

Uji Eksistensi *Zero Inflation* pada Regresi Poisson untuk Kasus Kematian Ibu

Zahra Hairiyah Putri *, Nusar Hajarisman

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Indonesia.

zahrahryp@gmail.com, nusarhajarisman@unisba.ac.id

Abstract. Maternal Mortality Rate (MMR) is a key indicator in evaluating health system performance, especially in achieving the Sustainable Development Goals (SDGs). MMR data, which are count data, are commonly analyzed using Poisson regression. However, the Poisson model has limitations because it assumes the same mean value and variance, while in reality there is often an overdispersion of many zeros (*excess zero*). This study investigates maternal death data in Bandung City in 2020 by addressing two main questions: (1) How to detect zero inflation in the data process modeled by the Naive test in the case of maternal deaths in Bandung City in 2020? (2) What is the impact caused by the zero inflation problem in Poisson regression modeling in the case of maternal deaths in Bandung City in 2020?. The research uses an inferential quantitative approach with secondary data from Dinas Kesehatan Bandung City. The results show that there is a significant presence of excess zeros in the data, making the Poisson model unsuitable. Instead, the Zero-Inflated Negative Binomial regression model provides a better fit by handling both overdispersion and zero inflation effectively by showing an increase in the standard error value.

Keywords: *Overdispersion, Zero Inflated Negative Binomial, Naive Test.*

Abstrak. Angka kematian ibu (AKI) merupakan indikator penting dalam menilai keberhasilan sistem kesehatan, terutama dalam pencapaian Tujuan Pembangunan Berkelanjutan (SDGs). Data AKI dapat dianalisis dengan metode regresi Poisson yang menghubungkan variabel respon berupa data cacahan dengan variabel prediktor. Namun, model Poisson mengasumsikan nilai rata-rata dan varians yang sama, sedangkan dalam kenyataannya sering terjadi overdispersion banyaknya data nol (*excess zero*). Untuk menguji keberadaan *excess zero* atau *zero inflation*, dapat dilakukan menggunakan uji *Naive*. Berdasarkan fenomena tersebut, maka permasalahan dalam penelitian ini dirumuskan sebagai berikut: (1) Bagaimana cara mendeteksi *zero inflation* dalam proses data yang dimodelkan oleh uji *Naive* pada kasus kematian ibu di Kota Bandung tahun 2020? (2) Bagaimana dampak yang ditimbulkan oleh permasalahan *zero inflation* dalam pemodelan regresi Poisson pada kasus kematian ibu di Kota Bandung tahun 2020?. Peneliti menggunakan jenis penelitian kuantitatif inferensial. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Dinas Kesehatan Kota Bandung. Hasil dari penelitian ini adalah: Terdapat *excess zero* pada data kasus kematian ibu di Kota Bandung tahun 2020 sehingga model Poisson tidak layak digunakan dan pemodelan data menggunakan regresi *Zero Inflated Negative Binomial* mampu mengatasi overdispersi pada data kematian ibu di Kota Bandung tahun 2020 dengan menunjukkan peningkatan nilai *standard error*.

Kata Kunci: *Overdispersi, Zero Inflated Negative Binomial, Naive Test.*

A. Pendahuluan

Angka Kematian Ibu (AKI) menjadi salah satu indikator utama untuk melihat keberhasilan program kesehatan ibu. AKI didefinisikan kejadian kematian seorang perempuan yang terjadi pada masa kehamilan, persalinan berlangsung, keguguran, aborsi atau selama 42 hari setelah pasca kehamilan yang disebabkan diluar kecelakaan (Dinkes Kota Bandung, 2021). Secara global, angka kematian ibu juga menjadi tolok ukur dalam pencapaian Tujuan Pembangunan Berkelanjutan (SDGs), yakni poin ketiga, “*Good Health and Well-being*” atau “Kehidupan Sehat dan Sejahtera”, yang ditujukan pada segala usia. Kematian ibu juga merupakan cerminan dari kualitas sistem kesehatan di suatu daerah (Jannah et al., 2024).

Pada tahun 2020, Provinsi Jawa Barat menduduki peringkat pertama dalam jumlah penduduk terbanyak, sebesar 48.274,20 (dalam ribu). Sama hal nya dalam angka kematian ibu, Provinsi Jawa Barat menduduki peringkat pertama dengan angka kematian ibu sebesar 745 kasus, dimana mengalami kenaikan 61 kasus kematian dari tahun sebelumnya. Di Provinsi Jawa Barat, Kota Bandung menduduki peringkat kedua belas dengan jumlah kematian ibu sebanyak 28 kasus. Kematian ibu 60% terjadi saat pascapersalinan dan hampir 50% diantaranya terjadi dalam 24 jam pascapersalinan (Saputri, 2020). Setengah dari jumlah AKI di Kota Bandung tahun 2020 terjadi pada masa nifas dan penyebabnya merupakan komplikasi perdarahan (12 kasus), disusul oleh kematian akibat hipertensi pada kehamilan dan infeksi (masing-masing 3 kasus), gangguan sistem peredaran darah (2 kasus) dan 8 kasus lainnya masuk ke dalam kematian dengan penyebab lain-lain (Dinkes Kota Bandung, 2021).

Jumlah kematian ibu di Kota Bandung, yang menampilkan berbagai faktor penyebab, merupakan contoh data cacahan karena data yang diperoleh dengan cara menghitung jumlah kasus dalam rentang waktu tertentu dan berbentuk bilangan bulat positif. Dalam analisis data cacahan, regresi Poisson merupakan metode yang umum digunakan untuk melihat hubungan antara variabel respon yang berupa data cacahan dengan satu atau lebih variabel prediktor. Dalam model dari regresi Poisson, variabel respon diasumsikan mengikuti distribusi Poisson, yakni peluang kejadian dalam periode waktu tertentu dan berupa bilangan bulat non negatif. Asumsi lainnya yaitu nilai rata-rata dan nilai ragam bernilai sama (*equidispersion*). Namun, seringkali asumsi tersebut tidak terpenuhi yang diakibatkan nilai varians yang lebih besar (*overdispersion*) atau lebih kecil (*underdispersion*) dari nilai rata-ratanya.

Overdispersion adalah kondisi dimana variabel respon memiliki *variens* yang lebih besar dari *mean* nya (Bal & Sodoudi, 2020). *Overdispersion* juga menghasilkan kesimpulan yang tidak sesuai yang diakibatkan nilai *standard error* lebih kecil (*underestimates*) dari yang seharusnya (Payne et al., 2018). Secara teoritis, *overdispersion* diakibatkan oleh terjadinya korelasi antar respon individu. Secara praktisnya, *overdispersion* diakibatkan terbentuknya pengelompokkan pada data dan banyak nilai nol pada data (*excess zero*). *Excess zero* sendiri disebabkan oleh beberapa faktor, seperti kejadian yang memang jarang terjadi atau hal lain yang menjadikan suatu kejadian tidak mungkin terjadi. Metode untuk mengatasi *overdispersion* akibat terjadinya korelasi antar respon dan terbentuknya pengelompokkan data (*clustering*) dapat menggunakan model *Random Effects*, model *Multilevel*, dan *Generalized Estimating Equations* (GEE). Sementara untuk *overdispersion* yang diakibatkan oleh *excess zero*, dapat menggunakan regresi *Zero-Inflated* dan regresi Hurdle. Dalam beberapa kasus, regresi *Zero-Inflated* dapat memberikan fit yang lebih baik terhadap data dibandingkan dengan regresi Hurdle.

Model regresi *Zero Inflated* dirancang untuk mengatasi masalah nol berlebih dengan cara memisahkan proses yang menghasilkan nol dan proses yang menghasilkan angka lain, sehingga hasil analisis menjadi lebih akurat. Dengan menggunakan model ini, akan memberikan pemahaman yang lebih baik dan prediksi yang lebih tepat terhadap data yang dimiliki. Untuk jenis data cacahan dapat menggunakan model regresi *Zero-Inflated Poisson* (ZIP) dan *Zero Inflated Negative Binomial*. Namun, regresi ZIP menjadi kurang tepat apabila nilai ragam jauh lebih besar dibandingkan dengan nilai rata-rata dari variabel respon (Astuti & Mulyanto, 2016). Oleh sebab itu, metode regresi *Zero Inflated Negative Binomial* (ZINB) bisa menjadi alternatif dalam memodelkan data dengan karakteristik data tersebut. Regresi *Zero Inflated Negative Binomial* (ZINB) juga mampu mengatasi *overdispersi* dan *excess zero* dalam data. Hasil dari penelitian terdahulu yang dilakukan oleh Prianggada (Prianggada, 2017), dalam kasus Angka Kematian Ibu di Kabupaten Bojonegoro, menunjukkan bahwa model regresi ZINB lebih mampu mengatasi *overdispersi* dibandingkan dengan

model regresi ZIP.

Dalam menguji keberadaan *excess zero* pada regresi Poisson, He, dkk. (He et al., 2019) memperkenalkan Uji *Naive* sebagai alternatif yang dapat digunakan. Uji *Naive* merupakan metode yang sederhana dan relatif cepat. Uji *Naive* melibatkan perbandingan langsung antara jumlah data nol yang diamati dengan jumlah nol yang diprediksi oleh model. Berdasarkan uraian tersebut, peneliti bermaksud untuk menerapkan uji *Naive* untuk melihat eksistensi *zero inflation* pada kasus kematian ibu di Kota Bandung tahun 2020. Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana cara mendeteksi *zero inflation* dalam proses data yang dimodelkan oleh uji *Naive* pada kasus kematian ibu di Kota Bandung tahun 2020?
2. Bagaimana dampak yang ditimbulkan oleh permasalahan *zero inflation* dalam pemodelan regresi Poisson pada kasus kematian ibu di Kota Bandung tahun 2020?

Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Untuk mengetahui cara mendeteksi *zero inflation* dalam proses data yang dimodelkan oleh uji *Naive* pada kasus kematian ibu di Kota Bandung tahun 2020.

Untuk mengetahui dampak yang ditimbulkan oleh permasalahan *zero inflation* dalam pemodelan regresi Poisson pada kasus kematian ibu di Kota Bandung tahun 2020.

B. Metode

Peneliti menggunakan jenis penelitian kuantitatif inferensial karena dalam penelitian ini, melibatkan pengumpulan data, penyajian data, analisa dengan menguji hipotesis hingga penarikan kesimpulan. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Dinas Kesehatan Kota Bandung. Variabel yang diamati terdiri dari satu variabel respon (Angka Kematian Ibu) dan empat variabel prediktor (Persentase K1, Persentase K4, Persentase Pelayanan Ibu Nifas dan Persentase Komplikasi Kebidanan yang Ditangani).

Regresi Poisson

Model regresi Poisson merupakan suatu model regresi untuk menganalisis data yang berbentuk data cacahan (*count data*). Jika data pada variabel respon tidak berdistribusi normal dan berjenis diskrit, maka salah satu metode yang dapat digunakan adalah regresi Poisson (Tendriyawati et al., 2023). Regresi Poisson menggunakan fungsi hubung log yang berbentuk eksponen $\exp(x_i^T \beta)$. Berikut fungsi masa peluang dari distribusi Poisson:

$$P(y_i; \beta) = e^{-s_i \lambda(y_i \beta)} \cdot \frac{(s_i \lambda(x_i \beta))^{y_i}}{y_i!} \quad \dots(1)$$

Fungsi masa peluang digunakan untuk menghitung jumlah kejadian atau peluang kejadian pada variabel acak X dengan λ merupakan rata-rata kejadian. Berdasarkan Persamaan (1), didapat nilai rata-rata dan varians dari variabel acak X , $E(X) = \text{Var}(X) = \lambda$. Model regresi Poisson menggunakan fungsi hubung log yang dinyatakan sebagai berikut:

$$\ln E(Y_i) = \ln(\mu_i) = \lambda(x_i) = x_i^T \beta = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad \dots(2)$$

Atau secara ekuivalen:

$$\mu_i = \exp(x_i^T \beta) + \varepsilon_i \quad \dots(3)$$

Penaksiran parameter

Penaksiran parameter yang umum digunakan adalah *Maximum Likelihood Esttimation* (MLE). Fungsi kemungkinan Y_i yang berdistribusi Poisson secara umum dapat dituliskan sebagai berikut:

$$L(\beta|y_i) = \prod_{i=1}^n f(y_i, \beta) = \prod_{i=1}^n e^{-s_i \lambda(x_i \beta)} \cdot \frac{(s_i \lambda(x_i \beta))^{y_i}}{y_i!} \quad \dots(4)$$

Logaritma dari fungsi likelihoodnya:

$$\ln L(\beta|y_i) = l(\beta|y_i) = \sum_{i=1}^n [y_i \ln s_i - \ln y_i! + y_i \lambda(x_i \beta) - s_i \lambda(x_i \beta)] \quad \dots(5)$$

Dimana MLE dapat diperoleh dengan menurunkan fungsi log likelihood tersebut terhadap parameter intersep β_0 dan parameter koefisien β_i , dimana $i = 1, 2, \dots, p$. Namun, karena Persamaan (5) tidak dapat secara langsung diselesaikan, maka dapat menggunakan metode iterasi *Fisher Scoring*.

Evaluasi Model

Evaluasi model digunakan untuk memastikan pemilihan variabel independen sudah tepat dalam memodelkan data. Dalam evaluasi model terdiri dari uji signifikansi parameter θ secara simultan (bersama-sama) dan parsial (terpisah). Parameter θ didefinisikan sebagai parameter yang digunakan di dalam model. Dalam pengujian parameter secara simultan, dapat menggunakan uji G . Hipotesis yang digunakan sebagai berikut:

$H_0 : \theta_1 = \theta_2 = \dots = \theta_p = 0$ (Tidak ada pengaruh antara variabel independen secara bersama-sama terhadap variabel dependen)

$H_1 : \text{Minimal terdapat satu } \theta_j \neq 0; j = 1, 2, \dots, p$ (Minimal ada pengaruh salah satu dari variabel independen terhadap variabel dependen)

Dengan perhitungan untuk uji simultan menggunakan menggunakan uji G yang dapat dilihat pada persamaan berikut:

$$G = -2 \ln L \left(\frac{L_0}{L_1} \right) = -2 \ln L(L_0 - L_1) \quad \dots(6)$$

dimana:

L_0 : *likelihood* model tanpa variabel independen

L_1 : *likelihood* model dengan semua variabel independen

Sementara itu, pengujian signifikansi parameter secara individu dapat menggunakan uji *Wald*. Uji signifikansi secara individu (parsial) digunakan untuk melihat kelayakan setiap variabel independen untuk masuk ke dalam model. Hipotesis dari uji *Wald* untuk model *count* sebagai berikut:

$H_0 : \hat{\theta}_j = 0; j = 1, 2, \dots, p$ (Tidak ada pengaruh antara variabel independen ke- j terhadap variabel dependen)

$H_1 : \hat{\theta}_j \neq 0; j = 1, 2, \dots, p$ (Ada pengaruh antara variabel independen ke- j terhadap variabel dependen)

Statistik uji *Wald* dirumuskan sebagai berikut:

$$W_j = \frac{\hat{\theta}_j}{se(\hat{\theta}_j)} \quad \text{dimana } se(\hat{\theta}_j) = \sqrt{Var(\hat{\theta}_j)} \quad \dots(7)$$

Kriteria uji dari pengujian ini yaitu hipotesis nol ditolak jika statistik uji $|W| > Z_{\alpha/2}$.

Overdispersi

Dalam regresi Poisson, asumsi equidispersi atau kondisi dimana nilai rata-rata dengan nilai varians pada variabel respon bernilai sama sangat penting. Namun, pada beberapa kasus, asumsi tersebut tidak dapat terpenuhi akibat nilai varians yang lebih besar dari nilai rata-ratanya yang disebut overdispersi. Hipotesis dalam pengujian overdispersi sebagai berikut:

$H_0 : \theta = 0$; Tidak terjadi overdispersi pada model regresi Poisson

$H_1 : \theta \neq 0$; Terjadi overdispersi pada model regresi Poisson

Pengujian overdispersi dilakukan dengan membandingkan nilai *Pearson Chi-Square* dengan nilai derajat bebasnya yang dinyatakan sebagai berikut:

$$\theta = \frac{\chi^2}{n - p} \quad \dots(8)$$

dimana rumus *Pearson Chi-Square* yakni sebagai berikut:

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{var(\hat{\mu}_i)} \quad \dots(9)$$

dimana:

χ^2 : *Pearson Chi-Square*

θ : parameter dispersi

Apabila nilai θ mendekati nilai 1, maka dapat disimpulkan bahwa tidak terjadi overdispersi. Dan sebaliknya, apabila nilai θ lebih besar dari 1, maka dapat disimpulkan bahwa terjadi overdispersi pada data (Famoye et al., 2004).

Zero Inflation

Zero Inflation merupakan istilah yang digunakan dalam dunia statistik atas ketidakseimbangan pada data yang diakibatkan kelebihan nilai nol (*excess zero*) dari yang diharapkan pada variabel respon.

Keadaan ini sering terjadi pada data diskrit. Keberadaan *zero inflation* dapat dilihat dengan pemeriksaan proporsi nilai nol. *Excess zero* lebih dari 50% yang menyebabkan terjadinya overdispersi, juga mengakibatkan ketidaktepatan penggunaan regresi Poisson dalam menggambarkan data yang sebenarnya (Famoye et al., 2004). Namun, untuk lebih meyakinkan lagi, dapat digunakan uji *Naive* untuk melihat keberadaan *zero inflation* pada data.

Uji *Naive*

Uji *Naive* merupakan metode yang menentukan apakah data mengandung *excess zero* dengan melihat perbandingan antara proporsi yang diamati dengan proporsi yang diharapkan dari nilai nol. Dalam mengendalikan kesalahan tipe I, uji *Naive* cenderung lebih konservatif atau berhati-hati dalam menolak hipotesis H_0 . Namun, kekurangan uji ini adalah kurang sensitif, terutama ketika efek *zero inflation* tidak terlalu kuat atau ketika ada variabel tambahan yang memengaruhi komponen *zero inflation*, sehingga terkadang uji *Naive* sulit mendeteksi keberadaan *zero inflation* dengan baik. Misalkan suatu sampel independen $y_i \sim \text{Poisson}(\mu_i)$ dengan $\log(\mu_i) = x_i^T \beta$ dan memisalkan r_i sebagai vektor indikator dari y_i , yakni merubah variabel Y menjadi data biner yang dapat dituliskan sebagai berikut:

$$r_i = \begin{cases} 1, & \text{untuk } y_i = 0 \\ 0, & \text{untuk } y_i > 0 \end{cases} \quad \dots(10)$$

Dengan probabilitas $y_i = 0$ dengan kovariat $x_i = \exp(-\exp(x_i^T \beta))$, maka proporsi nilai nol yang diharapkan adalah:

$$E(r_i) = \frac{1}{n} \sum_{i=1}^n \exp(-\exp(x_i^T \beta)) \quad \dots(11)$$

Sementara selisih antara proporsi yang diamati dengan proporsi yang diharapkan dari nilai nol adalah:

$$s = \frac{1}{n} \sum_{i=1}^n (r_i - E(r_i)) = \frac{1}{n} \sum_{i=1}^n r_i - \frac{1}{n} \sum_{i=1}^n \exp(-\exp(x_i^T \beta)) \quad \dots(12)$$

Dari Persamaan (2) didapat nilai $E(s) = 0$ dan $\text{Var}(s) = \frac{1}{n^2} \sum_{i=1}^n \exp(-\exp(x_i^T \beta))(1 - \exp(-\exp(x_i^T \beta)))$. Misalkan $\hat{\beta}$ merupakan estimasi parameter β melalui MLE, sehingga dapat diperoleh s' berdistribusi normal asimtotik $(0,1)$.

$$\hat{s}' = \frac{\frac{1}{n} \sum_{i=1}^n (r_i - \exp(-\exp(x_i^T \hat{\beta})))}{\left[\frac{1}{n^2} \sum_{i=1}^n \exp(-\exp(x_i^T \hat{\beta}))(1 - \exp(-\exp(x_i^T \hat{\beta}))) \right]^{\frac{1}{2}}} \sim N(0,1) \quad \dots(13)$$

Hipotesis yang digunakan dalam uji *Naive* yaitu:

H_0 : Tidak terdapat *zero inflation* atau jumlah nol yang diamati sesuai dengan yang diharapkan berdasarkan model Poisson (Data mengikuti distribusi Poisson).

H_1 : Terdapat *zero inflation* atau jumlah nol yang diamati lebih banyak daripada yang diprediksi oleh model Poisson.

Dengan statistika uji menggunakan Persamaan (13) dan kriteria uji yang digunakan yakni tolak H_0 apabila nilai $\hat{s}' > Z_{1-\alpha}$ atau $p\text{-value} < \alpha$.

Model Regresi *Zero Inflated Negative Binomial*

Regresi *Zero Inflated Negative Binomial* (ZINB) merupakan model campuran lainnya yang digunakan untuk mengatasi terjadinya overdispersi pada model yang mengandung *excess zero*. Model regresi ZINB mengandung distribusi Poisson dan distribusi Gamma (Rahmawati, 2018). Pada model regresi ZINB, terjadi dua keadaan (*state*) pada variabel respon yang diduga terjadi nol kejadian. Dua *state* tersebut yaitu *Zero State*, dimana data yang menghasilkan keadaan nol kejadian dengan peluang π_i dan *Negative Binomial State*, dimana data yang mengikuti distribusi Binomial Negatif dengan rata-rata λ dan peluang $(1 - \pi_i)$. Fungsi probabilitas untuk regresi ZINB dinyatakan sebagai berikut:

$$P(Y = y_i) = \begin{cases} \pi_i + (1 - \pi_i) \left(\frac{1}{1 + \kappa \lambda_i} \right)^{\frac{1}{\kappa}}, & \text{untuk } y_i = 0 \\ (1 - \pi_i) \frac{\Gamma(y_i + \frac{1}{\kappa})}{\Gamma(y_i + \frac{1}{\kappa}) y_i!} \left(\frac{1}{1 + \kappa \lambda_i} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa \lambda_i}{1 + \kappa \lambda_i} \right)^{y_i}, & \text{untuk } y_i > 0 \end{cases} \quad \dots(14)$$

dimana κ merupakan parameter dispersi (overdispersi) dan $\Gamma(\cdot)$ merupakan fungsi gamma. Diasumsikan bahwa λ_i dan π_i bergantung pada variabel prediktor x_i dengan fungsi hubung untuk regresi ZINB dinyatakan sebagai berikut:

$$\ln(\lambda_i) = x_i^T \beta \text{ dan } \text{logit}(\pi_i) = \ln \left[\frac{\pi_i}{1 - \pi_i} \right] = x_i^T \gamma \quad \dots(15)$$

Sehingga diperoleh parameter μ_i dan ω_i sebagai berikut:

$$\lambda_i = \exp(x_i^T \beta) \text{ dan } \pi_i = \frac{\exp(x_i^T \gamma)}{1 + \exp(x_i^T \gamma)} \quad \dots(16)$$

dengan rata-rata dan varians nya ditunjukkan sebagai berikut:

$$E(Y_i) = (1 - \pi_i) \lambda_i \text{ dan } \text{Var}(Y_i) = \lambda_i (1 - \pi_i) (1 + \kappa \lambda_i + \pi_i \lambda_i) \quad \dots(17)$$

sehingga model regresi ZINB untuk model *count* (model Binomial Negatif) adalah:

$$\ln(\lambda_i) = x_i^T \beta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad \dots(18)$$

dan untuk model *zero inflation* adalah:

$$\text{logit}(\pi_i) = x_i^T \gamma = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \dots + \gamma_p X_p \quad \dots(19)$$

Pada regresi *Zero Inflated Negative Binomial*, estimasi parameter β dan γ dapat menggunakan *Maximum Likelihood Estimation* (MLE). Dari Persamaan (16), disubstitusikan ke dalam Persamaan (14) sehingga didapat persamaan fungsi *likelihood* dengan β dan γ yang akan diestimasi:

$$L(\beta, \gamma | y_i) = \begin{cases} \prod_{i=1}^n \frac{\exp(x_i^T \gamma)}{1 + \exp(x_i^T \gamma)} + \frac{1}{1 + \exp(x_i^T \gamma)} \left(\frac{1}{1 + \kappa \exp(x_i^T \beta)} \right)^{\frac{1}{\kappa}}, & \text{untuk } y_i = 0 \\ \prod_{i=1}^n \frac{1}{1 + \exp(x_i^T \gamma)} \frac{\Gamma(y_i + \frac{1}{\kappa})}{\Gamma(y_i + \frac{1}{\kappa}) y_i!} \left(\frac{1}{1 + \kappa \exp(x_i^T \beta)} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa \mu_i}{1 + \kappa \exp(x_i^T \beta)} \right)^{y_i}, & \text{untuk } y_i > 0 \end{cases} \quad \dots(20)$$

dan logaritma dari fungsi *likelihood* nya adalah:

$$\ln(\beta, \gamma | y_i) = \begin{cases} \sum_{i=1}^n \ln \left(e^{x_i^T \gamma} + \exp(-e^{x_i^T \beta}) \right) - \sum_{i=1}^n \frac{1}{\kappa} \ln(1 + \kappa e^{x_i^T \beta}), & \text{untuk } y_i = 0 \\ - \sum_{i=1}^n \ln(1 + e^{x_i^T \gamma}) + \sum_{i=1}^n \ln \left(\frac{\Gamma(y_i + \frac{1}{\kappa})}{\Gamma(y_i + \frac{1}{\kappa}) y_i!} \right) - \sum_{i=1}^n \ln y_i! - \\ \sum_{i=1}^n \left(\frac{1}{\kappa} + y_i \right) \ln(1 + \kappa e^{x_i^T \beta}) + \sum_{i=1}^n y_i \ln(\kappa e^{x_i^T \beta}), & \text{untuk } y_i > 0 \end{cases} \quad \dots(21)$$

Sehingga, untuk mengestimasi parameter, dapat dilakukan dengan Persamaan (21). Namun, Persamaan (21) tidak dapat diselesaikan dengan metode numerik biasa, sehingga dalam memaksimalkan fungsi *log likelihood* dapat dilakukan dengan algoritma *Expectation Maximization* (EM).

C. Hasil Penelitian dan Pembahasan Analisis Deskriptif

Berikut merupakan hasil dari analisis statistika deskriptif yang mencakup informasi awal dari data kematian ibu di Kota Bandung tahun 2020.

Tabel 1. Hasil Ringkasan Analisis Deskriptif

Variabel	Jumlah	Rata-Rata	Standar Deviasi	Minimal	Maksimal
Kematian Ibu	80	0	1	0	5
Persentase K1 (%)	80	89,6391	1,0870	23,78	136,70
Persentase K4 (%)	80	82,6473	18,8161	22,90	117,01
Persentase Pelayanan Ibu Nifas (%)	80	83,4526	15,3212	31,23	109,93
Persentase Komplikasi Kebidanan yang Ditangani (%)	80	52,0810	32,1977	0,00	116,20

Berdasarkan Tabel 1, diperoleh nilai rata-rata terjadinya kasus kematian ibu di setiap UPT Puskesmas di Kota Bandung sebanyak 0 yang artinya secara umum tidak terjadi kematian ibu di UPT Puskesmas tersebut. Namun, dalam data diperoleh nilai maksimum sebesar 5 kasus. Nilai standar deviasi yang lebih besar dari nilai rata-rata sebesar 1, mengindikasikan adanya variasi yang cukup besar antar pengamatan. Pada variabel Persentase K1 (%), diperoleh nilai rata-rata sebesar 89,64%, yang artinya banyak ibu hamil yang menerima perawatan prenatal sesuai standar pada trimester pertama kehamilan dengan standar deviasi sebesar 18,0870, nilai minimum sebesar 23,78% dan nilai maksimum sebesar 136,70%. Pada variabel Persentase K4 (%), diperoleh nilai rata-rata sebesar 82,65%, yang artinya cukup banyak ibu hamil yang menerima perawatan prenatal sesuai standar minimal empat kali dengan standar deviasi sebesar 18,8161, nilai minimum sebesar 22,90% dan nilai maksimum sebesar 117,01%. Pada variabel Persentase Pelayanan Ibu Nifas (%), diperoleh nilai rata-rata sebesar 82,45%, yang artinya cukup banyak tenaga kesehatan yang memberikan pelayanan kesehatan kepada ibu melahirkan dalam masa nifas dengan standar deviasi sebesar 15,3212, nilai minimum sebesar 31,23% dan nilai maksimum sebesar 109,93%. Pada variabel Persentase Komplikasi Kebidanan yang Ditangani (%), diperoleh nilai rata-rata sebesar 52,08%, yang artinya hampir setengah komplikasi kehamilan yang menerima perawatan sesuai standar kebidanan oleh tenaga kesehatan terlatih pada tingkat perawatan primer dengan standar deviasi sebesar 32,1977, nilai minimum sebesar 0% dan nilai maksimum sebesar 116,20%.

Pemodelan Regresi Poisson

Untuk mengetahui faktor mana saja yang mempengaruhi kematian ibu di Kota Bandung tahun 2020, dilakukan pemodelan terlebih dahulu dengan menentukan taksiran parameter. Hasil estimasi dapat dilihat pada Tabel 2.

Tabel 2. Hasil Estimasi Parameter Model Regresi Poisson

Parameter	Estimasi	Standar Error	Z-value	P-value	Keputusan
β_0	-1,9306	1,2037	-1,604	0,109	Terima H_0
β_1	-0,0133	0,0185	-0,720	0,472	Terima H_0
β_2	0,0063	0,0176	0,357	0,721	Terima H_0
β_3	0,0123	0,0213	0,580	0,562	Terima H_0
β_4	0,0088	0,0066	1,342	0,180	Terima H_0

Model regresi Poisson pada data kematian ibu di Kota Bandung tahun 2020 yaitu:

$$\ln(\mu_i) = -1,9306 - 0,0133X_1 + 0,0063X_2 + 0,0123X_3 + 0,0088X_4$$

$$\mu_i = \exp(-1,9306 - 0,0133X_1 + 0,0063X_2 + 0,0123X_3 + 0,0088X_4)$$

Dari hasil pemodelan, dilakukan pengecekan asumsi overdispersi pada model.

Uji Overdispersi

Uji overdispersi dilakukan untuk melihat apakah asumsi equidispersi sudah terpenuhi atau belum. Hipotesis dalam pemeriksaan overdispersi yakni sebagai berikut:

$H_0 : \theta = 0$; Tidak terjadi overdispersi pada model regresi Poisson

$H_1 : \theta \neq 0$; Terjadi overdispersi pada model regresi Poisson

Berdasarkan hasil pengujian, diperoleh nilai $p - value = 0,09 < \alpha = 0,1$ (10%), maka dapat disimpulkan bahwa cukup bukti untuk menolak hipotesis H_0 atau terjadi overdispersi pada model regresi Poisson pada data kematian ibu di Kota Bandung tahun 2020 pada taraf signifikansi 10%. Sementara itu, nilai dispersi sebesar $1,8779 > 1$ mengindikasikan terjadinya overdispersi dalam data kematian ibu di Kota Bandung tahun 2020.

Uji Eksistensi Zero Inflation Melalui Uji Naive

Pemeriksaan *zero inflation* atau *excess zero* dapat dilihat dari persentase proporsi nilai nol pada data amatan. Apabila proporsi nilai nol pada data $> 50\%$, maka dapat dikatakan bahwa data amatan mengandung *excess zero*. Berdasarkan data kematian ibu di Kota Bandung tahun 2020, terdapat 80 data amatan dengan banyaknya amatan bernilai nol ada sebanyak 62 data. Maka, diperoleh nilai proporsi nilai nol sebesar $77,5\% > 50\%$, sehingga dapat disimpulkan bahwa terdapat *zero inflation* dalam data. Oleh karena itu, data kematian ibu di Kota Bandung tahun 2020 tidak dapat dimodelkan

menggunakan model regresi Poisson standar. Namun, untuk memastikannya, perlu dilakukan uji *Naive*. Uji *Naive* merupakan metode yang menentukan apakah data mengandung *excess zero* dengan melihat perbandingan antara proporsi yang diamati dengan proporsi yang diharapkan dari nilai nol. Hipotesis yang digunakan dalam uji ini yaitu:

H_0 : Tidak terdapat *zero inflation* atau jumlah nol yang diamati sesuai dengan yang diharapkan berdasarkan model Poisson (Data mengikuti distribusi Poisson)

H_1 : Terdapat *zero inflation* atau jumlah nol yang diamati lebih banyak daripada yang diprediksi oleh model Poisson

Kriteria dari uji *Naive* yaitu tolak H_0 apabila nilai $\hat{s}' > Z_{1-\alpha}$ atau $p - value < \alpha$. Berdasarkan hasil pengujian, dengan menggunakan taraf signifikansi $\alpha = 10\%$, diperoleh nilai $S = 1,2920 > Z_{1-\alpha} = Z_{1-0,10} = Z_{0,90} = 1,28$ dan $p - value = 0,0982 < \alpha = 0,1$ (10%). Oleh karena itu, dapat ditarik kesimpulan bahwa cukup bukti untuk menolak hipotesis H_0 atau terdapat *zero inflation* (jumlah nol yang diamati lebih banyak daripada yang diprediksi) sehingga model regresi *Zero Inflated Negative Binomial* dapat digunakan untuk memodelkan data kematian ibu di Kota Bandung tahun 2020 dibawah taraf signifikansi 10%.

Pemodelan Regresi *Zero Inflated Negative Binomial*

Model alternatif dalam mengatasi data *count* yang overdispersi akibat terdapat *zero inflation* di dalam data dapat menggunakan regresi *Zero Inflated Negative Binomial* (ZINB).

Tabel 3. Hasil Pemodelan Regresi *Zero Inflated Negative Binomial*

Model	Parameter	Estimasi	Standar Error	Z-value	P-value	Keputusan
<i>Negative Binomial</i>	β_0	0,2862	1,7910	0,160	0,8731	Terima H_0
	β_1	-0,0569	0,0294	-1,936	0,0528	Tolak H_0
	β_2	-0,0073	0,0263	-0,276	0,7827	Terima H_0
	β_3	0,0646	0,0338	1,911	0,0560	Tolak H_0
	β_4	-0,0077	0,0117	-0,659	0,5098	Terima H_0
<i>Zero Inflation</i>	γ_0	1,5111	3,1022	0,487	0,6262	Terima H_0
	γ_1	-0,1035	0,0783	-1,322	0,1860	Terima H_0
	γ_2	-0,0345	0,0482	-0,716	0,4737	Terima H_0
	γ_3	0,1481	0,0842	1,760	0,0785	Tolak H_0
	γ_4	-0,0476	0,0318	-1,498	0,1342	Terima H_0

Berdasarkan hasil pemodelan regresi ZINB Tabel 3, diperoleh model *count* regresi ZINB data kematian ibu di Kota Bandung tahun 2020 dengan fungsi hubung log (model Binomial Negatif) yaitu: $\ln(\lambda_i) = 0.2861 - 0.0569X_1 - 0.0073X_2 + 0.0646X_3 - 0.0077X_4$

Sementara itu, model regresi ZINB dengan fungsi hubung logit (model *zero inflation*) yaitu:

$$\text{logit}(\pi_i) = 1.5111 - 0.1035X_1 - 0.0345X_2 + 0.1481X_3 - 0.0476X_4$$

Dalam model regresi ZINB terdapat dua parameter yang digunakan, yakni β untuk model *count* (model Binomial Negatif) dan γ untuk model *zero inflation*. Hipotesis yang digunakan sebagai berikut.

$H_0 : \theta_p = 0; j = 1, 2, \dots, 4$ (Tidak ada pengaruh antara variabel independen ke- j terhadap variabel kematian ibu)

$H_1 : \theta_j \neq 0; j = 1, 2, \dots, 4$ (Ada pengaruh antara variabel independen ke- j terhadap variabel kematian ibu)

Dengan menggunakan taraf signifikansi 10%, diperoleh kriteria uji yaitu apabila nilai $Z - value < Z_{1-0,1/2} = Z_{0,95} = 1,645$ dan $p - value > \alpha(0,10)$, maka variabel independen tidak berpengaruh secara signifikan terhadap variabel kematian ibu. Berdasarkan Tabel 3, dalam model *count* (model *Negative Binomial*), dapat disimpulkan bahwa variabel Persentase K1 (X_1) dan Persentase Pelayanan Ibu Nifas (X_3) memberikan pengaruh terhadap variabel Kematian Ibu pada taraf signifikansi 10%. Sementara dalam model *Zero Inflation*, hanya variabel Persentase Pelayanan Ibu Nifas (X_3) yang memberikan pengaruh terhadap variabel Kematian Ibu pada taraf signifikansi 10%. Oleh karena itu, model *count* regresi ZINB akhir yang terbentuk yaitu:

$$\ln(\mu_i) = -0,0569X_1 + 0,0646X_3$$

Sehingga dapat digambarkan bahwa apabila terjadi peningkatan persentase K1 (X_1) pada ibu hamil sebanyak 1%, maka akan menurunkan rata-rata kejadian kematian ibu sebanyak $\exp(-0,0569) = 0,9447 \approx 1$ kali lipat dari sebelumnya dengan asumsi nilai variabel persentase pelayanan ibu nifas (X_3) dianggap konstan. Dan apabila terjadi peningkatan persentase pelayanan ibu nifas (X_3) sebanyak 1%, maka akan meningkatkan rata-rata kejadian kematian ibu sebanyak $\exp(0,0646) = 1,0667 \approx 1$ kali lipat dari sebelumnya dengan asumsi nilai variabel persentase K1 (X_1) dianggap konstan.

Sementara itu, untuk model *zero inflation* regresi ZINB akhir yang terbentuk yaitu $\text{logit}(\pi_i) = 0,1481X_3$. Sehingga dapat digambarkan bahwa apabila terjadi peningkatan persentase pelayanan ibu nifas (X_3) sebanyak 1%, maka akan meningkatkan *odds zero inflaton* pada kasus kematian ibu sebanyak $\exp(0,1481) = 1,1596 \approx 1$ kali lipat dari sebelumnya. Hasil pemodelan menunjukkan kenaikan persentase pelayanan ibu nifas (X_3) sebesar 1%, maka akan meningkatkan rata-rata kejadian dan *odds zero inflation* pada kasus kematian ibu. Hal ini mungkin disebabkan bahwa peningkatan pelayanan ibu nifas tidak menjamin kualitas dari pelayanan yang diberikan. Selain itu, wilayah dengan cakupan pelayanan ibu nifas yang tinggi menunjukkan banyaknya kasus komplikasi pada ibu dengan risiko kematian yang tinggi, sehingga meskipun pelayanan yang diberikan meningkat tetapi angka kematian tetap tinggi akibat risiko kematian.

D. Kesimpulan

Berdasarkan hasil analisis dalam penelitian ini, berdasarkan hasil dari uji eksistensi *zero inflation* menggunakan uji *Naive* diperoleh hasil yang signifikan di bawah taraf arti $\alpha = 10\%$. Sehingga diperoleh kesimpulan bahwa terjadinya *excess zero* atau *zero inflation* dalam data kematian ibu di Kota Bandung tahun 2020 yang menunjukkan bahwa model regresi Poisson standar tidak cukup layak untuk memodelkan data, sehingga perlu dilakukan pemodelan menggunakan metode alternatif yakni regresi *Zero Inflation Negative Binomial*. Regresi Poisson mengasumsikan varians sama dengan rata-rata pada data amatan sehingga terjadinya overdispersi dalam data mengakibatkan *standard error* dalam model menjadi sangat kecil (*underestimated*) dan membuat hipotesis nol mudah ditolak. Berdasarkan hasil pemodelan pada data kematian ibu di Kota Bandung tahun 2020 menggunakan regresi *Zero Inflation Negative Binomial*, diperoleh keputusan bahwa model regresi *Zero Inflation Negative Binomial* mampu mengatasi dampak overdispersi dengan peningkatan nilai *standard error* pada setiap parameter dalam model.

Ucapan Terimakasih

Penulis mengucapkan syukur kepada Allah SWT atas rahmat dan hidayah-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Ucapan terima kasih yang tulus penulis sampaikan kepada kedua orang tua atas segala doa, dukungan, semangat, dan motivasi yang tiada henti serta menjadi kekuatan tersendiri bagi penulis. Selain itu, terkhusus kepada Bapak Dr. Nusar Hajarisman, S.Si., M.Si., selaku pembimbing akademik yang telah memberikan bimbingan dan pengarahan kepada penulis. Tak lupa, penulis juga berterima kasih kepada teman-teman yang telah memberikan semangat, bantuan, dan motivasi selama penelitian.

Daftar Pustaka

- Astuti, C. C., & Mulyanto, A. D. (2016). Estimation Parameters And Modelling Zero Inflated Negative Binomial. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, 4(3), 115–119. <https://doi.org/10.18860/ca.v4i3.3656>
- Dinkes Kota Bandung. (2021). *Profil Kesehatan Kota Bandung Tahun 2020*. Dinas Kesehatan Kota Bandung.

- Famoye, F., Jr., J. T. W., & Singh, K. P. (2004). On the Generalized Poisson Regression Model with an Application to Accident Data. *Journal of Data Science*, 2, 287–295. [https://doi.org/10.6339/jds.2004.02\(3\).167](https://doi.org/10.6339/jds.2004.02(3).167)
- He, H., Zhang, H., Ye, P., & Tang, W. (2019). A test of inflated zeros for Poisson regression models. *Stat Methods Med Res.*, 28(4). <https://doi.org/10.1177/0962280217749991>
- Jannah, S. E., Muna, N., Rachmawati, E., & Roziqin, M. C. (2024). Sistem Informasi Geografis Angka Kematian Ibu dalam Upaya Surveilans Kesehatan di Kabupaten Bondowoso. *J-REMI: Jurnal Rekam Medik Dan Informasi Kesehatan*, 6(1), 88–99. <https://doi.org/10.25047/j-remi.v6i1.5120>
- Khoeriyah, R. Y., & Hajarisman, N. (2021). Regresi Terboboti Geografis Semiparametrik (RTG-S) untuk Pemodelan Indeks Pembangunan Kesehatan Masyarakat Kabupaten/Kota di Sumatera Utara. *Jurnal Riset Statistika*, 1(1), 43–50. <https://doi.org/10.29313/jrs.v1i1.145>
- Payne, E. H., Gebregziabher, M., Hardin, J. W., Ramakrishnan, V., & Egede, L. E. (2018). An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Communications in Statistics: Simulation and Computation*, 47(6), 1722–1738. <https://doi.org/10.1080/03610918.2017.1323223>
- Prianggada, Q. W. (2017). *Pemilihan Regresi Zero-Inflated Poisson (ZIP) dan Regresi Zero-Inflated Binomial Negatif (ZIBN) pada Angka Kematian Ibu Kabupaten Bojonegoro Tahun 2016 (Data Mengalami Overdispersi)* (Vol. 11).
- Rachmawati, K. R. (2018). *Perbandingan Regresi Zero Inflated Poisson dan Zero Inflated Negative Binomial Pada Data Overdispersi*.
- Saefuddin, A., & Setiabudi, N. A. (2011). The Effect of Overdispersion on Logistic Regression Analysis of Poverty in Indonesia. *International Journal for Statisticians*, 2(1), 1–4.
- Saputri, E. M. (2020). Asuhan Kebidanan Ibu Nifas Pada 6 Jam S/D 6 Hari Postpartum. *Jurnal Komunikasi Kesehatan*, XI(1), 85–96.
- Tendriyawati, Wibawa, G. N. A., & Abapihi, B. (2023). Pemodelan Regresi Poisson Terhadap Faktor-faktor Yang Mempengaruhi Terjadinya Hipertensi Di Kota Kendari. *Jurnal Matematika, Komputasi Dan Statistika*, 3, 255–262.
- Wildan, & Karyana, Y. (2021). Evaluasi Kesalahan Proyeksi Penduduk Tahun 2020 untuk Memproyeksikan Penduduk Tahun 2025 Provinsi Jawa Barat. *Jurnal Riset Statistika*, 1(2), 92–98. <https://doi.org/10.29313/jrs.v1i2.407>
- Zahara, A., & Kurniati, E. (2024). Penerapan Ekstrim Fungsi untuk Menentukan Profit Maksimum dalam Pasar Persaingan Sempurna untuk Short-Run dan Long-Run. *DataMath: Journal of Statistics and Mathematics*, 2(1), 33–40.