

PEMODELAN JUMLAH KEMATIAN BAYI DI KOTA PADANG TAHUN 2013 DAN 2014 DENGAN PENDEKATAN REGRESI BINOMIAL NEGATIF

WINDA OKTARI, HAZMIRA YOZZA, FERRA YANUAR

*Program Studi Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia,
email : windaoktari55@gmail.com*

Abstrak. *Millennium Development Goals (MDGS)* merupakan sebuah paradigma pembangunan global. Salah satu tujuan dari MDGs yang dilaksanakan oleh Negara Indonesia adalah menurunkan Angka Kematian Bayi (AKB). Penurunan Angka Kematian Bayi (AKB) mengindikasikan bahwa adanya peningkatan derajat kesehatan masyarakat. Tujuan penelitian ini adalah mendapatkan model terbaik dari jumlah kematian bayi dan menentukan faktor yang mempengaruhi jumlah kematian bayi di Kota Padang. Pada data jumlah kematian bayi menyebar mengikuti distribusi Poisson. Karakteristik penting dari distribusi Poisson yaitu nilai *mean* harus sama dengan *varians* atau disebut equidispersi, namun berdasarkan analisis, diketahui bahwa terdapat pelanggaran asumsi yaitu *varians* lebih besar dari *mean* yang disebut overdispersi pada data. Untuk menangani pelanggaran asumsi tersebut maka dilakukan pemodelan menggunakan regresi Binomial Negatif. Hasil penelitian ini menunjukkan bahwa faktor-faktor yang mempengaruhi jumlah kematian bayi di Kota Padang pada tahun 2013 dan 2014 adalah persentase Berat Bayi Lahir Rendah (BBLR).

Kata Kunci: Kematian bayi, overdispersi, regresi Binomial Negatif

1. Pendahuluan

Millennium Development Goals (MDGs) diartikan sebagai tujuan pembangunan *Millennium* yang merupakan sebuah paradigma pembangunan global yang dideklarasikan oleh anggota PBB di New York. Sebagai salah satu negara yang ikut serta menandatangani deklarasi MDGs, Indonesia memiliki komitmen untuk melaksanakan program pembangunan nasional. Salah satu tujuan dalam MDGs yang dilaksanakan oleh Indonesia adalah menurunkan Angka Kematian Bayi (AKB).

Indonesia merupakan negara yang memiliki jumlah kematian bayi relatif tinggi bila dibandingkan dengan negara ASEAN lain. Berdasarkan *Human Development Report 2010*, AKB di Indonesia mencapai 31 per 1.000 kelahiran hidup. Angka tersebut 5.2 kali lebih tinggi dibandingkan Malaysia, 1.2 kali lebih tinggi dibandingkan Filipina dan 2.4 kali lebih tinggi dibandingkan Thailand. Begitupun di Sumatera Barat, khususnya Kota Padang juga memiliki jumlah kematian bayi yang cenderung mengalami peningkatan. Jumlah kematian bayi di Kota Padang pada tahun 2013 dan 2014 mengalami peningkatan, dari 102 orang di tahun 2013 menjadi 108 orang di tahun 2014.

Kematian bayi merupakan indikator yang biasanya digunakan untuk menentukan derajat kesehatan masyarakat. Untuk mengetahui faktor-faktor yang berpengaruh terhadap angka kematian bayi, perlu dilakukan pemodelan hubungan angka kematian bayi dengan faktor-faktor yang diduga mempengaruhinya. Dalam statistika, analisis yang digunakan untuk memodelkan pengaruh faktor-faktor terhadap suatu variabel respon tertentu adalah analisis regresi. Namun karena jumlah kematian bayi merupakan peubah diskrit, pemodelan pengaruh faktor terhadap jumlah kematian bayi harus dilakukan secara khusus, yaitu dengan menggunakan regresi Poisson.

Model Regresi Poisson adalah model standar yang digunakan untuk menganalisis data count yaitu data yang nilainya nonnegatif dan menyatakan banyak kejadian dalam interval waktu, ruang, atau volume. Analisis regresi ini dilakukan dengan mengasumsikan bahwa data berdistribusi Poisson. Karakteristik penting dari distribusi Poisson ini yaitu *mean* (rata-rata) harus sama dengan *varians* (ragam) atau disebut dengan equidisersi. Namun, pada praktiknya, sering ditemui data diskrit dengan ragam lebih besar dibanding rata-rata atau bisa disebut dengan overdispersi.

Salah satu metode yang dapat digunakan untuk menangani masalah overdispersi adalah dengan menggunakan Regresi Binomial Negatif [4]. Regresi Binomial Negatif merupakan suatu model yang sesuai untuk data count dimana terjadi penyimpangan asumsi equidisersi pada distribusi Poisson. Pada penelitian ini akan dicari pemodelan jumlah kematian bayi di Kota Padang tahun 2013 dan 2014. Model yang digunakan adalah model regresi Binomial Negatif.

2. Landasan Teori

2.1. Distribusi Poisson dan Binomial Negatif

Distribusi Poisson

Dalam [6] dituliskan bahwa distribusi Poisson merupakan distribusi peluang bagi peubah acak Poisson Y , yang menyatakan banyaknya hasil percobaan yang terjadi dalam suatu selang waktu atau daerah tertentu. Fungsi kepekatan peluang dari distribusi ini adalah:

$$f(y; \mu) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, 2, \dots \quad (2.1)$$

dimana

μ : rata-rata kejadian pada selang waktu atau daerah tertentu

y : jumlah kejadian yang diobservasi pada waktu tertentu.

Pada distribusi Poisson nilai *mean* dan *varians* mempunyai nilai yang sama yaitu $E(Y) = Var(Y) = \mu$.

Distribusi Binomial negatif

Peubah acak Binomial Negatif Y didefinisikan sebagai banyaknya kegagalan yang terjadi hingga didapatkan p keberhasilan. Fungsi kepekatan peluang peubah acak

Binomial Negatif adalah [2]:

$$f(y; \alpha, p) = \binom{y + \alpha - 1}{\alpha - 1} p^\alpha q^y, \text{ untuk } y = 0, 1, 2, \dots \quad (2.2)$$

dimana

p : peluang sukses.

y : jumlah kegagalan sampai mendapatkan sukses ke α ,

α : jumlah keberhasilan yang muncul,

$y + \alpha$: banyaknya percobaan yang dilakukan.

Overdispersi

Model regresi Poisson mensyaratkan equidispersi, yaitu kondisi dimana nilai *mean* dan varians dari variabel respon bernilai sama. Namun, adakalanya terjadi overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi adalah kondisi pada saat varians lebih besar dari nilai *mean*. Overdispersi terjadi apabila rasio antara varians dan nilai harapan lebih dari 2.5 [8]. Hal ini mengindikasikan bahwa model regresi Poisson tidak cocok untuk data yang mengandung kondisi overdispersi.

Distribusi Binomial Negatif pada Kasus Overdispersi

Distribusi Binomial Negatif merupakan distribusi yang memiliki banyak cara dalam penurunannya, salah satunya dapat diturunkan sebagai distribusi campuran Poisson-Gamma.

Misalkan bahwa peubah acak $Y_i \sim \text{Poisson}(\lambda_i)$ dengan fungsi peluang $f(y | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^y}{y!}$, $y = 0, 1, 2, \dots$ dan $\lambda_i = 0$. Selanjutnya diasumsikan $\lambda_i \sim \text{Gamma}(\alpha, \beta)$ dengan fungsi kepekatan peluang:

$$f(\lambda_i) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda_i^{\alpha-1} \exp\left(-\frac{\lambda_i}{\beta}\right), & \lambda_i > 0 \\ 0, & \lambda_i \text{ yang lain} \end{cases} \quad (2.3)$$

Sehingga dapat dinyatakan

$$f(y_i; \mu_i, \kappa) = \binom{y + \kappa^{-1} - 1}{\kappa^{-1} - 1} \left(\frac{1}{1 + \kappa \mu_i}\right)^{\kappa^{-1}} \left(\frac{\kappa \mu_i}{1 + \kappa \mu_i}\right)^{y_i} \quad (2.4)$$

Fungsi peluang pada Persamaan (2.4) merupakan fungsi peluang dari sebaran binomial negatif dengan parameter μ_i dan κ dengan nilai *mean* $E(Y_i) = \alpha\beta = \mu_i$ dan varians $Var(Y_i) = \mu_i + \kappa\mu_i^2$.

2.2. Uji Kolmogorov Smirnov

Uji Kolmogorov-Smirnov adalah suatu uji *goodness-of-fit* yang menguji tingkat kesesuaian antara distribusi serangkaian nilai sampel (skor yang diobservasi) dengan suatu distribusi teoritis tertentu. Hipotesis uji Kolmogorov Smirnov adalah:

H_0 : Data mengikuti distribusi tertentu
 H_1 : Data tidak mengikuti distribusi tertentu
 Statistik uji yang digunakan adalah:

$$D = \max | F_0(x) - S_N(x) | .$$

Kriteria untuk pengujian ini adalah tolak H_0 jika nilai D_{hitung} lebih besar dari nilai D_{tabel} . Jika $\alpha = 0,05$ dengan banyak pengamatan N maka D_{tabel} sebagai berikut [7] :

$$D_{tabel} = \frac{1.36}{\sqrt{N}}. \tag{2.5}$$

2.3. Analisis Regresi Poisson

Generalized Linier Model (GLM) adalah sebuah pengembangan dari model linier klasik. Model linier klasik mengasumsikan variabel responnya berdistribusi normal. Namun pada kenyataannya sering ditemukan suatu kondisi dimana variabel responnya tidak berdistribusi normal. Untuk kondisi dimana variabel tidak berdistribusi normal, tetapi masih saling bebas maka para statistikawan yang dipelopori McCullagh dan Nelder (1989) telah mengembangkan model linear yang dikenal dengan *Generalized Linier Model*. Bentuk umum dari GLM adalah:

$$\begin{aligned} g(\mu) &= \eta_i = \beta_0 + \beta_1 x_i1 + \beta_2 x_i2 + \dots + \beta x_i p, \\ &= x_i^T \beta, \quad i = 1, 2, \dots, n. \end{aligned}$$

Regresi Poisson termasuk ke dalam *Generalized Linier Model* dan merupakan salah satu bentuk regresi yang digunakan untuk model data cacahan. Pada model regresi Poisson, biasanya link function yang digunakan adalah fungsi logaritma natural (\ln) [1], sehingga $g(\mu_i) = \ln(\mu_i)$. Dengan demikian model regresi Poisson dapat ditulis sebagai berikut

$$\ln(\mu_i) = x_i^T \beta, \quad i = 1, 2, \dots, n. \tag{2.6}$$

dimana: $\mu_i = \mu(X_i, \beta) = \text{Var}(Y_i) = \exp(x_i^T \beta)$, Model regresi Poisson dapat dituliskan sebagai berikut [2]:

$$\mu_i = x_i^T \beta + \varepsilon_i. \tag{2.7}$$

2.4. Regresi Binomial negatif

Model regresi Binomial Negatif pada umumnya menggunakan fungsi penghubung logaritma atau log *link* yaitu [1]:

$$\ln(\mu_i) = x_i^T \beta. \tag{2.8}$$

Setelah diperoleh fungsi penghubung yang tepat, maka selanjutnya dapat dinyatakan model regresi Binomial Negatif untuk memodelkan data count yaitu:

$$\mu_i = \exp\left(x_i^T \hat{\beta}\right) \tag{2.9}$$

2.5. Estimasi Parameter Model Regresi Binomial Negatif

Parameter-parameter dalam model regresi Binomial Negatif yang tidak diketahui nilainya, yaitu $\beta_0, \beta_1, \dots, \beta_p$ perlu ditaksir. Penaksiran parameter dilakukan dengan menggunakan metode maksimum *likelihood*. Persamaan log *likelihood* dari distribusi Binomial Negatif yaitu:

$$(l\beta^*) = \sum_{i=1}^n \left\{ \sum_{r=1}^{y_i-1} \ln(1 + \kappa r) - \ln(y_i) + y_i(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) - (y_i + \kappa^{-1}) \ln(1 + \kappa \exp(\mathbf{x}_i^T \boldsymbol{\beta})) \right\}. \quad (2.10)$$

Untuk mencari taksiran dari parameter-parameter, fungsi log *likelihood* pada Persamaan log *likelihood* diturunkan secara parsial terhadap masing-masing parameter yang bersesuaian kemudian disamakan dengan nol. Untuk mencari taksiran tersebut digunakan metode numerik, yaitu metode Newton-Raphson. Metode Newton-Raphson digunakan untuk menemukan solusi dari fungsi log *likelihood* sehingga diperoleh nilai konvergen untuk dijadikan sebagai taksiran untuk masing-masing parameter. Sehingga didapat taksiran untuk model regresi Binomial Negatif yaitu

$$\mu_i = \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}). \quad (2.11)$$

2.6. Uji Kesesuaian Model Regresi Binomial Negatif

2.6.1. Uji Rasio Likelihood

Uji Rasio *Likelihood* digunakan untuk menguji signifikansi dari taksiran model yang telah diperoleh. Statistik uji yang digunakan adalah:

$$G = \ln(L_1 - L_0).$$

Aturan keputusan yang akan diambil adalah H_0 ditolak pada tingkat signifikansi α jika $G > \chi_{\alpha, v}^2$. Penolakan H_0 pada tingkat signifikansi α memberi arti bahwa terdapat paling sedikit satu parameter yang berpengaruh terhadap model.

2.6.2. Uji Wald

Uji ini berguna untuk mengevaluasi signifikansi dari masing-masing variabel prediktor terhadap model. Statistik uji yang digunakan yaitu:

$$W_j = \left[\frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \right]^2. \quad (2.12)$$

Kriteria pengujian adalah tolak H_0 jika $W_j > \chi_{\alpha, v}^2$. Penolakan H_0 pada tingkat signifikansi α berarti bahwa variabel prediktor ke- j tertentu memiliki kontribusi yang signifikansi terhadap variabel respon Y .

2.7. Pemilihan Model Terbaik

Pemilihan model terbaik dari beberapa kemungkinan model regresi Binomial Negatif menggunakan nilai AIC (*Akaike's Information Criterion*). AIC digunakan untuk melihat kecocokan model terhadap data. Nilai AIC dapat dinotasikan sebagai berikut [1]:

$$AIC = -2 \ln L(\hat{\beta}) + 2k. \quad (2.13)$$

2.8. Angka Kematian Bayi

Kematian bayi adalah kematian yang terjadi antara saat setelah bayi lahir sampai bayi belum berusia tepat satu tahun. Angka Kematian Bayi (AKB) adalah banyaknya bayi yang meninggal sebelum mencapai umur satu tahun yang dinyatakan dalam 1000 kelahiran hidup pada tahun yang sama. Secara matematis dapat ditulis.

$$AKB = \frac{D_0}{L} \times 1000. \quad (2.14)$$

3. Metode Penelitian

Variabel data yang digunakan dalam penulisan ini yaitu banyaknya kasus jumlah kematian bayi di kota Padang yang dikelompokkan per kecamatan sebagai variabel respon (Y) dan variabel prediktor (X) meliputi:

1. Berat Bayi Lahir Rendah (X1) diukur dalam persentase.
2. Bayi mendapat vitamin A (X2) diukur dalam persentase.
3. Bayi yang diberi ASI eksklusif (X3) diukur dalam persentase.
4. Ibu bersalin ditolong tenaga kesehatan (X4) diukur dalam persentase.
5. Ibu hamil melaksanakan program K1 (X5) diukur dalam persentase.
6. Ibu hamil melaksanakan program K4 (X6) diukur dalam persentase.
7. Pelayanan kesehatan bayi (X7) diukur dalam persentase.
8. Imunisasi dasar lengkap (X8) diukur dalam persentase.
9. Rumah sehat (X9) diukur dalam persentase.

Langkah-langkah analisis data yang digunakan dalam penulisan ini adalah sebagai berikut:

1. Melakukan uji Kolmogorov-Smirnov untuk menguji apakah variabel respon Y mengikuti distribusi Poisson atau tidak.
2. Menghitung nilai koefisien overdispersi untuk mengidentifikasi overdispersi.
3. Jika terjadi overdispersi, kemudian dilanjutkan dengan pendekatan model regresi Binomial Negatif.
4. Menentukan model regresi Binomial Negatif.
5. Melakukan pengujian kesesuaian model regresi Binomial Negatif.
6. Melakukan pemilihan model terbaik.
7. Menginterpretasikan model yang terbentuk.

4. Pembahasan

4.1. Pengujian Distribusi Poisson pada Jumlah Kematian Bayi

Dilakukan pengujian distribusi Poisson untuk variabel respon Y yaitu banyaknya jumlah kasus kematian bayi. Pengujian ini dilakukan dengan uji *Kolmogorov-Smirnov*, dengan hipotesis

H_0 : Jumlah kasus kematian bayi di Kota Padang berdistribusi Poisson,

H_1 : Jumlah kasus kematian bayi di Kota Padang tidak berdistribusi Poisson.

Nilai D_{tabel} untuk taraf uji $\alpha = 5\%$ dan $n = 44$ adalah $D_{(44;0.05)} = \frac{1.36}{\sqrt{44}} = 0.205$. Berdasarkan hasil pengolahan data, nilai D_{hitung} adalah sebesar 0.196. Karena nilai $D_{hitung} < D_{tabel}$ maka diputuskan untuk tidak tolak H_0 yang berarti pada taraf nyata 5% tidak dapat disangkal bahwa jumlah kasus kematian bayi di Kota Padang tahun 2013 dan 2014 berdistribusi Poisson.

4.2. Pengujian Overdispersi

Setelah diketahui bahwa jumlah kematian bayi berdistribusi poisson, selanjutnya akan dilakukan pengujian overdispersi. Pengujian overdispersi dapat dilakukan dengan memperhatikan nilai *mean* (rata-rata) dengan *varians* (ragam).

Dapat dihitung bahwa rasio dari *mean* dan *varians* adalah 2.6786. Karena rasio antara *mean* dan *varians* jumlah kematian bayi lebih dari 2.5 maka dapat disimpulkan bahwa pada data jumlah kematian bayi telah terjadi overdispersi. Untuk mengatasi hal tersebut, maka dilakukan pemodelan menggunakan regresi Binomial Negatif.

4.3. Model Regresi Binomial Negatif

4.3.1. Hasil Pengolahan Data

Regresi Binomial Negatif merupakan salah satu cara untuk mengatasi overdispersi pada regresi Poisson. Dengan menggunakan SPSS 15.0 diperoleh output sebagai berikut:

Tabel 4.1 Hasil Estimasi Model Regresi Binomial Negatif I

Variabel	Parameter	Estimasi	SE	Wald	Db	P-value
Intersep	β_0	6.058	5.8093	1.087	1	0.297
Berat Bayi Lahir Rendah (%)	β_1	0.283	0.1573	3.229	1	0.072
Bayi Mendapat Vit A (%)	β_2	0.002	0.0110	0.049	1	0.825
Bayi yang Diberi ASI Eksklusif (%)	β_3	-0.007	0.0150	0.235	1	0.628
Ibu Bersalin Ditolong Tenaga Kesehatan (%)	β_4	0.029	0.0410	0.496	1	0.481
Ibu Hamil Melaksanakan Program K1 (%)	β_5	-0.080	0.0808	0.971	1	0.324
Ibu Hamil Melaksanakan Program K4 (%)	β_6	0.018	0.0689	0.066	1	0.797
Pelayanan Kesehatan Bayi (%)	β_7	-0.033	0.0381	0.747	1	0.387
Imunisasi Dasar Lengkap (%)	β_8	0.023	0.0197	1.418	1	0.234
Rumah Sehat (%)	β_9	-0.004	0.0061	0.340	1	0.550

Berdasarkan Tabel 1 dapat dilihat bahwa *p-value* untuk uji *Wald* pada variabel berat bayi lahir rendah adalah 0.072, maka dengan menggunakan taraf nyata $\alpha =$

10% diputuskan untuk tolak H_0 karena nilai $p\text{-value} < \alpha$.

4.3.2. Uji Signifikansi Model

Agar mendapatkan model terbaik maka variabel yang signifikan diuji kembali. Berikut adalah output parameter estimation pada model regresi Binomial Negatif II:

Tabel 2. Hasil Estimasi Model Regresi Binomial Negatif II

Variabel	Parameter	Estimasi	SE	Wald	Db	$p\text{-value}$
<i>Intercept</i>	β_0	1.178	0.2562	21.142	1	0.000
Berat Bayi Lahir Rendah (%)	β_1	0.248	0.1377	3.236	1	0.072

Pada Tabel 2 di atas dapat dilihat bahwa $p\text{-value}$ untuk uji *Wald* pada variabel berat bayi lahir rendah adalah 0.072, maka dengan menggunakan taraf nyata 10% dapat dilihat nilai $p\text{-value}$ dari parameter kurang dari taraf nyata 10%. Ini berarti persentase Berat Bayi Lahir Rendah (X_1) berpengaruh terhadap jumlah kematian bayi di Kota Padang, sehingga model regresi Binomial Negatif yang diperoleh untuk memodelkan jumlah kematian bayi di Kota Padang tahun 2013 dan 2014 adalah:

$$\mu_i = \exp(1.178 + 0.248x_{1i}). \tag{4.1}$$

Model di atas dapat diartikan bahwa rata-rata jumlah kematian bayi cenderung naik menjadi ($\exp(0.248) = 1.281$) kali jika persentase Berat Bayi Lahir Rendah (BBLR) naik 1%.

5. Kesimpulan

Dari analisis dan pembahasan diperoleh model terbaik untuk jumlah kasus kematian bayi di Kota Padang sebagai berikut:

$$\mu_i = \exp(1.178 + 0.248x_{1i}). \tag{5.1}$$

dimana μ_i adalah rata-rata jumlah kematian bayi dan x_1 adalah persentase Berat Bayi Lahir Rendah.

Daftar Pustaka

- [1] Joseph, M.H. 2011. *Negative Binomial Regression Second Edition*. Cambridge University Press, New York.
- [2] Myers, R.H. 1990. *Classical and Modern Regression with Applications*. PWS-Kent Publishing Company: Boston
- [3] Pratt, J.W, H. Raiffa, R. Schlaifer. 1995. *Statistical Decision Theory*. Cambridge, London
- [4] Sary, S.A. dan I. N. Latra. 2013. Pemodelan Jumlah Kematian Bayi di Provinsi Jawa Timur Tahun 2011 dengan Pendekatan Regresi Binomial Negatif. *Jurnal Sains dan Seni Pomits*. Vol. 2 No. 2. Hal. 282 – 287

- [5] Siegel, S. 1992. *Statistik Nonparametrik untuk Ilmu-Ilmu Sosial*. Gramedia Pustaka Utama, Jakarta
- [6] Walpole, R.E. *Pengantar Statistika Edisi ke-3*. PT. Gramedia Pustaka Utama, Jakarta
- [7] Winkelmann, R. 2008. *Econometric Analysis of Count Data*. 5th edition. Springer: Berlin
- [8] Yulianingsih, K.A, K.G Sukara dan L.P Suciptawati. 2012. Penerapan Regresi Poisson untuk Mengetahui Faktor-Faktor yang Mempengaruhi Jumlah Siswa SMA/SMK yang Tidak Lulus di Bali. *E-Jurnal Matematika* 1 : 59 – 63