{d}z. \qquad [2.21]$$

PROOF.– The proof is based on [BER 96, Prop. 2 page 76] (in the homogeneous case) (see also [ÇIN 11, theorem 7.18 pp. 372]). Let φ be any non-negative and measurable function. Based on the fact that the jumps are at

most countable, we have:

$$\mathbb{E}\left(\varphi\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)\right) = \sum_{s\geq 0} \mathbb{E}\left(\varphi\left(s, X_{s^-}, X_s\right) \mathbf{1}_{\left\{X_{s^-}<h\right\}} \mathbf{1}_{\{X_s \geq h\}}\right)$$
$$= \sum_{s\geq 0} \mathbb{E}\left(\varphi\left(s, X_{s^-}, X_{s^-} + \Delta X_s\right) \mathbf{1}_{\left\{X_{s^-}<h\right\}} \mathbf{1}_{\left\{X_{s^-}+\Delta X_s \geq h\right\}}\right)$$

where $\Delta X_s = X_s - X_{s^-}$. Due to the predictability of $(X_{s^-})_{s\geq 0}$ (as a left-continuous process) and remembering theorem 2.28, the compensation formula [BER 96, page 7] applied to the jump process $(\Delta X_s)_{s\geq 0}$ provides:

$$\mathbb{E}\left(\varphi\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)\right)$$
$$= \int_{\mathbb{R}_+^2} \mathbb{E}\left(\varphi\left(s, X_{s^-}, X_{s^-} + y\right) \mathbf{1}_{\left\{X_{s^-}<h\right\}} \mathbf{1}_{\left\{X_{s^-}+y\geq h\right\}}\right) \frac{e^{-y}}{y} \mathrm{d}y\, a\left(\mathrm{d}s\right).$$

As $a\left(\mathrm{d}s\right) = a\left(s\right)\ \mathrm{d}s$ and $X_s = X_{s^-}$ almost surely with respect of Lebesgue measure, we get:

$$\mathbb{E}\left(\varphi\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)\right)$$
$$= \int_{\mathbb{R}_+^2} \mathbb{E}\left(\varphi\left(s, X_s, X_s + y\right) \mathbf{1}_{\{X_s<h\}} \mathbf{1}_{\{X_s+y\geq h\}}\right) \frac{e^{-y}}{y} \mathrm{d}y\, a\left(s\right)\ \mathrm{d}s$$
$$= \int_{\mathbb{R}_+^3} \varphi\left(s, x, x+y\right) \mathbf{1}_{\{x<h\}} \mathbf{1}_{\{x+y\geq h\}} f_{A(s),1}\left(x\right) \mathrm{d}x\, \frac{e^{-y}}{y} \mathrm{d}y\, a\left(s\right)\ \mathrm{d}s$$
$$= \int_{\mathbb{R}_+^3} \varphi\left(s, x, z\right) \mathbf{1}_{\{x<h\leq z\}} f_{A(s),1}\left(x\right) \frac{e^{-(z-x)}}{z-x}\, a\left(s\right)\ \mathrm{d}s\, \mathrm{d}x\, \mathrm{d}z$$

setting $z = x + y$ in the last line. This provides [2.21]. □

COROLLARY 2.38.– Assume that $\lim_{t\to+\infty} A\left(t\right) = \infty$. Then:

1) the joint p.d.f. of $\left(X_{\tau_h^-}, X_{\tau_h}\right)$ is *independent* of the shape function $A\left(t\right)$ and is:

$$f_{\left(X_{\tau_h^-}, X_{\tau_h}\right)}\left(x, z\right) = \mathbf{1}_{\{0<x<h\leq z\}} \frac{e^{-(z-x)}}{z-x} U\left(x\right)$$

where

$$U\left(x\right) = \int_{\mathbb{R}_+} f_{t,1}\left(x\right)\ \mathrm{d}t \qquad [2.22]$$

is the *potential function*.

2) The p.d.f. and survival functions of X_{τ_h} are

$$f_{X_{\tau_h}}(z) = \mathbf{1}_{\{h \le z\}} \int_0^h \frac{e^{-(z-x)}}{z-x} U(x)\ \mathrm{d}x,$$
$$\bar{F}_{X_{\tau_h}}(z) = \int_0^h \mathrm{Ei}(z-x)\ U(x)\ \mathrm{d}x \text{ for all } z \ge h,$$

where we recall that Ei is the exponential integral function, see [2.16].

3) The p.d.f. of τ_h is

$$f_{\tau_h}(s) = a(s)\left(\mathrm{Ei} * f_{A(s),1}\right)(h) \text{ for all } s \ge 0. \qquad [2.23]$$

PROOF.– Starting from [2.21], we have:

$$\begin{aligned} f_{\left(X_{\tau_h^-}, X_{\tau_h}\right)}(x,z) &= \int_{\mathbb{R}_+} f_{\left(\tau_h, X_{\tau_h^-}, X_{\tau_h}\right)}(s,x,z)\ \mathrm{d}s \\ &= \mathbf{1}_{\{0<x<h\le z\}} \int_{\mathbb{R}_+} f_{A(s),1}(x)\ \frac{e^{-(z-x)}}{z-x} a(s)\,\mathrm{d}s \\ &= \mathbf{1}_{\{0<x<h\le z\}} \frac{e^{-(z-x)}}{z-x} \int_{\mathbb{R}_+} f_{t,1}(x)\ \mathrm{d}t \end{aligned}$$

setting $t = A(s)$ and remembering that $A(0) = 0$ and $\lim_{t\to+\infty} A(t) = \infty$. This provides the first point. For the second point, we have:

$$f_{X_{\tau_h}}(z) = \int_{\mathbb{R}_+} f_{\left(X_{\tau_h^-}, X_{\tau_h}\right)}(x,z)\ \mathrm{d}x = \mathbf{1}_{\{h\le z\}} \int_0^h \frac{e^{-(z-x)}}{z-x} U(x)\ \mathrm{d}x$$

and, for $z \ge h$:

$$\begin{aligned} \bar{F}_{X_{\tau_h}}(z) &= \int_z^{+\infty} f_{X_{\tau_h}}(u)\ \mathrm{d}u \\ &= \int_z^{+\infty} \left(\int_0^h \frac{e^{-(u-x)}}{u-x} U(x)\ \mathrm{d}x \right) \mathrm{d}u \\ &= \int_0^h \left(\int_z^{+\infty} \frac{e^{-(u-x)}}{u-x}\ \mathrm{d}u \right) U(x)\ \mathrm{d}x \\ &= \int_0^h \mathrm{Ei}(z-x)\ U(x)\ \mathrm{d}x \end{aligned}$$

due to Fubini's theorem. The third point is similar. □

REMARK 2.39.– Note that though the distribution of τ_h depends on the shape function $A(t)$, the joint distribution of the before/after jump locations $\left(X_{\tau_h^-}, X_{\tau_h}\right)$ is *independent* of $A(t)$. Up to our knowledge, this has not been noticed before.

Before going on with more classical results from the reliability literature [VAN 09] (see also [PAR 14]), we first provide the definition of some special functions to be used in the following proposition (for further details and properties of these functions, the reader can refer to [ABR 72] or to [GRA 07]).

DEFINITION 2.40.–

– The di-gamma function ψ_0 is the logarithmic derivative of the gamma function:

$$\forall x > 0, \quad \psi_0(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

– The lower incomplete gamma function $\gamma(\cdot,\cdot)$ is defined as follows, for any $a > 0$ and $x > 0$:

$$\gamma(a,x) = \int_0^a u^{x-1} e^{-u}\, \mathrm{d}u.$$

– The upper incomplete gamma function $\Gamma(\cdot,\cdot)$ is the complementary of $\gamma(\cdot,\cdot)$ to the gamma function:

$$\Gamma(a,x) = \Gamma(a) - \gamma(a,x) = \int_a^\infty u^{x-1} e^{-u}\, \mathrm{d}u$$

for all $a, x > 0$.

– Let $p, q \in \mathbb{N}^*$. The generalized hyper-geometric function ${}_pF_q$ of order (p,q) is defined as follows, for any $z > 0$ and any positive real numbers $a_1, \ldots, a_p, b_1, \ldots, b_q$:

$${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!},$$

where $(x)_n = \Gamma(x+n)/\Gamma(x)$ is the Pochammer symbol.

PROPOSITION 2.41.– With the previous notations, for all $h > 0$, we have:

1) The survival function of τ_h is:

$$\bar{F}_{\tau_h}(t) = \mathbb{P}(X_t < h) = F_{A(t),1}(h).$$

$$f_{\tau_h}(t) = a(t) \int_h^{+\infty} \{\log(u) - \psi_0(A(t))\} f_{A(t),1}(u)\, \mathrm{d}u \quad [2.24]$$

$$= a(t)\left(\psi_0(A(t)) - \log(h)\right) \frac{\gamma(A(t),h)}{\Gamma(A(t))} + \frac{a(t)h^{A(t)}}{A(t)^2\Gamma(A(t))} \times {}_2F_2(A(t), A(t); A(t)+1, A(t)+1; -h). \quad [2.25]$$

2) The mean time to failure is:

$$\mathbb{E}(\tau_h) = \int_0^{\infty} F_{A(s),1}(h)\ \mathrm{d}s \quad [2.26]$$

and $\mathbb{E}(\tau_h) = \frac{1}{a}\int_0^h U(x)\,\mathrm{d}x$ in the homogeneous case $A(t) = at$, where $U(t)$ is the potential function given by [2.22]. Also:

$$\mathbb{E}\left(\tau_h^2\right) = \int_0^{\infty} F_{A(\sqrt{s}),1}(h)\ \mathrm{d}s.$$

PROOF.– Based on the fact that X_t is non-decreasing and right continuous, we have

$$\mathbb{P}(\tau_h > t) = \mathbb{P}(X_t < h), \quad [2.27]$$

which provides the first point.

For the second point, we start from

$$\bar{F}_{\tau_h}(t) = F_{A(t),1}(h) = \frac{\int_0^h x^{A(t)-1}e^{-x}\,\mathrm{d}x}{\Gamma(A(t))} = 1 - \frac{\int_h^{+\infty} x^{A(t)-1}e^{-x}\,\mathrm{d}x}{\Gamma(A(t))} \quad [2.28]$$

and we first check that $\Gamma(t)$ is differentiable and compute its derivative. We write

$$\Gamma(t) = \int_0^{+\infty} g(t,x)\ \mathrm{d}x$$

with

$$g(t,x) = x^{t-1}e^{-x} = e^{(t-1)\log(x)}e^{-x}.$$

We have:

$$\frac{\partial}{\partial t}g(t,x) = \log(x)\ x^{t-1}e^{-x} \text{ for all } t > 0 \text{ and all } x > 0$$

and for $0 < t_1 < t < t_2$:

$$\left|\frac{\partial}{\partial t}g(t,x)\right| \leq h(x) = |\log(x)|\left(x^{t_1-1} + x^{t_2-1}\right)e^{-x}$$

with $\int_0^{+\infty} h(x)\,\mathrm{d}x < +\infty$. Based on Lebesgue's dominated convergence theorem, we may use differentiation under the integral sign, which provides the differentiability of $\Gamma(t)$ and

$$\Gamma'(t) = \int_0^{+\infty} \frac{\partial}{\partial t}g(t,x)\ \mathrm{d}x = \int_0^{+\infty} \log(x)\ x^{t-1}e^{-x}\,\mathrm{d}x.$$

Using similar arguments as for $\Gamma(t)$, we get:

$$\frac{\partial}{\partial t}\left(\int_h^{+\infty} x^{A(t)-1}e^{-x}\,\mathrm{d}x\right) = a(t)\int_h^{+\infty} \log(x)\ x^{A(t)-1}e^{-x}\,\mathrm{d}x$$

where $a(t) = A'(t)$ a.s.

Starting again from [2.28], we now have:

$$\begin{aligned} f_{\tau_h}(t) &= \frac{a(t)}{\Gamma(A(t))} \\ &\times \left(\int_h^{+\infty} \log(x)\ x^{A(t)-1}e^{-x}\,\mathrm{d}x - \int_h^{+\infty} x^{A(t)-1}e^{-x}\,\mathrm{d}x \times \frac{\Gamma'(A(t))}{\Gamma(A(t))}\right) \\ &= \frac{a(t)}{\Gamma(A(t))}\left(\int_h^{+\infty} \{\log(x) - \psi_0(A(t))\}\, x^{A(t)-1}e^{-x}\mathrm{d}x\right). \end{aligned} \qquad [2.29]$$

For the second expression of f_{τ_h}, we start again from equation [2.29] that we write as follows:

$$f_{\tau_h}(t) = a(t)g(A(t)) \qquad [2.30]$$

with

$$g(x) = \frac{1}{\Gamma(x)}\frac{\partial}{\partial x}\Gamma(x,h) - \frac{1}{\Gamma(x)}\psi_0(x)\Gamma(x,h). \qquad [2.31]$$

From equation (6.5.12) in [ABR 72], we have

$$\Gamma(x,h) = \Gamma(x) - \gamma(x,h) = \Gamma(x) - \frac{h^x}{x}\,{}_1F_1(x;x+1;-h). \qquad [2.32]$$

Moreover, the following identity holds:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{x+k}\right) = \frac{1}{x}\left(\frac{(x)_k}{(x+1)_k} - \left(\frac{(x)_k}{(x+1)_k}\right)^2\right).$$

From the definition of the generalized hyper-geometric function ${}_1F_1$ of order $(1,1)$ and using the above equation, we have

$$\begin{aligned}\frac{\partial}{\partial x}{}_1F_1(x;x+1;-h) \;=\; & \frac{1}{x}\left[{}_1F_1(x;x+1;-h)\right. \\ & \left. -{}_2F_2(x,x;x+1,x+1;-h)\right]. \end{aligned} \qquad [2.33]$$

From equations [2.32] and [2.33], we have

$$\begin{aligned}\frac{\partial}{\partial x}\Gamma(x,h) \;=\; & \Gamma(x)(\psi_0(x) - \log(h)) \\ & + \log(h)\Gamma(x,h) + \frac{h^x}{x}{}_2F_2(x,x;x+1,x+1;-h).\end{aligned}$$

Replacing this last expression in equation [2.31], we finally have

$$g(x) = (\psi_0(x) - \log(h))\left(1 - \frac{\Gamma(x,h)}{\Gamma(x)}\right) + \frac{h^x}{x^2}{}_2F_2(x,x;x+1,x+1;-h),$$

which provides the result based on [2.30].

For the third point, we write

$$\mathbb{E}(\tau_h) = \int_0^{+\infty} \bar{F}_{\tau_h}(t)\,\mathrm{d}t = \int_0^{+\infty} F_{A(t),1}(h)\ \mathrm{d}t,$$

$$\mathbb{E}(\tau_h^2) = \int_0^{+\infty} \mathbb{P}(\tau_h^2 > t)\,\mathrm{d}t = \int_0^{+\infty} \bar{F}_{\tau_h}\left(\sqrt{t}\right)\mathrm{d}t$$

$$= \int_0^{+\infty} F_{A(\sqrt{t}),1}(h)\ \mathrm{d}t.$$

Applying Fubini's theorem to the expression of $\mathbb{E}(\tau_h)$ in the homogeneous case and setting $s = at$, we get:

$$\mathbb{E}(\tau_h) = \int_0^{+\infty} \left(\int_0^h f_{at,1}(x)\,\mathrm{d}x\right)\mathrm{d}t = \frac{1}{a}\int_0^h \left(\int_0^{+\infty} f_{s,1}(x)\,\mathrm{d}s\right)\mathrm{d}x$$

$$= \frac{1}{a}\int_0^h U(x)\mathrm{d}x.$$

□

COROLLARY 2.42.– For all $h > 0$, we have:

$$\mathbb{E}(X_{\tau_h}) = \int_0^{\infty} F_{s,1}(h)\ \mathrm{d}s = \int_0^h U(x)\mathrm{d}x$$

(independent of $A(t)$) and in the homogeneous case $A(t) = at$, we have: $\mathbb{E}(X_{\tau_h}) = a\,\mathbb{E}(\tau_h)$.

PROOF.– As the distribution of X_{τ_h} does not depend on $A(t)$ (see corollary 2.38), it is sufficient to consider the standard case $\mathcal{G}(t,1)$. Using the fact that $(X_t - t)_{t\geq 0}$ is a martingale that is uniformly integrable on each $[0,t_0]$ and that τ_h is a stopping time, we obtain that

$$\mathbb{E}\left(X_{\min(\tau_h,t_0)}\right) = \mathbb{E}(\min(\tau_h,t_0))$$

for all $t_0 \geq 0$ (see, e.g. [DOO 53] for more details on these notions). Letting $t_0 \to \infty$, we get by monotonous convergence that

$$\mathbb{E}(X_{\tau_h}) = \lim_{t_0\to\infty} \mathbb{E}\left(X_{\min(\tau_h,t_0)}\right) = \mathbb{E}(\tau_h),$$

which provides the first result due to [2.26]. The remaining is clear. □

EXAMPLE 2.43.– To better understand the behavior of τ_h and of X_{τ_h}, their p.d.f.s are plotted in Figure 2.3 in the case of a standard gamma process $\mathcal{G}(t,1)$

and $h = 2$. In this case, we have $\mathbb{E}(\tau_h) = \mathbb{E}(X_{\tau_h}) \simeq 2.50$, $\mathrm{Var}(\tau_h) \simeq 1.94$, $\mathrm{Var}(X_{\tau_h}) \simeq 0.45$ and $\mathbb{P}(X_{\tau_h} > 3) \simeq 0.15$. Based on the figure and on these results, we can see that X_{τ_h} may be largely beyond h (and it is not a rare event). Considering that $X_{\tau_h} \simeq h$ (namely, not taking the overshoot of the gamma process into account) as is sometimes done in reliability engineering may hence lead to a clear underestimation of the deterioration level at time τ_h and hence to some eventually hazardous situation.

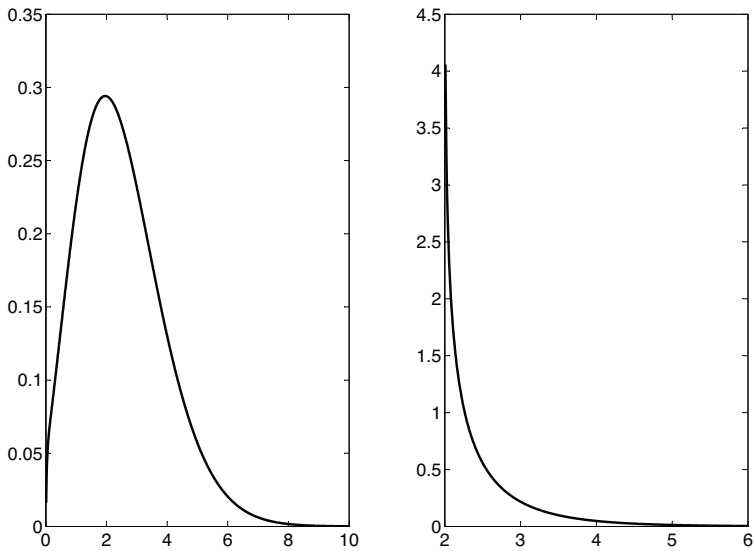

Figure 2.3. *p.d.f. of τ_h (left) and of X_{τ_h} (right) for a standard gamma process $\mathcal{G}(t, 1)$ and $h = 2$*

We next come to some *aging* properties of the hitting time τ_h, which indicates that a system with non-decreasing gamma deterioration becomes more prone to failures when it gets older. More details on these notions may be found in [SHA 07] for example. These aging properties justify the use of preventive actions to rejuvenate the system and prevent failures, when it is getting too old (or too deteriorated).

Let us consider a system with lifetime U and let U_t stand for the *remaining (residual) lifetime* of the system at time t (given that it has not failed at time

t). We recall that U_t is identically distributed as the conditional distribution of $U - t$ given that $U > t$. The survival function of U_t is

$$\bar{F}_{U_t}(u) = \mathbb{P}\left(U - t > u | U > t\right) = \frac{\bar{F}_U\left(t+u\right)}{\bar{F}_U\left(t\right)} \text{ for all } u \geq 0$$

and all t such that $\bar{F}_U\left(t\right) > 0$. Now, let us recall that a non-negative random variable U is said to have an increasing failure rate (IFR) as soon as $\frac{\bar{F}_U(t+u)}{\bar{F}_U(t)}$ is non-increasing in $t \geq 0$ (on the open support of $\bar{F}_U$), for each $u \geq 0$. Based on the previous interpretation, the IFR property means that

$$\bar{F}_{U_{t_1}}(u) \leq \bar{F}_{U_{t_2}}(u)$$

for all $t_1 \geq t_2$ and all $u \geq 0$. Equivalently, it means that the remaining lifetime of the system at time t stochastically decreases with respect to t in the sense of the usual stochastic order (see definition 2.7). If U admits a p.d.f. f_U with respect of Lebesgue measure, the IFR property is equivalent to the non-decreasingness of the failure rate function

$$h_U\left(u\right) = \begin{cases} \frac{f_U(u)}{\bar{F}_U(u)} \text{ if } \bar{F}_U\left(u\right) > 0, \\ \quad \infty \text{ elsewhere,} \end{cases}$$

hence the name.

In the same way, a non-negative random variable U is said to be new better than used (NBU) if, for all $t \geq 0$, the remaining lifetime of the system at time t is stochastically smaller than the lifetime of a new system, namely if

$$\bar{F}_{U_t}(u) = \frac{\bar{F}_U\left(t+u\right)}{\bar{F}_U\left(t\right)} \leq \bar{F}_U(u)$$

for all $u \geq 0$ and all t in the open support of $\bar{F}_U$. We recall that the IFR property implies the NBU property.

We are now ready to give aging properties of the hitting time τ_h of a gamma process. An alternate proof of the following proposition may be found in [PAR 14].

PROPOSITION 2.44.– Let $h > 0$.

1) If $A(t)$ is superadditive (namely $A(t+u) \geq A(u) + A(t)$ for all $u, t \geq 0$), then τ_h is NBU;

2) If $A(t)$ is convex, then τ_h is IFR.

PROOF.– Based on the fact that $X_t \geq 0$, $X_{t+u} - X_t \geq 0$ and the independent increments, we have:

$$\begin{aligned}\bar{F}_{\tau_h}(t+u) &= \mathbb{P}(X_t + (X_{t+u} - X_t) \leq h) \\ &\leq \mathbb{P}(\{X_t \leq h\} \cap \{X_{t+u} - X_t \leq h\}) \\ &\leq \mathbb{P}(X_t \leq h)\ \mathbb{P}(X_{t+u} - X_t \leq h) \end{aligned} \quad [2.34]$$

where $X_{t+u} - X_t$ is $\mathcal{G}(A(t+u) - A(t), 1)$ distributed. Assuming $A(t)$ to be superadditive, we have $A(t+u) - A(t) \geq A(u)$. Based on proposition 2.8, we know that $X_{t+u} - X_t \succ_{sto} X_u$ and consequently that

$$\mathbb{P}(X_{t+u} - X_t \leq h) \leq \mathbb{P}(X_u \leq h).$$

Starting again from [2.34], we obtain that $\bar{F}_{\tau_h}(t+u) \leq \bar{F}_{\tau_h}(t)\,\bar{F}_{\tau_h}(u)$. Hence τ_h is NBU, which provides the first point.

As for the second point, let us write

$$\bar{F}_{\tau_h}(t+u) = \mathbb{P}(X_{t+u} \leq h) = \mathbb{E}\left(\mathbf{1}_{\{X_t \leq h - (X_{t+u} - X_t)\}}\right).$$

Conditioning by $X_{t+u} - X_t$ and using the independent increments of $(X_t)_{t \geq 0}$, we obtain:

$$\bar{F}_{\tau_h}(t+u) = \mathbb{E}\left(F_{A(t),1}(h - (X_{t+u} - X_t))\right)$$

and

$$\frac{\bar{F}_{\tau_h}(t+u)}{\bar{F}_{\tau_h}(t)} = \frac{\mathbb{E}\left(F_{A(t),1}(h - (X_{t+u} - X_t))\right)}{F_{A(t),1}(h)}.$$

Let $t_1 < t_2$. We then have $A(t_1) \leq A(t_2)$ and $X_{t_1} \prec_{rhr} X_{t_2}$ (see proposition 2.8). We derive that

$$\frac{F_{A(t_1),1}(h - (X_{t_1+u} - X_{t_1}))}{F_{A(t_1),1}(h)} \geq \frac{F_{A(t_2),1}(h - (X_{t_1+u} - X_{t_1}))}{F_{A(t_2),1}(h)}$$

and consequently that

$$\frac{\bar{F}_{\tau_h}(t_1+u)}{\bar{F}_{\tau_h}(t_1)} = \frac{\mathbb{E}\left(F_{A(t_1),1}\left(h-\left(X_{t_1+u}-X_{t_1}\right)\right)\right)}{F_{A(t_1),1}(h)}$$
$$\geq \frac{\mathbb{E}\left(F_{A(t_2),1}\left(h-\left(X_{t_1+u}-X_{t_1}\right)\right)\right)}{F_{A(t_2),1}(h)}. \quad [2.35]$$

Assuming $A(t)$ to be convex, we have $A(t_1+u)-A(t_1) \leq A(t_2+u)-A(t_2)$ so that $X_{t_1+u}-X_{t_1} \prec_{sto} X_{t_2+u}-X_{t_2}$. As $\varphi(x) = F_{A(t_2),1}(h-x)$ is non-increasing, we derive (see [SHA 07]) that

$$\mathbb{E}\left(F_{A(t_2),1}\left(h-\left(X_{t_1+u}-X_{t_1}\right)\right)\right) \geq \mathbb{E}\left(F_{A(t_2),1}\left(h-\left(X_{t_2+u}-X_{t_2}\right)\right)\right)$$
$$= \bar{F}_{\tau_h}(t_2+u).$$

Starting again from [2.35], we now have

$$\frac{\bar{F}_{\tau_h}(t_1+u)}{\bar{F}_{\tau_h}(t_1)} \geq \frac{\bar{F}_{\tau_h}(t_2+u)}{\bar{F}_{\tau_h}(t_2)}$$

and τ_h is IFR. □

2.6.3. *Simulation*

We here simulate trajectories of a gamma process on a time interval $[0,T]$, where T is chosen such that $\mathbb{P}(\tau_h > T)$ is very small. This allows us to observe τ_h for nearly almost all simulated trajectories. An alternate method would be (1) to generate step by step the trajectory by the RWA method, (2) to test whether h has been exceeded at each step and (3) to stop as soon as h has been exceeded. However, this entails a lot of tests and cannot be implemented in a matrix way. This consequently entails longer computing times.

We take $h = 2$ and several choices for $A(t)$. We simulate 1,000 independent data sets composed each of 1,000 independent observations of τ_h. For each data set of size 1,000, we test the adequation of the observations to the distribution of τ_h with the help of a Kolmogorov–Smirnov test at the significance level $\alpha = 0.05$. For each of these 1,000 tests, we compute the p-value and the test statistic. In the following tables, we provide the proportion of times the null assumption is rejected among the 1,000 tests (the coverage probability), which should be around 0.05, as well as the mean of the 1,000 computed p-values and both the mean and standard deviation of the 1,000 test statistics. The c.p.u. times are also given.

As a first case, we consider $\mathrm{A}(t) = t$. With $h = 2$, this provides $\mathbb{E}(\tau_h) = 2.4961$ and $\mathrm{Var}(\tau_h) = 1.9375$. The trajectories are simulated on $[0, T]$ with $T = 15$, so that $\mathbb{P}(\tau_h > T) = 3.87 \times 10^{-9}$ and the probability that all the $10^3 \times 10^3 = 10^6$ trajectories reach h is 0.9961. The results are provided in Table 2.12.

As a second case, we take $\mathrm{A}(t) = t^{0.8}$. With $h = 2$, we get $\mathbb{E}(\tau_h) = 3.2896$ and $\mathrm{Var}(\tau_h) = 5.0364$. The trajectories are simulated on $[0, T]$ with $T = 30$, so that $\mathbb{P}(\tau_h > T) = 2.59 \times 10^{-9}$ and the probability that all the 10^6 trajectories reach h is 0.9974. The results are provided in Table 2.13.

		Probability of rejecting $\mathcal{H}_0$		p-value		Test statistic	
n_0	K	RWA	Bond.	RWA	Bond.	RWA	Bond.
8	5	0.1720	0.0530	0.2854	0.4779	0.0342(0.0097)	0.0279(0.0085)
9	6	0.0930	0.0440	0.3930	0.5027	0.0303(0.0084)	0.0272(0.0080)
10	7	0.0520	0.0510	0.4569	0.4864	0.0284(0.0081)	0.0277(0.0083)
11	8	0.0420	0.0560	0.4964	0.4894	0.0273(0.0079)	0.0276(0.0082)

c.p.u. times				
(n_0, K)	$(8, 5)$	$(9, 6)$	$(10, 7)$	$(11, 8)$
(RWA, Bond.)	(150, 58)	(242, 53)	(380, 63)	(645, 70)

Table 2.12. *Coverage probability, p-value, mean(std) of the test statistic and c.p.u. times; $h = 2$; cutoff value = 0.0428; $A(t) = t$*

		Probability of rejecting $\mathcal{H}_0$		p-value		Test statistic	
n_0	K	RWA	Bond.	RWA	Bond.	RWA	Bond.
8	5	0.1120	0.0450	0.3527	0.4682	0.0316(0.0087)	0.0280(0.0080)
9	6	0.0640	0.0410	0.4662	0.5047	0.0284(0.0087)	0.0272(0.0082)
10	7	0.0660	0.0480	0.4687	0.4949	0.0282(0.0086)	0.0274(0.0083)
11	8	0.0510	0.0470	0.5020	0.5140	0.0272(0.0081)	0.0270(0.0084)

c.p.u. times				
(n_0, K)	$(8, 5)$	$(9, 6)$	$(10, 7)$	$(11, 8)$
(RWA, Bond.)	(178, 59)	(293, 63)	(417, 68)	(694, 78)

Table 2.13. *Coverage probability, p-value and mean(std) of the test statistic; $h = 2$; cutoff value = 0.0428; $A(t) = t^{0.8}$*

As a third and last case, we take $\mathrm{A}(t) = t^{1.5}$. With $h = 2$, we get $\mathbb{E}(\tau_h) = 1.7721$ and $\mathrm{Var}(\tau_h) = 0.4782$. The trajectories are simulated on $[0, T]$ with $T = 6$, so that $\mathbb{P}(\tau_h > T) = 7.20 \times 10^{-9}$ and the probability that all the 10^6 trajectories reach h is 0.9928. The results are provided in Table 2.14.

		Probability of rejecting $\mathcal{H}_0$		p-value		Test statistic	
n_0	K	RWA	Bond.	RWA	Bond.	RWA	Bond.
8	5	0.1220	0.0660	0.3360	0.4938	0.0322(0.0089)	0.0276(0.0087)
9	7	0.0870	0.0460	0.4311	0.5046	0.0294(0.0089)	0.0271(0.0080)
10	9	0.0670	0.0440	0.4872	0.5188	0.0277(0.0083)	0.0268(0.0080)
11	11	0.0470	0.0560	0.4831	0.4964	0.0278(0.0083)	0.0275(0.0085)

c.p.u. times				
(n_0, K)	$(8, 5)$	$(9, 7)$	$(10, 9)$	$(11, 11)$
(RWA, Bond.)	(154, 64)	(231, 87)	(369 ,108)	(631 ,125)

Table 2.14. *Coverage probability, p-value and mean(std) of the test statistic; $h = 2$; cutoff value = 0.0428; $A(t) = t^{1.5}$*

Based on these tables, the RWA method requires much longer computing times than Bondesson's (or the Rejection) method to get a similar accuracy for the hitting time τ_h. Though the results are not provided here, the conclusion is similar for the levels reached just before ($X_{\tau_{h-}}$) or just after (X_{τ_h}) the level h is crossed.

In the case where it is desirable to observe hitting times or the corresponding deterioration levels, we consequently suggest to use Bondesson's (or the Rejection) method for simulating a gamma process.

2.7. Statistical inference in parametric models

In this section, we consider the degradation model based on a gamma process X that initiates at time t_0 with initial degradation level equal to x_0:

$$\forall t \geq 0, \quad X_t = (x_0 + \tilde{X}_{t-t_0})\mathbf{1}_{t \geq t_0}, \qquad [2.36]$$

where $\tilde{X}$ is a standard gamma process (i.e. starting from $\tilde{X}_0 = 0$ at time $t = 0$) with shape function $A(\cdot)$ and rate parameter b as defined previously.

The expectation is then equal to:

$$\mathbb{E}[X_t] = \begin{cases} 0 & \text{if } t < t_0 \\ x_0 + \frac{A(t-t_0)}{b} & \text{otherwise} \end{cases}$$

and the variance is equal to:

$$\text{Var}[X_t] = \begin{cases} 0 & \text{if } t < t_0 \\ \frac{A(t-t_0)}{b^2} & \text{otherwise} \end{cases}$$

We can notice that, in particular, we have $\mathbb{E}[X_{t_0}] = x_0$ and $\text{Var}[X_{t_0}] = 0$ since $A(0) = 0$.

For such a model, we provide estimators of the parameters and some of their asymptotic properties. The case where $t_0 = 0$ and/or $x_0 = 0$ is detailed. Two kinds of data can be distinguished: degradation measurements and failure times. We will consider three different cases: only degradation data are available, or only failure data are available or both kinds of data are available.

Here, we assume that the shape function depends on some unknown parameters $\theta \in \Theta \subseteq \mathbb{R}^p$. Thus, we will denote by $A(\cdot;\theta)$ the shape function in this framework. The two principal choices for $A(\cdot;\theta)$ are the following ones: (1) the power shape function: $A(t;\theta) = at^\gamma$ and thus $\theta = (a, \gamma)$ with $a, \gamma > 0$; (2) the exponential shape function: $A(t;\theta) = e^{at} - 1$ and thus $\theta = a > 0$. In the former case, when $\gamma = 1$ (linear shape function), it turns to be the homogeneous gamma process. In the later case, for a small value of a and at the beginning of the degradation, the "behavior" of the process is close to that of a homogeneous gamma process.

In some cases, the critical level h is an unknown parameter that should be estimated too. It follows that we will consider the problem of estimating either (t_0, x_0, θ, b) or (t_0, x_0, θ, b, h).

We now introduce some notations. We assume that we observe n independent and identically distributed copies of the stochastic process X described by equation [2.36], denoted by $(X_t^{(1)})_{t\geq 0}, \ldots, (X_t^{(n)})_{t\geq 0}$. The i-th sample path is observed m_i times at the instants $t_0 < t_{i,1} < \cdots < t_{i,m_i}$.

For any $i \in \{1, \ldots, n\}$ and any $j \in \{1, \ldots, m_i\}$, we set:

– $X_{i,j} = X_{t_{i,j}}^{(i)}$;

– $s_{i,j} = t_{i,j} - t_{i,j-1}$ with $t_{i,0} = t_0$;

– $Y_{i,j} = X_{i,j} - X_{i,j-1}$ with $X_{i,0} = x_0$;

– $\tau_h^{(i)} = \inf\{t \geq 0 \ ; \ X_t^{(i)} \geq c\}$;

– $A_{i,j}(\theta) = A(t_{i,j}; \theta)$;

– $\Delta A_{i,j}(\theta) = A(t_{i,j}; \theta) - A(t_{i,j-1}; \theta)$;

– $\mathbf{D}_{obs} = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m_i} \{(t_{i,j}, X_{i,j})\}$ (set of all the observed degradation data).

Clearly, for any $i \in \{1, \ldots, n\}$ and any $j \in \{1, \ldots, m_i\}$, $Y_{i,j}$ is gamma distributed with shape parameter $\Delta A_{i,j}(\theta)$ and rate parameter b. For a linear shape function (homogeneous gamma process), $Y_{i,j}$ is simply gamma distributed with shape parameter $as_{i,j}$ and rate parameter b. In addition, when the degradation processes are observed the same number m of times at regular instants (called, hereafter, balanced case), all increments are identically distributed. Indeed, we have $m_i = m$ for any $i \in \{1, \ldots, n\}$ and $s_{i,j} = \delta$ for any $i \in \{1, \ldots, n\}$ and for any $j \in \{1, \ldots, m\}$. Thus, $Y_{1,1}, \ldots, Y_{1,m}, \ldots, Y_{n,1}, \ldots, Y_{n,m}$ are nm independent and identically distributed gamma distributed random variables with shape function $a\delta$ and rate parameter b.

2.7.1. *Estimation based only on degradation data*

Here, we will estimate parameters of the degradation model using only degradation data. Two different approaches are considered: the MLE and the moments method estimator (MME). In this case, the parameter h cannot be estimated (this is obvious since no information on failure times is available). We will consider two different situations. The first situation corresponds to the general case where (t_0, x_0, θ, b) are unknown, while the second situation corresponds to the simpler case where only (θ, b) are unknown. For the sake of simplicity and without loss of generality, in this situation, we assume that $t_0 = 0$ and $x_0 = 0$. For all these situations, we first consider the general case and next detail the balanced case.

2.7.1.1. *Maximum likelihood estimator*

For the different models described above, we provide the likelihood function and the equations satisfied by the MLE. Moreover, in some cases, we also give asymptotic results:

– *Case 1*: the estimation of t_0 and x_0 is based on the first increments, as we did for the Wiener process (see section 1.6.1). For the parameters θ and b, all the observations are used. The likelihood function is given by

$$L_1(t_0, x_0, \theta, b|\mathbf{D}_{obs}) = \prod_{i=1}^{n} \frac{b^{A(t_{i,1};\theta)-A(t_0;\theta)}}{\Gamma(A(t_{i,1};\theta) - A(t_0;\theta))} (X_{i,1} - x_0)^{A(t_{i,1};\theta)-A(t_0;\theta)-1} e^{-b(X_{i,1}-x_0)} \times \prod_{i=1}^{n}\prod_{j=2}^{m_i} \frac{b^{\Delta A_{i,j}(\theta)}}{\Gamma(\Delta A_{i,j}(\theta))} Y_{i,j}^{\Delta A_{i,j}(\theta)-1} e^{-bY_{i,j}}. \quad [2.37]$$

The MLE can be computed numerically by maximizing the log-likelihood function:

$$\ell_1(t_0, x_0, \theta, b|\mathbf{D}_{obs}) = \log L_1(t_0, x_0, \theta, b|\mathbf{D}_{obs}).$$

Let us consider now the special case of a homogeneous gamma process: $A(t;\theta) = at$ for any $t \geq 0$. For such a model, the expression of the likelihood function can be a little simplified:

$$L_1(t_0, x_0, a, b|\mathbf{D}_{obs}) = \prod_{i=1}^{n} \frac{b^{a(t_{i,1}-t_0)}}{\Gamma(a(t_{i,1} - t_0))} \times (X_{i,1} - x_0)^{a(t_{i,1}-t_0)-1} e^{-b(X_{i,1}-x_0)} \prod_{i=1}^{n}\prod_{j=2}^{m_i} \frac{b^{as_{i,j}}}{\Gamma(as_{i,j})} Y_{i,j}^{as_{i,j}-1} e^{-bY_{i,j}} \quad [2.38]$$

and thus the log-likelihood function is given by

$$\begin{aligned}\ell_1(t_0, x_0, a, b|\mathbf{D}_{obs}) &= \sum_{i=1}^{n} \{a(t_{i,1} - t_0)\log b - \log\Gamma(a(t_{i,1} - t_0)) + (a(t_{i,1} - t_0) - 1) \\ &\log(X_{i,1} - x_0) - b(X_{i,1} - x_0)\} \\ &+ \sum_{i=1}^{n}\sum_{j=2}^{m_i} \{as_{i,j}\log b - \log\Gamma(as_{i,j}) + (as_{i,j} - 1)\log Y_{i,j} - bY_{i,j}\}.\end{aligned}$$

By differentiating the log-likelihood function ℓ_1 with respect to the four variables, the MLE $(\widehat{t_0}, \widehat{x_0}, \widehat{a}, \widehat{b})$ of (t_0, x_0, a, b) is the solution of the following

set of four equations:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} \psi_0(\widehat{a}(t_{i,1} - \widehat{t}_0)) = n \log \widehat{b} + \sum_{i=1}^{n} \log(X_{i,1} - \widehat{x}_0) & \text{[2.39a]} \\ \displaystyle\widehat{b} = \sum_{i=1}^{n} \frac{\widehat{a}(t_{i,1} - \widehat{t}_0) - 1}{X_{i,1} - \widehat{x}_0} & \text{[2.39b]} \\ \displaystyle\sum_{i=1}^{n} (t_{i,1} - \widehat{t}_0)\psi_0(\widehat{a}(t_{i,1} - \widehat{t}_0)) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} s_{i,j}\psi_0(\widehat{a}s_{i,j}) \\ \displaystyle= \sum_{i=1}^{n}\sum_{j=2}^{m_i} s_{i,j} \log Y_{i,j} \\ \displaystyle+ \log \widehat{b} \left(\sum_{i=1}^{n} (t_{i,1} - \widehat{t}_0) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} s_{i,j} \right) + \sum_{i=1}^{n} (t_{i,1} - \widehat{t}_0) \log(X_{i,1} - \widehat{x}_0) & \text{[2.39c]} \\ \displaystyle\widehat{b} = \widehat{a} \frac{\displaystyle\sum_{i=1}^{n} (t_{i,1} - \widehat{t}_0) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} s_{i,j}}{\displaystyle\sum_{i=1}^{n} (X_{i,1} - \widehat{x}_0) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} Y_{i,j}} & \text{[2.39d]} \end{cases}$$

This offers two ways to determine the MLE numerically, either by numerically maximizing the log-likelihood function ℓ_1 or by solving the above set of equations. From equations [2.39b] and [2.39d], $\widehat{a}$ can be expressed with respect of $\widehat{t}_0$ and $\widehat{x}_0$ as follows:

$$\widehat{a} \left(\sum_{i=1}^{n} \frac{t_{i,1} - \widehat{t}_0}{X_{i,1} - \widehat{x}_0} - \frac{\displaystyle\sum_{i=1}^{n} (t_{i,1} - \widehat{t}_0) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} s_{i,j}}{\displaystyle\sum_{i=1}^{n} (X_{i,1} - \widehat{x}_0) + \sum_{i=1}^{n}\sum_{j=2}^{m_i} Y_{i,j}} \right) = \sum_{i=1}^{n} \frac{1}{X_{i,1} - \widehat{x}_0}. \quad \text{[2.40]}$$

Next, from equations [2.40] and [2.39b] (or equivalently [2.39d]), we can also express $\widehat{b}$ as a function of $\widehat{t}_0$ and $\widehat{x}_0$. Finally, replacing these expressions for $\widehat{a}$ and $\widehat{b}$ in equations [2.39a] and [2.39c], we obtain two equations depending only on $\widehat{t}_0$ and $\widehat{x}_0$. However, these equations cannot be simplified anymore and the solution cannot be easily determined. Thus, in practice, it is often easier to numerically maximize the log-likelihood function ℓ_1.

– *Case 2*: we now consider the situation where $t_0 = 0$ and $x_0 = 0$. The MLE $(\widehat{\theta}, \widehat{b})$ of (θ, b) can be obtained by maximizing the following likelihood function:

$$L_1(\theta, b|\mathbf{D}_{obs}) = \prod_{i=1}^{n}\prod_{j=1}^{m_i} \frac{b^{\Delta A_{i,j}(\theta)}}{\Gamma(\Delta A_{i,j}(\theta))} Y_{i,j}^{\Delta A_{i,j}(\theta)-1} e^{-bY_{i,j}} \quad [2.41]$$

and the log-likelihood function is given as (up to an additive constant):

$$\ell_1(\theta, b|\mathbf{D}_{obs}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i} (\Delta A_{i,j}(\theta)\log(b) - \log\Gamma(\Delta A_{i,j}(\theta)) + \Delta A_{i,j}(\theta)\log(Y_{i,j}) - bY_{i,j}). \quad [2.42]$$

When the shape function is linear, $(\widehat{a}, \widehat{b})$ satisfies the following equations:

$$\begin{cases} \widehat{b} = \widehat{a}\dfrac{\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}}{\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m_i} Y_{i,j}} & [2.43a] \\ \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}\psi_0(\widehat{a}s_{i,j}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}\log(\widehat{b}Y_{i,j}) & [2.43b] \end{cases}$$

Substituting the expression of $\widehat{b}$ from equation [2.43a] into equation [2.43b], we get

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}\left(\psi_0(\widehat{a}s_{i,j}) - \log\widehat{a}\right) = \left(\sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}\right)\log\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{m_i} Y_{i,j}}\right) + \sum_{i=1}^{n}\sum_{j=1}^{m_i} s_{i,j}\log Y_{i,j}. \quad [2.44]$$

Hence, the estimator $\widehat{a}$ can be numerically computed by solving this last equation and $\widehat{b}$ can next be computed using equation [2.43a].

Now, let us consider the balanced case where $m_i = m$ for any $i \in \{1, \ldots, n\}$ and $s_{i,j} = \delta$ for any $i \in \{1, \ldots, n\}$ and any $j \in \{1, \ldots, m\}$.

As noticed earlier, this situation is the basic one in statistical inference since we deal with i.i.d. random variables. In such a case, equation [2.44] can be written in terms of the empirical arithmetic and geometric averages of the $Y_{i,j}$'s as:

$$\begin{cases} \psi_0(\widehat{a}\delta) - \log(\widehat{a}\delta) = \overline{\log Y} - \log \overline{Y} & \text{[2.45a]} \\ \widehat{b} = \dfrac{\widehat{a}\delta}{\overline{Y}}, & \text{[2.45b]} \end{cases}$$

where

$$\overline{Y} = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{i,j} \quad \text{and} \quad \overline{\log Y} = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} \log Y_{i,j}.$$

Moreover, the MLE is known to be asymptotically normal (see [KOT 06], for instance). More specifically, we have:

$$\sqrt{nm}\begin{pmatrix} \widehat{a} - a \\ \widehat{b} - b \end{pmatrix} \xrightarrow[nm\to\infty]{d} \mathcal{N}(0, \Sigma)$$

where

$$\Sigma = \frac{1}{a\psi_1(a) - 1}\begin{pmatrix} \dfrac{a}{\delta^2} & -\dfrac{b}{\delta} \\ -\dfrac{b}{\delta} & b^2\psi_1(a) \end{pmatrix}$$

and where ψ_1 is the tri-gamma function defined as the first derivative of the di-gamma function.

2.7.1.2. *Moments method estimator*

Here, we consider the estimator based on the moments method, only for the case of a homogeneous gamma process. The first increments are used to estimate t_0 and x_0, after estimating a and b using the other increments. When $t_0 \neq 0$ and $x_0 \neq 0$, we assume that $t_{i,1} = t_1$ for any $i \in \{1, \ldots, n\}$, meaning that the first observation times is the same for all units (this particular situation has been also considered when looking at the inference for the Wiener process, see section 1.6.1).

Let $\mu_{(1)} = \mathbb{E}[X_{t_1}^{(i)}]$ and $\sigma^2_{(1)} = \text{Var}[X_{t_1}^{(i)}]$, for any $i \in \{1, \ldots, n\}$. We have:

$$\mu_{(1)} = \frac{a}{b}(t_1 - t_0) + x_0 \quad \text{and} \quad \sigma^2_{(1)} = \frac{a}{b^2}(t_1 - t_0).$$

Let also

$$\mu_{(2)} = a/b = \mathbb{E}[Y_{i,j}]/s_{i,j} \quad \text{and} \quad \sigma^2_{(2)} = a/b^2 = \text{Var}[Y_{i,j}]/s_{i,j},$$

where both quantities are independent of $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m_i\}$.

Substituting $\mu_{(1)}$, $\sigma^2_{(1)}$, $\mu_{(2)}$ and $\sigma^2_{(2)}$ by their empirical estimators in the previous equations provides a set of four equations that are easily solved. We get:

$$\begin{cases} \widehat{t_0} = t_1 - \widehat{\sigma}^2_{(1)}/\widehat{\sigma}^2_{(2)} \\ \widehat{x}_0 = \widehat{\mu}_{(1)} - \widehat{\mu}_{(2)}\widehat{\sigma}^2_{(1)}/\widehat{\sigma}^2_{(2)} \\ \widehat{a} = \widehat{\mu}^2_{(2)}/\widehat{\sigma}^2_{(2)} \\ \widehat{b} = \widehat{\mu}_{(2)}/\widehat{\sigma}^2_{(2)} \end{cases}$$

where

$$\begin{cases} \widehat{\mu}_{(1)} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_{i,1} \\ \widehat{\sigma}^2_{(1)} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_{i,1} - \widehat{\mu}_{(1)})^2 \\ \widehat{\mu}_{(2)} = \dfrac{1}{n(\bar{m}-1)}\displaystyle\sum_{i=1}^{n}\sum_{j=2}^{m_i}\frac{Y_{i,j}}{s_{i,j}} \\ \widehat{\sigma}^2_{(2)} = \dfrac{1}{n(\bar{m}-1)}\displaystyle\sum_{i=1}^{n}\sum_{j=2}^{m_i}\frac{(Y_{i,j} - s_{i,j}\widehat{\mu}_{(2)})^2}{s_{i,j}} \end{cases}$$

where $\bar{m} = \frac{1}{n}\sum_{i=1}^{n} m_i$ (see [BOR 15] where similar empirical quantities are introduced for a related model). This leads to asymptotically unbiased estimators that may however present some bias, when considering finite sample size. When the number n of individuals is "small", we could

consequently prefer unbiased estimators of $\sigma^2_{(1)}$ and $\sigma^2_{(2)}$, which are the following one:

$$\begin{cases} \widetilde{\sigma}^2_{(1)} = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_{i,1} - \widehat{\mu}_{(1)})^2 \\ \widetilde{\sigma}^2_{(2)} = \dfrac{1}{n(\bar{m}-1)-1}\displaystyle\sum_{i=1}^{n}\sum_{j=2}^{m_i}\frac{(Y_{i,j} - s_{i,j}\widehat{\mu}_{(2)})^2}{s_{i,j}} \end{cases}$$

Of course, $\widehat{\sigma}^2_{(i)}$ and $\widetilde{\sigma}^2_{(i)}$ are asymptotically equivalent for $i \in \{1, 2\}$. Below, we go on with $\widehat{\sigma}^2_{(1)}$ and $\widehat{\sigma}^2_{(2)}$.

One can notice that $(\widehat{\mu}_{(1)}, \widehat{\sigma}^2_{(1)})$ and $(\widehat{\mu}_{(2)}, \widehat{\sigma}^2_{(2)})$ are two independent random vectors since they are defined on different independent increments. The following theorem states the asymptotic normality of $(\widehat{\mu}_{(1)}, \widehat{\sigma}^2_{(1)})$, which is given without proof since it is a classical result.

THEOREM 2.45.– Estimators $\widehat{\mu}_{(1)}$ and $\widehat{\sigma}^2_{(1)}$ are asymptotically normal:

$$\sqrt{n}\begin{pmatrix} \widehat{\mu}_{(1)} - \mu_{(1)} \\ \widehat{\sigma}^2_{(1)} - \sigma^2_{(1)} \end{pmatrix} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \Sigma_1)$$

where

$$\Sigma_1 = \begin{pmatrix} \dfrac{a}{b^2} & \dfrac{a(a+1)(a+2)}{b^3} \\ \dfrac{a(a+1)(a+2)}{b^3} & \dfrac{a((a+1)(a+2)(a+3)-a)}{b^4}. \end{pmatrix}$$

We now provide asymptotic results for $(\widehat{\mu}_{(2)}, \widehat{\sigma}^2_{(2)})$. Proofs are similar to that of Bordes *et al.* [BOR 15, theorems 2.1 and 2.3] devoted to a close degradation model and they are hence not given here. We begin with a strong consistency result.

THEOREM 2.46.– Under the following assumptions:

– (H_1) $\displaystyle\sum_{n\geq 1}\sum_{j=2}^{m_n} s^{-1}_{n,j}\left(\sum_{i=1}^{n} m_i - n\right)^{-2} < \infty$;

– (H_2) $\exists \delta_u, \forall i \in \mathbb{N}^\star; \forall j \in \{2, \ldots, m_i\}$, $s_{i,j} \leq \delta_u$,