

Penerapan Metode Regresi Binomial Negatif terhadap Jumlah Perceraian di Provinsi Jawa Barat Tahun 2017

Airnarni Harni Murni*, Anneke Iswani Achmad

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

*airnarni72@gmail.com, annekeiswani11@gmail.com

Abstract. Regression analysis is the most widely used analysis in a study. The primary purpose of regency analysis is to identify the predictor variabels that affect the response variabels. One of the statistic methods to use is the regression of negative binomial before using the negative regression method used first, to see if the data experiences the equidispersion (a state where the variance and rate are as high) or overdispersion (a state where variances is greater than rate). By the time the data is overdispersion, it's used a negative binomial using the data for divorces in west Java in 2017. Obtained by regression model negative $\hat{y}_i = \exp(3,59007 + 0,22107 (X_2) + 0,06026 (X_3))$, from that binary regression model that it coefficient regression increases in means that more people are poor, more unemployment is open, and school participation rates are going to raise the number of divorces. But for the β_1 not significantly affected.

Keywords: Regression Poisson, Overdispersion, Regression Negative Binomial, and Divorce.

Abstrak. Analisis regresi merupakan analisis yang paling banyak digunakan dalam suatu penelitian. Tujuan utama dari analisis regresi adalah untuk mengidentifikasi variabel-variabel prediktor yang berpengaruh terhadap variabel respon. Salah satu metode statistika yang akan digunakan adalah Regresi Binomial Negatif. Sebelum menggunakan metode Regresi Binomial Negatif digunakan Regresi Poisson terlebih dahulu, untuk melihat apakah data tersebut mengalami equidispersion (keadaan dimana nilai variansi dan rataannya sama besar) atau overdispersion (keadaan dimana nilai variansinya lebih besar dari rataannya). Pada saat data mengalami overdispersi maka digunakan metode Regresi Binomial Negatif. Pada penelitian ini dilakukan pemodelan regresi binomial negatif menggunakan data jumlah perceraian di Jawa Barat Tahun 2017. Diperoleh model regresi binomial negative $\hat{y}_i = \exp(3,59007 + 0,22107 (X_2) + 0,06026 (X_3))$, dari model regresi binomial negatif tersebut koefisien regresi bertambah artinya semakin banyak penduduk miskin, tingkat pengangguran terbuka, dan angka partisipasi sekolah akan semakin tinggi jumlah perceraian. Namun untuk β_1 tidak mempengaruhi secara signifikan.

Kata Kunci: Regresi Poisson, Overdispersion, Regresi Binnomial Negatif, dan Perceraian.

1. Pendahuluan

Analisis regresi merupakan suatu metode yang digunakan untuk mengetahui pengaruh dari suatu variabel terhadap variabel lainnya. Pada analisis regresi suatu variabel yang mempengaruhi disebut variabel bebas (X), sedangkan variabel yang dipengaruhi disebut variabel tak bebas (Y). Regresi Poisson dan Binomial Negatif merupakan salah satu analisis regresi nonlinear yang dimodelkan dengan GLM. Analisis Regresi Poisson dan Binomial Negatif digunakan untuk menganalisis hubungan antara sebuah variabel tidak bebas (Y) yang berupa data diskrit dengan satu atau lebih variabel bebas (X). Dalam model Regresi Poisson digunakan dalam keadaan nilai variansi dan rataannya sama (equidisperse), sedangkan Regresi Binomial Negatif dapat baik digunakan dalam keadaan equidispersi dan overdispersi (nilai variansinya lebih besar dari rataannya).

Untuk data dengan keadaan overdispersi pada regresi binomial negatif bermakna bahwa metode analisis ini dapat mengakomodasi overdispersi. Karena, regresi binomial negatif dapat memodelkan data yang mengalami overdispersi tanpa harus menghilangkan data pencilan dari data tersebut. Model yang digunakan dalam penelitian ini untuk mengetahui pengaruh perceraian yang terjadi di Provinsi Jawa Barat tahun 2017. Tingginya tingkat Perceraian di Provinsi Jawa Barat pada tahun 2017 mencapai 22.013 kasus. Dilihat dari jumlah perceraian yang sangat tinggi membuat pihak pengadilan untuk mempersulit perceraian, agar tidak terjadi peningkatan di setiap tahunnya. Faktor-faktor yang mempengaruhi tingkat Perceraian antara lain Penduduk Miskin, Tingkat Pengangguran, dan Angka Partisipasi Sekolah (APS).

Dalam penelitian ini akan dijelaskan pengaruh antara jumlah perceraian variabel Y (tak bebas) dengan faktor-faktor yang berpengaruh variabel X (bebas) yaitu Penduduk Miskin, Tingkat Pengangguran, dan Angka Partisipasi Sekolah (APS). Sehingga hubungan antara jumlah perceraian dengan faktor-faktor yang berpengaruh dapat dilihat melalui regresi Binomial Negatif.

Tujuan penelitian ini ingin mengetahui data dari kasus Perceraian di Provinsi Jawa Barat tahun 2017 mengandung *equidispersion* atau *overdispersion*, memperoleh model dari metode analisis Regresi Binomial Negatif untuk kasus Perceraian di Provinsi Jawa Barat tahun 2017, dan mengetahui jumlah perceraian apabila variabel bebasnya diketahui.

2. Landasan Teori

Uji Multikolinearitas

Salah satu syarat yang harus terpenuhi dalam pembuatan model regresi adalah tidak adanya multikolinearitas antar variabel bebas. Untuk mendeteksi adanya multikolinearitas pada model regresi dapat menggunakan VIF (*Variance Inflation Factor*). Jika nilai $VIF > 10$ maka mengindikasikan bahwa terjadi multikolinearitas antar variabel bebas. Nilai VIF dapat dinyatakan sebagai berikut: [6]

$$VIF = \frac{1}{1-R_j^2} \quad (2.1)$$

Dengan $j = 1, 2, \dots, k$ dan k adalah banyaknya variabel bebas, sedangkan R_j^2 adalah koefisien determinasi yang dihasilkan dari regresi variabel bebas X_j dengan variabel bebas lainnya. H_0 pengujian multikolinearitas adalah tidak terdapat multikolinearitas, dengan kriteria keputusan jika nilai $VIF < 10$ maka H_0 diterima artinya tidak terdapat multikolinearitas.

Regresi Poisson

Regresi Poisson merupakan analisis regresi nonlinear dari distribusi Poisson, dimana analisis ini sangat cocok digunakan dalam menganalisis data diskrit. Model regresi Poisson merupakan *Generalized Linear Model* (GLM) yang variabel tak bebasnya diasumsikan berdistribusi Poisson. Model regresi Poisson dapat ditulis sebagai berikut. [4]

$$Y_i \sim \text{Poisson}(\mu_i), \mu_i = \exp(x_i^T \beta)$$

Maka,

$$y_i = \mu_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}) \quad (2.2)$$

Overdispersi

Pada model regresi Poisson terdapat asumsi yang harus dipenuhi. Salah satunya adalah asumsi kesamaan antara rata-rata dan variansi dari variabel tak bebas, yang disebut juga equidispersi. Pengujian hipotesis tentang kasus overdispersi dilakukan sebagai berikut. [5]

Hipotesis:

$$H_0 : \theta = 0 \text{ (tidak terjadi overdispersi)}$$

$$H_1 : \theta \neq 0 \text{ (terjadi overdispersi)}$$

Statistik uji:

$$\Phi = \frac{X^2}{df} \quad (2.3)$$

$$X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \sim \chi^2_{(n-p)}$$

Dengan kriteria pengujiannya yaitu tolak H_0 apabila nilai $\chi^2 > \chi^2_{(n-p)}$ atau $p\text{-value} < \alpha$. Taksiran dispersi diukur dengan nilai deviansi atau *Pearson's Chi-Square* yang dibagi derajat bebas, jika hasil pembagian menghasilkan nilai yang lebih besar 1 maka dapat disimpulkan data mengalami overdispersi.

Regresi Binomial Negatif

Regresi Binomial Negatif (RBN) berdistribusi binomial negatif untuk memodelkan data diskrit yang mengalami kondisi ekuidispersi maupun yang mengalami overdispersi. Untuk membentuk suatu model regresi diperoleh rata-rata dan varians dalam bentuk $E(y) = \mu$ dan $Var(y) = \mu(1 + \theta\mu)$ dengan fungsi massa peluang binomial negatif sebagai berikut. [2]

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

Model regresi Binomial Negatif dapat digunakan untuk memodelkan data Poisson yang mengalami overdispersi karena distribusi Binomial Negatif merupakan perluasan dari distribusi Poisson-Gamma yang membuat parameter dispersi θ . Model Binomial Negatif merupakan alternatif yang sering digunakan untuk kasus overdispersi pada regresi. Model regresi Binomial Negatif dinyatakan sebagai berikut. (Hilbe, J., 2011)

$$\hat{y}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki}) \quad (2.4)$$

Pengujian Parameter Regresi Binomial Negatif

Dari hasil pembentukan model regresi Binomial Negatif, estimasi parameter belum tentu berpengaruh secara signifikan terhadap model, sehingga perlu dilakukan pengujian parameter secara parsial untuk melihat signifikansi parameter terhadap model tersebut menggunakan uji Wald. Pengujiannya sebagai berikut: [3]

Hipotesis:

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0; \quad k = 0, 1, 2, \dots, p$$

Statistik uji:

$$W_t = \left(\frac{\hat{\beta}_k}{se(\hat{\beta}_k)}\right)^2 \quad (2.5)$$

Daerah Penolakan:

Tolak H_0 jika $W_t > \chi^2_{(\alpha,1)}$ artinya bahwa jika tolak H_0 parameter ke $-j$ signifikan terhadap model yang dibentuk. $se(\hat{\beta}_k)$ diperoleh dari akar kuadrat variansnya yang didapatkan dari nilai diagonal minus invers dari matriks *Hessian*.

Pemilihan Model Terbaik

Untuk melihat model terbaik pada regresi Binomial Negatif bisa dilakukan dengan melihat nilai AIC (*Akaike Information Criteria*) yang terkecil. Indikator pemilihan model terbaik dari beberapa model yang umum digunakan pada model regresi adalah nilai *Akaike Information*

Criterion (AIC). Metode ini didasarkan pada metode *Maximum Likelihood Estimate* (MLE). Perhitungan AIC dapat dilakukan dengan rumus sebagai berikut: [1]

$$AIC = -2 \ln L(\hat{\beta}) + 2k \quad (2.6)$$

Dimana $L(\hat{\beta})$ adalah nilai likelihood dan k adalah jumlah parameter variabel bebas. Model regresi terbaik adalah model regresi yang memiliki nilai AIC terkecil.

3. Hasil Penelitian dan Pembahasan

Data yang digunakan adalah data sekunder yang bersumber dari hasil pencatatan yang diperoleh dari Badan Pusat Statistika (BPS) Jawa Barat. Pada penelitian ini yang akan dijadikan unit penelitian adalah kabupaten/kota di Jawa Barat. Untuk data yang digunakan penulis disajikan pada Tabel 1:

Tabel 1. Data Kasus Perceraian beserta Variabel lainnya tahun 2017

Kota/kab	Perceraian	Penduduk Miskin	Tingkat Pengangguran	Angka Partisipasi Sekolah
Kab. Bandung	3102	4.32	8.44	43.23
Kab. Bekasi	772	11.05	10.97	14.1
Kab. Bogor	940	8.83	9.55	14.64
Kab. Ciamis	373	7.61	5.17	20.38
Kab. Cianjur	215	11.62	10.1	5.85
...
Kota Tasikmalaya	592	15.6	6.89	21.74

Pemeriksaan Multikolinearitas

Sebelum melakukan analisis regresi poisson maka akan dilakukan pengujian multikolinearitas terhadap data yang digunakan untuk mengetahui apakah diantara variabel bebas terjadi multikolinearitas ataukah tidak terjadi multikolinearitas. Salah satu cara yang dapat digunakan untuk mendeteksi multikolinearitas adalah dengan nilai VIF (*Variance Inflation Factor*). Tabel 2 adalah nilai VIF dari variabel bebas yang diperoleh sebagai berikut:

Tabel 2. Nilai VIF dari Variabel Bebas

Variabel	VIF
X_1	1,225468
X_2	1,081432
X_3	1,146613

Multikolinearitas terjadi apabila nilai VIF dari variabel bebas lebih dari 10. Pada Tabel 2 menunjukkan bahwa, nilai VIF dari masing-masing variabel bebas kurang dari 10. Maka dapat disimpulkan bahwa dari masing-masing variabel bebas tidak terjadi kasus multikolinearitas. Sehingga dapat dilanjutkan ke pemodelan regresi poisson.

Model Regresi Poisson

Hasil pendugaan parameter untuk model regresi poisson dapat dilihat pada Tabel 3. Hasil ini diperoleh menggunakan software *R Studio*.

Tabel 3. Nilai dugaan parameter model Regresi Poisson

Parameter	Nilai Dugaan	Std.Error	Nilai Z	Pr(> Z)
β_0	3,8890583	0,0506447	76,791	<2e-16***
β_1	-0,0066552	0,0013891	-4,791	1,66e-06***
β_2	0,1895954	0,0057571	32,933	<2e-16***
β_3	0,0622338	0,0006202	100,351	<2e-16***

Deviance: 8689,2; derajat bebas: 20; Rasio dispersi: 434,46. Tanda “***” menunjukkan peubah penjelas yang signifikan pada taraf nyata 5%.

Model regresi Poisson untuk semua penjelas dapat ditulis sebagai berikut:

$$y_i = \exp(3,8890583 - 0,0066552 X_1 + 0,1895954 X_2 + 0,0622338 X_3)$$

Berdasarkan Tabel 3 terlihat bahwa dari lima peubah penjelas semuanya berpengaruh terhadap variabel terikat dengan taraf nyata 5%, yaitu variabel (X_1) Penduduk Miskin, (X_2) Tingkat Pengangguran Terbuka, dan (X_3) Angka Partisipasi Sekolah.

Pengujian Overdispersi

Pada Tabel 3 menunjukkan bahwa nilai rasio dispersi antara *Deviance* dengan derajat bebas sebesar 434,46. Nilai ini berarti bahwa model regresi poisson mengalami overdispersi. Sehingga hasil yang diperoleh merupakan hasil semu. Kondisi ini menyebabkan faktor-faktor yang mempengaruhi jumlah perceraian tidak dapat dipastikan berdasarkan model regresi poisson. Pendekatan model yang dilakukan adalah menggunakan regresi binomial negatif sebagai model alternatif untuk menghindari masalah overdispersi pada regresi poisson.

Model Regresi Binomial Negatif

Hasil pendugaan parameter untuk model regresi Binomial Negatif dapat dilihat pada Tabel 4. Hasil ini diperoleh menggunakan *software R Studio*.

Tabel 4. Nilai dugaan parameter model Regresi Binomial Negatif

Parameter	Nilai Dugaan	Std.Error	Nilai Z	Pr(> Z)
β_0	3,612796	0,986103	3,664	0,000249***
β_1	-0,002049	0,031951	-0,064	0,948875
β_2	0,222341	0,104538	2,127	0,033429*
β_3	0,059881	0,017778	3,368	0,000757***

Deviance: 26,150; derajat bebas: 20; rasio dispersi: 1,307. Tanda “***” menunjukkan peubah penjelas yang signifikan pada taraf nyata 5%.

Model regresi binomial negatif untuk semua penjelas dapat ditulis sebagai berikut:

$$\hat{y}_i = \exp(3,612796 - 0,002049 X_1 + 0,222341 X_2 + 0,059881 X_3)$$

Regresi binomial negatif untuk mengatasi adanya overdispersi pada regresi poisson.

Pemilihan Model Terbaik

Pemodelan dilakukan dengan meregresikan kombinasi dari kelima variabel yang digunakan untuk kombinasi 1 variabel, 2 variabel, dan 3 variabel.. Regresi binomial negatif dapat mengatasi overdispersi karena memuat parameter dispersi dalam distribusi binomial negatif. Nilai parameter ini ditentukan oleh peneliti, sehingga nilai *deviance* yang dihasilkan dibagi dengan derajat bebasnya sama dengan 1. Berikut adalah ringkasan dari 7 pemodelan yang dilakukan, dan yang telah terpilih berdasarkan AIC terkecil.

Tabel 5. Model Regresi Binomial Negatif

Variabel	Parameter Signifikan	AIC	Parameter Dispersi
X_3	β_0, β_3	370,78	1,5887
X_2, X_3	$\beta_0, \beta_2, \beta_3$	368,77	1,8317
X_1, X_2, X_3	$\beta_0, \beta_2, \beta_3$	370,77	1,8319

Tabel 5 menunjukkan model terbaik dari setiap kombinasi variabel. Nilai parameter dispersi yang lebih besar dari 0 membuktikan bahwa terjadi overdispersi yang telah diatasi dengan parameter tersebut. Tabel 5 menunjukkan bahwa nilai AIC yang paling kecil adalah 368,77, sehingga model yang dipilih untuk digunakan adalah model ke dua.

Tabel 7. Model terbaik Regresi Binomial Negatif

Parameter	Nilai Dugaan	Std.Error	Nilai Z	Pr(> Z)
β_0	3,59007	0,93827	3,826	0,000130***
β_2	0,22107	0,10112	2,186	0,028794*
β_3	0,06026	0,01670	3,608	0,000309***

Model regresi binomial negatif untuk semua penjelas dapat ditulis sebagai berikut:

$$\hat{y}_i = \exp(3,59007 + 0,22107X_2 + 0,06026X_3)$$

4. Kesimpulan

Berdasarkan hasil dan analisis yang sudah dilakukan pada bab sebelumnya, penulis dapat menarik kesimpulan diantaranya. Bahwa pada analisis data untuk model regresi Poisson, ternyata data ini mengalami overdispersi.

Berdasarkan nilai AIC dari setiap model dengan nilai AIC terkecil didapat model terbaik, dengan membandingkan nilai AIC dari setiap model. Sehingga didapatkan nilai AIC terkecil sebesar 368,77. Dengan model regresi binomial negatif diperoleh sebagai berikut:

$$\hat{y}_i = \exp(3,59007 + 0,22107X_2 + 0,06026X_3)$$

Dari model regresi binomial negatif tersebut koefisien regresi bertambah artinya semakin banyak penduduk miskin, tingkat pengangguran terbuka, dan angka partisipasi sekolah akan semakin tinggi jumlah perceraian. Namun untuk β_1 tidak mempengaruhi secara signifikan.

Daftar Pustaka

- [1] Bozdogan, H. (2000). *Akaike' Information Criterion and Recent Development in Information Complexity*, Mathematical Psychology, 44. 62-91.
- [2] Hilbe, J. (2011). *Negative Binomial Regression, Second Edition*. New York: Cambridge University Press.
- [3] Hocking, R. (1996). *Methods and Application of Linear Models*. New York: John Wiley & Sons, Inc.
- [4] Irawati, & Purhadi. (2012). *Perbandingan Analisis Generalized Poisson Regression (GPR) dan Regresi Binomial Negatif untuk mengatasi overdispersi*.
- [5] Joseph, M. (2011). *Negative Binomial Regression Second Edition*. New York: Cambridge University Press.
- [6] Montgomery, D., & Peck, E. (1992). *Introduction to Linear Regression Analisis Second Edition*. New York: John Wiley & Sons, Inc.