

MENGATASI OVERDISPERSI DENGAN REGRESI BINOMIAL NEGATIF PADA ANGKA KEMATIAN IBU DI KOTA BANDUNG

Hilma Mutiara Winata^{1*}

¹ UIN Sunan Gunung Djati Bandung

*e-mail : hilmamutiawinata@uinsgd.ac.id

DOI: 10.14710/j.gauss.11.4.616-622

Article Info:

Received: 2023-01-23

Accepted: 2023-02-20

Available Online: 2023-02-25

Keywords:

Poisson regression; Negative

Binomial Regression;

Overdispersion.

Abstract: The maternal mortality rate in the city of Bandung is still a concern for the government, even though various health programs have been held to handle it. The very slight reduction in maternal mortality is a reason for further research to look for factors that have a significant effect. The data on maternal mortality cases usually contain a lot of zeros and follow the Poisson distribution so that they are solved with a Poisson regression model, however the model formed cannot be used because the model shows overdispersion with a deviation value of more than one. Therefore, to overcome this problem, negative binomial regression is used as a solution. This negative binomial regression model produces three predictor variables out of seven variables that have a significant effect on maternal mortality in the city of Bandung including pregnant women receiving FE1 (30 tablets), deliveries assisted by health personnel and postpartum service coverage. Then tested the goodness of the model from the negative binomial regression model by looking at the AIC value. The true negative binomial regression model is better because the AIC value is 109.4 which is smaller than 121.65 which is the AIC value of the Poisson regression model.

1. PENDAHULUAN

Angka kematian ibu masih menjadi masalah di Kota Bandung. Berdasarkan data dari Dinas Kesehatan Kota Bandung di tahun 2019 terdapat 29 kasus kematian ibu dan ditahun 2020 hanya turun 1 yaitu menjadi 28 kasus (infopublik.id, 2021). Kematian ibu sendiri merupakan kasus kematian perempuan akibat dari kehamilan, persalinan, abortus, dan masa nifas yaitu waktu 42 hari setelah melahirkan, tidak termasuk usia gestasi serta kematian akibat kecelakaan atau kejadian *incidental* (Tim Penulis Kemenkes RI, 2020). Pemerintah Indonesia sudah berupaya untuk menurunkan angka kematian ibu dengan mengadakan berbagai program kesehatan yaitu program Kesehatan Ibu dan Anak (KIA). Namun masih saja ada permasalahan yang mendasar dan mengganggu keberhasilan program ini. Untuk mengatasi hal tersebut, maka dicari gambaran terkait dengan faktor apa saja yang dapat mempengaruhi kematian ibu dilihat dari program yang sudah dibuat sehingga bisa jadi hal yang dapat membantu pemerintah dalam menentukan faktor mana yang harus di prioritaskan untuk lebih ditingkatkan, sehingga dapat melebihi menurunkan atau bahkan menghapus kejadian seperti itu.

Salah satu cara dalam mencari faktor yang mempengaruhi kematian ibu bisa dengan menerapkan analisis statistika. Karena statistika merupakan alat untuk mengumpulkan, menyajikan, menganalisis data serta mengambil kesimpulan yang bersifat objektif berdasarkan data. Salah satu manfaat dari analisis statistika adalah untuk melihat pengaruh suatu peubah prediktor terhadap peubah respon. Analisis tersebut dinamakan analisis regresi. Data yang digunakan dalam analisis ini biasanya adalah data kontinu. Namun dalam beberapa kasus ada juga yang berupa data diskrit terutama pada variable responnya, dimana

data diskritnya berupa data count yaitu data yang nilainya nonnegatif serta menerangkan banyaknya kejadian dalam selang waktu, ruang atau volume tertentu. Data seperti ini adalah data yang biasanya ada pada kasus angka kematian ibu. Kasus seperti itu dapat diselesaikan dengan analisis regresi poisson.

Analisis regresi poisson yaitu metode yang digunakan untuk melihat hubungan antara peubah respon yang berdistribusi poisson dengan beberapa peubah prediktor. Terdapat asumsi yang harus dipenuhi pada metode ini yaitu nilai variansi datanya sama dengan nilai reratanya (ekuidispersi). Tetapi nyatanya itu adalah hal yang jarang terjadi karena kecenderungannya data count memiliki nilai variansi lebih besar dari reratanya (overdispersi) atau sebaliknya (Utami, 2013). Oleh karena itu regresi binomial negatif bisa menjadi salah satu alternatif dalam menyelesaikan masalah tersebut (Cameron & Trivedi, 2013). Regresi binomial negatif bisa mengatasi overdispersi karena asumsi rerata dan variansinya tidak harus sama. maka dari itu dirasa lebih fleksibel dari regresi poisson.

2. TINJAUAN PUSTAKA

Tinjauan pustaka memuat tinjauan singkat dan jelas mengenai teori-teori, keterangan-keterangan atau segala sesuatu yang berkaitan dan mendukung masalah yang diajukan baik bersumber pada kepustakaan formal (buku, jurnal, laporan ilmiah tertulis) maupun kondisi nyata yang dapat dibuktikan/diamati. Bagian ini hanya berisikan dasar teori yang sangat dibutuhkan, yang sangat mendukung metode penelitian dan pembahasan. Teori yang bersifat umum tidak perlu dituliskan.

Regresi Poisson adalah teknik analisis statistik yang digunakan pada kasus dimana peubah responnya memiliki distribusi poisson serta memiliki varians yang sama dengan reratanya.

$$Var(Y|x) = E(Y|x) = \mu \quad (1)$$

Regresi poisson menjadi solusi ketika peubah responnya adalah bilangan bulat nonnegatif. Model regresi poisson dapat ditulis sebagai berikut:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}, i = 1, 2, \dots, n \quad (2)$$

dimana $g(\mu_i) = \ln(\mu_i)$ yaitu fungsi hubung log, x_{ik} merupakan peubah penjelas ke-k pada pengamatan ke-i, dan μ adalah nilai tengah banyaknya kejadian (Agresti, 2018). Untuk menduga koefisien parameter dalam model regresi poisson maka diterapkan metode maksimum likelihood pada parameter yang akan diguga, kemudian diiterasikan dengan metode Newton-Rapshon, Berikut adalah fungsi log-likelihoodnya:

$$\ln L(\beta; y) = \sum_{i=1}^n y_i \ln(\mu_i) - \sum_{i=1}^n \mu_i - \sum_{i=1}^n \ln(y_i) \quad (3)$$

Diberbagai kasus sering sekali terjadi overdispersi yang mana pada akhirnya tidak bisa diterapkan regresi poisson. Overdispersi ialah kondisi ketika varians dari peubah respon lebih besar dari reratanya. Overdispersi dapat terjadi ketika data pada variable respon banyak yang bernilai nol. Hal ini yang harus dihindari dalam melakukan analisis regresi poisson karena dalam regresi poisson, asumsi kesamaan antara nilai tengah dengan varians dari peubah respon harus terpenuhi. Jika dipaksakan untuk digunakan maka model yang dibangun akan menghasilkan hasil yang bias atau dengan kata lain dapat menghasilkan peubah yang seharusnya tidak signifikan menjadi signifikan (Hilbe, 2011).

Overdispersi ditentukan dari nilai *deviance* yang dibagi dengan derajat bebasnya. Apabila hasil perhitungan nilai *deviance* dibagi derajat bebas lebih besar dari 1, maka

terindikasi adanya overdispersi. Dengan kata lain menunjukkan bahwa nilai varians lebih besar daripada nilai rerata. Berikut adalah perumusan untuk menghitung nilai Deviance:

$$D = 2 \sum_{i=1}^n \left(y_i \ln \left(\frac{y_i}{\hat{y}_i} \right) - (y_i - \hat{y}_i) \right) \quad (4)$$

Dimana D adalah nilai deviance, y_i adalah nilai peubah respon amatan ke- i dan \hat{y}_i adalah nilai duga peubah respon amatan ke- i (Ulfa, Sholeh, & Sartono, 2021).

Overdispersi diatasi dengan menggunakan regresi binomial negative. Regresi binomial negative merupakan model yang mengasumsikan bahwa peubah responnya mengikuti distribusi binomial negative. Salah satu keunggulan dari regresi binomial negative ialah dapat digunakan dalam keadaan equidispersion ataupun overdispersion. Selain itu karena peubah respon pada regresi binomial negatif diasumsikan mengikuti distribusi binomial negative maka tidak mengharuskan nilai variansnya sama dengan reratanya.

Adapun model dari regresi binomial negatif dapat dituliskan sebagai berikut:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}, i = 1, 2, \dots, n \quad (5)$$

dimana $g(\mu_i) = \ln(\mu_i)$ yaitu fungsi hubung log. Diguakan metode maksimum *likelihood* untuk menduga parameter pada model dan melakukan iterasi menggunakan metode Newton-Raphson agar memperoleh dugaan parameter regresinya (Hilbe, 2011).

Setelah model terentuk untuk mendapatkan model regresi yang optimal dapat diperoleh dengan terlebih dahulu menentukan model yang terbaik. Adapaun cara memilih model mana yang terbaik dapat dilihat dari nilai Akaike Information Criteria (AIC). AIC merupakan suatu kriteria yang mempertimbangkan banyak parameter dalam perhitungan ukuran kebaikan model.

$$AIC = 2k - 2\log(L) \quad (6)$$

dimana L adalah fungsi likel dan k adalah banyaknya parameter penduga pada model. Model yang baik adalah model yang memiliki nilai AIC yang kecil (Hilbe, 2011).

3. METODE PENELITIAN

Data mengenai angka kematian ibu di Kota Bandung pada tahun 2020 ini berasal dari pangkalan data yang bernama Open Data Jabar. Dalam penelitian ini terdapat 30 kecamatan di Kota Bandung sebagai lokasi penelitian. Adapun peubah yang digunakan dalam penelitian ini adalah sebagai berikut (opendata.jabarprov.go.id, 2022):

Tabel 1. Peubah Penelitian

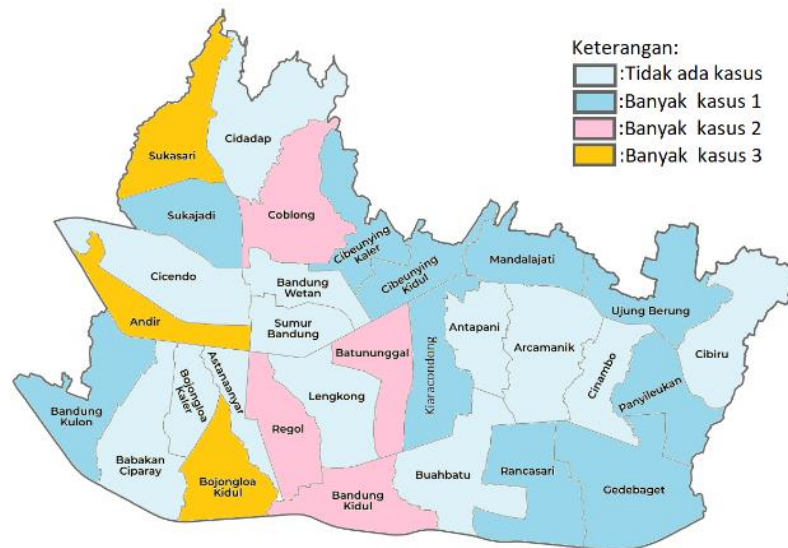
Kode	Peubah
Y	Angka Kematian Ibu
X1	Ibu Hamil Yang Mendapat FE1 (30 tablet)
X2	Ibu Hamil Yang Mendapat FE3 (90 tablet)
X3	Cakupan Imuniasasi TT-2+
X4	Cakupan Imunisasi TT-5
X5	Persalinan Yang Ditangani Oeh Tenaga Kesehatan
X6	Cakupan Pelayanan Nifas
X7	Ibu Hamil Yang Mendapat Kunjungan K4

Tahapan analisis yang dilakukan dalam mencari model terbaik dalam penelitian ini menggunakan bantuan program R dengan tahapan sebagai berikut:

1. Mengumpulkan data dari berbagai pangkalan data.
2. Melakukan eksplorasi data dengan mendeskripsikan data yang diperoleh.
3. Menguji sebaran pada peubah respon, apakah benar memiliki sebaran poisson atau tidak.
4. Menentukan model regresi poisson.
5. Mengidentifikasi overdispersi dengan menghitung nilai koefisien disperse.
6. Menentukan model regresi binomial negatif.
7. Menghitung nilai AIC untuk memilih model terbaik.

4. HASIL DAN PEMBAHASAN

Berdasarkan gambar 1 terlihat bahwa dari 30 kecamatan yang ada di Kota Bandung terdapat 3 kecamatan yang memiliki angka kematian ibu tertinggi sebesar tiga kasus diantaranya Kecamatan Andir, Sukasari dan Bojongloa Kidul. Selanjutnya disusul dengan Kecamatan Bandung Kidul, Batununggal, Coblong dan Regol sebanyak dua kasus, Kecamatan Bandung Kulon, Cibeunying Kaler, Cibeunying Kidul, Gedebage, Kiaracondng, Mandalajati, Panyileukan, Rancasari, Sukajadi dan Ujung Berung masing-masing satu kasus, dan sisanya tidak terjadi kematian ibu di tahun tersebut. Total kasus yang terjadi pada tahun 2020 ini sebanyak 28 kasus.



Gambar 1. Peta Angka Kematian Ibu Di Kota Bandung Tahun 2020

Tahapan selanjutnya adalah memastikan terlebih dahulu bahwa sebaran pada peubah responnya mengikuti sebaran poisson dengan hipotesis sebagai berikut:

H_0 : Angka kematian ibu di kota bandung berdistribusi poisson

H_1 : Angka kematian ibu di kota bandung tidak berdistribusi poisson

Tabel 2. Uji Kolmogorov Smirnov

N		30
Parameter Poisson	Rerata	1.63
Kolmogorov-Smirnov Z		1.304
Nilai-P		0.067

Berdasarkan hasil pengujian pada Tabel 2 diperoleh nilai signifikansi sebesar $0.067 > 0.05$ yaitu taraf nyata, yang berarti belum cukup bukti untuk menolak H_0 , Sehingga sebaran pada peubah respon mengikuti sebaran poisson. Oleh karena sebaran pada peubah responnya mengikuti sebaran poisson maka dapat dilakukan analisis regresi poisson.

Model regresi poisson yang terbetuk adalah sebagai berikut:

$$\begin{aligned} \ln(\mu_i) = & 1.281 + 3.230 \times 10^{-5}x_1 - 3.675 \times 10^{-3}x_2 - 1.866 \times 10^{-4}x_3 \\ & - 2.271 \times 10^{-4}x_4 - 2.391 \times 10^{-1}x_5 - 2.358 \times 10^{-1}x_6 \\ & + 3.836 \times 10^{-3}x_7 \end{aligned} \quad (7)$$

Setelah diperoleh model regresi poisson selanjutnya diuji parameternya, untuk mencari peubah prediktor mana yang berpengaruh secara signifikan terhadap peubah respon.

Tabel 3. Nilai Duga Parameter Model Regresi Poisson

Parameter	Nilai Duga	Nilai-P
β_0	1.281	0.4602
β_1	3.230×10^{-5}	4.67×10^{-5}
β_2	-3.675×10^{-3}	0.04451
β_3	-1.866×10^{-4}	0.38189
β_4	-2.271×10^{-4}	0.79215
β_5	-2.391×10^{-1}	0.00758
β_6	-2.358×10^{-1}	0.00754
β_7	3.836×10^{-3}	0.09199

Berdasarkan Tabel 3 diperoleh bahwa terdapat 4 peubah yang signifikan pada taraf nyata ($\alpha = 5\%$) yaitu variabel x_1 , x_2 , x_5 dan x_6 artinya peubah ibu hamil yang mendapat FE1 (30 tablet), ibu hamil yang mendapat FE3 (90 tablet), persalinan yang ditangani oleh tenaga kesehatan dan cakupan pelayanan nifas berpengaruh signifikan terhadap angka kematian ibu.

Setelah diperoleh model regresi poisson langkah selanjutnya adalah menguji overdispersi, karena salah satu asumsi yang harus terpenuhi jika ingin menggunakan model poisson maka harus tidak ada overdispersi. Maka dari itu dilakukan pengujian dengan menghitung nilai deviance, berikut hasil perhitungannya:

Tabel. 4 Perhitungan Nilai Deviance

Nilai Deviance	Derajat Bebas	Nilai Deviance/ Derajat Bebas
59.25253	29	2.693297

Terlihat bahwa hasil bagi antara nilai deviance dengan derajat bebasnya adalah $2.69 > 1$, ini berarti terjadi overdispersi pada model regresi yang dibangun. Adanya overdispersi menyebabkan model regresi Poisson menjadi kurang baik atau menghasilkan hasil yang bias. Kondisi ini menyebabkan peubah prediktor yang mempengaruhi jumlah kematian ibu tidak dapat dipastikan berdasarkan regresi poisson. Oleh karena hal tersebut maka harus

dilanjutkan dengan model regresi binomial negative untuk mengatasi masalah overdispersi yang ada. Adapun model regresi binomial yang terbentuk sebagai berikut:

$$\begin{aligned} \ln(\mu_i) = & 2.204 + 4.134 \times 10^{-5}x_1 - 5.925 \times 10^{-3}x_2 - 2.947 \times 10^{-4}x_3 \\ & -4.551 \times 10^{-4}x_4 - 3.353 \times 10^{-1}x_5 - 3.320 \times 10^{-1}x_6 \\ & +4.133 \times 10^{-3}x_7 \end{aligned} \quad (8)$$

Tetapi setelah diuji peubah prediktor mana yang berpengaruh secara signifikan terhadap peubah respon diperoleh hasil sebagai berikut:

Tabel 5. Nilai Duga Parameter Model Regresi Poisson

Parameter	Nilai Duga	Nilai-P
β_0	2.204	0.4274
β_1	4.134×10^{-5}	0.00363
β_2	-5.925×10^{-3}	0.06325
β_3	-2.947×10^{-4}	0.44445
β_4	-4.551×10^{-4}	0.76573
β_5	-3.353×10^{-1}	0.01494
β_6	-3.320×10^{-1}	0.01464
β_7	4.133×10^{-3}	0.26526

Berdasarkan table 5 hanya terdapat 3 peubah prediktor saja yang signifikan mempengaruhi angka kematian ibu yaitu ibu hamil yang mendapat FE1 (30 tablet), persalinan yang ditangani oeh tenaga kesehatan dan cakupan pelayanan nifas. Maka interpretasi dari hasil peubah prediktor yang mempengaruhi peubah respon masing-masing adalah sebagai berikut:

- $\beta_1 = 4.134 \times 10^{-5}$ artinya untuk setiap kenaikan satu satuan ibu hamil yang mendapat FE1 (30 tablet) maka angka kematian ibu akan naik sebesar $\exp(4.134 \times 10^{-5})$ dengan asumsi peubah lain dianggap tetap.
- $\beta_5 = -3.353 \times 10^{-1}$ artinya untuk setiap kenaikan satu satuan persalinan yang ditangani oleh tenaga kesehatan maka angka kematian ibu akan turun sebesar $\exp(-3.353 \times 10^{-1})$ dengan asumsi peubah lain dianggap tetap.
- $\beta_6 = -3.320 \times 10^{-1}$ artinya untuk setiap kenaikan satu satuan cakupan pelayanan nifas maka angka kematian ibu akan turun sebesar $\exp(-3.320 \times 10^{-1})$ dengan asumsi peubah lain dianggap tetap.

Langkah terakhir setelah diperoleh model regresi binomial negatif maka akan diuji kebaikan modelnya. Model yang baik adalah model dengan nilai AIC yang lebih kecil Berikut adalah hasil perhitungan AIC dari model regresi poisson dan dari model regresi binomial negatif:

Tabel 6. Nilai AIC

Model	Nilai AIC
Regresi Poisson	121.65
Regresi Binomial Negatif	109.4

5. KESIMPULAN

Berawal dari pemodelan menggunakan regresi poisson karena akan memodelkan bagaimana pengaruh indikator program kesehatan ibu dan anak yang terdiri dari 7 peubah prediktor terhadap angka kematian ibu di Kota Bandung, dimana angka kematian ibu merupakan peubah respon diskrit yang merupakan data count. Setelah diuji overdispersi pada model regresi poisson terindikasi adanya overdispersi dengan nilai deviance yang lebih dari 1. Sehingga untuk mengatasi masalah tersebut diterapkanlah regresi binomial negatif yang mana menghasilkan model dengan hanya terdapat 3 peubah prediktor yang signifikan terhadap angka kematian ibu yaitu peubah ibu hamil yang mendapat FE1 (30 tablet), persalinan yang ditangani oleh tenaga kesehatan dan cakupan pelayanan nifas. Jika dituliskan persamaan regresinya menjadi seperti berikut:

$$y = \exp(4.134 \times 10^{-5}x_1 - 3.353 \times 10^{-1}x_5 - 3.320 \times 10^{-1}x_6) \quad (9)$$

Berdasarkan hasil perhitungan nilai AIC untuk mendapatkan model terbaik, model regresi binomial negatiflah yang terbaik karena menghasilkan nilai AIC yang lebih kecil dari model regresi poisson. Dari model yang terbentuk ini dapat memberikan informasi terhadap pemerintah bahwa terdapat tiga faktor yang perlu menjadi perhatian agar bisa menekan atau bahkan terus mengurangi angka kematian ibu di Kota Bandung.

DAFTAR PUSTAKA

- Agresti, A., 2018. *An Introduction to Categorical Data Analysis*. New York: John Wiley & Sons.
- Cameron, A. C. & Trivedi, P. K., 2013. *Regression Analysis of Count Data*. New York: Cambridge University Press.
- Hilbe, J. M., 2011. *Negative Binomial Regression Second Edition*. New York: Cambridge University Press.
- MC Kota Bandung, 2021. [Online] Available at: <https://infopublik.id/kategori/nusantara/504080/angka-kematian-ibu-dan-bayi-di-kota-bandung-masih-tinggi>
- Tim Open Data Jabar, 2022. [Online] Available at: <https://opendata.jabarprov.go.id/id/dataset?topic=2>
- Tim Penulis Kemenkes RI, 2020. *Indikator Program Kesehatan Masyarakat Dalam RPJMN dan Renstra Kementerian Kesehatan 2020-2024*. Jakarta: Kementerian Kesehatan RI.
- Ulfa, Y. A., Sholeh, A. M. & Sartono, B., 2021. Penanganan Overdispersi pada Model Regresi Poisson dengan Binomial Negatif untuk Jumlah Kasus Baru Kusta di Jawa. *Indonesian Journal of Statistics and Its Applications*, pp. Vol. 5, No. 1, 1-13.
- Utami, T. W., 2013. Analisis Regresi Binomial Negatif Untuk Mengatasi Overdispersion Regresi Poisson Pada Kasus Demam Berdarah Dengue. *Jurnal Statistika Universitas Muhammadiyah Semarang*, pp. Vol. 1, No. 2.