Modeling the Count Data of Public Health Service Visits with Overdispersion Problem by Using Negative Binomial Regression

Eha Espinoza, Ully Saputri, Faizal Hafiz Fadilah, Dodi Devianto*

Departement of Mathematics, Faculty of Mathematics and Natural Sciences, Andalas University, Padang 25163, Indonesia

* ddevianto@sci.unand.ac.id

Abstract. The count data of health service visits can be modeled into Poisson regression analysis, where there is no overdispersion assumption by looking at the comparison between mean and variance. The overdispersion test is performed by using the ratio of the sum of Pearson residuals over the number of degrees of freedom that must be less than one. The overdispersion problem can be corrected accurately by building mixture distribution where the parameter of Poisson distribution is made to have Negative Binomial distribution as the theoretical model. The data used in this study are the number of visits to public health service at Padang city as many as 460 data, where the predictor variables are age, gender, education level, occupation, income, home health status, individual health status, health insurance type, distance to health service, and diet type. The best model of negative binomial regression is selected by considering the values of AIC, BIC, Log-likelihood, and overdispersion tests that occur between the resulting models. The final result of this count data model with negative binomial regression fits better and overcomes the overdispersion problem with the significant variable is individual health status for this population, and it can be explained that the more individual has a history of having severe illness the more often the number of visits to the health service, meanwhile the other predictor variables have no effect to the number of visits.

1. Introduction

Regression analysis is a statistical method that is widely used in various fields of science with the aim to know the relationship between the response variable and the predictor variable. In general, regression analysis is used to analyze the response variable which is continuous data and following the assumption of normality [1]. However, in some of its applications the response variables to be analyzed can be discrete data such as number of visits to health services, therefore we need a regression model that can accommodate discrete type data responses.

Modeling demand for health services is a key area of application of arithmetic regression since observed outcome variables retrieve only non-negative integer data, for example, the number of visits to doctors or hospitals. The dominant regression techniques for modeling health services demand are fully parametric, for example, the Poisson model, negative binomial, zero-inflated, and hurdle regression models [2]. The regression modeling with count data as response variable will be a very good alternative which is expected to be able to provide a more representative and robust regression model.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd
As shown in \([2]\), the linearity assumption contained in the conditional mean is probably frequently violated, for example, due to nonlinear or heterogeneous effects. In the context of demand for health services, non linearity may emerge for specific characteristics. For example, a person's number of hospital visits may increase more sharply if the person is suffering from several chronic conditions at the same time, compared to a situation without the previous illness. In addition, heterogeneous effects can lead to violations of the linearity assumption and may not be detected by parametric arithmetic data models \([2,3]\). Heterogeneity refers to different impacts for individuals with different characteristics.

Several models are used to overcome the heterogeneous effect are negative binomial model \([4]\), Penalized Conway-Maxwell-Poisson regression \([5]\), multilevel Zero-Inflated Generalized Poisson regression \([6]\), generalized Poisson with time varying population sizes \([7]\), fitting N-mixture models \([8]\), and generalized Poisson integer-valued GARCH model \([9]\). While the methods used to analyze count data from several previous studies are bell distribution \([10]\), Bayesian additive regression \([11]\), approximate filtering of conditional intensity process \([12]\), gamma block effect \([13]\), Bayesian forecasting \([14]\), longitudinal zero-inflated count data \([15]\), Poisson-Lindey count data \([16]\), modified Akaike's Information Criterion \([17]\).

In the context of public health service visits, nonlinearity may emerge for certain characteristics. Additionally, heterogeneous effects can lead to violations of the linearity assumption. The Poisson regression model for count data has often violated the equidispersion assumption. One model that can be used to overcome the overdispersion problem in count data is the negative binomial model. This paper proposed negative binomial model for public health service visits as the response variable to overcome overdispersion problem in modeling count data.

2. Literature Review

Regression analysis is a method used to analyze the relationship between the dependent variable and several independent variables. In general, the regression analysis method is used to analyze the dependent variable data in the form of continuous data, but in some applications, the dependent variable data to be analyzed can be in the form of discrete data or count data. One of the regression models that can be used to analyze the relationship between the discrete dependent variable in the form of calculated data and the independent variable in the form of discrete, continuous, categorical or mixed data is the Poisson regression model \([1]\).

The Poisson distribution is the distribution of values for a discrete random variable \(X\), which is the number of experimental results that occur in a specific time interval or in a specific area. The Poisson distribution belongs to the exponential family \([10]\) with probability functions

\[
P(X = x) = P(x; \mu) = \frac{e^{-\mu} \mu^x}{x!}; \quad x = 0, 1, 2, \ldots
\]  

with \(\mu\) is the average number of successes that occur in an interval or specific area and \(e = 2.71828\ldots\). The Poisson distribution has the same variance as \(\mu\).

The negative binomial distribution is a model for counting the number of events. The classic approach of the negative binomial distribution that is often used is the negative binomial distribution as a Bernoulli experiment sequence. Suppose that the random variable \(Y\) states the number of experiments required to have \(k\) successes, then \(Y\) has a negative binomial distribution with the probability function as follows \([1]\):

\[
f(y) = \binom{y - 1}{k - 1} p^k (1 - p)^{y - k}; \quad y = k, k + 1, \ldots
\]  

To overcome conditions with response variables that are not normally distributed, but still independent of each other, then the statisticians pioneered by \([19]\) have developed a linear model known as the Generalized Linear Model (GLM). There are three main components in GLM analysis, namely the random element, the systematic component, and the link function.
Poisson regression is a regression model that can be used on data whose dependent variables are not normally distributed and discrete type, which is Poisson distributed as the main condition. Poisson regression is a good choice when the dependent variable \( Y \) is a small integer. The Poisson regression model is written as follows:

\[
y_i = \mu_i + \epsilon_i , \quad i = 1,2,3,...,n
\]  

(3)

where \( y_i \) is the number of occurrences and \( \mu_i \) is the mean number of events for which \( \mu_i \) is assumed not to change from data to data.

Poisson regression uses the Generalized Linear Model (GLM) so that the model can be used in observations, where the dependent variable does not require a normal distribution. In the Generalized Linear Model (GLM), there is a linear function \( g \) that connects the mean of the dependent variable with the independent variable as follow

\[
g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_p x_{pi}
\]  

(4)

the \( g \) function is a link function. The Poisson regression model with log links can be written as follows:

\[
y_i = \exp \left( \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_p x_{pi} \right) + \epsilon_i
\]  

(5)

In the Poisson regression model, there is an underlying assumption in conducting Poisson regression analysis, namely that in the dependent variable there must be equidispersion (mean is equal to variance), \( E(Y) = Var(Y) \). However, in the analysis what sometimes happens is that the variance of the dependent variable is greater than the mean, \( Var(Y) > E(Y) \), this condition is also called overdispersion. The presence or absence of overdispersion can be seen from the Deviance or Pearson Chi-Square values divided by the degrees of freedom. If the Pearson Chi-Square value divided by the degrees of freedom is greater than 1, this indicates an overdispersion of the data [19]. Thus, the negative binomial distribution has an important role in parametric statistical analysis to overcome the overdispersion of data. This negative binomial distribution is obtained from the integration process of the Poisson-Gamma mixture distribution with respect to a parameter \( \delta_i \). Suppose that the \( y_i \) a random variable has a Poisson distribution with the parameter \( \mu_i \). However, \( \mu_i \) itself is a random variable and it is assumed to have a Gamma distribution, namely:

\[
y_i \sim \text{Poisson} (\mu_i)
\]

\[
\mu_i \sim \text{Gamma} (\alpha , \beta)
\]

If a Poisson distribution with parameter \( \mu_i \), where the parameter \( \mu_i \) is the value of a random variable with a Gamma distribution, and a mixed distribution called the Negative Binomial distribution would be generated. The negative binomial regression model assumes the variable \( \delta_i \) has Gamma distribution with mean of 1 and a variance of \( 1/\theta_i \). So that we have \( E(\delta_i) = 1 \) if the parameter \( \beta = \theta_i \) is the part of the mean from Poisson distribution. Let \( v_i \) be the source of the unobserved variability, so that the mean of the Poisson-Gamma mixture distribution is

\[
E(y_i) = \tilde{\mu}_i = \exp (x_i^T \beta + v_i) = \exp (x_i^T \beta) + \exp (v_i) = \mu_i \delta_i
\]  

(6)

where \( \mu_i = \exp (x_i^T \beta) \) is the mean of the Poisson model and \( \delta_i = \exp (v_i) \). Assuming \( E(\delta_i) = 1 \), the Poisson and Negative Binomial models have the same mean, namely \( E(y_i) = E(\delta_i) = \mu_i \). The Poisson-Gamma mixed probability density function can be written as follows:

\[
f(y_i) = \frac{(\mu_i \delta_i)^{y_i} \exp(\mu_i \delta_i)}{y_i!}
\]  

(7)

The variable \( \delta_i \) has Gamma distribution with \( \alpha \) and \( \beta \) parameters [20] as the following

\[
g(\delta_i) = \frac{\beta^\alpha}{\Gamma(\alpha)} \delta_i^{\alpha-1} \exp(-\delta_i \beta)
\]  

(8)
with $E(\delta_i) = \alpha / \beta$ so that it is obtained $E(\delta_i) = 1$, the parameter $\alpha = \beta$ is determined to be $\theta_i$, then we can rewrite the Gamma probability density function becomes

$$g(\delta_i) = \frac{\theta_i^\delta_i}{\Gamma(\theta_i)} \delta_i^{\theta_i - 1} \exp(-\delta_i \theta_i)$$

So it can be written in the form of the marginal function of the Poisson-Gamma mixed distribution is

$$f(y_i) = \int f(y_i | \mu_i, \delta_i) g(\delta_i) d\delta_i$$

From the integral results for the marginal function of the Poisson-Gamma mixed distribution, the general form of the negative binomial regression model is obtained as follows:

$$f(y_i) = \frac{\Gamma(y_i+\theta_i)}{y_i!\Gamma(\theta_i)} \left( \frac{\theta_i}{\mu_i+\theta_i} \right)^{\theta_i} \left( \frac{\mu_i}{\mu_i+\theta_i} \right)^{y_i}$$

After obtaining the regression model, the two models will be compared to get the best mode that can be used to describe the relationship between the dependent variable and the independent variable. The measurement commonly used for selecting the best model is by using Akaike Information Criteria (AIC) and Bayesian Schwartz Information Criteria (BIC). When comparing two or more regression models, the model with the smallest AIC and BIC values is the best model.

3. Data and Method

This article uses primary data from the results of distributing questionnaires to the community in Padang City, West Sumatra, in 2020. The data used is data from Padang City community visits to health services with a total of 460 responses. Data on the number of visits to public health services are the response variable ($Y$), and the predictor variable ($X$) is as follows:

a. Age ($X_1$)

b. Gender ($X_2$) consists of two categories. The first group (0) for men and the second group (1) for women.

c. Education level ($X_3$) with three groups. The first group (0) for basic education, the second group (1) for secondary education, and the third group (2) for higher education.

d. Occupation ($X_4$) consists of seven groups. The first, second, and so on groups are in groups with dummy variables 0, 1, 2, 3, 4, 5, and 6 respectively with information from government employees/police, farmers, laborers, employees, entrepreneurs, not working/retiring, and students/college student.

e. Income ($X_5$) are grouped into seven groups. The first group (0) has a salary below 0.75 million rupiahs, the second group (1) has a salary of 0.75 - 1.5 million rupiahs, the third group (2) has a salary of 1.50 - 3 million rupiahs, the fourth group (3) has a salary of 3 - 6 million rupiahs, groups five, six and seven respectively have salaries of 6 - 10 million rupiahs, 10 - 20 million rupiahs, and above 20 million rupiahs.

f. Home health status ($X_6$) is divided into two groups. The first group (0) with a house status with home health status and the second group (1) with a house status without home health status.

g. Individual health status ($X_7$) consisted of three groups, consecutively had never been seriously ill (0), had been seriously ill (1), and were frequently seriously ill (2).

h. Health insurance ($X_8$) are grouped into three. The first group (0) has no/general insurance, the second group (1) other insurance, and the third group (2) has BPJS insurance.

i. Distance to health service ($X_9$), consists of three groups. Each group, respectively, is less than 2 km (0), a distance of 2 - 5 km (1), and a distance of more than 5 km (2).
j. Diet type ($X_{10}$), consists of three groups. The first group (0) always ate healthy foods, the second group (1) ate fast food occasionally, and the third group (2) had no rules to always eat healthy foods.

In this research, SPSS software was used to analyze Poisson regression and negative binomial. First, the Kolmogorov Smirnov test was conducted to test whether the data on the number of visits to health services as a response variable ($Y$) followed the Poisson distribution or not.

### 4. Regression Model for Count Data of Public Health Service Visits

The first step of count data regression model for public health service visits is performed by using Poisson regression model. The following are the results of parameter estimation in the Poisson regression model using SPSS software.

| Parameter | Estimation Value | Std. Error | P-Value |
|-----------|------------------|------------|---------|
| (Intercept) | .856 | .5576 | .125 |
| Age ($X_{1}$) | .007 | .0035 | .045 |
| [Gender=0] ($X_{50}$) | -.136 | .0743 | .067 |
| [Education=0] ($X_{50}$) | -.331 | .1930 | .086 |
| [Education=1] ($X_{31}$) | -.252 | .1031 | .014 |
| [Occupation=0] ($X_{90}$) | -.138 | .1830 | .451 |
| [Occupation=1] ($X_{41}$) | .053 | .1968 | .789 |
| [Occupation=2] ($X_{42}$) | .084 | .1906 | .661 |
| [Occupation=3] ($X_{43}$) | -.171 | .1538 | .266 |
| [Occupation=4] ($X_{44}$) | .012 | .2021 | .952 |
| [Occupation=5] ($X_{45}$) | -.003 | .1566 | .984 |
| [Income=0] ($X_{50}$) | .916 | .4945 | .064 |
| [Income=1] ($X_{51}$) | 1.046 | .4816 | .030 |
| [Income=2] ($X_{52}$) | 1.118 | .4699 | .017 |
| [Income=3] ($X_{53}$) | 1.144 | .4648 | .014 |
| [Income=4] ($X_{54}$) | 1.228 | .4663 | .008 |
| [Income=5] ($X_{55}$) | .640 | .4881 | .190 |
| [Home health status=0] ($X_{60}$) | -.127 | .1109 | .252 |
| [Individual health status=0] ($X_{70}$) | -1.272 | .1448 | .000 |
| [Individual health status=1] ($X_{71}$) | -.447 | .1433 | .002 |
| [Health insurance=0] ($X_{80}$) | -.277 | .1001 | .006 |
| [Health insurance=1] ($X_{81}$) | .089 | .1032 | .390 |
| [Distance to health service=0] ($X_{90}$) | .022 | .1028 | .828 |
| [Distance to health service=1] ($X_{91}$) | -.044 | .1045 | .675 |
| [Diet type =0] ($X_{100}$) | -.067 | .0996 | .501 |
| [Diet type =1] ($X_{101}$) | -.042 | .0813 | .604 |

From the table above, it can be seen that the significant parameters are $X_1$, $X_{31}$, $X_{51}$, $X_{52}$, $X_{53}$, $X_{54}$, $X_{70}$, $X_{71}$, and $X_{80}$. This means that age, secondary education level, income parameters, individual health status of never having been seriously ill and having been seriously ill, as well as non-public or non-existent insurance status, affect the number of visits to health services in Padang City. So that the Poisson regression model is obtained as follows:

$$y_i = \exp(0.856 + 0.007X_1 - 0.252X_{31} + 1.046X_{51} + 1.118X_{52} + 1.144X_{53} + 1.228X_{54} - 1.272X_{70} - 0.447X_{71} - 0.277X_{80})$$
Data that spreads according to the Poisson distribution has a requirement that the variance is equal to the expected value (equidispersion). To ensure that the data meets these assumptions, the following are the results of the SPSS goodness of fit output from the data on the number of health service visits using the Poisson regression model.

**Table 2. Goodness of Fit of Poisson Regression Model**

| Value       | Df  | Value/df |
|-------------|-----|----------|
| Deviance    | 544.126 | 434 | 1.254 |
| Pearson Chi-Square | 500.059 | 434 | 1.152 |

Because the Deviance value and the Pearson Chi-Square value divided by the degrees of freedom are greater than 1 so that the health service data is an overdispersion in the resulting Poisson regression model, this is not good because the resulting error rate is higher. Therefore, one way to overcome the overdispersion data case is to analyze using the negative binomial regression model.

**Table 3. Parameter Estimation of Negative Binomial Regression**

| Parameter | Estimation Value | Std. Error | Sig. |
|-----------|------------------|------------|------|
| (Intercept) | .762 | .8684 | .380 |
| Age (X_1) | .009 | .0059 | .111 |
| [Gender=0] (X_{20}) | -.146 | .1332 | .274 |
| [Education=0] (X_{30}) | -.472 | .3957 | .233 |
| [Education=1] (X_{31}) | -.256 | .1847 | .165 |
| [Occupation=0] (X_{40}) | .205 | .3129 | .512 |
| [Occupation=1] (X_{41}) | .100 | .3660 | .785 |
| [Occupation=2] (X_{42}) | .100 | .3438 | .770 |
| [Occupation=3] (X_{43}) | -.174 | .2558 | .496 |
| [Occupation=4] (X_{44}) | -.010 | .3397 | .977 |
| [Occupation=5] (X_{45}) | -.142 | .2724 | .603 |
| [Income=0] (X_{50}) | 1.093 | .7228 | .130 |
| [Income=1] (X_{51}) | 1.202 | .7040 | .088 |
| [Income=2] (X_{52}) | 1.328 | .6804 | .051 |
| [Income=3] (X_{53}) | 1.361 | .6723 | .043 |
| [Income=4] (X_{54}) | 1.441 | .6822 | .035 |
| [Income=5] (X_{55}) | .830 | .7142 | .245 |
| [Home health status=0] (X_{60}) | -.129 | .2141 | .546 |
| [Individual health status=0] (X_{70}) | -.1410 | .3418 | .000 |
| [Individual health status=1] (X_{71}) | -.557 | .3460 | .107 |
| [Health insurance=0] (X_{80}) | -.279 | .1800 | .121 |
| [Health insurance=1] (X_{81}) | .151 | .1973 | .444 |
| [Distance to health service=0] (X_{90}) | -.046 | .1848 | .805 |
| [Distance to health service=1] (X_{91}) | -.105 | .1867 | .573 |
| [Diet type=0] (X_{100}) | -.020 | .1761 | .908 |
| [Diet type=1] (X_{101}) | -.007 | .1451 | .962 |

From the table above, it can be seen that the significant parameters are \( X_{53}, X_{54}, \) and \( X_{70} \). This means that the income parameter of the fourth group (3) with an interval of 3 – 6 million rupiahs, the parameter of the salary of the fifth group with an interval of 6 – 10 million rupiahs and the parameter of individual health with the status of never being seriously ill, affects the number of visits to public health services in Padang City. So that the negative binomial regression model is obtained as follows:
Then performed the SPSS goodness of fit analysis for the negative binomial regression model, the following output was obtained:

| Table 4. Goodness of Fit |
|--------------------------|
|                          |
| Deviance                 |
| 230.383                  |
| df                      |
| 434                     |
| Value/df                |
| .531                    |
| Pearson Chi-Square       |
| 174.886                  |
| df                      |
| 434                     |
| Value/df                |
| .403                    |

Based on the table above, it can be seen that if the Deviance value and the Pearson Chi-Square value are divided by the degrees of freedom, it results in a smaller value of 1. When compared with the Deviance value and Pearson Chi-Square value in the Poisson regression, it has decreased after using the Negative Binomial regression.

So that the data on the number of visits to Padang City public health services that have overdispersion cases in the Poisson regression data can be overcome by analysis using negative binomial regression. After estimating using the negative binomial regression, there are several parameters that are not significant, so these parameters are removed from the model and tested again until the estimation results are obtained as follows:

| Table 5. Parameter Estimates |
|------------------------------|
| Parameter                   |
| B                            |
| Std. Error                   |
| Sig.                         |
| (Intercept)                  |
| .916                        |
| .6744                       |
| .174                        |
| [Income=0]                   |
| .767                        |
| .6310                       |
| .224                        |
| [Income =1]                  |
| .926                        |
| .6317                       |
| .143                        |
| [Income =2]                  |
| 1.039                       |
| .6313                       |
| .100                        |
| [Income =3]                  |
| 1.129                       |
| .6329                       |
| .074                        |
| [Income =4]                  |
| 1.194                       |
| .6463                       |
| .065                        |
| [Income =5]                  |
| .682                        |
| .6750                       |
| .312                        |
| [Individual health status=0]|
| -1.413                      |
| .2706                       |
| .000                        |
| [Individual health status =1]|
| -.561                       |
| .2907                       |
| .054                        |

Based on the table above, the negative binomial regression model with significant parameters is obtained as follows:

\[ y_i = \exp(0.916 - 1.413X_{70}) \]

The last step on modeling the count data is to select the best model. In determining the best model is to compare the smallest AIC and BIC values among the three models above, namely the Poisson regression model, the negative binomial regression model, and the negative binomial regression model with significant parameters. Below are the AIC and BIC values of each model:

| Table 6. AIC and BIC values of each model |
|------------------------------------------|
| Poisson Regression Model                  |
| AIC                                       |
| 1610.846                                  |
| BIC                                       |
| 1718.258                                  |
| Negative Binomial Regression Model         |
| AIC                                       |
| 1782.885                                  |
| BIC                                       |
| 1890.297                                  |
| Negative Binomial Regression Model with significant parameters |
| AIC                                       |
| 1759.228                                  |
| BIC                                       |
| 1796.409                                  |
According to the Table 6, it is found that the smallest AIC and BIC values are a model with Poisson regression. But this violate the assumption that the data must meet the equidispersion case so that the Binomial Negative regression model with significant parameters is the best model because it has the smallest AIC and BIC values when compared to the Negative Binomial regression model with all parameters.

5. Conclusion

Poisson regression model and negative binomial regression model were used to model count data of public health service visits. If an overdispersion occurs but Poisson regression is still used, the standard error will increase, so the conclusion will be invalid. Based on the AIC and BIC values, even though Poisson regression has the smallest values compared to other models, but the overdispersion is occurred. The final result of this count data model with negative binomial regression fits better and overcomes the overdispersion problem with the significant variable is individual health status for this population and it can be explained that the more individual has a history of having severe illness the more often the number of visits to the health service, meanwhile the other predictor variables have no effect to the number of visits.

References

[1] Cameron A C and Trivedi P K 1998 Regression Analysis of Count Data Cambridge University Press
[2] Winkelmann R 2008 Econometric Analysis of Count Data 5th edition Berlin Springer
[3] Bach P, Farbmacher H and Spindler M 2018 Semiparametric count data modeling with an application to health service demand Econometrics and Statistics 8 125-140
[4] Saez-Castillo A J and Conde-Sanchez A 2013 A hyper-Poisson regression model for overdispersed and underdispersed count data Computational Statistics and Data Analysis 61 141-157
[5] Abdella G M, Kim J, Al-Khalifa K N and Hamouda A M 2019 Penalized Conway-Maxwell-Poisson regression for modelling dispersed discrete data: The case study of motor vehicle crash frequency Safety Science 120 157-163
[6] Almasi A, Eshraghian M R, Moghimbeigi A, Rahimi A, Mohammad K and Fallahigilan S 2016 Multilevel zero-inflated Generalized Poisson regression modeling for dispersed correlated count data Statistical Methodology 30 1-14
[7] Ding D and Li J 2020 Efficient monitoring of overdispersed counts with time varying population sizes Computers and Industrial Engineering 143 106-409
[8] Duarte A, Adams M J and Peterson J T 2018 Fitting N-mixture models to count data with unmodeled heterogeneity: Bias, diagnostics, and alternative approaches Ecological Modelling 374 51-59
[9] Zhu F 2012 Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models Journal of Mathematical Analysis and Applications 389 58-71
[10] Castellares F, Ferrari S L P and Lemonte A J 2018 On the bell distribution and its associated regression model for count data Applied Mathematical Modelling 56 172-185
[11] Krueger R, Bansal P and Buddhavarapu P 2020 A new spatial count data model with Bayesian additive regression trees for accident hot spot identification Accident Analysis and Prevention 144 105-623
[12] Santitissadeekorn N, Lloyd J B, Short M B and Delahaies S 2019 Approximate filtering of conditional intensity process for Poisson count data: application to urban crime Computational Statistics and Data Analysis
[13] Schmidt M and Schwabe R 2020 Optimal designs for Poisson count data with Gamma Block effects Journal of Statistical Planning and Inference 204 128-140
[14] Nariswari R and Pudjihastuti H 2019 Bayesian forecasting for time series of count data
[15] Feng J and Zhu Z 2011 Semiparametric analysis of longitudinal zero-infated count data \textit{Journal of Multivariate Analysis} \textbf{102}

[16] Fernandez A J 2019 Optimal lot disposition from Poisson-Lindley count data \textit{Applied Mathematical Modelling} \textbf{70} 595-604

[17] Latief AHMH, Hossain M Z and Islam M A 2008 Model selection using modified Akaike's Information Criterian: An application to maternal morbidity data \textit{Australian Journal of Statistics} \textbf{37} 175-184

[18] Devianto D 2018 On The Class of Infinitely Divisible Exponential Distribution \textit{Journal of Physics C}onf. Ser. \textbf{1097} 1-6

[19] McCullagh and Nelder J A 1987 \textit{Generalized Linear Models} London Chapman and Hall 2nd Edition

[20] Devianto D, Maiyastri M and Azmi K 2020 On the Levy Measure Property of Gamma Distribution \textit{AP Conf.Proc.} \textbf{2242} 1-8