Far East Journal of Mathematical Sciences (FJMS)

Published Online: September 2016
http://dx.doi.org/10.17654/MS100060911
Volume 100, Number 6, 2016, Pages 911-922 **ISSN: 0972-0871**

ESTIMATING THE VARIANCE FUNCTION OF A COMPOUND CYCLIC POISSON PROCESS

Fitriani Ida Makhmudah, I Wayan Mangku and Hadi Sumarno

Department of Mathematics
Bogor Agricultural University
Jalan Meranti, Kampus IPB Dramaga
Bogor 16680, Indonesia
e-mail: fitrianiida11@gmail.com
wayan.mangku@gmail.com
hadisumarno@yahoo.com

Abstract

An estimator of the variance function of a compound cyclic Poisson process is constructed and investigated. We do not assume any particular parametric form for the intensity function except that it is periodic. Moreover, we consider the case when there is only a single realization of the Poisson process observed in a bounded interval. The proposed estimator is proved to be weakly and strongly consistent when the size of the interval indefinitely expands. In addition, asymptotic approximations to the bias and variance of the proposed estimator are computed.

1. Introduction

Let $\{N(t),\ t \geq 0\}$ be a Poisson process with (unknown) locally integrable

Received: May 30, 2016; Accepted: June 24, 2016
2010 Mathematics Subject Classification: 62E20, 62G20, 62M20.
Keywords and phrases: compound Poisson process, cyclic intensity function, variance function, bias, variance.
Communicated by K. K. Azad

intensity function λ. We consider the case when the intensity function λ is a periodic function with (known) period $\tau > 0$. We do not assume any (parametric) form of λ except that it is periodic, that is, the equality

$$\lambda(s + k\tau) = \lambda(s) \tag{1.1}$$

holds for all $s \geq 0$ and $k \in \mathbb{N}$, where $\mathbb{N}$ denotes the set of natural numbers. This condition of intensity function is also considered in [4].

Let $\{Y(t), t \geq 0\}$ be a process with

$$Y(t) = \sum_{i=1}^{N(t)} X_i, \tag{1.2}$$

where $\{X_i, i \geq 1\}$ is a sequence of independent and identically distributed nonnegative random variables with mean μ and variance σ^2, which is also independent of the process $\{N(t), t \geq 0\}$. The process $\{Y(t), t \geq 0\}$ is said to be a *compound cyclic Poisson process*.

The model presented in (1.2) is a generalization of the (well known) compound Poisson process, which assumes that $\{N(t), t \geq 0\}$ is a homogeneous Poisson process.

There are many applications of the compound Poisson model. Some examples are as follow. Application of the compound Poisson model in physics can be seen in [2], while its application in insurance and financial problem can be found in [1]. In addition, application of the compound Poisson model in demography can be seen in [5], in geology can be seen in [7], and in biology can be found in [8].

Suppose that, for some $\omega \in \Omega$, a single realization $N(\omega)$ of the cyclic Poisson process $\{N(t), t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with intensity function λ is observed, though only within a bounded interval $[0, n]$. Furthermore, suppose that for each data point in the observed realization $N(\omega) \cap [0, n]$, say ith data point $i = 1, 2, ..., N([0, N])$, its corresponding random variable X_i is also observed. In [9], an estimator for the mean function of the process $\{Y(t), t \geq 0\}$ has been constructed and its consistency has been proved. Our goals in this paper are to construct an estimator for the variance function of the process $\{Y(t), t \geq 0\}$ using the observed realization, to prove its consistency, and to compute asymptotic approximations to the bias and variance of the proposed estimator.

The variance function of $Y(t)$, denoted by $V(t)$, is given by

$$V(t) = E[N(t)]E[X_1^2] = \Lambda(t)\mu_2$$

with $\Lambda(t) = \int_0^t \lambda(s)ds$. Let $t_r = t - \left\lfloor \frac{t}{\tau} \right\rfloor \tau$, where for any real number x, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x, and let also $k_{t,\tau} = \left\lfloor \frac{t}{\tau} \right\rfloor$. Then, for any given real number $t \geq 0$, we can write $t = k_{t,\tau}\tau + t_r$, with $0 \leq t_r < \tau$. Let $\theta = \frac{1}{\tau}\int_0^\tau \lambda(s)ds$, that is, the global intensity of the cyclic Poisson process $\{N(t), t \geq 0\}$. We assume that

$$\theta > 0. \tag{1.3}$$

Then, for any given $t \geq 0$, we have

$$\Lambda(t) = k_{t,\tau}\tau\theta + \Lambda(t_r)$$

which implies

$$V(t) = (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2.$$

2. The Estimator and Main Results

The estimator of the variance function $V(t)$ using the available data set at hand is given by

$$\hat{V}_n(t) = (k_{t,\tau}\tau\hat{\theta}_n + \hat{\Lambda}_n(t_r))\hat{\mu}_{2,n}, \tag{2.1}$$

where

$$\hat{\theta}_n = \frac{1}{k_{n,\tau}\tau} \sum_{k=0}^{k_{n,\tau}-1} N([k\tau, k\tau + \tau]),$$

$$\hat{\Lambda}_n(t_r) = \frac{1}{k_{n,\tau}} \sum\nolimits_{k=0}^{k_{n,\tau}-1} N([k\tau, k\tau + t_r])$$

and

$$\hat{\mu}_{2,n} = \frac{1}{N([0, n])} \sum_{i=1}^{N([0, n])} X_i^2,$$

with $k_{n,\tau} = \left\lfloor \frac{n}{\tau} \right\rfloor$, and the understanding that $\hat{\mu}_{2,n} = 0$, when $N([0, n]) = 0$. Thus, $\hat{V}_n(t) = 0$, when $N([0, n]) = 0$.

Our main results are presented in the following four theorems.

Theorem 1 (Weak consistency)**.** *Suppose that the intensity function* λ *satisfies* (1.1) *and is locally integrable. If, in addition,* $Y(t)$ *satisfies condition of* (1.2), *then*

$$\hat{V}_n(t) \xrightarrow{p} V(t)$$

as $n \to \infty$. *Hence,* $\hat{V}_n(t)$ *is a weakly consistent estimator of* $V(t)$.

Theorem 2 (Strong consistency)**.** *Suppose that the intensity function* λ *satisfies* (1.1) *and is locally integrable. If, in addition,* $Y(t)$ *satisfies condition of* (1.2), *then*

$$\hat{V}_n(t) \xrightarrow{a.s} V(t)$$

as $n \to \infty$. *Hence,* $\hat{V}_n(t)$ *is a strongly consistent estimator of* $V(t)$.

Theorem 3 (Asymptotic approximation to the bias)**.** *Suppose that the intensity function* λ *satisfies* (1.1) *and is locally integrable. If, in addition,* $Y(t)$ *satisfies condition of* (1.2), *then*

$$Bias[\hat{V}_n(t)] = -\frac{V(t)}{e^{n\theta}} + o(e^{-n})$$

as $n \to \infty$.

Theorem 4 (Asymptotic approximation to the variance)**.** *Suppose that the intensity function* λ *satisfies* (1.1) *and is locally integrable*. *If*, *in addition*, $Y(t)$ *satisfies condition of* (1.2), *then*

$$Var[\hat{V}_n(t)] = \frac{1}{n}\left(\frac{\sigma_2^2 V(t)}{\theta\mu_2} + 2\mu_2^2 t\Lambda(t_r) + \mu_2^2\tau\Lambda(t_r) + \mu_2^2 t^2\theta\right) + \mathcal{O}\left(\frac{1}{n^2}\right) \quad (2.2)$$

as $n \to \infty$.

We note that, since the $bias^2(\hat{V}_n(t))$ is of smaller order than $\mathcal{O}(n^{-2})$, as $n \to \infty$, we also have that asymptotic approximation to the *Mean Squared Error* of our estimator is given by the r.h.s. of (2.2).

3. Some Technical Lemmas

In this section, we present some results which are needed in the proofs of our theorems.

Throughout this paper, for any random variables X_n and X on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, we write $X_n \xrightarrow{c} X$ to denote that X_n converges completely to X as $n \to \infty$. We say that X_n converges completely to X if, for every $\varepsilon > 0$, $\sum_{n=1}^{\infty} P(| X_n - X | > \varepsilon) < \infty$.

Lemma 1. *Suppose that the intensity function* λ *satisfied* (1.1) *and is locally integrable*. *Then*

$$\hat{\theta}_n \xrightarrow{p} \theta, \quad (3.1)$$

$$\hat{\theta}_n \xrightarrow{c} \theta, \quad (3.2)$$

$$\hat{\Lambda}_n(t_r) \xrightarrow{p} \Lambda(t_r), \quad (3.3)$$

$$\hat{\Lambda}_n(t_r) \xrightarrow{c} \Lambda(t_r), \tag{3.4}$$

as $n \to \infty$.

Proof. Proofs of (3.1) and (3.2) are similar to the proofs of Lemma 2 and Lemma 3 of [9], while proofs of (3.3) and (3.4) are similar to the proofs of Lemma 4 and Lemma 5 of [9]. Hence, these are omitted.

Lemma 2. *Suppose that the intensity function* λ *satisfied* (1.1) *and is locally integrable*. *If, in addition*, (1.3) *holds, then with probability* 1,

$$N([0, N]) \to \infty \tag{3.5}$$

as $n \to \infty$.

Proof. We refer to [9].

4. Proofs of Theorems 1, 2, 3 and 4

Proof of Theorem 1. By (2.1), (3.1) and (3.3), to prove Theorem 1, it remains to check

$$\hat{\mu}_{2,n} \xrightarrow{p} \mu_2 \tag{4.1}$$

as $n \to \infty$. Since $\{X_i, i \geq 1\}$ is a sequence of independent and identically distributed nonnegative random variables, we also have that $\{X_i^2, i \geq 1\}$ is a sequence of independent and identically distributed random variables. By Lemma 2 and the weak law of large numbers, we have (4.1). This completes the proof of Theorem 1.

Proof of Theorem 2. By (2.1), to prove Theorem 2, it suffices to check

$$\hat{\theta}_n \xrightarrow{a.s} \theta, \tag{4.2}$$

$$\hat{\Lambda}_n(t_r) \xrightarrow{a.s} \Lambda(t_r) \tag{4.3}$$

and

$$\hat{\mu}_{2,n} \xrightarrow{a.s} \mu_2 \tag{4.4}$$

as $n \to \infty$ (cf. [3]). By (3.2) of Lemma 1 and the Borell-Cantelli Lemma, we have (4.2). Similarly, (3.4) of Lemma 1 and the Borell-Cantelli Lemma also lead to (4.3). By Lemma 2 and the strong law of large numbers, we obtain (4.4). This completes the proof of Theorem 2.

Proof of Theorem 3. Expected value of $\hat{V}_n(t)$ can be computed as follows:

$$E[\hat{V}_n(t)] = E[E[\hat{V}_n(t) \mid N([0, n])]]]$$

$$= \sum_{m=0}^{\infty} E[\hat{V}_n(t) \mid N([0, n]) = m] P(N([0, n]) = m)$$

$$= \sum_{m=1}^{\infty} E[\hat{V}_n(t) \mid N([0, n]) = m] P(N([0, n]) = m),$$

due to $\hat{V}_n(t) = 0,$ when $N([0, n]) = 0.$

Now we note that, by Lemma 1 and Lemma 2 of [9], we have $E[\hat{\theta}_n] = \theta$ and $E[\hat{\Lambda}_n(t_r)] = \Lambda(t_r).$ A simple argument also shows that

$$E[\hat{\mu}_{2,n} \mid N([0, n]) = m] = \mu_2.$$

By (2.1) and nothing that $\hat{\mu}_{2,n}$ is independent of $\hat{\theta}_n$ and $\hat{\Lambda}_n(t_r),$ we have

$$E[\hat{V}_n(t)] = \sum_{m=1}^{\infty} (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2 P(N([0, n]) = m)$$

$$= (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2 \sum_{m=1}^{\infty} P(N([0, n]) = m)$$

$$= (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2(1 - P(N([0, n]) = 0))$$

$$= (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2(1 - e^{-\Lambda(n)}).$$

Since $\Lambda(n) = \int_0^n \lambda(s)ds = n\theta + \mathcal{O}(1),$ as $n \to \infty,$ we obtain

$$E[\hat{V}_n(t)] = (k_{t,\tau}\tau\theta + \Lambda(t_r))\mu_2(1 - e^{-n\theta+\mathcal{O}(1)})$$

$$= V(t)(1 - e^{-n\theta+\mathcal{O}(1)})$$

$$= V(t) - \frac{V(t)}{e^{n\theta}} + o(e^{-n})$$

as $n \to \infty,$ and Theorem 3 follows.

Proof of Theorem 4. The variance of $\hat{V}_n(t)$ can be computed as follows:

$$Var[\hat{V}_n(t)] = E[(\hat{V}_n(t))^2] - (E[\hat{V}_n(t)])^2. \tag{4.5}$$

The first term on the r.h.s. of (4.5) can be computed as follows:

$$E[(\hat{V}_n(t))^2] = E[E[(\hat{V}_n(t))^2 \mid N([0, N])]]$$

$$= \sum_{m=0}^{\infty} E[(\hat{V}_n(t))^2 \mid N([0, N]) = m]P(N([0, N]) = m)$$

$$= \sum_{m=1}^{\infty} E[(\hat{V}_n(t))^2 \mid N([0, N]) = m]P(N([0, N]) = m),$$

due to $\hat{V}_n(t) = 0,$ when $N([0, N]) = 0.$

First, we compute

$$E[(\hat{V}_n(t))^2 \mid N([0, N]) = m]$$

$$= E\left[\left((k_{t,\tau}\tau\hat{\theta}_n + \hat{\Lambda}_n(t_r))\frac{1}{m}\sum_{i=1}^{m} X_i^2\right)^2\right]$$

$$= E[((k_{t,\tau}\tau\hat{\theta}_n + \hat{\Lambda}_n(t_r)))^2]E\left[\left(\frac{1}{m}\sum_{i=1}^{m} X_i^2\right)^2\right]. \tag{4.6}$$

The first expected value on the r.h.s. of (4.6) can be computed as follows:

$$k_{t,\tau}^2\tau^2 E[\hat{\theta}_n^2] + 2k_{t,\tau}\tau E[\hat{\theta}_n\hat{\Lambda}_n(t_r)] + E[(\hat{\Lambda}_n(t_r))^2]. \tag{4.7}$$

Now we note that, by Lemma 1 and Lemma 2 of [6], we have $E[\hat{\theta}_n^2] = \frac{\theta}{k_{n,\tau}\tau} + \theta^2$ and $E[(\hat{\Lambda}_n(t_r))^2] = \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2$. By (4.8) of [9], we have $E[\hat{\theta}_n\hat{\Lambda}_n(t_r)] = \Lambda(t_r)\theta + \frac{\Lambda(t_r)}{k_{n,\tau}\tau}$. Then, the quantity in (4.7) can be written as

$$k_{t,\tau}^2\tau^2\left(\frac{\theta}{k_{n,\tau}\tau} + \theta^2\right) + 2k_{t,\tau}\tau\left(\Lambda(t_r)\theta + \frac{\Lambda(t_r)}{k_{n,\tau}\tau}\right) + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2$$

$$= \frac{k_{t,\tau}^2\tau\theta}{k_{n,\tau}} + k_{t,\tau}^2\tau^2\theta^2 + 2k_{t,\tau}\tau\theta\Lambda(t_r) + \frac{2k_{t,\tau}\Lambda(t_r)}{k_{n,\tau}} + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2.$$

Next, we compute

$$E\left[\left(\frac{1}{m}\sum_{i=1}^{m} X_i^2\right)^2\right] = \frac{1}{m^2}\left[\sum_{i=1}^{m} E[X_i^4] + \sum_{i=1}^{m}\sum_{\substack{j=1 \\ j\neq i}}^{m} E[X_i^2]E[X_j^2]\right]$$

$$= \frac{1}{m^2}(mE[X_1^4] + (m^2 - m)(E[X_1^2])^2)$$

$$= \mu_2^2 + \frac{1}{m}\sigma_2^2,$$

where $\sigma_2^2 = Var(X_1^2)$. Then, we have

$$E[(\hat{V}_n(t))^2 \mid N([0, N]) = m]$$

$$= \left(\frac{k_{t,\tau}^2 \tau\theta}{k_{n,\tau}} + k_{t,\tau}^2 \tau^2 \theta^2 + 2k_{t,\tau}\tau\theta\Lambda(t_r) + \frac{2k_{t,\tau}\Lambda(t_r)}{k_{n,\tau}} + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2\right)$$

$$\times \left(\mu_2^2 + \frac{1}{m}\sigma_2^2\right).$$

Therefore,

$$E[(\hat{V}_n(t))^2]$$

$$= E[E[(\hat{V}_n(t))^2 \mid N([0, N])]]$$

$$= \sum_{m=1}^{\infty}\left(\frac{k_{t,\tau}^2 \tau\theta}{k_{n,\tau}} + k_{t,\tau}^2 \tau^2 \theta^2 + 2k_{t,\tau}\tau\theta\Lambda(t_r) + \frac{2k_{t,\tau}\Lambda(t_r)}{k_{n,\tau}} + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2\right)$$

$$\times \left(\mu_2^2 + \frac{1}{m}\sigma_2^2\right) P(N([0, N]) = m)$$

$$= \left(\frac{k_{t,\tau}^2 \tau\theta}{k_{n,\tau}} + k_{t,\tau}^2 \tau^2 \theta^2 + 2k_{t,\tau}\tau\theta\Lambda(t_r) + \frac{2k_{t,\tau}\Lambda(t_r)}{k_{n,\tau}} + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2\right)$$

$$\times \left[\mu_2^2 \sum_{m=1}^{\infty} P(N([0, N]) = m) + \sigma_2^2 \sum_{m=1}^{\infty} \frac{1}{m} P(N([0, N]) = m)\right].$$

A simple calculation shows that

$$\sum_{m=1}^{\infty} P(N([0, N]) = m) = 1 + \mathcal{O}(e^{-n})$$

and

$$\sum_{m=1}^{\infty} \frac{1}{m} P(N([0, N]) = m) = \frac{1}{n\theta} + \mathcal{O}\left(\frac{1}{n^2}\right), \tag{4.8}$$

as $n \to \infty$. Proof (4.8) can be seen in [9].

Thus, we obtain

$$E[(\hat{V}_n(t))^2]$$

$$= \left(\frac{k_{t,\tau}^2 \tau\theta}{k_{n,\tau}} + k_{t,\tau}^2 \tau^2\theta^2 + 2k_{t,\tau}\tau\theta\Lambda(t_r) + \frac{2k_{t,\tau}\Lambda(t_r)}{k_{n,\tau}} + \frac{\Lambda(t_r)}{k_{n,\tau}} + (\Lambda(t_r))^2\right)$$

$$\times\left[\mu_2^2(1 + \mathcal{O}(e^{-n})) + \sigma_2^2\left(\frac{1}{n\theta} + \mathcal{O}\left(\frac{1}{n^2}\right)\right)\right].$$

Simplifying this quantity, and after some algebra, we obtain

$$E[(\hat{V}_n(t))^2]$$

$$= (V(t))^2 + \frac{1}{n}\left(\frac{\sigma_2^2 V(t)}{\theta\mu_2} + 2\mu_2^2 t\Lambda(t_r) + \mu_2^2\tau\Lambda(t_r) + \mu_2^2 t^2\theta\right) + \mathcal{O}\left(\frac{1}{n^2}\right)$$

as $n \to \infty$. Combining this result and Theorem 3, we obtain Theorem 4.

References

[1] V. E. Bening and V. Y. Korolev, Generalized Poisson Models and their Application in Insurance and Finance, VSP International Science Publishers, Utrecht, Boston, 2002.

[2] J. Byrne, Properties of compound Poisson processes with application in statistical physics, Physica 41 (1969), 575-587.

[3] A. Dasgupta, Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics, Springer, New York, 2011.

[4] R. Helmers, I. W. Mangku and R. Zitikis, Statistical properties of a kernel-type estimator of the intensity function of a cyclic Poisson process, J. Multivariate Anal. 84(1) (2003), 19-39.

[5] S. R. Kegler, Applying the compound Poisson process model to the reporting of injury-related mortality rates, Epidemiologic Perspective and Innovations 4 (2007), 1-9.

[6] I W. Mangku, Ruhiyat and I. G. P. Purnaba, Statistical properties of an estimator for the mean function of a compound cyclic Poisson process, Far East J. Math. Sci. (FJMS) 82(2) (2013), 227-237.

[7] G. Ozel and C. Inal, The probability function of the compound Poisson process and an application to aftershock sequence in Turkey, Environmentrics 19 (2008), 79-85.

[8] P. Puig and J. F. Barquinero, An Application of compound Poisson modeling to biological dosimetry, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 467(2127) (2011), 897-910.

[9] Ruhiyat, I W. Mangku and I. G. P. Purnaba, Consistent estimation of the mean function of a compound cyclic Poisson process, Far East J. Math. Sci. (FJMS) 77(2) (2013), 183-194.