Ann Inst Stat Math (2012) 64:991–1007
DOI 10.1007/s10463-011-0344-7

Some problems in nonparametric inference for the stress release process related to the local time

Takayuki Fujii · Yoichi Nishiyama

Received: 5 November 2010 / Revised: 31 May 2011 / Published online: 11 November 2011

Abstract This paper is concerned with nonparametric statistics for the stress release process. We propose the local time estimator (LTE) for the stationary density and show that it is unbiased and uniformly consistent. The LTE is used in constructing an estimator for the intensity function. A goodness of fit test for the intensity function is also presented. In these studies, the local time of the stress release process plays an important role.

Keywords Stress release process · Local time · Stationary density · Uniform consistency · Goodness of fit test

1 Introduction

Consider the stationary ergodic stress release process

$$X_t = X_0 + t - N_t. \tag{1}$$

Here X_0 is an initial random variable distributed according to the stationary probability density f and N_t is a point process with the conditional intensity defined by:

T. Fujii (✉)
Center for the Study of Finance and Insurance, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
e-mail: fujii@sigmath.es.osaka-u.ac.jp

Y. Nishiyama
The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
e-mail: nisiyama@ism.ac.jp

$$\lambda_t = \lim_{\Delta \downarrow 0} \frac{P(N_{t+\Delta} - N_t = 1 \mid \mathcal{F}_{t-})}{\Delta} = \phi(X_{t-}),$$

where $\mathcal{F}_t$ is the σ-field generated by $\{X_s; 0 \leq s \leq t\}$. For the intensity function ϕ, we suppose that

(i) $0 \leq \phi(x) < \infty$ for any $x \in \mathbf{R}$,
(ii) there exists a constant $c_0 > 0$ such that $\phi(x) \geq c_0$ for all $x > 0$,
(iii) $\liminf_{x\to\infty} \phi(x) > 1$ and $\limsup_{x\to-\infty} \phi(x) < 1$,
(iv) for any $K > 0$, there is a positive constant M such that $\phi(x) \leq M$, for every $x \leq K$.

Then the stress release process (1) is ergodic, so it holds that

$$\frac{1}{T}\int_0^T g(X_t)\,\mathrm{d}t \xrightarrow{p} \int g\,\mathrm{d}\mu \qquad \text{for every } g \in L^1(\mu),$$

where μ is the probability measure on $\mathbf{R}$ given by $\mu(\mathrm{d}x) = f(x)\,\mathrm{d}x$ (see Hayashi 1986; Vere-Jones 1988). Fujii (2010) showed that the density function f satisfies

$$f(x) = \int_x^{x+1} \phi(y) f(y)\,\mathrm{d}y \tag{2}$$

with

$$\int_{-\infty}^{\infty} \phi(y) f(y)\,\mathrm{d}y = 1. \tag{3}$$

In this paper, we first consider the nonparametric estimation problem for the stationary (invariant) density function f of the stress release process (1) in the situation where the intensity function ϕ is unknown. We give a formula for the local time estimator (LTE) which is first introduced in Kutoyants (1996) for the study of the invariant density estimation of ergodic diffusion processes, see details in Kutoyants (1998, 2004). For more general stationary processes, Bosq (1998) and Bosq and Davydov (1999) also provide some important properties of the LTE. In the case of the stress release process (1), we show that the LTE is unbiased and uniformly consistent. Based on this result, we propose a uniformly consistent estimator for the unknown intensity function in terms of the kernel methods and the LTE. The uniform convergence result of the LTE for the diffusion can be seen in van Zanten (2000).

As another nonparametric statistical problem, we consider a goodness of fit test for the intensity function ϕ of the stress release process (1). We propose a test statistic in terms of the score marked empirical process (see Koul and Stute 1999). Negri and Nishiyama (2009) also used this statistic in the similar problem for a diffusion process. As a related work, Dachian and Kutoyants (2009) studied the problem of testing Poisson versus stress release.

To close this section, we recall the definition of the bracketing number (see e.g., van der Vaart and Wellner 1996). Let Ψ be a class of real functions defined on a space

I on which a seminorm $||\cdot||$ is defined. For every $\varepsilon > 0$, the bracketing number $N_{[\,]}(\varepsilon, \Psi, ||\cdot||)$ is the smallest integer N such that there exist N pairs $[l_k, u_k]$, $k = 1, \dots, N$, of functions on I, where each l_k and u_k may not belong to Ψ, such that for every $\psi \in \Psi$ the inequality $l_k \le \psi \le u_k$ holds for some k and that $||u_k - l_k|| < \varepsilon$. We refer to van der Vaart and Wellner (1996) for the weak convergence theory in $\ell^\infty(\Psi)$ space, the space of bounded functions on Ψ. We denote by $\xrightarrow{p}$ and $\xrightarrow{d}$ the convergence in probability and the convergence in distribution, respectively. The limit notations always mean to take a limit as $T \to \infty$. In Appendix, we prepare some limit theorems for more general point processes via the bracketing method.

2 Local time

We define the local time $\Lambda_T(x)$ of the stress release process (1) in a similar manner of Shorack and Wellner (1986, Section 9.8), who studied the empirical process.

Definition 1 The local time $\Lambda_T(x)$ of the stress release process (1) is defined by

$$\Lambda_T(x) = \#\{t : X_t = x,\ 0 \le t \le T\}, \tag{4}$$

where $\#A$ denotes the number of elements in the set A.

Let $\tau_n,\ n \ge 1$ be the time of the nth jump of the stress release process (1) and $\tau_0 = 0$. Then for any bounded measurable function $g(x)$,

$$\begin{aligned}
\int_0^T g(X_t)\,\mathrm{d}t &= \sum_{n=1}^{\infty} \int_{[\tau_{n-1}\wedge T, \tau_n \wedge T)} g(X_t)\,\mathrm{d}t \\
&= \sum_{n=1}^{\infty} \int_{[\tau_{n-1}\wedge T, \tau_n \wedge T)} g(X_{\tau_{n-1}} + (t - \tau_{n-1}))\,\mathrm{d}t \\
&= \sum_{n=1}^{\infty} \int_{[X_{\tau_{n-1}\wedge T}, X_{\tau_n \wedge T})} g(x)\,\mathrm{d}x \\
&= \int_{\mathbf{R}} g(x) \cdot \Lambda_T(x)\,\mathrm{d}x.
\end{aligned}$$

Therefore the local time $\Lambda_T(x)$ is thought as the occupation density.

Remark The local time defined by (4) corresponds to the so-called Banach indicatrix, introduced in Banach (1925). In addition, the last equation above is a consequence of the Banach theorem.

We note that for all $T > 0$, the local time $\Lambda_T(x)$ is right continuous and has left limit in x. Let us denote the local time $\tilde{\Lambda}_T(x)$ for the left limit process X_{t-} as in (4), then we have:

$$\begin{aligned}\int_0^T g(X_{t-})\,\mathrm{d}t &= \int_{-\infty}^{\infty} g(x)\tilde{\Lambda}_T(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} g(x)\Lambda_T(x)\,\mathrm{d}x + \int_{-\infty}^{\infty} g(x)\left(\tilde{\Lambda}_T(x) - \Lambda_T(x)\right)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} g(x)\Lambda_T(x)\,\mathrm{d}x, \end{aligned} \tag{5}$$

since $\Lambda_T(x) = \tilde{\Lambda}_T(x)$ for almost all x. Thus the local time $\Lambda_T(x)$ for the stress release process (1) can be used as the occupation density of its left limit process.

We represent the local time $\Lambda_T(x)$ by the stochastic integral in the following theorem, since the representation (4) is not suitable for studying some statistical properties concerning the local time $\Lambda_T(x)$. Note that this result corresponds to the Tanaka–Meyer formula for semimartingales.

Theorem 1 *The local time* (4) *is represented as:*

$$\Lambda_T(x) = 1_{\{X_T > x\}} - 1_{\{X_0 > x\}} + \int_0^T 1_{\{x < X_{t-} \le x+1\}}\,\mathrm{d}N_t, \tag{6}$$

where 1_A *is the indicator function of a set* A.

Proof Let $g(u)$ be a non-negative differentiable function with support [0, 1] satisfying that

$$\int_0^1 g(u)\,\mathrm{d}u = 1.$$

Introduce the function

$$\psi_n(y, x) = n\int_{-\infty}^{y} g(n(z-x))\,\mathrm{d}z,$$

which approximates the indicator function, i.e., $\psi_n(y, x) \to 1_{\{y>x\}}$ as $n \to \infty$. By Ito's formula and (5), we have:

$$\begin{aligned}\psi_n(X_T, x) - \psi_n(X_0, x) &= \int_0^T \psi_n'(X_{t-}, x)\,\mathrm{d}t + \int_0^T \psi_n(X_{t-}-1, x) - \psi_n(X_{t-}, x)\,\mathrm{d}N_t \\ &= \int_{-\infty}^{\infty} \psi_n'(y, x)\cdot\Lambda_T(y)\,\mathrm{d}y + \int_0^T \psi_n(X_{t-}-1, x) \\ &\quad - \psi_n(X_{t-}, x)\,\mathrm{d}N_t.\end{aligned}$$

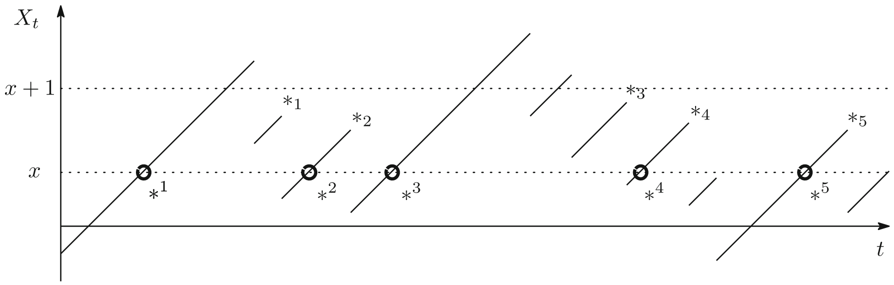

Fig. 1 Relation between the visiting at x and the jump

For the first term on the right hand side, it follows that

$$\begin{aligned}\int_{-\infty}^{\infty} \psi_n'(y,x)\cdot\Lambda_T(y)\,\mathrm{d}y &= n\int_x^{x+\frac{1}{n}} g(n(y-x))\cdot\Lambda_T(y)\,\mathrm{d}y\\ &= \int_0^1 g(u)\cdot\Lambda_T(x+\tfrac{u}{n})\,\mathrm{d}u\\ &\to \Lambda_T(x)\quad \text{as } n\to\infty,\end{aligned}$$

since for all fixed $T>0$, $\Lambda_T(x)$ is right continuous with left limit in x. Therefore, as $n\to\infty$ we obtain

$$1_{\{X_T>x\}} - 1_{\{X_0>x\}} = \Lambda_T(x) + \int_0^T 1_{\{X_{t-}-1>x\}} - 1_{\{X_{t-}>x\}}\,\mathrm{d}N_t.$$

This yields the assertion (6) immediately. □

Remark The interpretation of the representation (6) for the local time (4) is as follows: each jump which occurs in $(x, x+1]$ corresponds to one visit to x before the occurrence time, see Fig. 1. Further note that the first jump in $(x, x+1]$ does not related to any crossing if $X_0 > x$, and the last crossing does not too, if $X_T > x$.

3 Stationary density estimation

We consider the nonparametric estimation of the stationary density of the stress release process (1), in the situation where the intensity function ϕ is unknown. Here we define the LTE $f_T^\circ(x)$ for the stationary density of the stress release process (1) by:

$$f_T^\circ(x) = \frac{\Lambda_T(x)}{T}. \tag{7}$$

The unbiasedness is a well-known property of the LTE (see e.g., Bosq and Davydov 1999). Equations (2) and (6) help us to understand this fact in our particular model.

Theorem 2 *The LTE f_T° is unbiased.*

Proof By the stationarity, it follows that

$$\begin{aligned} E[f_T^\circ(x)] &= \frac{1}{T} E\left[\int_0^T 1_{\{x<X_{t-}\le x+1\}}\phi(X_{t-})\,\mathrm{d}t\right] \\ &= \int_x^{x+1} \phi(y) f(y)\,\mathrm{d}y \\ &= f(x), \end{aligned}$$

where we use Eqs. (2) and (6). □

In order to show the uniform continuity of the LTE, we provide a lemma.

Lemma 1

$$\sup_{x\in\mathbf{R}} \left| \frac{1}{T}\int_0^T 1_{\{x<X_{t-}\le x+1\}}\phi(X_{t-})\,\mathrm{d}t - f(x)\right| \xrightarrow{p} 0.$$

Proof Let us consider the class $\Phi = \{1_{\{x<y\le x+1\}},\ x \in \mathbf{R}\}$, then it follows from Theorem 2.7.5 in van der Vaart and Wellner (1996) that for any $\varepsilon > 0$, the class Φ satisfies $N_{[\,]}(\varepsilon, \Phi, L^1(\phi(x)\mu(\mathrm{d}x))) < \infty$. Hence by using the relation (2), the claim of this lemma can be proved quite similarly to Theorem 6 in Appendix. □

We state our main result in this section.

Theorem 3 *The LTE f_T° is uniformly consistent, i.e.,*

$$\sup_{x\in\mathbf{R}} |f_T^\circ(x) - f(x)| \xrightarrow{p} 0.$$

Proof It holds that

$$\begin{aligned} |f_T^\circ(x) - f(x)| \le{}& \left|\frac{1_{\{X_T>x\}} - 1_{\{X_0>x\}}}{T}\right| + \left|\frac{1}{T}\int_0^T 1_{\{x<X_{t-}\le x+1\}}\phi(X_{t-})\,\mathrm{d}t - f(x)\right| \\ &+ \left|\frac{1}{T}\int_0^T 1_{\{x<X_{t-}\le x+1\}}\,\mathrm{d}M_t\right|, \end{aligned}$$

where $M_t = N_t - \int_0^t \phi(X_{s-})\,\mathrm{d}s$. By taking the supremum over all $x \in \mathbf{R}$, the first term obviously converges to 0 and the second term is estimated by Lemma 1. Finally, Theorem 6 yields that

$$\sup_{x\in\mathbf{R}} \left|\frac{1}{T}\int_0^T 1_{\{x<X_{t-}\le x+1\}}\,\mathrm{d}M_t\right| \xrightarrow{p} 0.$$

Hence our claim has been proved. □

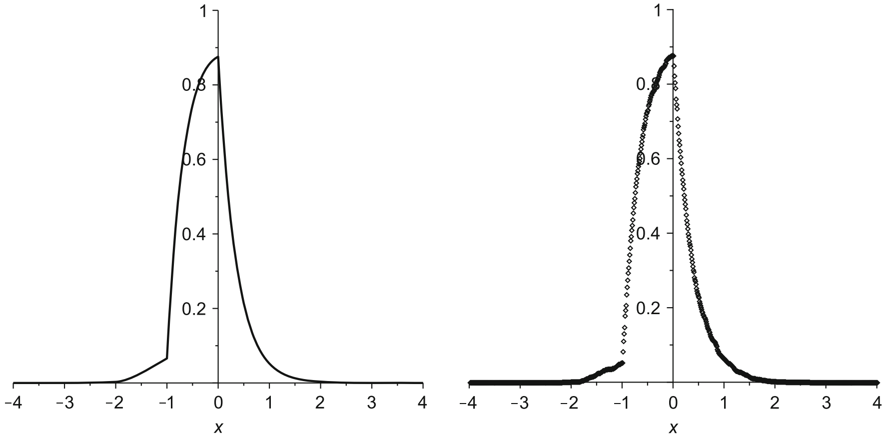

Fig. 2 The invariant density (*left*) and its LTE (*right*)

Remark Bosq and Davydov (1999, Proposition 4.3) give an almost sure convergence result for the LTE on bounded intervals in rather general settings. However, their setup implies that there exists a continuous version of the local time [see their assumption (3.12)]. In contrast, our local time (4) is not continuous, and moreover our result is concerned with the uniformity on the whole interval. Hence our result is not a special case of Bosq and Davydov (1999). On the other hand, a natural question may be whether it is possible to extend our result to the almost sure convergence one. The answer is affirmative (use the Borel–Cantelli theorem with help from Burkholder's inequality), however, we omit the details here for brevity.

We illustrate a numerical example for the LTE, briefly. Let us consider the case where the point process N_t is the simple self-correcting point (SSCP) process whose conditional intensity is given by:

$$\lambda_t = \phi(X_{t-}) = \begin{cases} 0.1, & X_{t-} < 0, \\ 3.0, & X_{t-} \geq 0, \end{cases}$$

that is introduced in Inagaki and Hayashi (1990). For a simulation of X_t on $t \in [0, 1{,}000]$, we construct the LTE (7) at points $x = \frac{i}{100}$, $i = -400 \ldots 400$. The result is shown in Fig. 2.

4 Intensity estimation

We consider the nonparametric estimation problem of the intensity function ϕ of the stress release process (1). We suppose in this section that the intensity function ϕ is continuous on any compact interval on $\mathbf{R}$. Then by Eq. (2), the stationary density f is so too. In order to estimate the intensity function, we use the following statistic:

$$\hat{\phi}_T(x) = \frac{\tilde{A}_T(x)}{f_T^\circ(x)}, \tag{8}$$

where $f_T^\circ(x)$ is the LTE for the stationary density and

$$\tilde{A}_T(x) = \frac{1}{T}\int_0^T \frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right) \mathrm{d}N_t,$$

where K is the uniform kernel on $[-\frac{1}{2}, \frac{1}{2}]$ and b_T is the bandwidth satisfying that

$$b_T \to 0 \quad \text{and} \quad Tb_T^2 \to \infty \qquad \text{as } T \to \infty.$$

We need the following auxiliary result for the asymptotic behavior of $\tilde{A}_T$.

Lemma 2 *For any compact interval* $[\alpha, \beta]$*,*

$$\sup_{x\in[\alpha,\beta]} |\tilde{A}_T(x) - \phi(x) f(x)| \xrightarrow{p} 0.$$

Proof We write

$$\tilde{A}_T(x) = \frac{1}{T}\int_0^T \frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right) \phi(X_{t-})\,\mathrm{d}t + \frac{1}{T}\int_0^T \frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right) \mathrm{d}M_t.$$

For the first term, we have:

$$\begin{aligned}
\frac{1}{T}\int_0^T \frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right) \phi(X_{t-})\,\mathrm{d}t &= \frac{1}{T}\int_{x-\frac{b_T}{2}}^{x+\frac{b_T}{2}} \frac{1}{b_T}\phi(y)\Lambda_T(y)\,\mathrm{d}y \\
&= \int_{x-\frac{b_T}{2}}^{x+\frac{b_T}{2}} \phi(y) f(y)\,\mathrm{d}y \\
&\quad + \int_{x-\frac{b_T}{2}}^{x+\frac{b_T}{2}} \frac{1}{b_T}\phi(y)\left(\frac{\Lambda_T(y)}{T} - f(y)\right)\mathrm{d}y \\
&\xrightarrow{p} \phi(x) f(x) \quad \text{uniformly in } x \in [\alpha, \beta],
\end{aligned}$$

where we use Theorem 3 and the fact that both of ϕ and f are uniformly continuous on any compact interval.

To show that the second term converges to zero uniformly, we check the conditions in Theorem 3.2 of Nishiyama (2000). For the condition [C2], by the assumption

$Tb_T^2 \to \infty$ we have:

$$\left\langle \frac{1}{T}\int_0^{\cdot} \frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right) \mathrm{d}M_t \right\rangle_T = \frac{1}{T^2}\int_0^T \frac{1}{b_T^2} K\left(\frac{X_{t-}-x}{b_T}\right) \phi(X_{t-})\,\mathrm{d}t$$

$$\leq \frac{1}{Tb_T^2}\cdot\frac{1}{T}\int_0^T \phi(X_{t-})\,\mathrm{d}t \overset{p}{\to} 0.$$

The condition [L2]:

$$\int_0^T \left|\overline{W}_t\right|^2 \cdot 1_{\{\overline{W}_t > \varepsilon\}} \phi(X_{t-})\,\mathrm{d}t \overset{p}{\to} 0,$$

where

$$\overline{W}_t = \sup_{x\in[\alpha,\beta]} \left\{\frac{1}{T}\cdot\frac{1}{b_T} K\left(\frac{X_{t-}-x}{b_T}\right)\right\} = \frac{1}{T}\cdot\frac{1}{b_T},$$

is derived from the assumption $Tb_T^2 \to \infty$ as follows:

$$\frac{1}{T^2}\cdot\frac{1}{b_T^2}\int_0^T 1_{\{\frac{1}{T}\frac{1}{b_T} > \varepsilon\}}\phi(X_{t-})\,\mathrm{d}t \leq \frac{1}{Tb_T^2}\cdot\frac{1}{T}\int_0^T \phi(X_{t-})\,\mathrm{d}t \overset{p}{\to} 0.$$

Finally, we check the condition [PE]. For any $\varepsilon > 0$, choose some finite points $\alpha = x_0 < x_1 < \cdots < x_{N_\varepsilon} = \beta$ such that $x_k - x_{k-1} \leq \varepsilon^2$ with $N_\varepsilon \leq \text{const.}\cdot\varepsilon^{-2}$. Then the entropy condition $\int_0^T \sqrt{\log N_\varepsilon}\,\mathrm{d}\varepsilon < \infty$ also holds. So our proof is finished by showing that

$$\sup_{\varepsilon>0}\max_{1\leq k\leq N_\varepsilon} \frac{1}{\varepsilon^2}\frac{1}{(Tb_T)^2}\int_0^T \sup_{x,y\in[x_{k-1},x_k]} \left|K\left(\frac{X_{t-}-x}{b_T}\right) - K\left(\frac{X_{t-}-y}{b_T}\right)\right|^2 \phi(X_{t-})\,\mathrm{d}t = O_P(1). \tag{9}$$

If $\varepsilon^2 \leq b_T$, then

$$\max_{1\leq k\leq N_\varepsilon}\frac{1}{T}\int_0^T \sup_{x,y\in[x_{k-1},x_k]}\left|K\left(\frac{X_{t-}-x}{b_T}\right) - K\left(\frac{X_{t-}-y}{b_T}\right)\right|^2 \phi(X_{t-})\,\mathrm{d}t$$

$$\leq \max_{1\leq k\leq N_\varepsilon}\int_{-\infty}^{\infty}\sup_{x,y\in[x_{k-1},x_k]}\left|K\left(\frac{z-x}{b_T}\right) - K\left(\frac{z-y}{b_T}\right)\right|^2 \phi(z)\cdot\frac{\Lambda_T(z)}{T}\,\mathrm{d}z$$

$$\leq \sup_{z\in[\alpha-\frac{b_T}{2},\beta+\frac{b_T}{2}]}\left\{\phi(z)\cdot\frac{\Lambda_T(z)}{T}\right\}\max_{1\leq k\leq N_\varepsilon}\int_{-\infty}^{\infty} 1_{\{x_{k-1}-\frac{b_T}{2}\leq z\leq x_k-\frac{b_T}{2}\}} + 1_{\{x_{k-1}+\frac{b_T}{2}\leq z\leq x_k+\frac{b_T}{2}\}}\,\mathrm{d}z$$

$$\leq \sup_{z\in[\alpha-\frac{b_T}{2},\beta+\frac{b_T}{2}]}\left\{\phi(z)\cdot\frac{\Lambda_T(z)}{T}\right\}\cdot 2\varepsilon^2.$$

Similarly, if $\varepsilon^2 > b_T$, then

$$
\begin{aligned}
&\max_{1\le k\le N_\varepsilon} \frac{1}{T}\int_0^T \sup_{x,y\in[x_{k-1},x_k]} \left| K\left(\frac{X_{t-}-x}{b_T}\right) - K\left(\frac{X_{t-}-y}{b_T}\right)\right|^2 \phi(X_{t-})\,\mathrm{d}t \\
&\le \sup_{z\in[\alpha-\frac{b_T}{2},\beta+\frac{b_T}{2}]} \left\{\phi(z)\cdot\frac{\Lambda_T(z)}{T}\right\}\cdot \max_{1\le k\le N_\varepsilon}\int_{-\infty}^{\infty} 1_{\{x_{k-1}-\frac{b_T}{2}\le z\le x_k+\frac{b_T}{2}\}}\,\mathrm{d}z \\
&\le \sup_{z\in[\alpha-\frac{b_T}{2},\beta+\frac{b_T}{2}]} \left\{\phi(z)\cdot\frac{\Lambda_T(z)}{T}\right\}\cdot(\varepsilon^2+b_T) \\
&< \sup_{z\in[\alpha-\frac{b_T}{2},\beta+\frac{b_T}{2}]} \left\{\phi(z)\cdot\frac{\Lambda_T(z)}{T}\right\}\cdot 2\varepsilon^2.
\end{aligned}
$$

Hence (9) is derived from the fact that $\sup_z\{\frac{\Lambda_T(z)}{T}\}$ is bounded in probability. □

This lemma and Theorem 3 yield the following theorem.

Theorem 4 *For any* $\varepsilon > 0$,

$$
\sup_{x\in I_\varepsilon} |\hat{\phi}_T(x) - \phi(x)| \xrightarrow{p} 0,
$$

where $I_\varepsilon = \{x : f(x) \ge \varepsilon\}$.

Remark It is possible to use more general kernels. However, it demands more complicated calculation when we control bracketing numbers, so we consider only the uniform kernel here.

The simulation result is provided in the case that the intensity function is:

$$
\phi(x) = \exp\left\{\frac{1}{2}x - 1\right\}.
$$

This type of the stress release process was studied in Ogata and Vere-Jones (1984). Now we observe the stress release process $X_t,\ t \in [0, 50{,}000]$ with this intensity and construct the estimator (8) with bandwidth $b_T = 0.15$ at points $x = \frac{i}{100}$, $i = -150\ldots450$. Figure 3 illustrates the result.

5 Goodness of fit test

In this section, we consider the one sample problem of the intensity function ϕ of the stress release process (1), i.e., consider testing hypotheses:

$$
\begin{aligned}
H_0 &: \quad \phi = \phi_0, \\
H_1 &: \quad \phi \ne \phi_0,
\end{aligned}
$$

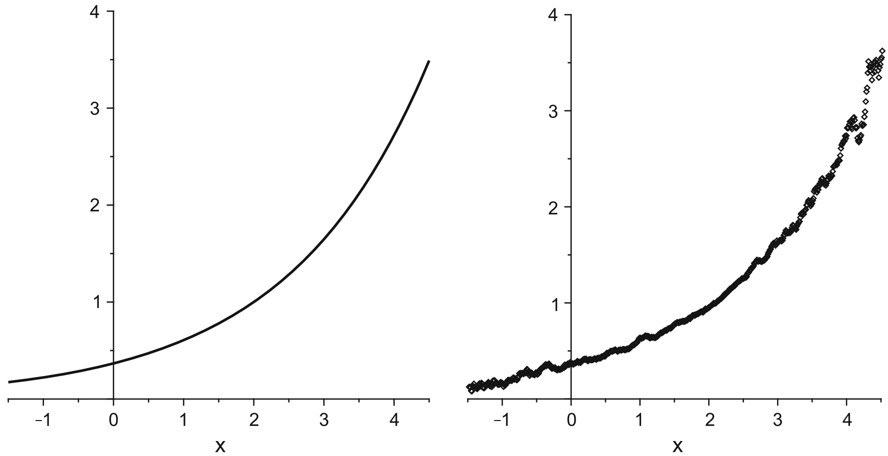

Fig. 3 The true intensity function (*left*) and the estimator (*right*)

where the hypothesis H_1 means that

$$\left|\int_{-\infty}^{x_0} (\phi(x) - \phi_0(x))\ \mu_\phi(\mathrm{d}x)\right| > 0, \quad \text{for some } x_0.$$

Here we rewrite the invariant measure μ as μ_ϕ for emphasis on the intensity ϕ. In the following, the intensity function ϕ is supposed to be globally bounded, i.e., there exists a constant $C > 0$ such that $0 \le \phi(x) \le C$ for any $x \in \mathbf{R}$. For every $x \in \mathbf{R}$, let us introduce the score marked empirical process

$$V_T(x) = \frac{1}{\sqrt{T}} \int_0^T 1_{(-\infty,x]}(X_{t-})\ (\mathrm{d}N_t - \phi_0(X_{t-})\,\mathrm{d}t)$$

and define the test statistic

$$\mathcal{S}_T = \sup_{x\in\mathbf{R}} |V_T|.$$

Then we have the following:

Theorem 5 *For the test statistic* $\mathcal{S}_T$,

(i) *under the hypothesis* H_0, *it holds that*

$$\mathcal{S}_T \xrightarrow{d} \sup_{s\in[0,1]} |B_s|$$

where B_s *is a standard Brownian motion, i.e., the test is asymptotically distribution-free,*

(ii) *under the hypothesis* H_1*, it holds that*

$$P(\mathcal{S}_T \leq K) \to 0 \quad \textit{for any } K.$$

Remark The limit distribution under the null hypothesis is given by:

$$P\left(\sup_{s\in[0,1]} |B_s| \leq x\right) = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left(-\frac{\pi^2(2k+1)^2}{8x^2}\right).$$

See e.g., 343 page of Feller (1971) for this formula. Before proving this theorem, we prepare a technical lemma.

Lemma 3

$$\sup_{x\in\mathbf{R}}\left\{\frac{\Lambda_T(x)}{T}\cdot(1+x^2)\right\} = O_P(1).$$

Proof It holds that for all $x \geq 0$,

$$\left|\frac{\Lambda_T(x)}{T}\cdot(1+x^2)\right| \leq \frac{1+X_T^2}{T} + \frac{1+X_0^2}{T} + \frac{1}{T}\int_0^T (1+X_{t-}^2)\,\mathrm{d}N_t.$$

Markov's inequality yields

$$P\left\{\frac{1}{T}\int_0^T (1+X_{t-}^2)\,\mathrm{d}N_t \geq a\right\} \leq \frac{C}{a}E[1+X_0^2] \to 0, \quad \text{as } a\to\infty,$$

since the stationary distribution has finite moments under our assumptions, see Hayashi (1986) and Vere-Jones (1988). The other terms can be estimated similarly. Therefore we have:

$$\sup_{x\geq 0}\left\{\frac{\Lambda_T(x)}{T}\cdot(1+x^2)\right\} = O_P(1).$$

For $x < 0$, our claim can be proved by noticing that

$$\Lambda_T(x) = 1_{\{X_0\leq x\}} - 1_{\{X_T\leq x\}} + \int_0^T 1_{\{x<X_{t-}\leq x+1\}}\,\mathrm{d}N_t,$$

in the same way. □

Proof of Theorem 5 For the proof of the claim (i), notice that by Lemma 3, the class $\Psi = \{1_{(-\infty,z]};\, z\in\mathbf{R}\}$ satisfies all requirements in Theorem 7 (see Appendix), including the bracketing entropy condition (see e.g., Theorem 2.7.5 of van der Vaart and Wellner 1996) for the finite measure $\nu(\mathrm{d}z) = \frac{1}{1+z^2}\mathrm{d}z$ on $\mathbf{R}$ and the corresponding local time $\overline{\Lambda}_T(z) = \Lambda_T(z)\cdot(1+z^2)$. Therefore, the score marked empirical process

$V_T(x)$ weakly converges to $\{B_{S(x)}; x \in \mathbf{R}\}$ in $\ell^\infty(\mathbf{R})$, where $s \rightsquigarrow B_s$ is a standard Brownian motion and

$$S(x) = \int_{-\infty}^{x} \phi_0(z)\mu_{\phi_0}(\mathrm{d}z). \tag{10}$$

The continuous mapping theorem leads to

$$\mathcal{S}_T = \sup_{x\in\mathbf{R}} |V_T(x)| \xrightarrow{d} \sup_{x\in\mathbf{R}} |B_{S(x)}|.$$

Hence the claim (i) follows from (3) and (10).

To show (ii), observe that

$$\begin{aligned}
&\sup_{x\in\mathbf{R}} \left| \frac{1}{\sqrt{T}} \int_0^T 1_{(-\infty,x]}(X_{t-})\, (\mathrm{d}N_t - \phi_0(X_{t-})\, \mathrm{d}t) \right| \\
&\quad\geq \sup_{x\in\mathbf{R}} \sqrt{T} \cdot \left| \frac{1}{T} \int_0^T 1_{(-\infty,x]}(X_{t-})(\phi(X_{t-}) - \phi_0(X_{t-}))\, \mathrm{d}t \right| \\
&\quad\quad - \sup_{x\in\mathbf{R}} \left| \frac{1}{\sqrt{T}} \int_0^T 1_{(-\infty,x]}(X_{t-})\, (\mathrm{d}N_t - \phi(X_t)\mathrm{d}t) \right|,
\end{aligned}$$

Since by Theorem 7, the process

$$\frac{1}{\sqrt{T}} \int_0^T 1_{(-\infty,x]}(X_{t-})\, (\mathrm{d}N_t - \phi(X_{t-})\mathrm{d}t)$$

converges weakly in $\ell^\infty(\mathbf{R})$ to a tight law, the second term on the right hand side is tight. The first term on the right hand side tends to ∞ in probability because there exists some $x_0 \in \mathbf{R}$ such that

$$\int_{\mathbf{R}} 1_{(-\infty,x_0]}(z)(\phi(z) - \phi_0(z))\, \mu_\phi(\mathrm{d}z) \neq 0$$

and for such a x_0 it holds that

$$\left| \frac{1}{T}\int_0^T 1_{(-\infty,x_0]}(X_{t-})(\phi(X_{t-}) - \phi_0(X_{t-}))\, \mathrm{d}t \right| \xrightarrow{p} \left| \int_{-\infty}^{x_0} (\phi(z) - \phi_0(z))\, \mu_\phi(\mathrm{d}z) \right| > 0.$$

Thus we have proved the claim (ii). □

Appendix A

Suppose that a point process $t \rightsquigarrow N_t$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ admits the predictable intensity

$$\alpha(Z_t),$$

where $t \rightsquigarrow Z_t$ is a predictable process which take values in a measurable space $(I, \mathcal{I})$. We consider the following two properties:

A1 *Ergodicity* There exists a probability measure μ on $(I, \mathcal{I})$ such that for every μ-integrable function ψ on $(I, \mathcal{I})$

$$\frac{1}{T}\int_0^T \psi(Z_t)\,\mathrm{d}t \xrightarrow{p} \int_I \psi(z)\,\mu(\mathrm{d}z), \quad \text{as } T \to \infty.$$

A2 *Existence of a good local time* There exists a non-negative predictable processes $t \rightsquigarrow \Lambda_t(z)$, $z \in I$ and a measure ν on $(I, \mathcal{I})$ such that for any measurable function ψ on $(I, \mathcal{I})$

$$\int_0^T \psi(Z_t)\,\mathrm{d}t = \int_I \psi(z)\Lambda_T(z)\,\nu(\mathrm{d}z), \quad \text{almost surely,}$$

provided the integrals on the both sides exist, and that

$$\sup_{z\in I} \Lambda_T(z) = O_P(T), \quad \text{as } T \to \infty.$$

Theorem 6 *Assume A1 Suppose that the class $\Psi \subset L_1(\alpha(z)\mu(dz))$ satisfies that for every $\varepsilon > 0$, $N_{[\,]}(\varepsilon, \Psi, L_1(\alpha(z)\mu(dz))) < \infty$. Then it holds that*

$$\sup_{\psi\in\Psi} \frac{1}{T}\left|\int_0^T \psi(Z_t)\,dN_t - \int_0^T \psi(Z_t)\alpha(Z_t)\,\mathrm{d}t\right| \xrightarrow{p} 0.$$

Proof Choose any $\varepsilon > 0$. Then, there exists $N_\varepsilon = N_{[\,]}(\varepsilon, \Psi, L_1(\alpha(z)\mu(dz)))$ brackets $[l_k, u_k]$, $k = 1, \ldots, N_\varepsilon$, which cover Ψ such that

$$\varepsilon_k = \int_I |u_k(z) - l_k(z)|\alpha(z)\,\mu(\mathrm{d}z) < \varepsilon.$$

Notice that for any $\psi \in [l_k, u_k]$

$$\begin{aligned}
&\frac{1}{T}\left\{\int_0^T \psi(Z_t)\,\mathrm{d}N_t - \int_0^T \psi(Z_t)\alpha(Z_t)\,\mathrm{d}t\right\}\\
&\quad \le \frac{1}{T}\left\{\int_0^T u_k(Z_t)\,(\mathrm{d}N_t - \alpha(Z_t)\mathrm{d}t) + \int_0^T (u_k(Z_t) - l_k(Z_t))\alpha(Z_t)\,\mathrm{d}t\right\}.
\end{aligned}$$

Considering also the lower bound, we finally get

$$\begin{aligned}
&\sup_{\psi\in\Psi}\frac{1}{T}\left|\int_0^T \psi(Z_t)\,\mathrm{d}N_t-\int_0^T \psi(Z_t)\alpha(Z_t)\,\mathrm{d}t\right|\\
&\quad\le \max_{1\le k\le N_\varepsilon}\frac{1}{T}\left\{\int_0^T u_k(Z_t)\,(dN_t-\alpha(Z_t)\mathrm{d}t)\right\}\\
&\qquad+\min_{1\le k\le N_\varepsilon}\frac{1}{T}\left\{\int_0^T l_k(Z_t)\,(dN_t-\alpha(Z_t)\mathrm{d}t)\right\}\\
&\qquad+\max_{1\le k\le N_\varepsilon}\frac{1}{T}\int_0^T (u_k(Z_t)-l_k(Z_t))\alpha(Z_t)\,\mathrm{d}t.
\end{aligned}$$

The first two terms on the right converges to zero in probability (use Lenglart's inequality). As for the third term on the right, since each $\frac{1}{T}\int_0^T (u_k(Z_t)-l_k(Z_t))\alpha(Z_t)\,\mathrm{d}t$ converges in probability to ε_k, by Slutsky's lemma (see e.g., Example 1.4.7 of van der Vaart and Wellner 1996) they converge in probability, jointly. So the third term converges in probability to $\max_{1\le k\le N_\varepsilon}\varepsilon_k$ which is smaller than ε. So we have proved that

$$\sup_{\psi\in\Psi}\frac{1}{T}\left|\int_0^T \psi(Z_t)\,\mathrm{d}N_t-\int_0^T \psi(Z_t)\alpha(Z_t)\,\mathrm{d}t\right| < \varepsilon + o_P(1).$$

Since the choice of ε is arbitrary, we may conclude that the left hand side converges to zero in probability. The proof is finished. □

Theorem 7 *Assume A1 and A2. Suppose that the class $\Psi\subset L_2(\alpha(z)\nu(dz))$ has an envelope function $\bar\psi\in L_{2+\delta}(\alpha(z)\nu(dz))$ for some $\delta>0$, i.e., $|\psi|\le\bar\psi$ for all $\psi\in\Psi$, and satisfies the metric entropy condition with L_2-bracketing:*

$$\int_0^1 \sqrt{\log N_{[\,]}(\varepsilon,\Psi,L_2(\alpha(z)\nu(dz)))}\,\mathrm{d}\varepsilon<\infty.$$

Then, it holds that the random fields $Y_T=\{Y_T(\psi);\ \psi\in\Psi\}$ defined by

$$Y_T(\psi)=\frac{1}{\sqrt{T}}\left\{\int_0^T \psi(Z_t)\,dN_t-\int_0^T \psi(Z_t)\alpha(Z_t)\,\mathrm{d}t\right\}$$

converges weakly in $\ell^\infty(\Psi)$ to a zero-mean Gaussian random field $G=\{G(\psi);\ \psi\in\Psi\}$ with the covariance

$$EG(\psi)G(\psi')=\int_I \psi(z)\psi'(z)\alpha(z)\,\mu(\mathrm{d}z).$$

Furthermore, almost all paths of $\psi\rightsquigarrow G(\psi)$ are uniformly ρ-continuous with respect to the semimetric $\rho(\psi,\psi')=\sqrt{E|G(\psi)-G(\psi')|^2}$.

Proof We apply Theorem 3.4 of Nishiyama (2000) to the terminal variables $\{M_T^{T,\psi}; \psi \in \Psi\}$ of the martingales $t \rightsquigarrow M_t^{T,\psi}$ given by:

$$M_t^{T,\psi} = \frac{1}{\sqrt{T}} \int_0^t \psi(Z_s)(\mathrm{d}N_s - \alpha(Z_s)\,\mathrm{d}s).$$

His condition [C2] is satisfied because

$$\langle M^{T,\psi}, M^{T,\psi'} \rangle_T = \frac{1}{T} \int_0^T \psi(Z_s)\psi'(Z_s)\alpha(Z_s)\,\mathrm{d}s \xrightarrow{p} \int_I \psi(z)\psi'(z)\alpha(z)\,\mu(\mathrm{d}z).$$

To check the Lindberg condition [L2], take any $\varepsilon > 0$ and observe that

$$\begin{aligned}
\frac{1}{T} \int_0^T \bar{\psi}(Z_s)^2 1_{\{\frac{1}{\sqrt{T}}\bar{\psi}(Z_s)>\varepsilon\}} \alpha(Z_s)\,\mathrm{d}s &\leq \frac{1}{T} \int_0^T \bar{\psi}(Z_s)^2 \left| \frac{1}{\varepsilon} \cdot \frac{1}{\sqrt{T}} \bar{\psi}(Z_s) \right|^\delta \alpha(Z_s)\,\mathrm{d}s \\
&= \frac{1}{\varepsilon^\delta} \cdot \frac{1}{T^{\frac{\delta}{2}}} \cdot \frac{1}{T} \int_0^T \bar{\psi}(Z_s)^{2+\delta} \alpha(Z_s)\,\mathrm{d}s \\
&\xrightarrow{p} 0.
\end{aligned}$$

To check [PE], for every $\varepsilon > 0$ choose some ε-brackets $[l_k, u_k]$, $k = 1, \dots, N_\varepsilon$, where $N_\varepsilon = N_{[\,]}(\varepsilon, \Psi, L^2(\alpha(z)\nu(\mathrm{d}z)))$. Construct the partition $\Psi = \bigcup_{k=1}^{N_\varepsilon} \Psi(\varepsilon; k)$, which corresponds to Nishiyama's notation, as

$$\Psi(\varepsilon; k) = \{\psi \in \Psi : l_k \leq \psi \leq u_k\}.$$

The square of the *quadratic modulus* is bounded by

$$\begin{aligned}
&\sup_{\varepsilon>0} \max_{1\leq k\leq N_\varepsilon} \frac{1}{T} \cdot \frac{1}{\varepsilon^2} \int_0^T |u_k(Z_s) - l_k(Z_s)|^2 \alpha(Z_s)\,\mathrm{d}s \\
&\quad = \sup_{\varepsilon>0} \max_{1\leq k\leq N_\varepsilon} \frac{1}{T} \cdot \frac{1}{\varepsilon^2} \int_I |u_k(z) - l_k(z)|^2 \alpha(z)\Lambda_T(z)\,\nu(\mathrm{d}z) \\
&\quad \leq \frac{1}{T} \sup_{z\in I} \Lambda_T(z) = O_P(1).
\end{aligned}$$

The proof is finished. □

Acknowledgments The authors would like to thank the Associate Editor and two anonymous referees for their valuable comments. This research was supported by Grant-in-Aid for Young Scientists (B), No. 21700320 [T.F.] and Grant-in-Aid for Scientific Research (C), No. 21540157 [Y.N.] from Japan Society for the Promotion of Science (JSPS).

References

Banach, S. (1925). Sur les lignes rectifiables et les surfaces dont l'aire est finie. *Fundamenta Mathematicae, 7*, 225–236.

Bosq, D. (1998). *Nonparametric statistics for stochastic processes* (2nd ed.). Lecture notes in statistics (Vol. 110). New York: Springer.

Bosq, D., Davydov, Yu. (1999). Local time and density estimation in continuous time. *Mathematical Methods of Statistics, 8*, 22–45.

Dachian, S., Kutoyants, Yu. A. (2009). Hypotheses testing: Poisson versus stress-release. *Journal of Statistical Planning and Inference, 139*, 1668–1684.

Feller, W. (1971). *An introduction to probability theory and its applications* (2nd ed., Vol. II). Wiley: New York.

Fujii, T. (2010) On the invariant density of the stress release process (submitted).

Hayashi, T. (1986). Laws of large numbers in self-correcting point processes. *Stochastic Processes and Applications, 23*, 319–326.

Inagaki, N., Hayashi, T. (1990). Parameter estimation for the simple self-correcting point process. *Annals of the Institute of Statistical Mathematics, 42*, 89–98.

Koul, H. L., Stute, W. (1999). Nonparametric model checks for time series. *The Annals of Statistics, 27*, 204–236.

Kutoyants, Yu. A. (1996). Efficient density estimation for ergodic diffusion. *Research Memorandum* No. 607. The Institute of Statistical Mathematics, Tokyo.

Kutoyants, Yu. A. (1998). Efficient density estimation for ergodic diffusion processes. *Statistical Inference for Stochastic Processes, 1*, 131–155.

Kutoyants, Yu. A. (2004). *Statistical inference for ergodic diffusion processes*. London: Springer.

Negri, I., Nishiyama, Y. (2009). Goodness of fit test for ergodic diffusion processes. *Annals of the Institute of Statistical Mathematics, 61*, 919–928.

Nishiyama, Y. (2000). Weak convergence of some classes of martingales with jumps. *The Annals of Probability, 28*, 685–712.

Ogata, Y., Vere-Jones, D. (1984). Inference for earthquake models: a self-correcting model. *Stochastic Processes and Applications, 17*, 337–347.

Shorack, G. R., Wellner, J. A. (1986). *Empirical processes with applications to statistics*. New York: Wiley.

van der Vaart, A. W., Wellner, J. A. (1996). *Weak convergence and empirical processes*. New York: Springer.

van Zanten, J. H. (2000). On the uniform convergence of the empirical density of an ergodic diffusion. *Statistical Inference for Stochastic Processes, 3*, 251–262.

Vere-Jones, D. (1988). On the variance properties of stress release models. *The Australian Journal of Statistics, 30A*, 123–135.