## Daftar Pustaka


Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley and Sons, Inc.

Astuti, C. Cindy. 2015. *Pemodelan Regresi Zero Inflated Negative Binomial (ZINB) pada Kasus Tetanus Neonatorium di Provinsi Jawa Timur*. ISBN : 978.602.361.002.0

Depkes. 2018. *Profil Kesehatan Indonesia*. Jakarta : Kementrian Kesehatan Republik Indonesia

Depkes. 2019. *Profil Kesehatan Indonesia*. Jakarta : Kementrian Kesehatan Republik Indonesia

Famalasari, Kiki. 2019. *Gambaran Kasus Difteri Tahun 2009-Agustus 2019 di Kabupaten Bojonegoro. Media Gizi & Kesehatan Masyarakat*. Vol. 8, No. 2. Hal. 67-76.

Famoye, F., & Singh, K. P. (2006). Zero Inflated *Poisson* Regression Model with an Applications Domestic Violence to Accident Data. Journal of Data Science, 117-130.

Gujarati, D. (1991). Ekonometrika Dasar (S. Zain, Trans.). Jakarta: Penerbit Erlangga.

Huwae IR, Shodikin MA. Clinical appearance of diphtheria in children in Saiful Anwar General Hospital, Malang, East Java, Indonesia. *International Journal of Infectious Diseases.* 2012;16(Supplement 1):e261

Hidayati, Ridha. 2017. *Faktor-Faktor yang Mempengaruhi Angka Kejadian Penyakit Difteri di Kota Padang*. Social and Economics Research. Vol. 02, No. 02. Hal. 180-187.

Hinde, J., & Demetrio, C. G. B. (2007). *Overdispersi: Model and Estimation*. Exeter: Deparment of Laver Building.

Hilbe, J. M. (2011). Negative Binomial Regression. New York: Cambridge University Press.

Hocking, R. (1996). Methods and Application of Linier Models. New York: John Wiley and Sons.

Hogg, R. V., McKean, J. W., & Craig, A. T. 2013. Introduction to Mathematical Statistics, Seventh Edition. United States of America: Pearson Education, Inc.

Izza, N., dan Soenarnatalia. 2015. *Analisis Data Spasial Penyakit Difteri di Provinsi Jawa Timur Tahun 2010 dan 2011*. Bulletin Penelitian Sistem Kesehatan.

Kementerian Kesehatan Republik Indonesia. 2018. *Pedoman Pencegahan dan Pengendalian Difteri*.

Kurniasih, Rini. 2015. *Model Klaim Agregasi dengan Besar Klaim Berdistribusi Erlang dan Jumlah Klaim Berdistribusi COM-Poisson*. Academia Accelerating The World's Research.

Lord, D., B.-J. Park, and N. Levine. 2013. *CrimeStat IV: Chapter 16 – Poisson Regression Modeling (Part V: Spatial Modeling II)*. Ned Levine & Associates and the National Institute of Justice, Washington,D.C. https://nij.ojp.gov/sites/g/files/xyckuh171/files/media/document/CrimeStat%2520IV%2520Chapter%252016.pdf

MacGill, Markus. 2018. *Everything You Need to Know About Diphteria*. Medical News Today.

McCullagh, P. and Nelder, J.A. 1989. Generalized Linear Models. 2nd Edition, Chapman and Hall, London. http://dx.doi.org/10.1007/978-1-4899-3242-6

Mustika, Diva A., Nooraeni, Rani. 2019. *Kajian Efek Spasial Pada Kasus Difteri Dengan Metode Geographically Weighted Negative Binomial Regression (Gwnbr)*. Indonesian Journal of Statistics and Its Applications, Vol 3 No 1 (2019), 91 – 104.

Myers, R.H., Douglas C. Montgomery, G. Geoffrey Vining, & Timothy J, Robinson. 2010. *Generalized Linear Models with Applications in Engineering and The Sciences*. Second edition. New Jersey: John Wiley and Sons

Nadhiroh, I. M. (2009). *Zero Inflated Negative Binomial Models in Small Area Estimation*. Bogor: IPB

Pusat Data dan Informasi. 2021. *Profil Kesehatan Indonesia 2020*.

Pradana, Dimas A., Lestari, Trianingsih. 2020. *Estimasi Parameter Regresi Zero-Inflated Negative Binomial Dengan Metode Algoritma Expectation Maximization (Em) (Studi Kasus: Penyakit Difteri Di Jawa Barat Tahun 2016)*. Jurnal Kajian Matematika dan Aplikasinya, Vol. 1 No. 1, e-ISSN: 2722-7650. **http://journal2.um.ac.id/index.php/jkma**

Pratama, Yovi. 2017. *Analisis Statistik Pemilihan Bidang Skripsi Menggunakan Ekspectation Maximization.* Jurnal Mediasisfo. Vol.11, No.1.

Rahayu, Rini L., Asril, Asnafiya. 2018. *Perbandingan Regresi Zero Inflated Negatif Binomial dan Regresi Hurdle Negatif Binomial pada Data Overdispersi (Studi Kasus: Kejadian Difteri di Indonesia)*. Jurnal Sains Matematika dan Statistika, Vol. 4, No. 1, ISSN 2460-4542.

Saifudin, Nanang. 2016. *Faktor Resiko Kejadian Difteri di Kabupaten Blitar Tahun 2015*. Jurnal Wiyata. Vol 3, No. 1. Hal. 61-66.

Sellers, Kimberly F., Raim, Andrew. 2016. *A flexible zero-inflated model to address data dispersion*. Computational Statistics and Data Analysis 99 (2016) 68–80.

Sim, Shin Z., Gupta, Ramesh C., Ong, Seng H. 2018. *Zero-inflated Conway-Maxwell Poisson Distribution to Analyze Discrete Data*. The International Journal of Biostatistics. 2018; 20160070

Tosepu, R., Gunawan, J., Effendy, D. S., & Farzan, A. (2018). The outbreak of diphtheria in Indonesia. *Pan African Medical Journal*, *31*(1).

WHO. 2017. *Review of the epidemiology of diphtheria 2000-2016*. World Health Organization. Geneva.

World Health Organization. 2018. *Diphteria*.

Lampiran 1. Data Penelitian Jumlah Kasus Mortalita Pasien Difteri di Indonesia Tahun 2019

| No | Provinsi | Y | X1 | X2 | X3 | X4 | X5 | X6 |
|---|---|---|---|---|---|---|---|---|
| 1 | Aceh | 0 | 15.32 | 6.3 | 70.7 | 10.08 | 69.76 | 75 |
| 2 | Sumatera Utara | 3 | 8.83 | 5.8 | 73.86 | 10.24 | 58.66 | 199 |
| 3 | Sumatera Barat | 0 | 6.42 | 3.9 | 73.09 | 37 | 71.59 | 81 |
| 4 | Riau | 0 | 7.08 | 4.4 | 69.56 | 43.24 | 62 | 78 |
| 5 | Jambi | 0 | 7.6 | 4.9 | 62.94 | 40.25 | 67.18 | 40 |
| 6 | Sumatera Selatan | 0 | 12.71 | 3.9 | 63.26 | 37.94 | 69.68 | 73 |
| 7 | Bengkulu | 0 | 15.23 | 2.4 | 49.52 | 35.05 | 66.63 | 29 |
| 8 | Lampung | 0 | 12.62 | 2.7 | 59.65 | 31.6 | 70.58 | 99 |
| 9 | Kep. Bangka Belitung | 1 | 4.62 | 4.6 | 84.09 | 59.37 | 80.21 | 23 |
| 10 | Kepulauan Riau | 0 | 5.9 | 2.7 | 91.53 | 37.73 | 57.2 | 29 |
| 11 | DKI Jakarta | 1 | 3.47 | 1.5 | 88.44 | 61.25 | 64.55 | 191 |
| 12 | Jawa Barat | 10 | 6.91 | 2.5 | 72.38 | 28.84 | 65.44 | 330 |
| 13 | Jawa Tengah | 0 | 10.8 | 3.5 | 78.86 | 44.69 | 83.15 | 295 |
| 14 | D I Yogyakarta | 0 | 11.7 | 1.4 | 75.3 | 66.21 | 71.95 | 75 |
| 15 | Jawa Timur | 1 | 10.37 | 3.6 | 76.1 | 44.5 | 37.18 | 372 |
| 16 | Banten | 2 | 5.09 | 3 | 73.68 | 24.79 | 46.88 | 121 |
| :<br>: | | :<br>: | :<br>: | :<br>: | :<br>: | :<br>: | :<br>: | :<br>: |
| 21 | Kalimantan Tengah | 0 | 4.98 | 4.7 | 69.18 | 42.97 | 67.27 | 30 |
| 22 | Kalimantan Selatan | 0 | 4.55 | 3.9 | 72.05 | 34.64 | 69.07 | 44 |
| 23 | Kalimantan Timur | 0 | 5.94 | 3.4 | 93.77 | 47.79 | 43.46 | 53 |
| 24 | Kalimantan Utara | 0 | 6.63 | 1.7 | 80.58 | 57.47 | 73.17 | 8 |
| 25 | Sulawesi Utara | 0 | 7.66 | 6.3 | 81.72 | 17.2 | 19.47 | 47 |
| 26 | Sulawesi Tengah | 0 | 13.48 | 3.8 | 74.67 | 50.84 | 58.41 | 38 |
| 27 | Sulawesi Selatan | 0 | 8.69 | 4.8 | 79.32 | 36.63 | 45.26 | 111 |
| 28 | Sulawesi Tenggara | 2 | 11.24 | 6.3 | 79.32 | 39.39 | 39.68 | 36 |
| 29 | Gorontalo | 0 | 15.52 | 8.1 | 82.33 | 26.36 | 45.72 | 14 |
| 30 | Sulawesi Barat | 0 | 11.02 | 5.7 | 63.55 | 38.32 | 45.28 | 12 |
| 31 | Maluku | 0 | 17.69 | 10.3 | 73.17 | 45.01 | 48.15 | 30 |
| 32 | Maluku Utara | 0 | 6.77 | 6.5 | 64.82 | 44.67 | 78.7 | 21 |
| 33 | Papua Barat | 1 | 22.17 | 4.1 | 70.18 | 42.55 | 55.97 | 19 |
| 34 | Papua | 0 | 27.53 | 4.5 | 39.05 | 20 | 34.12 | 44 |

Lampiran 2. Turunan Pertama dan kedua

- Fungsi $Z_i$

$Z_i = Z(\lambda_i, \phi) = \sum_{j=0}^{\infty} \frac{\exp(\boldsymbol{X_i^T \beta})^j}{(j!)^{\phi}}$

Turun pertama Fungsi $Z_i$ terhadap $\phi$

$$
\begin{aligned}
\frac{\partial Z_i}{\partial \phi} &= \frac{\partial Z_i}{\partial \phi}\left(\sum_{j=0}^{\infty} \frac{e^{(\boldsymbol{X_i^T \beta})^j}}{(j!)^{\phi}}\right) \\
&= \frac{\partial Z_i}{\partial \phi}\left(\sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} (j!)^{-\phi}\right) \\
&= \sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} \left(\frac{\partial Z_i}{\partial \phi}\left(\sum_{j=0}^{\infty} (j!)^{-\phi}\right)\right) \\
&= \sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} \left(\frac{\partial Z_i}{\partial \phi} e^{(-\phi)\ln(j!)}\right) \\
&= \sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} \left(e^{(-\phi)\ln(j!)} \left(\frac{\partial Z_i}{\partial \phi}\left((-\phi)\ln(j!)\right)\right)\right) \\
&= \sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} \left(e^{(-\phi)\ln(j!)} \left(\frac{\partial Z_i}{\partial \phi}\left((-\phi)\ln(j!)\right)\right)\right) \\
&= \sum_{j=0}^{\infty} e^{(\boldsymbol{X_i^T \beta})^j} (j!)^{-\phi} (\ln(j!)) \\
&= \sum_{j=0}^{\infty} \frac{-\exp(\boldsymbol{X_i^T \beta})^j \ln(j!)}{(j!)^{\phi}}
\end{aligned}
$$

Turunan Kedua Fungsi $Z_i$ terhadap $\phi$

$$
\frac{\partial^2 Z_i}{\partial \phi^2} = \frac{\partial Z_i}{\partial \phi}\left(\sum_{j=0}^{\infty} \frac{-\exp(\boldsymbol{X_i^T \beta})^j \ln(j!)}{(j!)^{\phi}}\right)
$$

$$= \sum_{j=0}^{\infty} -\mathrm{e}^{\left(\boldsymbol{X}_i^T \boldsymbol{\beta}\right)^j} \ln(j!) \left( \frac{\partial Z_i}{\partial \phi} \left( (j!)^{-\phi} \right) \right)$$

$$= \sum_{j=0}^{\infty} -\mathrm{e}^{\left(\boldsymbol{X}_i^T \boldsymbol{\beta}\right)^j} \ln(j!) \left( \frac{\partial Z_i}{\partial \phi} \left( e^{(-\phi)\ln(j!)} \right) \right)$$

$$= \sum_{j=0}^{\infty} -\mathrm{e}^{\left(\boldsymbol{X}_i^T \boldsymbol{\beta}\right)^j} \ln(j!) \left( e^{(-\phi)\ln(j!)} \right) \left( \frac{\partial Z_i}{\partial \phi} \left( (-\phi)\ln(j!) \right) \right)$$

$$= \sum_{j=0}^{\infty} -\mathrm{e}^{\left(\boldsymbol{X}_i^T \boldsymbol{\beta}\right)^j} \ln(j!) \left( (j!)^{-\phi} \right) (-\ln(j!))$$

$$= \sum_{j=2}^{\infty} \frac{\exp\left(\boldsymbol{X}_i^T \boldsymbol{\beta}\right)^j \ln(j!)^2}{(j!)^{\phi}}$$

Lampiran 3A. Pengujian Distribusi Poisson

**One-Sample Kolmogorov-Smirnov Test**

| | | Y |
|---|---|---|
| N | | 34 |
| Poisson Parameter[a,b] | Mean | .6765 |
| Most Extreme Differences | Absolute | .227 |
| | Positive | .227 |
| | Negative | -.029 |
| Kolmogorov-Smirnov Z | | 1.323 |
| Asymp. Sig. (2-tailed) | | .060 |

a. Test distribution is Poisson.

b. Calculated from data.

Lampiran 3B. Output SAS Overdispersi

**Model Information**

| | |
|---|---|
| **Data Set** | WORK.DIFTERI |
| **Distribution** | Poisson |
| **Link Function** | Log |
| **Dependent Variable** | Y |
| **Observations Used** | 34 |

**Criteria For Assessing Goodness Of Fit**

| Criterion | DF | Value | Value/DF |
|---|---|---|---|
| **Deviance** | 27 | 33.9971 | 1.2592 |
| **Scaled Deviance** | 27 | 27.0000 | 1.0000 |
| **Pearson Chi-Square** | 27 | 126.5310 | 4.6863 |
| **Scaled Pearson X2** | 27 | 100.4892 | 3.7218 |
| **Log Likelihood** | | -7.5590 | |
| **Full Log Likelihood** | | -28.4936 | |
| **AIC (smaller is better)** | | 70.9871 | |
| **AICC (smaller is better)** | | 75.2948 | |
| **BIC (smaller is better)** | | 81.6717 | |

**Lampiran 4. Multikolinieritas**

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Unstandardized Coefficients Std. Error | Standardized Coefficients Beta | t | Sig. | Collinearity Statistics Tolerance | Collinearity Statistics VIF |
|---|---|---|---|---|---|---|---|---|
| 1 | (Constant) | 2.536 | 4.107 | | .618 | .542 | | |
| | X1 | -.059 | .063 | -.184 | -.940 | .356 | .633 | 1.581 |
| | X2 | -.049 | .186 | -.052 | -.261 | .796 | .607 | 1.647 |
| | X3 | -.023 | .028 | -.161 | -.831 | .414 | .645 | 1.551 |
| | X4 | -.018 | .027 | -.135 | -.677 | .505 | .612 | 1.633 |
| | X5 | -.006 | .023 | -.051 | -.268 | .791 | .678 | 1.476 |
| | X6 | .009 | .004 | .468 | 2.585 | .016 | .740 | 1.351 |

a. Dependent Variable: Y

**Lampiran 5.** Hasil Estimasi Parameter Model Regresi NB Menggunakan *Software* SAS

```
Data Zicmp;
Input y X1 X2 X3 X4 X5 X6;
Datalines;
Run;
proc nlmixed data= Zicmp tech= newrap;
parameters b0=0 b1=0 b2=0 b3=0 b4=0 b5=0 b6=0 nu=0;
bpart=b0+b1*X1+b2*X2+b3*X3+b4*X4+b5*X5+b6*X6;
lambda=exp(bpart);
Z = 1;
DO s = 1 to 100;
Q = 1;
DO r = 1 to s;
Q = Q * (lambda / (r**nu));
END;
Z = Z + Q;
END;
ll = (Y*log(lambda))-(nu*lgamma(Y+1))-(log(Z));
model Y ~ general(ll);
predict _ll out=LL_1;
run;
```

**Lampiran** Hasil Estimasi Parameter Model Regresi NB Menggunakan *Software* SAS **(lanjutan)**

The NLMIXED Procedure

**Specifications**

| | |
|---|---|
| **Data Set** | WORK.ZICMP |
| **Dependent Variable** | Y |
| **Distribution for Dependent Variable** | General |
| **Optimization Technique** | Newton-Raphson |

**Fit Statistics**

| | |
|---|---|
| **-2 Log Likelihood** | 59.7 |
| **AIC (smaller is better)** | 87.6 |
| **AICC (smaller is better)** | 114.5 |
| **BIC (smaller is better)** | 105.9 |

**Parameter Estimates**

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | 95% Confidence Limits | | Gradient |
|---|---|---|---|---|---|---|---|---|
| b0 | 2.3431 | 2.7724 | 34 | 0.85 | 0.4039 | -3.2911 | 7.9774 | 3.697E-6 |
| b1 | -0.1358 | 0.08162 | 34 | -1.66 | 0.1053 | -0.3017 | 0.03005 | 0.000028 |
| b2 | 0.04772 | 0.1848 | 34 | 0.26 | 0.7978 | -0.3279 | 0.4233 | 0.000011 |
| b3 | -0.02994 | 0.02136 | 34 | -1.40 | 0.1702 | -0.07336 | 0.01348 | 0.000268 |
| b4 | -0.02229 | 0.01795 | 34 | -1.24 | 0.2228 | -0.05877 | 0.01419 | 0.000111 |
| b5 | -0.00209 | 0.01515 | 34 | -0.14 | 0.8912 | -0.03288 | 0.02870 | 0.000228 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **b6** | 0.006043 | 0.002539 | 34 | 2.38 | 0.0231 | 0.000882 | 0.01120 | 0.001174 |
| **Nu** | 0.2916 | 0.3531 | 34 | 0.83 | 0.4145 | -0.4259 | 1.0091 | -9.66E-6 |

**Lampiran 6.** Sintaks MATLAB untuk Estimasi Parameter

```
%% Ambil dataset
dataset = xlsread('F:\DIFTERI1.xlsx');
[n0,p0] = size(dataset);
datafull = dataset(:,1:6);
%% Definisi X dan Y
Y = datafull(:,1);
X = datafull(:,2:6);
[n,p] = size(X);
%% Matrix X baru
Xbaru = [ones(n,1),X];
X1 = Xbaru(:,1);
[nb,kb] = size(Xbaru);
%% V, Beta dan Gamma
B= (Xbaru'*Xbaru)\(Xbaru'*Y);
GM = (Xbaru'*Xbaru)\(Xbaru'*Y);
T0 = [v,B',GM'];
%% iterasi & selisi
iterasi1 = 0;
iterasi2 = 0;
selisih1 = 0.1;
selisih2 = 0.1;

while selisih1 >= 0.0001
while selisih2 >= 0.0001
    XB = Xbaru*B;
    XG = Xbaru*GM;
    LM= exp(XB);

    %% E-Step
    for j=1:kb
        for i= 1:n
            Zi(i)= LM(i)^i/(factorial(i))^v;
            % Turunan pertama
            dldv(j)=-log(factorial(Y(j)))-(dZdv(j)/Zi(i));
            % Turunan pertama B
            dZdB(i,j)= j*Xbaru(i,j)*LM(i).^j/(factorial (j))^v;
            dldB(i,j)= Y(i)*Xbaru(i,j)-(dZdB(i,j))/Zi(i);
        end
        for ii= 1:n
            L(ii) = XG(ii)+LM(ii);
            z(ii) = exp(XG(ii))/(1+exp(-L(ii)));
            dBB(ii,j) = (z(ii)-1)*dldB(ii,j);
            dVV(j)= (z(ii)-1)*dldv(j);
            dGG(ii,j)= z(ii)*Xbaru(ii,j);
        end
        for iii=1:n
        dG(iii,j) = -Xbaru(iii,j)*exp(XG(iii))/(1+exp(XG(iii)));
        end
    d1v= sum(dVV);
    d1B(j)= sum(dBB(:,j));
    d1G(j)= sum(dGG(:,j)) + sum(dG(:,j));
```

```
    end
for k=1:kb
    for i=1:kb
        for jj=2:kb
          Zi(i) = (LM(i)^i)/(factorial(i))^v;
          % Turunan kedua v^2
          d2ldv(k)= (((dZdv(k))^2)/(Zi(i)^2))-((d2ZdV(k))/Zi(i));
          % Turunan pertama B^2
          d2ldB(i,k)= ((dZdB(i,k)^2)/Zi(i))-d2ZdB(k)/Zi(i);
          %turunan kedua v dan B
          d2VB(i,k) = ((dZdv(k)*dZdB(i,k))/(Zi(i)^2))-(d2ZvB(i,k)/Zi(i));
        end
        for ii=1:n
            L(ii)=XG(ii)+LM(ii);
            z(ii) = exp(XG(ii))/(1+exp(-L(ii)));
            d2BB(ii,k) =(z(ii)-1)*d2ldB(i,k);
            d2VV(ii,k)=(z(ii)-1)*d2ldv(k);
            d2VBB(ii,k)=(z(ii)-1)*d2VB(k);
        end
        for iii= 1:n
            EG(iii) = exp(XG(iii))/(1+exp(XG(iii)));
            d2G(iii,j,k) = Xbaru(iii,j)*Xbaru(iii,k)*(EG(iii)^2- EG(iii));
        end
                d2LV= sum(d2ldv);
                d2LVB(j,k) = sum(d2VB(:,k));
                d2LBB(j,k) = sum(d2BB(:,k));
                d2LGG(j,k) = sum(d2G(:,j,k));
        end
end
    T = [v,B'];
    G = [d1v;d1B'];
    H1= [d2LV,d2LVB(6,:)];
    H2= [d2LVB(6,:)',d2ldB];
    H = [H1;H2];

       %% M-step
   Tlama= T';
   iterasi1=iterasi1+1;
   c =inv(H)*G;
   Theta= Tlama-c;
   Selisih1 =norm(Theta-Tlama);
   disp([iterasi1' Theta' Selisih1'])
   Glama = GM;
   iterasi2 = iterasi2 + 1;
   d =inv(d2LGG)*d1G';
   Gamma = Glama - d;
   Selisih2 = norm(Gamma - Glama);
   disp([iterasi2' Gamma' Selisih2'])

   LM = Theta(1,:);
   B = Theta(1,:);
   GM = Gamma;
 end
```

```
end

%% Uji Parsial Parameter Model
VB= -(inv(H));
VGM = -(inv(d2LGG));
% Uji Parsial Parameter Beta
for w=1:kb
    SEB(w) = sqrt(VB(j,j));
    WB(w) = (Gamma(w,:)/SEB(:,w))^2;
end
% Uji Parsial Parameter Gamma
for j=1:kb
    SEGM(j) = sqrt(VGM(j,j));
     WGM(j) = (Gamma(j,:)/SEGM(:,j))^2;
end
```