Supplement to ”Semiparametric Estimation of Gamma Processes for Deteriorating Products”

Z.S. YE[†], M. XIE[†‡], L.C. TANG[†], and N. CHEN[†]

[†]Department of Industrial & Systems Engineering, National University of Singapore

[‡]Department of Systems Engineering & Engineering Management, City University of Hong Kong

November 13, 2013

1 Introduction

This supplement provides the proof of Theorem 1. We also establish the consistency of the MLEs of the simple gamma-process model and the gamma-process model with random effects.

2 Proof of Theorem 1

Consider Theorem 1(a). It is easy to see that for $1 \leq l \leq h$,

$$E\left[\ln \Delta Y_{j+l}|\mathbf{D}_{obs}, \Theta\right] = E\left\{E\left[\ln \Delta Y_{j+l}|\mathbf{D}_{obs}, \Theta, \gamma\right]\right\}. \tag{S.1}$$

Therefore, we can compute $E[\ln \Delta Y_{j+l}|\mathbf{D}_{obs}, \Theta, \gamma]$ first. When $l = 1$, the conditional PDF for ΔY_{j+1} given the observed data $\mathbf{D}_{obs}$ and the random effect term γ is

$$f(y_{j+1}|\mathbf{D}_{obs}, \Theta, \gamma) \propto \left(\frac{\Delta Y_{j+1}}{Y(T_{j+h}) - Y(T_j)}\right)^{\eta(T_{j+1})-\eta(T_j)-1} \left(1 - \frac{\Delta Y_{j+1}}{Y(T_{j+h}) - Y(T_j)}\right)^{\eta(T_{j+h})-\eta(T_{j+1})-1}$$

This relationship implies that the variable X_{beta} follows a beta distribution as

$$X_{beta} = \left[\frac{\Delta Y_{j+1}}{Y(T_{j+h}) - Y(T_j)}\middle| \mathbf{D}_{obs}, \Theta, \gamma\right] \sim \text{beta}(\eta(T_{j+1}) - \eta(T_j), \eta(T_{j+h}) - \eta(T_{j+1})).$$

The moment generating function for $\log X_{beta}$ can be shown to be

$$M_{\log X_{beta}}(t) = \frac{B\left(\Delta\eta_{j+1} + t, \eta(T_{j+h}) - \eta(T_{j+1})\right)}{B\left(\Delta\eta_{j+1}, \eta(T_{j+h}) - \eta(T_{j+1})\right)}, \tag{S.2}$$

where $B(\cdot, \cdot)$ is the beta function defined by

$$B(x, y) = \int_0^1 u^{x-1}(1-u)^{y-1} du.$$

The expectation of $\ln X_{beta}$ can be readily obtained based on (S.2) as

$$E[\ln X_{beta}] = \psi(\Delta\eta_{j+1}) - \psi\left(\eta(T_{j+h}) - \eta(T_j)\right).$$

Note that $E[\log \Delta Y_{j+1}|\mathbf{D}_{obs}, \Theta, \gamma]$ can be written as

$$E[\log \Delta Y_{j+1}|\, \mathbf{D}_{obs}, \Theta, \gamma] = E[\log X_{beta}] + \log\left[Y(T_{j+h}) - Y(T_j)\right]. \tag{S.3}$$

Substituting (S.3) into (S.1) yields (11). Similarly, we can verify the case when $l = h$.

Next, we proceed to verify the case when $1 < l < h$. consider the conditional joint PDF for $\Delta Y_{j+1}, \cdots, \Delta Y_{j+h}$ given the observed data $\mathbf{D}_{obs}$ and the random effect term γ. When

$\Delta Y_{j+i} \geq 0$ for $1 \leq i \leq h$ and $\sum_{i=1}^{h} \Delta Y_{j+i} = Y(T_{j+h}) - Y(T_j)$, the joint PDF is

$$f(\Delta Y_{j+1}, ..., \Delta Y_{j+h} | \mathbf{D}_{obs}, \Theta, \gamma) \propto \prod_{i=1}^{h} \Delta Y_{j+i}^{\Delta \eta_{j+i} - 1}.$$

Integrating $\Delta Y_{j+1}, \cdots, \Delta Y_{j+l-1}, \Delta Y_{j+l+1}, \cdots, \Delta Y_{j+h}$ out of $f(\Delta Y_{j+1}, \cdots, \Delta Y_{j+h} | \mathbf{D}_{obs}, \Theta, \gamma)$ yields the marginal distribution of ΔY_{j+l} given $\mathbf{D}_{obs}$ and γ as

$$f(\Delta Y_{j+l} | \mathbf{D}_{obs}, \Theta, \gamma) \propto \left(\frac{\Delta Y_{j+l}}{Y(T_j) - Y(T_{j-h})} \right)^{\Delta \eta_{j+l} - 1} \left(1 - \frac{\Delta Y_{j+l}}{Y(T_j) - Y(T_{j-h})} \right)^{\eta(T_j + h) - \eta(T_j) - \Delta \eta_{j+l} - 1}.$$

It is immediately seen from the above relation that given $\mathbf{D}_{obs}$ and γ, $\Delta Y_{j+l} / \left[Y(T_j) - Y(T_{j+h}) \right]$ follows a beta distribution with the shape parameters $(\Delta \eta_{j+l}, \eta\,(T_{j+h}) - \eta(T_j) - \Delta \eta_{j+l})$. Parallel to the above calculation for $E\left[\log \Delta Y_{j+1} | \mathbf{D}_{obs}, \gamma\right]$, we can reach (11).

At last, consider Theorem 1(b). Similar to the computation of $\upsilon_{i,N}$, it can be shown that for $j < l \leq m$,

$$E[\ln \Delta Y_l | \mathbf{D}_{obs}, \Theta, \gamma] = \psi(\Delta \eta_l) - \ln \gamma. \tag{S.4}$$

Taking the expectation of Equation (S.4) with respect to γ yields the desired results.

3 Consistency of the MLE

Consider a unit that is inspected J times with inspection epochs $\underline{\boldsymbol{T}} = \{0 = T_0, T_1, \cdots, T_J\}$ and the associated degradation levels $\underline{\boldsymbol{Y}} = \{0 = Y_0, Y_1, \cdots, Y_J\}$. Suppose that J is a positive and integer-valued random variable with $E[J] < \infty$ and that $\underline{\boldsymbol{T}}$ is a random vector with $0 = T_0 < T_1 < \cdots < T_J$ taking values in the bounded set $[0, \tau]$ where $\tau < \infty$. Further assume J and $\underline{\boldsymbol{T}}$ are independent. Suppose n testing units are observed. Let $\hat{\Theta}_n$ and $\hat{\Theta}_n^{sg}$ be the ML estimators under the gamma process with random effects model and the simple gamma process model, respectively. To prove the consistency, we define some measures similar to Wellner and Zhang (2000).

Define $\mathcal{B}$ as the collection of Borel sets in $\mathcal{R}$ and let $\mathcal{B}_{[0,\tau]} = \{B \cap [0,\tau] : B \in \mathcal{B}\}$. For any $B \in \mathcal{B}_{[0,\tau]}$, define the measure $\hbar$ on the measurable space $\big([0,\tau], \mathcal{B}_{[0,\tau]}\big)$ as

$$\hbar(B) = \sum_{l=1}^{\infty} P(J=l) \sum_{j=1}^{l} P(T_j \in B | J = l) = E_{J,\underline{\boldsymbol{T}}}\left[\sum_{j=1}^{J} \mathbf{1}_B(T_j)\right],$$

Further define the L_2 metric $d_1(\Theta_1, \Theta_2)$ for the random effects model as

$$d_1^2(\Theta_1, \Theta_2) = (k_1 - k_2)^2 + (\lambda_1 - \lambda_2)^2 + \int [\eta_1(x) - \eta_2(x)]^2 \, d\hbar(x).$$

Similarly, define $d_2(\Theta_1^{sg}, \Theta_2^{sg})$ for the simple model as

$$d_2^2(\Theta_1^{sg}, \Theta_2^{sg}) = (\mu_1 - \mu_2)^2 + \int [\eta_1(x) - \eta_2(x)]^2 \, d\hbar(x).$$

Let $\Theta_0 = (k_0, \lambda_0, \eta_0(\cdot))$ and $\Theta_0^{sg} = (\gamma_0, \eta_0(\cdot))$ be the true values of the parameters for the random effects model and the simple model, respectively. Theorem 3 establishes consistency of the ML estimators with respect to $\hbar$ for both models. It should be pointed out that based on the definition of d_1 and d_2, the consistency for $\eta(t)$ is only meaningful when there is a positive density/mass of observation at time t, i.e., $\hbar([t-\varepsilon, t+\varepsilon]) > 0$ for any $\varepsilon > 0$.

Theorem 3.

(a) Suppose the underlying degradation follows a gamma process with random effects. If $E[J] < \infty, \eta_0(\tau) < \infty$, and (k_0, λ_0) is in the interior of $\mathcal{R}_+^2$, then for every $b \leq \tau$ with $\hbar([b,\tau]) > 0$, we have

$$d_1\left((\hat{\eta}_n \mathbf{1}_{[0,b]}, \hat{k}_n, \hat{\lambda}_n), (\eta_0 \mathbf{1}_{[0,b]}, k_0, \lambda_0)\right) \xrightarrow{\text{a.s.}} 0 \quad \text{as } n \to \infty.$$

(b) Suppose the underlying degradation follows a simple gamma process. If μ_0 is in the interior of $\mathcal{R}_+, E[J] < \infty$, and $\eta_0(\tau) < \infty$, then for every $b \leq \tau$ with $\hbar([b,\tau]) > 0$, we have

$$d_2\left((\hat{\eta}_n \mathbf{1}_{[0,b]}, \hat{\mu}_n), (\eta_0 \mathbf{1}_{[0,b]}, \mu_0)\right) \xrightarrow{\text{a.s.}} 0 \quad \text{as } n \to \infty.$$

Proof: We shall only focus on Theorem 3(a). Theorem 3(b) can be proved in a similar vein. Proof of this theorem can be established similar to that for the pseudo-likelihood estimation method developed in Wang (2008). The empirical process theory is used here. Denote by $\mathbb{P}_n$ the empirical measure and

$$m_\Theta(\underline{\boldsymbol{Y}}) = \ln \frac{\lambda^k \Gamma(\eta_J + k)}{\Gamma(k)} - (\eta_J + k)\ln(\lambda + Y_J) + \sum_{j=1}^{J} [(\Delta\eta_j - 1)\ln \Delta Y_j - \ln \Gamma(\Delta\eta_j)],$$

where $\eta_J = \eta(T_J), \Delta\eta_j = \eta(T_j) - \eta(T_{j-1}), Y_J = Y(T_J)$ and $\Delta Y_j = Y(T_j) - Y(T_{j-1})$. Let $\mathbb{M}_n(\Theta) = n^{-1} l(\Theta; \mathbf{D}_{obs}) = \mathbb{P}_n m_\Theta(\underline{\boldsymbol{Y}})$ and $\mathbb{M}(\Theta) = \mathbb{P} m_\Theta(\underline{\boldsymbol{Y}})$, where $l(\Theta; \mathbf{D}_{obs})$ is given by Equation (6) in the original paper. The basic idea is to show that $\{\hat{\Theta}_n\}$ is sequentially compact a.e. and that every pointwise limit $\Theta^\dagger$ of $\{\hat{\Theta}_n\}$ satisfies $\mathbb{M}(\Theta_0) \le \mathbb{M}(\Theta^\dagger)$ and $\mathbb{M}(\Theta_0) \ge \mathbb{M}(\Theta^\dagger)$. If we can further show that Θ_0 is the unique maximum of $\mathbb{M}(\Theta)$, then $\Theta_0 = \Theta^\dagger$ (Wellner and Zhang 2000).

The proof of $\mathbb{M}(\Theta_0) \le \mathbb{M}(\Theta^\dagger)$ invokes the Helly's selection theorem and the one-sided Glivenko-Cantelli theorem (Wellner and Zhang 2000, Thm A.1). Therefore, we need to show that $m_\Theta(\underline{\boldsymbol{Y}})$ is upper bounded by an integrable function and that $\{\hat{\Theta}_n\}$ is uniformly bounded. First, we will find an envelope function for $m_\Theta(\underline{\boldsymbol{Y}})$ as follows.

$$\begin{aligned} m_\Theta(\underline{\boldsymbol{Y}}) &\le \ln \frac{\lambda^k \Gamma(\eta_J + k)}{\Gamma(k)} - (\eta_J + k)\ln(\lambda + Y_J) + \sum_{j=1}^{J} [(\Delta\eta_j - 1)\ln Y_J - \ln \Gamma(\eta_J)] \\ &< \ln \frac{\lambda^k \Gamma(\eta_J + k)}{\Gamma(k)} + (\eta_J + J)\ln Y_J + C \equiv \mathbb{M}_0(\underline{\boldsymbol{Y}}), \end{aligned} \tag{S.5}$$

where C is a constant that changes from line to line. $\mathbb{M}_0(\underline{\boldsymbol{Y}})$ is integrable because

$$E[J] < \infty \quad \text{and} \quad E[\ln Y_J] \le E[\ln Y(\tau)] = \ln \lambda_0 - \psi(k_0) + \psi(\eta_0(\tau) + k_0).$$

Next, we show $\hat{\Theta}_n$ is bounded. Let $\Theta = (\eta, 0, 0)$. It follows that

$$\begin{aligned} 0 &\leq \lim_{\varepsilon \downarrow 0} \frac{\mathbb{M}_n(\hat{\Theta}_n) - \mathbb{M}_n(\varepsilon\Theta + (1-\varepsilon)\hat{\Theta}_n)}{\varepsilon} \\ &= \mathbb{P}_n \sum_{j=1}^{J} \left[(\Delta\hat{\eta}_j - \Delta\eta_j) \left(-\ln \frac{\lambda + Y_J}{y_j} + \psi(\hat{\eta}_J + k) - \psi(\Delta\hat{\eta}_j) \right) \right]. \end{aligned} \tag{S.6}$$

Therefore, we have

$$\mathbb{P}_n \sum_{j=1}^{J} \Delta\hat{\eta}_j \log \frac{\lambda + Y_J}{\Delta Y_j} \leq \mathbb{P}_n \sum_{j=1}^{J} \left[\Delta\eta_j \ln \frac{\lambda + Y_J}{\Delta Y_j} \right] + \mathbb{P}_n \sum_{j=1}^{J} \left[(\Delta\hat{\eta}_j - \Delta\eta_j)(\psi(\hat{\eta}_J + k) - \psi(\Delta\hat{\eta}_j)) \right].$$

The first term is obviously bounded, while the second term on the right-hand side is upper bounded because

$$\psi(\hat{\eta}_J + k) - \psi(\Delta\hat{\eta}_j) = O(1/\Delta\hat{\eta}_j).$$

By the strong law of large numbers, we can see that $\hat{\eta}_J$ is upper bounded. On the other hand,

$$\begin{aligned} \limsup_{n\to\infty} \mathbb{P}_n\{\hat{\eta}_J\} &\geq \limsup_{n\to\infty} \mathbb{P}_n\{\hat{\eta}_J \mathbf{1}_{[b,\tau]}(T_J)\} \\ &\geq \limsup_{n\to\infty} \hat{\eta}(b) P_n\{\mathbf{1}_{[b,\tau]}(T_J)\} \geq \hbar([b,\tau]) \limsup_{n\to\infty} \hat{\eta}(b). \end{aligned} \tag{S.7}$$

Therefore, $\limsup_{n\to\infty} \hat{\eta}(b) \leq C/\hbar([b,\tau])$ and thus $\hat{\eta}(t)$ is uniformly bounded a.s. for $t \in [0, b]$ if $\hbar([b,\tau]) > 0$. Then by the Helly's selection theorem, $\{\hat{\Theta}_n\}$ has a sequence $\{\hat{\Theta}_{n'}\}$ converging to some point $\Theta^\dagger = (\eta^\dagger, k^\dagger, \lambda^\dagger)$ where $\eta^\dagger$ is a non-negative and increasing function on $[0, b]$. Consider the class of functions

$$\mathcal{T}_b = \{m_\Theta(\underline{\boldsymbol{Y}}) : (k, \lambda) \in \mathcal{R}^2, \eta \in \mathcal{H}_b\},$$

where $\mathcal{H}_b = \{\eta \in \mathcal{H} : \eta(b) \leq 1 + C/\hbar([b,\tau])\}$ is compact under the metric d_1. Since $m_\Theta(\underline{\boldsymbol{Y}})$

has an integrable envelope, an application of the Glivenko-Cantelli theorem yields that

$$\limsup_{n\to\infty} \sup_{\Theta: m_\Theta(\underline{\boldsymbol{Y}})\in\mathcal{T}_b} (\mathbb{P}_n - \mathbb{P}) m_\Theta(\underline{\boldsymbol{Y}}) \le 0 \quad \text{a.s.} \tag{S.8}$$

Furthermore, we have $\mathbb{M}_n(\hat{\Theta}_n) \ge \mathbb{M}(\Theta_0)$ and $\mathbb{M}_n(\Theta_0) \to \mathbb{M}(\Theta_0)$ a.s. by the strong law of large numbers. Therefore, $\mathbb{M}(\Theta_0) \le \liminf_{n\to\infty} \mathbb{M}_n(\hat{\Theta}_n)$ a.s. and $\limsup_{n'\to\infty} \mathbb{M}_{n'}(\hat{\Theta}_{n'}) \le \mathbb{M}(\Theta^\dagger)$ because of the semi-continuity of $\mathbb{M}(\Theta)$ on Θ.

Next, we proceed to show the uniqueness of the maximum. We need to show that for every Θ in the parameter space, $\mathbb{M}(\Theta_0) - M(\Theta) \ge 0$. This can be done by first conditional on $(J, \underline{\boldsymbol{T}})$ and then taking the unconditional expectation, which yields

$$\begin{aligned} \mathbb{M}(\Theta_0) - \mathbb{M}(\Theta) &= \int E_{\Theta_0}\left[f(\underline{\boldsymbol{Y}}|\Theta_0, J, \underline{\boldsymbol{T}}) - f(\underline{\boldsymbol{Y}}|\Theta, J, \underline{\boldsymbol{T}})\right] du(J, \underline{\boldsymbol{T}}) \\ &= -\int E_{\Theta_0}\left[\ln \frac{f(\underline{\boldsymbol{Y}}|\Theta, J, \underline{\boldsymbol{T}})}{f(\underline{\boldsymbol{Y}}|\Theta_0, J, \underline{\boldsymbol{T}})}\right] du(J, \underline{\boldsymbol{T}}) \\ &\ge -\int \log E_{\Theta_0}\left[\frac{f(\underline{\boldsymbol{Y}}|\Theta, J, \underline{\boldsymbol{T}})}{f(\underline{\boldsymbol{Y}}|\Theta_0, J, \underline{\boldsymbol{T}})}\right] = 0. \end{aligned} \tag{S.9}$$

The inequality follows from the strict concavity of $\log(x)$ and the Jensen's inequality. It is easy to see that when $\Theta \ne \Theta_0$, $f(\underline{\boldsymbol{Y}}|\Theta, J, \underline{\boldsymbol{T}}) \ne (\underline{\boldsymbol{Y}}|\Theta_0, J, \underline{\boldsymbol{T}})$. Therefore, the equality holds only when $\Theta = \Theta_0$. This means that Θ_0 is the unique maximum of $M(\Theta)$. Therefore, $\Theta^\dagger = \Theta_0$ a.s., and hence the consistency of $\hat{\Theta}_n$ under the d_1 matrix follows from the dominated convergence theorem.

References

Wang, X. (2008), "A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects," *Statistica Sinica*, 18(3), 1153–1163.

Wellner, J. A. and Zhang, Y. (2000), "Two estimators of the mean of a counting process with panel count data," *The Annals of Statistics*, 28(3), 779–814.