On the Cramér-von Mises test with parametric hypothesis for Poisson process

Ali Souleymane DABYE[1] email: dabye_ali @ yahoo.fr

[1]Laboratoire d'Études et de Recherches en Statistique et Développement(LERSTAD)

Université Gaston BERGER, Saint-Louis (Sénégal)

1. Introduction

☞Inhomogeneous Poisson process is a mathematecal model of the series of events which is quite rich to fit in many applied problems and at the same time this class of processes is sufficiently simple and allows to solve many statistical problems. Therefore the problem of goodness of fit testing for this model is important and our work is devoted to this problem. In this work we consider the problem of goodness of fit testing for inhomogeneous Poisson process with composite parametric basic hypothesis.
The problems of goodness of fit for Poisson processes were considered by many authors. Let us mention here the works [5], [2],[3], [6], [1],. In the last work it was shown that for the model of inhomogeneous Poisson processes in the case of singular estimation the limit distribution of the Cramér-von Mises tests is asymptotically distribution free.
☞ In the present work we study the goodness of fit test with parametric basic hypotheses in the situation when the unknown parameter is a shift parameter. We show that the limit distribution of the Cramér-von Mises statistics under hypothesis does not depend on the unknown parameter.

2. A CLASSICAL RESULT

Suppose that we observe $X^{(n)} = (X_1, \dots, X_n)$, n independent Poisson processes, where $X_j = (X_j(t), t \in)$ are trajectories of the Poisson process with mean function $\Lambda(t) = X_j(t)$. Remind that if the basic hypothesis is simple, say,

$$H_0 \quad : \quad \Lambda(\cdot) = \Lambda_0(\cdot) \qquad H_1 \quad : \quad \Lambda(\cdot) \neq \Lambda_0(\cdot)$$

where $\Lambda_0(\cdot)$ is a known function with $\Lambda_0(\infty) < \infty$ and alternative then we can introduce the Cramér-von Mises type statistic

$$\tilde{\Delta}_n = \frac{n}{\Lambda_0(\infty)^2} \int_{-\infty}^{+\infty} \left[\widehat{\Lambda}_n(t) - \Lambda_0(t)\right]^2 d\Lambda_0(t), \quad \widehat{\Lambda}_n(t) = \frac{1}{n}\sum_{j=1}^{n} X_j(t)$$

is the empirical mean of the Poisson process. It can be verified that this statistic converges to the following limit

$$\tilde{\Delta}_n \Longrightarrow \Delta \equiv \int_0^1 W(s)^2 \, ds.$$

Here $W(s), 0 \le s \le 1$ is a standard Wiener process. Therefore the test

$$\tilde{\psi}_n(X^n) = 1\!\!1_{\{\tilde{\Delta}_n > c_\varepsilon\}}, \qquad \mathbb{P}\{\Delta > c_\varepsilon\} = \varepsilon,$$

is asymptotically distribution free (see, e.g., Dachian and Kutoyants [3]).
We are interested by the same problem but with the parametric basic hypothesis, i.e., we suppose that under hypothesis H_0 the mean function belongs to a parametric family of functions and we propose a test of Cramér-von Mises type which is "partially distribution free". This means that the limit distribution of the statistic does not depend on the unknown parameter. This result allows us to construct a test with asymptotically chosen probability of errors (under hypothesis). This test is as well consistent against any fixed alternative.

3. MAIN RESULT

NOTATION

Let us introduce a parametric family

$$\mathcal{L}(\Theta) = \{\Lambda_0(t - \vartheta), \vartheta \in \Theta\}, \quad \Theta = (\alpha, \beta), -\infty < \alpha < \beta < \infty$$

where $\Lambda_0(\cdot)$ is some known nondecreasing function with properties:

$$\Lambda_0(-\infty) = 0, \quad \Lambda_0(\infty) < \infty,$$

and absolutely continuous :

$$\Lambda_0(t) = \int_{-\infty}^{t} \lambda_0(s) \, ds.$$

We observe $X^{(n)} = (X_1, \dots, X_n)$, n independent inhomogeneous Poisson processes, $X_j = \{X_j(t), t \in\}$ with the same mean function $\Lambda(\cdot)$ and we have to test a composite parametric hypothesis

$$H_0 \quad : \quad \Lambda(\cdot) \in \mathcal{L}(\Theta) \qquad H_1 \quad : \quad \Lambda(\cdot) \notin \mathcal{L}(\Theta),$$

More precizely, we suppose that under this alternative

$$\inf_{\vartheta \in \Theta} \|\Lambda(t) - \Lambda_0(t - \vartheta)\|_\vartheta > 0.$$

Here and in the sequel $\|\cdot\|$ is the following L_2-norme:

$$\|h\|_\vartheta^2 = \int_{-\infty}^{\infty} h(t)^2 \lambda_0(t - \vartheta) \, dt.$$

We show that for such alternatives the test is consistent. This test will be uniformly consistent against another class of alternatives:

$$H_1^\rho \quad : \quad \Lambda(\cdot) \in \mathcal{F}_\rho = \left\{\Lambda(\cdot) : \inf_{\vartheta \in \Theta} \|\Lambda(t) - \Lambda_0(t - \vartheta)\|_\vartheta > \rho\right\}, \qquad .$$

Here $\rho > 0$ is some given number. We suppose as well that $\mathcal{F}_\rho$ is such that $\sup_{\Lambda \in \mathcal{F}_\rho} \Lambda(\infty) < \infty$.
In our case when the value ϑ is unknown and we replace it by its maximum likelihood estimator $\hat{\vartheta}_n$:

$$\Delta_n = n \int_{-\infty}^{+\infty} \left[\widehat{\Lambda}_n(t) - \Lambda_0(t - \hat{\theta}_n)\right]^2 \lambda_0(t - \hat{\theta}_n) \, dt.$$

Therefore the Cramér-von Mises type test is

$$\hat{\Psi}_n(X^n) = 1\!\!1_{\{\Delta_n > c_\varepsilon\}}.$$

The threshold c_ε must be chosen so that this test belongs to the class of tests of asymptotic level ε:

$$\mathcal{K}_\varepsilon = \left\{\bar{\Psi}_n : \qquad \lim_{n \to \infty} {}_\vartheta \bar{\Psi}_n = \varepsilon, \quad \vartheta \in \Theta\right\}$$

As we use the MLE $\hat{\vartheta}_n$ we need the following regularity conditions : we suppose that the intensity function $\lambda_0(\cdot)$ is strictly positive and sufficiently smooth. Under these conditions the MLE is consistent, asymptotically normal and the polynomial moments converge (see [5]).
Note that the Fisher information

$$I_n(\vartheta) = n \int_{-\infty}^{+\infty} \frac{\dot{\lambda}_0^2(t - \vartheta)}{\lambda_0(t - \vartheta)} dt = n \int_{-\infty}^{+\infty} \frac{\dot{\lambda}_0^2(t)}{\lambda_0(t)} dt = n\mathrm{I}_0$$

i.e., does not depend on ϑ. Here and in the sequel dot means differentiation with respect to θ. Let us introduce the following random variable

$$\Delta_0 = \int_{-\infty}^{\infty} \left[W(\Lambda_0(t)) + \frac{\lambda_0(t)}{\mathrm{I}_0} \int_{-\infty}^{\infty} \frac{\dot{\lambda}_0(s)}{\lambda_0(s)} dW(\Lambda_0(s))\right]^2 d\Lambda_0(t)$$

where $W(\cdot)$ is a standard Wiener process. The constant c_ε is solution of the equation

$$\mathbb{P}\{\Delta_0 > c_\varepsilon\} = \varepsilon.$$

CONDITIONS

- **a1.** *The function* $\sqrt{\lambda_0(\cdot)} \in \mathcal{L}_2(R)$ *is strictly positive and three times continuously differentiable.*
- **a2.** *Its derivatives belong to* $\mathcal{L}_2(R)$ *. The Fisher information* $\mathrm{I}_0 > 0$.
- **a3.** *For any* $\nu > 0$ *we have*

$$\inf_{|\vartheta - \vartheta_0| > \nu} \left\|\sqrt{\lambda_0(\cdot - \vartheta)} - \sqrt{\lambda_0(\cdot - \vartheta_0)}\right\| > 0.$$

- **a4.** *The derivative* $\dot{\Lambda}_0(\cdot) \in \mathcal{L}_1(R)$.

Here $\|\cdot\|$ is the usual $\mathcal{L}_2(R)$ norm. Note that under conditions **a1**- **a4** the MLE $\hat{\vartheta}_n$ is consistent, asymptotically normal

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \mathrm{I}_0^{-1}\right)$$

and the moments converge : for any $p > 0$

$$n_\vartheta^{p/2} \left|\hat{\vartheta}_n - \vartheta\right|^p \longrightarrow |\zeta|^p, \quad \zeta \sim \mathcal{N}\left(0, \mathrm{I}_0^{-1}\right).$$

Moreover, it admits the representation

$$\hat{\vartheta}_n = \vartheta + \frac{1}{\sqrt{n}\mathrm{I}_0} \int_{-\infty}^{\infty} \frac{\dot{\lambda}_0(t - \vartheta)}{\lambda_0(t - \vartheta)} dW_n(t) + O\left(n^{-3/4}\right)$$

where $W_n(t) = \sqrt{n}\left(\hat{\Lambda}_n(t) - \Lambda_0(t - \vartheta)\right)$. For the proofs see [5].

THEOREM1

Let the conditions **a1**- **a4** be fulfilled. Then the test

$$\hat{\Psi}_n = 1\!\!1_{\{\Delta_n > c_\varepsilon\}} \in \mathcal{K}_\varepsilon.$$

PROOF

Let us put $h(v) = I_0^{-1} \frac{\dot{\lambda}_0(v)}{\lambda_0(v)}$ and denote by ϑ_0 the true value. Then we can write

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \int_{-\infty}^{+\infty} h(s - \vartheta_0) \, dW_n(s) + O(n^{-1/4}).$$

Hence

$$\begin{aligned} u_n(t) &= \sqrt{n}\left(\widehat{\Lambda}_n(t) - \Lambda_0(t - \hat{\theta}_n)\right) \\ &= W_n(t) - \sqrt{n}\left(\Lambda_0(t - \hat{\theta}_n) - \Lambda_0(t - \theta_0)\right). \end{aligned} \tag{1}$$

According (1), we have the representation

$$\begin{aligned} u_n(t) &= W_n(t) + \lambda_0(t - \theta_0) \cdot (\hat{\theta}_n - \theta_0)\sqrt{n} + o\left(\sqrt{n}(\hat{\theta}_n - \theta_0)\right) \\ &= W_n(t) + \lambda_0(t - \theta_0) \int_{-\infty}^{+\infty} h(s - \theta_0) \, dW_n(s) + r_n(t). \end{aligned}$$

Therefore,

$$u_n(t) = W_n(t) + \lambda_0(t - \theta_0) \, \hat{v}_n + r_n(t), \tag{2}$$

where we have set $\hat{v}_n = \int_{-\infty}^{+\infty} h(s - \theta_0) \, dW_n(s)$ Futhermore, we put

$$\hat{u}_n(t) = W_n(t) + \lambda_0(t - \theta_0) \, \hat{v}_n \tag{3}$$

and introduce the stochastic process

$$\hat{u}(t) = W\big(\Lambda_0(t - \theta_0)\big) + \lambda_0(t - \theta_0) \int_{-\infty}^{+\infty} h(s - \vartheta_0) \, dW\big(\Lambda_0(s - \vartheta_0)\big). \tag{4}$$

It is easy to see that if we change the variables $t - \vartheta_0 = u$ and $s - \vartheta_0 = v$ in the integrals then we obtain the following equality

$$\begin{aligned} &\int_{-\infty}^{\infty} \hat{u}(t)^2 \lambda_0(t - \vartheta_0) \, dt \\ &\quad = \int_{-\infty}^{\infty} \left[W\left(\Lambda_0(u) + \lambda_0(u) \int_{-\infty}^{\infty} h(v) \, dW(\Lambda_0(v))\right)\right]^2 \lambda_0(u) \, du = \Delta_0. \end{aligned}$$

Lemma 3.2

Let the conditions **a1**- **a4** be satisfied, then the finite dimensional distributions of the process $\hat{u}_n(t), t \in$ converge to the finite dimensional distributions of the process $\hat{u}(t), t \in$ as $n \to \infty$.

Lemma 3.3

Let the conditions **a1**- **a4** be satisfied, then for any $L > 0$, we have

$$\sup_{|t| + |s| < L} E\_\theta_0 \left|\hat{u}_n^2(t) - \hat{u}_n^2(s)\right|^2 \le C\left(1 + L^{7/2}\right) \sqrt{|t - s|}, \tag{5}$$

where $C = C(L) > 0$ does not depend on n. Lemmas (3.2) and (3.3) allow us for any $L > 0$ to establish the convergence in distribution

$$\int_{-L}^{L} \hat{u}_n^2(t) \lambda_0(t - \vartheta_0) \, dt \Longrightarrow \int_{-L}^{L} \hat{u}^2(t) \lambda_0(t - \vartheta_0) \, dt, \tag{6}$$

Lemma 3.4

Let the conditions **a1**- **a4** be fulfilled. Then for any $\varepsilon > 0$ there exist $L > 0$ and n_0 such that for all $n \ge n_0$, we have

$$\mathbb{P}_{\vartheta_0}\left(\int_{|s| > L} \hat{u}_n^2(s) \lambda_0(s - \vartheta_0) \, ds > \varepsilon\right) \le \varepsilon. \tag{7}$$

The last step is the following Lemma.

Lemma 3.5

Let the conditions **a1**- **a4** be fulfilled. Then

$$n \int {}_{\theta_0} \left[\widehat{\Lambda}_n(t) - \Lambda_0(t - \hat{\theta}_n)\right]^2 \left|\lambda_0(t - \hat{\theta}_n) - \lambda_0(t - \theta_0)\right| \, dt \xrightarrow[n \to \infty]{} 0. \tag{8}$$

THEOREM2

The test

$$\hat{\Psi}_n(X^{(n)}) = 1\!\!1_{\{\Delta_n > c_\varepsilon\}}$$

is consistent under alternative H_1, that is, for any $\Lambda \notin \mathcal{L}(\Theta)$ we have:

$$\beta\left(\hat{\Psi}_n, \Lambda\right) \xrightarrow[n \to \infty]{} 1,$$

and it is uniformly consistent under alternatives $H_1{}^\rho$, that is,

$$\inf_{\Lambda(\cdot) \in \mathcal{F}_\rho} \beta\left(\hat{\Psi}_n, \Lambda\right) \xrightarrow[n \to \infty]{} 1.$$

4. References

[1] Dabye A.S., Diakhaby A. and Diop F.N. (2011) On the Goodness-of-fit tests for Poisson processes with a discontinuous intensity. *Far East Journal of Theoretical Statistics*, **37**, 1, 23-43

[2] Dachian S. and Kutoyants, Yu. A. (2006) Hypotheses testing : Poisson versus self-exciting. *Scandinavian J. of Statistics*, **33**, 391-408.

[3] Dachian S. and Kutoyants,Yu. A. (2007) On the goodness-of-fit testing for some continuous time processes, in Statistical Models and Methods for Biomedical and Technical Systems. *F.Vonta et al. (Eds)*, Birkhäuser, Boston, 395-413.

[4] Ibragimov I.A. and Khasminskii R.Z. (1981) Statistical Estimation. *Springer*, New York.

[5] Kutoyants Yu.A. (1998) Statistical Inference for Spatial Poisson Processes. *Lecture Notes in Statistics*, **134**, Springer-Verlag, New York.

[6] Martynov, G. (2010) Note on the Cramér-von Mises test with estimated parameters. *Publ. Math. Debrecen*, **76/3**, 341-346.