
Nurul Indah Safitri*, I Wayan Mangku and Hadi Sumarno
Department of Mathematics, IPB University, Bogor 16680, Indonesia
Email: pasca_nurul@apps.ipb.ac.id

Abstract
This article discusses the estimation for the expected value, also called the mean function, of a compound periodic Poisson process with a power function trend. The aims of our study are, first, to modify the existing estimator to produce a new estimator that is normally distributed, and, second, to determine the smallest observation interval size such that our proposed estimator is still normally distributed. Basically, we formulate the estimator using the moment method. We use Monte Carlo simulation to check the distribution of our new estimator. The result shows that a new estimator for the expected value of a compound periodic Poisson process with a power function trend is normally distributed and the simulation result shows that the distribution of the new estimator is already normally distributed at the length of 100 observation interval for a period of 1 unit. This interval is the smallest size of the observation interval. The Anderson-Darling test shows that when the period is getting larger, the p-value is also getting bigger. Therefore, the larger period requires a wider observation interval to ensure that the estimator still has a normal distribution.

Keywords: moment method; normal distribution; Poisson process; the smallest observation interval.

1. INTRODUCTION

We consider the problem of estimating the expected value, also called the mean function, of a compound Poisson process when the intensity function of the Poisson component is a periodic or cyclic function plus a power function trend. Since the periodic compound Poisson process has wide applications in applied sciences, it is needed to know the distribution of this process. However, it is
A Study on the Estimator Distribution for the Expected Value of a Compound Periodic Poisson Process


Erliana et al. [11], Maulidi, et al. [12] and Mulidi et al. [13] discussed the power function trend of a non-homogeneous Poisson process. Mangku [14], Maulidi et al. [15], and Abdullah et al. [16] estimated the intensity of a cyclic Poisson process with the power function of a non-homogeneous Poisson process. The confidence interval for the expected value function had been estimated in [17] and [18]. The Poisson process had been applied to ore flow in a belt conveyor system [19], in the probability distribution of COVID-19 [20], and in pricing catastrophe bonds [21].

Formulation of consistent estimation of the mean (expected value) function of purely cyclic compound Poisson process was first done in [1]. In [2], this estimator was slightly modified in order to obtain asymptotic bias and asymptotic variance of the estimator. In [4], a consistent estimator for the mean (expected value) function of a compound periodic Poisson process with power function trend was formulated. The asymptotic bias and asymptotic variance of this estimator also have been formulated in [4], [9], and [10].

This research is a continuation of Sari et al. [4]. Sari et al. [4] do not assume any parametric form for the estimator for the mean function whereas in this research the estimator has a distribution. We will formulate a new estimation for the expected value of a compound periodic Poisson process with power function trend with the objectives as follows. (i) To modify the estimator formulated in [4] to obtain a new estimator which is normally distributed. (ii) To find the smallest size of the observation interval of the Poisson process such that the estimator is still normally distributed.

2. METHODS

Figure 1 shows our research framework. We formulate the new estimator using the moment method and use Monte Carlo simulation to check the distribution of the new estimator.

\[
\hat{\psi}_{n,b}(t) = \left( (k_{t+1} + 1)\tilde{\Lambda}_{c,n,b}(t_r) + k_{t,r}\tilde{\Lambda}_{c,n,b}(t_r) + \frac{a_{m,b}}{b+1} t^{b+1} \right) \hat{\mu}_n
\]

Figure 1. The research framework.
The steps in this research include: (1) preliminary study and (2) the main study. The preliminary study comprises studying the compound periodic Poisson process with power function trend and exploring mathematical foundations to develop a new estimator from the previous theories. The main study involves modifying the existing estimator and investigating the normality of the new estimator; as well as obtaining the smallest observation interval size of the process where the new estimator is still normally distributed.

3. RESULTS AND DISCUSSION

In this section, formulation for the expected value (mean) function of a compound cyclic Poisson process with power function trend is discussed.

3.1 Formulation of the Mean Function

Let \( \{N(t), t \geq 0\} \) be a nonhomogeneous Poisson process having (unknown) intensity function \( \lambda \). It is assumed that intensity function \( \lambda \) has two components, namely a cyclic (periodic) component, \( \lambda_c \), with (known) period \( \tau > 0 \) and a power function trend component. In other words, for all \( s \geq 0 \), the intensity function \( \lambda \) can be written as ([4])

\[
\lambda(s) = \lambda_c(s) + as^b,
\]

with \( \lambda_c \) is the cyclic component with period \( \tau \) and \( as^b \) is the power function trend, with \( a > 0 \) is the slope of the power function trend. It is assumed that \( b \) is a known real number and \( 0 < b < \frac{1}{2} \). It is not assumed any parametric form for the cyclic component \( \lambda_c \), except that it is cyclic or periodic, which satisfies

\[
\lambda_c(s) = \lambda_c(s + k\tau)
\]

for all \( s \geq 0 \) and \( k \in \mathbb{N} \), where \( \mathbb{N} \) denotes the set of natural numbers.

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

\[
Y(t) = \sum_{i=1}^{N(t)} X_i
\]

with \( \{X_i, i \geq 1\} \) is a sequence of i.i.d. (independent and identically distributed) random variables having mean \( \mu < \infty \) and variance \( \sigma^2 < \infty \), and also independent of the process \( \{N(t), t \geq 0\} \). The Process \( \{Y(t), t \geq 0\} \) is called a compound cyclic Poisson process with power function trend [4].

Let for some \( \omega \in \Omega \), a single realization \( N(\omega) \) of the process \( \{N(t), t \geq 0\} \) is observed on an interval \( [0, m] \), \( m \) is a real number. Furthermore, suppose that for each data point in the realization \( N(\omega) \cap [0, m] \), say \( i \)-th data point, \( i = 1, 2, \ldots, N[0,m] \), its corresponding random variable \( X_i \) is also observed.
A Study on the Estimator Distribution for the Expected Value of a Compound Periodic Poisson Process

Let \( \psi (t) \) denotes the expected value (mean) function of the considered process. \( \psi (t) \) can be written as

\[
\psi (t) = E[Y(t)] = E[N(t)]E[X] = \Lambda (t) \mu ,
\]

with \( \mu = E(X) \) and

\[
\Lambda (t) = \int_0^t \lambda (s) ds.
\]

Let \( t = t - \left[ \frac{t}{\tau} \right] \tau \) where \( \left[ x \right] \) denotes the biggest integer that is less than or equal to \( x \in \mathbb{N} \), and \( k_{t,\tau} = \left[ \frac{t}{\tau} \right] \). Then for any real number \( t \geq 0 \), we can write \( t = k_{t,\tau} \tau + t_r \) with \( 0 \leq t_r < \tau \). Let \( \tau \theta = \int_0^t \lambda (s) ds \), which can be written as \( \Lambda_r (t_r) + \Lambda'_r (t_r) \) with \( \Lambda_r (t_r) = \int_0^{t_r} \lambda (s) ds \) and \( \Lambda'_r (t_r) = \int_{t_r}^t \lambda (s) ds \). It is assumed that \( \theta > 0 \). By using the above notations and substituting (1) into (5), for any \( t \geq 0 \), \( \Lambda (t) \) can be formulated as

\[
\Lambda (t) = (k_{t,\tau} + 1) \Lambda_r (t_r) + k_{t,\tau} \Lambda'_r (t_r) + \frac{a}{b+1} t_r^{b+1}.
\]

By substituting (6) into the r.h.s. of (4), we have

\[
\psi (t) = \left( (k_{t,\tau} + 1) \Lambda_r (t_r) + k_{t,\tau} \Lambda'_r (t_r) + \frac{a}{b+1} t_r^{b+1} \right) \mu .
\]

3.2 Formulation of the Estimator for the Mean Function

After formulating the mean function, we formulate the estimator for the expected value (mean) function of the process as follows. Let \( n^{1+b} = m \), where \( \delta \) be an arbitrarily small positive real number. For some reason, to estimate \( a \) we use the realization of the Poisson process observed in \( [0,m] \), while to estimate the other quantities we use the realization of the Poisson process observed in \( [0,n] \). The estimator of the mean function can be written as

\[
\hat{\psi}_{n,b} (t) = \left( (k_{t,\tau} + 1) \hat{\Lambda}_{c,n,b} (t_r) + k_{t,\tau} \hat{\Lambda}'_{c,n,b} (t_r) + \frac{\hat{\mu}_m}{b+1} t_r^{b+1} \right) \hat{\mu}_n ,
\]

where

\[
\hat{\Lambda}_{c,n,b} (t_r) = \frac{(1-b) \tau^{1-b}}{n^{1-b}} \sum_{k=1}^k \frac{1}{k} N([k \tau, k \tau + t_r]) - \hat{\mu}_m (1-b) n^b t_r ,
\]

\[
\hat{\Lambda}'_{c,n,b} (t_r) = \frac{(1-b) \tau^{1-b}}{n^{1-b}} \sum_{k=1}^k \frac{1}{k} N([k \tau + t_r, k \tau + \tau]) - \hat{\mu}_m (1-b) n^b (\tau - t_r) ,
\]

\[
\hat{\mu}_n = \frac{1}{m} \sum_{j=0}^{m-1} Y_j ,
\]

\[
\hat{\mu}_m = \frac{1}{m} \sum_{j=0}^{m-1} Y_j .
\]
\[
\hat{\theta}_{m,b} = \frac{(1+b)N([0,m])}{m^{1+b}} - \frac{(1+b)}{m^b} \tilde{\theta}_n,
\]
\[
\tilde{\theta}_n = \frac{(1-b)}{n^{1+b}b^2} \sum_{k=1}^{N([0,n])} 1 - \frac{(1-b)(1+b)n^b N([0,n])}{n^{1+b}b^2}, \quad \text{and}
\]
\[
\hat{\mu}_n = \frac{1}{N[0,n]} \sum_{i=1}^{N[0,n]} X_i, \quad \text{if } N([0,n]) > 0 \quad \text{and} \quad \hat{\mu}_n = 0 \quad \text{if } N([0,n]) = 0.
\]

The estimator for the global intensity function \( \theta \) is a modification of the estimator formulated in [4]. In this paper, the normality of the estimator given in (8) was investigated using Monte Carlo simulation. We also determine the smallest size of the observation interval such that the distribution of the estimator is already normal.

### 3.3 The Monte Carlo Simulation

The simulation was carried out with the help of R software. The simulation is aimed at checking normality distribution and determining the smallest \( n \) value where the distribution of the estimator is close to the normal distribution. The programming stage begins with generating the realization of a compound periodic Poisson process with a power function trend with the formulation of the intensity function as follows.

\[
\lambda(t) = \sin\left(\frac{2\pi t}{\tau}\right) + 1 + 0.07t^{2/5}.
\]

In this study, the normality test was carried out using a QQ-plot graph analysis approach. The data is normally distributed if the spread of each point approaches the diagonal line, hence the assumption of normality is met. The simulation test results can be seen in Figures 2-5.

![Normal Q-Q Plot](image1)

![Normal Q-Q Plot](image2)

**Figure 2. Cont.**
A Study on the Estimator Distribution for the Expected Value of a Compound Periodic Poisson Process

Figure 2. Normality of $\hat{\Psi}_{a,b}(t)$ for $\tau = 1$. (a) Observation interval at [0, 25]; (b) Observation interval at [0, 75]; (c) Observation interval at [0, 100].

Figure 3. Normality of $\hat{\Psi}_{a,b}(t)$ for $\tau = 5$. (a) Observation interval at [0, 100]; (b) Observation interval at [0, 200].

Figure 4. Normality of $\hat{\Psi}_{a,b}(t)$ for $\tau = 10$. (a) Observation interval at [0, 300]; (b) Observation interval at [0, 500].
Figure 5. Normality of $\hat{\Psi}_{\alpha, \beta}(t)$ for $\tau = 30$. (a) Observation interval at $[0, 900]$; (b) Observation interval at $[0, 1200]$.

Based on the simulation results, in the observation intervals $[0, 25]$, $\tau = 1$ (Figure 2(a)) and $[0, 75]$, $\tau = 1$ (Figure 2(b)), the spread of each estimator point is still visible away from the diagonal line. Meanwhile, in the observation interval $[0, 100]$, $\tau = 1$ (Figure 2(c)), the spread of each estimator point already approaches the diagonal line. This illustration shows that the estimator for the mean function of a compound periodic Poisson process with a power function trend is getting closer to the normal distribution for a longer interval of observation. Furthermore, the Anderson-Darling test was carried out to check the normality of the estimator based on the $p$-value. The $p$-value of the normality simulation results of the estimator for the mean function is shown in Table 1.

Table 1 shows that the $p$-value when the period of 1 unit for the observation intervals $[0, 25]$ and $[0, 75]$ is less than 0.05. It means that the estimator of the mean function is not normally distributed. Meanwhile, for the same period of 1 unit with an observation interval at $[0, 100]$, the $p$-value is more than 0.05. That is, the estimator of the mean function begins to spread normally. So, $[0, 100]$ is the smallest length of the observation interval where the distribution of the estimator for the mean function is normally distributed. Furthermore, note that the $p$-value at the length of the observation interval $[0, 100]$ when the period of 1 unit is greater than 5. This is because when the period is getting bigger, the error generated is also getting bigger. When the period is getting longer, then the $p$-value is also getting bigger. This means that a longer period requires a wider observation interval to ensure that the estimator is still normally distributed.

Table 1. The simulation of normality for the estimator of mean function

<table>
<thead>
<tr>
<th>Period ($\tau$)</th>
<th>Observation Interval $[0, n]$</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$[0,25]$</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td>$[0,75]$</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td>$[0,100]$</td>
<td>0.1231</td>
</tr>
<tr>
<td>5</td>
<td>$[0,100]$</td>
<td>0.0617</td>
</tr>
<tr>
<td></td>
<td>$[0,200]$</td>
<td>0.0955</td>
</tr>
<tr>
<td>10</td>
<td>$[0,300]$</td>
<td>0.2994</td>
</tr>
<tr>
<td></td>
<td>$[0,500]$</td>
<td>0.3564</td>
</tr>
<tr>
<td>30</td>
<td>$[0,900]$</td>
<td>0.4582</td>
</tr>
<tr>
<td></td>
<td>$[0,1200]$</td>
<td>0.9608</td>
</tr>
</tbody>
</table>
4. CONCLUSIONS

Our study proposes a new estimator for the expected value (mean) function of a compound periodic Poisson process with a power function trend that has a normal distribution. The simulation result shows that our new estimator is already normally distributed at the length of 100 observation interval for a period of 1 unit. This is the smallest size of observation interval. The Anderson-Darling test shows that as the period gets larger, the p-value also gets larger. It means that, the larger the period, it requires wider observation interval to ensure that the estimator is still normally distributed.

REFERENCES


