

# Pemodelan Jumlah Kematian Bayi di Provinsi Jawa Barat dengan Pendekatan Geographically Weighted Negative Binomial Regression (GWNBR)

Irfana Mualana Ismail<sup>1</sup>, Tiani Wahyu Utami<sup>2</sup>, M. Al Haris<sup>3</sup>

<sup>1,2,3</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

Alamat e-mail : [irfanamaulanablora@gmail.com](mailto:irfanamaulanablora@gmail.com)

## ABSTRAK

Angka kematian bayi (AKB) merupakan indikator yang sangat berguna untuk mengetahui status kesehatan anak dan kondisi ekonomi penduduk suatu wilayah dan untuk melihat target penurunan angka kematian bayi. Salah satu cara yang dapat dilakukan untuk menurunkan angka kematian bayi adalah dengan mengetahui faktor-faktor penyebabnya. Pemodelan pada data count berdistribusi Poisson sering menggunakan metode regresi Poisson. Banyaknya pelanggaran asumsi Regresi Poisson salah satunya adalah Overdispersi. Salah satu metode yang dapat menganalisis adanya overdispersi pada data count adalah dengan regresi Binomial Negatif yang menghasilkan penaksir parameter bersifat global. Akan tetapi data spasial dari berbagai lokasi mewakili perbedaan kondisi di setiap lokasi. Hal itu dapat dipengaruhi oleh karakteristik masyarakat (sosial budaya), kondisi geografis dan ekonomi sebagai gambaran adanya faktor spasial. Tujuan dari penelitian ini adalah untuk memodelkan jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat dan faktor-faktor yang diduga mempengaruhinya dengan metode Geographically Weighted Negative Binomial Regression (GWNBR). Hasil penelitian menunjukkan bahwa metode GWNBR dalam kasus angka kematian bayi lebih baik dibandingkan dengan metode Regresi binomial dengan AIC sebesar 271.721. Hasil analisis pemodelan tingkat angka kematian bayi di Provinsi Jawa Barat tahun 2017 hasil pemodelan GWNBR dengan fungsi pembobot kernel adaptive bisquare didapatkan pengelompokan sebanyak 2 kelompok berdasarkan variabel-variabel yang signifikan.

**Kata Kunci :** Angka Kematian Bayi, Overdispersi, Geographically Weighted Negative Binomial Regression (GWNBR)

## ABSTRACT

*The infant mortality rate (IMR) is a very useful indicator to determine the status of children's health and economic conditions of the population of a region and to see the infant mortality goals. One way that can be done to lower the infant mortality rate is to determine the contributing factors. Factors suspected to affect infant mortality rate in the province of West Java is the number of Low Birth Weight, Neonatal Visits 1 and Scope vitamin A. Modeling the data count Poisson distributed often using Poisson regression methods. Assuming the number of violations Poisson regression one is overdispersi. One method that can analyze the data count is overdispersi with Negative Binomial regression parameter estimator that generates global nature. However, the spatial data from different locations representing different conditions in each location. It can be influenced by the characteristics of the community (social and cultural), geographical and economic conditions as a picture of the spatial factor. The purpose of this study was to model the number of cases of infant mortality in West Java province and suspected factors influencing the Negative Binomial method Geographically Weighted Regression (GWNBR). The results showed that the method GWNBR in the case of infant mortality better than the binomial regression method to the AIC at 271.721. The result of modeling analysis of infant mortality rates in West Java province in 2017 of the modeling results of GWNBR with the adaptive bisquare kernel weighting function obtained by grouping as many as 2 groups based on significant variables.*

**Keywords :** Number of Infant Mortality, Overdispersi, Geographically Weighted Negative Binomial Regression (GWNBR)

## PENDAHULUAN

Angka kematian merupakan indikator outcome pembangunan kesehatan. Angka kematian dapat menggambarkan seberapa tinggi derajat kesehatan masyarakat di suatu wilayah. Pada dasarnya penyebab kematian ada yang langsung dan tidak langsung, walaupun dalam kenyataan terdapat interaksi dari berbagai faktor yang mempengaruhi terhadap tingkat kematian di masyarakat. Angka kematian bayi (AKB) merupakan indikator yang sangat berguna untuk mengetahui status kesehatan anak dan kondisi ekonomi penduduk suatu wilayah dan untuk melihat target penurunan angka kematian bayi. Salah satu cara yang dapat dilakukan untuk menurunkan angka kematian bayi adalah dengan mengetahui faktor-faktor penyebabnya (Elyna, 2012).

Berbagai faktor yang berkaitan dengan penyebab kematian, baik langsung maupun tidak langsung, antara lain dipengaruhi oleh tingkat sosial ekonomi, kualitas lingkungan hidup, upaya pelayanan kesehatan dan lain-lain. Di Provinsi Jawa Barat beberapa faktor penyebab kematian perlu mendapat perhatian khusus, diantara yang berhubungan dengan kematian ibu dan bayi yaitu besarnya tingkat kelahiran, umur masa paritas, jumlah anak yang dilahirkan serta penolong persalinan. Angka Kematian Bayi (AKB) atau Infant Mortality Rate (IMR) merupakan indikator yang sangat sensitif terhadap upaya pelayanan kesehatan terutama yang berhubungan dengan bayi baru lahir perinatal dan neonatal.

Dalam penelitian ini yang akan diteliti yaitu jumlah kematian bayi di Provinsi Jawa Barat. Dimana jumlah kematian bayi merupakan salah satu indikator penting bagi pemerintah dalam mengevaluasi dibidang kesehatan. Dengan mengetahui faktor-faktor yang mempengaruhi, tentu saja pemerintah memiliki gambaran langkah apa saja yang harus dilakukan dalam menekan angka kematian bayi. Mengingat dalam tahun terakhir yaitu

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinearitas. Pendeteksian kasus multikolinearitas yaitu dengan koefisien korelasi dan nilai VIF (Variance Inflation Factor). Terdapat kasus multikolinearitas jika nilai VIF lebih besar dari 10.

Analisis regresi merupakan salah satu metode untuk menentukan adanya hubungan sebab

akibat antara satu variabel dan variabel yang lain. Analisis regresi sendiri sangat luas pemakaiannya karena ada model pada analisis regresi yang dapat digunakan secara baik hanya pada kondisi tertentu. Salah satu contoh kondisinya ketika data variabel respon yang dijumpai adalah data cacah seperti jumlah kematian bayi dengan sebaran Poisson maka regresi Poisson menjadi cocok untuk digunakan. Namun, regresi Poisson mempunyai asumsi yang harus dipenuhi yaitu rata-rata variabel respon harus sama dengan variansinya yang dikenal dengan istilah equidispersi (Dobson dan Barnett, 2008). Namun, dalam kenyataannya tidak semua data cacah ini memiliki nilai rata-rata yang sama dengan variansinya. Giuffre et al. (2011) menyatakan bahwa yang sering terjadi pada data poisson adalah kondisi rata-rata yang lebih kecil dari variansinya atau lebih dikenal dengan istilah overdispersi.

Overdispersi dapat terjadi karena ada data yang berkelompok dalam populasi (McCullagh dan Nelder, 1989). Jika data dalam kelompok tersebut berkorelasi positif maka analisis dengan metode yang mengasumsikan kebebasan antar elemen akan menghasilkan penduga yang underestimate atau varians yang lebih kecil dari nilai sebenarnya (Astuti, 2006). Dari permasalahan overdispersi tersebut maka dibutuhkan model regresi yang lain, karena regresi Poisson menjadi kurang cocok digunakan untuk menganalisis data tersebut. Salah satu regresi yang tidak harus memenuhi asumsi equidispersi pada regresi Poisson adalah regresi Negative Binomial (NB). Regresi ini merupakan salah satu metode campuran poisson-gamma yang distribusi gamma-nya digunakan untuk mengatasi data overdispersi yang terjadi pada regresi Poisson (Hardin dan Hilbe 2007).

Pada dasarnya regresi Poisson maupun regresi Negative Binomial (NB) mempunyai kegunaan untuk menganalisis hubungan antara variabel respon (Y) data cacah dan satu atau lebih variabel penjelas (X). Akan tetapi, metode ini kurang representatif jika diterapkan pada data spasial atau data yang mengandung kondisi geografis (Widodo dkk, 2013). Di dalam penelitiannya Afri (2013) dijelaskan bahwa salah satu alat analisis untuk mengatasi

kasus data spasial tersebut dikembangkan oleh McMillen dan McDonald tahun 1997 dan model ini dinamai Geographically Weighted Regression (GWR) oleh Fotheringham et al. (2002). Kemudian, untuk kasus penanganan fenomena overdispersi dan pengaruh lokasi atau

aspek spasial pada data, penelitian selanjutnya menggunakan Geographically Weighted Binomial Regression (GWNBR).

GWNBR adalah salah satu solusi yang tepat untuk membentuk analisis regresi yang bersifat lokal untuk setiap lokasi pengamatan. Hasil analisisnya adalah nilai-nilai parameternya berlaku hanya pada tiap lokasi pengamatan dan berbeda dengan lokasi lainnya (Rahmawati dan Djuraidah, 2010). GWR merupakan bagian dari analisis spasial yang mempunyai pembobot berdasarkan posisi atau jarak satu lokasi dengan lokasi pengamatan lain. Unsur pembobot tersebut menentukan semakin dekat suatu lokasi, bobot pengaruhnya akan semakin besar.

Penelitian sebelumnya mengenai kasus Angka Kematian Bayi telah dilakukan oleh Afri (2013) menggunakan Model GWNBR untuk Data Kematian Bayi pada tahun 2008 di Jawa Timur. Jumlah kematian bayi adalah kejadian langka yang terjadi pada satuan waktu yang satu daerah dengan daerah lainnya memberikan pengaruh yang berbeda, sehingga variabel-variabel penelitian ini mengacu pada penelitian Afri (2013). Variabel yang didapat kemudian disesuaikan dengan kondisi data yang ada pada tahun 2012 terkait ketersediaan dan kelengkapan data. Kemudian wilayah yang dipilih untuk penelitian ini adalah kabupaten/kota di Pulau Jawa karena jumlah pengamatan lebih banyak dan karakteristik pengamatan lebih beragam.

Berdasarkan uraian yang telah dijelaskan, maka akan dibuat perbandingan dari dua metode yaitu Pemodelan Jumlah Kematian Bayi Di Provinsi Jawa Barat Tahun 2017 Dengan Pendekatan Geographically Weighted Negative Binomial Regression (GWPR). Untuk mendapatkan hasil model yang terbaik, maka perbandingan akan menggunakan perhitungan nilai AIC pada kasus Angka Kematian Bayi di Provinsi Jawa Barat.

## TINJAUAN PUSTAKA

### 1. Multikolinearitas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinieritas. Pendeteksian kasus multikolinieritas yaitu dengan koefisien korelasi dan kriteria nilai VIF (Variance Inflation Factor). Terdapat kasus multikolinieritas jika nilai VIF lebih besar dari 10. Nilai VIF dinyatakan sebagai berikut.

$$VIF_k = \frac{1}{1 - R_k^2} \quad (2.1)$$

$$\text{dimana } R_k^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (y_i - \bar{y})_k^2}{\sum_{i=1}^n (y_i - \bar{y})_k^2}$$

dengan  $R_k^2$  adalah koefisien determinasi antara satu variabel prediktor dengan variabel prediktor lainnya (Hocking, 1996).

### 2. Regresi Poisson

Model regresi Poisson merupakan model standar untuk data diskrit dan termasuk dalam model regresi nonlinier (Cameron, et al., 1998). Regresi Poisson digunakan pada data yang berdistribusi poisson. Probabilitas distribusi Poisson diberikan oleh Myers (1990).

$$p(y; \mu) = \frac{e^{-\mu} \mu^y}{y!} \quad (y = 0, 1, 2, \dots) \quad (2.2)$$

dimana  $\mu$  adalah mean distribusi Poisson. Parameter  $\mu$  sangat bergantung pada beberapa unit yang ditetapkan atau periode waktu, jarak, luas, volume, dan lain-lain. Distribusi Poisson digunakan untuk memodelkan peristiwa yang jarang terjadi selama periode tertentu. Probabilitas banyak kejadian  $y$  dalam periode waktu  $t$  yaitu:

$$p(y; \mu) = \frac{e^{-\mu t} (\mu t)^y}{y!} \quad (y = 0, 1, 2, \dots) \quad (2.3)$$

Persamaan tersebut digunakan untuk probabilitas kejadian  $y$ , dan rata-rata jumlah kejadian, berdasarkan asumsi bahwa rata-rata jumlah kejadian per periode waktu adalah konstan.

model regresi Poissonnya dapat ditulis sebagai berikut Myers, (1990):

$$y_i = \mu_i + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (2.4)$$

dimana  $y_i$  adalah jumlah kejadian, dan  $\mu_i$  adalah rata-rata jumlah kejadian dalam periode  $t_i$ .  $\mu_i$  diasumsikan tidak berubah dari data ke data. Persamaan distribusi Poisson dinyatakan dengan persamaan sebagai berikut:

$$p(y_i; \hat{\beta}) = \frac{e^{-[\mu(x_i; \hat{\beta})]} [\mu(x_i; \hat{\beta})]^{y_i}}{y_i!} \quad (2.5)$$

### 3. Overdispersi

Metode regresi Poisson mewajibkan equidispersi, yaitu kondisi dimana nilai mean dan varians dari variabel respon bernilai sama. Namun, adakalanya terjadi fenomena overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi berarti varians lebih besar daripada mean. Taksiran dispersi diukur dengan devians atau Pearson's Chi-Square yang dibagi derajat bebas. Data overdispersi jika taksiran dispersi lebih



besar dari 1 dan underdispersi jika taksiran dispersi kurang dari 1 (Khoshgoftaar, 2004 dalam Camelia P.S, Nur I.M dan Darsyah M.Y, 2016).

#### 4. Efek Spasial

Pemodelan pada data spasial dapat dikelompokkan berdasarkan tipe data spasial yang digunakan yaitu spasial titik dan spasial area. Masing-masing tipe data spasial tersebut dapat dikelompokkan lagi berdasarkan jenis data yang digunakan yaitu cross sectional dan time series. Pemodelan data spasial selalu melibatkan matriks pembobot spasial. Sedangkan efek spasial pada data dapat berupa error yang saling berkorelasi (dependensi spasial) maupun keragaman (heterogenitas) spasial antar lokasi.

Pengujian heterogenitas spasial dilakukan untuk melihat ciri khas pada setiap lokasi pengamatan yang akan mengakibatkan parameter regresi yang dihasilkan berbeda secara spasial. Pengujian heterogenitas spasial dilakukan dengan statistik uji Breusch-Pagan (BP).

#### 5. Matriks Pembobot Spasial

Pembobot memiliki peranan penting pada data spasial, karena nilai suatu pembobot merupakan perwakilan dari lokasi dimana masing-masing data diambil. Informasi mengenai suatu lokasi dapat direpresentasikan oleh sebuah titik koordinat, seperti Garis Lintang dan Garis Bujur. Berdasarkan informasi spasial tersebut dapat diperhitungkan jarak titik koordinat antar lokasi sehingga diharapkan kekuatan dari dependensi spasial akan menurun dengan adanya jarak tersebut. Lokasi yang berdekatan seharusnya menunjukkan hubungan kemiripan, begitu juga sebaliknya. Lokasi yang berjauhan juga memperlihatkan adanya keragaman spasial. Keragaman spasial antara lokasi yang satu dengan lokasi yang lain ditunjukkan dengan adanya matriks pembobot,  $W(u_i, v_i)$ , yang entri-entrinya merupakan fungsi dari jarak Euclidian antar lokasi. Besarnya pembobotan untuk model GWR di setiap lokasi dapat ditentukan dengan menggunakan fungsi kernel Adaptive Bisquare kernel, Penggunaan metode Adaptive Kernel sangat cocok apabila suatu pengamatan tersebar dengan pola yang acak, tidak beraturan dan berkelompok. Metode Adaptive Kernel memungkinkan untuk mendapatkan nilai bandwidth yang berbeda untuk setiap titik pengamatan. Hal ini dikarenakan metode Adaptive Kernel dapat menyesuaikan dengan kondisi titik pengamatan (Fotheringham, dkk,2002).

#### 6. Regresi Binomial Negatif

Regresi binomial negatif merupakan salah satu solusi untuk mengatasi adanya kasus overdispersi. Model regresi binomial negatif memiliki fungsi massa peluang sebagai berikut (Greene, 2008).

$$f(y, \mu, \theta) = \frac{\Gamma(y + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y+1)} \left(\frac{1}{1 + \theta\mu}\right)^{\theta^{-1}} \left(\frac{\theta\mu}{1 + \theta\mu}\right)^y, y = 0, 1, 2, \dots$$

Pendugaan parameter model regresi binomial negatif menggunakan metode Maximum Likelihood Estimation (MLE). Kemudian untuk mendapatkan nilai dari parameternya digunakan iterasi Newton Raphson. Fungsi likelihood dari regresi binomial negatif adalah sebagai berikut.

$$L(\beta, \theta) = \prod_{j=1}^n \left\{ \frac{\Gamma(y_j + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y_j+1)} \left(\frac{1}{1 + \theta\mu_j}\right)^{\theta^{-1}} \left(\frac{\theta\mu_j}{1 + \theta\mu_j}\right)^{y_j} \right\} \quad (2.24)$$

#### 7. Geographically Weighted Negative Binomial Regression (GWNBR)

Geographically Weighted Negative Binomial Regression (GWNBR) merupakan salah satu metode yang cukup efektif menduga data yang memiliki heterogenitas spasial untuk data count yang over dispersi. Model GWNBR merupakan pengembangan dari model regresi binomial negatif. Model GWNBR menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Model GWNBR dapat dirumuskan sebagai berikut (Ricardo & Carvalho, 2013)

$$y_i \sim NB[\exp(\beta_j(u_i, v_i)x_k), \theta(u_i, v_i)], i = 1, 2, \dots, n \quad (2.26)$$

dimana,

$y_i$ : nilai observasi respon ke- $i$

$x_k$ : nilai observasi variabel prediktor ke- $k$  pada pengamatan lokasi ( $u_i, v_i$ )

$\beta_j(u_i, v_i)$ : koefisien regresi variabel prediktor ke- $j$  untuk setiap lokasi ( $u_i, v_i$ )

$\theta(u_i, v_i)$ : parameter disperse untuk setiap lokasi ( $u_i, v_i$ )

Estimasi model GWNBR menggunakan Maximum Likelihood Estimation (MLE). Langkah awal dari metode Likelihood adalah membentuk fungsi Likelihood. Pengujian signifikansi parameter model GWNBR terdiri dari uji serentak dan parsial. Uji signifikansi serentak menggunakan Maximum Likelihood Ratio Test (MLRT).

#### 8. Kriteria Kebaikan Model

Akaike Information Criterion (AIC) adalah kriteria kesesuaian model dalam menduga model secara statistik. Kriteria AIC digunakan apabila pemodelan regresi bertujuan untuk

mengidentifikasi faktor-faktor yang berpengaruh terhadap model. Besarnya nilai AIC sejalan dengan nilai devians dari model. Semakin kecil nilai devians maka akan semakin kecil pula tingkat kesalahan yang dihasilkan model sehingga model yang diperoleh menjadi semakin tepat. Nilai devians akan semakin kecil apabila rasio antara fungsi likelihood di bawah  $H_0$  dengan fungsi likelihood di bawah populasi semakin besar. Hal ini mengindikasikan bahwa parameter yang diuji semakin mendekati nilai parameter populasi yang sebenarnya yang berarti dugaan model semakin baik. Oleh karena itu, model terbaik adalah model dengan AIC terkecil dan dengan devians terkecil pula.

## METODE PENELITIAN

### 1. Sumber Data

Data yang akan digunakan dalam penelitian ini merupakan data sekunder pada tahun 2017 yang diperoleh dari publikasi Badan Pusat Statistik (BPS) dan data kesehatan dari Dinas Kesehatan yaitu Profil Kesehatan Provinsi Jawa Barat. Unit observasi yang digunakan dalam penelitian ini adalah sebanyak 27 Kabupaten/Kota di Jawa Barat.

### 2. Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini merupakan data jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat menurut Kabupaten/Kota tahun 2017 sebagai variabel respon dan 3 faktor yang berhubungan dengan penyakit angka kematian bayi sebagai variabel predictor.

$Y$  = Jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat menurut Kabupaten/Kota tahun 2017

$X_1$  = Banyaknya Berat Badan Lahir Rendah di Jawa Barat menurut Kabupaten/Kota tahun 2017

$X_2$  = Persentase Kunjungan Neonatal 1 di Jawa Barat menurut Kabupaten/Kota tahun 2017

$X_3$  = Persentase Cakupan Vitamin A di Jawa Barat menurut Kabupaten/Kota tahun 2017

### 3. Analisis Data

Berikut merupakan langkah-langkah untuk analisis faktor-faktor yang mempengaruhi jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat:

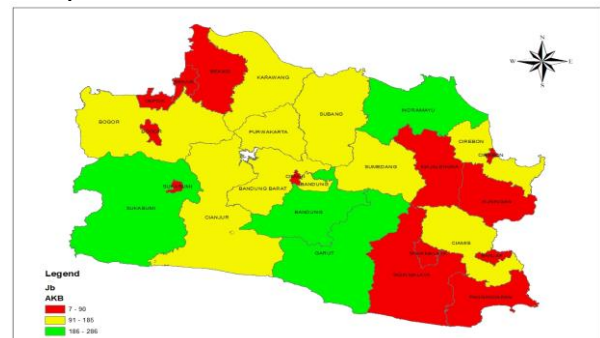
1. Mendeskripsikan gambaran umum dari data jumlah kasus angka kematian bayi di Provinsi Jawa Barat serta variabel-variabel yang diduga mempengaruhi.
2. Mendeteksi adanya multikolinieritas.
3. Melakukan pemodelan regresi Poisson.

4. Melakukan pengujian overdispersi.
5. Jika terjadi overdispersi, kemudian dilanjutkan dengan pendekatan model regresi Binomial Negatif.
6. Menentukan model regresi Binomial Negatif
7. Melakukan pengujian kesesuaian model regresi Binomial Negatif
8. Melakukan pengujian aspek data spasial. Uji Moran's I untuk dependensi spasial data dan Breusch Pagan untuk heterogenitas spasial data.
9. Mendapatkan model Geographically Weighted Negative Binomial Regression (GWNBR) dengan langkah-langkah sebagai berikut.
10. Pengujian signifikansi parameter secara serentak
11. Pengujian signifikansi parameter secara parsial
12. Interpretasi model dan membentuk peta pengelompokan.
13. Membandingkan model Regresi Binomial Negative dengan GWNBR dengan menggunakan nilai AIC.
14. Menarik kesimpulan.

## HASIL DAN PEMBAHASAN

1. Persebaran Jumlah Kasus Angka Kematian Bayi per Kabupaten dan Kota di Jawa Barat

Jumlah kasus angka kematian bayi dan faktor yang mempengaruhinya dikategorikan menjadi tiga kelompok yaitu tinggi, sedang, dan rendah. Hal ini bertujuan untuk mempermudah menginterpretasi hasil dari pemetaan. Berikut ini hasil dari pemetaan variabel respon yang digunakan dalam penelitian dengan menggunakan Kabupaten dan kota yang ada di Kabupaten/Kota Jawa Barat.



### Persebaran Angka Kematian Bayi di Jawa Barat Berdasarkan Kabupaten/Kota tahun 2017

Wilayah yang masuk dalam kategori tinggi dalam jumlah kasus angka kematian bayi adalah Kabupaten atau Kota Jawa Barat adalah Kota Sukabumi, Bandung, Garut, dan Indramayu dengan jumlah kasus antara 186-286 kasus. Untuk jumlah

Parameter	Estimate	z-value	P-value
$\beta_0$	4.869	42.519	$< 2e - 16$ (***)
$\beta_1$	0.0818	6.504	$7.84e - 11$ (***)
$\beta_2$	-0.0068	-5.715	$1.09e - 08$ (***)
$\beta_3$	0.0096	13.959	$< 2e - 16$ (***)
Devians	1073.3		

kasus angka kematian bayi kategori sedang, terdapat di Cianjur, Bandung Barat, Purwakarta, Ciamis, Sumedang, Kabupaten Bandung, Cirebon, Subang, Karawang, dan Kota Bogor dengan interval kasus 91-185 kasus. Kabupaten atau Kota lain yang tidak disebutkan di atas merupakan Kabupaten atau Kota dengan jumlah kasus Angka Kematian Bayi kategori rendah dengan interval 7-90 kasus.

## 2. Multikolinearitas

Uji multikolinearitas merupakan penyujian untuk melihat apakah terdapat hubungan antara masing-masing variabel bebas dengan variabel bebas lainnya. Apabila jika terdapat multikolinearitas, maka harus ada variabel yang harus direduksi sampai tidak terdapat korelasi antar variabel bebasnya. Asumsi Multikolinearitas ini harus terpenuhi agar bisa dilakukan pengujian menggunakan pendekatan Binomial Negatif Seperti yang sudah dijelaskan pada subbab 2.2.1 terdapat tiga kriteria yang bisa digunakan untuk melihat adanya multikolinearitas, salah satunya adalah nilai VIF. Berikut adalah table hasil nilai VIF dari masing-masing variabel bebas.

Tabel 4.1 Hasil VIF

Variabel Bebas	Nilai VIF
X1	1,327
X2	1,110
X3	1,358

Dari tabel diatas, dapat dilihat bahwa tidak terjadi multikolinearitas atau tidak terdapat hubungan antar variabel bebas, karena semua variabel bebas menunjukkan nilai VIF yang kurang dari 10 sehingga variabel di atas dapat digunakan untuk membentuk regresi poisson, maka seluruh variabel prediktor dapat digunakan untuk pembentukan model regresi poisson dan regresi binomial negatif.

## 3. Regresi poisson

Digunakan jika data Angka Kematian Bayi di Jawa Barat berdistribusi poisson karena datanya

berubah data diskrit atau data count (jumlahan), hal ini dapat dibuktikan dengan Tabel dibawah ini: Tabel 4.2 Nilai Estimasi Parameter Model Regresi Poisson

Pengujian distribusi Poisson pada variabel dependen membuktikan bahwa variabel dependen berdistribusi Poisson. Pengujian ini menggunakan uji Kolmogorof-Smirnov, dimana hipotesis pengujiannya sebagai berikut :

$H_0$  :  $F(y) = F^*(y)$  (Variabel random y mengikuti distribusi Poisson)

$H_1$  :  $F(y) \neq F^*(y)$  (variabel random y tidak mengikuti distribusi Poisson)

Taraf signifikansi yang digunakan untuk menguji distribusi dan pengujian selanjutnya adalah  $\alpha = 0.1$ , Kriteria pengujian dalam uji Kolmogorof-Smirnov adalah tolak  $H_0$  jika  $\alpha$  pada taraf signifikansi atau nilai signifikansi. Nilai p-value untuk uji Kolmogorof-Smirnov adalah  $< 2e - 16$ (\*\*\*) sehingga gagal tolak  $H_0$  yang artinya variabel dependen atau variabel respon berdistribusi Poisson.

Setelah mendapatkan nilai estimasi parameter model regresi poisson, maka langkah selanjutnya adalah melakukan pengujian Overdispersi.

## 4. Overdispersi

Salah satu asumsi yang harus dipenuhi dalam regresi poisson adalah kondisi ekuidispersi. Namun dapat diketahui bahwa nilai residual devians sebesar 1073.3 dengan derajat bebas sebesar 23. Nilai hasil bagi antara residual devians dan derajat bebas pada penelitian ini adalah 46.665. Sehingga dapat disimpulkan bahwa pada model regresi poisson untuk jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat terdapat kasus overdispersi.

## 5. Pengujian Aspek Data Spial

Pengujian aspek data spasial dilakukan dengan pengujian, yaitu pengujian heterogenitas spasial. Dari uji Breusch-Pagan diperoleh nilai p-value sebesar 0.03777. Jika dibandingkn dengan  $\alpha$  sebesar 10% maka dapat disimpulkan bahwa  $H_0$  ditolak atau variansi antar lokasi berbeda. Sehingga terdapat perbedaan karakteristik antara satu titik pengamatan dengan titik pengamatan lainnya.

## 6. Regresi Binomial Negatif

Metode regresi binomial negatif juga dapat digunakan untuk menangani kasus overdispersi yang terjadi pada regresi poisson. Semua variabel prediktor dalam penelitian akan digunakan untuk memodelkan jumlah kasus Angka Kematian Bayi di Provinsi Jawa Barat menggunakan regresi



binomial negatif. Hasil estimasi ditunjukkan dalam tabel 4.3.

Parameter	Estimasi	SE	Z hitung
$\widehat{\beta}_0$	4.740	0.733	6.467
$\widehat{\beta}_1$	0.144	0.085	1.692
$\widehat{\beta}_2$	-0.010	0.008	-1.355
$\widehat{\beta}_3$	0.019	0.006	3.325
Devians	28.781		
AIC	306.81		

Pengujian parameter model regresi binomial negatif secara serentak menggunakan nilai devians ditunjukkan yang oleh tabel 4.3 sebesar 28.781. Dengan menggunakan  $\alpha$  sebesar 10% sehingga nilai  $\chi^2_{(0,1;3)}$  sebesar 6.251. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai  $\chi^2_{(0,1;3)}$ . Sehingga dapat disimpulkan bahwa  $H_0$  ditolak yang berarti minimal terdapat satu variabel prediktor yang berpengaruh terhadap variabel respon.

Pengujian parameter secara parsial menggunakan statistik uji  $|Z_{hitung}|$  yang dibandingkan dengan  $Z_{0,1/2} = 1,634$ . Tabel 4.3 menunjukkan bahwa nilai  $|Z_{hitung}|$  dari variabel  $X_1$  dan  $X_3 > 1,634$ . Maka dapat diambil keputusan bahwa dari ketiga variabel prediktor variabel  $X_1$  dan  $X_3$  yang gagal tolak  $H_0$  atau variabel  $X_1$  dan  $X_3$  berpengaruh secara signifikan terhadap variabel  $Y$  dengan menggunakan model regresi binomial negatif. Sehingga persamaan model regresi binomial negatif yang diperoleh sebagai berikut:

$$\mu = \exp(4.740 + 1.44X_1 + 0.019X_3)$$

#### 7. Pemodelan Geographically Weighted Negative Binomial Regression (GWNBR)

Pemodelan GWNBR merupakan bentuk lokal dari regresi binomial negatif karena memperhatikan letak geografis sehingga perlu adanya pembobot lokasi. Sebelum menghitung matriks pembobot yang harus dilakukan yaitu mencari jarak Euclidean ( $d_{ij}$ ) antar provinsi terlebih dahulu. Untuk nilai bandwidth optimum (h) pemodelan GWNBR didapatkan dengan melihat hasil Cross Validation (CV) terkecil. Dari hasil pada lampiran didapatkan bandwidth nilai CV terkecil yaitu 0.658653. Setelah mendapatkan nilai  $d_{ij}$  pada wilayah yang ditaksir parameternya, selanjutnya matriks pembobot spasial disusun berdasarkan persamaan (2.19) dengan menggunakan bandwidth optimum.

matriks pembobot Adaptive Bisquare Kernel adalah sebagai berikut:

$$W(i)(u_i, v_i) = \text{diag}[1; 0.297849; 0.560544; 0; \dots \dots \dots; W_{27}(u_i, v_i)]$$

Dalam penentuan matriks pembobot akan dipilih salah satu pembobot yang terbaik dengan melihat nilai AIC terkecil.

Pembentukan matriks pembobot digunakan untuk menaksir parameter berdasarkan lokasi ( $u_1, v_1$ ). Jadi untuk penaksiran parameter dari lokasi ( $u_2, v_2$ ) sampai ( $u_{35}, v_{35}$ ) menggunakan matriks yang berbeda.

#### 8. Estimasi Parameter GWNBR

##### a. Pengujian Kesamaan Model

Pengujian pertama yang dilakukan adalah pengujian kesamaan model antara model regresi binomial negatif dengan model GWNBR, dengan hipotesis sebagai berikut.

Model	Devians	$F_{hitung}$
NBR	35.320	2.54
GWNBR	4239.32	

Berdasarkan hasil penghitungan didapatkan nilai  $F_{hitung}$  sebesar 4239.326 dengan menggunakan alpha sebesar 10% didapatkan nilai  $F_{(0,1;2;24)} = 2.54$ , yang berarti bahwa nilai  $F_{hitung}$  lebih besar daripada  $F_{tabel} = 2.54$  sehingga dapat diputuskan Tolak  $H_0$  atau terdapat perbedaan signifikan antara model regresi Binomial Negatif dengan Model GWNBR.

##### b. Pengujian Serentak

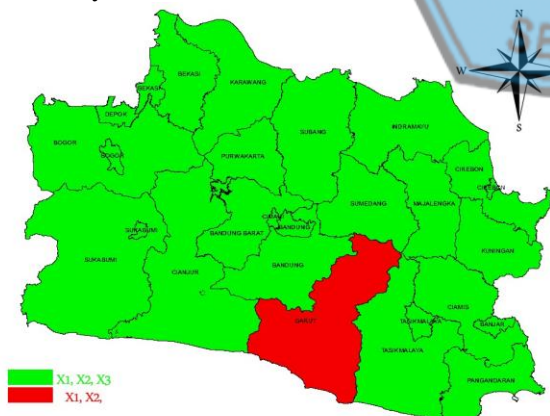
Nilai devians yang dihasilkan sebesar 6.924 dan menggunakan  $\alpha$  sebesar 10% sehingga nilai  $\chi^2_{(0,1;3)}$  sebesar 6.25139. Hal ini menunjukkan bahwa nilai devians lebih besar dari  $\chi^2_{(0,1;3)}$ . Maka dapat disimpulkan bahwa  $H_0$  ditolak yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model regresi GWNBR.

##### c. Pengujian Parsial

Hasil pengujian secara parsial menghasilkan parameter yang berbeda untuk setiap kabupaten/kota. Nilai  $|Z_{hitung}|$  yang dibandingkan dengan  $Z_{0,1/2} = 1,634$ .  $H_0$  ditolak jika nilai  $|Z_{hitung}| > 1,634$ . Variabel yang signifikan di setiap kabupaten/kota di Jawa Barat disajikan dalam tabel 4.4.

Kelompok	Kabupaten/ Kota	Variabel yang signifikan
1	Bogor, Sukabumi, Cianjur, Bandung, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar	$X_1, X_2$ dan $X_3$
2	Garut	$X_1$ dan $X_2$

Variabel yang berpengaruh signifikan terhadap Angka Kematian Bayi di 26 Kabupaten/kota di Jawa Barat adalah Variabel  $X_1, X_2$  dan  $X_3$ . Variabel yang berpengaruh signifikan di satu Kabupaten/ Kota lainnya adalah variabel  $X_1$  dan  $X_2$ .



#### 9. Pemilihan Model Terbaik

Pemilihan model terbaik didasarkan pada nilai AIC. Model yang baik akan memiliki nilai AIC yang kecil. Nilai AIC untuk setiap model regresi ditampilkan dalam table berikut.

Metode	AIC
Regresi Binomial Negatif	306.81
GWNBR	271.7172

Dari tabel di atas dapat diketahui bahwa nilai AIC yang dihasilkan metode GWNBR lebih kecil dibandingkan dengan AIC Regresi Binomial Negatif, sehingga metode yang sesuai untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus Angka Kematian Bayi di setiap kabupaten/kota di Jawa Barat adalah metode GWNBR.

#### KESIMPULAN

- Berdasarkan hasil pemodelan GWNBR dengan fungsi pembobot kernel *adaptive bisquare* didapatkan pengelompokan sebanyak 2 kelompok berdasarkan variabel-variabel yang signifikan. Kelompok pertama adalah Bogor, Sukabumi, Cianjur, Bandung, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar. Kelompok kedua adalah Kota Garut.
- Model untuk *Geographically Weighted Negative Binomial Regression* (GWNBR) yang dihasilkan adalah:  $\ln(\hat{\mu}) = 1.317 + 0.002X_1 - 0.00X_2 + 6.93E - 05X_3$  untuk Kabupaten Bogor.
- Perbandingan nilai AIC yang dihasilkan oleh metode Regresi Binomial Negatif dan GWNBR dengan nilai AIC terkecil terdapat pada metode GWNBR sebesar 271.712, sehingga metode yang terbaik untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus Angka Kematian Bayi di setiap kabupaten/kota di Jawa Barat adalah metode GWNBR

#### DAFTAR PUSTAKA

- Afri, L. E., 2013, Model Regresi Binomial Negatif Terboboti Geografis Untuk Data Kematian Bayi. *Jurnal Edu Research*, 2(1), pp 15-26.
- Astuti, E. T., 2006, Uji Trend/Regresi untuk Data Over Dispersi, *Jurnal Statistika Ikatan Perstatistikan Indonesia*, 2, pp 179-186



- Cameron, A.C, and Trivedi, P.K. 1998. *Regression Analysis of Count Data*. Cambridge : Cambridge University Press.
- Departemen Kesehatan RI. 2010. Revisi Buku Panduan Penyelidikan dan Penanggulangan Kejadian Luar Biasa (Pedoman Epidemiologi Penyakit). Jakarta: Ditjen PP & PL.
- Dinas Kesehatan Provinsi Jawa Barat. *Profil Kesehatan Provinsi Jawa Barat Tahun 2012*. Semarang: Dinas Kesehatan Provinsi Jawa Barat
- Dobson, A. J., dan Barnett, A., 2008, *An Introduction to Generalized Linear Models*. CRC press Departemen Kesehatan RI. 2007. Revisi Buku Panduan Penyelidikan dan Penanggulangan Kejadian Luar Biasa (Pedoman Epidemiologi Penyakit). Jakarta: Ditjen PP & PL.
- Fotheringham A S, C. M. 1997. Geographically Weighted Regression: a Natural Evolution of the Expansion Method for Spatial Data Analysis. *Environment and Planning A* 1998 vol 30 , 1905-1927
- Giuffrè, O., Granà, A., Roberta, M., dan Corriere, F., 2011, Handling Underdispersion in Calibrating Safety Performance Function at Urban, Four-Leg, Signalized Intersections. *Journal of Transportation Safety & Security*, 3(3), pp 174-188
- Greene, W. 2008. *Functional Forms for the Negative Binomial Model for Count Data, Foundation, and Trends in Econometrics*, 99, 585-590. New York: New York University.
- Gujarati, D. N. 2004. *Basic Econometrics, 4th Edition*. New York: The McGraw-Hill Companies.
- Hardin, J. W., Hilbe, J. M., dan Hilbe, J., 2007, *Generalized Linear Models and Extensions*. Stata Press.
- Hocking, R. (1996), *Methods and Applications of Linear Models*, New York : John Wiley & Sons.
- Khoshgoftaar, T.m., Gao, K, dan Szabo, R.M. 2004. *Comparing software fault prediction of pure and zero-inflated Poisson regression models*. *Internasional Jurnal of System Science* 36,11 : 705-715
- McCullagh, P., dan J. A. Nelder, 1989, *Generalized Linear Models, 2nd Ed.*, Chapman and Hall, New York.
- Myers RH., 1990, *Classical and Modern Regression with Applications Second Edition*, PWS-KENT, New York
- Rahmawati, R., dan Djuraidah, A., 2010, Regresi Terboboti Geografis dengan Pembobot Kernel Kuadrat Ganda untuk Data Kemiskinan di Kabupaten Jember. In *Forum Statistika dan Komputasi*, Vol. 15, No. 2.
- Ricardo, A., & Carvalho, T. 2013. *Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion*. Business Media New York: Springer Science.
- Widodo. T.C, Sunaryo. S, dan Purhadi, 2013, *Pemodelan Spasial Balita Gizi buruk dengan Geographically Weighted Negative Binomial Regression dan Flexibly Shaped Spatial Scan Statistic*. Surabaya: ITS [Thesis]. Tidak dipublikasikan