

# Nonparametric tests for Cox processes

Benoît CADRE\*  
IRMAR, ENS Rennes, CNRS  
Campus de Ker Lann  
Avenue Robert Schuman, 35170 Bruz, France  
benoit.cadre@ens-rennes.fr

Gaspar MASSIOT and Lionel TRUQUET  
IRMAR, Ensai, CNRS  
Campus de Ker Lann  
Avenue Robert Schuman, 35170 Bruz, France  
massiot,truquet@ensai.fr

## Abstract

In a functional setting, we elaborate and study two test statistics to highlight the Poisson nature of a Cox process when  $n$  copies of the process are available. Our approach involves a comparison of the empirical mean and the empirical variance of the functional data and can be seen as an extended version of a classical overdispersion test for count data. The limiting distributions of our statistics are derived using a functional central limit theorem for càdlàg martingales. Our procedures are easily implementable and do not require any knowledge on the covariate. We address a theoretical comparison of the asymptotic power of our tests under some local alternatives. A numerical study reveals the good performances of the method. We also present two applications of our tests to real data sets.

*Index Terms* — Functional Statistic; Cox Process; Test Statistic; Local alternative; Nonparametric statistics

*AMS 2010 Classification* — 62C12, 62M07, 60G44.

---

\*Corresponding author

# 1 Introduction

Count process formulation is commonly used to describe and analyze many kind of data in sciences and engineering. A widely used class of such processes is the family of the so-called Cox processes or doubly stochastic Poisson processes. Compared to the standard Poisson process, the key feature of a Cox process is that its arrival rate is stochastic, depending on some covariate. In other words, if  $T > 0$  denotes the observation period,  $N = (N_t)_{t \in [0, T]}$  the Cox process and  $\Lambda = (\Lambda(t))_{t \in [0, T]}$  the (stochastic) cumulative arrival rate then, conditioning on  $\Lambda$ , the distribution of  $N$  is that of a Poisson process with cumulative intensity  $\Lambda$ . Using randomness in the cumulative intensity, the statistician can take into account some auxiliary information, thus leading to a more realistic model. For general references, we refer the reader to the books by Cox and Isham (1980), Karr (1991), Kingman (1993) or Snyder and Miller (1991).

In actuarial sciences and risk theory for instance, the number of claims in the risk model may be represented by a Cox process. In this area, the central quantity is the ruin probability, i.e. the probability that the surplus of the insurer is negative at some time (see e.g., Björk and Grandell, 1988; Grandell, 1991; Schmidli, 1996). Cox process also appears in biophysics and physical chemistry (see e.g., Kou et al., 2005; Kou, 2008; Zhang and Kou, 2010). In these fields, experimental data consist of photon arrival times with the arrival rate depending on the stochastic dynamics of the system under study (for example, the active and inactive states of an enzyme can have different photon emission intensities). By analyzing the photon arrival data, one aims to learn the system's biological properties. Cox process data arise in neuroscience, to analyze the form of neural spike trains, defined as a chain of action potentials emitted by a single neuron over a period of time (see e.g., Gerstner and Kistler, 2002; Reynaud-Bouret et al., 2014). Finally, let us mention astrophysics as another area where Cox process data often occur (see e.g., Scargle, 1998; Carroll and Ostlie, 2007).

In general, it is tempting to incorporate abusively numerous covariates in the statistical model, though a Poisson process model might be satisfactory. In this paper, we elaborate two nonparametric test statistics to highlight the Poisson nature of a Cox process. More precisely, based on i.i.d. copies of  $N$ , we construct two nonparametric test statistics for  $\mathbf{H}_0$ :  $N$  is a Poisson process vs  $\mathbf{H}_1$ :  $N$  is not a Poisson process. This setting of i.i.d. copies of  $N$  is justified by the fact that in many situations, the duration of observation is limited but the number of observed paths is large.

Among the various possibilities for constructing a test statistic devoted to this

problem, a naive approach consists in first estimating both functions  $t \mapsto \mathbb{E}[N_t|\Lambda]$  and  $t \mapsto \mathbb{E}N_t$  and then testing whether these functions are equal. However, this approach suffers from two main drawbacks: the curse of dimensionality (whenever  $\Lambda$  takes values in a high-dimensional space) and the knowledge a priori on  $\Lambda$ . Another approach is to test whether the time-jumps of  $N$  are Poisson time-jumps; in this direction, we refer the reader to the paper by Reynaud-Bourret et al. (2014), in which a modified Kolmogorov-Smirnov statistic is used.

In this paper, we elaborate and study two test statistics, both based on the observation that a Cox process is a Poisson process if, and only if its mean and variance functions are equal. As we shall see, this approach leads to very simple and easily implementable tests.

The paper is organized as follows. In Section 2, we first present the test statistics, then we establish their asymptotic properties. We also compare the asymptotic properties of the tests by using a local alternative. Section 3 is devoted to a simulation study. An application to real data is presented in Section 4. The proofs of our results are postponed to the three last sections of the paper.

## 2 Tests for Cox processes

### 2.1 Principle of the tests

Throughout the paper,  $T > 0$  is the (deterministic) duration of observation, and  $N = (N_t)_{t \in [0, T]}$  is a Cox process with cumulative intensity process  $\Lambda = (\Lambda(t))_{t \in [0, T]}$ , such that the fourth moment of  $N_T$  is finite, i.e.  $\mathbb{E}N_T^4 < \infty$ , and  $\mathbb{E}N_t \neq 0$  for some  $t \in ]0, T[$ . Note that the mean function  $t \mapsto \mathbb{E}N_t = \mathbb{E}\Lambda(t)$  might not be absolutely continuous, so it is not necessarily the integral of an intensity.

We let  $m$  and  $\sigma^2$  the mean and variance functions of  $N$ , i.e. for all  $t \in [0, T]$ :

$$m(t) = \mathbb{E}N_t \text{ and } \sigma^2(t) = \text{var}(N_t).$$

Recall that for all  $t \in [0, T]$  (see p. 66 in the book by Kingman, 1993):

$$\sigma^2(t) = m(t) + \text{var}(\mathbb{E}[N_t|\Lambda]) = m(t) + \text{var}(\Lambda(t)). \quad (2.1)$$

Hence,  $\sigma^2(t) \geq m(t)$  that is, each  $N_t$  is overdispersed. Moreover, if  $m = \sigma^2$ , then  $\mathbb{E}[N_t|\Lambda] = \mathbb{E}N_t$  for all  $t \in [0, T]$ , thus  $N$  is a Poisson process. As a consequence,  $N$  is a Poisson process if, and only if  $m = \sigma^2$ . This observation is the key feature

for the construction of our test statistics. With this respect, the problem can be written as follows:

$$\mathbf{H}_0 : \sigma^2 = m \text{ vs } \mathbf{H}_1 : \exists t \leq T \text{ with } \sigma^2(t) > m(t).$$

From now on, we let the data  $N^{(1)}, \dots, N^{(n)}$  to be independent copies of  $N$ . By above, natural test statistics are based on the process  $\hat{\sigma}^2 - \hat{m} = (\hat{\sigma}^2(t) - \hat{m}(t))_{t \in [0, T]}$ , where  $\hat{m}$  and  $\hat{\sigma}^2$  are the empirical counterparts of  $m$  and  $\sigma^2$ :

$$\hat{m}(t) = \frac{1}{n} \sum_{i=1}^n N_t^{(i)} \text{ and } \hat{\sigma}^2(t) = \frac{1}{n-1} \sum_{i=1}^n (N_t^{(i)} - \hat{m}(t))^2.$$

In this paper, convergence in distribution of stochastic processes is intended with respect to the Skorokhod topology (see Chapter VI in the book by Jacod and Shiryaev, 2003).

Our first main result gives the asymptotic distribution of the process  $\hat{\sigma}^2 - \hat{m}$  under  $\mathbf{H}_0$  (see Section 5 for the proof).

**Theorem 2.1.** *Let  $B = (B_t)_{t \in \mathbb{R}_+}$  be a standard Brownian Motion on the real line. Under  $\mathbf{H}_0$ ,  $\hat{\sigma}^2 - \hat{m}$  is a martingale and*

$$\sqrt{n} (\hat{\sigma}^2 - \hat{m}) \xrightarrow{\text{(law)}} (B_{2m(t)^2})_{t \leq T}.$$

□

As far as we know, the martingale property for  $\hat{\sigma}^2 - \hat{m}$  has not been observed yet. This property, which is interesting by itself, plays a crucial role in the derivation of the asymptotic result.

## 2.2 Testing $\mathbf{H}_0$ vs $\mathbf{H}_1$

Various test statistics based on the process  $\hat{\sigma}^2 - \hat{m}$  can be derived for our purpose. In this paper, we shall concentrate on two test statistics for testing whether a Cox process is Poisson or not. These statistics, denoted by  $\hat{S}_1$  and  $\hat{S}_2$ , involve either the supremum or the integral of the process  $\hat{\sigma}^2 - \hat{m}$ . They are defined as follows:

$$\hat{S}_1 = \sup_{t \leq T} (\hat{\sigma}^2(t) - \hat{m}(t)), \text{ and } \hat{S}_2 = \int_0^T (\hat{\sigma}^2(t) - \hat{m}(t)) dt.$$

These test statistics are natural for testing if a nonnegative function is equal to 0. Moreover they are compatible with the unilateral nature of the problem, since the alternative hypothesis may be written  $\mathbf{H}_1 : \sigma^2(t) > m(t)$  for some  $t \leq T$ .

We now present the asymptotic properties of  $\hat{S}_1$  and  $\hat{S}_2$ .

**Corollary 2.2.** Let  $\hat{I}^2 = \int_0^T (T-t)\hat{m}(t)^2 dt$ .

(i) Under  $\mathbf{H}_0$ ,

$$\sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} \xrightarrow{\text{(law)}} |\mathcal{N}(0,2)|, \text{ and } \sqrt{n} \frac{\hat{S}_2}{\hat{I}} \xrightarrow{\text{(law)}} \mathcal{N}(0,4).$$

(ii) Under  $\mathbf{H}_1$ ,

$$\sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} \xrightarrow{\text{prob.}} +\infty, \text{ and } \sqrt{n} \frac{\hat{S}_2}{\hat{I}} \xrightarrow{\text{prob.}} +\infty$$

□

This result is proved in Section 6. By above, the test statistics  $\hat{S}_1/\hat{m}(T)$  and  $\hat{S}_2/\hat{I}$  define tests with maximal asymptotic power. The rejection regions for tests of level  $\alpha \in ]0, 1[$  are:

$$R_1(\alpha) = \left\{ \frac{\hat{S}_1}{\hat{m}(T)} \geq \sqrt{\frac{2}{n}} q_{1-\alpha/2} \right\} \text{ and } R_2(\alpha) = \left\{ \frac{\hat{S}_2}{\hat{I}} \geq \frac{2}{\sqrt{n}} q_{1-\alpha} \right\}, \quad (2.2)$$

where for each  $\beta \in ]0, 1[$ ,  $q_\beta$  is the  $\mathcal{N}(0,1)$ -quantile of order  $\beta$ .

**Remark.** A close inspection of the proof of Corollary 2.2 reveals that a more general setting may be reached. Indeed, for the test of  $\mathbf{H}_0$  vs  $\mathbf{H}_1$ , we only need to assume that  $N$  is in some class of overdispersed counting processes (i.e.  $\text{var}(N_t) \geq \mathbb{E}N_t$  for all  $t \in [0, T]$ ) which satisfies the property :  $\text{var}(N_t) = \mathbb{E}N_t$  for all  $t \in [0, T]$  if, and only if  $N$  is a Poisson process. The archetype of such a class of counting processes is given by the Cox process. It is also satisfied by other classes, such as some subclasses of Hawkes process for instance. Thus, our test is more or less a functional version of the classical overdispersion test, that is used for testing the Poisson distribution of a sequence of count data (see for instance Rao and Chakravarti, 1956 or Bohning, 1994). Recall that overdispersion tests are widely used in actuarial science for the study of claims counts (e.g. Denuit et al., 2007) .

### 2.3 Local alternative

For testing  $\mathbf{H}_0$  vs  $\mathbf{H}_1$ , we have seen that the statistics  $\hat{S}_1$  and  $\hat{S}_2$  provide tests with maximal asymptotic power. The aim of this section is to proceed to a theoretical comparison between the two tests.

One of the most common way to analyze the asymptotic properties of test statistics when the null hypothesis  $\mathbf{H}_0$  is false is to consider an artificial situation,

represented by a so-called *local alternative*  $\mathbf{H}_1^n$ , under which the process which actually generates the data changes systematically as the sample size is increased. Specifically, it is necessary that as the sample size tends to infinity, the distance between the null hypothesis  $\mathbf{H}_0$  and the alternative  $\mathbf{H}_1^n$  should diminish. With this respect, the asymptotic power under the local alternative turns to be an interesting quantity for comparing several tests. We refer the reader to Engle (1984) and the book by van der Vaart (1998) for a general overview on local alternatives.

In this section, we assume in addition that the Cox process  $N$  has an intensity  $\lambda = (\lambda(t))_{t \in [0, T]}$ , i.e. with probability 1,  $\Lambda$  is absolutely continuous, and

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad \forall t \in [0, T].$$

A classical way to define the local alternative is to consider a *Pitman drift* (see Russell and MacKinnon, 2006, and the references therein), which specifies the intensity as a local deviation from the null hypothesis. We introduce a vanishing sequence of positive numbers  $(d_n)_n$ .

$\mathbf{H}_1^n$  : There exists a bounded non-null function  $\lambda_0 : [0, T] \rightarrow \mathbb{R}_+$  and a process  $\Delta = (\Delta_t)_{t \in [0, T]}$  with  $\sup_{t \in [0, T]} \mathbb{E} \Delta_t^6 < \infty$  and  $\text{var}(\int_0^{t_0} \Delta_s ds) > 0$  for some  $t_0 \in [0, T]$ , and such that  $\lambda = \lambda_0 + d_n \Delta$  w.p. 1.

Here, the assumption on the sixth moment  $\sup_{t \in [0, T]} \mathbb{E} \Delta_t^6 < \infty$  ensures that the Cox process  $N$  satisfies  $\mathbb{E} N_T^6 < \infty$  as well. Moreover, the condition  $\text{var}(\int_0^{t_0} \Delta_s ds) > 0$  for some  $t_0 \in [0, T]$  implies  $\sigma^2(t_0) > m(t_0)$  (see subsection 7.2 for the proof), thus the hypothesis  $\mathbf{H}_1^n$  is contained in  $\mathbf{H}_1$ .

Also observe that when  $n$  gets larger and  $\mathbf{H}_1^n$  holds,  $N$  becomes closer to a Poisson process. Thus,  $(d_n)_n$  has to be understood as a separation rate from  $\mathbf{H}_1^n$  to the null hypothesis  $\mathbf{H}_0$ . In particular, the next result -whose proof has to be found in Section 2.3- states that in view of a consistent test for  $\mathbf{H}_0$  vs  $\mathbf{H}_1^n$ , it is necessary and sufficient that  $d_n^2$  tends to 0 slower than  $1/\sqrt{n}$ .

**Theorem 2.3.** *Let  $B = (B_t)_{t \in \mathbb{R}_+}$  be a standard Brownian Motion on the real line. Assume that  $\mathbf{H}_1^n$  holds, and denote by  $m_0$  and  $v$  the functions defined for all  $t \in [0, T]$  by*

$$m_0(t) = \int_0^t \lambda_0(s) ds, \quad \text{and} \quad v(t) = \text{var}(\int_0^t \Delta_s ds).$$

Moreover, we let  $I_0^2 = \int_0^T (T-t)m_0(t)^2 dt$ .

(i) If  $\sqrt{n}d_n^2 \rightarrow \infty$ , then

$$\sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} \xrightarrow{\text{prob.}} +\infty, \text{ and } \sqrt{n} \frac{\hat{S}_2}{\hat{I}} \xrightarrow{\text{prob.}} +\infty.$$

(ii) If  $\sqrt{n}d_n^2 \rightarrow d < \infty$ , then

$$\begin{aligned} \sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} &\xrightarrow{\text{(law)}} \frac{1}{m_0(T)} \sup_{t \leq T} (B_{2m_0(t)^2} + dv(t)), \text{ and} \\ \sqrt{n} \frac{\hat{S}_2}{\hat{I}} &\xrightarrow{\text{(law)}} 2\mathcal{N}(0, 1) + \frac{d}{I_0} \int_0^T v(t) dt. \end{aligned}$$

□

In the problem  $\mathbf{H}_0$  vs  $\mathbf{H}_1^n$ , we consider the tests defined by the rejection regions in (2.2), with  $\alpha \in ]0, 1[$ . For a power study, we assume from now on that  $\mathbf{H}_1^n$  holds. By above, if  $\sqrt{n}d_n^2 \rightarrow d < \infty$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_2(\alpha)) = 1 - \Phi\left(q_{1-\alpha} - \frac{d}{2I_0} \int_0^T v(t) dt\right) < 1, \quad (2.3)$$

where  $\Phi$  stands for the cumulative distribution function of the  $\mathcal{N}(0, 1)$ -distribution. However, we only have

$$\limsup_{n \rightarrow \infty} \mathbb{P}(R_1(\alpha)) \leq \mathbb{P}\left(\frac{1}{m_0(T)} \sup_{t \leq T} (B_{2m_0(t)^2} + dv(t)) \geq \sqrt{2}q_{1-\alpha/2}\right) < 1,$$

according to the Portmanteau Theorem, as the limit distribution may have a mass at point  $\sqrt{2}q_{1-\alpha/2}$ . At least, we deduce from above and part (i) of Theorem 2.3 that both tests defined by  $R_1(\alpha)$  and  $R_2(\alpha)$  are consistent if, and only if  $\sqrt{n}d_n^2 \rightarrow \infty$ .

In the rest of the section, we assume that  $\sqrt{n}d_n^2 \rightarrow d < \infty$ . For a comparison of the tests, we need an additional assumption ensuring that the limit distribution associated with statistic  $\hat{S}_1$  is continuous. To this aim, we restrict  $\mathbf{H}_1^n$  to the cases where  $\lambda_0(t) > 0$  for all  $t \in [0, T]$ , and we let  $\ell_0$  be the function such that  $\ell_0(t) = 2m_0(t)^2$  for all  $t \in [0, T]$ . Then,  $\ell_0$  is a continuous and increasing function with  $\ell_0(0) = 0$ , and

$$\sup_{t \leq T} (B_{\ell_0(t)} + dv(t)) = \sup_{s \leq \ell_0(T)} (B_s + dv \circ \ell_0^{-1}(s)). \quad (2.4)$$

Now observe that the function  $v \circ \ell_0^{-1}$  is absolutely continuous. Thus, by the Girsanov Theorem (Revuz and Yor, 1999), there exists a probability measure  $Q$  such that the distribution of the random variable in (2.4) equals the distribution under  $Q$  of the supremum of a standard Brownian Motion over  $[0, \ell_0(T)]$ . According to Proposition III.3.7 in the book by Revuz and Yor (1999), this distribution is continuous, which proves that the distribution of the random variable in (2.4) is also continuous. As consequence, the Portmanteau Theorem and Theorem 2.3 (ii) give

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_1(\alpha)) = \mathbb{P}\left(\frac{1}{m_0(T)} \sup_{s \leq \ell_0(T)} (B_s + dv \circ \ell_0^{-1}(s)) \geq \sqrt{2}q_{1-\alpha/2}\right).$$

Unfortunately, the latter probability is not known, except for some special cases that we now study.

We now restrict  $\mathbf{H}_1^n$  to the cases where  $\lambda_0(t) = \lambda_0 > 0$  and  $\Delta_t = Z$  for all  $t \in [0, T]$ , where  $Z$  is a random variable with variance  $w^2$ . In particular, we study the case of a small deviation from an homogeneous Poisson process. Then, formula (2.4) writes

$$\sup_{t \leq T} (B_{\ell_0(t)} + dv(t)) = \sup_{s \leq 2\lambda_0^2 T^2} (B_s + \frac{dw^2}{2\lambda_0^2} s).$$

Obviously, we have in this case  $v(t) = w^2 t^2$ ,  $m_0(t) = \lambda_0 t$  and  $I_0^2 = \lambda_0^2 T^4 / 12$  for all  $t \in [0, T]$ . Setting  $x = dw^2 T$ , we obtain with Theorem 2.3 (ii) and the distribution of the supremum of a drifted Brownian Motion (see p. 250 in the book by Borodin and Salminen, 2002):

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(R_1(\alpha)) &= \exp\left(\frac{\sqrt{2}}{\lambda_0} x q_{1-\alpha/2}\right) \left(1 - \Phi\left(q_{1-\alpha/2} + \frac{x}{\sqrt{2}\lambda_0}\right)\right) \\ &\quad + 1 - \Phi\left(q_{1-\alpha/2} - \frac{x}{\sqrt{2}\lambda_0}\right). \end{aligned}$$

Moreover, by (2.3):

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_2(\alpha)) = 1 - \Phi\left(q_{1-\alpha} - \frac{x}{\sqrt{3}\lambda_0}\right).$$

Assume  $\lambda_0 = 1$  and respectively denote by  $g_1$  and  $g_2$  the above functions of  $x$ . Figure 1 shows a comparison of these two quantities as functions of  $x$  and for  $\alpha = 0.05$  or  $\alpha = 0.1$ . As a conclusion, it suggests a better power for the test induced by



$R_1(\alpha)$ . The numerical experiments given in the next section also suggest a better power for the first test.

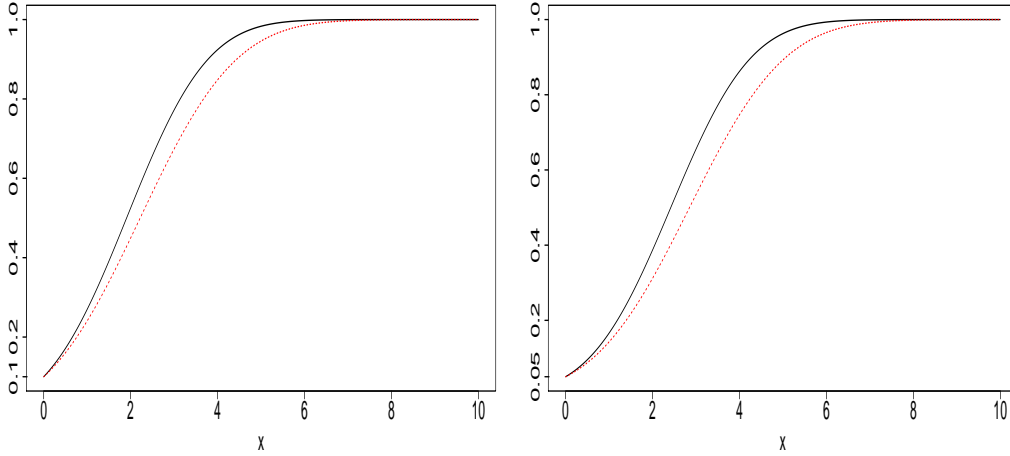


Figure 1: Plots of  $g_1$  (full lines) and  $g_2$  (dashed lines), with  $\alpha = 0.1$  for the curves on the left,  $\alpha = 0.05$  for the curves on the right.

### 3 Simulation study

In this section we illustrate the good properties of our tests with a simulation study. We consider  $n_{MC}$  replications of Monte Carlo simulations and we study the performances of our tests in terms of asymptotic level and rejection power.

Throughout this section, we fix  $T = 1$ . In the following,  $\lambda$  denotes the intensity function of  $N$ , *i.e.* the first derivative of its cumulative intensity function  $\Lambda$ .

#### 3.1 Asymptotic level study

*Model 1.* We first consider the following model for the asymptotic level study,

$$\lambda(t) = \beta t^{\beta-1}, \text{ with } \beta > 0. \quad (3.1)$$

This is the intensity function of a so-called Weibull process, which is frequently used in Reliability Theory for instance. Remark that the function  $\lambda$  is decreasing for  $\beta < 1$ , constant for  $\beta = 1$  and increasing for  $\beta > 1$ .

In Table 1 we evaluate the empirical rejection frequency for both tests using the rejection regions defined in (2.2) for levels  $\alpha = 5\%$  and  $\alpha = 10\%$ . This evaluation is based on  $n_{MC} = 10,000$  Monte Carlo simulations (for each value of  $\beta \in \{1/2, 1, 2\}$ ), with sample sizes  $n = 100$  or  $n = 500$ . We note that test statistic  $\hat{S}_1$  has a similar behavior than test statistic  $\hat{S}_2$  at this range. For both statistics, the empirical rejection frequency is close to the nominal value even with the smallest sample size  $n = 100$ .

$n$	$\beta = 1/2$		$\beta = 1$		$\beta = 2$	
	100	500	100	500	100	500
$\alpha = 5\%$						
$\hat{S}_1$	6.55	5.84	5.99	5.52	5.69	5.56
$\hat{S}_2$	6.27	5.38	5.74	5.44	6.00	5.71
$\alpha = 10\%$						
$\hat{S}_1$	10.96	10.29	10.25	9.95	10.45	10.25
$\hat{S}_2$	10.55	10.05	10.38	10.02	10.29	10.06

Table 1: Empirical rejection frequency (%) for  $n_{MC} = 10,000$  under the model (3.1).

*Model 2.* We define the second model as follows,

$$\lambda(t) = \exp(\theta t), \text{ with } \theta > 0. \quad (3.2)$$

Again, we evaluate in Table 2 the empirical rejection frequency for both tests using the rejection regions defined in (2.2) for levels  $\alpha = 5\%$  and  $\alpha = 10\%$ , and this evaluation is based on  $n_{MC} = 10,000$  Monte Carlo simulations, for each value of  $\theta \in \{1/2, 1, 2\}$  and sample sizes  $n = 100$  or  $n = 500$ . The same remarks as for model (3.1) apply here.

### 3.2 Rejection power study

*Model 1.* For the asymptotic power study, we first consider the model defined by

$$\lambda(t) = \exp(\theta Zt), \quad (3.3)$$

$n$	$\theta = 1/2$		$\theta = 1$		$\theta = 2$	
	100	500	100	500	100	500
$\alpha = 5\%$						
$\hat{S}_1$	5.95	5.55	5.81	5.74	6.15	5.41
$\hat{S}_2$	5.57	5.26	5.67	5.54	5.68	5.43
$\alpha = 10\%$						
$\hat{S}_1$	10.74	10.23	10.42	10.61	11.20	10.00
$\hat{S}_2$	10.19	10.40	10.14	10.56	10.60	9.77

Table 2: Empirical rejection frequency (%) for  $n_{MC} = 10,000$  under the model (3.2).

with  $Z \sim 2 + \text{Beta}(\frac{1}{2}, \frac{1}{2})$  and  $\theta \in [0, 1]$ . Figure 2 represents the empirical rejection frequency for different values of  $\theta$  between 0 and 1 in model (3.3). For  $\theta = 0$ , the simulated process is an homogeneous Poisson process with intensity 1. The process deviates from the homogeneous Poisson process as  $\theta$  increases. We observe on Figure 2 that both tests catch this behavior for small  $\theta$ 's. In both cases the power goes to 1 for higher values of  $\theta$ . The test statistic  $\hat{S}_1$  has better performances, its power curve being always above that of test statistic  $\hat{S}_2$ .

*Model 2.* We define the second model as follows,

$$\lambda(t) = \exp(\theta \sin(Z_t)), \quad (3.4)$$

with  $(Z_t)_{t \in [0,1]}$  a standard Brownian Motion and  $\theta \in [0, 1]$ . This model differs from the previous one as the covariate depends on the time variable. Figure 3 represents the empirical rejection frequency for  $\theta$  varying in  $[0, 3]$  in model (3.4). The same remarks as for model (3.3) apply here. Note that the two test statistics look closer on this example.

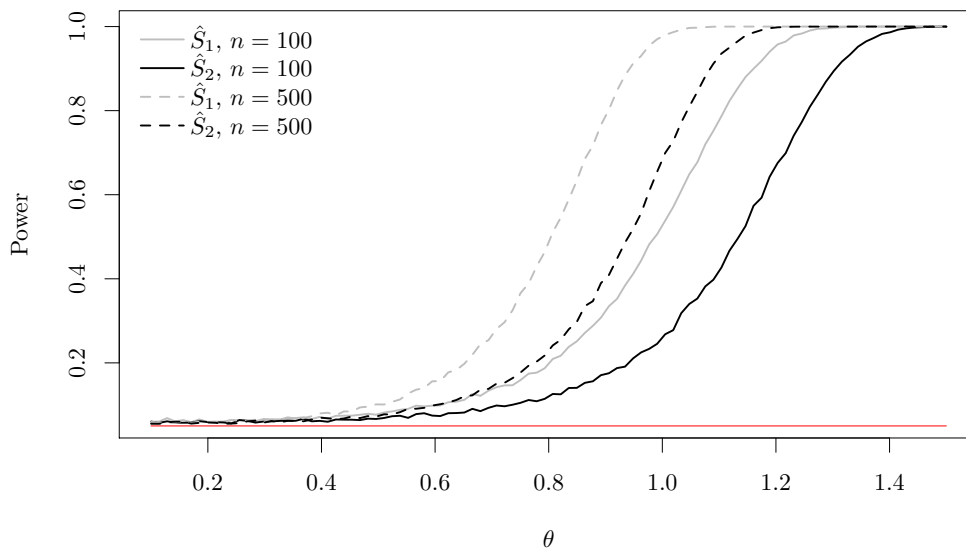


Figure 2: Empirical rejection frequency under (3.3) for  $n_{MC} = 10,000, 100$  and  $500$  trajectories and  $\alpha = 5\%$ . The red horizontal line represents the value of  $\alpha$ .

## 4 Application to real data

### 4.1 Analysis of some arrival times in a call center

The use of Poisson processes has been often considered as a first approach for modeling the arrival times in call centers and more generally queue systems. See Asmussen (2003) for the nice theory developed around this assumption.

As in the papers by Brown et al. (2005) and Mandelbaum et al. (2000), we consider a call center for an anonymous Israel's bank. A description of the calls received over the year 1999 is available online<sup>1</sup>. The call center is open on week-days (Sunday to Thursday in Israel), from 7 A.M. to midnight, and the calls are separated in different classes, depending on the needs of the customers. Each call can be described as follows. A customer calls one of the phone numbers of the

<sup>1</sup><http://ie.technion.ac.il/serveng/callcenterdata/>

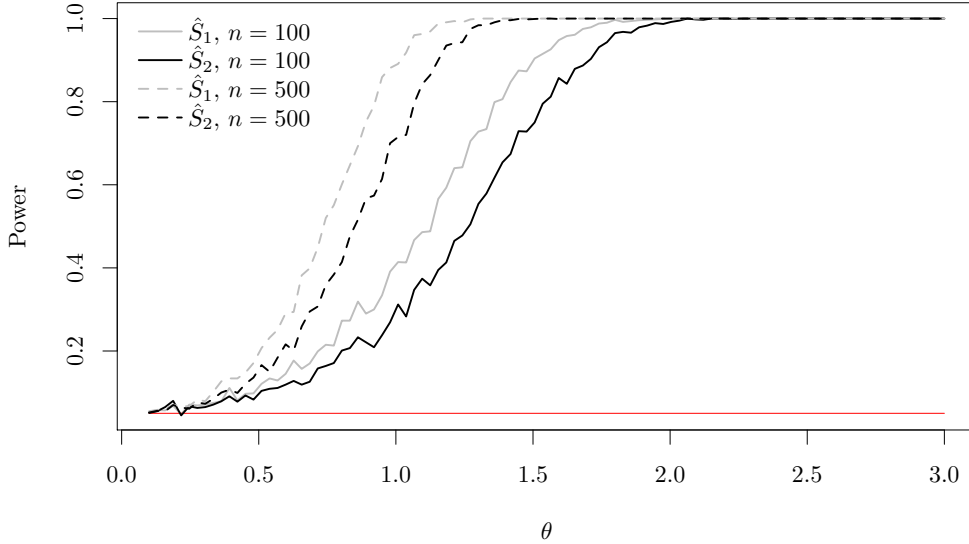


Figure 3: Empirical rejection frequency under (3.4) for  $n_{MC} = 10,000, 100$  and  $500$  trajectories and  $\alpha = 5\%$ . The red horizontal line represents the value of  $\alpha$ .

call center. Except for rare busy signals, the customer is then connected to an interactive voice respond unit (IVR or VRU) and identifies himself/herself. While using the VRU, the customer receives recorded information. He/She can choose to perform some self-service transactions ( $\sim 65\%$  of the calls) or indicate the need to speak with an agent ( $\sim 35\%$ ). Here, we are interested in the latter, which represents roughly 30,000 to 40,000 calls per month. Each call record in the database also includes a categorical description of the type of service requested. The main call types are regular (PS in the database), stock transaction (NE), new or potential customer (NW), and Internet assistance (IN). Mandelbaum et al. (2000) and Brown et al. (2005) described the process of collecting and cleaning the data and provided complete descriptive analysis of the data.

In this study, we concentrate on IN calls recording between 3:25 P.M. and 3:35 P.M. all weekdays of year 1999. Times at which calls enter the VRU represent the arrival times of a counting process. The dataset then consists in 258 trajectories

that we can assume to be realizations of *i.i.d.* Cox processes.

The results of the statistical study are presented in Table 3. One can see that the null hypothesis  $\mathbf{H}_0$  is highly rejected using both statistics. This result suggests that even on a short period of time, these arrival times, which depend on complex human behavior, are strongly influenced by some covariates. One might easily imagine that weather conditions or other company intrinsic variables (*e.g.* number of recent opened accounts) could reduce this overdispersion and help to explain the number of IN phone calls.

	$\hat{S}_1$	$\hat{S}_2$
$p$ -values	$1.95 \times 10^{-6}$	$1.05 \times 10^{-6}$

Table 3:  $p$ -values of both tests for the call center dataset.

Finally mention that Brown et al. (2005) also studied IN calls but did not reject the Poisson assumption. However, their study consists in testing the exponential distribution for the interarrival times of IN calls occurring in a single day, which is not compatible with our asymptotic and cannot help to determine if some covariates influence the daily calls process.

## 4.2 Analysis of the scoring times of a soccer team

As seen in Heuer et al. (2010), Poisson processes may also be used to model scoring goals during a soccer match. Nevertheless, one could suspect the influence of some covariates such as the behavior of the spectators or fitness fluctuations of the team under study. Thus, we propose to test the Poisson process assumption  $\mathbf{H}_0$  for the scoring times of Arsenal soccer club first team.

To this end, we collected on the SoccerSTATS.com website their scoring times (in minutes) for each match in "Premier League" over three seasons (from 2012 to 2015). For each match, the scoring times of the team define the jump times of a counting process, giving a total of 229 counting paths. We can assume that these data are *i.i.d.* realizations of Cox processes.

The results of the statistical study are presented in Table 4. For both statistics  $\hat{S}_1$  and  $\hat{S}_2$ , we cannot reject the null hypothesis  $\mathbf{H}_0$  and the Poisson process seems to be a reasonable approximation for these counting processes.

Recall that the analysis in Heuer et al. (2010) also suggests that the Poisson process is relevant for modeling the scoring goals of a given team in the German Bundesliga.

	$\hat{S}_1$	$\hat{S}_2$
$p$ -values	0.419	0.298

Table 4:  $p$ -values of both tests for the soccer goals dataset.

## 5 Proof of Theorem 2.1

In the rest of the paper, we assume for notational simplicity that  $T = 1$ . Let  $\hat{M} = (\hat{M}_t)_{t \in [0,1]}$  be the process defined by  $\hat{M}_t = \hat{\sigma}^2(t) - \hat{m}(t)$  and  $\hat{\tau} = (\hat{\tau}_t)_{t \in [0,1]}$  the process such that for all  $t \in [0, 1]$ ,

$$\hat{\tau}_t = \frac{4}{n-1} \int_0^t \hat{\sigma}^2(s) dm(s).$$

Martingale properties and predictable  $\sigma$ -field are implicitly with respect to the natural filtration generated by the sample  $N^{(1)}, \dots, N^{(n)}$ . As usual,  $\langle X \rangle$  stands for the predictable quadratic variation of a martingale  $X$ .

### 5.1 Auxiliary results

**Lemma 5.1.** *Under  $\mathbf{H}_0$ , the process  $\hat{M}$  is a martingale, and  $\langle \hat{M} \rangle = \hat{\tau}$ .*

**Proof.** First we prove that  $\hat{M}$  is a martingale. Observe that for all  $t \in [0, 1]$ ,

$$\hat{\sigma}^2(t) = \frac{1}{n-1} \sum_{k=1}^n (N_t^{(k)})^2 - \frac{n}{n-1} \hat{m}(t)^2. \quad (5.1)$$

We now decompose  $\hat{M}_t$  as a sum of stochastic integrals. In the sequel, we let  $\bar{N}^{(k)} = N^{(k)} - m$ . Note that  $\bar{N}^{(k)}$  is a martingale. According to the integration by parts formula (Proposition 0.4.5 in the book by Revuz and Yor, 1999), we have

$$\begin{aligned} \sum_{k=1}^n (N_t^{(k)})^2 &= \sum_{k=1}^n \left[ 2 \int_0^t N_{s^-}^{(k)} dN_s^{(k)} + N_t^{(k)} \right] \\ &= 2 \sum_{k=1}^n \int_0^t N_{s^-}^{(k)} dN_s^{(k)} + n \hat{m}(t) \\ &= 2 \sum_{k=1}^n \int_0^t N_{s^-}^{(k)} d\bar{N}_s^{(k)} + 2n \int_0^t \hat{m}(s^-) dm(s) + n \hat{m}(t). \end{aligned} \quad (5.2)$$

Moreover, by the integration by parts formula,

$$\hat{m}(t)^2 = 2 \int_0^t \hat{m}(s^-) d\hat{m}(s) + \sum_{s \leq t} (\Delta \hat{m}(s))^2.$$

Using the fact that two independent Poisson processes do not jump at the same time (Proposition XII.1.5 in the book by Revuz and Yor, 1999), we deduce that

$$\begin{aligned} \sum_{s \leq t} (\Delta \hat{m}(s))^2 &= \frac{1}{n^2} \sum_{s \leq t} \left( \sum_{k=1}^n \Delta N_s^{(k)} \right)^2 = \frac{1}{n^2} \sum_{s \leq t} \sum_{k=1}^n \Delta N_s^{(k)} \\ &= \frac{1}{n} \hat{m}(t). \end{aligned}$$

Hence,

$$\hat{m}(t)^2 = 2 \int_0^t \hat{m}(s^-) d\hat{m}(s) + \frac{1}{n} \hat{m}(t).$$

Then, combining (5.1) and (5.2), we obtain

$$\begin{aligned} \hat{M}_t &= -\hat{m}(t) + \frac{1}{n-1} \sum_{k=1}^n (N_t^{(k)})^2 - \frac{n}{n-1} \hat{m}(t)^2 \\ &= \frac{2}{n-1} \sum_{k=1}^n \int_0^t N_{s^-}^{(k)} d\bar{N}_s^{(k)} - \frac{2n}{n-1} \int_0^t \hat{m}(s^-) d(\hat{m}(s) - m(s)). \end{aligned} \quad (5.3)$$

Since  $\hat{m} - m$  and each of the  $\bar{N}^{(k)}$ 's are martingales and the integrands are predictable, we deduce that all the integrals in this formula are local martingales. It is a classical exercise to prove that they are of class DL (see Definition IV.1.6 in the book by Revuz and Yor, 1999), so that they are martingales, as well as  $\hat{M}$ .

In view of computing the predictable quadratic variation of  $\hat{M}$ , we first observe that by the integration by parts formula,

$$\hat{M}_t^2 = 2 \int_0^t \hat{M}_{s^-} d\hat{M}_s + \sum_{s \leq t} (\Delta \hat{M}_s)^2. \quad (5.4)$$

But, by (5.3),

$$\Delta \hat{M}_s = \frac{2}{n-1} \sum_{k=1}^n N_{s^-}^{(k)} \Delta N_s^{(k)} - \frac{2n}{n-1} \hat{m}(s^-) \Delta \hat{m}(s).$$

Again, we shall make use of the fact that two Poisson processes do not jump at the same time. Hence, if  $s$  is a time-jump for  $N^{(k)}$ ,

$$\Delta \hat{M}_s = \frac{2}{n-1} (N_{s^-}^{(k)} - \hat{m}(s^-)) = \frac{2}{n-1} (N_{s^-}^{(k)} - \hat{m}(s^-)) \Delta N_s^{(k)}. \quad (5.5)$$



Thus,

$$\begin{aligned}\sum_{s \leq t} (\Delta \hat{M}_s)^2 &= \sum_{s \leq t} \sum_{k=1}^n (\Delta \hat{M}_s)^2 \mathbf{1}_{\{\Delta N_s^{(k)}=1\}} \\ &= \frac{4}{(n-1)^2} \sum_{k=1}^n \int_0^t (N_{s^-}^{(k)} - \hat{m}(s^-))^2 dN_s^{(k)}.\end{aligned}$$

By (5.4), we have

$$\begin{aligned}\hat{M}_t^2 &= 2 \int_0^t \hat{M}_{s^-} d\hat{M}_s + \frac{4}{(n-1)^2} \sum_{k=1}^n \int_0^t (N_{s^-}^{(k)} - \hat{m}(s^-))^2 d\bar{N}_s^{(k)} \\ &\quad + \frac{4}{(n-1)^2} \sum_{k=1}^n \int_0^t (N_{s^-}^{(k)} - \hat{m}(s^-))^2 dm(s).\end{aligned}$$

As above, we can conclude from the fact that both  $\hat{M}$  and  $\bar{N}^{(k)}$  are martingales that the first two terms on the right-hand side are martingales. Last term, namely

$$\frac{4}{n-1} \int_0^t \hat{\sigma}^2(s^-) dm(s) = \frac{4}{n-1} \int_0^t \hat{\sigma}^2(s) dm(s),$$

where equality holds by continuity of  $m$ , is predictable. Hence, it is the predictable quadratic variation of  $\hat{M}$ .  $\square$

**Lemma 5.2.** *Under  $\mathbf{H}_0$ , we have :*

$$n \mathbb{E} \sup_{t \leq 1} |\Delta \hat{M}_t|^2 \rightarrow 0.$$

**Proof.** First observe that for all  $u \in ]0, 1[$  :

$$\begin{aligned}n \mathbb{E} \sup_{t \leq 1} |\Delta \hat{M}_t|^2 &\leq u + n \mathbb{E} \sup_{t \leq 1} |\Delta \hat{M}_t|^2 \mathbf{1}_{\{\sqrt{n} \sup_{t \leq 1} |\Delta \hat{M}_t| > u\}} \\ &\leq u + 2 \int_0^\infty x \mathbb{P}(\sqrt{n} \sup_{t \leq 1} |\Delta \hat{M}_t| \mathbf{1}_{\{\sqrt{n} \sup_{t \leq 1} |\Delta \hat{M}_t| > u\}} > x) dx \\ &\leq 2u + 2 \int_u^\infty x \mathbb{P}(\sqrt{n} \sup_{t \leq 1} |\Delta \hat{M}_t| > x) dx.\end{aligned}\tag{5.6}$$

But, according to (5.5), if  $t$  is a time-jump for  $N^{(k)}$ , we have

$$\Delta \hat{M}_t = \frac{2}{n-1} (N_{t^-}^{(k)} - \hat{m}(t^-)),$$

and hence,

$$\sup_{t \leq 1} |\Delta \hat{M}_t| \leq \frac{2}{n-1} \sup_{k \leq n} \sup_{t \leq 1} |N_t^{(k)} - \hat{m}(t)|.$$

Thus, for all  $x > 0$  :

$$\begin{aligned} \mathbb{P}(\sqrt{n} \sup_{t \leq 1} |\Delta \hat{M}_t| \geq x) &\leq n \mathbb{P}\left(\sup_{t \leq 1} |N_t^{(1)} - \hat{m}(t)| \geq \frac{(n-1)x}{2\sqrt{n}}\right) \\ &\leq 8 \sup_{t \leq 1} \mathbb{E}|N_t^{(1)} - \hat{m}(t)|^3 \frac{n^{5/2}}{(n-1)^3 x^2}, \end{aligned}$$

where the last inequality is due to Doob's Inequality (see Revuz and Yor, 1999) applied to the martingale  $N^{(1)} - \hat{m}$ . A direct calculation shows that exists a constant  $C > 0$  (independent of  $n$ ) such that

$$\sup_{t \leq 1} \mathbb{E}|N_t^{(1)} - \hat{m}(t)|^3 \leq C.$$

Consequently, by (5.6) :

$$\begin{aligned} n \mathbb{E} \sup_{t \leq 1} |\Delta \hat{M}_t|^2 &\leq 2u + \frac{16Cn^{5/2}}{(n-1)^3} \int_u^\infty \frac{dx}{x^2} \\ &\leq 2u + \frac{16Cn^{5/2}}{u(n-1)^3}. \end{aligned}$$

Taking for instance  $u = n^{-1/4}$  gives the result.  $\square$

## 5.2 Proof of Theorem 2.1

According to Theorem VIII.3.22 in the book by Jacod and Shiryaev (2003), the sequence of square integrable martingales  $(\sqrt{n}\hat{M})_n$  converges in distribution to a continuous Gaussian martingale  $M$  such that  $\langle M \rangle = 2m^2$  if, for all  $t \in [0, 1]$  and  $\varepsilon > 0$ ,

$$\langle \sqrt{n}\hat{M} \rangle_t \rightarrow 2m(t)^2 \text{ and } \int_{\mathbb{R} \times [0, t]} |x|^2 \mathbf{1}_{\{|x| > \varepsilon\}} \nu_n(dx, ds) \rightarrow 0, \quad (5.7)$$

both in probability, where  $\nu_n$  stands for the predictable compensator of the random jump measure associated to the martingale  $\sqrt{n}\hat{M}$ . Regarding the first property, we know from Lemma 5.1 that

$$\lim_{n \rightarrow \infty} \langle \sqrt{n}\hat{M} \rangle_t = \lim_{n \rightarrow \infty} n \hat{\tau}_t = 4 \int_0^t \sigma^2(s) dm(s),$$

in probability. Since  $\sigma^2 = m$  under  $\mathbf{H}_0$ , we deduce that  $\lim_{n \rightarrow \infty} \langle \sqrt{n} \hat{M} \rangle_t = 2m(t)^2$  in probability. In order to prove the second property in (5.7), we fix  $\varepsilon > 0$  and we let  $U$  and  $V$  be the processes defined for all  $t \in [0, 1]$  by

$$U_t = \int_{\mathbb{R} \times [0, t]} |x|^2 \mathbf{1}_{\{|x| > \varepsilon\}} \nu_n(dx, ds) \text{ and } V_t = n \sum_{s \leq t} |\Delta \hat{M}_s|^2 \mathbf{1}_{\{\sqrt{n} |\Delta \hat{M}_s| > \varepsilon\}}.$$

Observing that  $U$  is L-dominated by the increasing adapted process  $V$ , we deduce from the Lenglart Inequality (see p.35 in the book by Jacod and Shiryaev, 2003) that for all  $t \in [0, 1]$  and  $\alpha, \eta > 0$ :

$$\mathbb{P}(U_t \geq \eta) \leq \frac{1}{\eta} (\alpha + \mathbb{E} \sup_{s \leq t} \Delta V_s) + \mathbb{P}(V_t \geq \alpha).$$

But,  $\{V_t > 0\} = \{\sqrt{n} \sup_{s \leq t} |\Delta \hat{M}_s| > \varepsilon\}$  and  $\sup_{s \leq t} \Delta V_s \leq n \sup_{s \leq t} |\Delta \hat{M}_s|^2$ . Thus, letting  $\alpha \searrow 0$ , we obtain with the help of Markov's Inequality:

$$\begin{aligned} \mathbb{P}(U_t \geq \eta) &\leq \frac{n}{\eta} \mathbb{E} \sup_{s \leq t} |\Delta \hat{M}_s|^2 + \mathbb{P}(\sqrt{n} \sup_{s \leq t} |\Delta \hat{M}_s| > \varepsilon) \\ &\leq \left( \frac{1}{\eta} + \frac{1}{\varepsilon^2} \right) n \mathbb{E} \sup_{s \leq t} |\Delta \hat{M}_s|^2. \end{aligned}$$

We conclude from Lemma 5.2 that  $U_t$  converges to 0 in probability. Hence, both properties in (5.7) are satisfied so that the sequence of square integrable martingales  $(\sqrt{n} \hat{M})_n$  converges in distribution to a continuous Gaussian martingale  $M$ , with  $\langle M \rangle = 2m^2$ . The Dambis-Dubins-Schwarz Theorem (see Theorem V.1.6 in the book by Revuz and Yor, 1999) then gives  $M = B_{2m^2}$ , where  $B$  is a standard real Brownian Motion.  $\square$

## 6 Proof of corollary 2.2

(i) Let  $\mathbb{D}$  be the space of càdlàg functions from  $[0, 1]$  to  $\mathbb{R}$ , equipped with the Skorokhod topology. By continuity of the application  $\mathbb{D} \ni x \mapsto \sup_{t \leq T} x(t)$ , we deduce from Theorem 2.1 that

$$\sqrt{n} \hat{S}_1 = \sqrt{n} \sup_{t \leq 1} (\hat{\sigma}^2(t) - \hat{m}(t)) \xrightarrow{\text{(law)}} \sup_{t \leq 1} B_{2m(t)^2} = \sup_{t \leq 2m(1)^2} B_t.$$

According to the reflection principle (Proposition III.3.7 in the book by Revuz and Yor, 1999), the distribution of the latter term is  $\sqrt{2}m(1) |\mathcal{N}(0, 1)|$ . Hence the

result follows for the statistics  $\hat{S}_1$ . Similarly, by continuity of  $\mathbb{D} \ni x \mapsto \int_0^1 x(t)dt$ , we have

$$\sqrt{n}\hat{S}_2 = \sqrt{n} \int_0^1 (\hat{\sigma}^2(t) - \hat{m}(t))dt \xrightarrow{(\text{law})} \int_0^1 B_{2m(t)}^2 dt,$$

and the distribution of the limit is  $\mathcal{N}(0, 4 \int_0^1 (1-t)m(t)^2 dt)$ . Moreover, using the fact that  $\hat{m} - m$  is a martingale, we easily prove with Doob's Inequality that  $\sup_{t \leq 1} |\hat{m}(t) - m(t)|$  converges in probability to 0. Putting all pieces together and applying Slutsky's Lemma gives the result.

(ii) Under  $\mathbf{H}_1$ , there exists  $t_0 \in [0, 1]$  such that  $\sigma^2(t_0) > m(t_0)$ . Then,

$$\begin{aligned} \sqrt{n}\hat{S}_1 &\geq \sqrt{n}(\hat{\sigma}^2(t_0) - \hat{m}(t_0)) \\ &\geq \sqrt{n}(\hat{\sigma}^2(t_0) - \sigma^2(t_0)) + \sqrt{n}(m(t_0) - \hat{m}(t_0)) + \sqrt{n}(\sigma^2(t_0) - m(t_0)). \end{aligned}$$

The latter term tends to  $+\infty$ , while the Central Limit Theorem (that can be used because  $\mathbb{E}N_1^4 < \infty$ ) shows that the sequences induced by the first two terms on the right-hand side are stochastically bounded, hence the result with  $\hat{S}_1$ . Regarding  $\hat{S}_2$ , we first observe that under  $\mathbf{H}_1$ ,  $\int_0^1 (\sigma^2(t) - m(t))dt > 0$ , because  $\sigma$  and  $m$  are right-continuous functions and  $\sigma^2 \geq m$ . Thus, we only need to prove that the sequences  $(\sqrt{n} \int_0^1 (m(t) - \hat{m}(t))dt)_n$  and  $(\sqrt{n} \int_0^1 (\hat{\sigma}^2(t) - \sigma^2(t))dt)_n$  are stochastically bounded. Let us focus on the second sequence. We have

$$\begin{aligned} \sqrt{n} \int_0^1 (\hat{\sigma}^2(t) - \sigma^2(t))dt &= \sqrt{n} \left( \frac{1}{n-1} \sum_{i=1}^n \int_0^1 (N_t^{(i)})^2 dt - \mathbb{E} \int_0^1 N_t^2 dt \right) \\ &\quad - \sqrt{n} \int_0^1 \left( \frac{n}{n-1} \hat{m}(t)^2 - m(t)^2 \right) dt. \end{aligned} \quad (6.1)$$

Since  $\mathbb{E}N_1^4 < \infty$ , the sequence induced by the first term on the right-hand side is stochastically bounded according to the Central Limit Theorem. Regarding the latter term in (6.1), we observe that

$$\begin{aligned} \sqrt{n} \left| \int_0^1 \left( \frac{n}{n-1} \hat{m}(t)^2 - m(t)^2 \right) dt \right| &\leq 2\sqrt{n}(\hat{m}(1) + m(1)) \int_0^1 |\hat{m}(t) - m(t)| dt \\ &\quad + m(1)^2. \end{aligned}$$

By the Cauchy-Schwarz Inequality, there exists a constant  $C > 0$  such that the  $L^1$ -norm of the leftmost term is bounded by

$$\begin{aligned} C \left[ 1 + \sqrt{n} \mathbb{E}^{1/2} \left( \int_0^1 |\hat{m}(t) - m(t)| dt \right)^2 \right] &\leq C \left[ 1 + \left( \int_0^1 \text{var}(N_t) dt \right)^{1/2} \right] \\ &\leq C(1 + \mathbb{E}N_1^2). \end{aligned}$$

Thus,  $(\sqrt{n} \int_0^1 (\hat{\sigma}^2(t) - \sigma^2(t)) dt)_n$  is stochastically bounded.  $\square$

## 7 Proof of Theorem 2.3

In this section, we assume that  $\mathbf{H}_1^n$  holds, hence in particular  $N$  is a Cox process with intensity  $\lambda = \lambda_0 + d_n \Delta$  that depends on  $n$ .

For simplicity, we let  $Z^{(n)} = (Z_t^{(n)})_{t \in [0,1]}$  the centered process defined for all  $t \in [0, 1]$  by

$$Z_t^{(n)} = \sqrt{n}(\hat{\sigma}^2(t) - \sigma^2(t) + m(t) - \hat{m}(t)).$$

### 7.1 Auxiliary results

**Lemma 7.1.** *There exists  $C > 0$  independent of  $n$  such that for all  $s, t \in [0, 1]$ ,*

$$\mathbb{E}|N_t - N_s|^6 \leq C|t - s|^6.$$

**Proof.** Without loss of generality, we assume that  $s \leq t$ . Recall that the distribution of  $N_t - N_s$  given  $\Lambda$  follows a Poisson distribution with parameter  $\Lambda(t) - \Lambda(s) = \int_s^t \lambda(u) du$ , and that the sixth moment of a Poisson distribution with parameter  $\mu > 0$  is bounded by some constant  $C > 0$  multiplied by  $\mu^6$ . Thus, using Jensen's Inequality, we get

$$\begin{aligned} \mathbb{E}(N_t - N_s)^6 &= \mathbb{E} \mathbb{E}[(N_t - N_s)^6 | \Lambda] \\ &\leq C \mathbb{E} \left( \int_s^t \lambda(u) du \right)^6 \\ &\leq C(t-s)^5 \int_s^t \mathbb{E} \lambda^6(u) du \\ &\leq 2^6 C \left( \sup_{t \in [0,1]} \lambda_0(t)^6 + \sup_{t \in [0,1]} \mathbb{E} |\Delta_t^6| \right) (t-s)^6, \end{aligned}$$

hence the lemma.  $\square$

**Lemma 7.2.** *For all  $t \in [0, 1]$ ,  $((Z_t^{(n)})^2)_n$  is a uniformly integrable sequence.*

**Proof.** According to the Rosenthal Inequality and Lemma 7.1, there exists a constant  $C > 0$  that does not depend on  $n$  such that

$$n^{3/2} \mathbb{E} |\hat{m}(t) - m(t)|^3 \leq C \text{ and } n^{3/2} \mathbb{E} |\hat{\sigma}^2(t) - \sigma^2(t)|^3 \leq C.$$

Thus, we deduce that  $\sup_n \mathbb{E} |Z_t^{(n)}|^3 < \infty$ , which implies that  $((Z_t^{(n)})^2)_n$  is a uniformly integrable sequence.  $\square$

**Lemma 7.3.** *The sequence of processes  $(Z^{(n)})_n$  is tight.*

**Proof.** For an integrable real random variable  $Z$ , we let  $\{Z\} = Z - \mathbb{E}Z$ . First observe that we have, for all  $t \in [0, 1]$  :

$$Z_t^{(n)} = \sqrt{n} \left( \frac{1}{n-1} \mathbb{E}N_t^2 + \frac{1}{n-1} \sum_{i=1}^n \{(N_t^{(i)})^2\} - \{\hat{m}(t)\}^2 + \{\hat{m}(t)\}(1 - 2m(t)) \right).$$

We shall make use of the classical criterion of tightness (see e.g. Theorem 13.6 in the book by Billingsley, 1999). Clearly, the sequence will be proved to be tight if each of the sequences of processes defined by

$$X^{1,n} = \sqrt{n}\{\hat{m}\}, X^{2,n} = \sqrt{n}\{\hat{m}\}^2 \text{ and } X^{3,n} = \frac{\sqrt{n}}{n-1} \sum_{i=1}^n \{(N_t^{(i)})^2\}$$

satisfy the inequality

$$\mathbb{E}(X_t^{k,n} - X_s^{k,n})^2 \leq C(F(t) - F(s))^2, \quad \forall 0 \leq s \leq t \leq 1, \quad (7.1)$$

for a constant  $C > 0$  and some nondecreasing and continuous function  $F$ , both independent of  $n$ . We only prove it for  $X^{3,n}$ . In the sequel,  $C > 0$  is a constant, independent of  $n$ , and whose value may change from line to line. Observe that

$$\begin{aligned} \mathbb{E}(X_t^{3,n} - X_s^{3,n})^2 &= \frac{n}{n-1} \mathbb{E}(\{N_t^2\} - \{N_s^2\})^2 \\ &\leq C \mathbb{E}(N_t^2 - N_s^2)^2 + C(\mathbb{E}N_t^2 - \mathbb{E}N_s^2)^2 \\ &\leq C\{\mathbb{E}(N_t - N_s)^4\}^{1/2} + C\mathbb{E}(N_t - N_s)^2, \end{aligned}$$

by Cauchy-Schwarz's Inequality. Then, Lemma 7.1 gives

$$\mathbb{E}(X_t^{3,n} - X_s^{3,n})^2 \leq C(t-s)^2.$$

Consequently, (7.1) holds for  $k = 3$ , with the continuous and nondecreasing function  $F(t) = t$ .  $\square$

**Lemma 7.4.** *Let  $B = (B_t)_{t \in \mathbb{R}_+}$  be a real and standard Brownian Motion. Then, if  $m_0$  is the function defined for all  $t \in [0, 1]$  by  $m_0(t) = \int_0^t \lambda_0(u) du$ , we have*

$$Z^{(n)} \xrightarrow{\text{(law)}} (B_{2m_0^2(t)})_{t \in [0,1]}.$$

**Proof.** In the sequel, for  $p = 1$  or  $2$ , we let

$$\hat{m}_p(t) = \frac{1}{n} \sum_{i=1}^n (N_t^{(i)})^p \text{ and } m_p(t) = \mathbb{E}N_t^p.$$

Let  $k \geq 1$  and  $0 \leq t_1 < \dots < t_k \leq 1$ . According to the Central Limit Theorem for triangular arrays (for instance the Lyapunov condition is easily seen to be true according to Lemma 7.1), we know that the  $2k$ -dimensional random vector defined by

$$\sqrt{n} \begin{pmatrix} \hat{m}_1(t_j) - m_1(t_j) \\ \hat{m}_2(t_j) - m_2(t_j) \end{pmatrix}_{j=1, \dots, k}$$

converges to a normal distribution. Now apply the  $\delta$ -method to deduce that the  $3k$ -dimensional random vector

$$\sqrt{n} \begin{pmatrix} \hat{m}_1(t_j) - m_1(t_j) \\ \hat{m}_2(t_j) - m_2(t_j) \\ \hat{m}_1(t_j)^2 - m_1(t_j)^2 \end{pmatrix}_{j=1, \dots, k}$$

also converges to a normal distribution. Thus,

$$\sqrt{n} (\hat{m}_2(t_j) - \hat{m}_1(t_j)^2 - m_2(t_j) + m_1(t_j)^2 - \hat{m}_1(t_j) + m_1(t_j))_{j=1, \dots, k},$$

converges to a  $k$ -dimensional normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ , as well as  $(Z_{t_j}^{(n)})_{j=1, \dots, k}$ . Since for all  $j = 1, \dots, k$ ,  $\hat{m}(t_j)$  and  $\hat{\sigma}^2(t_j)$  are unbiased estimators of  $m(t_j)$  and  $\sigma^2(t_j)$ ,  $\mathbb{E}Z_{t_j}^{(n)} = 0$ . Thus, by Lemma 7.2,  $\mu = 0$ . We now proceed to compute the variance matrix  $\Sigma$ . Let  $i, j = 1, \dots, k$ . With the notation  $\{Z\} = Z - \mathbb{E}Z$  for an integrable real random variable  $Z$ , we easily see that the difference between  $\mathbb{E}Z_{t_i}^{(n)}Z_{t_j}^{(n)}$  and

$$n\mathbb{E}(\{\hat{m}_2(t_i)\} - \{\hat{m}_1^2(t_i)\} - \{\hat{m}_1(t_i)\})(\{\hat{m}_2(t_j)\} - \{\hat{m}_1^2(t_j)\} - \{\hat{m}_1(t_j)\})$$

tends to 0 as  $n \rightarrow \infty$ . Moreover, the difference between the latter term and

$$\frac{1}{n} \mathbb{E} \left( \sum_{\ell=1}^n \{(N_{t_i}^{(\ell)})^2 - N_{t_i}^{(\ell)}(2m(t_i) + 1)\} \right) \left( \sum_{\ell=1}^n \{(N_{t_j}^{(\ell)})^2 - N_{t_j}^{(\ell)}(2m(t_j) + 1)\} \right),$$

denoted by  $A$ , vanishes as well. But, by independence of the processes  $(N^{(\ell)})_{\ell \leq n}$ , we have

$$\begin{aligned} A &= \mathbb{E}(\{N_{t_i}^2\} - \{N_{t_i}\}(2m(t_i) + 1))(\{N_{t_j}^2\} - \{N_{t_j}\}(2m(t_j) + 1)) \\ &= \text{cov}(N_{t_i}^2 - (2m(t_i) + 1)N_{t_i}, N_{t_j}^2 - (2m(t_j) + 1)N_{t_j}). \end{aligned}$$

Recall that, under  $\mathbf{H}_1^n$ ,  $N$  is a Cox process with intensity  $\lambda = \lambda_0 + d_n \Delta$ . Since  $(d_n)_n$  tends to 0, easy calculations prove that, as  $n \rightarrow \infty$ ,  $A$  converges to

$$\text{cov}(P_{t_i}^2 - (2m_0(t_i) + 1)P_{t_i}, P_{t_j}^2 - (2m_0(t_j) + 1)P_{t_j}), \quad (7.2)$$

where  $P$  is a Poisson process with intensity  $\lambda_0$ . Now use the properties of the Poisson process and the fact that, for a random variable  $Q$  that follows a Poisson distribution with parameter  $\mu > 0$ , we have

$$\mathbb{E}(Q - \mu)^2 = \mathbb{E}(Q - \mu)^3 = \mu \text{ and } \mathbb{E}(Q - \mu)^4 = \mu + 3\mu^2.$$

Moreover, for all  $t \in [0, 1]$ ,  $P_t^2 - (2m_0(t) + 1)P_t + m_0(t)^2 = \{P_t\}^2 - \{P_t\}$ . We easily deduce from the above properties and the independence of the increments of a Poisson process that the covariance in (7.2) equals  $2m_0(t_i \wedge t_j)^2$ . Thus, by Lemma 7.2, the  $(i, j)$  term of matrix  $\Sigma$  is given by the previous formula. The sequence of processes  $(Z^{(n)})_n$  being tight according to Lemma 7.3, we have proved that  $Z^{(n)}$  converges in distribution to a centered Gaussian process  $Z$  such that if  $s, t \in [0, 1]$ ,  $\mathbb{E}Z_s Z_t = 2m_0(s \wedge t)^2$ . Such a Gaussian process can be written  $B_{2m_0^2}$  where  $B$  is a standard Brownian Motion on the line, hence the result.  $\square$

## 7.2 Proof of Theorem 2.3

Recall that, by assumption,  $v(t) \neq 0$  for some  $t \in [0, 1]$ . Observe that according to (2.1), we have under  $\mathbf{H}_1^n$ , for all  $t \in [0, 1]$  :

$$\begin{aligned} \sigma^2(t) - m(t) &= \text{var}(\Lambda(t)) \\ &= \text{var}\left(\int_0^t (\lambda_0(s) + d_n \Delta_s) ds\right) \\ &= d_n^2 v(t). \end{aligned}$$

Thus,

$$\sqrt{n} \hat{S}_1 = \sup_{t \leq 1} (Z_t^{(n)} + \sqrt{n} d_n^2 v(t)).$$

Moreover, it is an easy exercise to prove that  $\hat{m}(T) \rightarrow m_0(T)$  and  $\hat{I} \rightarrow I_0$ , both in probability. If  $\sqrt{n} d_n^2 \rightarrow \infty$ , we then have by Lemma 7.4,

$$\sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} \xrightarrow{\text{prob.}} +\infty,$$



and similarly for  $\hat{S}_2$ , hence (i). We now assume that  $\sqrt{nd_n^2} \rightarrow d < \infty$ . By Lemma 7.4 and Slutsky's Lemma, the continuity of the underlying functional gives :

$$\sqrt{n} \frac{\hat{S}_2}{\hat{I}} \xrightarrow{\text{(law)}} \frac{1}{I_0} \int_0^1 (B_{2m_0(t)^2} + dv(t)) dt.$$

Observing now that the distribution of the latter term equals

$$2\mathcal{N}(0, 1) + \frac{d}{I_0} \int_0^1 v(t) dt$$

gives the result for  $\hat{S}_2$ . Regarding  $\hat{S}_1$ ,

$$\sqrt{n} \frac{\hat{S}_1}{\hat{m}(T)} \xrightarrow{\text{(law)}} \frac{1}{m_0(T)} \sup_{t \leq 1} (B_{2m_0(t)^2} + dv(t)),$$

hence the theorem.  $\square$

**Acknowledgments.** We thank the Action Editor and two referees for valuable comments and insightful suggestions, which led to a substantial improvement of the paper.

## References

- Billingsley, P. (1999). *Convergence of Probability Measures, 2nd Ed.*, Wiley, New-York.
- Asmussen, S. (2003). *Applied Probability and Queues, 2nd Ed.*, Springer, New-York.
- Björk, T. and Grandell, J. (1988). Exponential inequalities for ruin probabilities in the Cox case, *Scandinavian Actuarial Journal*, 77-111.
- Bohning, D. (1994). A Note on a Test for Poisson Overdispersion, *Biometrika*, 418-419.
- Borodin, A.N, Salminen, P. (2002). *Handbook of Brownian Motion-Facts and Formulae, 2nd Ed.*, Springer, New-York.
- Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., and Zhao, L. (2005). Statistical analysis of a telephone call center: A queueing-science perspective, *Journal of the American statistical association*, 100(469), 36-50.

- Carroll, B. and Ostlie, D. (2007). *An Introduction to Modern Astrophysics*, 2nd ed. Benjamin Cummings, Reading, MA.
- Cox, D.R. and Isham, V. (1980). *Point Processes*, Chapman and Hall, London.
- Russell, D. and MacKinnon, J.G. (2006). The Power of Bootstrap and Asymptotic Tests, *Journal of Econometrics*, 133, 421-441.
- Denuit, M., Maréchal, X., Pitrebois, S., and Walhin, J. F. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*, John Wiley & Sons.
- Engle, R. F. (1984). Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics, in *Handbook of Econometrics*, Vol. 2, ed. by Z. Griliches and M. Intriligator, Amsterdam.
- Gerstner, W. and Kistler, W. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge Univ. Press, Cambridge.
- Grandell, J. (1991). *Aspects of Risk Theory*, Springer-Verlag, New-York.
- Heuer, A., Mueller, C., and Rubner, O. (2010). Soccer: Is scoring goals a predictable Poissonian process ? *EPL (Europhysics Letters)*, 89(3), 380-7.
- Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd Ed., Springer, New-York.
- Karr, A.F. (1991). *Point Processes and their Statistical Applications*, 2nd Ed., Marcel Dekker, New-York.
- Kingman, J.F.C. (1993). *Poisson Processes*, Oxford Studies in Probability, Oxford.
- Kou, S. C. (2008). Stochastic networks in nanoscale biophysics: Modeling enzymatic reaction of a single protein, *J. Amer. Statist. Assoc.*, 961-5.
- Kou, S. C., Xie, X. S. and Liu, J. S. (2005). Bayesian analysis of single-molecule experimental data (with discussion), *J. Roy. Statist. Soc. Ser. C*, 469-6.
- Mandelbaum, A., Sakov, A., and Zeltyn, S. (2000). Empirical analysis of a call center, URL <http://iew3.technion.ac.il/serveng/References/ccdata.pdf>. *Technical Report*.
- Rao, C. R., and Chakravarti, I. M. (1956). Some small sample tests of significance for a Poisson distribution, *Biometrics*, 12(3), 264-282.
- Reynaud-Bourret, P., Rivoirard, V., Grammont, F., Tuleau-Malot, C. (2014). Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis, *Journal of Mathematical Neuroscience*, 4:3.

- Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion, 3rd Ed.*, Springer, New-York.
- Scargle, J. D. (1998). Studies in astronomical time series analysis. V. Bayesian blocks, a new method to analyze structure in photon counting data, *Astrophys. J.*, 405-8.
- Schmidili, H. (1996). Lundberg inequalities for a Cox model with a piecewise constant intensity, *J. Applied Probability*, 196-210.
- Snyder, D. L. and Miller, M. I. (1991). *Random Point Processes in Time and Space*, Springer, New-York.
- van der Vaart, A.W. (1998). *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics.
- Zhang, T. and Kou, S.C. (2010). Nonparametric inference of doubly stochastic Poisson process data via the kernel method, *Ann. Applied Statist.*, 1913-1941.