Algorithm for Generating Compound Poisson Process Which has Nonhomogeneous Poisson Process and Exponential Distribution Components

Abstract—Compound Poisson process (CPP) is one of the models of stochastic processes in which this model can model a real phenomenon that has an element of uncertainty in the process. CPP has two main components, those is a Poisson process on the component of the poses of an event that occurs and a sequence component of the magnitude as a result of the process of events that occur. This research aims to develop an algorithm to generate random numbers from CPP with a component in the Poisson process in the form of a nonhomogeneous Poisson process (NHPP) and a component in the magnitude of the effect in the form of an exponential distribution (CPP-NHPP-ED). The method used is using the acceptance and rejection method in the form of Thinning process techniques. The results of the study obtained several algorithms, namely the algorithm for CPP-HPP-ED, CPP-NHPP-ED type 1, and CPP-NHPP-ED type 2. These algorithms can be used for computer simulation analysis that can be applied to various fields of science and engineering.

Keywords: algorithm, Compound Poisson Process (CPP), exponential distribution, Nonhomogeneous Poisson Process (NHPP), Thinning process

I. INTRODUCTION

In real phenomena, there are a lot of events that contain random events. Not necessarily events that have occurred in the past will recur in the present and future. But it could be that the event will recur, such as cyclic or periodic events. So that from these real phenomena a model is created that can resemble the actual model. The appropriate model is expected to be able to predict future events in terms of both profits and losses. In probability statistics, if a real phenomenon contains an element of uncertainty and contains rules of opportunity, then the event is classified as a stochastic process event.

In terms of time, stochastic processes can be divided into discrete stochastic processes and continuous stochastic processes. One continuous stochastic process that is often used to count the number of events that occur, both success or failure at a certain time or a certain time interval, the Poisson process. According to its intensity function, this process is divided into a homogeneous Poisson process (HPP) and a nonhomogeneous Poisson process (NHPP). HPP has an intensity function that is assumed to have a constant time-intensity function (not time-dependent), whereas NHPP has an intensity function that is not constant (time-dependent). One example of HPP and NHPP in engineering can be seen in [1], [2], and [3].

Stochastic process modeling that not only counts the number of events that occur at certain time intervals but also calculates the effects or effects of events that occur is called a compound process. A compound process consists of at least 2 components, namely the number of events (frequency) and the magnitude component (severity). In this case, the frequency distribution component in question is the distribution of the number of events in a given period and the distribution of severity is a large distribution of consequences caused when the event occurs, can be in the form of money or other amounts. If it is assumed that the frequency component is a Poisson process, then the compound process is called the compound Poisson process (CPP). Whereas the severity component can be assumed to be in the form of the exponential distribution, gamma, Pareto, and others which correspond to the magnitude of the consequences of the event.

The CPP model with the assumption on the frequency component has an HPP component, where the HPP is
assumed to have a constant intensity function, has many A.
applications in various fields, for example in the fields of
biology, seismography, demographics, insurance and finance
(see reference in [4]), and engineering ([5] and [6]). However,
if in an event that happens to have an event that increase or
decrease at a certain period, which has an intensity function
that is no longer constant (time-dependent), then the model is
not appropriate. So it is necessary to develop a new model
where the frequency component in CPP is assumed to have
the form of NHPP. Furthermore, if in an event that has an
event that tends to be repeated so that it has cyclic or periodic
properties, then the model must be changed to the form of
NHPP with an intensity function in the form of periodic
functions ([7], [8] and [9]). Because this intensity function
has a repetitive function in the next period, this model is good
for describing phenomena that occur periodically. Furthermore,
if the model has properties that not only have periodic properties but also have periodic properties that have
linear trends, then the model must be changed to the form of
NHPP with intensity functions in the form of periodic functions with linear trends ([10] and [4]).

In formulating a stochastic model to describe real
phenomena, people used to agree to choose between a model
that is a realistic replica of an actual situation and a model for
which mathematical analysis is easy to do. However, the
relatively new progress of rapid computing opens up a new
approach, namely by trying to make phenomena as precisely
as possible and then relying on simulation studies to analyze
them. In particular, the computer can be used to produce
random numbers (pseudo-random) and then these random
numbers can be used to produce random variable values from
arbitrary distributions. So from the right random variable can
produce the behavior of stochastic models at the right time.

The algorithm for conducting simulations on CPP models
with frequency components in the form of NHPP and
assumptions on the severity component in the form of
exponential distribution has not been made. So in this study
aims to create algorithms to generate a random number in the
compound Poisson process (CPP) model which has a non-
homogeneous Poisson process (NHPP) and exponential
distribution components. The results of this algorithm can be
used in the analysis of models in computational simulations,
for example, for the analysis of reliability, total loss
(aggregate loss) and risk size (VaR) and expected shortfall
(ES) in risk theory. So the company can provide protection or
protection against the possibility of future events that can
cause financial losses.

II. COMPONENTS OF COMPOUND POISSON PROCESS (CPP)

This section discusses the theory of compound Poisson
processes (CPP) that have homogeneous Poisson processes
(HPP) and exponential distribution components (CPP-HPP-
ED), and compound Poisson processes (CPP) that have
nonhomogeneous Poisson processes (NHPP) and exponential
distribution components (CPP-NHPP-ED).

The counting process \{N(t), t ≥ 0\} represents the number
of events that occurred at time \(T\). This process is said to be
a homogeneous Poisson process (HPP) if it has a Poisson
distribution and has an intensity function \(\lambda > 0\). Because it
has a constant intensity function, it is said that this process is
not time-dependent. Formal definitions and theorems of HPP
can be seen in Ross [11]. The expectation and variance of
HPP respectively as follows.

\[
E[N(t)] = \lambda t \quad \text{and} \quad \text{var}[N(t)] = \lambda t.
\]  

(1)

Poisson compound process has two basic components, namely the component number of events (frequency) which
is assumed to have the Poisson process and the component of
the magnitude of the result (severity). The definition of a
Poisson compound process is as follows.

Definition 2.1. A stochastic process \{Y(t), t ≥ 0\} is said to be
a CPP if it can be represented as

\[
Y(t) = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0,
\]

(2)

where \{N(t), t ≥ 0\} is a Poisson process, and \{X_i, i ≥ 1\}
is a family of independent and identically distributed (i.i.d)
random variables that is also independent of \{N(t), t ≥ 0\}

The compound Poisson process with its frequency component
in the form of the Poisson process with a constant
intensity function has an expectation value (denoted by
\(\psi_1(t)\) ) and variants (denoted by \(V_1(t)\) ) respectively as
follows Ross [11]:

\[
\psi_1(t) = E[Y(t)] = E[N(t)]E[X_1]
\]

(3)

\[
V_1(t) = \text{var}[Y(t)] = E[N(t)]E[X_1^2].
\]

(4)

Assumed that \(X_i\) has an exponent distribution with the
parameter \(\theta\). Then the expectation, variance and second-
order of \(X_i\) are as follows.

\[
E[X_1] = \frac{1}{\theta}, \quad \text{var}[X_1] = \frac{1}{\theta^2} \quad \text{and} \quad E[X_1^2] = \frac{1}{\theta^2}.
\]

(5)

By substituting equations (1), and (5) to equations (3) and
(4), we get, respectively,

\[
\psi_1(t) = \frac{\lambda t}{\theta}, \quad \text{and}
\]

(6)
\[
V_1(t) = \frac{\lambda t}{\theta^2},
\]

(7)

B. Compound Poisson process having nonhomogeneous Poisson process (NHPP) and exponential distribution components (CPP-NHPP-ED).

The counting process \( \{N^*(t), t \geq 0\} \) is said to be a nonhomogeneous Poisson process (NHPP) if it has a Poisson distribution and has a non-constant intensity function \( \lambda(t) > 0 \). Because it has a non-constant intensity function, it is said to be a time-dependent process. Formal definitions and theorems of NHPP can be seen in Ross [11]. The expectation and variance of NHPP respectively as follows:

\[
E[N^*(t)] = \Lambda(t) = \text{var}[N^*(t)]
\]

(8)

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

Compound Poisson process with its frequency component in the form of Poisson process with the function of the non-constant intensity (denoted by \( \psi^\prime (t) \)) has an expected value (denoted by \( \psi^\prime (t) \)) and variants (denoted by \( V^\prime (t) \)) respectively as follows ([12] and [13]):

\[
\psi^\prime (t) = E[Y^*(t)] = E[N^*(t)]E[X_1]
\]

(9)

\[
V^\prime (t) = \text{var}[Y^*(t)] = E[N^*(t)]E[X_1^2]
\]

(10)

By substituting equations (8) and (5) to equations (9) and (10), we get, respectively,

\[
\psi^\prime (t) = \frac{\Lambda(t)}{\theta}, \quad \text{and}
\]

(11)

\[
V^\prime (t) = \frac{\Lambda(t)}{\theta^2}.
\]

(12)

It should be noted because this process depends on time, the mean and its variance are functions of time as well.

The algorithm for generating random numbers contained in [14] and [15] is only limited to HPP and NHPP. The algorithm for generating random numbers using the thinning process method for CPP-HPP and CPP-NHPP has been discussed in Abdullah [16], but it has not been discussed if \( X_i \) is an exponential distribution. Furthermore, in this paper, the algorithm and its generalizations will be modified and developed for generating CPP-HPP-ED and CPP-NHPP-ED.

III. RESULTS AND DISCUSSION

A. CPP-HPP-ED Algorithm

The algorithm generates a compound Poisson process with HPP components and exponential distribution (CPP-HPP-ED) has 3 main steps as follows:

1. Generating an HPP process (denoted \( Z_i \)).
2. Generating a random number \( X_i \) that has an exponential distribution (ED).
3. Calculate CPP (denoted \( Y(t) \)).

Per the definition of HPP and Definition 1 (equation (2)), suppose the process \( \{N(t), t \geq 0\} \) is defined in space \( \Omega \), so that for each \( \omega \in \Omega \) is defined in the probability space \( (\Omega, F, P) \). This process has an intensity function \( \lambda \) and is observed. The function \( N(\omega) \) is a single realization of the number of events that occur at intervals \( (0, t] \) with \( N(0) = 0 \). Suppose that for each data point observed in correspondence with a family of random variables i.i.d with an exponential distribution with the parameter \( \theta \) \( (X \sim \exp(\theta)) \). Each \( Z_i \) is observed, then \( Z_i \) produces a single realization to start generating \( X_i \) and is given a value with an exponential distribution. Following is the modification of the algorithm to generate CPP-HPP-ED which is illustrated in the flowchart of Figure 1.

Fig 1. Flowchart of CPP-HPP-ED algorithm

The following are the algorithm modification steps for CPP-HPP-ED.

<table>
<thead>
<tr>
<th>Step</th>
<th>Instructions</th>
<th>Step</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( t = 0, I = 0, Y = 0 )</td>
<td>5</td>
<td>Generate ( X ) (( X \sim \exp(\theta) ))</td>
</tr>
<tr>
<td>2</td>
<td>Generate ( U )</td>
<td>6</td>
<td>Set ( I = I + 1 ), ( Y = Y + X )</td>
</tr>
<tr>
<td>3</td>
<td>( t = t - \frac{1}{\lambda} \log(U) )</td>
<td>7</td>
<td>Go to step 2</td>
</tr>
<tr>
<td>4</td>
<td>If ( t &gt; T ), stop.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The meanings of the symbols in the flowchart and the algorithm above are respectively: \( t \) represents time, \( I \) represents the number of events that occur at time \( t \), \( Y \) represents CPP-HPP-ED, \( U \) represents random numbers with uniform distribution, \( \lambda \) represents constant intensity.
function, $T$ represents the first time unit and $X$ represents
an exponential random number with parameter $\theta$.

B. CPP-NHPP-ED Algorithm type 1

The algorithm generates a compound Poisson process
with NHPP components and exponential distribution (CPP-
HPP-ED) has 3 main steps as follows:

- Generating an NHPP process (denoted $Z_i$).
- Generating a random number $X_i$ that has an exponential
distribution (ED).
- Calculate CPP (denoted $Y(t)$).

Per the definition of NHPP and Definition 1 (equation
(2)), suppose the process $\{N(t), t \geq 0\}$ is defined in space
$\Omega$, so that for each $\omega \in \Omega$ is defined in the probability
space $(\Omega, F, P)$. This process has an intensity function $\lambda$
and is observed. The function $N(\omega)$ is a single realization
of the number of events that occur at intervals $(0,t]$ with
$N(0) = 0$. Suppose that for each data point observed in
correspondence with a family of random variables i.i.d with
an exponential distribution with the parameter $\theta$ ($X \sim \exp(\theta)$). Remember that to generate NHPP, the value
of $\lambda$ must be chosen so that $\lambda(t) \leq \lambda$ for all $t \leq T$
and with the probability value of $\lambda(t)/\lambda$. Each $Z_i$ is observed,
then $Z_i$ produces a single realization to start generating $X_i$
and is given a value with an exponential distribution. The
following preposition is the basis of the method for
generating CPP-NHPP-ED with a thinning process.

Preposition 1. Consider form of CCP-NHPP in equation (2).
Suppose $Y(T)$ is representing CCP, $X_1^*, X_2^*, ..., X_n^*$ are
representing a family of i.i.d random variables that is
independent of $\{N(t), t \geq 0\}$, and $T_1^*, T_2^*, ..., T_n^*$ are random
variables representing event times from the NHPP with
intensity function, $\lambda^*(t) \geq 0$, and lying in the fixed interval
$(0, t_0]$. Let $\lambda(t)$ be an intensity functions such that
$0 \leq \lambda(t) \leq \lambda^*(t)$ for all $t \in [0, t_0]$. If the ith event time $T_i^*$
is independently deleted with probability $1 - \lambda(t)/\lambda^*(t)$ for
$i = 1, 2, ..., n$, then the remaining event time of NHPP with the
intensity function $\lambda(t)$ in the interval $(0, t_0]$, bring in an
observed realization of $X_{ji}^*$ for $j = 1, 2, ..., n$ and $Y(T)$ can
also be calculated at that interval [16].

Following is the modification of the algorithm to generate
CPP-NHPP-ED which is illustrated in the flowchart of Figure
2.

The meaning of the symbols in the flowchart and the
algorithm are: $t$ represented as time, $I$ represented a number
of events that have occurred by time $t$, $Y^*$ represented as
CPP-NHPP, $U_1$ and $U_2$ represented as random number,
$\lambda(t)$ represented as intensity function of NHPP that is non-
constant function with $\lambda(t) \leq \lambda$, $T$ represented as first $T$
time units and $X$ represented as random variables that
exponential distributed which has i.i.d properties and $\theta$
represented as parameter of exponential distribution.

C. CPP-NHPP-ED algorithm type 2

The algorithm in type 1 is modified and developed by
dividing the interval into subintervals. After being divided,
procedures are used such as type 1. The algorithm generates a
compound Poisson process with NHPP components and
exponential distribution (CPP-HPP-ED) type 2 has 4 main
steps as follows:

- Generating an NHPP process (denoted $Z_j$).
- Generating a random number $X_j$ that has an exponential
distribution (ED).
- Calculate CPP (denoted $Y(t)$).
Per the definition of NHPP and Definition 1 (equation (2)), suppose the process \( \{N(t), t \geq 0\} \) is defined in space \( \Omega \), so that for each \( \omega \in \Omega \) is defined in the probability space \((\Omega, F, P)\). This process has an intensity function \( \lambda \) and is observed. The function \( N(\omega) \) is a single realization of the number of events that occur at intervals \((0, t]\) with \( N(0) = 0 \). Suppose that for each data point observed in correspondence with a family of random variables i.i.d with an exponential distribution with the parameter \( \theta \) (\( X \sim \exp(\theta) \)). Remember that to generate NHPP, the value of \( \lambda_i \) must be chosen so that \( \lambda(s) \leq \lambda_i \) for all \( s \in (t_{i-1} - t_i) \) and with the probability value of \( \lambda(t)/\lambda_i \).

Because exponential has no memory, use \( \lambda_i [Z - (t_i - t)]/\lambda_{i+1} \) for the next exponential form with the rate \( \lambda_{i+1} \). Each \( Z_i \) is observed, then \( X_i \) produces a single realization to start generating \( X_i \) and is given a value with an exponential distribution. The steps to generate the CPP-NHPP-ED algorithm type 2 describe in the flowchart in figure 3 as follows.

![Flowchart of CPP-NHPP-ED algorithm type 2](image)

The steps of the algorithm to generate the CPP-NHPP-ED type 2 as follows:

<table>
<thead>
<tr>
<th>Ste</th>
<th>Instructions</th>
<th>Ste</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( t = 0, J = 1, I = 0, Y^* = ) Generate ( X ) (( X \sim \exp(\theta) )).</td>
<td>9</td>
<td>( \lambda_i ) Set ( Y^* = Y^* + X ).</td>
</tr>
<tr>
<td>2</td>
<td>Generate ( U_1 ).</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( Z = -\frac{1}{\lambda_j} \log(U_1) ) Go to Step 2.</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \text{If } t + Z &gt; t_j, \text{go to Step 12.} )</td>
<td>12</td>
<td>( \text{If } J = k + 1, \text{stop.} )</td>
</tr>
<tr>
<td>5</td>
<td>( \text{Set } t = t + Z ).</td>
<td>13</td>
<td>( Z = \frac{(Z - t_j + t)}{\lambda_{j+1}} )</td>
</tr>
<tr>
<td>6</td>
<td>Generate ( U_2 ).</td>
<td>14</td>
<td>( \text{Set } t = t_j )</td>
</tr>
<tr>
<td>7</td>
<td>( \text{If } U_2 \leq \lambda(t)/\lambda_j ).</td>
<td>15</td>
<td>( \text{Set } J = J + 1 ).</td>
</tr>
<tr>
<td>8</td>
<td>( \text{Set } I = I + 1 ).</td>
<td>16</td>
<td>Go to Step 4.</td>
</tr>
</tbody>
</table>

The meaning of the symbols in the flowchart and the algorithm are: \( t \) represented as present time, \( J \) represented as present interval, \( I \) represented a number of events that have occurred by time \( t \), \( Y^* \) represented as CPP-NHPP, \( U_1 \) and \( U_2 \) represented as uniform random number, \( \lambda(t) \) represented as intensity function of NHPP that is non-constant function with \( \lambda(t) \leq \lambda \), \( X \) represented as random variables that has exponential distributed which has i.i.d properties, and \( \theta \) represented as parameter of exponential distribution.

In the last algorithm, it can be seen that time \( t \) is broken down into several subintervals. This results in the intensity function \( \lambda(t) \) being corrected and also broken down, resulting in fairly complex processes and consuming a lot of memory, but produces more accurate results and the model approaches the desired real phenomenon.

From the results of the random number generation algorithms discussed above, then the expected value and variance can be calculated. Furthermore, these results can be used for model analysis in computational simulations, for example, for reliability analysis, total loss (aggregate loss) and risk measures, for example, the value of risk (VaR) and expected shortfall (ES), in risk theory in risk theory, in various fields of science and engineering.

### IV. CONCLUSIONS

The results of the study obtained several algorithms, namely CPP-HPP-ED algorithm, CPP-NHPP-ED algorithm type 1, and CPP-NHPP-ED algorithm type 2. The first algorithm is generated with components in the Poisson process in the form of constant functions and distributed components exponential. The second algorithm is generated with components in the Poisson process in the form of a non-constant function and an exponential distribution component. The third algorithm is produced by broken down the time interval in the second algorithm and producing more accurate results and the model approaches the desired real phenomenon. These algorithms can be used for computer simulation analysis that can be applied to various fields of science and engineering.

### ACKNOWLEDMENTS

The authors would like to thank referees for remarks and suggestions for improving the quality of this paper, and also thank the editor for improving the appearance of this paper. The work was published with supported by Penelitian Dosen Pemula (PDP) grant scheme from Faculty of Engineering Universitas Sultan Ageng Tirtayasa, Indonesia.

**Brief biography of first author:** Syarif Abdullah, Lecturer from Department of Mechanical Engineering, University of Sultan Ageng Tirtayasa, Indonesia. Specifications of the research is in mathematics, stochastic process, Poisson process, computer analysis, mechanical engineering and computer analysis for engineering.

### REFERENCES

[1] Pahlajani CD, Poulakakis I, Herbert G and Tanner HG 2014. Networked decision making for Poisson processes: Application to...


