

# Stein estimation of Poisson process intensities

Nicolas Privault\*

Department of Mathematics  
City University of Hong Kong  
83 Tat Chee Avenue  
Kowloon Tong  
Hong Kong

Anthony Réveillac†

Laboratoire de Mathématiques  
Université de La Rochelle  
Avenue Michel Crépeau  
17042 La Rochelle Cedex  
France

October 29, 2007

## Abstract

We construct superefficient estimators of Stein type for the intensity parameter  $\lambda > 0$  of a Poisson process, using integration by parts and superharmonic functionals on the Poisson space.

**Key words:** Poisson process, Intensity estimation, Stein estimation, Malliavin calculus.

*Mathematics Subject Classification:* 62G05, 60J75, 60H07, 31B05.

## 1 Introduction

Consider a Poisson process  $(X_t)_{t \in [0, T]}$  with intensity  $u$  of the form  $u(t) = \lambda h(t)$ ,  $t \in [0, T]$ , under a probability  $\mathbb{P}_u$ , where  $(h(t))_{t \in [0, T]}$  is a given deterministic function. As is well-known, cf. [6], or [8], p. 351, Example 2, Ch. XIX, the classical parametric maximum likelihood estimator (MLE)

$$\hat{\lambda}_T := \frac{X_T}{h(T)}$$

---

\*nprivault@cityu.edu.hk

†anthony.reveillac@univ-lr.fr

of  $\lambda$  on the time interval  $[0, T]$  is obtained by maximization of the Girsanov density, i.e. under the condition:

$$\frac{d}{d\lambda} \left( \lambda^{X_T} e^{-(\lambda h(T)-T)} \prod_{k=1}^{X_T} h(T_k) \right) = \left( \frac{X_T}{\lambda} - h(T) \right) \lambda^{X_T} e^{-(\lambda h(T)-T)} \prod_{k=1}^{X_T} h(T_k) = 0.$$

The MLE  $\hat{\lambda}_T$  is efficient in the sense that it attains the Cramer-Rao bound

$$\mathbb{E}_u \left[ |\hat{\lambda}_T - \lambda|^2 \right] = \frac{\lambda}{h(T)}$$

over all unbiased estimators  $\zeta_T$  satisfying  $\mathbb{E}_u[\zeta_T] = \lambda$ , for all  $\lambda > 0$ , where  $\mathbb{E}_u$  denotes expectation under  $\mathbb{P}_u$ .

In this paper we construct superefficient estimators for the intensity parameter  $\lambda > 0$  when the intensity  $(u(t))_{t \in [0, T]}$  of  $(X_t)_{t \in [0, T]}$  is constrained to have the form  $u(t) = \lambda h(t)$ ,  $t \in [0, T]$ .

We use integration by parts and harmonic analysis on the Poisson space, via the technique introduced by Stein [15] for the estimation of the mean of a standard Gaussian random vector  $Z$  in  $\mathbb{R}^d$ , and extended to drift estimation on the Wiener space in [14], [13]. Recall that Stein's argument relies on:

a) the integration by parts

$$\mathbb{E}_\mu[(Z_i - \mu_i)g_i(Z)] = \mathbb{E}_\mu[\partial_i g_i(Z)], \quad (1.1)$$

where  $\mathbb{E}_\mu$  denotes expectation under the standard Gaussian measure with mean  $\mu \in \mathbb{R}^d$ ,

b) the chain rule of derivation for the partial derivative  $\partial_i$  on  $\mathbb{R}^d$ ,

c) the existence and properties of non-negative superharmonic functions on  $\mathbb{R}^d$  for  $d \geq 3$ .

Precisely, given an estimator of  $\mu \in \mathbb{R}^d$  of the form  $Z + \text{grad log } f(Z)$ , where  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is sufficiently smooth, one gets, using the chain rule of derivation,

$$\begin{aligned}
& \mathbb{E}_\mu[\|Z + \text{grad log } f(Z) - \mu\|_{\mathbb{R}^d}^2] \\
&= \mathbb{E}_\mu[\|Z - \mu\|_{\mathbb{R}^d}^2] + \mathbb{E}_\mu[\|\text{grad log } f(Z)\|_{\mathbb{R}^d}^2] + 2 \sum_{i=1}^d \mathbb{E}_\mu[(Z_i - \mu_i) \partial_i \log f(Z)] \\
&= d + \mathbb{E}_\mu[\|\text{grad log } f(Z)\|_{\mathbb{R}^d}^2] + 2 \mathbb{E}_\mu \left[ \sum_{i=1}^d \partial_i^2 \log f_i(Z) \right] \tag{1.2} \\
&= d + 4 \sum_{i=1}^d \mathbb{E}_\mu \left[ \frac{\partial_i^2 \sqrt{f}(Z)}{\sqrt{f}(Z)} \right],
\end{aligned}$$

i.e.  $Z + \text{grad log } f(Z)$  improves in the mean square sense over the maximum likelihood estimator (MLE)  $Z$  if  $d \geq 3$  and  $\sqrt{f}$  is superharmonic on  $\mathbb{R}^d$ .

Integration by parts for  $g : \mathbb{N} \rightarrow \mathbb{R}$  with respect to the discrete Poisson distribution  $P(X = k) = e^{-\lambda} \lambda^k / k!$ ,  $k \in \mathbb{N}$ , can be written as

$$\mathbb{E}_\lambda[(X - \lambda)g(X)] = \lambda \mathbb{E}_\lambda[g(X + 1) - g(X)].$$

where  $\mathbb{E}_\lambda$  denotes expectation under the Poisson distribution with parameter  $\lambda > 0$ , and has been used to derive Stein identities for jump processes, such as

$$\begin{aligned}
\mathbb{E}_\lambda[|X - \lambda + g(X)|^2] &= \lambda + \mathbb{E}_\lambda[|g(X)|^2] + 2 \mathbb{E}_\lambda[(X - \lambda)g(X)] \\
&= \lambda + \mathbb{E}_\lambda[|g(X)|^2] + 2\lambda \mathbb{E}_\lambda[g(X + 1) - g(X)],
\end{aligned}$$

cf. [1], [3]. However the absence of chain rule for the finite difference operation  $g \mapsto g(\cdot + 1) - g(\cdot)$  prevents us from continuing the calculation as in (1.2) above, and from using superharmonic functions as in the Gaussian case. On a more general level the derivation property requirement prevents us from using finite difference gradients on Poisson functionals cf. e.g. [9].

In this paper we apply Stein's argument on the Poisson space, and construct superefficient estimators for the discrete Poisson law, by replacing the Stein equation (1.1) with the integration by parts formula of [2], [4], extended to arbitrary intensity

functions on the Poisson space as in [11], in which the gradient  $\nabla$  satisfies the chain rule of derivation. When  $u(t)$  has the form  $u(t) = \lambda h(t)$  we apply our result to the parametric estimation of the Poisson process intensity  $\lambda > 0$  via estimators of the form

$$\hat{\lambda}_T + \frac{c}{h(T)} \mathbf{1}_{\{X_T=0\}} + \frac{1}{h(T)} \nabla_T \log F,$$

where  $F$  is a positive superharmonic random variable on the Poisson space,  $c \in \mathbb{R}$  is a suitably chosen constant, and  $\nabla_T$  is a gradient operator on the Poisson space.

Unlike in the Gaussian case, the Laplacian considered here contains first order terms and is not the standard Laplacian on  $\mathbb{R}^d$ . As a consequence the  $d \geq 3$  dimension condition imposed in the Gaussian case can be waived and superharmonic functionals can be constructed as functions of  $d$  jump times for  $d \geq 1$ .

We proceed as follows. In Section 2 we introduce the Poisson space and derive the Cramer-Rao bound for a non-parametric estimator of the intensity. Our proof uses stochastic calculus, and in this respect it differs from the ones usually found in the literature, cf. e.g. § 1.2 of [7]. In Section 3 we recall the elements of analysis and integration by parts on the Poisson space which will be needed in Section 4 to construct superefficient estimators for the intensity of a Poisson process. In case  $u$  has the form  $u(t) = \lambda t$ , numerical applications and simulations are given in Section 5 using simple examples of (pseudo) superharmonic functionals.

## 2 Preliminaries

In this section we state some notation on the Poisson space and Poisson process, and derive the Cramer-Rao bound. Let  $T > 0$  and consider  $(X_t)_{t \in [0, T]}$  the canonical process on

$$\Omega = \left\{ \omega = \sum_{k=1}^n \delta_{t_k} : 0 \leq t_1 < \dots < t_n \leq T, \quad n \in \mathbb{N} \cup \{\infty\} \right\},$$

defined as

$$X_t(\omega) = \omega([0, t]), \quad t \in [0, T],$$

where  $\delta_x$  denotes the Dirac measure at  $x \in [0, T]$ . Let  $(T_k)_{k \geq 1}$  denote the jump times of  $(X_t)_{t \in [0, T]}$ , i.e. any  $\omega \in \{X_T = n\}$  is written as

$$\omega = \sum_{k=1}^n \delta_{T_k}.$$

Let  $\mathbb{P}$  denote the standard Poisson measure on  $\Omega$ , under which  $(X_t)_{t \in \mathbb{R}_+}$  is a standard Poisson process, and let  $(\mathcal{F}_t)_{t \in [0, T]}$  denote the filtration generated by  $(X_t)_{t \in [0, T]}$ .

**Definition 2.1.** Let  $\mathcal{P}$  denote the set of functions of the form

$$u(t) = \int_0^t \dot{u}(s) ds, \quad t \in [0, T],$$

where  $\dot{u} : [0, T] \rightarrow [0, \infty)$  is a non-negative function.

Let now  $u \in \mathcal{P}$ . By the Girsanov theorem, the measure  $\mathbb{P}_u$  on  $\Omega$ , under which the canonical process  $(X_t)_{t \in [0, T]}$  is a Poisson process with intensity  $\dot{u}(t)dt$ , is absolutely continuous with respect to  $\mathbb{P}$  with

$$d\mathbb{P}_u = \Lambda(u) d\mathbb{P},$$

where

$$\Lambda(u) = \exp\left(-\int_0^T (\dot{u}(s) - 1) ds\right) \prod_{k=1}^{X_T} \dot{u}(T_k)$$

denotes the Girsanov density. In the sequel we will denote by  $\mathbb{E}_u$  the expectation under  $\mathbb{P}_u$  and let  $L_u^2(\Omega) = L^2(\Omega, \mathbb{P}_u)$ .

We close this section with a derivation of the Cramer-Rao inequality using stochastic calculus, for non-parametric estimation of the intensity. In case the intensity is constrained to be constant on intervals, our bound can be recovered from the Cramer-Rao inequality for arbitrary finite dimensional estimators, cf. Theorem 1.5 of [7].

**Definition 2.2.** An estimator  $\xi_t$  of  $u \in \mathcal{P}$  is called unbiased if

$$\mathbb{E}_u[\xi_t] = u(t), \quad t \in [0, T],$$

and adapted if the process  $(\xi_t)_{t \in [0, T]}$  is adapted to the filtration  $(\mathcal{F}_t)_{t \in [0, T]}$  generated by  $(X_t)_{t \in [0, T]}$ .

Here,  $X_t$  can be considered as an unbiased maximum likelihood estimator of its own intensity  $u(t)$  under  $\mathbb{P}_u$ ,  $t \in [0, T]$ . From the next proposition, this estimator is efficient since its mean square error is equal to

$$\mathbb{E}_u [|X_t - u(t)|^2] = u(t), \quad t \in [0, T]. \quad (2.1)$$

**Proposition 2.3.** *Cramer-Rao inequality.* Let  $u \in \mathcal{P}$  and  $t \in [0, T]$ . For any unbiased and adapted estimator  $\xi_t$  of  $u(t)$  we have

$$\mathbb{E}_u [|\xi_t - u(t)|^2] \geq u(t), \quad u \in \mathcal{P}, \quad (2.2)$$

where for all  $u \in \mathcal{P}$  the lower bound  $u(t)$  is attained by  $\xi_t = X_t$ .

*Proof.* Since  $\xi_t$  is unbiased, for all  $v \in \mathcal{P}$  and  $\varepsilon \in \mathbb{R}$  we have

$$\mathbb{E}_{u+\varepsilon v}[\xi_t] = u(t) + \varepsilon v(t) = \mathbb{E}_{u+\varepsilon v}[u(t)] + \varepsilon v(t),$$

hence

$$\begin{aligned} v(t) &= \frac{d}{d\varepsilon} \mathbb{E}_{u+\varepsilon v}[\xi_t - u(t)]|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \mathbb{E}[(\xi_t - u(t))\Lambda(u + \varepsilon v)]|_{\varepsilon=0} \\ &= \mathbb{E} \left[ (\xi_t - u(t)) \frac{d}{d\varepsilon} \Lambda(u + \varepsilon v)|_{\varepsilon=0} \right] \\ &= \mathbb{E}_u \left[ (\xi_t - u(t)) \frac{d}{d\varepsilon} \log \Lambda(u + \varepsilon v)|_{\varepsilon=0} \right] \\ &= \mathbb{E}_u \left[ (\xi_t - u(t)) \int_0^T \mathbf{1}_{\{\dot{u}(s) \neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} (dX_s - \dot{u}(s)ds) \right] \\ &= \mathbb{E}_u \left[ (\xi_t - u(t)) \int_0^t \mathbf{1}_{\{\dot{u}(s) \neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} (dX_s - \dot{u}(s)ds) \right]. \end{aligned}$$

Note that the adaptedness hypothesis on the estimator  $\xi_t$  was used to get the last equality above, and that the exchange between expectation and derivative can be justified by standard uniform integrability arguments. Thus, by the Cauchy-Schwarz inequality and the Itô isometry, we have

$$v^2(t) \leq \mathbb{E}_u \left[ \left( \int_0^t \mathbf{1}_{\{\dot{u}(s) \neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} (dX_s - \dot{u}(s)ds) \right)^2 \right] \mathbb{E}_u [|\xi_t - u(t)|^2]$$

$$= \int_0^t \mathbf{1}_{\{\dot{u}(s) \neq 0\}} \frac{|\dot{v}(s)|^2}{\dot{u}(s)} ds \mathbb{E}_u[|\xi_t - u(t)|^2].$$

It then suffices to take

$$\dot{v}(s) := \dot{u}(s), \quad s \in [0, t],$$

to get

$$\text{Var}_u[\xi_t] = \mathbb{E}_u[|\xi_t - u(t)|^2] \geq u(t), \quad (2.3)$$

which leads to (2.2). As noted in (2.1),  $\hat{u}_t = X_t$  is clearly unbiased under  $\mathbb{P}_u$  and attains the lower bound  $u(t)$ .  $\square$

### 3 Analysis on the Poisson space

In this section we recall the elements of analysis and integration by parts on the Poisson space which will be needed for the construction of Stein estimators.

**Definition 3.1.** *We denote by  $\mathcal{S}$  the space of Poisson functionals of the form*

$$F = f_0 \mathbf{1}_{\{X_T=0\}} + \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} f_n(T_1, \dots, T_n), \quad (3.1)$$

where  $f_0 \in \mathbb{R}$  and  $f_n$ ,  $n \geq 1$ , is  $\mathcal{C}^1$  on  $\{0 \leq t_1 \leq \dots \leq t_n \leq T\}$ , and satisfying the continuity condition

$$f_n(t_1, \dots, t_n) = f_{n+1}(t_1, \dots, t_n, T), \quad 0 \leq t_1 \leq \dots \leq t_n \leq T, \quad n \in \mathbb{N}. \quad (3.2)$$

Recall that for all  $F \in \mathcal{S}$  of the form (3.1), letting

$$\tilde{f}_n(t_{(1)}, \dots, t_{(n)}) := f_n(t_{(1)}, \dots, t_{(n)}), \quad n \geq 1,$$

where  $(t_{(1)}, \dots, t_{(n)})$  represents the arrangement of  $(t_1, \dots, t_n) \in [0, T]^n$  in increasing order, we have:

$$\begin{aligned} \mathbb{E}_u[F] & \\ &= e^{-u(T)} f_0 + e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \dots \int_0^T \tilde{f}_n(t_1, \dots, t_n) \dot{u}(t_1) \dots \dot{u}(t_n) dt_1 \dots dt_n \end{aligned} \quad (3.3)$$

$$= e^{-u(T)} f_0 + e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \dots, u^{-1}(t_n)) dt_1 \cdots dt_n.$$

From now we assume that  $u \in \mathcal{P}$  is such that  $\dot{u}(t)$  is lower bounded by a (strictly) positive constant for all  $t \in \mathbb{R}_+$  in order to satisfy the integrability conditions needed in the sequel.

**Definition 3.2.** For  $F \in \mathcal{S}$  of the form (3.1), let

$$\dot{D}_t F = - \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \mathbf{1}_{[0, T_k]}(t) \frac{1}{\dot{u}(T_k)} \partial_k f_n(T_1, \dots, T_n), \quad t \in [0, T],$$

for  $F$  of the form (3.1), where  $\partial_k f_n$  denotes the partial derivative of  $f_n$  with respect to its  $k$ -th variable.

Let

$$H = \left\{ v : [0, T] \rightarrow \mathbb{R} : v(t) := \int_0^t \dot{v}(s) ds, \quad t \in [0, T], \quad \dot{v} \in L^2([0, T], \dot{u}(t) dt) \right\}$$

denote the Cameron-Martin space with inner product

$$\langle v, w \rangle_H = \int_0^T \dot{v}(s) \dot{w}(s) \dot{u}(s) ds, \quad v, w \in H.$$

We have

$$\langle DF, v \rangle_H = \int_0^T \dot{v}(t) \dot{D}_t F \dot{u}(t) dt, \quad F \in \mathcal{S}, \quad v \in H.$$

Let  $L_u^2(\Omega; H)$  denote the space of processes  $(v(t))_{t \in [0, T]}$  of the form

$$v(t) = \int_0^t \dot{v}(s) ds, \quad t \in [0, T],$$

such that

$$\mathbb{E}_u \left[ \int_0^T |\dot{v}(s)|^2 \dot{u}(s) ds \right] < \infty.$$

We now turn to the definition of the operator  $\delta$  adjoint of  $D$ . Note that as  $D$  has the derivation property, the operator  $\delta$  is different from the Kabanov-Skorohod integral [5], whose adjoint is a finite difference operator [9]. See [10], [12] for a comparative study of these gradient and Skorohod type integral operators.



**Proposition 3.3.** *i) The operator  $D$  is closable and admits a closable adjoint  $\delta : L_u^2(\Omega; H) \rightarrow L_u^2(\Omega)$  under  $\mathbb{P}_u$ , which satisfies the integration by parts formula*

$$\mathbb{E}_u[F\delta(v)] = \mathbb{E}_u[\langle v, DF \rangle_H], \quad F \in \text{Dom}(D), \quad v \in \text{Dom}(\delta). \quad (3.4)$$

*ii) We have*

$$\delta(v) = \int_0^T \dot{v}(t)(dX_t - \dot{u}(t)dt), \quad (3.5)$$

for every  $\mathcal{F}_t$ -adapted process  $v \in L_u^2(\Omega; H)$ .

*Proof.* By standard integration by parts we first prove (3.4) when  $v \in H$ :

$$\begin{aligned} & \mathbb{E}_u[\langle DF, v \rangle_H] \\ &= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^n \int_0^T \cdots \int_0^T \int_0^{t_k} \frac{\dot{u}(s)}{\dot{u}(t_k)} \dot{v}(s) ds \partial_k \tilde{f}_n(t_1, \dots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\ &= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^n \int_0^{u(T)} \cdots \int_0^{u(T)} \int_0^{u^{-1}(t_k)} \dot{v}(s) \dot{u}(s) ds \frac{\partial}{\partial t_k} \tilde{f}_n(u^{-1}(t_1), \dots, u^{-1}(t_n)) dt_1 \cdots dt_n \\ &= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^n \int_0^{u(T)} \cdots \int_0^{u(T)} \int_0^{t_k} \dot{v}(u^{-1}(s)) ds \frac{\partial}{\partial t_k} \tilde{f}_n(u^{-1}(t_1), \dots, u^{-1}(t_n)) dt_1 \cdots dt_n \\ &= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^n \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \dots, u^{-1}(t_n)) \dot{v}(u^{-1}(t_k)) dt_1 \cdots dt_n \\ &\quad - e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^{u(T)} \dot{v}(u^{-1}(s)) ds \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \dots, u^{-1}(t_{n-1}), T) dt_1 \cdots dt_{n-1} \\ &= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^n \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \dots, t_n) \dot{v}(t_k) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\ &\quad - e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^T \dot{v}(s) \dot{u}(s) ds \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \dots, t_{n-1}, T) \dot{u}(t_1) \cdots \dot{u}(t_{n-1}) dt_1 \cdots dt_{n-1}. \end{aligned}$$

The continuity condition (3.2), i.e.

$$\tilde{f}_{n-1}(t_1, \dots, t_{n-1}) = \tilde{f}_n(t_1, \dots, t_{n-1}, T), \quad n \geq 1,$$

yields

$$\begin{aligned} & \mathbb{E}_u[\langle DF, v \rangle_H] \\ &= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \dots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) \sum_{k=1}^n \dot{v}(t_k) dt_1 \cdots dt_n \end{aligned}$$

$$\begin{aligned}
& -e^{-u(T)} \int_0^T \dot{v}(s) \dot{u}(s) ds \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \dots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\
&= \mathbb{E}_u \left[ F \left( \sum_{k=1}^{X_T} \dot{v}(T_k) - \int_0^T \dot{v}(s) \dot{u}(s) ds \right) \right] \\
&= \mathbb{E}_u \left[ F \left( \int_0^T \dot{v}(s) (dX(s) - \dot{u}(s) ds) \right) \right].
\end{aligned}$$

Next we define  $\delta(Gv)$ ,  $G \in \mathcal{S}$ ,  $v \in H$ , by

$$\delta(Gv) = G \int_0^T \dot{v}(t) (dX_t - \dot{u}(t) dt) - \langle v, DG \rangle_H, \quad (3.6)$$

with for all  $G \in \mathcal{S}$ :

$$\begin{aligned}
\mathbb{E}_u[G \langle DF, v \rangle_H] &= \mathbb{E}_u[\langle D(FG), v \rangle_H - F \langle DG, v \rangle_H] \\
&= \mathbb{E}_u \left[ F \left( G \int_0^T \dot{v}(t) dX_t - \langle DG, v \rangle_H \right) \right] \\
&= \mathbb{E}_u[F \delta(Gv)],
\end{aligned}$$

which proves (3.4). The closability of  $D$  then follows from the integration by parts formula (3.4): if  $(F_n)_{n \in \mathbb{N}} \subset \mathcal{S}$  is such that  $F_n \rightarrow 0$  in  $L_u^2(\Omega)$  and  $DF_n \rightarrow U$  in  $L_u^2(\Omega; H)$ , then (3.4) implies

$$\begin{aligned}
|\mathbb{E}_u[\langle U, Gv \rangle_H]| &\leq |\mathbb{E}_u[F_n \delta(Gv)] - \mathbb{E}_u[\langle U, Gv \rangle_H]| + |\mathbb{E}_u[F_n \delta(Gv)]| \\
&= |\mathbb{E}_u[\langle DF_n - U, Gv \rangle_H]| + |\mathbb{E}_u[F_n \delta(Gv)]| \\
&\leq \|\langle DF_n, v \rangle_H - \langle U, v \rangle_H\|_{L_u^2(\Omega)} \|G\|_{L_u^2(\Omega)} + \|F_n\|_{L_u^2(\Omega)} \|\delta(Gv)\|_{L_u^2(\Omega)},
\end{aligned}$$

$n \in \mathbb{N}$ , hence  $\mathbb{E}_u[\langle U, Gv \rangle_H] = 0$ ,  $G \in \mathcal{S}$ ,  $v \in H$ , i.e.  $U = 0$ . The proof of the closability of  $\delta$  is similar. Finally, by standard arguments we consider processes of the form  $\dot{v} = G \mathbf{1}_{[t, T]}$  where  $G \in \mathcal{S}$  is  $\mathcal{F}_t$ -measurable,  $t \in [0, T]$ , for which we have  $\mathbf{1}_{[t, T]}(s) D_s G = 0$ ,  $s \in [0, T]$ , which shows from (3.6) that

$$\delta(v) = G \int_0^T \mathbf{1}_{[t, T]}(s) (dX_s - \dot{u}(s) ds) = \int_0^T \dot{v}_s (dX_s - \dot{u}(s) ds),$$

hence  $\delta$  extends the Itô integral on all square-integrable  $\mathcal{F}_t$ -adapted processes, and (3.5) is proved.  $\square$

For all  $t \in [0, T]$  we let  $\chi_t(s) = \min(s, t)$ ,  $s \in [0, T]$ .

**Definition 3.4.** *Let*

$$\nabla_t F := \langle DF, \chi_t \rangle_H = \int_0^t \dot{u}(s) \dot{D}_s F ds, \quad F \in \text{Dom}(D).$$

For  $F$  of the form (3.1) we have:

$$\nabla_t F = \int_0^t \dot{D}_s F \dot{u}(s) ds = - \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \frac{u(t \wedge T_k)}{\dot{u}(T_k)} \partial_k f_n(T_1, \dots, T_n).$$

In the parametric case  $u(t) = \lambda h(t)$ ,  $t \in [0, T]$ ,  $h \in \mathcal{P}$ , we have

$$\nabla_t F = - \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \frac{h(t \wedge T_k)}{\dot{h}(T_k)} \partial_k f_n(T_1, \dots, T_n), \quad (3.7)$$

which is independent of  $\lambda$ .

We close this section by introducing a Laplacian on the Poisson space.

**Definition 3.5.** *We define the Laplacian  $\Delta_t$  by*

$$\Delta_t F = \nabla_t \nabla_t F, \quad F \in \mathcal{S}.$$

The operator  $\Delta_t$  is easily shown to be closable, i.e. for any sequence  $(F_n)_{n \in \mathbb{N}}$  of random variables converging to 0 in  $L_u^2(\Omega)$  and such that  $(\Delta_t F_n)_{n \in \mathbb{N}}$  converges in  $L_u^2(\Omega)$ , we have

$$\lim_{n \rightarrow \infty} \Delta_t F_n = 0.$$

This allows one to define the domain of  $\Delta_t$ , denoted by  $\text{Dom}(\Delta_t)$ , as the set of functionals  $F$  for which there exists a sequence of cylindrical functionals  $(F_n)_{n \in \mathbb{N}}$ , which converges in  $L_u^2(\Omega)$  to  $F$  and such that the sequence  $(\Delta_t F_n)_{n \in \mathbb{N}}$  converges in  $L_u^2(\Omega)$ . We will say that a random variable  $F$  in  $\text{Dom}(\Delta_t)$  is  $\Delta_t$ -superharmonic on  $\Omega$  if

$$\Delta_t F(\omega) \leq 0, \quad \omega \in \Omega. \quad (3.8)$$

For example if  $u(t) = \lambda t$ , then for any  $F \in \mathcal{S}$  of the form (3.1) we have

$$\Delta_t F = - \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \nabla_t \sum_{k=1}^n \partial_k f_n(T_1, \dots, T_n) (t \wedge T_k)$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k,l=1}^n (t \wedge T_l)(t \wedge T_k) \partial_k \partial_l f_n(T_1, \dots, T_n) \\
&\quad + \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \mathbf{1}_{[0,t]}(T_k) T_k \partial_k f_n(T_1, \dots, T_n),
\end{aligned}$$

which is independent of  $\lambda$ . Note that due to the presence of first order terms this Laplacian differs from the canonical Laplacian used in the Gaussian case, as a consequence the existence of associated positive superharmonic functions is not conditioned by a lower bound (such as  $n \geq 3$ ) on the number of variables, see the examples in Section 5.

## 4 Stein estimators

Our aim is to construct a superefficient estimator  $\tilde{\lambda}_T$  of  $\lambda$  of the form

$$\hat{\lambda}_T + \frac{\xi_T}{h(T)},$$

whose mean square error will be strictly smaller than the Cramer-Rao bound when  $\xi_T \in L_u^2(\Omega)$  is suitably chosen, where  $\hat{\lambda}_T = X_T/h(T)$  is the MLE of  $\lambda$ . In agreement with Proposition 2.3, this estimator will be biased and anticipating with respect to the Poisson filtration.

The next proposition is our main result on estimation of the intensity parameter  $\lambda > 0$ .

**Proposition 4.1.** *In the parametric case  $u(t) = \lambda h(t)$ ,  $t \in [0, T]$ , for any  $F \in \mathcal{S}$  of the form (3.1) the estimator*

$$\tilde{\lambda}_T := \hat{\lambda}_T - \frac{1}{h(T)} \frac{f_1'(T)}{f_1(T)} \mathbf{1}_{\{X_T=0\}} + \frac{1}{h(T)} \nabla_T \log F,$$

of  $\lambda$ , where  $\nabla_T F$  is given in (3.7), has risk

$$\mathbb{E}_{\lambda h} [|\tilde{\lambda}_T - \lambda|^2] = \frac{\lambda}{h(T)} + \frac{1}{h^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-\lambda h(T)} + \frac{4}{h^2(T)} \mathbb{E}_{\lambda h} \left[ \frac{|\nabla_T \nabla_T \sqrt{F}|}{\sqrt{F}} \right]. \quad (4.1)$$

The proof of Proposition 4.1 will rely on the following two lemmas. First in the next lemma we construct an unbiased risk estimator by applying Stein's integration by parts argument in which we replace (1.1) by the duality relation (3.4) between the gradient and divergence operators on the Poisson space.

**Lemma 4.2.** *Let  $t \in [0, T]$ . For any  $\xi_t \in \text{Dom}(D)$  we have*

$$\mathbb{E}_u [|X_t + \xi_t - u(t)|^2] = u(t) + \|\xi_t\|_{L_u^2(\Omega)}^2 + 2 \mathbb{E}_u [\nabla_t \xi_t]. \quad (4.2)$$

*Proof.* We have

$$\begin{aligned} \mathbb{E}_u [|X_t - u(t) + \xi_t|^2] &= \mathbb{E}_u [|X_t - u(t)|^2] + \|\xi_t\|_{L_u^2(\Omega)}^2 + 2 \mathbb{E}_u [(X_t - u(t))\xi_t] \\ &= u(t) + \|\xi_t\|_{L_u^2(\Omega)}^2 + 2 \mathbb{E}_u [(X_t - u(t))\xi_t]. \end{aligned}$$

We now use the duality relation (3.4) and Relation (3.5) to get

$$\begin{aligned} \mathbb{E}_u [(X_t - u(t))\xi_t] &= \mathbb{E}_u [\delta(\chi_t)\xi_t] \\ &= \mathbb{E}_u [\langle \chi_t, D\xi_t \rangle_H] \\ &= \mathbb{E}_u \left[ \int_0^t \dot{u}(s) \dot{D}_s \xi_t ds \right] \\ &= \mathbb{E}_u [\nabla_t \xi_t], \end{aligned}$$

which yields (4.2). □

The proof of Proposition 4.1 is then a consequence of the following result which applies Lemma 4.2 to processes  $(\xi_t)_{t \in [0, T]}$  of the form

$$\xi_t = c \frac{u(t)}{\dot{u}(T)} \mathbf{1}_{\{X_T=0\}} + \nabla_t \log F, \quad t \in [0, T],$$

where  $c$  is chosen in such a way that  $\xi_t \in \text{Dom}(D)$ ,  $t \in [0, T]$ , and  $F \in \text{Dom}(D)$  is such that  $F > 0$  and  $\sqrt{F} \in \text{Dom}(\Delta_t)$ .

**Lemma 4.3.** *Let  $t \in [0, T]$  and let  $F \in \mathcal{S}$  of the form (3.1) such that  $F > 0$ ,  $\mathbb{P}$ -a.s.,  $F \in \text{Dom}(\Delta_t)$ , and*

$$\partial_n f_n(t_1, \dots, t_{n-1}, T) = 0 \quad \text{and} \quad \partial_k f_n(t_1, \dots, t_n) = \partial_k f_{n+1}(t_1, \dots, t_n, T),$$

$0 \leq t_1 \leq \dots \leq t_n \leq T$ ,  $1 \leq k < n$ ,  $n \geq 2$ . Let also

$$\xi_t := -\frac{u(t)}{\dot{u}(T)} \frac{f_1'(T)}{f_1(T)} \mathbf{1}_{\{X_T=0\}} + \nabla_t \log F. \quad (4.3)$$

Then  $\xi_t \in \mathcal{S} \subset \text{Dom}(D)$  and

$$\mathbb{E}_u [ |X_t + \xi_t - u(t)|^2 ] = u(t) + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} + 4 \mathbb{E}_u \left[ \frac{\Delta_t \sqrt{F}}{\sqrt{F}} \right], \quad t \in [0, T]. \quad (4.4)$$

*Proof.* By construction we have  $\xi_t \in \mathcal{S}$ ,  $t \in [0, T]$ , and from Lemma 4.2:

$$\begin{aligned} \mathbb{E}_u [ |X_t + \xi_t - u(t)|^2 ] &= u(t) + \|\xi_t\|_{L_u^2(\Omega)}^2 + 2 \mathbb{E}_u [ \nabla_t \xi_t ] \\ &= u(t) + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} + \mathbb{E}_u \left[ \left| \frac{\nabla_t F}{F} \right|^2 + 2 \nabla_t \nabla_t \log F \right] \\ &= u(t) + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} + \mathbb{E}_u \left[ 2 \frac{\nabla_t \nabla_t F}{F} - \left| \frac{\nabla_t F}{F} \right|^2 \right] \\ &= u(t) + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} + 4 \mathbb{E}_u \left[ \frac{\nabla_t \nabla_t \sqrt{F}}{\sqrt{F}} \right]. \end{aligned} \quad (4.5)$$

□

*Proof of Proposition 4.1.* Apply (4.3) and (4.4) above at  $t = T$  with  $u(t) = \lambda h(t)$ . □

As a consequence, the  $\Delta_t$ -superharmonicity of  $F$  may imply the superefficiency of  $X + \xi$ . Note also that  $X_t + \xi_t$  may not be positive and replacing  $\xi_t$  with  $\max(X_t + \xi_t, 0)$  will yield a lower risk since the intensity  $\dot{u}$  is known to be positive.

We close this section with some additional remarks.

### Remarks

- a) Relation (4.5) established in proof of Lemma 4.3 shows that the  $\Delta_t$ -superharmonicity of  $F$  implies

$$\mathbb{E}_u [ |X_t + \xi_t - u(t)|^2 ] \quad (4.6)$$

$$\leq u(t) + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} - \mathbb{E}_u [|\nabla_t \log F|^2], \quad t \in [0, T],$$

with equality in (4.6) when  $F$  is  $\Delta_t$ -harmonic. Nevertheless the  $\Delta_t$ -superharmonicity of  $\sqrt{F}$  is a weaker condition.

- b) Note that the risk of any non-adapted estimator  $\zeta_t$  of  $u(t)$  can be lowered by adapted projection, indeed we have

$$\begin{aligned} \mathbb{E}_u[|\mathbb{E}_u[\zeta_t | \mathcal{F}_t] - u(t)|^2] &= \mathbb{E}_u[|\zeta_t - u(t)|^2] - \mathbb{E}_u[|\mathbb{E}_u[\zeta_t | \mathcal{F}_t] - \zeta_t|^2] \\ &< \mathbb{E}_u[|\zeta_t - u_t|^2], \end{aligned} \quad (4.7)$$

for all  $u \in \mathcal{P}$ , and in particular

$$\mathbb{E}_u[X_t + \zeta_t | \mathcal{F}_t] = X_t - \frac{u(t)}{\dot{u}(T)} \frac{f_1'(T)}{f_1(T)} \mathbf{1}_{\{X_t=0\}} e^{-(u(T)-u(t))} + \mathbb{E}_u[\nabla_t \log F | \mathcal{F}_t],$$

$t \in [0, T]$ , which is however dependent on the intensity  $u$ .

- c) Both estimators  $X_t + \xi_t$  and  $X_t + \mathbb{E}_u[\xi_t | \mathcal{F}_t]$  have bias

$$b(t) = \mathbb{E}_u[X_t + \xi_t - u(t)] = \mathbb{E}_u[\xi_t], \quad t \in [0, T],$$

which, using the relation

$$\left| \frac{\nabla_t F}{F} \right|^2 = 2 \frac{\nabla_t \nabla_t F}{F} - \frac{4}{\sqrt{F}} \nabla_t \nabla_t \sqrt{F},$$

can be bounded as follows:

$$\begin{aligned} b^2(t) &= |\mathbb{E}_u[\xi_t]|^2 \\ &\leq \mathbb{E}_u[|\xi_t|^2] \\ &= 2 \mathbb{E}_u \left[ \frac{\nabla_t \nabla_t F}{F} \right] + \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} - 4 \mathbb{E}_u \left[ \frac{\nabla_t \nabla_t \sqrt{F}}{\sqrt{F}} \right], \end{aligned}$$

hence when  $F$  is  $\Delta_t$ -superharmonic we have

$$b^2(t) \leq \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} - 4 \mathbb{E}_u \left[ \frac{\nabla_t \nabla_t \sqrt{F}}{\sqrt{F}} \right], \quad t \in [0, T].$$

d) By integration over  $[0, T]$ , Proposition 2.3 immediately yields, for any unbiased and adapted estimator  $\zeta$  of  $u \in \mathcal{P}$ :

$$\mathbb{E}_u \left[ \int_0^T |\zeta_t - u(t)|^2 dt \right] \geq \|u\|_{L^1([0, T])}, \quad u \in \mathcal{P}, \quad (4.8)$$

where the lower bound  $\|u\|_{L^1([0, T])}$  is attained by  $\zeta = X$ . This bound can be used to derive a nonparametric estimation result for the process  $(u(t))_{t \in [0, T]}$ .

On the other hand, formal maximization of the Girsanov density  $\Lambda(u)$  gives

$$\frac{d}{d\varepsilon} \Lambda(u + \varepsilon v)|_{\varepsilon=0} = \Lambda(u) \int_0^T \frac{\dot{v}(s)}{\dot{u}(s)} (dX_s - du(s)) = 0$$

for all  $v \in H$ , i.e.  $\hat{u}_t = X_t$ . Hence the canonical process  $\hat{u} = (X_t)_{t \in [0, T]}$  be considered as an unbiased maximum likelihood estimator of its own intensity  $(u(t))_{t \in [0, T]}$  under  $\mathbb{P}_u$ , which is efficient in the sense that it attains the Cramer-Rao bound

$$\mathbb{E}_u \left[ \|X - u\|_{L^2([0, T])}^2 \right] = \|u\|_{L^1([0, T])}. \quad (4.9)$$

Given  $(X_t^{(1)})_{t \in [0, T]}, \dots, (X_t^{(N)})_{t \in [0, T]}$ ,  $N$  independent samples of  $(X_t)_{t \in [0, T]}$ , the process

$$\bar{X}_t = \frac{1}{N} \left( X_t^{(1)} + \dots + X_t^{(N)} \right)$$

is a point process with intensity  $u$  under  $\mathbb{P}_u$ , which is consistent as  $T$  tend to 0 and as  $N$  goes to infinity, since by independence we have

$$\mathbb{E}_u \left[ \|\bar{X} - u\|_{L^2([0, T])}^2 \right] = \frac{1}{N^2} \mathbb{E}_u \left[ \sum_{i=1}^N \int_0^T |X_t^{(i)} - u(t)|^2 dt \right] = \frac{1}{N} \int_0^T u(t) dt.$$

Similarly to the above, integration over  $[0, T]$  and Lemma 4.3 show that for  $\xi_t$  defined as in (4.3),  $t \in [0, T]$ ,  $\xi_t \in \mathcal{S}$ ,  $t \in [0, T]$ , and

$$\begin{aligned} \mathbb{E}_u \left[ \|X + \xi - u\|_{L^2([0, T])}^2 \right] &= \|u\|_{L^1([0, T])} + \frac{\|u\|_{L^2([0, T])}^2}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} \\ &\quad + 4 \mathbb{E}_u \left[ \frac{1}{\sqrt{F}} \int_0^T \Delta_t \sqrt{F} dt \right], \end{aligned}$$



hence the  $\Delta_t$ -superharmonicity of  $F$ ,  $t \in [0, T]$ , may imply the superefficiency of  $(X_t + \xi_t)_{t \in [0, T]}$ . Note however that in the general non-parametric case, the estimator  $(X_t + \xi_t)_{t \in [0, T]}$  of  $u$  is dependent on  $u$ .

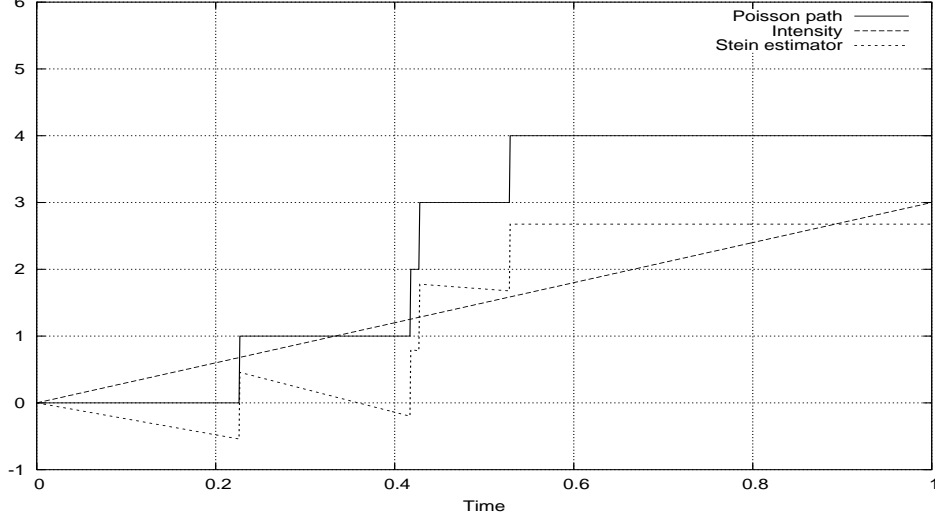


Figure 4.1:  $u(t) = 3t$ ,  $t \in [0, T]$ ;  $N = 5$ .

Figure 4.1 represents a sample path of the process  $X_t + \xi_t$ ,  $t \in [0, T]$  when  $u(t) = \lambda t$ ,  $\lambda = 3$ .

## 5 Examples

In this section we present some examples of estimators satisfying the hypotheses of the previous sections, and we test their superefficiency. In the parametric case  $u(t) = \lambda h(t)$ ,  $t \in [0, T]$ , the percentage gain of an estimator  $\tilde{\lambda}_T$  of  $\lambda$  over the MLE  $\hat{\lambda}_T = X_T/h(T)$  is defined as

$$100 \times \frac{\mathbb{E}_u[|\hat{\lambda}_T - \lambda|^2] - \mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2]}{\mathbb{E}_u[|\hat{\lambda}_T - \lambda|^2]} = 100 \times \frac{\lambda/h(T) - \mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2]}{\lambda/h(T)}.$$

In the sequel we assume that  $u(t) = \lambda t$ ,  $t \in [0, T]$ , hence (4.1) reads

$$\mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2] = \frac{\lambda}{T} + \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} + \frac{4}{T^2} \mathbb{E}_u \left[ \frac{\nabla_T \nabla_T \sqrt{F}}{\sqrt{F}} \right]$$

and  $\tilde{\lambda}_T$  is superefficient, i.e. its gain is positive, provided  $\sqrt{F}$  is  $\Delta_T$ -superharmonic and  $f'_1(T)/f_1(T)$  vanishes or is small enough.

The positive  $\Delta_t$ -superharmonic functionals we consider are of the form

$$\sqrt{F} = \int_0^T g_{N_t}(t) dN_t = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n g_k(T_k),$$

where  $g_k : [0, T] \rightarrow (0, \infty)$ ,  $k \geq 1$ , and

$$g_1(t_1) + \cdots + g_n(t_n) \geq 0, \quad 0 \leq t_1 \leq \cdots \leq t_n \leq T, \quad n \geq 1. \quad (5.1)$$

Then,  $\xi_t$  defined from (4.3) as

$$\begin{aligned} \xi_t &:= -2t \frac{g'_1(T)}{g_1(T)} \mathbf{1}_{\{X_T=0\}} + \nabla_t \log F \\ &= -2t \frac{g'_1(T)}{g_1(T)} \mathbf{1}_{\{X_T=0\}} - 2 \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \frac{(t \wedge T_k) g'_k(T_k)}{g_1(T_1) + \cdots + g_n(T_n)} \\ &= -2t \frac{g'_1(T)}{g_1(T)} \mathbf{1}_{\{X_T=0\}} - \frac{2}{\sqrt{F}} \int_0^T (t \wedge s) g'_{N_s}(s) dN_s, \end{aligned}$$

belongs to  $\text{Dom}(D)$  provided

$$g_k(T) = 0, \quad \text{and} \quad g'_k(T) = 0, \quad k \geq 2, \quad (5.2)$$

and for the condition  $\Delta_T \sqrt{F} \leq 0$  to hold it suffices that

$$g'_k(x) + x g''_k(x) \leq 0, \quad x \in [0, T], \quad k \geq 1. \quad (5.3)$$

a) Let  $g_1(x) = T(1 + \beta) - x$  and  $g_k = 0$ ,  $k \geq 2$ , i.e.

$$F = \mathbf{1}_{\{X_T \geq 1\}} (\beta T + T - T_1)^2,$$

with  $\beta > 0$ . We have from (4.3):

$$\xi_t = \mathbf{1}_{\{X_T=0\}} \frac{2}{\beta T} t + \mathbf{1}_{\{X_T \geq 1\}} \frac{2}{T + \beta T - T_1} (t \wedge T_1),$$

and

$$\nabla_t \nabla_t \sqrt{F} = -T_1 \mathbf{1}_{[0, t]}(T_1) \mathbf{1}_{\{X_T \geq 1\}} \leq 0,$$

hence

$$\begin{aligned}\mathbb{E}_{\lambda h}[|\tilde{\lambda}_T - \lambda|^2] &= \frac{\lambda}{T} + \frac{4}{\beta^2 T^2} e^{-\lambda T} - \frac{4}{T^2} \mathbb{E}_{\lambda h} \left[ \frac{T_1}{\beta T + T - T_1} \right] \\ &= \frac{\lambda}{T} + \frac{4}{\beta^2 T^2} e^{-\lambda T} - \frac{4\lambda}{T^2} \int_0^T \frac{x}{\beta T + T - x} e^{-\lambda x} dx.\end{aligned}$$

The gain of this estimator is equal to the function of  $\lambda T$ :

$$\begin{aligned}\frac{4}{T} \int_0^T \frac{x}{\beta T + T - x} e^{-\lambda x} dx - \frac{4}{\lambda T \beta^2} e^{-\lambda T} &= 4 \int_0^1 \frac{x}{1 + \beta - x} e^{-\lambda T x} dx - \frac{4}{\lambda T \beta^2} e^{-\lambda T} \\ &\geq 4e^{-\lambda T} \left( \int_0^1 \frac{1-x}{1+\beta} e^{\lambda T x} dx - \frac{1}{\lambda T \beta^2} \right) \\ &= 4 \frac{e^{-\lambda T}}{\beta \lambda T} \left( \frac{((\lambda T)^2 - \lambda T + 1)e^{\lambda T} - 1}{(1 + 1/\beta)\lambda T} - \frac{1}{\beta} \right),\end{aligned}$$

which is strictly positive (i.e.  $\tilde{\lambda}_T$  is superefficient) provided  $\beta \geq 2\lambda^{-1}T^{-1}$ . Figure 5.1 represents the gain of  $\tilde{\lambda}_T$  as a function of  $\beta$ .

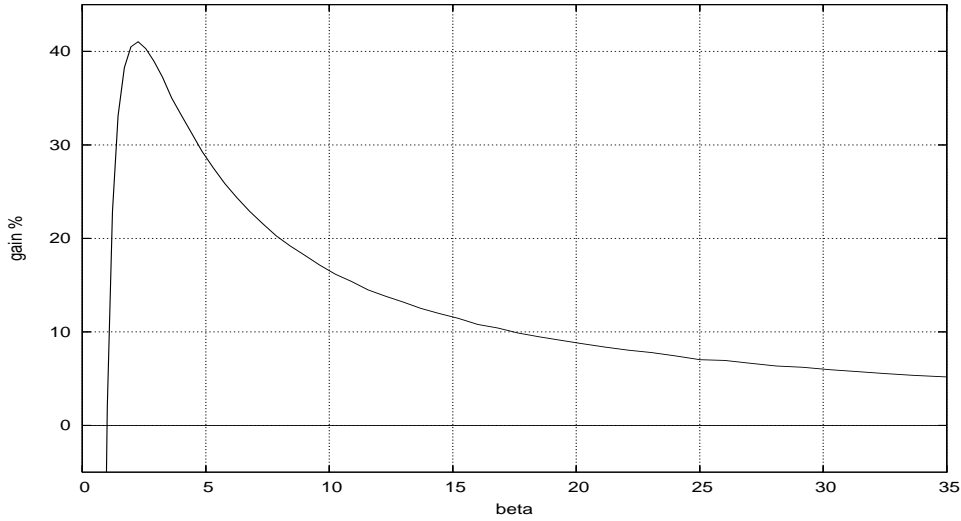


Figure 5.1: Gain as a function of  $\beta$  with  $\lambda = 1$  and  $N = 1$ .

- b) When  $k \geq 2$ , conditions (5.2) and (5.3) are not compatible and as a consequence, superefficiency of  $\tilde{\lambda}_T$  will be dependent on the value of  $\lambda$ . We take

$$g_1(x) = C, \quad g_k(x) = -(-\log((c+x)/(c+T)))^{\alpha_k}, \quad 2 \leq k \leq N,$$

$g_k = 0$ ,  $k > N$ , with  $C \geq \sum_{k=1}^N (-\log(c/(c+T)))^{\alpha_k}$  and  $\alpha_k > 1$ ,  $2 \leq k \leq N$ . In this case,  $\Delta_T g_k$  is not everywhere negative as shown in Figure 5.2, with  $\alpha_2 = 2$ ,

$k = 2$ ,  $T = 1$ , and  $c = 0.01$ , but this suffices to achieve superefficiency for most values of  $\lambda$ , see below.

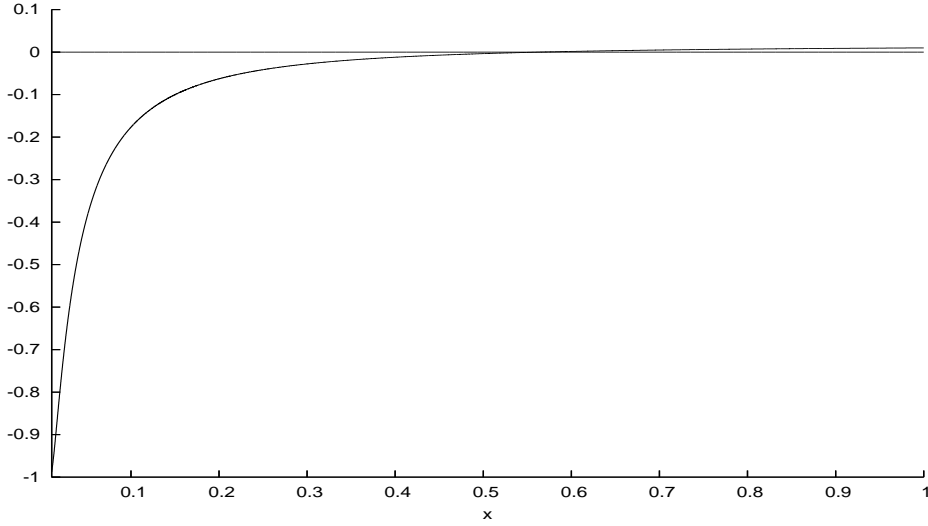


Figure 5.2: Graph of  $\Delta_T g_2$

Figure 5.3 represents the gain of  $\tilde{\lambda}_T$  as a function of  $\lambda$ , with  $10^6$  samples,  $\alpha_2 = \alpha_3 = 2$ ,  $N = 3$ , and  $C = g_2(0) + g_3(0)$ .

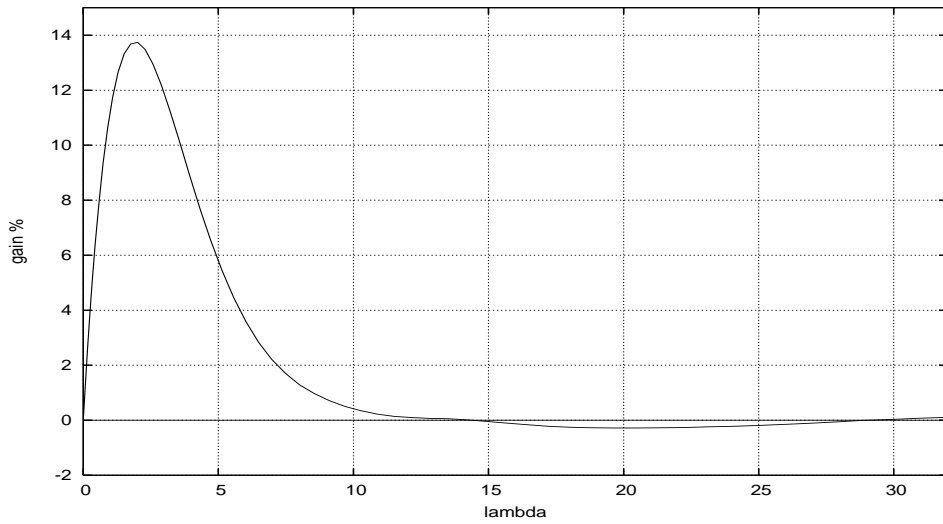


Figure 5.3: Gain as a function of  $\lambda$ .

## References

- [1] R. Averkamp and C. Houdré. Stein estimation for infinitely divisible laws. *ESAIM Probab. Stat.*, 10:269–276 (electronic), 2006.
- [2] E. Carlen and E. Pardoux. Differential calculus and integration by parts on Poisson space. In S. Albeverio, Ph. Blanchard, and D. Testard, editors, *Stochastics, Algebra and Analysis in Classical and Quantum Dynamics (Marseille, 1988)*, volume 59 of *Math. Appl.*, pages 63–73. Kluwer Acad. Publ., Dordrecht, 1990.
- [3] L.H.Y. Chen and A. Xia. Stein’s method, Palm theory and Poisson process approximation. *Ann. Probab.*, 32(3B):2545–2569, 2004.
- [4] R.J. Elliott and A.H. Tsoi. Integration by parts for Poisson processes. *J. Multivariate Anal.*, 44(2):179–190, 1993.
- [5] Y.M. Kabanov. On extended stochastic integrals. *Theory of Probability and its Applications*, XX(4):710–722, 1975.
- [6] Y. Kutoyants. Intensity parameter estimation of an inhomogeneous Poisson process. *Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform.*, 8(2):137–149, 1979.
- [7] Y. Kutoyants. *Statistical inference for spatial Poisson processes*, volume 134 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1998.
- [8] R.S. Liptser and A.N. Shiryaev. *Statistics of random processes. II*, volume 6 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 2001.
- [9] D. Nualart and J. Vives. A duality formula on the Poisson space and some applications. In R. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progress in Probability*, pages 205–213. Birkhäuser, Basel, 1995.
- [10] N. Privault. Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stochastics and Stochastics Reports*, 51:83–109, 1994.
- [11] N. Privault. Calcul des variations stochastique pour les martingales. *C. R. Acad. Sci. Paris Sér. I Math.*, 321:923–928, 1995.
- [12] N. Privault. A calculus on Fock space and its probabilistic interpretations. *Bull. Sci. Math.*, 123(2):97–114, 1999.
- [13] N. Privault and A. Réveillac. Stein estimation for the drift of Gaussian processes using the Malliavin calculus. Preprint, 2006, to appear in the *Annals of Statistics*.
- [14] N. Privault and A. Réveillac. Superefficient estimation on the Wiener space. *C. R. Acad. Sci. Paris Sér. I Math.*, 343:607–612, 2006.
- [15] C. Stein. Estimation of the mean of a multivariate normal distribution. *Ann. Stat.*, 9(6):1135–1151, 1981.