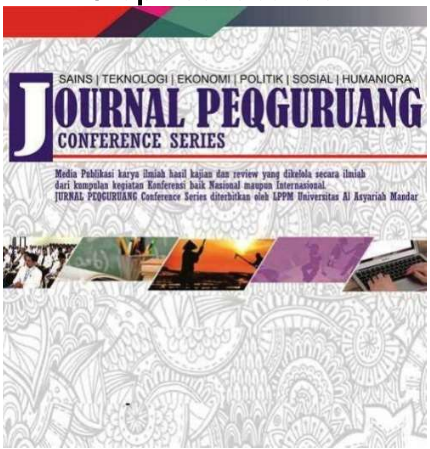
Journal
Peqguruang: Conference Series
eISSN: 2686-3472

JPCS
Vol. 2 No. 1 Mei 2020

Graphical abstract



MODEL-MODEL REGRESI UNTUK MENGATASI MASALAH OVERDIPERSI PADA REGRESI POISSON

[1*]Ayu Rahayu
[1]Universitas Al Asyariah Mandar

*Corresponding author
ayurahayu_makmur@mail.unasman.ac.id

Abstract

The Poisson Regression Model is a standard model used to analyze data that contains a discrete dependent variable (count data). In Poisson regression, there are assumptions that must be met, namely the similarity between the mean and the variance. However, in discrete data analysis using Poisson regression, overdispersion often occurs, that is, the state of the variance value is greater than the mean value. One of the causes of overdispersion is that there is an excess of zero values in the dependent variable. The existence of overdispersion in the data causes the predicted value to be incorrect so that the Poisson distribution is not suitable for use. Alternative regression models that can be used to solve the overdispersion problem are to use the Generalized Poisson (GP), Zero Inflated Poisson (ZIP), and Zero Inflated Generalized Poisson (ZIGP) regression models. By comparing the AIC, Log-likelihood, Pearson Chi Square / DB values, the ZIP model is better used to solve the overdispersion problem caused by excess zero values on the dependent variable compared to the GP and ZIGP models.

Keywords: *Overdispersion, Poisson Regression, Generalized Poisson, ZIP, ZIGP*

Abstrak

Model Regresi Poisson merupakan model standar yang digunakan untuk menganalisis data yang memuat variabel dependen berupa diskrit (count data). Pada regresi Poisson terdapat asumsi yang harus dipenuhi yaitu kesamaan antara nilai mean dan variansinya. Akan tetapi, pada analisis data diskrit yang menggunakan regresi Poisson sering terjadi overdispersi (overdispersion) yaitu keadaan nilai variansnya lebih besar dari nilai meannya. Salah satu penyebab terjadinya overdispersi adalah terdapat kelebihan nilai nol pada variabel dependennya. Adanya overdispersi dalam data menyebabkan nilai prediksi menjadi tidak tepat sehingga distribusi Poisson tidak layak digunakan. Model-model regresi alternatif yang dapat digunakan untuk mengatasi masalah overdispersi adalah dengan menggunakan model regresi *Generalized Poisson* (GP), *Zero Inflated Poisson* (ZIP), dan *Zero Inflated Generalized Poisson* (ZIGP). Dengan membandingkan nilai AIC, *Log-likelihood, Pearson Chi Square/DB* maka model ZIP lebih baik digunakan untuk mengatasi masalah overdispersi yang disebabkan oleh kelebihan nilai nol pada varaibel dependennya dibandingkan dengan model GP dan ZIGP.

Kata kunci: Overdispersi, Regresi Poisson, Generalized Poisson, ZIP, ZIGP

Article history
DOI: http://dx.doi.org/10.35329/jp.v2i1.1866
Received : *06 Januari 2020* | **Received in revised form :** *10 Februari 2020* | **Accepted :** *1 April 2020*

1. PENDAHULUAN

1.1. Latar Belakang

Dalam beberapa penelitian yang memerlukan penerapan statistika, sering dilakukan pengkajian mendalam terkait hubungan antara variabel dependen dengan variabel independen. Hubungan fungsional antara variabel dependen dan variabel independen dapat dijelaskan oleh analisis regresi. Analisis regresi merupakan analisis yang melibatkan satu variabel dependen dan satu variabel independen dengan asumsi bahwa variabel dependen berupa data kontinu yang mengikuti sebaran normal menggunakan analisis regresi linear. Sedangkan, jika melibatkan beberapa variabel independen maka digunakan analisis regresi linear berganda. Apabila variabel dependen berupa diskrit atau data cacah, maka analisis ini tidak dapat digunakan. (Zamani & Ismail, 2014) mengungkapkan bahwa jika analisis regresi linear digunakan untuk variabel dependen berupa data cacah maka akan menyebabkan hasil yang tidak efisien, tidak efektif, dan pendugaan parameter yang tidak berbias.

Metode paling efektif yang dapat digunakan untuk menganalisis hubungan variabel dependen dan variabel independen dalam bentuk data cacah adalah analisis regresi Poisson dengan variabel dependen menyebar Poisson. Pada model regresi Poisson disyaratkan keadaan yang equidispersi yaitu nilai mean (nilai rata-rata) dan variansi dari variabel dependen bernilai sama. Pada praktiknya, sering ditemui *count data* dengan varinasi lebih besar dibandingkan rata-ratannya atau bisa disebut dengan overdispersi. Salah satu penyebab terjadinya overdispersi adalah terjadi kelebihan nilai nol pada data.

(Kirtee K. Kamalja & Wagh, 2018) menyatakan bahwa ketika model Poisson diaplikasikan untuk data overidspersi, efisiensi dari estimasi parameter masih tinggi tetapi estimasi standar erronya akan tidak tepat. Sehingga probabilitas dari interval konfidensi dan tingkat signifikansinya tidak valid dan menghasilkan suatu hasil yang tidak tepat. Penanganan overdispersi pada regresi Poisson dapat menggunakan model regresi *Generalized Poisson, Zero Inflated Poisson,* dan *Zero Inflated Generalized Poisson* (Lee & Kim, 2017; Wagh & Kamalja, 2018; Zamani & Ismail, 2014).

1.2. Landasan Teori

1.2.1. Regresi Poisson

Model regresi Poisson adalah model yang paling sering digunakan untuk menganalisis count data. Karakteristik penting dari distribusi yang sering digunakan dalam pemodelan *rare event* (kasus jarang terjadi) ini yaitu mean sama dengan variansi. Kasus seperti ini biasa disebut dengan *equidispersion.* Namun, kondisi ini seperti ini sulit untuk dipenuhi. Model regresi Poisson merupakan model yang digunakan untuk memodelkan banyaknya kemunculan dari suatu kejadian dalam interval waktu tertentu (Yang et al., 2009) mengungkapkan bahwa model regresi digunakan sebagai pendekatan untuk analisis data cacah dan tergantung pada asumsi munculnya data tersebut.

Pada regresi Poisson diasumsikan bahwa variabel dependen Y yang menyatakan jumlah (cacah) kejadian berdistribusi Poisson, diberikan sejumlah variabel independen $x_1, x_2, \dots, x_k$.

$$P(Y = y|x_1, x_2, \dots, x_k) = \frac{e^{-\mu}\mu^y}{y!},\ \ y = 0,1,2, \dots \tag{1}$$

atau dengan kata lain, $Y_i \sim POI(\mu_i),\ i = 1,2, \dots, n$.

Salah satu tujuan dari analisis regresi adalah untuk menentukan pola hubungan antara variabel respon dengan variabel penjelas. Selanjutnya, dalam regresi Poisson hubungan tersebut dapat dituliskan dalam bentuk

$$E(Y_i|X_i) = \mu_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon_i \tag{2}$$

atau

$$\mu_i = x_i{}^T\beta \tag{3}$$

Karena nilai $\mu_i > 0$, maka digunakan fungsi *link* $\eta_i = \log(x_i{}^T\beta)$ atau $\eta_i = \log(\mu_i) = x_i{}^T\beta$ untuk menghubungkan $E(Y_i|X_i)$ dengan fungsi linier $x_i{}^T\beta$ sehingga hubungan antara $E(Y_i|X_i)$ dan $x_i{}^T\beta$ menjadi tepat. Dengan demikian model regresi Poisson dapat ditulis dalam bentuk

$$E(Y_i|X_i) = \mu_i = \exp(x_i{}^T\beta)\, , i = 1,2, \dots, n \tag{4}$$

dengan β merupakan parameter yang tidak diketahui dalam model dan perlu diestimasi.

1.2.2. Overdispersi

Dalam regresi Poisson terdapat asumsi yang harus dippenuhi. Asumsi tersebut adalah kesamaan nilai mean dan vairansi variabel dependen atau yang dikenal dengan sebutan equidispersi. Rahayu (2014) menyatakan bahwa overdispersi dapat terjadi karena adanya nilai nol yang berlebihan pada variabel dependenya, sumber keragaman yang tidak teramati pada data atau adanya pengaruh peubah lain yang mengakibatkan peluang suatu kejadian bergantung pada kejadian sebelumnya. Selain itu, overdispersi dapat pula terjadi karena adanya pencilan pada data dan kesalahan spesifikasi fungsi penghubung. Apabila regresi Poisson digunakan untuk kondisi overdispersi, maka terjadi keragaman data yang terdapat pada variabel dependen. Keragaman data dapat ditunjukkan dengan adanya rasio dispersi (ϕ), yaitu:

$$E(Y) = \phi\ Var\ (Y)$$

Untuk menguji asumsi equidispersi pada regresi Poisson dilakukan dengan melihat nilai statistic *Pearson's Chi Square* yang dibagi dengan derajat bebasnya (n-p-1)

$$\widehat{\emptyset} = \frac{X^2}{n-p-1} \quad \text{dengan} \quad X^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{var(\hat{\mu}_i)} \tag{5}$$

(Fitriani et al., 2019; Zamani & Ismail, 2013) mengungkapkan bahwa ada dua acara untuk melihat adanya overdispersi, yaitu:

1. Memplotkan nilai *fitted*nya terhadap niliamutlak residualnya. Jika hasil plot menunjukkan tidak adanya trend maka fungsi variansi sesuai dengan meannya. Tetapi jika plot menunjukkan adanya trend maka hal ini mngindikasikan adanya overdispersi.
2. Dilihat dari nilai *statistics deviance* atau Pearson Chi-Square yang dibagi dengan derajat bebasnya, nilai hasil bagi tersebut mendekati satu. Jika nilai hasil bagi lebih besar dari satu maka hal ini mengindikasikan adanya overdispersi pada data.

1.2.3. Model Regresi Generalized Poisson (GP)

Model *Generalized Poisson* merupakan perluasan dari model regresi Poisson. Bentuk ini umumnya digunakan untuk menjelaskan sejumlah data cacah yang memperlihatkan sifat-sifat eqiudispersi, underdispersi, maupun overdispersi Putra (2013). Andaikan variabel dependen Y_i adalah variabel random yang berdistribusi GP, fungsi kepadatan peluang dari Y_i adalah

$$P(Y_i = y_i) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{(1+\varphi y_i)^{y_i - 1}}{y_i!} \exp\left(-\frac{\mu_i(1+\varphi y_i)}{1+\varphi\mu_i}\right) \tag{6}$$

Mean diasumsikan ke dalam persamaan $E(Y_i) = \mu_i = e_i \exp(X_i^T\beta)$, sedangkan variansi bersyarat ekuivalen dengan $Var(Y_i) = \mu_i(1 + \varphi\mu_i)^2$. Jika parameter β dan φ diestimasi menggunakan *Maximum Likelihood Estimation (MLE)*, maka fungsi yang terkait adalah:

$$\frac{\partial(\beta,\varphi)}{\partial\varphi} = -\sum_{i=1}^{n} \frac{y_i \exp(X_i^T\beta)}{1+\varphi \text{ex } (X_i^T\beta)} + \sum_{i=1}^{n} \frac{y_i(y_i-1)}{1+\varphi y_i} - \sum_{i=1}^{n} \frac{\exp(X_i^T\beta)(y_i - \exp(X_i^T\beta))}{(1+\varphi\exp(X_i^T\beta))} = 0 \tag{7}$$

Dan

$$\frac{\partial(\beta,\varphi)}{\partial\beta} = \sum_{i=1}^{n} \frac{(y_i - \text{ex } (X_i^T\beta))}{(1+\varphi\exp(X_i^T\beta))^2} x_i \tag{8}$$

Persamaan (7) dan (8) memberikan penyelesaiaan yang cukup rumit, karena persamaan likelihood tersebut adalah persamaan non linear dalam parameter β dan φ. Penaksiran parameter β dapat diselesaikan dengan menggunalan metode *fisher scoring* dan φ dapat disetimasi menggunakan iterasi Newton Raphson.

1.2.4. Model Regresi Zero Inflated Poisson (ZIP)

Model regresi *Zero Inflated Poisson* merupakan model campuran yang sederhana untuk data cacah dengan banyak peristiwa nol (Rodriguez dan Maria, 2016). Jika Y_i adalah variabel random dependen yang mempunyai distribusi ZIP, maka observasi nol dapat dikembangkan dalam dua langkah. Langkah pertama disebut *zero state* terjadi dengan probabilitas w_i dan hanya menghasilkan obeservasi bernilai nol, sedangkan langkah kedua disebut *Poisson state* terjadi dengan probabilitas $1 - w_i$ dan berdistribusi Poisson dengan mean μ_i (Moghimbeigi, 2015; et al., 2018).

Dua keadaan ini memberikan dua komponen distirbusi campuran dengan fungsi probabilitas sebagai berikut.

$$P(Y_i = y_i) = \begin{cases} w_i + (1 - w_i)e^{-\mu_i}, y_i = 0 \\ \frac{(1-w_i)e^{-\mu_i}\mu_i^{y_i}}{y_i!}, y_i > 0, 0 \le w_i \le 1 \end{cases} \tag{9}$$

Yang dinotassikan dengan $Y_i \sim ZIP\ (\mu_i, w_i)$ dengan parameter $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$dan $w = (w_1, w_2, \dots, w_n)^T$ yang memenuhi

$$\log(\mu) = X\beta \text{ dan } logit\ (w) = \log\left[\frac{w}{1-w}\right] = Z\gamma \tag{10}$$

Dengan X dan Z merupakan matriks kovariat yang terdiri dari variabel-variabel penjelas yang masing-masing memepengaruhi mean Poisson. Estimasi parameter regresi ZIP dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation*. Berdasarkan persamaan (9) diperoleh :

$$\mu_i = \exp(x_i^T\beta), w_i = \frac{\exp(x_i^T\gamma),}{1 + \exp(x_i^T\gamma),}, dan$$

$$(1 - w_i) = \frac{1}{1+\exp(x_i^T\gamma),} \tag{11}$$

Persamaan log likelihood dari ZIP :

$$\log \ell\ (\beta, \alpha) = \sum_{y_i=0} \log(w_i(1 - w_i)\exp(\mu_i)) \quad + \sum_{y_i>0}\{\log(1 - w_i)\mu_i - \log(y_i!) + (y_i - 1)\log(\mu_i) - \mu_i \tag{12}$$

1.2.5. Model Regresi Zero Inflanted Generalizee Poisson (ZIGP)

Fungsi kepadatan peluang model regresi ZIGP dapat dituliskan sebagai berikut.

$$P(Y_i = y_i) = \begin{cases} w_i + (1 - w_i)e^{-\mu_i}, y_i = 0 \\ (1 - w_i)\frac{\mu_i(\mu_i + (\varphi_i - 1)y_i)^{y_i - 1}}{y_i!}\varphi_i^{-y_i}\exp(-\frac{\mu_i + (\varphi_i - 1)y_i}{\varphi_i}), y_i \neq 0 \end{cases} \tag{13}$$

Dengan, $y_i = 0,1,2,\dots$ fungsi μ_i memenuhi $\mu_i = \exp x_i^T\beta$ dimana x_i merupakan vector covariat (p x 1) dan β adalah vector (p x 1) yang tidak diketahui parameternya. Fungsi disperse dinyatakan sebagai φ_i, dimana bila $\varphi_i > 1$ dikatakan overdispersi (Fisicaro et al., 2016; Masfian et al., 2016). Fungsi log likelihood ZIGP adalah

$$\log \ell\,(\beta, \alpha) = \sum_{y_i=0} \log(w_i(1-w_i)\exp(\frac{\mu_i}{\varphi_i})) + \sum_{y_i>0}\{\log\frac{(1-w_i)\mu_i}{\varphi_i} - \log(y_i!) + (y_i-1)\log(\frac{\mu_i(\varphi_i-1)y_i}{\varphi_i}) - \frac{\mu_i(\varphi_i-1)y_i}{\varphi_i} \quad (14)$$

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data simulasi yang dibangkitkan dan dianalisis dengan menggunakan software R. Adapun langkah-langkah simulasi datanya adalah sebagai berikut:

1. Membangkitkan variabel dependen Y berdistribusi Poisson dengan n sebanyak 1000, lambda = 7 dengan peluang munculnya angka nol adalah 40%.
2. Membandingkan nilai mean dan variansi variabel dependen tersebut untuk melihat adanya overdispersi.
3. Membangkitkan variabel $X_1 \sim UNIF(1{,}3)$, $X_2 \sim N(0{,}2)$, $X_3 \sim N(0.5{,}1)$, $X_4 \sim UNIF(0{,}2)$, dan $X_5 \sim N(0.3{,}1.5)$ dengan n masing-masing 1000.
4. Meregresikan variabel Y dan variable X dengan model regresi Poisson dan menghitung nilai loglikelihood, AIC.
5. Meregresikan variabel Y dan variabel X dengan model regresi GP serta menghitung nilai loglikelihood, AIC.
6. Meregresikan variabel Y dan variabel X dengan model regresi ZIP serta menghitung nilai loglikelihood, AIC.
7. Meregresikan variabel Y dan variabel X dengan model regresi ZIGP serta menghitung nilai loglikelihood, AIC.
8. Memilih model terbaik berdasarkan kriteria pemilihan model.

3. HASIL DAN PEMBAHASAN

Deskripsi data variabel independen X_1, X_2, …, X_5 serta variabel dependen Y yang digunakan dalam penelitian ini dapat dilihat ada tabel berikut.

Tabel 1. Deskripsi data

Variabel	N	Mean	Variansi	Minimum	Maksimum
X_1	1000	1.982	0.3415	1	2.998
X_2	1000	0.0674	3.5743	-5.6550	6.174
X_3	1000	0.4571	0.9833	-2.6650	3.9
X_4	1000	0.983	0.3370	0.00197	1.999
X_5	1000	0.290	2.157	3.896	5.334
Y	1000	1.461	3.023	0	9

Berdasarkan tabel di atas dapat dilihat bahwa nilai variansi variabel dependen lebih besar dibandingkan nilai meannya. Hal ini mengindikasikan adanya overdispersi data. Hasil analisis regresi Poisson untuk data tersebut diperoleh model sebagai berikut.

$$\mu = \exp(0.3356 + 0.02X_1 - 0.009X_2 - 0.002X_3 - 0.014X_3 - 0.054X_5) \quad (15)$$

Karena data tersebut mengalami overdispersi berarti penggunaan model regresi Poisson kurang tepat untuk menganalisis data tersebut, untuk itu dilakukan analisis model regresi *Generalized Poisson, Zero Inflated Poisson, dan Zero Inflated Generalized Poisson* sebagai model regresi alternatif. Nilai estimasi parameter dari masing-masing model disajikan pada tabel berikut.

Tabel 2. Estimasi masing-masing parameter regresi

	GP	ZIP	ZIGP
Intercept	0.3680	0.8753	0.8732
X_1	-0.0017	0.0294	0.0288
X_2	-0.0148	-0.0052	-0.0061
X_3	0.0008	0.0016	0.0019
X_4	-0.0067	-0.0205	-0.0204
X_5	0.0599	0.0290	0.0296

Berdasarkan hasil yang diperoleh pada tabel 2 maka dapat dituliskan masing-masing model alternative sebagai berikut.

Model regresi *Generalized Poisson*

$$\mu = \exp(0.3680 - 0.0017X_1 - 0.0148X_2 + 0.0008X_3 - 0.0067X_3 + 0.0599X_5) \quad (16)$$

Model regresi *Zero Inflated Poisson*

$$\mu = \exp(0.8753 + 0.0294X_1 - 0.0052X_2 + 0.0016X_3 - 0.0205X_3 + 0.0290X_5) \quad (17)$$

Model regresi *Zero Inflated Generalized Poisson*

$$\mu = \exp(0.8732 + 0.0288X_1 - 0.0061X_2 - 0.0019X_3 - 0.0204X_3 + 0.0296X_5) \quad (18)$$

Intrepretasi model yang terbentuk pada model regresi GP adalah apabila terjadi penigkatan sebesar satu satuan pada X_5 maka Y akan meningkat sebesar $100(e^{0.0599(1)} - 1)\% = 100(1.0071 - 1)\% = 6.17\%$.

Pada model regresi ZIP, apabila terjadi peningkatan sebesar satu satuan pada X_1 maka variabel Y akan meningkat sebesar $100(e^{0.0294(1)} - 1)\% = 100(1.0294 - 1)\% = 2.98\%$.

Sedangkan pada model regresi ZIGP, apabila terjadi penigkatan sebesar satu satuan pada X_1 maka Y akan meningkat sebesar $100(e^{0.0288(1)} - 1)\% = 100(1.0292 - 1)\% = 2.92\%$.

Tahapan selanjutnya adalah melakukan pengujian model terbaik untuk memabndingkan metode analisis yang lebih baik digunakan untuk mengatasi overdispersi. Untuk melihat model yang terbaik dalam mengatasi overdispersi pada regresi Poisson, akan digunakan salah satu kriteria pemilihan model terbaik yaitu dengan menggunakan AIC. Hasil uji model terbaik disajikan pada tabel berikut.

Tabel 3. Uji Model terbaik

Kriteria	Poisson	GP	ZIP	ZIGP
AIC	3726.2	3355	3195	3197
Pearson Chi Square/DB	2.072	2.077	2.072	2.072
Log-likelihood	-1857.09	-1670.4	-1590	-1590

Tabel 3 memperlihatkan bahwa dari kriteria model terbaik berdasarkan nilai AIC maka model regresi *Zero inflated Poisson* memiliki nilai AIC terkecil. Sedangkan berdasarkan nilai *log-Likelihood* model regresi ZIP dan ZIGP memiliki nilai yang lebih besar dari model regresi Poisson dan GP. Sedangkan berdasarkan nilai *Pearson Chi Square/DB* model Poisson, ZIP, dan ZIGP memiliki nilai yang sama dan merupakan nilai yang paling mendekati 1. Dengan demikian model ZIP lebih baik digunakan dalam mengatasi masalah overdispersi yang disebabkan oleh kelebihan nilai nol pada variabel dependen dibandingkan dengan model regresi Poisson, *Generalized Poisson,* dan *Zero Inflated Generalized Poisson.*

4. SIMPULAN

1. Salah satu penyebab terjadinya overdispersi pada regresi Poisson adalah karena kelebihan nilai nol pada variabel dependen.
2. Model regresi *Generalized Poisson, Zero Inflated Poisson, Zero Inflated Generalized Poisson* merupakan model-model regresi alternatif yang dapat digunakan untuk mengatasi overdispersi pada regresi Poisson yang disebabkan kelebihan nilai nol pada variabel dependennya.
3. Model regresi *Zero Inflated Poisson* merupakan model yang lebih baik digunakan untuk mengatasi masalah overdispersi yang disebabkan oleh kelebihan nilai nol pada variabel dependennya dibandingkan model regresi Poisson, GP, dan ZIGP dilihat dari nilai AIC, *Log-Likelihood, Pearson Chi-Square/DB.*

DAFTAR PUSTAKA

Farhadi Hassankiadeh, R., Kazemnejad, A., Gholami Fesharaki, M., & Kargar Jahromi, S. (2018). Efficiency of Zero-Inflated Generalized Poisson Regression Model on Hospital Length of Stay Using Real Data and Simulation Study. *Caspian Journal of Health Research.* https://doi.org/10.29252/cjhr.3.1.5

Fisicaro, G., Genovese, L., Andreussi, O., Marzari, N., & Goedecker, S. (2016). A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments. *Journal of Chemical Physics.* https://doi.org/10.1063/1.4939125

Fitriani, R., Chrisdiana, L. N., & Efendi, A. (2019). Simulation on the Zero Inflated Negative Binomial (ZINB) to Model Overdispersed, Poisson Distributed Data. *IOP Conference Series: Materials Science and Engineering.* https://doi.org/10.1088/1757-899X/546/5/052025

Kirtee K. Kamalja, & Wagh, Y. S. (2018). Estimation in zero-inflated Generalized Poisson distribution. *Journal of Data Science.*

Lee, C. E., & Kim, S. U. (2017). Applicability of zero-inflated models to fit the torrential rainfall count data with extra zeros in South Korea. *Water (Switzerland).* https://doi.org/10.3390/w9020123

Masfian, I., Yuniarti, D., & Hayati, M. N. (2016). Penerapan Generalized Poisson Regression I Untuk Mengatasi Overdispersi Pada Regresi Poisson (Studi Kasus: Pemodelan Jumlah Kasus Kanker Serviks di Provinsi Kalimantan Timur). *Jurnal Eksponensial.*

Moghimbeigi, A. (2015). Two-part zero-inflated negative binomial regression model for quantitative trait loci mapping with count trait. *Journal of Theoretical Biology.* https://doi.org/10.1016/j.jtbi.2015.02.016

Wagh, Y. S., & Kamalja, K. K. (2018). Zero-inflated models and estimation in zero-inflated Poisson distribution. *Communications in Statistics: Simulation and Computation.* https://doi.org/10.1080/03610918.2017.1341526

Yang, Z., Hardin, J. W., & Addy, C. L. (2009). Testing overdispersion in the zero-inflated Poisson model. *Journal of Statistical Planning and Inference.* https://doi.org/10.1016/j.jspi.2009.03.016

Zamani, H., & Ismail, N. (2013). Score test for testing zero-inflated Poisson regression against zero-inflated generalized Poisson alternatives. *Journal of Applied Statistics.* https://doi.org/10.1080/02664763.2013.804904

Zamani, H., & Ismail, N. (2014). Functional form for the zero-inflated generalized poisson regression model. *Communications in Statistics - Theory and Methods.* https://doi.org/10.1080/03610926.2012.665553