

**PENERAPAN REGRESI *QUASI-LIKELIHOOD* PADA DATA
CACAH (*COUNT DATA*) YANG MENGALAMI
OVERDISPERSI DALAM REGRESI POISSON
(Studi Kasus: Jumlah Kasus Perceraian di Tiap Desa/Kelurahan
Kota Denpasar Tahun 2011)**

**DESAK PUTU PRAMI MEITRIANI¹, I KOMANG GDE SUKARSA²,
I PUTU EKA NILA KENCANA³**

^{1,2,3}Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali
e-mail: ¹desak_mei@yahoo.co.id, ²sukarsakomang@yahoo.com, ³i.putu.enk@gmail.com

Abstract

Poisson regression can be used to analyze count data, with assuming equidispersion. However, in the case of overdispersion often occur in the count data. The implementation of Poisson Regression can not be applied on this data because the data having overdispersion, that will lead to underestimate the standard error. Thus, use Quasi-Likelihood regression on this data. Quasi-Likelihood regression was also could not handle the overdispersion, but Quasi-Likelihood regression can improve the value of the standard error becomes greater than the value of the standard error on Poisson regression. Thus, by using the Quasi-Likelihood regression obtained three independent variables that affect the number of divorce cases in each urban city of Denpasar in 2011.

Keywords: *Poisson Regression, Overdispersion, Quasi-Likelihood Regression, standard error.*

1. Pendahuluan

Regresi Poisson merupakan model standar untuk *count data* dan termasuk dalam model regresi nonlinier (Cameron&Trivedi [2]). Regresi Poisson mengasumsikan keadaan yang equidispersi, namun sering terjadi kasus overdispersi yaitu nilai variansi lebih besar dari nilai *mean* (Ver Hoef, J. M and Boveng, P. L [5]). Penggunaan yang tidak tepat dari regresi Poisson pada data yang mengalami *overdispersi* dapat berakibat fatal dalam interpretasi model, khususnya parameter model karena diperoleh *standard error* yang terlalu rendah (*underestimate*) dan dapat memberikan kesimpulan yang keliru tentang signifikan atau tidaknya parameter model regresi (Darnah [3]).

Secara statistik, overdispersi dapat dideteksi dari rasio dispersinya, yang diukur dari nilai *Deviance* pada data, dengan hipotesis uji:

$$\begin{aligned} H_0 &: \alpha = 1 \\ H_1 &: \alpha > 1 \end{aligned}$$

¹ Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

^{2,3} Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

Kriteria uji: tolak H_0 jika nilai $G > \chi_{(n-k-1),\alpha}^2$ yang mengindikasikan terjadi overdispersi. Pada penulisan ini, α yang digunakan adalah 5%.

Salah satu penerapan regresi yang mampu mengatasi akibat dari overdispersi adalah regresi *Quasi-Likelihood* yaitu dengan memperbaiki nilai *standard error* pada regresi Poisson. Nilai *standard error* pada regresi *Quasi-Likelihood* disesuaikan dengan nilai parameter dispersinya, sehingga pada regresi *Quasi-Likelihood* akan selalu menghasilkan *standard error* yang lebih besar dari regresi Poisson dan regresi *Quasi-Likelihood* lebih tepat digunakan untuk mengatasi akibat dari kasus overdispersi dibanding dengan regresi Poisson (Agresti [1]).

Sehubungan dengan hal di atas, secara teori diberikan Tabel 1 mengenai perbedaan antara regresi Poisson dengan regresi *Quasi-Likelihood*:

Tabel 1. Perbedaan Regresi Poisson dan Regresi *Quasi-Likelihood*

Kriteria	Regresi Poisson	Regresi <i>Quasi-Likelihood</i>
Variabel respons	$Y_i \stackrel{iid}{\sim} Poi(\mu_i)$	$Y_i \stackrel{iid}{\sim} Poi(\mu_i, \phi)$
Parameter dispersi	Tidak memiliki	ϕ
Fungsi distribusi peluang	$p(y; \mu) = \frac{e^{-\mu} \mu^y}{y!}$ (Myers [4])	Tidak memperhatikan distribusi dari variabel respon (Ver Hoef, J. M and Boveng, P. L [5])
Mean dan variansi	$E(Y_i) = \mu_i$ $Var(Y_i) = \mu_i$	$E(Y_i) = \mu_i$ $Var(Y_i) = \phi \mu_i$
Fungsi penghubung	$\log(\mu_i) = \eta_i$	$\log(\mu_i) = \eta_i$
Penaksir parameter	<i>Maksimum Likelihood Estimation (MLE)</i>	<i>Quasi-Likelihood Estimation (QLE)</i>
Uji signifikansi parameter Kriteria Uji: $H_0: \beta_j = 0$ $H_1: \beta_j \neq 0$	$W_j = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2$ Kriteria uji: tolak H_0 apabila $W_j > \chi_{(\alpha;v)}^2$.	$W_j = \left(\frac{\hat{\beta}_j}{\sqrt{\phi} SE(\hat{\beta}_j)} \right)^2$ $= \left(\frac{\hat{\beta}_j}{\sqrt{\hat{\phi}} (SE(\hat{\beta}_j))} \right)^2$ Kriteria uji : tolak H_0 apabila $W_j > \chi_{(\alpha;v)}^2$
Uji kesesuaian model	Statistik uji <i>Deviance</i> Kriteria : tolak H_0 , jika $G > \chi_{(n-k-1),\alpha}^2$	Statistik uji <i>Deviance</i> Kriteria : tolak H_0 , jika $G > \chi_{(n-k-1),\alpha}^2$
Model regresi	$\mu_i = x_i' \beta + \varepsilon_i$	$\mu_i = x_i' \beta + \varepsilon_i$

Adapun tujuan dari penelitian ini adalah: (1) memperoleh model regresi Poisson dalam kasus perceraian di Kota Denpasar tahun 2011, (2) memperoleh model regresi *Quasi-Likelihood* dalam kasus perceraian di Kota Denpasar tahun 2011, (3) mengetahui peranan dari *standard error* pada regresi *Quasi-Likelihood* untuk data yang mengalami overdispersi, (4) melihat faktor-faktor yang

memengaruhi perceraian di Kota Denpasar tahun 2011 pada regresi *Quasi-Likelihood*.

2. Metode Penelitian

Sumber data pada penelitian ini adalah data sekunder yang diperoleh di Badan Pusat Statistik Provinsi Bali dan Kepolisian Kota Besar (Poltabes) kota Denpasar berupa data kuantitatif dengan skala rasio. Data-data tersebut adalah data variabel respon (Y) dan variabel bebas (X). Dengan variabel respon adalah jumlah kasus perceraian (Y) dan variabel bebas adalah jumlah kasus kekerasan dalam rumah tangga (X_1), jumlah kasus perzinahan (X_2), jumlah kasus perjudian (X_3), jumlah penduduk yang bekerja sebagai buruh tani (X_4), jumlah penduduk yang bekerja sebagai pegawai negeri (X_5), jumlah penduduk dengan tingkat pendidikan tamatan SMU (X_6), jumlah penduduk dengan tingkat pendidikan tamatan sarjana/Diploma IV (X_7).

Langkah-langkah analisis data yang digunakan adalah: (1) mempersiapkan data pada *Microsoft Excel 2007* yang akan diolah dengan menggunakan *software SAS 9.1.3*; (2) menganalisis data dengan menggunakan regresi Poisson; (3) memeriksa adanya overdispersi pada regresi Poisson; (4) menganalisis data dengan menggunakan regresi *Quasi-Likelihood*; (5) mengidentifikasi *standard error* pada regresi Poisson dan regresi *Quasi-Likelihood* untuk menentukan signifikansi parameter regresi.

3. Hasil dan Pembahasan

Secara teori, overdispersi dapat dideteksi dengan melakukan uji rasio dispersinya. Dengan hipotesis sebagai berikut:

$$H_0 : \alpha = 1$$

$$H_1 : \alpha > 1$$

Diperoleh nilai *Deviance* sebesar 81,6891, dengan $\alpha = 5\%$ dan diperoleh nilai $\chi^2_{(0.05;35)} = 49,7655$. Dengan demikian nilai *Deviance* (81,6891) > nilai tabel $\chi^2_{(0.05;35)} = 49,7655$, sehingga keputusan tolak H_0 yang mengindikasikan bahwa pada model regresi Poisson terjadi overdispersi.

Pada hasil analisis data dengan menggunakan regresi Poisson menunjukkan bahwa model regresi Poisson tidak cocok digunakan pada data, terlihat dari nilai *Deviance* (81,6891) > $\chi^2_{(0.05;35)} = 49,655$.

Oleh karena terjadi kasus overdispersi pada regresi Poisson, maka selanjutnya dilakukan analisis data dengan regresi *Quasi-Likelihood*. Model pada regresi *Quasi-Likelihood* menghasilkan model yang sama dengan model regresi Poisson, yaitu :

$$\hat{\mu}_i = \exp(3,7783 + 0,2994X_1 - 0,0802X_2 + 0,0259X_3 - 0,0000X_4 + 0,0001X_5 - 0,0001X_6 - 0,0004X_7).$$

Walaupun regresi *Quasi-Likelihood* tidak mampu mengatasi overdispersi, tetapi regresi *Quasi-Likelihood* akan berperan dalam memperbaiki nilai *standard error*.

Berikut disajikan Tabel 2 tentang perbedaan nilai *standard error* dan signifikansi parameter dengan menggunakan regresi Poisson dan regresi *Quasi-Likelihood*:

Tabel 2. Perbedaan Nilai *Standard Error* dan Signifikansi Parameter

Parameter	Regresi Poisson		Regresi <i>Quasi-Likelihood</i>	
	<i>Standard Error</i>	Pr>ChiSq	<i>Standard Error</i>	Pr>ChiSq
Intercept	0,1138	<,0001	0,1743	<,0001
X ₁	0,0163	<,0001	0,025	<,0001
X ₂	0,0962	0,4045	0,1473	0,5861
X ₃	0,0105	0,0137	0,0161	0,1075
X ₄	0,0004	0,9429	0,0006	0,9627
X ₅	0	0,0275	0,0001	0,1501
X ₆	0	<,0001	0	<,0001
X ₇	0,0001	0,0007	0,0002	0,0265

Dari Tabel 2 terlihat nilai *standard error* regresi *Quasi-Likelihood* lebih besar atau sama dengan nilai *standard error* regresi Poisson. Karena pada regresi Poisson terdapat 5 variabel bebas yang bernilai lebih kecil dari α , maka terdapat 5 variabel bebas yang berpengaruh terhadap variabel respon. Nilai *standard error* regresi *Quasi-Likelihood* yang lebih besar dari regresi Poisson, mengakibatkan jumlah variabel bebas yang berpengaruh berubah menjadi tiga variabel bebas, yaitu X₁, X₆, dan X₇.

4. Kesimpulan

Berdasarkan hasil analisis diperoleh kesimpulan bahwa regresi Poisson dan regresi *Quasi-Likelihood* tidak mampu mengatasi kasus overdispersi karena nilai *Pearson Chi-Square* dan *Deviance* yang masing-masing dibagi db masih menghasilkan nilai yang lebih besar dari satu bahkan masih menghasilkan nilai *Pearson Chi-Square* dan *Deviance* yang sama pada regresi Poisson.

Diperoleh model regresi Poisson sama dengan model regresi *Quasi-Likelihood*, yaitu:

$$\hat{\mu}_i = \exp(3.7783 + 0.2994X_1 - 0.0802X_2 + 0.0259X_3 - 0.0000X_4 + 0.0001X_5 - 0.0001X_6 - 0.0004X_7)$$

Akan tetapi regresi *Quasi-Likelihood* dapat memperbaiki nilai *standard error*, dengan menghasilkan nilai *standard error* yang lebih besar dari regresi Poisson sehingga hanya terdapat 3 variabel bebas yang berpengaruh terhadap jumlah kasus perceraian di Kota Denpasar tahun 2011 adalah jumlah kasus kekerasan dalam rumah tangga, jumlah penduduk dengan tingkat pendidikan tamatan SMU, dan jumlah penduduk dengan tingkat pendidikan tamatan sarjana/Diploma IV.

Daftar Pustaka

- [1] Agresti, A., 2002. *Categorical Data Analysis*. 2nd ed. New York: A John Wiley & Sons, Inc.
- [2] Cameron, A.C. & Trivedi, P.K., 1999. *Essentials of Count Data Regression*. [Online]<http://www.econ.ucdavis.edu/faculty/cameron/research/cte01preprint.pdf>. Diakses : 29 Mei 2013.
- [3] Darnah, 2011. Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson Regression I. *Ekspensial*, 2.
- [4] Myers, R.H., 1990. *Classical and Modern Regression with Applications*. 2nd ed. Boston: PWS-KENT.
- [5] Ver Hoef, J.M. & Boveng, P.L., 2007. Quasi-Poisson Vs. Negative Binomial Regression : How Should We Model Overdispersed Count Data ? *Agencies and Staff of the U.S. Departement of Commerce*, pp.2766-72.