Implementasi *Zero Inflated Beta Regression Model* pada Proporsi Kematian Ibu di Kota Bandung Tahun 2020

Labana Kaulika*, Nusar Hajarisman

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

*labanakaulika22@gmail.com, nusarhajarisman@unisba.ac.id

Abstract. Zero inflated beta is a mixture of the continous distribution on (0, 1) and the generated distribution which can produce a non-negative probability to 0. Zero Inflated Beta Regression (BeZI) is a method that can handle or model data that has a high proportion of zeros or there are excess zeros in the data. In this study, the response variable y has a mix between the beta distribution and the point mass at zero. Estimation of the regression parameters from the zero inflated beta regression model uses the Maximum Likelihood Estimation (MLE), where the estimation process is solved numerically. The numerical method used is the Fisher's scoring method based on the score vector and the Fisher Information matrix to estimate the parameters of the maternal mortality rate in the city of Bandung in 2020. The results of the research on the count regression model show that the percentage variable K1 has a negative effect on the proportion of maternal deaths in Bandung City in 2020, while in the zero inflation model it is found that there are no variables that have an influence on the proportion when maternal deaths do not occur.

Keywords: *Maximum Likelihood Estimation, Proportion of Maternal Mortality, Zero Inflated Beta Regression* .

Abstrak. *Zero inflated beta* merupakan campuran distribusi kontinu pada (0, 1) dan distribusi yang dibangkitkan dimana dapat menghasilkan probabilitas non-negatif ke 0. *Zero Inflated Beta Regression* (BeZI) merupakan metode yang dapat menangani atau memodelkan suatu data yang memiliki proporsi nol yang tinggi atau terdapat *excess zero* dalam data. Dalam skripsi ini, variabel respon y memiliki percampuran antara distribusi beta dan titik massa pada nol. Penaksiran parameter regresi dari model regresi *zero inflated beta* menggunakan *Maximum Likelihood Estimation* (MLE), dimana proses penaksirannya diselesaikan secara numerik. Metode numerik yang digunakan yaitu metode *Fisher's scoring* berdasarkan pada vektor skor dan matriks informasi Fisher untuk menaksir parameter dari angka kematian ibu di kota bandung tahun 2020. Hasil penelitian pada model *count regression* diperoleh bahwa variabel persentase K1 memiliki pengaruh negatif terhadap proporsi kematian ibu Kota Bandung tahun 2020, sedangkan pada model *zero inflation* diperoleh bahwa tidak ada variabel yang memiliki pengaruh terhadap proporsi pada saat tidak terjadinya kematian ibu.

Kata Kunci: *Maximum Likelihood Estimation, Proporsi Kematian Ibu, Zero Inflated Beta Regression.*

Corresponding Author
Email: nusarhajarisman@unisba.ac.id

A. Pendahuluan

Analisis regresi klasik adalah metode untuk menentukan hubungan sebab akibat suatu variabel dengan variabel lainnya. Analisis regresi klasik ini salah satu analisis yang paling popular dan luas dalam pemakaiannya. Namun, penggunaan model regresi klasik ini tidak sepenuhnya dapat diterapkan pada bidang penelitian tertentu [1].

Apabila dihadapkan dengan variabel respon yang berbentuk proporsi, jika menggunakan regresi klasik untuk menangani kasus ini maka hal ini bukan solusi yang tepat untuk menangani masalah tersebut karna masalah ketidaklinieran, di mana model regresi linier klasik memberikan estimasi proporsi di luar interval $(0, 1)$, dan masalah heteroskedastisitas di mana varians dari proporsi tersebut memiliki sifat yang tidak konstan. Respon rasio akan berdistribusi beta, karena model regresi beta adalah model yang dibandingkan dengan regresi klasik, memberikan penaksir parameter yang akurat dan efisien ketika variabel respon yang diamati memiliki distribusi yang tidak simetris atau ketika terdapat masalah heteroskedastisitas.

Distribusi beta memiliki berbagai bentuk yang berbeda tergantung pada nilai dua parameter, model regresi beta digunakan ketika variabel responnya kontinu dan terbatas pada interval $(0, 1)$ [2]. Distribusi beta sangat fleksibel dan dapat digunakan untuk memodelkan berbagai fenomena ketidakpastian. Pada kasus tertentu, dalam memodelkan data dengan variabel respon berbentuk proporsi ini juga ditemukan suatu kondisi dimana peristiwa yang diamati terlalu jarang terjadi sehingga memberikan variabel respon yang terlalu banyak bernilai nol. Keadaan seperti ini seringkali disebut juga sebagai *zero inflation*. Distribusi beta tidak berlaku ketika terdapat data yang banyak dalam batas-batas pengamatan atau memiliki data yang lebih banyak nol. Jika jumlah nol tidak dapat diabaikan dan nilai nol berlebih yang menjadi perhatian khusus untuk analisis, mengubah data asli bukan solusi yang paling tepat [3]. Penerapan model zero-inflated dalam model regresi sudah mulai dikembangkan, dimana berdasarkan asumsi variabel respon mengikuti distribusi beta yaitu disebut dengan *Zero Inflated Beta Regression Model* (BeZI). Ospina dan Ferrari (2010) dalam penelitiannya mereka mengatakan untuk data yang diamati pada $[0, 1)$ mereka menggunakan campuran distribusi kontinu pada $(0, 1)$ dan distribusi yang dibangkitkan dimana dapat menghasilkan probabilitas non-negatif ke 0.

Metode ini akan di implementasikan pada kasus kematian ibu di Kota Bandung tahun 2020 yang dimana pada data kematian banyak mengandung nilai nol berlebih. Salah satu masalah terbesar di Indonesia yang masih menjadi sorotan pemerintah adalah tingginya angka kematian ibu hamil. Berdasarkan data dari Dinas Kesehatan (Dinkes) Kota Bandung, pada tahun 2019 angka kematian ibu terdapat 29 kasus dan mengalami penurunan pada tahun 2020 yaitu berjumlah 28 kasus [14].

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana penerapan metode regresi *Zero Inflated Beta* (BeZI) pada kasus kematian ibu di Kota Bandung tahun 2020?” dan “Bagaimana pengaruh dari masing-masing variabel prediktor terhadap angka kematian ibu di Kota Bandung tahun 2020?”. Selanjutnya, tujuan dalam penelitian ini sebagai berikut: “Untuk mengetahui penerapan metode regresi *Zero Inflated Beta* (BeZI) pada kasus kematian ibu di Kota Bandung tahun 2020” dan “Untuk mengetahui pengaruh dari masing-masing variabel prediktor terhadap angka kematian ibu di Kota Bandung tahun 2020”.

B. Metodologi Penelitian

Distribusi *Zero Inflated Beta* (BeZI)

Dalam teori probabilitas dan statistik, distribusi beta adalah distribusi probabilitas yang didefinisikan pada interval (0, 1) oleh dua parameter positif, bentuk parameter biasanya dilambangkan dengan a dan b.

Fungsi densitas dari suatu variabel acak Y yang berdistribusi beta, dengan parameter a dan b masing-masing adalah [4]:

$$f(y; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} \qquad (1)$$

dimana $0 < y < 1$, $a > 0$ dan $b > 0$ dan symbol Γ menunjukkan fungsi Gamma.

Untuk memodelkan regresi beta dengan menyertakan rata-rata respon bersamaan dengan parameter dispersinya, maka perlu dilakukan reparameterisasi dari fungsi densitas beta. Misalkan parameter baru yaitu $\pi = a/(a+b)$, sehingga $0 < \pi < 1$ dan misalkan $\phi = a + b$ maka $\phi > 0$, sehingga substitusi kebalikannya yaitu $a = \phi\pi$ dimana $a > 0$ dan $b = \phi(1-\pi)$ dimana $b > 0$. Dengan mempertimbangkan parameterisasi baru, fungsi densitas peluang dari distribusi beta yaitu $B(a, b)$ dinotasikan menjadi $B(\pi, \phi)$ dan dapat dituliskan sebagai berikut [5]:

$$f_{Be}(y; \pi, \phi) = \frac{\Gamma(\phi)}{\Gamma(\pi\phi)\Gamma\big((1-\pi)\phi\big)} y^{\pi\phi-1}(1-y)^{(1-\pi)\phi-1} \tag{2}$$

dimana:
$0 < y < 1$, $\pi \in (0,1)$, $\phi > 0$
π = parameter lokasi
ϕ = parameter dispersi
Γ = fungsi Gamma
y = variabel respon

Dengan nilai harapan, dan variannya, dapat dilihat pada persamaan berikut:

$$E(Y) = \pi \text{ dan } V(Y) = \frac{\pi(1-\pi)}{(\phi+1)} \tag{3}$$

Jika respon memiliki nilai nol yang berlebih dan pula memiliki nilai dalam interval (0,1), maka untuk memodelkan data dalam situasi tersebut menggunakan campuran dari dua distribusi yaitu distribusi beta dan distribusi yang dibangun dari nilai c dimana dalam kasus ini $(c = 0)$ distribusi ini disebut distribusi *Zero Inflated Beta*.

Selanjutnya fungsi densitas dari $BeZI(\delta, \pi, \phi)$ yaitu dapat dituliskan sebagai berikut:

$$f_{BeZI}(y; \delta, \pi, \phi) = \begin{cases} \delta, & jika\ y = 0 \\ [(1-\delta) f_{Be}(y; \pi, \phi)], & jika\ y \in (0,1) \end{cases} \tag{4}$$

dimana:
$0 \le y < 1$, $\pi \in (0,1)$, $\phi > 0$
$f_{Be}(y; \pi, \phi)$ = fungsi densitas distribusi beta
π = parameter lokasi
ϕ = parameter dispersi
δ = peluang sukses yang terjadi pada suatu daerah dimana $y = 0$
$(1-\delta)$ = peluang gagal yang terjadi pada suatu daerah dimana $y \in (0,1)$
y = variabel respon

Berikut merupakan rata-rata dan varians dari distribusi $BeZI(\delta, \mu, \phi)$:

$$E(Y) = (1-\delta)\pi \text{ dan } Var(Y) = (1-\delta)\left[\frac{\pi(1-\pi)}{\phi+1} + \delta\pi^2\right] \tag{5}$$

Excess Zero

Excess Zero dikenal sebagai data pada variabel *dependent* yang memiliki nilai nol berlebih. *Excess Zero* terjadi saat variabel respon dari suatu data memiliki nilai nol lebih banyak dibandingkan nilai data diskrit lain, dimana spesifikasinya $> 50\%$ dari variabel respon berisikan nilai nol [6].

Definisi Model dan Taksiran Parameter Regresi BeZI

Misalkan $y_1, \ldots, y_n$ adalah variabel respon acak sehingga setiap y_i, untuk $i = 1, \ldots, n$, memiliki fungsi densitas peluang pada persamaan (2.11) dengan parameter $\delta = \delta_i$, $\pi = \pi_i$, dan $\phi = \phi_i$. Diasumsikan bahwa $\delta = \delta_i$, $\pi = \pi_i$, dan $\phi = \phi_i$ didefinisikan sebagai fungsi hubung logit dari parameter δ, π, dan ϕ dapat dimodelkan menjadi [5]:

ISSN: 2828-206X

$$\begin{aligned} h_1(\pi_i) &= logit(\pi_i) = f_1(x_i, \beta) = x_i^\top \beta \\ h_2(\phi_i) &= log(\phi_i) = f_2(z_i, \rho) = z_i^\top \rho \\ h_3(\delta_i) &= logit(\delta_i) = f_3(v_i, \gamma) = v_i^\top \gamma \end{aligned} \tag{6}$$

dimana:

1. $\beta = (\beta_1, \dots, \beta_k)^\top$, $\rho = (\rho_1, \dots, \rho_m)^\top$, dan $\gamma = (\gamma_1, \dots, \gamma_p)^\top$ merupakan vektor dari parameter model regresi BeZI
2. $x_i = (x_{i1}, \dots, x_{ik})$, $z_i = (z_{i1}, \dots, v_{im})$, dan $v_i = (v_{i1}, \dots, v_{ip})$ adalah variabel prediktor masing-masing parameter

Inferensi *maximum likelihood* untuk (π, ϕ) dapat dilakukan secara terpisah dari parameter δ, dapat dimisalkan seolah-olah nilai δ diketahui dan sebaliknya [7]. Berikut merupakan fungsi *likelihood* untuk parameter $\theta = (\gamma^\top, \beta^\top, \rho^\top)^\top$ berdasarkan sampel pengamatan yang saling bebas [8]:

$$L(\theta) = \prod_{i=1}^{n} f_{BeZI}\,(y_i; \delta, \pi, \phi) = L_1(\delta) L_2(\pi, \phi) \tag{7}$$

dimana:

$$L_1(\delta) = \prod_{i=1}^{n} \delta^{\mathbb{1}_{\{c\}}(y_i)} (1-\delta)^{1-\mathbb{1}_{\{c\}}(y_i)}\ ;\ L_2(\pi, \phi) = \prod_{i=1}^{n} f(y_i; \pi, \phi)^{1-\mathbb{1}_{\{c\}}(y_i)}$$

dimana $\mathbb{1}_{\{c\}}$ = fungsi indikator yang sama dengan 1 jika $y_i \in A$ dan 0 jika $y_i \notin A$.

Selanjutnya fungsi *log-likelihood* dapat dituliskan sebagai berikut:

$$l(\theta) = \log\big(L(\theta)\big) = l_1(\delta) + l_2(\pi, \phi) = \sum_{i=1}^{n} l_i(\delta_i) + \sum_{i: y_i \in (0,1)} l_i(\pi_i, \phi_i) \tag{8}$$

dimana:

$$l_1(\delta_i) = \mathbb{1}_{\{c\}}(y_i) \log \delta_i + \Big(1 - \mathbb{1}_{\{c\}}(y_i)\Big) \log(1 - \delta_i)$$

$$l_2(\pi_i, \phi_i) = \log \Gamma(\phi_i) - \log \Gamma(\pi_i, \phi_i) - \log\big((1-\pi_i)\phi_i\big) + (\pi_i\, \phi_i - 1) y_i^* + (\phi_i - 2) y_i^\dagger$$

dimana:

y_i = pengamatan ke-i pada variabel respon dengan $i = 1, 2, \dots, n$

$y_i^\dagger$ = $\log(1 - y_i)$

y_i^* = $\log\{y_i/(1 - y_i)\}$

Pada kenyataannya MLE tidak dapat dilakukan secara langsung, hal ini terjadi karena fungsi *log-likelihood* pada distribusi BeZI yang *non linear*. Oleh karena itu, untuk mengatasi kondisi tersebut dapat diselesaikan dengan menggunakan metode numerik yaitu salah satunya adalah *Fisher's scoring*.

Dalam *Fisher's scoring* dibutuhkan vektor skor, untuk vektor skor turunan pertama dari fungsi *log-likelihood* pada parameter $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$, dan $\boldsymbol{\rho}$ diberikan oleh:

$$U_s = \frac{\partial l(\theta)}{\partial \gamma_s} = \sum_{i=1}^{n} \frac{\partial l_i(\delta_i)}{\partial \delta_i} \frac{\partial \delta_i}{\partial \eta_{1i}} \frac{\partial \eta_{1i}}{\partial \gamma_s} = \sum_{i=1}^{n} \frac{\mathbb{1}_{\{c\}}(y_i) - \delta_i}{\delta_i(1-\delta_i)} \frac{1}{h_1'(\delta_i)} \frac{\partial \eta_{1i}}{\partial \gamma_s}$$

$$U_r = \frac{\partial l(\theta)}{\partial \beta_r} = \sum_{i: y_i \in (0,1)} \frac{\partial l_i(\pi_i, \phi_i)}{\partial \pi_i} \frac{\partial \pi_i}{\partial \eta_{2i}} \frac{\partial \eta_{2i}}{\partial \beta_r} = \sum_{i: y_i \in (0,1)} \phi_i (y_i^* - \pi_i^*) \frac{1}{h_2'(\pi_i)} \frac{\partial \eta_{2i}}{\partial \beta_r}$$

$$U_R = \frac{\partial l(\theta)}{\partial \rho_R} = \sum_{i: y_i \in (0,1)} \frac{\partial l_i(\pi_i, \phi_i)}{\partial \phi_i} \frac{\partial \phi_i}{\partial \eta_{3i}} \frac{\partial \eta_{3i}}{\partial \rho_R} = \sum_{i: y_i \in (0,1)} \left[\pi_i (y_i^* - \pi_i^*) + \left(y_i^\dagger - \pi_i^\dagger\right)\right] \frac{1}{h_3'(\phi_i)} \frac{\partial \eta_{3i}}{\partial \rho_R}$$

dimana:

$\pi_i^\dagger = \log(1 - \pi_i)$

$\pi_i^* = \log\{\pi_i/(1 - \pi_i)\}$

untuk $s = 1, \dots, p$, $r = 1, \dots, q$, dan $R = 1, \dots, m$

Tahapan selanjutnya masuk pada tahap maksimalisasi. Memaksimalkan $\boldsymbol{\beta}$ dan $\boldsymbol{\gamma}$ pada persamaan dengan menghitung $\boldsymbol{\beta}^{(i+1)}$ dan $\boldsymbol{\gamma}^{(i+1)}$ dengan metode *Fisher's scoring*. Iterasi dilakukan sampai diperoleh penduga parameter yang konvergen, dikatakan konvergen jika $\left|\boldsymbol{\beta}^{(i+1)} - \boldsymbol{\beta}^{(i)}\right|$ dan $\left|\boldsymbol{\gamma}^{(i+1)} - \boldsymbol{\gamma}^{(i)}\right|$ sangat kecil.

Model Taksiran Regresi BeZI
Model taksiran pada regresi BeZI terdapat dua model yaitu model taksiran pada saat proporsi berada dalam selang (0, 1) atau disebut model *count regression* dapat dituliskan seperti persamaan berikut [9]:

$$\hat{\pi}_i = \frac{\exp\{\hat{\beta}_0+\hat{\beta}_1 x_{m+1}+\cdots+\hat{\beta}_{k-m} x_k\}}{1+\exp\{\hat{\beta}_0+\hat{\beta}_1 x_{m+1}+\cdots+\hat{\beta}_{k-m} x_k\}} \text{ atau } logit(\hat{\pi}_i) = \log\left(\frac{\hat{\pi}_i}{1-\hat{\pi}_i}\right) = \boldsymbol{\beta X} \quad (9)$$

Selanjutnya model taksiran pada saat proporsi sama dengan nol atau disebut model *zero inflation* dituliskan seperti persamaan berikut:

$$\hat{\delta}_i = \frac{\exp\{\hat{\gamma}_0+\hat{\gamma}_1 x_1+\cdots+\hat{\gamma}_m x_m\}}{1+\exp\{\hat{\gamma}_0+\hat{\gamma}_1 x_1+\cdots+\hat{\gamma}_m x_m\}} \text{ atau } logit(\hat{\delta}_i) = \log\left(\frac{\hat{\delta}_i}{1-\delta_i}\right) = \boldsymbol{\gamma X} \quad (10)$$

Pemilihan Model Terbaik
Dalam memilih model terbaik, dapat digunakan metode *Akaike Information Criterion* (AIC) yang ditemukan oleh Akaike. Untuk mendapatkan nilai AIC digunakan dengan rumus sebagai berikut:

$$AIC = 2k - 2\ln L(\theta)$$

Dengan k merupakan jumlah parameter yang diestimasi dalam model. Model regresi terbaik merupakan model regresi yang memiliki nilai AIC paling kecil [10].

Uji Signifikansi Parameter Secara Simultan
Untuk mengetahui signifikansi parameter regresi secara serempak atau bersama-sama dapat dilakukan dengan cara pengujian parameter simultan dengan menggunakan statistik uji G seperti berikut [11]:

$$G = -2\ln[L_0 - L_1] \sim \chi^2_{p(\alpha)}$$

dimana:
L_0 = Model intersep atau ln *likelihood* model tanpa semua variabel prediktor
L_1 = Model model penuh atau ln *likelihood* model dengan k variabel prediktor

Dengan kriteria, jika statistik uji $G > \chi^2_{(\alpha,k)}$, maka hipotesis nol ditolak dengan α adalah tingkat signifikansi yang digunakan dalam penelitian dan k merupakan jumlah variabel prediktor.

Uji Signifikansi Parameter Secara Parsial
Berbeda dengan uji simultan, uji parsial adalah uji parameter pada model secara individual yang terdiri dari masing-masing variabel prediktor. Terdapat dua parameter yang masing-masing harus uji yaitu β_j dan γ_j. Statistik uji yang digunakan adalah statistik uji t [12]:

Pengujian parsial parameter pada masing-masing model *count regression* dan *zero inflation* di hipotesiskan sebagai berikut:
$H_0 : \beta_j = 0$ atau $\gamma_j = 0 \ vs \ H_1 : \ \beta_j \neq 0 \ atau \ \gamma_j \neq 0, j = 1,2, \ldots q/p$
dimana:
$\hat{\beta}_j$ dan $\hat{\gamma}_j$: nilai estimasi parameter variabel prediktor ke-j pada masing-masing model
$se(\hat{\beta}_j)$ dan $se(\hat{\gamma}_j)$: nilai standar *error* dari estimasi variabel prediktor ke-j pada masing-masing model
dengan statistik uji:

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \ ; \ t = \frac{\hat{\gamma}_j}{se(\hat{\gamma}_j)} \text{ dengan } se(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)} \ ; \ se(\hat{\gamma}_j) = \sqrt{Var(\hat{\gamma}_j)}$$

Kedua model pada regresi BeZI memiliki kriteria jika nilai $|t| > t_{(\frac{\alpha}{2};n-k-1)}$ maka H_0 ditolak.

Angka Kematian Ibu
Kematian ibu menurut WHO adalah kematian selama kehamilan atau dalam periode 42 hari

setelah berakhirnya kehamilan, akibat semua sebab yang terkait dengan atau diperberat oleh kehamilan atau penanganannya, tetapi bukan disebabkan oleh kecelakaan/cedera.

Penyebab kematian ibu diakibatkan oleh komplikasi kehamilan, persalinan dan nifas yang disebabkan oleh intervensi, kegagalan, penanganan yang tidak tepat atau rangkaian semua peristiwa tersebut. Penyebab tidak langsung dari kematian ibu yaitu masih banyaknya kasus 3 terlambat. Kasus 3 terlambat, meliputi: terlambat mengenali tanda bahwa persalinan dan mengambil keputusan, terlambat dirujuk ke fasilitas kesehatan, terlambat ditangani oleh tenaga kesehatan di fasilitas pelayanan kesehatan [13].

Data yang akan digunakan yaitu data Angka Kematian Ibu pada setiap Puskesmas Kecamatan di Kota Bandung tahun 2020, berikut variabel yang akan digunakan dalam penelitian ini:

Tabel 1. Variabel yang digunakan dalam penelitian

Variabel	Nama Variabel	Keterangan
Y	Proporsi Kematian Ibu	Proporsi kematian ibu yang berhubungan dengan kehamilan atau persalinan
X_1	Presentase K1 (%)	Persentase ibu hamil yang menerima perawatan prenatal sesuai standar pada trimester pertama kehamilan
X_2	Persentase K4 (%)	Persentase ibu hamil yang menerima perawatan prenatal sesuai standar minimal empat kali
X_3	Persentase Pelayanan Ibu Nifas (%)	Persentase dari tenaga kesehatan yang memberikan pelayanan kesehatan kepada ibu melahirkan dalam masa nifas
X_4	Persentase Komplikasi Kebidanan yang Ditangani (%)	Persentase komplikasi kehamilan yang menerima perawatan sesuai standar kebidanan oleh tenaga kesehatan terlatih pada tingkat perawatan primer

C. Hasil Penelitian dan Pembahasan

Excess Zero

Berdasarkan data angka kematian ibu di Kota Bandung tahun 2020, banyaknya amatan bernilai nol yaitu sebesar 62 dengan banyaknya data yaitu sebesar 80. Maka dari itu, persentase nilai nol pada variabel respon adalah 77,5%, nilai tersebut lebih besar dari syarat terjadinya *excess zero* yaitu 50%. Maka dapat diambil kesimpulan bahwa terjadi *excess zero* pada variabel respon yaitu Proporsi Kematian Ibu pada data Angka Kematian Ibu di Kota Bandung tahun 2020.

Evaluasi Model

Sebelum melakukan pemodelan dalam metode BeZI ini akan dilakukan pemilihan dengan mempertimbangkan kombinasi variabel yang cocok untuk model *count regression* dan *zero inflation* dengan melakukan pemilihan model terbaik yang dimana model dengan nilai AIC yang terkecil ini akan dijadikan sebagai model yang akan dianalisis pada tahap selanjutnya.

Dari hasil yang diperoleh, model regresi yang memiliki nilai AIC paling kecil yaitu $logit(\pi_i) = X1 + X3$ dan $logit(\delta_i) = X2 + X4$ dengan nilai AIC sebesar $-77{,}1$. Maka dari itu, model tersebut merupakan model terbaik. Selanjutnya, model tersebut akan dilakukan pengujian secara simultan dan parsial untuk melihat pengaruh dari variabel prediktor terhadap masing-masing model.

Tabel 2. Hasil Estimasi Parameter Model Regresi BeZI

Parameter	Model Count Regression				
	Koefisien	**SE**	$\|t\|$	$t_{0,05;75}$	**Keputusan**
β_0	-3,9539	0,7489	5,28	1,66543	Tolak H_0
β_1	-0,03345	0,01859	1,80		Tolak H_0
β_3	0,01524	0,01816	0,84		Terima H_0
ϕ	638,33	224,84			
	Model Zero Inflation				
	Koefisien	**SE**	$\|t\|$	t_{tabel}	**Keputusan**
γ_0	2,7509	1,4288	1,93	1,66543	Tolak H_0
γ_2	-0,01121	0,01696	0,66		Terima H_0
γ_4	-0,01026	0,008992	1,14		Terima H_0

Berdasarkan hasil perhitungan Tabel 2 diperoleh nilai-nilai estimasi untuk model *count regression* dan *zero inflation* masing-masing adalah:

$logit(\hat{\pi}_i) = -3,9539 - 0,03345X_1 + 0,01524X_3$

$logit(\hat{\delta}_i) = 2,7509 - 0,01121X_2 - 0,01026X_4$

Selanjutnya akan dilakukan uji simultan, uji simultan digunakan untuk menguji apakah variabel prediktor berpengaruh signifikan terhadap variabel respon secara bersamaan, hipotesis yang digunakan adalah

$H_0 : \beta_1 = \beta_3 = \gamma_2 = \gamma_4 = 0$

H_1 : minimal terdapat satu $\beta_j \neq 0$ atau $\gamma_j \neq 0$, untuk $j = 1,2,3,4$

Statistik uji yang digunakan adalah statistik G, berdasarkan output dari bantuan *software* SAS *OnDemand for Academics*, diperoleh nilai $G_{hitung} = 19,7 > \chi^2_{0,1;4} = 7,779$ maka tolak H_0, dengan taraf nyata 10% yang berarti terdapat minimal satu variabel prediktor yang berpengaruh signifikan terhadap proporsi kasus kematian ibu di Kota Bandung tahun 2020.

Kemudian untuk uji parsial akan menggunakan nilai statistik t yang terlampir pada Tabel 2. Uji parsial adalah uji parameter pada model secara individu yang terdiri dari masing-masing variabel prediktor. Pertama, uji parameter untuk model *count regression* dengan hipotesis:

$H_0 : \beta_j = 0$, dimana $j = 1,3$; variabel prediktor ke-j tidak berpengaruh secara signifikan terhadap proporsi kematian ibu.

$H_1 : \beta_j \neq 0$, dimana $j = 1,3$; variabel prediktor ke-j berpengaruh secara signifikan terhadap proporsi kematian ibu.

Dengan menggunakan $\alpha = 10\%$ dan $db = n - k - 1 = 80 - 4 - 1 = 75$, diperoleh $t_{\frac{\alpha}{2};n-k-1)} = t_{\frac{0,1}{2};75} = t_{0,05;75} = 1,66543$. Berdasarkan Tabel 2 untuk model *count regression* diperoleh variabel yang berpengaruh signifikan terhadap kasus kematian ibu di kota Bandung tahun 2020 adalah variabel X_1 (Persentase K1). Sedangkan untuk variabel X_3 (Persentase Pelayanan Ibu) tidak berpengaruh signifikan terhadap kasus kematian ibu di Kota Bandung tahun 2020.

Selanjutnya, uji parameter untuk model *zero inflation* dengan hipotesis:

$H_0 : \gamma_j = 0$, dimana $j = 2,4$; variabel prediktor ke-j tidak berpengaruh secara signifikan terhadap proporsi pada saat tidak terjadinya kematian ibu.

$H_1 : \gamma_j \neq 0$, dimana $j = 2,4$; variabel prediktor ke-j berpengaruh secara signifikan terhadap proporsi pada saat tidak terjadinya kematian ibu.

Dengan menggunakan $\alpha = 10\%$ dan $db = n - k - 1 = 80 - 4 - 1 = 75$, diperoleh $t_{\frac{\alpha}{2};n-k-1)} = t_{\frac{0,1}{2};75} = t_{0,05;75} = 1,66543$. Berdasarkan Tabel 2 untuk model *zero inflation* diperoleh bahwa variabel X_2 (Persentase K4) dan variabel X_4 (Persentase Komplikasi Kebidanan yang Ditangani) tidak berpengaruh signifikan terhadap kasus tidak terjadinya kematian ibu di kota Bandung tahun 2020.

Berdasarkan hasil uji parameter simultan dan parsial, maka hasil dari estimasi parameter model *zero inflation beta* diperoleh. Model *count regression* dinyatakan sebagai berikut:

$$logit(\hat{\pi}_i) = \beta_0 + \beta_1 X_1 + \beta_3 X_3$$
$$logit(\hat{\pi}_i) = -3{,}9539 - 0{,}03345 X_1 + 0{,}01524 X_3$$

Berdasarkan hasil yang dilakukan pada pengujian parsial, parameter yang signifikan yaitu parameter β_0 dan β_1. Untuk nilai konstanta β_0 sebesar -3,9539, parameter tersebut dapat diinterpretasikan bahwa apabila variabel persentase K1 (X1) dan variabel persentase pelayanan ibu nifas (X3) bernilai nol, maka nilai peluang angka kematian ibu di tiap puskesmas kota Bandung berkurang sebesar $\exp(-3{,}9539) = 0{,}01917975 \approx 0{,}0192$.

Kemudian variabel (X_1) dengan nilai koefisien sebesar -0,03345 bernilai negatif yang menyatakan bahwa apabila suatu puskesmas terjadi peningkatan persentase pelayanan antenatal ibu hamil sesuai standar pada triwulan pertama kehamilan sebesar satu persen dengan asumsi nilai variabel X_3 dianggap konstan, maka peluang terjadinya kematian ibu akan turun sebesar $\exp(-0{,}03345) = 0{,}9671033 \approx 0{,}9671$ kali lipat dibandingkan dengan ibu hamil yang mendapatkan persentase pelayanan sesuai dengan standar pada triwulan pertama kehamilan pada suatu puskesmas satu satuan lebih rendah.

Model *zero inflation* dinyatakan sebagai berikut:

$$logit(\hat{\delta}_i) = \gamma_0 + \gamma_2 X_2 + \gamma_4 X_4$$
$$logit(\hat{\delta}_i) = 2{,}7509 - 0{,}01121 X_2 - 0{,}01026 X_4$$

Berdasarkan hasil yang dilakukan pada pengujian parsial, parameter yang signifikan yaitu hanya parameter intersep γ_0. Untuk konstanta γ_0 sebesar 2,7509, parameter tersebut dapat diinterpretasikan bahwa apabila variabel persentase K4 (X2) dan variabel persentase komplikasi kebidanan yang ditangani (X4) bernilai nol, maka nilai proporsi pada saat tidak terjadinya kematian ibu di kota Bandung meningkat sebesar $\exp(2{,}7509) = 15{,}65671659 \approx 15{,}6567$.

Selanjutnya, dari hasil Tabel 2 dapat dilihat pada model *count regression*, variabel prediktor yang signifikan hanya variabel X1 saja. Sedangkan pada model *zero inflation*, tidak terdapat variabel prediktor yang berpengaruh terhadap model. Maka dari itu, disini penulis ingin melihat peluang dari model *count regression* dimana hanya mengandung variabel X1 saja. Berikut ini perhitungan peluang kematian ibu dari model *count regression*:

$$\hat{\pi}_i = \frac{\exp(\beta_0 + \beta_1 X_1)}{1 + \exp(\beta_0 + \beta_1 X_1)}$$
$$\hat{\pi}_i = \frac{\exp(-3{,}9539 - 0{,}03345 X_1)}{1 + \exp(-3{,}9539 - 0{,}03345 X_1)}$$

dimana $i = 1,2, \dots ,80$ (Puskesmas)

Adapun hasil yang diperoleh dengan menggunakan alat bantu *software* Microsoft Excel, berikut merupakan tabel peluang dua puskesmas terendah dan tertinggi:

Tabel 3. Peluang Kematian Ibu di Kota Bandung Tahun 2020

Peluang Terendah				
No	**Kecamatan**	**Puskesmas**	**Persentase K1 (%) (X1)**	**Peluang**
1	Panyieukan	Panyileukan	136,7	0.0001981
2	Batununggal	Ibrahim Adjie	121.63	0.0003279
Peluang Tertinggi				
No	**Kecamatan**	**Puskesmas**	**Persentase K1 (%) (X1)**	**Peluang**
1	Bandung Wetan	Salam	23.78	0.0085831
2	Regol	Pasundan	41.72	0.0047284

Berdasarkan Tabel 3, dapat disimpulkan bahwa semakin tinggi persentase pelayanan antenatal ibu hamil sesuai standar pada triwulan pertama kehamilan (Persentase K1), maka semakin rendah peluang terjadinya kematian ibu. Sebaliknya, semakin rendah persentase pelayanan antenatal ibu hamil sesuai standar pada triwulan pertama kehamilan (Persentase K1), maka semakin tinggi peluang terjadinya kematian ibu.

D. Kesimpulan

Secara umum dalam penelitian ini telah ditunjukkan mengenai hubungan antara variabel prediktor dan variabel respon yang berbentuk proporsi, yang pada umumnya dimodelkan dengan model regresi beta, akan tetapi dalam penelitian ini kasus yang diangkat yaitu kasus yang jarang terjadi sehingga memiliki nilai nol yang berlebih, sehingga menggunakan pemodelan alternative yaitu model regresi *zero inflated beta*. Berdasarkan hasil dari pemodelan regresi *zero inflated beta* bahwa untuk model *count regression* diperoleh faktor yang berpengaruh secara signifikan terhadap proporsi angka kematian ibu di Kota Bandung tahun 2020 yaitu hanya faktor persentase K1. Sedangkan untuk model *zero inflation* diperoleh tidak terdapat faktor yang berpengaruh secara signifikan terhadap proporsi pada saat tidak terjadinya kematian ibu di Kota Bandung tahun 2020.

Selanjutnya dilakukan pemodelan baru pada model *count regression* yang digunakan untuk menghitung peluang terjadinya kematian ibu, peluang yang memiliki jumlah kasus tertinggi terdapat di puskesmas Salam Kecamatan Bandung Wetan dengan nilai variabel X_1 (Persentase K1) sebesar 23,78 didapat nilai peluang sebesar 0,0085831 atau sebanyak 9 angka kematian ibu per 1.000 kelahiran hidup. Sedangkan, peluang terjadinya kematian ibu yang terendah terdapat di puskesmas Panyileukan Kecamatan Panyileukan dengan nilai variabel X_1 (Persentase K1) sebesar 136,7 didapat nilai peluang terjadinya kematian ibu sebesar 0,0001981 atau sebanyak 1 angka kematian ibu per 10.000 kelahiran hidup.

Acknowledge

Penelitian ini dapat terlaksana dengan baik atas bantuan, bimbingan, saran dan dukungan dari berbagai pihak. Oleh karena itu, peneliti mengucapkan terima kasih kepada Yang Maha Kuasa Allah SWT atas Rahmat dan Hidayah-Nya, kedua orang tua serta keluarga yang selalu mendoakan dan memberi semangat, Bapak Nusar Hajarisman, S.Si., M.Si., yang telah memberikan sumbangan pikiran, pengetahuan, ilmu, serta motivasi, para dosen Statistika Unisba yang telah memberikan ilmu pengetahuannya, dan teman-teman yang telah memberikan kritikan, bantuan, dan dorongannya.

Daftar Pustaka

[1] Hajarisman, N. (2012, Mei). Penaksiran Parameter Model Regresi Beta untuk Memodelkan Data Proporsi. *Statistika, Vol. 12 No. 1*, 9-18.

[2] Peng, X., Li, G., & Liu, Z. (2015). Zero-Inflated Beta Regression for Differential Abundance Analysis with Metagenomics Data. *Journal of Computational Biology, Volume 23, Number 2*, 1 - 9.

[3] Masserini, L., Bini, M., & Pratesi, M. (2016). Effectiveness of Non-selective Evaluation Test Scores for Predicting First-year Performance in University Career: a Zero-inflated Beta Regression Approach. *Springer*, 1-16.

[4] Walpole, R., Myers, R., Myers, S., & Ye, K. (2012). *Probability & Statistics for Engineers & Scientists, 9th ed.* Boston: Pearson Education, Inc.

[5] Junior, O. A., Guedes, T. A., Zuliani, A. M., & Nunes, W. M. (2017). Zero-inflated Beta Regression Model for Leaf Citrus Canker Incidence in Orange Genotypes Grafted Onto Different Rootstocks. *Acta Scientiarum, Vol. 39, No. 2*, 161-171.

[6] Famoye, F., & Singh, K. P. (2006). Zero-Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data. *Journal of Data Science, 4*, 117-130.

[7] Ospina, R., & Ferrari, S. L. (2012). A General Class of Zero-or-one Inflated Beta Regression Models. *Elsevier*, 1609-1623.

[8] Ospina, R., & Ferrari, S. L. (2010). Inflated Beta Distributions. *Springer*, 111-126.

[9] Jumiartanti, M., Indahwati, & Kurnia, A. (2017). Zero Inflated Beta Model in Small Area Estimation to Estimate Poverty Rates on Village Level in Langsa Municipality. *IJCSN - International Journal of Computer Science and Network, Volume 6*, 812-819.

[10] Anderson, D., Burnham, K., & White, G. (1998). Comparison of Akaike Information Criterion and Consistent Akaike Information Criterion for Model Selection and Statistical Inference from Capture-Recapture Studies. *Journal of Applied Statistics, Vol. 25 No. 2*, 263-282.

[11] Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression.* Canada: John Wiley & Sons, Inc.

[12] Kurniawan, R., & Yuniarto, B. (2016). *Analisis Regresi: Dasar dan Penerapannya dengan R.* Jakarta: Kencana.

[13] Lisnawati, & Prahastuti, D. L. (2021, Desember 2). Pengaruh Pendidikan Kesehatan Menggunakan Media Video Animasi Terhadap Kesiapan Ibu Hamil Dalam Menghadapi Persalinan. *Jurnal Media Kesehatan, Vol. 14, No. 2*, 146-154.

[14] Tawangki Sri Fadilah and Abdul Kudus, “Penerapan Metode Regresi Kernel Smoothing untuk Imputasi Data Lama Waktu Terinfeksi Covid-19,” *Jurnal Riset Statistika*, pp. 51–60, Jul. 2023, doi: 10.29313/jrs.v3i1.1802.