Pemodelan *Generalized Poisson Regression* (GPR) dan *Negative Binomial Regression* (NBR) untuk Mengatasi Overdispersi pada Jumlah Kematian Bayi di Kabupaten Probolinggo

Amara Deviana Chaniago, dan Sri Pingit Wulandari
Departemen Statistika Bisnis, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: sri_pingit@statistika.its.ac.id

***Abstrak*—Data Dinas Kesehatan Kabupaten Probolinggo tahun 2020 menyatakan Angka Kematian Bayi (AKB) di Kabupaten Probolinggo sebesar 8,11. Angka tersebut tergolong tinggi jika dibandingkan dengan AKB Provinsi Jawa Timur sebesar 6,3. Sehingga perlu dilakukan analisis untuk mengetahui faktor-faktor yang mempengaruhi jumlah kematian bayi di Kabupaten Probolinggo tahun 2020. Regresi poisson merupakan analisis yang sesuai dalam pemodelan kasus data diskrit. Namun regresi poisson mensyaratkan kondisi equidispersi yang sulit dipenuhi. Pada umumnya sering ditemui kondisi overdispersi. Metode yang dapat digunakan untuk mengatasi kondisi overdispersi diantaranya adalah metode *Generalized Poisson Regression* (*GPR*) dan *Negative Binomial Regression* (*NBR*). Objek penelitian terdiri dari variabel respon yaitu jumlah kematian bayi serta variabel prediktor yang diduga mempengaruhi jumlah kematian bayi diantaranya yaitu faktor kesehatan ibu hamil dan bayinya, faktor kebersihan dan gizi, serta faktor peningkatan imunitas bayi yang terdiri dari 9 variabel serta 2 variabel tambahan yaitu jumlah tenaga kesehatan dan jumlah fasilitas kesehatan. Hasil analisis diperoleh bahwa terjadi kasus overdispersi sehingga analisis GPR dan NBR perlu dilakukan. Berdasarkan analisis GPR dan NBR, model yang layak untuk digunakan adalah model dengan kombinasi variabel prediktor jumlah bayi lahir rendah (X_7), jumlah ibu hamil mendapat imunisasi Td2+ (X_8), dan jumlah tenaga kesehatan (X_{10}). Keseluruhan variabel berpengaruh signifikan terhadap model. Diperoleh hasil bahwa metode yang paling baik digunakan untuk memodelkan jumlah kematian bayi untuk mengatasi overdispersi adalah metode GPR karena memiliki kriteria kebaikan model AIC, AICc, BIC, dan BICc yang lebih kecil dibandingkan dengan metode NBR.**

***Kata Kunci*—Generalized Poisson Regression, Kematian Bayi, Negative Binomial Regression, Overdispersi.**

I. PENDAHULUAN

KESEHATAN bayi sangat berpengaruh terhadap pertumbuhan dan perkembangan anak sehingga variabel yang mempengaruhi kesehatan bayi perlu diperhatikan. Kematian bayi merupakan kematian yang terjadi antara kelahiran bayi sampai sebelum hari ulang tahun pertama [1]. Angka kematian dinilai dapat merepresentasikan keberhasilan fasilitas dan program pelayanan kesehatan masyarakat. Selain itu, salah satu indikator yang lazim digunakan untuk mengukur derajat kesehatan masyarakat adalah Angka Kematian Bayi (AKB). Angka Kematian Bayi (AKB) menunjukkan jumlah kematian bayi pada umur kurang dari satu tahun per 1000 kelahiran bayi. Profil Kesehatan Jawa Timur menyebutkan bahwa pada selama tahun 2019, Kabupaten Probolinggo menempati urutan ketiga dengan jumlah kasus kematian bayi terbanyak yaitu sebesar 180 kasus. Berdasarkan data [2], angka kematian bayi (AKB) Kabupaten Probolinggo sebesar 8,11. Angka tersebut masih tergolong tinggi jika dibandingkan dengan nilai AKB Provinsi Jawa Timur yaitu sebesar 6,3.

Setiap daerah selalu berusaha untuk meningkatkan kualitas kesehatan begitu pula dengan Kabupaten Probolinggo. Oleh karena itu, kualitas kesehatan bayi perlu ditingkatkan untuk mewujudkan derajat kesehatan masyarakat yang lebih baik. Penyebab kematian bayi terdiri dari faktor eksogen dan faktor endogen. Penyebab langsung kematian bayi berdasarkan [2] paling besar disebabkan oleh kasus Berat Bayi Lahir Rendah yaitu sebesar 23,12%. Penelitian terdahulu terkait kesehatan bayi dilakukan oleh [3] yang menyimpulkan bahwa terdapat tiga indikator kesehatan bayi yaitu faktor kesehatan ibu hamil dan bayinya, faktor kebersihan dan gizi, serta faktor peningkatan imunitas bayi. Penelitian yang dilakukan oleh [4], [5] tentang kematian bayi memperoleh hasil bahwa jumlah tenaga kesehatan dan jumlah fasilitas kesehatan berpengaruh signifikan terhadap jumlah kematian bayi. Jumlah tenaga dan fasilitas kesehatan yang memadai dapat meningkatkan kualitas pelayanan kesehatan sehingga dapat menurunkan risiko terjadinya kematian bayi.

Usaha peningkatan derajat kesehatan di Kabupaten Probolinggo, dapat dilakukan dengan mencari faktor-faktor yang mempengaruhi jumlah kematian bayi. Regresi poisson merupakan analisis regresi yang digunakan untuk menganalisis data diskrit seperti jumlah kematian bayi di Kabupaten Probolinggo tahun 2020. Dalam regresi poisson salah satu karakteristik yang harus dipenuhi adalah equidispersi. Equidispersi adalah kejadian dimana terdapat kesamaan nilai *mean* dan varians. Namun, kondisi tersebut tidak selalu bisa dipenuhi. Konsekuensi atas tidak terpenuhinya equidispersi menyebabkan nilai deviansi model menjadi sangat besar sehingga regresi poisson tidak sesuai untuk memodelkan data, dan model yang terbentuk akan menghasilkan estimasi parameter yang bias. Untuk menangani masalah overdispersi, dapat dilakukan pemodelan dengan *Generalized Poisson Regresion* (*GPR*) dan *Negative Binomial Regression (NBR)*. Model ini dapat mengatasi masalah overdispersi karena tidak mengharuskan nilai *mean* yang sama dengan nilai varians seperti model regresi poisson.

Penelitian ini dilakukan untuk mendapatkan model terbaik dengan membandingkan hasil pemodelan menggunakan analisis *GPR* dan *NBR* untuk mengatasi adanya kasus overdispersi pada studi kasus jumlah kematian bayi di Kabupaten Probolinggo tahun 2020 menggunakan regresi poisson. Hasil penelitian diharapkan dapat memberikan informasi mengenai model terbaik untuk menangani overdispersi serta faktor yang mempengaruhi terjadinya kematian bayi.

Tabel 6.
Variabel penelitian

No.	Variabel	Keterangan
1.	Y	Jumlah Kematian Bayi
2.	X_1	Jumlah Ibu Hamil Yang Mendapat 90 Tablet Tambah Darah
3.	X_2	Jumlah Ibu Hamil Yang Melaksanakan Program K1
4.	X_3	Jumlah Ibu Hamil Yang Melaksanakan Program K4
5.	X_4	Jumlah Ibu Bersalin Ditolong Tenaga Kesehatan
6.	X_5	Jumlah Bayi Dapat Vitamin A
7.	X_6	Jumlah Bayi Mendapat Imunisasi Dasar Lengkap
8.	X_7	Jumlah Bayi Berat Lahir Rendah
9.	X_8	Jumlah Ibu Hamil Mendapat Imunisasi Td2+
10.	X_9	Jumlah Bayi Mendapat ASI Eksklusif
11.	X_{10}	Jumlah Tenaga Kesehatan
12.	X_{11}	Jumlah Fasilitas Kesehatan

Tabel 7.
Karakteristik data

Variabel	Rata-rata	Varians	Minimum	Maksimum
Y	6,125	22,723	0	20
X_1	667,167	70358,058	138	1345
X_2	793,167	82072,841	230	1420
X_3	697,625	67540,940	196	1335
X_4	749,417	73559,993	210	1346
X_5	699,292	62152,476	208	1106
X_6	677,125	63617,766	203	1207
X_7	40,583	517,297	5	81
X_8	846,542	441877,650	77	2976
X_9	81,458	12943,476	9	599
X_{10}	80,958	12246,476	19	548
X_{11}	69,417	369,819	32	107

Tabel 1.
Nilai VIF 11 variabel

Variabel	Rata-rata	Varians	Minimum
Y	6,125	22,723	0
X_1	667,167	70358,058	138
X_2	793,167	82072,841	230
X_3	697,625	67540,940	196
X_4	749,417	73559,993	210
X_5	699,292	62152,476	208
X_6	21,28		

Tabel 2.
Nilai VIF 6 variabel

Variabel	VIF
X_1	4,28
X_7	3,23
X_8	3,83
X_9	1,31
X_{10}	3,02
X_{11}	5,91

Tabel 3.
Parameter GPR

Parameter	Koefisien	t_{hitung}	*P-value*
β_0	0,7673	2,45	0,0221
β_7	0,0295	3,44	0,0021
β_8	-0,0006	-1,65	0,1116
β_{10}	0,0033	1,62	0,1174

Tabel 4.
Pemodelan NBR

Parameter	Koefisien	Z_{hitung}	*P-value*
β_0	0,7979	2,673	0,0007
β_7	0,0283	3,779	0,0001
β_8	-0,0006	-1,613	0,1067
β_{10}	0,0031	1,656	0,0977

Tabel 5.
Pemilihan model terbaik

Metode	Variabel Signifikan	AIC	AICc	BIC	BICc
GPR	$\beta_0, \beta_7, \beta_8, \beta_{10}$	131,4	134,73	137,29	140,62
NBR	$\beta_0, \beta_7, \beta_8, \beta_{10}$	131,69	135,02	137,58	140,91

II. TINJAUAN PUSTAKA

A. Multikolinieritas

Multikolinearitas adalah kondisi terdapatnya hubungan linear atau korelasi yang tinggi antara masing-masing variabel prediktor dalam model regresi. Multikolinearitas biasanya terjadi ketika sebagian besar variabel yang digunakan saling terkait dalam suatu model regresi [6]. Oleh karena itu masalah multikolinearitas tidak terjadi pada regresi linear sederhana yang hanya melibatkan satu variabel independen. Pendeteksian adanya kasus multikolinearitas menurut [7] dapat dilihat melalui.

1. *Variance Inflation Factors (VIF) > 10*

$$\mathrm{VIF} = \frac{1}{1-{R_j}^2} \quad (1)$$

${R_j}^2$ merupakan nilai koefisien determinasi antara variabel x_j dengan variabel x lainnya. Nilai VIF yang lebih besar dari 10 menunjukkan adanya kasus multikolinearitas antar variabel prediktor.

2. Nilai koefisien korelasi Pearson (r_{ij}) antar variabel-variabel prediktor.

Apabila nilai koefisien korelasi Pearson antar variabel prediktor cukup besar (lebih dari 0,95) maka mengindikasikan adanya multikolinearitas.

$$r_{x_a,x_b} = \frac{\sum_{i=1}^{n}(x_i-\bar{x}_a)(x_i-\bar{x}_b)}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x}_a)^2\sum_{i=1}^{n}(x_i-\bar{x}_b)^2}}, a \neq b \quad (2)$$

B. Regresi Poisson

Regresi Poisson merupakan analisis regresi yang digunakan untuk memodelkan variabel respon yang diasumsikan berdistribusi poisson. Menurut Walpole (2012), distribusi poisson menyatakan banyaknya sukses yang terjadi dalam selang waktu atau daerah tertentu. Variabel acak poisson adalah variabel respon atau bilangan Y yang menunjukkan banyaknya hasil percobaan dalam suatu percobaan poisson, sebaran peluangnya disebut dengan sebaran poisson. Fungsi peluang dari distribusi poisson dinyatakan pada Persamaan 3 [8].

$$f(y,\mu) = \frac{e^{-\mu}\mu^y}{y!} \quad (3)$$

Distribusi poisson adalah adalah salah satu distribusi sederhana dalam pemodelan data yang berupa *count* (jumlah).. Regresi poisson merupakan *Generalized Linear Model* (*GLM*) yang dikembangkan dari model linear klasik khususnya dalam mengatasi variabel respon tidak normal [9].

Distribusi poisson termasuk dalam keluarga eksponensial. Hal ini ditunjukkan oleh [9] fungsi peluang densitasnya sebagai berikut.

$$f(y,\mu) = \frac{e^{-\mu}\mu^y}{y!} = \frac{exp(-\mu)\,exp(y\,ln\,\mu)}{y!} \quad (4)$$

1) Estimasi Parameter

Maximum Likelihood Estimation (MLE) digunakan sebagai metode penaksiran parameter model regresi poisson. MLE dilakukan dengan cara memaksimumkan fungsi *likelihood*.

$$\begin{aligned} ln\,L(\boldsymbol{\beta}) &= ln\left(\prod_{i=1}^{n}\frac{exp(-\mu_i)\mu_i^{y_i}}{y_i!}\right) \\ &= \sum_{i=1}^{n} ln\left(\frac{exp(-\mu_i)\mu_i^{y_i}}{y_i!}\right) \\ &= \sum_{i=1}^{n}(ln(e^{-\mu_i}) + ln\mu_i^{y_i}) - ln(y_i!)) \\ &= \sum_{i=1}^{n}(-\mu_i + y_i ln\mu_i - ln(y_i!)) \\ &= \sum_{i=1}^{n}\left(-e^{x_i^T\boldsymbol{\beta}} + y_i\boldsymbol{x_i^T\beta} - ln(y_i!)\right) \end{aligned}$$

$$ln\,L(\boldsymbol{\beta}) = -\sum_{i=1}^{n} e^{x_i^T\boldsymbol{\beta}} + \sum_{i=1}^{n} y_i\boldsymbol{x_i^T\beta} - \sum_{i=1}^{n} ln(y_i!) \quad (5)$$

Persamaan 5 masih berbentuk implisit sehingga perlu dilakukan iterasi Newton Raphson untuk melakukan iterasi munerik dengan algoritma optimisasi sebagai berikut:

a. Menentukan nilai taksiran awal parameter ($\hat{\beta}_{(0)}$) yang diperoleh dengan metode *Ordinary Least Square* (OLS) yaitu:

$$\boldsymbol{\hat{\beta}_{(0)}} = (\boldsymbol{X^TX})^{-1}\boldsymbol{X^TY} \quad (6)$$

dimana:

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n,1} & \cdots & x_{nk} \end{bmatrix} \text{ dan } \mathbf{Y} = [Y_1 \quad Y_2 \quad \ldots \quad Y_n]^T \quad (7)$$

b. Membentuk vektor gradien $\boldsymbol{g}$,

$$\boldsymbol{g(\beta)} = -\sum_{i=1}^{n} x_i.\mu_i + \sum_{i=1}^{n} y_i x_i = \sum_{i=i}^{n} x_i(y_i - \mu_i)$$

Sehingga vektor $\boldsymbol{g}^T(\boldsymbol{\beta}_{(m)})_{(k+1)x1}$ adalah sebagai berikut:

$$\left[\frac{\partial\,ln\,L(\boldsymbol{\beta})}{\partial\beta_0} \quad \frac{\partial\,ln\,L(\boldsymbol{\beta})}{\partial\beta_1} \quad \frac{\partial\,ln\,L(\boldsymbol{\beta})}{\partial\beta_2} \quad \ldots \quad \frac{\partial\,ln\,L(\boldsymbol{\beta})}{\partial\beta_k}\right]_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}(m)}$$

dimana *k* adalah banyaknya parameter yang ditaksir.

c. Membentuk matriks Hessian **H** yang disebut juga matriks informasi.

$$\boldsymbol{H(\beta)} = -\sum_{i=1}^{n} x_i.x_i^T.\mu_i = -\sum_{i=1}^{n}\mu_i.x_i.x_i^T$$

$$\boldsymbol{H}(\boldsymbol{\beta}_{(m)})_{(k+1)x(k+1)} = \begin{bmatrix} \frac{\partial^2 ln\,L(\beta)}{\partial\beta_0^2} & \frac{\partial^2 ln\,L(\beta)}{\partial\beta_0\beta_1} & \cdots & \frac{\partial^2 ln\,L(\beta)}{\partial\beta_0\beta_k} \\ & \frac{\partial^2 ln\,L(\beta)}{\partial\beta_1^2} & \cdots & \frac{\partial^2 ln\,L(\beta)}{\partial\beta_1\beta_k} \\ & & \ddots & \vdots \\ simetris & & & \frac{\partial^2 ln\,L(\beta)}{\partial\beta_k^2} \end{bmatrix}_{\beta=\beta_{(m)}}$$

$$\boldsymbol{Var(\hat{\beta})} = -E[\boldsymbol{H^{-1}(\hat{\beta})}]$$

d. Memasukkan nilai $\boldsymbol{\hat{\beta}}_{(0)}$ ke dalam elemen-elemen vektor $\boldsymbol{g}$ dan matriks $\boldsymbol{H}$, sehingga diperoleh vektor $\boldsymbol{g}(\boldsymbol{\hat{\beta}}_{(0)})$ dan matriks $\boldsymbol{H}(\boldsymbol{\hat{\beta}}_{(0)})$.

e. Mulai dari $m = 0$ dilakukan iterasi pada persamaan: $\boldsymbol{\hat{\beta}}_{(m+1)} = \boldsymbol{\hat{\beta}}_{(m)} - \boldsymbol{H^{-1}}(\boldsymbol{\beta}_{(m)})\boldsymbol{g}(\boldsymbol{\beta}_{(m)})$. Nilai $\boldsymbol{\hat{\beta}}_{(m)}$ merupakan sekumpulan penaksir parameter yang konvergen pada iterasi ke-*m*.

f. Kekonvergenan penaksir parameter dapat dilihat ketika nilai $\|\boldsymbol{\hat{\beta}}_{(m+1)} - \boldsymbol{\hat{\beta}}_{(m)}\| \le \varepsilon$ dimana ε merupakan suatu bilangan yang sangat kecil sekali. Sehingga hampir tidak ada perbedaan antara $\boldsymbol{\beta}_{(m+1)}$ dan $\boldsymbol{\beta}_{(m)}$. Jika belum didapatkan penaksir parameter yang konvergen, maka dilanjutkan kembali langkah 5 hingga iterasi ke $m = m + 1$.

2) Pengujian Parameter

Salah satu metode yang digunakan untuk menentukan statistik uji dalam pengujian parameter model regresi poisson adalah dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis. $L(\hat{\Omega})$ adalah nilai *likelihood* untuk model lengkap dengan melibatkan variabel prediktor dan $L(\hat{\omega})$ yaitu nilai *likelihood* untuk model *saturated*. Model *saturated* adalah model teoritis dimana setiap observasi memiliki parameter masing-masing.

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_j = 0$

H_1 : paling sedikit ada satu $\beta_j \neq 0$; j=1, 2, ... , k

Statistik uji untuk kelayakan model regresi poisson adalah sebagai berikut:

$$D(\boldsymbol{\hat{\beta}}) = -2\,ln\left[\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right] = 2\left[ln\left(L(\hat{\Omega})\right) - ln\left(L(\hat{\omega})\right)\right] \quad (8)$$

Keputusan tolak H0 jika $D(\boldsymbol{\hat{\beta}}) > \chi^2_{k;\alpha}$ dengan *k* adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter dibawah H0. Parameter model regresi poisson yang telah dihasilkan dari estimasi perlu dilakukan pengujian terhadap parameter model regresi poisson secara individu dengan hipotesis sebagai berikut:

H_0 : $\beta_1 = 0$ (pengaruh variabel ke-j tidak signifikan)

H_1 : $\beta_j \neq 0$ (pengaruh variabel ke-j signifikan)

Statistik uji yang digunakan adalah:

$$z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (9)$$

$se(\hat{\beta}_j)$ adalah nilai *standar error* atau tingkat kesalahan dari parameter β_j. $se(\hat{\beta}_j)$ didapatkan dari elemen diagonal ke (j+1) dari $[-\boldsymbol{H^{-1}}(\boldsymbol{\hat{\beta}})]$. Diputuskan tolak H0 jika $|z_{hit}| > z_{\frac{\alpha}{2}}$ dimana *α* adalah tingkat signifikansi.

3) Overdispersi pada Regresi Poisson

Overdispersi terjadi apabila nilai varians variabel respon lebih besar dari nilai meannya. Apabila overdispersi terjadi pada regresi poisson, maka nilai dugaan dari koefisien parameter regresi tetap konsisten tetapi tidak efisien, sehingga kesimpulannya menjadi tidak valid. Fenomena overdispersi [10] dapat dituliskan $Var(Y) > E(Y)$. Misal θ merupakan parameter dispersi maka jika θ = 0 artinya tidak terjadi kasus overdispersi.

Nilai θ dapat diindikasikan dengan nilai dispersi *deviance* yang dibagi dengan derajat bebasnya sebagai berikut:

$$\theta = \frac{D(\hat{\beta})}{db} = \frac{2\left[ln\left(L(\hat{\Omega})\right) - ln\left(L(\hat{\omega})\right)\right]}{db} \quad (10)$$

Jika nilai $\theta > 0$ maka dikatakan terjadi overdispersi pada data [11]. Overdispersi dapat terjadi karena adanya pengamatan *missing* pada variabel prediktor, adanya pencilan pada data, perlunya interaksi dalam model, variabel prediktor perlu ditransformasi atau kesalahan spesifikasi *link function* [12].

C. Generalized Poisson Regression (GPR)

Model GPR merupakan suatu model yang sesuai digunakan untuk data count dimana terjadi over/under dispersi. Sehingga selain parameter μ juga terdapat θ sebagai parameter dispersi. GPR hampir mirip dengan regresi poisson yaitu merupakan suatu model GLM. Akan tetapi pada model GPR mengasumsikan bahwa komponen randomnya berdistribusi *Generalized Poisson* (GP). Misal, y = 0,1,2,... merupakan variabel respon. Distribusi GP diberikan sebagai berikut [13].

$$f(y;\mu;\theta) = \left(\frac{\mu}{1+\theta\mu}\right)^{y} \frac{(1+\theta y)^{y-1}}{y!} \exp\left(\frac{-\mu(1+\theta y)}{1+\theta\mu}\right) \quad (11)$$

dengan mean dan varians model GPR adalah sebagai berikut:

$$E(y) = \mu \text{ dan } Var(y) = \mu(1+\theta\mu)^2$$

Apabila $\theta = 0$ maka model GPR akan menjadi regresi poisson biasa, apabila $\theta > 0$ maka model GPR merepresentasikan data count yang overdispersion, dan jika $\theta < 0$ maka GPR merepresentasikan data count yang underdispersion.

1) Penaksiran Parameter Generalized Poisson Regression (GPR)

Penaksir parameter model *GPR* pada Persamaan 11 dilakukan dengan metode MLE *(Maximum Likelihood Estimator)*. Fungsi likelihood untuk model *GPR* adalah sebagai berikut:

$$L(\boldsymbol{\beta},\theta) = \prod_{i=1}^{n} f(\boldsymbol{\beta},\theta)$$

$$L(\boldsymbol{\beta},\theta) = \prod_{i=1}^{n}\left\{\left(\frac{\mu_i}{1+\theta\mu_i}\right)^{y_i} \frac{(1+\theta y_i)^{(y_i-1)}}{y_i!} exp\left(\frac{-\mu_i(1+\theta y_i)}{1+\theta\mu_i}\right)\right\} \quad (12)$$

Selanjutnya persamaan (12) diubah dalam bentuk fungsi logaritma natural menjadi:

$$ln\,L(\boldsymbol{\beta},\theta) = \sum_{i=1}^{n}\{y_i\, ln(\mu_i) - y_i\, ln(1+\theta\mu_i) + \Delta\} \quad (13)$$

dimana:

$$\Delta = (y_i - 1)\, ln(1+\theta y_i) - ln(y_i!) - \frac{\mu_i(1+\theta y_i)}{1+\theta\mu_i}$$

Kemudian persamaan logaritma natural dari fungsi likelihood diturunkan terhadap $\boldsymbol{\beta}^T$ dan disamakan dengan nol untuk mendapatkan parameter $\hat{\boldsymbol{\beta}}$. Jika ingin mendapatkan penaksir parameter $\hat{\theta}$ maka persamaan tersebut diturunkan terhadap θ dan disamakan dengan nol. Selanjutnya digunakan iterasi Newton-Raphson sampai didapatkan penaksir parameter yang konvergen.

2) Pengujian Parameter Generalized Poisson Regression (GPR)

Pengujian parameter *GPR* dilakukan dengan metode *Maximum Likelihood Ratio Test (MLRT)*. Hipotesis yang digunakan adalah sebagai berikut:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_j = 0$

H_1 : paling sedikit ada satu $\beta_j \neq 0; j = 1,2,\ldots,k$

Statistik uji:

$$D(\hat{\beta}) = -2\, ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right) = 2\left(ln\, L\left(\hat{\Omega}\right) - ln\, L\left(\hat{\omega}\right)\right) \quad (14)$$

Tolak H_0 jika $D(\hat{\beta}) > \chi^2_{(k;\alpha)}$ sehingga paling tidak ada satu $\beta_j \neq 0$ yang menunjukkan bahwa X_j berpengaruh secara signifikan terhadap model. Pengujian dilanjutkan dengan uji secara parsial dengan hipotesis sebagai berikut:

H_0 : $\beta_1 = 0$ (pengaruh variabel ke-j tidak signifikan)

H_1 : $\beta_j \neq 0$ (pengaruh variabel ke-j signifikan)

Statistik uji yang digunakan mengikuti distribusi z yaitu:

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (15)$$

$se(\hat{\beta}_j)$ adalah nilai *standar error* atau tingkat kesalahan dari parameter β_j. $se(\hat{\beta}_j)$ didapatkan dari elemen diagonal ke (j+1) dari $[-\boldsymbol{H}^{-1}(\hat{\boldsymbol{\beta}})]$. Diputuskan tolak H_0 jika $|t_{hit}| > t_{\frac{\alpha}{2},df}$ dimana α adalah tingkat signifikansi.

D. Negative Binomial Regression (NRB)

Negative Binomial Regression merupakan salah satu model regresi terapan dari GLM. Pada *Negative Binomial Regression* variabel respon Y_i diasumsikan berdistribusi binomial negatif yang dihasilkan dari distribusi *mixture* poisson-gamma. $P(Y|\alpha,\beta)$ merupakan fungsi kepadatan peluang binomial negatif yang dihasilkan dari distribusi *mixture* poisson-gamma. Nilai mean dan varians poisson-gamma *mixture* adalah $E[Y] = \alpha\beta$ dan $V[Y] = \alpha\beta + \alpha\beta^2$. Untuk membentuk suatu model regresi pada distribusi binomial negatif, maka nilai parameter dari distribusi poisson-gamma *mixture* dinyatakan dalam $\mu = \alpha\beta$ dan $\theta = \frac{1}{\alpha}$ sehingga diperoleh *mean* dan varians sebagai berikut:

$$E[Y] = \mu \text{ dan } V[Y] = \mu + \theta\mu^2$$

Kemudian fungsi massa peluang binomial negatif menjadi [14]:

$$f(y;\mu,\theta) = \frac{\Gamma\left(y+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y!}\left(\frac{1}{1+\theta\mu}\right)^{\frac{1}{\theta}}\left(\frac{\theta\mu}{1+\theta\mu}\right)^{y} \quad (16)$$

Saat $\theta \to 0$ maka distribusi binomial negatif memiliki varians $V[Y] \to \mu$. Distribusi binomial negatif akan mendekati suatu distribusi poisson yang mengasumsikan mean dan varians sama yaitu $E[Y] = V[Y] = \mu$. Fungsi distribusi keluarga eksponensial dari distribusi binomial negatif adalah sebagai berikut:

$$f(y;\mu,\theta) = exp\left\{y\, ln\left(\frac{\theta\mu}{1+\theta\mu}\right) + \frac{1}{\theta} ln\left(\frac{1}{1+\theta\mu}\right) + \Delta\right\} \quad (17)$$

dimana:

$$\Delta = ln\left(\frac{\Gamma\left(y+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y!}\right)$$

1) Penaksir Parameter Negative Binomial Regression (NRB)

Estimasi parameter dari *Negative Binomial Regression* digunakan metode MLE dengan prosedur Newton Rhapson. Metode ini membutuhkan turunan pertama dan kedua dari fungsi *likelihood*. Y_i mempunyai fungsi massa probabilitas distribusi binomial negatif sebagai berikut:

$$f(y_i|\mu_i,\theta)=\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)\Gamma(y_i+1)}\left(\frac{1}{1+\theta\mu_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta\mu_i}{1+\theta\mu_i}\right)^{y_i} \quad (18)$$

Karena fungsinya saling bebas, maka fungsi likelihood adalah sebagai berikut [14]:

$$L(\boldsymbol{\beta},\theta)=\prod_{i=1}^{n}\frac{\Gamma\left(y+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)\Gamma(y+1)}\left(\frac{1}{1+\theta\mu_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta\mu_i}{1+\theta\mu_i}\right)^{y} \quad (19)$$

dengan:

$$\frac{\Gamma\left(y+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)}=\prod_{j=1}^{y-1}(j+\theta^{-1})$$

Selanjutnya persamaan diubah dalam bentuk fungsi logaritma natural menjadi:

$$L(\boldsymbol{\beta},\theta)=ln\{l(\boldsymbol{\beta},\theta)\}=\sum_{i=1}^{n}\left[\left(\sum_{r=0}^{y_i-1}ln(j+\theta^{-1})\right)-ln(\Gamma(y+1))-\Delta\right]$$

dimana:

$$\Delta=(y_i+\theta^{-1})\,ln(1+\theta\mu_i)+y_i\,ln(\theta)+y_i ln(\mu_i)$$

Estimasi parameter *Negative Binomial Regression* dilakukan dengan langkah sebagai berikut:

a. Menentukan taksiran awal dari θ, misal $\hat{\theta}_1=0$
b. Menentukan estimasi maksimum likelihood dari parameter β menggunakan prosedur iterasi *Fisher Scoring* dengan asumsi $\theta=\hat{\theta}_1$

$$\hat{\beta}_{i+1}=\hat{\beta}_i+(X^TW_iX)^{-1}X^TW_iz_i$$

Iterasi berakhir jika diperoleh $\left\|\hat{\beta}_{i+1}-\hat{\beta}_i\right\|\le\varepsilon$
c. Menggunakan $\hat{\beta}$ untuk menghasilkan estimasi dari parameter θ dengan menggunakan prosedur iterasi Newthon-Rhapson satu dimensi

$$c\hat{\theta}_{i+1}=\hat{\theta}_i-\frac{f'(\theta_i)}{f''(\theta_i)}$$

Iterasi berakhir jika diperoleh $\left|\hat{\theta}_{i+1}-\hat{\theta}_1\right|<\varepsilon$
d. Jika $\left|\hat{\theta}_{i+1}-\hat{\theta}_1\right|<\varepsilon$ selesai; bila tidak, gunakan parameter $\theta=\hat{\theta}_{i+1}$ dan kembali ke langkah 2. Nilai ε merupakan nilai bilangan positif yang sangat kecil

2) Pengujian Parameter Generalized Poisson Regression (GPR)

Pengujian parameter *NBR* menggunakan dengan uji deviansi megunakan hipotesis sebagai berikut:

H_0 : $\beta_1=\beta_2=\cdots=\beta_j=0$
H_1 : paling sedikit ada satu $\beta_j\neq 0; j=1,2,\dots,k$
Statistik uji:

$$D(\hat{\beta})=-2\,ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right)=2\left(ln\,L\left(\hat{\Omega}\right)-ln\,L\left(\hat{\omega}\right)\right) \quad (20)$$

Tolak H_0 jika $D(\hat{\beta})>\chi^2_{(k;\alpha)}$. Pengujian dilanjutkan dengan uji secara parsial dengan hipotesis sebagai berikut:
H_0 : $\beta_1=0$ (pengaruh variabel ke-j tidak signifikan)
H_1 : $\beta_j\neq 0$ (pengaruh variabel ke-j signifikan)
Statistik uji yang digunakan mengikuti distribusi z yaitu:

$$z=\frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (21)$$

$se(\hat{\beta}_j)$ adalah nilai *standar error* atau tingkat kesalahan dari parameter β_j. $se(\hat{\beta}_j)$ didapatkan dari elemen diagonal ke (j+1) dari $[-\boldsymbol{H}^{-1}(\hat{\boldsymbol{\beta}})]$. Diputuskan tolak H_0 jika $|z_{hitung}|>z_{\frac{\alpha}{2}}$ dimana α adalah tingkat signifikansi.

E. Pemilihan Model Terbaik

Kriteria pemilihan model terbaik dalam analisis regresi diantaranya terdapat beberapa metode yaitu *Akaike Information Criterion* (*AIC*) dan *Bayes Information Criterion* (*BIC*). *Akaike Information Criterion* (*AIC*) menurut [15] didefinisikan sebagai berikut:

$$AIC=-2\,ln\,L\left(\tilde{\theta}\right)+2k \quad (22)$$

dimana $L(\hat{\theta})$ nilai *likelihood*, dan *k* jumlah parameter.

Model terbaik adalah model yang mempunyai nilai *AIC* terkecil. Faktor koreksi pada *AIC* yaitu *Corrected Akaike Information Criterion* (AIC_C), dimana n menunjukkan jumlah data pada penelitian dan k adalah jumlah parameter.

$$AIC_C=AIC+\frac{2k\,(k+1)}{n-k-1} \quad (23)$$

Salah satu modifikasi dari AIC adalah BIC. BIC sering digunakan karena memiliki keunggulan dalam kecepatan serta kesederhanaan komputasi [14].

$$BIC=-2\,ln\,L\left(\tilde{\theta}\right)+k\,ln(n) \quad (24)$$

$$BIC_C=BIC+\frac{2k\,(k+1)}{n-k-1} \quad (25)$$

III. METODOLOGI PENELITIAN

A. Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah data sekunder. Data diperoleh dari publikasi Dinas Kesehatan Kabupaten Probolinggo yang dapat diakses melalui website https://dinkesprobolinggokab.go.id dan publikasi Badan Pusat Statistik Kabupaten Probolinggo yang dapat diakses melalui website https://probolinggokab.bps.go.id. Variabel prediktor yang diduga berpengaruh terhadap jumlah kematian bayi merupakan indikator kesehatan bayi yang sebanyak 9 variabel serta 2 variabel tambahan yaitu jumlah tenaga kesehatan dan jumlah fasilitas kesehatan. Jumlah data yang digunakan sebanyak 24 data yang terdiri dari kecamatan di Kabupaten Probolinggo.

B. Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini adalah jumlah kematian bayi sebagai variabel respon dan variabel prediktor yang terdiri dari 11 variabel dengan skala data rasio disajikan pada Tabel 1.

C. *Langkah Analitis*

Langkah-langkah analisis pada penelitian ini diantaranya yaitu sebagai berikut:

1. Melakukan analisis statistika deskriptif jumlah kematian bayi yang terjadi di Kabupaten Probolinggo beserta dengan variabel prediktor.
2. Menganalisis korelasi antar variabel-variabel prediktor untuk mendeteksi adanya kasus multikolinearitas menggunakan nilai VIF dan koefisien korelasi.
3. Mendeteksi adanya overdispersi pada data dengan melihat nilai *Deviance* yang dibagi dengan derajat bebasnya.
4. Mendapatkan model terbaik untuk *GPR* pada pemodelan jumlah kematian bayi yang terjadi pada kecamatan di Kabupaten Probolinggo.
   a. Menaksir parameter model *GPR*.
   b. Menguji signifikansi parameter model *GPR* secara serentak dan parsial.
5. Mendapatkan model terbaik untuk *NBR* pada pemodelan jumlah kematian bayi yang terjadi pada kecamatan di Kabupaten Probolinggo.
   a. Menaksir parameter model NBR.
   b. Menguji signifikansi parameter model NBR secara serentak dan parsial.
6. Membandingkan nilai *AIC*, *AICC, BIC,* dan *BICc* dari *GPR* dan *NBR* yang dihasilkan untuk mendapatkan model terbaik. Model terbaik adalah model dengan nilai *AIC*, *AICC, BIC,* dan *BICc* terkecil.

IV. ANALISIS DAN PEMBAHASAN

A. *Karakteristik Data*

Tabel 2 menunjukkan rata-rata jumlah kematian bayi pada 24 kecamatan di Kabupaten Probolinggo sebesar 6,125 kasus. Pada tahun 2020, di Kecamatan Sukapura tidak terdapat kasus kematian bayi. Jumlah kematian bayi tertinggi terdapat pada Kecamatan Paiton yaitu sejumlah 20 kematian bayi. Keragaman dari jumlah kematian bayi di Kabupaten Probolinggo cukup kecil yaitu sebesar 22,732. Hal tersebut menunjukkan bahwa persebaran jumlah kematian bayi di tiap kecamatan di Kabupaten Probolinggo besarnya hampir sama.

Nilai keragaman tertinggi pada variabel yang diduga berpengaruh terhadap kematian bayi di Kabupaten Probolinggo terdapat pada variabel X_8 (Jumlah Ibu Hamil Mendapat Imunisasi Td2+) yaitu sebesar 441877,65 dengan nilai minimum 77 terdapat pada Kecamatan Sukapura dan nilai maksimum 2976 ibu hamil terdapat pada Kecamatan Kraksaan. Kecamatan Kraksaan diduga dapat menjadi kecamatan dengan jumlah ibu hamil yang mendapat imunisasi Td2+ tertinggi karena daerah tersebut merupakan daerah pusat pemerintahan daerah sehingga program imunisasi dapat lebih terakomodir.

B. *Pemeriksaan Multikolinieritas*

Hasil pemeriksaan multikolinearitas pada 11 variabel terdapat pada Tabel 3. Multikolinearitas mengindikasikan adanya hubungan antar variabel prediktor. Nilai VIF pada variabel X_1, X_2, X_3, X_4, X_5, dan X_6 memiliki nilai lebih dari 10 sehingga variabel terindikasi terjadi kasus multikolinearitas pada variabel tersebut. Deteksi kasus multikolinearitas diperkuat dengan hasil nilai koefisien korelasi dimana terdapat koefisien korelasi yang nilainya besar (>0,95) yang memperkuat indikasi terjadinya kasus multikolinearitas pada variabel X_2, X_3, X_4, X_5, dan X_6. Oleh karena itu perlu dilakukan penanganan dengan menghapus variabel X_2, X_3, X_4, X_5, dan X_6 yang kemudian dilakukan perhitungan nilai VIF kembali yang disajikan pada Tabel 4.

Hasil penanganan multikolinearitas menunjukkan bahwa dengan cara menghapus variabel X_2, X_3, X_4, X_5, dan X_6 telah mampu menangani kasus multikolinearitas. Hal tersebut dibuktikan dengan keseluruhan variabel memiliki nilai VIF kurang dari 10.

C. *Pemeriksaan Overdispersi*

Pemeriksaan overdispersi dapat dilakukan menggunakan statistik uji serentak yang dibagi dengan derajat bebas yang disimbolkan dengan θ. Jika nilai θ lebih dari nol maka terdapat kondisi overdispersi. Didapatkan nilai residual devians pada model regresi poisson sebesar 23,8595 dengan derajat bebas 17 sehingga didapatkan nilai θ sebesar 1,4035 sehingga terjadi kondisi overdispersi.

D. *Pemodelan Generalized Poisson Regression (GPR)*

Pemodelan dilakukan dengan menggunakan 6 variabel prediktor dengan memilih model terbaik berdasarkan jumlah variabel signifikan dalam model. Penelitian ini menggunakan 6 variabel prediktor sehingga terdapat 63 kombinasi.

Nilai estimasi parameter yang diperoleh pada Tabel 5 selanjutnya dilakukan uji signifikansi secara serentak dan parsial. Uji serentak dari model GPR dijelaskan dengan hipotesis sebagai berikut:

H_0 : $\beta_7, \beta_8, \beta_{10} = 0$ atau tidak ada variabel prediktor yang berpengaruh signifikan terhadap model.

H_1 : Paling sedikit ada satu $\beta_i \neq 0, i = 7, 8, 10$ atau minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap model.

Pada tingkat signifikansi α sebesar 0,15, jika nilai $D(\widehat{\boldsymbol{\beta}})$ lebih besar dari $\chi^2_{(3;0.15)}$ maka diputuskan tolak H_0. Nilai $D(\widehat{\boldsymbol{\beta}})$ yang didapatkan adalah sebesar 13,49 lebih besar dari $\chi^2_{(3;0,15)}$ sebesar 5,32 sehingga diputuskan tolak H_0 yang artinya minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap model.

Selanjutnya dilakukan pengujian secara parsial dengan hipotesis sebagai berikut:

H_0 : $\beta_i = 0, i = 7, 8, 10$ atau variabel prediktor ke-i tidak memberikan pengaruh yang signifikan terhadap model secara parsial.

H_1 : $\beta_i \neq 0$ atau variabel variabel prediktor ke-i memberikan pengaruh yang signifikan terhadap model secara parsial.

Hipotesis diputuskan tolak H_0 jika statistik uji t_{hitung} yang didapatkan nilainya lebih dari $t_{(\frac{0,15}{2},1)}$ yang diperkuat dengan nilai *p-value* yang nilainya kurang dari α sebesar 0,15.

Tabel 5 menunjukkan bahwa nilai $|t_{hitung}|$ variabel pada prediktor X_7, X_8, dan X_{10} lebih besar daripada $t_{(\frac{0.15}{2},1)}$ sebesar 1,49 yang diperkuat dengan nilai *p-value* yang nilainya kurang dari taraf signifikan (α) sebesar 0,15 sehingga diputuskan tolak H_0 yang artinya variabel jumlah bayi lahir rendah (X_7), jumlah ibu hamil mendapat imunisasi Td2+ (X_8), dan umlah tenaga kesehatan (X_{10}) berpengaruh signifikan terhadap jumlah kematian bayi di Kabupaten Probolinggo tahun 2020 dengan model GPR sebagai berikut:

$$\hat{\mu} = e^{0,7673+0,02955X_7-0,0006X_8+0,0033X_{10}}$$

Didapatkan nilai parameter dispersi sebesar 0,078 yang nilainya lebih besar daripada nol sehingga pemodelan GPR sesuai dan dapat digunakan untuk mengatasi overdispersi. Berdasarkan model yang terbentuk maka dapat diketahui bahwa jika setiap peningkatan 1 jumlah bayi lahir rendah akan meningkatkan rata-rata jumlah kematian bayi sebesar exp(0,02955) = 1,0299 kali lipat nilai sebelumnya dengan asumsi variabel lain konstan, setiap peningkatan 1 ibu hamil yang mendapat imunisasi Td2+ akan mengurangi rata-rata jumlah kematian bayi sebesar exp(-0,0006) = 0,999 kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan, dan setiap peningkatan 1 tenaga kesehatan akan meningkatkan rata-rata jumlah kematian bayi sebesar exp(0,0033)=1,0033 kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan.

Prediksi dapat dilakukan berdasarkan hasil model yang didapatkan. Jika terdapat satu bayi lahir rendah, 1 ibu hamil yang mendapat imunisasi Td2+ dan terdapat 1 tenaga kesehatan makadiprediksi akan terdapat 2,22 ≈ 2 kasus kematian bayi di Kabupaten Probolinggo tahun 2020.

E. Pemodelan Negative Binomial Regression (NBR)

Pemodelan Pemodelan dilakukan dengan memilih model terbaik berdasarkan jumlah variabel signifikan dalam model. Penelitian ini menggunakan 6 variabel prediktor sehingga terdapat 63 kombinasi.

Nilai estimasi parameter yang diperoleh selanjutnya dilakukan uji signifikansi secara serentak dan parsial. Uji serentak dari model NBR dijelaskan dengan hipotesis berikut:

H_0 : $\beta_7, \beta_8, \beta_{10} = 0$ atau tidak ada variabel prediktor yang berpengaruh signifikan terhadap model.

H_1 : Paling sedikit ada satu $\beta_i \neq 0, i = 7, 8, 10$ atau minimal terdapat satu variabel prediktor yang berpengaruh signifikan terhadap model.

Pada tingkat signifikansi α sebesar 0,15, jika nilai $D(\hat{\boldsymbol{\beta}})$ lebih besar dari ${\chi^2}_{(3;0.15)}$ maka diputuskan tolak H_0. Nilai $D(\hat{\boldsymbol{\beta}})$ yang didapatkan adalah sebesar 25,208 lebih besar dari ${\chi^2}_{(3;0,15)}$ sebesar 5,32 sehingga diputuskan tolak H_0 yang artinya minimal terdapat satu variabel prediktor yang berpengaruh signifikan jumlah kematian bayi di Kabupaten Probolinggo tahun 2020. Selanjutnya untuk mengetahui variabel yang berpengaruh secara parsial terhadap jumlah kematian bayi di Kabupaten Probolinggo tahun 2020 dilakukan pengujian dengan hipotesis sebagai berikut:

H_0 : $\beta_i = 0, i = 7, 8, 10$ atau variabel prediktor ke-i tidak memberikan pengaruh yang signifikan terhadap model secara parsial.

H_1 : $\beta_i \neq 0$ atau variabel variabel prediktor ke-i memberikan pengaruh yang signifikan terhadap model secara parsial.

Hipotesis diputuskan tolak H_0 jika statistik uji Z_{hitung} yang didapatkan nilainya lebih dari $Z_{\left(\frac{0,15}{2}\right)}$ yang diperkuat dengan nilai *p-value* yang nilainya kurang dari α sebesar 0,15.

Tabel 6 menunjukkan bahwa nilai $|Z_{hitung}|$ variabel pada prediktor X_7, X_8, dan X_{10} lebih besar daripada $Z_{(0.15/2)}$ sebesar 1,43 yang diperkuat dengan nilai *p-value* yang nilainya kurang dari taraf signifikan (α) sebesar 0,15 sehingga diputuskan tolak H_0 yang artinya artinya variabel jumlah bayi lahir rendah (X_7), jumlah ibu hamil mendapat imunisasi Td2+ (X_8), dan umlah tenaga kesehatan (X_{10}) berpengaruh signifikan terhadap jumlah kematian bayi di Kabupaten Probolinggo tahun 2020 dengan model NBR sebagai berikut:

$$\hat{\mu} = e^{0,7070+0,0283X_7-0,0006X_8+0,0031X_{10}}$$

Didapatkan nilai parameter dispersi sebesar 5,13 yang nilainya lebih besar daripada nol sehingga pemodelan NBR sesuai dan dapat digunakan untuk mengatasi overdispersi. Berdasarkan model yang terbentuk maka dapat diketahui bahwa jika setiap peningkatan 1 jumlah bayi lahir rendah akan meningkatkan rata-rata jumlah kasus kematian bayi sebesar exp(0,0283) = 1,0287 kali lipat dari jumlah sebelumnya dengan asumsi variabel lain konstan, setiap peningkatan 1 ibu hamil yang mendapat imunisasi Td2+ akan mengurangi rata-rata jumlah kematian bayi sebesar exp(-0,0006) = 0,999 kali lipat nilai sebelumnya dengan asumsi variabel lain konstan, dan setiap peningkatan 1 tenaga kesehatan akan meningkatkan rata-rata kematian bayi sebesar exp(0,00311)=1,0031 kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan.

Prediksi dapat dilakukan berdasarkan hasil model yang didapatkan. Jika terdapat satu bayi lahir rendah, 1 ibu hamil yang mendapat imunisasi Td2+ dan terdapat 1 tenaga kesehatan makadiprediksi akan terdapat 2,23 ≈ 2 kasus kematian bayi di Kabupaten Probolinggo tahun 2020.

F. Pemilihan Model Terbaik GPR dan NBR

Pemodelan Pemilihan model terbaik dari pemodelan dengan metode GPR dan NBR didasarkan pada kriteria kebaikan model dengan nilai AIC, AICc, BIC, dan BICc terkecil. Hasil nilai kebaikan model pada masing-masing metode disajikan pada Tabel 7 yang menunjukkan bahwa baik pada metode GPR maupun NBR memiliki model dengan variabel signifikan yang sama yaitu β_0, β_7, β_8, dan β_{10}. Pemodelan menggunakan metode GPR memiliki keseluruhan nilai kriteria kebaikan model yang lebih kecil dibandingkan dengan metode NBR. Berdasarkan hal tersebut maka pemodelan jumlah kematian bayi di Kabupaten Probolinggo untuk mengatasi overdispersi lebih baik menggunakan metode GPR. Namun, perbedaan nilai kriteria kebaikan model dari metode GPR dan NBR tidak berbeda secara signifikan.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis statistika deskriptif, GPR, dan NBR, serta perbandingan hasil analisis maka data diperoleh kesimpulan sebagai berikut: (1) Secara umum, gambaran jumlah kematian bayi di Kabupaten Probolinggo tahun 2020 menunjukkan tingkat keragaman yang tinggi. Kecamatan dengan jumlah kasus kematian bayi tertinggi adalah Kecamatan Paiton. Sedangkan daerah yang tidak memiliki jumlah kematian bayi adalah Kecamatan Sukapura. Daerah dataran tinggi yang berada pada bagian barat Kabupaten Probolinggo yaitu Kecamatan Sukapuran, Sumber, Lumbang, dan Kuripan mayoritas termasuk dalam kategori rendah pada seluruh variabel. (2) Analisis Generalized Poisson Regression dan Negative Binomial Regression menunjukkan hasil yang sama berdasarkan kombinasi dari variabel prediktor yang telah terbebas dari multikolinearitas. Model terpilih adalah model dengan variabel prediktor adalah kombinasi dari variabel jumlah bayi lahir rendah (X_7), jumlah ibu hamil mendapat imunisasi Td2+ (X8), dan jumlah tenaga kesehatan (X10). Hasil pengujian secara serentak maupun parsial menunjukkan bahwa keseluruhan parameter berpengaruh signifikan terhadap model, sehingga dapat disimpulkan bahwa jumlah bayi lahir rendah (X7), jumlah ibu hamil mendapat imunisasi Td2+ (X8), dan jumlah tenaga kesehatan (X10) berpengaruh terhadap jumlah kematian bayi (Y). (3) Hasil perbandingan metode Generalized Poisson Regression dan Negative Binomial Regression menunjukkan hasil yang tidak berbeda secara signifikan. Parameter signifikan yang diperoleh dari kedua metode sama. Namun, dilihat dari kebaikan model yang dihasilkan, metode Generalized Poisson Regression menunjukkan nilai kriteria kebaikan model yang lebih kecil baik pada nilai AIC, AICc, BIC, maupun BICc. Sehingga untuk mengatasi kondisi overdispersi pada pemodelan kasus jumlah kematian bayi di Kabupaten Probolinggo lebih baik menggunakan metode Generalized Poisson Regression. hasil analisis dan pembahasan dapat diambil kesimpulan bahwa lulusan SMK Negeri 3 Boyolangu tahun 2018 dengan keputusan setelah lulus adalah melanjutkan pendidikan sebesar 42.7%. Sedangkan berdasarkan model yang terbentuk peluang lulusan SMK melanjutkan pendidikan sebesar 93.7% dan ketepatan klasifikasinya sebesar 79.2%. Faktor yang berpengaruh signifikan terhadap keputusan setelah lulus untuk melanjutkan pendidikan adalah harapan orang tua dan peringkat 5 besar.

Berdasarkan hasil analisis didapatkan variabel yang berpengaruh terhadap jumlah kematian bayi adalah jumlah bayi berat lahir rendah, jumlah ibu hamil mendapat imunisasi Td2+, dan jumlah tenaga kesehatan. Sehingga saran untuk pemerintah khususnya dinas kesehatan Kabupaten Probolinggo untuk lebih memeperhatian variabel tersebut.

Saran mengenai upaya yang dapat dilakukan diantaranya dengan mengurangi kasus terjadinya BBLR (Berat Bayi Lahir Rendah) sehingga jumlah kematian bayi akibat BBLR dapat berkurang. Selain itu perlu dilakukan penambahan jumlah tenaga kesehatan khusus nya tenaga kesehatan seperti dokter, bidan, dan perawat yang menangani langsung masalah bayi, kehamilan, serta kelahiran. Pemerataan jumlah tenaga kesehatan di setiap kecamatan juga perlu diperhatikan, karena berdasarkan analisis statistika deskriptif hanya terdapat 4 kecamatan dari total 24 kecamatan yang memiliki jumlah tenaga kesehatan dalam kategori sedang hingga tinggi sementara sisanya yaitu 20 kecamatan memiliki jumlah tenaga kesehatan yang tergolong rendah.

Saran untuk penelitian selanjutnya terkait dengan kematian bayi adalah melakukan penelitian dengan menggunakan data individu dengan metode regresi logistik sehingga faktor yang mempengaruhi kematian bayi beserta peluang terjadinya kematian bayi pada ibu hamil dapat diketahui secara lebih mendalam.

DAFTAR PUSTAKA

[1] A. K. Wandira and R. Indawati, “Faktor penyebab kematian bayi di kabupaten sidoarjo,” *Jurnal Biometrika dan Kependudukan*, vol. 1, no. 1, pp. 33–42, 2012.

[2] Dinas Kesehatan Kabupaten Probolinggo, *Profil Kesehatan Kabupaten Probolinggo Tahun 2021*. Kabupaten Probolinggo: Dinas Kesehatan Kabupaten Probolinggo, 2021.

[3] R. A. Putri, “Penentuan Indikator Kesehatan Bayi dengan Menggunakan Analisis Faktor,” Departemen Matematika, Universitas Andalas, Padang, 2017.

[4] S. N. Aulele, “Pemodelan jumlah kematian bayi di provinsi maluku tahun 2010 dengan menggunakan regresi poisson,” *Jurnal Barekeng*, vol. 5, no. 2, pp. 23–27, 2012.

[5] N. S. Umami, D. Ispriyanti, and T. Widiharih, “Aplikasi model regresi poisson tergeneralisasi pada kasus angka kematian bayi di jawa tengah tahun 2007,” *Jurnal Gaussian*, vol. 2, no. 4, pp. 361–368, Oct. 2013.

[6] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed. New York: John Wiley & Sons, 1998. ISBN 9780471170822.

[7] R. R. Hocking, “Methods and applications of linear models: regression and analysis of variance,” *Eur J Orthod*, vol. 19, no. 2, p. 232, Apr. 1996, doi: 10.1093/ejo/19.2.232.

[8] R. E. Walpole, *Probability and Statistics for Engineers and Scientists*. 2012. ISBN 9780321629111.

[9] A. Agresti, *Categorical Data Analysis*, 2nd ed. New York: John Wiley & Sons, 2002. doi: 10.1002/0471249688. ISBN 9780471249689.

[10] P. McCullagh and J. A. Nedler, *Generalized Linear Models*, 2nd ed. New York: CRC Press, 1989. ISBN 9780412317606.

[11] P. C. Consul and F. Famoye, “Generalized poisson regression model,” *Commun Stat Theory Methods*, vol. 21, no. 1, pp. 89–109, Jan. 1992, doi: 10.1080/03610929208830766.

[12] J. W. H. J. M. H. J. H. James W. Hardin, *Generalized Linear Model and Extentisions*, 2nd ed. United States of America: Stata Press, 2007. ISBN 9781597180146.

[13] W. Wang and F. Famoye, “Modeling household fertility decisions with generalized poisson regression,” *J Popul Econ*, vol. 10, no. 3, pp. 273–283, 1997, doi: 10.1007/s001480050043.

[14] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*, 2nd ed. United States of America: Cambridge University Press, 2013. ISBN 9781107014169.

[15] H. Bozdogan, “Akaike’s information criterion and recent developments in information complexity,” *J Math Psychol*, vol. 44, no. 1, pp. 62–91, 2000, doi: 10.1006/jmps.1999.1277.