Dhaka Univ. J. Sci. 72(1): 24-29, 2024 (January)

ISSN 1022 - 2502, E-ISSN 2408- 8528
DOI: https://doi.org/10.3329/dujs.v72i1.71181

Analyzing Correlated Road Accident Count Data Using Zero Truncated Bivariate Poisson Regression Model

Trishna Saha and Anamul Haque Sajib[*]

Department of Statistics, University of Dhaka, Dhaka-1000, Bangladesh.

(Received : 18 June 2023; Accepted : 13 December 2023)

Abstract

This paper aims to determine the significant factors which influence two correlated count responses, namely the total number of cars involved in an accident and the total number of fatalities due to that accident, of United Kingdom (UK) road accident count data. The bivariate Poisson (BVP) of two different forms and zero truncated bivariate Poisson regression (ZTBVP) models are considered to analyze UK road accident count data and the best model is selected based on the AIC and BIC values. From the data analysis, it is observed that the ZTBVP model provides the best fit (AIC value: 20563.26) for the UK road accident count data compared to all two variants of the BVP model (AIC value: >20563.26). From the results obtained from ZTBVP model, it is also observed that sex of driver, area, serious severity, and light condition are the significant covariates for the total number of cars involved in an accident while area, fatal severity, serious severity, light condition and year 2021 are the significant covariates for the total number of fatalities due to that accident.

Keywords: Bivariate Poisson Regression, Zero Truncated Bivariate Poisson Regression, Akaike Information Criterion, and Bayesian Information Criterion.

I. Introduction

Accidents are unfortunate occurrences that happen unexpectedly and unintentionally due to a variety of factors, causing harm to people and property. Such unfortunate events can be prevented or mitigated if the risk factors responsible for such unfortunate events are identified ahead of time and effective counter measures are implemented. Therefore, public health researchers or higher authorities of a state want to identify the factors which are responsible for such unfortunate events. In addition to identifying the responsible factors for such unfortunate events, they also want to know the total number of fatalities or total amount of accident related cost which depends on how many vehicles are involved in an accident. Data obtained from such types of real life problems are known as correlated or paired data. This paired data is called paired count data when both of the response variables denote the total number of counts.

Paired count data arise vastly in our everyday life from different disciplines, most importantly, medical science, engineering and public health. For example, the total number of cars involved in an accident and the total number of fatalities as a result of this accident are considered as paired count variables. Furthermore, the frequency of antenatal care (ANC) visits and number of antenatal care services received by a pregnant mother is another example of paired count data.

Researchers' main aim is to investigate the effects of covariates on these count outcomes through suitable statistical modeling. Choosing a suitable statistical model to model a phenomenon depends on the context of the problem. The univariate and bivariate Poisson regression model can be used as a primary model to model a phenomenon with single and correlated count responses respectively. When the count response is overdispersed, the Quasi-Poisson or the negative binomial regression model can be used as a primary model. Furthermore, zero inflated Poisson and zero inflated negative binomial regression model can be used when response possess excess zero count for equidispersed and overdispersed count data respectively. The road accident data considered in this paper are correlated zero truncated count data and suitable models for such types of data are any zero truncated bivariate count models.

Chowdhury and Islam (2016) proposed covariate dependent zero truncated bivariate Poisson model: marginal conditional approach to analyze UK road safety data collected from 2005–2013, published by the Department of Transport, United Kingdom and compared the performance of their proposed model with the performance of zero truncated bivariate Poisson model without covariates (null model). They showed that their proposed model offered better fittings to the UK road safety data compared to the null model. However, they did not explore how other competing models perform in such situations compared to the performance of their proposed model.

Motivated by the lack of their study, we consider two other competing models along with their proposed model to analyze UK road safety data and compare the performance of all models considered in this paper. This paper also considers an updated road safety data set recorded from 2017–2021 instead of data recorded from 2005–2013. All

[*] Author for correspondence. e-mail: sajibstat@du.ac.bd

the results presented in this paper are produced for updated road safety data, and the detailed discussion about the similarities and dissimilarities between the updated and previous data sets are provided in the data description section.

The rest of the paper is organized as follows. In section 2, the related methodologies of all competing models considered in this paper are presented in detail. A detailed overview of the data and variables are presented in section 3. Finally, we discuss the results of our study which is followed by the conclusion presented in section 4 and section 5, respectively.

II. Methodology

In this study, the bivariate Poisson (BVP) of two different forms and zero truncated bivariate Poisson regression (ZTBVP) models are considered to analyze UK road accident count data.

Bivariate Poisson model (BVP-1)

Consider the random variables X_i, i= 1, 2, 3 which are independent Poisson distribution with parameters λ_i, respectively. Let the random variables $X = X_1 + X_3$ and $Y = X_2 + X_3$. Here, $X \sim Pois(\lambda_1 + \lambda_3)$ with $E(X) = \lambda_1 + \lambda_3$ and $Y \sim Pois(\lambda_1 + \lambda_3)$ with $E(Y) = \lambda_2 + \lambda_3$. The joint probability mass function of X and Y is bivariate Poisson distribution, $BP(\lambda_1, \lambda_2, \lambda_3)$, which was proposed by Holgate (1964), written as:

$$f_{BP}(x, y|\lambda_1, \lambda_2, \lambda_3)$$
$$= e^{-(\lambda_1+\lambda_2+\lambda_3)} \frac{{\lambda_1}^x}{x!} \frac{{\lambda_2}^y}{y!} \sum_{i=0}^{\min(x,y)} i! \binom{x}{i} \binom{y}{i} \left(\frac{\lambda_3}{\lambda_1 \lambda_2}\right)^i.$$

The two random variables can positively depend on one another according to the bivariate distribution shown above. Moreover, $COV(X, Y) = \lambda_3$, and λ_3 represents the degree of dependence between the two random variables. The bivariate Poisson distribution simplifies to the double Poisson distribution if $\lambda_3 = 0$, indicating that the two variables are independent. The correlation coefficient ρ between X and Y is $\frac{\lambda_3}{\sqrt{(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)}}$. Under generalized linear model framework, $ln\lambda_1 = \boldsymbol{W}_1{}'\boldsymbol{\beta}_1$, $ln\lambda_2 = \boldsymbol{W}_2{}'\boldsymbol{\beta}_2$ and $ln\lambda_3 = \boldsymbol{W}_3{}'\boldsymbol{\beta}_3$ are considered as the link functions where $\boldsymbol{\beta}$ denotes the regression coefficients. To estimate these regression coefficients, we use the Expectation Maximization (EM) estimation technique. EM algorithm is an iterative procedure for maximum likelihood estimation. It is primarily useful when there is missing data. Missing data can be (i) actual missing data (ii) hypothetical variable which makes likelihood function simpler to solve. EM algorithm produces a sequence of values that converge to a stationary value. Each iteration consists of two steps: E-step and M-step. In the E step missing value is replaced by its conditional expectation while expected log likelihood is maximized in the M step. In order to construct the EM-algorithm we need to estimate the unobserved data by their conditional expectations and then fit Poisson regression models to the pseudo values obtained by the E-step. Denoting as $\varPhi$ the entire vector of parameters, that is $\varPhi = (\boldsymbol{\beta}_1{}', \boldsymbol{\beta}_2{}', \boldsymbol{\beta}_3{}')$, the complete data log-likelihood is given by

$$L(\varPhi) = -\sum_{i=1}^{n} \sum_{k=1}^{3} \lambda_{ki} + \sum_{i=1}^{n} \sum_{k=1}^{3} x_{ki} \ln(\lambda_{ki}) - \sum_{i=1}^{n} \sum_{k=1}^{3} \ln(x_{ki}!).$$

The EM-algorithm works in two steps:

E-step: Using the current parameter values of k iteration noted by $\Phi^{(\mathrm{k})}$, ${\lambda_{1i}}^{(\mathrm{k})}$, ${\lambda_{2i}}^{(\mathrm{k})}$ and ${\lambda_{3i}}^{(\mathrm{k})}$ calculate the conditional expected values of X_{3i}, for $i = 1, 2, 3, \dots \dots, n$ which is denoted by $s_i = E\left(X_{3i} \middle| X_i, Y_i, \Phi^{(\mathrm{k})}\right)$

$$= \begin{cases} {\lambda_{3i}}^{(k)} \dfrac{f_{BP}\left(x_i - 1, y_i - 1|{\lambda_{1i}}^{(\mathrm{k})}, {\lambda_{2i}}^{(\mathrm{k})}, {\lambda_{3i}}^{(\mathrm{k})}\right)}{f_{BP}\left(x_i, y_i|{\lambda_{1i}}^{(\mathrm{k})}, {\lambda_{2i}}^{(\mathrm{k})}, {\lambda_{3i}}^{(\mathrm{k})}\right)} \\ 0 \end{cases}$$

M-step: Update the estimates

$$\boldsymbol{\beta}_1^{(k+1)} = \widehat{\boldsymbol{\beta}}(\boldsymbol{x} - \boldsymbol{s}, \boldsymbol{W}_1),$$
$$\boldsymbol{\beta}_2^{(k+1)} = \widehat{\boldsymbol{\beta}}(\boldsymbol{y} - \boldsymbol{s}, \boldsymbol{W}_2),$$
$$\boldsymbol{\beta}_3^{(k+1)} = \widehat{\boldsymbol{\beta}}(\boldsymbol{s}, \boldsymbol{W}_3),$$
$$\boldsymbol{\lambda}_{ki}^{(k+1)} = \exp\left(\boldsymbol{W}_{ki}' \widehat{\boldsymbol{\beta}}_k^{(k+1)}\right) \; for \;\; k = 1, 2, 3,$$

where $s = (s_1, \dots, s_n)'$ is the $n \times 1$ vector calculated in the E-step, $\widehat{\boldsymbol{\beta}}(\boldsymbol{x}, \boldsymbol{W})$ are the maximum likelihood estimates of a Poisson model with response vector **x** and design or data matrix given by **W**. Each data matrix $\boldsymbol{W}_k$ is a $n \times p_k$ matrix and W_{ki}' is its corresponding i^{th} row $(for\ i = 1, \dots, n)$. If we wish to have common (or equal) parameters among different λ_k then we should construct a common design matrix **W** and the corresponding parameter vector $\boldsymbol{\beta}$ will be estimated as $\boldsymbol{\beta}^{k+1} = \widehat{\boldsymbol{\beta}}(\boldsymbol{u}, \boldsymbol{W})$ with $u' = (x' - s',\ y' - s',\ s')$. In the functions provided, we have considered the possibility to have common parameters only between λ_1 and λ_2. It is also noted that standard GLM procedures can be used for the M-step despite the fact that the responses are not any more integers. The latter does not cause any numerical problems.

Bivariate Poisson Model (BVP-2)

The joint distribution of BVP proposed by Islam and Chowdhury (2016) is

$$f(x,y) = \frac{e^{-\lambda_1}\lambda_1{}^{x} e^{-\lambda_2 X}(\lambda_2\, x)^{y}}{y!\, x!},$$

where $x = 0,1,2,\ldots, y = 0,1,2,\ldots,\lambda_1, \lambda_2 > 0$. The exponential version of the above equation can be expressed as $f(x,y) = \exp(xln\lambda_1 - ln(x!) - \lambda_1 + yln(\lambda_2 x) - \lambda_2\, x - ln(y!))$, where the link functions are $ln\lambda_1 = \boldsymbol{W'\beta}_1$ and $\lambda_2 = \boldsymbol{W'\beta_2}$ and $\boldsymbol{\beta}$ denotes the regression coefficients. The maximum likelihood estimation technique is used to estimate these regression coefficients. The estimates of the regression parameters vectors β_1 and β_2 can be obtained iteratively by using Newton-Raphson method as follows

$$\widehat{\boldsymbol{\beta}}_t = \widehat{\boldsymbol{\beta}}_{t-1} + {I_0}^{-1}(\widehat{\boldsymbol{\beta}}_{t-1})\, U(\widehat{\boldsymbol{\beta}}_{t-1}),$$

where $\widehat{\boldsymbol{\beta}}_t$ denotes the estimate at t^{th} iteration, ${I_0}^{-1}\,(\widehat{\boldsymbol{\beta}}_{t-1})$ is the information matrix of the parameters and $U(\widehat{\boldsymbol{\beta}}_{t-1})$ is the score function of the parameters.

Zero Truncated Bivariate Poisson (ZTBVP) Model*:*

The joint distribution of ZTBVP proposed by Chowdhury and Islam (2016) is

$$f(x,y) = \frac{(\lambda_2\, x)^{y}\lambda_1{}^{x}}{y!\, x!\, (e^{\lambda_2\, X} - 1)(e^{\lambda_1} - 1)},$$

where $x = 1,2,\ldots, y = 1,2,\ldots,\lambda_1, \lambda_2 > 0$.

The ZTBVP's exponential form can be written as

$$f(x,y) = exp\big[xln\lambda_1 \text{-} ln(x!) - ln\big(e^{\lambda_1} - 1\big) + yln\lambda_2 + ylnx - ln(y!) - ln\big(e^{\lambda_2 X} - 1\big)\big],$$

where the link functions are $ln\lambda_1 = W'\boldsymbol{\beta}_1$ and $\lambda_2 = W'\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}$ denotes the regression coefficients. To estimate these regression coefficients we use maximum likelihood estimation technique. The estimates of the regression parameters vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ can be obtained iteratively by using Newton-Raphson method as follows

$$\widehat{\boldsymbol{\beta}}_t = \widehat{\boldsymbol{\beta}}_{t-1} + {I_0}^{-1}(\widehat{\boldsymbol{\beta}}_{t-1})\, U(\widehat{\boldsymbol{\beta}}_{t-1}),$$

where $\widehat{\boldsymbol{\beta}}_t$ denotes the estimate at t^{th} iteration, ${I_0}^{-1}(\widehat{\boldsymbol{\beta}}_{t-1})$ is the information matrix of the parameters and $U(\widehat{\boldsymbol{\beta}}_{t-1})$ is the score function of the parameters.

AIC and BIC

'Information criterion' (IC) is used to select optimal model. Most of the ICs are calculated using logarithm of likelihood and a penalized term. Based on this 'penalty', they are known with different name. AIC and BIC are the most common ICs found in the literature. The mathematical form of AIC and BIC are:

$$AIC = -2\ln(L) + 2k,$$

$$BIC = -2\ln(L) + \ln(N) \times k,$$

where N is number of observations and K is number of parameters to be estimated. Both of the above ICs make same decision for simple models. Different conditions, however, lead to different conclusions, and arguably none of them captures the full complexity of real model selection problems. So, it is better to choose optimal model when both of them makes same conclusion [Kuha' 2004].

Table 1. Type of the selected variables

Variables	Type	Leveling
Outcome ($\boldsymbol{X}$)	Count	-
Outcome ($\boldsymbol{Y}$)	Count	-
Sex of Driver	Categorical	Male and Female
Area	Categorical	Rural and Urban
Accident Severity	Categorical	Fatal, Serious and Slight
Light Condition	Categorical	Daylight, and Others
Year	Categorical	2017, 2018, 2019, 2020, 2021

III. Data and Variables

Chowdhury and Islam (2016) showed the application of their proposed ZTBVP model on UK road safety data collected from 2005-2013, published by the Department of Transport, United Kingdom. In this paper, we will also consider ZTBVP along with other two variants of the BVP model for analyzing updated road safety data recorded from 2017−2021. In the updated data set, there are 562439 observations with many variables. A random sample of 5624 accident reports is chosen, representing approximately 1% of all accident records. In the context of our problem, we have two outcome variables namely number of cars involved in an accident (X) and the total number of fatalities (Y) while sex of driver, accidental area, accident severity, lighting condition and years are explanatory variables. All of these factors were taken into account by Chowdhury and Islam (2016) as an explanatory factors. It is noted that all the categorical variables considered here are dummy variables. Table 1 shows the number of variables and their corresponding type considered from the road safety data set.

Dummy variables for the above five categorical variables are created as: sex of the driver (male = 1; female = 0), area (rural = 1; urban = 0), accident severity (fatal severity = 1, else 0; serious severity = 1, else = 0; slight severity is the reference category), light condition (daylight = 1, others = 0) and four dummy variables for year 2018 to year 2021, where year 2017 is considered as reference category.

Bivariate count data considered in this paper need to be tested whether they are overdispersed or not. The Likelihood Ratio test and the Dean‘s test are considered to test whether there exists

Overdispersion in the count data. The hypothesis is defined as H_0: count data is not overdispersed versus H_1: there exists overdispersion in the count data. To test the overdispersion, we use DCluster R package in R programming language. From the result (shown in Table 2), we see that the p-value of this three tests are very high (1.00) that means we may not reject the null hypothesis. So, there is no overdispersion in the road accident data set.

Table 2. Results of Overdispersion test

Variable	Test Name	Test statistic	p-value
X	Likelihood Ratio Test	-0.062	1
	Dean's P_B	-38.882	1
	Dean's $\acute{P}_B$	-38.789	1
Y	Likelihood Ratio Test	-0.074224	1
	Dean's P_B	-32.13	1
	Dean's $\acute{P}_B$	-32.031	1

IV. Results and Discussion

In this section, firstly, we decide which covariates are used as regressors to model λ_3 in BVP-1 model. Here, we have fitted seven BVP-1 models under GLM setup of λ_3: 1) a model without any covariance; 2) a model with constant covariance term; 3) a model with covariate sex of driver on the covariance term λ_3; 4) a model with covariates sex of driver and area on the covariance term λ_3; 5) a model with covariates sex of driver, area, and fatal severity on the covariance term λ_3; 6) a model with covariates sex of driver, area, fatal severity, and serious severity on the covariance term λ_3; 7) a model with covariates sex of driver, area, fatal severity, serious severity, and light condition on the covariance term λ_3; 8) a model with covariates sex of driver, area, fatal severity, serious severity, light condition and all years on the covariance term λ_3. For λ_1 and λ_2 we consider sex of driver, area, fatal severity, serious severity, light condition and all years as covariates. Now, different performance criteria of the 8 fitted models are presented in Table 3.

Table 3. Results from the fitted bivariate Poisson (BVP-1) models for the accident data

Models	AIC	BIC	Log-Like	Par
1	29092.05	29238.60	-14526.02	20
2	27255.67	27409.56	-13606.84	21
3	27257.51	27418.73	-13604.63	22
4	27255.26	27423.80	-13604.63	23
5	27256.33	27432.20	-13604.16	24
6	27258.33	27449.67	-13604.16	25
7	27259.14	27449.67	-13603.57	26
8	27266.28	27486.11	-13603.14	30

Table 3 shows that BIC is the minimum for model 2, which is 27409.56. But AIC is the minimum for model 4, which is 27255.26 and 2nd minimum value of AIC is got for model 2, which is 27255.67. The AIC values of the 2nd and 4th models differ by a little amount in this area, and the 4th model has 2 covariates with constant terms, but the 2nd model contains just one constant term which is estimated from the data as a constant covariate. If we consider model 4, we can see that to gain a modest quantity of AIC value which is 0.41, a large number of covariates are required to estimate, and the interpretation will be challenging. Due to this, we decided to choose the second one as the optimal model among these 8 models and compare this second model (BVP-1) to the BVP-2 and ZTBVP models. Now the question is which model is the best to analyze zero truncated bivariate count data? Here, we use AIC and BIC to select the optimal model among these three models. Table 3 represents the results of the log likelihood, AIC, BIC, and number of parameters of the BVP-1, BVP-2, and ZTBVP models.

Table 4. Test statistics results of BVP and ZTBVP models

Model **Statistics**	**BVP-1**	**BVP-2**	**ZTBVP**
Log likelihood	-14526.02	-14727.21	-10261.63
AIC	27255.67	29494.42	20563.26
BIC	27409.56	29627.11	20695.96
Par	21	20	20

From Table 4, it is observed that the AIC and BIC values for BVP-1, BVP-2, and ZTBVP models are 27255.67 and 27409.56, 29494.42 and 29627.11, and 20563.26 and 20695.96, respectively. This demonstrates that the ZTBVP model has lower AIC and BIC values than the BVP-1 and BVP-2 models. Therefore, the ZTBVP model can be considered as a better model to analyze UK road accident data compared to all two variants of the BVP model as far as AIC and BIC are concerned.

Finally, the ZTBVP model was chosen as a better model to analyze UK road accident data over the BVP-1 and BVP-2 models based on the above performance comparison of the BVP-1, BVP-2, and ZTBVP models. Therefore, only regression outputs obtained from the ZTBVP model were considered to determine the significant factors of road accident which are presented in Table 5.

Table 5 . Parameter estimates of ZTBVP model for road safety data

ZTBVP			
Variables	Estimate	Std. Error	p-value
Marginal Model for X			
constant	0.193198	0.042801	0.000006^{***}
sex of driver	0.068552	0.028521	0.016268^{*}
area	0.077869	0.028342	0.006024^{**}
fatal severity	-0.111077	0.114856	0.333539
serious severity	-0.170630	0.036921	0.000004^{***}
light condition	0.115211	0.030778	0.000183^{***}
year 2018	0.002491	0.040179	0.950559
year 2019	0.016183	0.040403	0.688786
year 2020	-0.012440	0.043519	0.775010
year 2021	0.009116	0.042737	0.831096
Conditional Model for Y			
constant	-1.283057	0.071122	0.000000^{***}
sex of driver	-0.051581	0.048905	0.291595
area	0.559077	0.046801	0.000000^{***}
fatal severity	0.790931	0.113374	0.000000^{***}
serious severity	0.174728	0.057834	0.002529^{**}
light condition	-0.119616	0.050741	0.018439^{*}
year 2018	-0.105068	0.067058	0.117213
year 2019	-0.116095	0.067947	0.087578
year 2020	-0.049916	0.071890	0.487500
year 2021	-0.277163	0.077233	0.000335^{***}

$^{*}p<0.05, ^{**}p<0.01, ^{***}p<0.001$

From Table 5, it is observed that in the marginal model, the explanatory variables such as area, sex of a driver, serious severity and light condition have statistically significant effects on the total number of cars involved in an accident while fatal severity, year 2018, year 2019, year 2020 and year 2021 do not have statistically significant effects on the total number of cars involved in an accident. More specifically, the effects of serious severity and light condition on the total number of cars involved in an accident are statistically significant at 0.1% level while the effects of area on the total number of cars involved in an accident and the effects of sex of driver on the total number of cars involved in an accident are statistically significant at 1% and 5% levels, respectively. On the other hand, Chowdhury and Islam's found all the variables are statistically significant except years from their analysis. The variable fatal severity was found to be significant in the analysis conducted by Chowdhury and Islam's but it is not found significant in the analysis shown in this paper which is the only dissimilarity between these two analysis as far as marginal model is concerned.

Moreover, in the conditional model, the explanatory variables such as area, fatal severity, serious severity, light condition and year 2021 have statistically significant effects on the total number of fatalities while sex of a driver, year 2018, year 2019 and year 2020 do not have statistically significant effects on the total number of fatalities. More specifically, the effects of area, fatal severity and year 2021 on the number of fatalities are statistically significant at 0.1% level while the effects of serious severity on the number of fatalities and the effects of light condition on the number of fatalities are statistically significant at 1% and 5% levels, respectively. Like earlier, one dissimilarity has been found in both analyses: the variable sex of driver is not found to be significant in our analysis but it was significant in previous analysis as far as marginal model is concerned.

V. Conclusion

In this study, an attempt is made to explore the performance of the BVP-1, the BVP-2, and the ZTBVP models to analyze the zero-truncated bivariate pair count data based on AIC and BIC values. We use UK road safety data collected from 2017–2021 which is zero truncated pair count data. According to our data analysis, it is observed that the ZTBVP model provides the best fit for the UK road accident count data compared to all two variants of the BVP model. From the results obtained from ZTBVP model, it is also observed that sex of driver, area, serious severity and light condition are the significant covariates for the total number of cars involved in an accident while area, fatal severity, serious severity, light condition and year 2021 are the significant covariates for the total number of fatalities due to that accident.

References

1. Campbell, J. T., 1934. The Poisson correlation function. *Proceedings of the Edinburgh Mathematical Society*, **4(1)**, 18-26.
2. Chowdhury, R.I. and Islam, M.A., 2016. Zero truncated bivariate Poisson model: Marginal-conditional modeling approach with an application to traffic accident data. *Applied Mathematics*, **7(14)**, 1589.
3. Islam, M.A. and Chowdhury, R.I., 2015. A Bivariate Poisson Models with Covariate Dependence. *Bulletin of Calcutta Mathematical Society*, **107(1)**, 11-20.
4. Karlis, D. and Ntzoufras, I., 2005. Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R. *Journal of statistical Software*, **14**, 1-36.
5. Kocherlakota, S. and K., Kocherlakota. 1992. Discrete Multivariate Distributions.
6. Kuha, J., 2004. AIC and BIC: Comparisons of assumptions and performance. *Sociological methods & research*, **33(2)**, 188-229.
7. Holgate, P., 1964. Estimation for the bivariate Poisson distribution. *Biometrika*, **51 (1-2)**, 241-287.