Modifikasi *Regression Correlation Coefficient* (RCC) untuk Model Regresi Poisson

Modification Regression Correlation Coefficient (RCC) for Poisson Regression Model

[1]Puzi Yatni Dwiastuti, [2]Lisnur Wachidah
[1,2]Prodi Statistika, Fakultas MIPA, Universitas Islam Bandung,
Jl. Tamansari No.1 Bandung 40116
email: 1puziyda@gmail.com, 2lisnur_w@yahoo.co.id

Abstract. The Poisson regression model originating from data must be evaluated by using a certain prediction strength so that the best candidate model is selected. Some measures of predictive power that have been proposed, most have limitations in certain situations such as interpretation, application, consistency and affinity. To overcome these limitations, Zheng and Agresti (2000) proposed Regression Correlation Coefficient (RCC) for the Generalized Linear Model (GLM). Then, Takahashi and Kurosawa (2016) propose RCC for Poisson regression models developed from RCC for GLM. However, the RCC for the Poisson regression model is considered to still have deficiencies, namely if there is multicollinearity in the predictor variable. Therefore, to overcome these shortcomings, Kaengthong and Domthong (2017) propose modification the RCC for Poisson regression models that can work well in the case of two or more predictor variables and have multicollinearity in predictor variables. This paper discusses the application of the modified RCC for the Poisson regression model. The data used in the application is the number of maternal deaths in West Kalimantan Province in 2017. The results obtained the best model for Poisson regression models with 1 variable, namely the percentage variable of pregnant women implementing the K4 (X_2) program with a value of $\hat{R}^*$ of 0.5826.

Keywords: Poisson Regression Model. Measure Predictive Power, Modification RCC.

Abstrak. Model regresi Poisson yang berasal dari data harus dievaluasi dengan menggunakan ukuran kekuatan prediksi tertentu agar kandidat model terbaik terpilih. Beberapa ukuran kekuatan prediksi yang telah diusulkan, sebagian besar memiliki keterbatasan dalam situasi tertentu seperti interpretasi, penerapan, konsistensi dan afinitas. Untuk mengatasi keterbatasan tersebut, Zheng dan Agresti (2000) mengusulkan Koefisien Korelasi Regresi atau *Regression Correlation Coefficient* (RCC) untuk *Generalized Linear Model* (GLM). Kemudian, Takahashi dan Kurosawa (2016) mengusulkan RCC untuk model regresi Poisson yang dikembangkan dari RCC untuk GLM. Namun, RCC untuk model regresi Poisson tersebut dianggap masih memiliki kekurangan yaitu jika terdapat multikoliniearitas dalam variabel prediktor. Oleh karena itu, untuk mengatasi kekurangan tersebut, Kaengthong dan Domthong (2017) mengusulkan modifikasi RCC untuk model regresi Poisson yang dapat bekerja dengan baik dalam kasus dua atau lebih variabel prediktor dan memiliki multikolinieritas dalam variabel prediktor. Makalah ini membahas penerapan modifikasi RCC untuk model regresi Poisson. Data yang digunakan dalam pengaplikasiannya adalah jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017. Hasilnya diperoleh model terbaik untuk model regresi Poisson dengan 1 variabel yaitu variabel persentase ibu hamil melaksanakan program K4 (X_2) dengan nilai $\hat{R}^*$ sebesar 0,5826.

Kata Kunci: Model Regresi Poisson, Ukuran Kekuatan Prediksi, Modifikasi RCC.

A. Pendahuluan

Analisis regresi merupakan suatu analisis statistika yang bertujuan untuk memodelkan hubungan antara variabel respon (Y) dengan variabel prediktor (X). Salah satu analisis statistika yang menggunakan analisis regresi, diantaranya regresi Poisson. Model regresi Poisson merupakan *Generalized Linear Model* (GLM) dengan data responnya (Y) diasumsikan berdistribusi Poisson.

Model regresi Poisson yang berasal dari data harus dievaluasi dengan menggunakan ukuran kekuatan prediksi tertentu agar kandidat model terbaik terpilih. Beberapa ukuran kekuatan prediksi yang telah diusulkan, sebagian besar memiliki keterbatasan dalam situasi tertentu.

Untuk mengatasi keterbatasan tersebut, dikembangkan suatu ukuran kekuatan prediksi yang diusulkan oleh Zheng dan Agresti (2000) yang disebut *Regression Correlation Coefficient* (RCC) untuk GLM. RCC adalah nilai populasi yang didefinisikan oleh korelasi antara variabel respon (Y) dan harapan kondisional dari variabel respon $\left(E(Y|X)\right)$.

Takahashi dan Kurosawa (2016) mengusulkan RCC untuk model regresi Poisson yang dikembangkan dari RCC untuk GLM. RCC untuk model regresi Poisson dapat menemukan bentuk eksplisit dari RCC, dengan mengasumsikan bahwa variabel prediktor berdistribusi normal multivariat.

RCC untuk model regresi Poisson di atas, dianggap masih memiliki kekurangan yaitu jika terdapat multikoliniearitas dalam variabel prediktor. Oleh karena itu, untuk mengatasi kekurangan tersebut, Kaengthong dan Domthong (2017) mengusulkan modifikasi RCC untuk model regresi Poisson yang dapat bekerja dengan baik dalam kasus ketika dua atau lebih variabel prediktor dan ketika variabel prediktor memiliki multikolinieritas.

Berdasarkan latar belakang masalah yang telah diuraikan diatas, maka permasalahan yang dapat diidentifikasi adalah bagaimana mengaplikasikan modifikasi RCC untuk model regresi Poisson ke dalam kasus jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017?.

Selanjutnya, tujuan yang ingin dicapai dalam penulisan ini adalah: mengaplikasikan modifikasi RCC untuk model regresi Poisson ke dalam kasus jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017.

B. Landasan Teori

Distribusi Poisson merupakan suatu distribusi untuk mencacah suatu peristiwa yang peluang kejadiannya kecil, dimana kejadiannya acak menurut ruang dan waktu, dengan hasil pengamatan berupa variabel diskrit dan antar variabel saling bebas. Fungsi massa peluang dari distribusi Poisson diberikan oleh:

$$P(Y=y;\lambda)=\frac{e^{-\lambda}\lambda^{y}}{y!} \qquad (1)$$

dengan $y = 0,1,2,3,\ldots$, sedangkan $e =$ 2,7183 dan $\lambda > 0$. Ciri-ciri dari distribusi Poisson adalah mempunyai rata-rata dan varians yang sama yaitu λ,

Model regresi Poisson merupakan GLM dengan variabel respon Y diasumsikan berdistribusi Poisson. Berdasarkan Takahashi dan Kurosawa (2016), GLM dapat dituliskan sebagai berikut:

$$g\left(E(Y|X)\right)=\alpha+\boldsymbol{\beta}^{T}\boldsymbol{X};Y|X\sim D(\theta)$$

Model menghubungkan $E(Y|X)$ dengan variabel penjelas oleh $g(\cdot)$, di mana $g(\cdot)$ adalah suatu fungsi yang dapat diturunkan, α adalah intersep, $\boldsymbol{\beta}$ adalah vektor koefisien dan $D(\theta)$ adalah distribusi dalam keluarga eksponensial dengan parameter θ.

Model regresi Poisson mengasumsikan distribusi Poisson $P(\lambda)$ sebagai struktur error dan fungsi logaritmik sebagai fungsi penghubung. Kemudian, model regresi Poisson ditulis sebagai berikut:

$$log\left(E(Y|X)\right)=\alpha+\boldsymbol{\beta}^{T}\boldsymbol{X} \qquad (2)$$

atau

$$log(\lambda)=\alpha+\boldsymbol{\beta}^{T}\boldsymbol{X} \qquad (3)$$

Ketika $Y|X\sim P(\lambda)$, maka:

$$E(Y|X)\ =Var(Y|X)=\lambda$$
$$=exp(\alpha+\boldsymbol{\beta}^{T}\boldsymbol{X}) \qquad (4)$$

Untuk penaksiran parameter regresi Poisson dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Adapun fungsi *log likelihood*-nya adalah sebagai berikut:

$$\ell n\, L(\alpha,\boldsymbol{\beta}) = \sum_{i=1}^{n}\left(-e^{(\alpha+\boldsymbol{\beta}^T\boldsymbol{X})} + y_i(\alpha+\boldsymbol{\beta}^T\boldsymbol{X}) - \ell n(y_i!)\right)$$

Untuk menguji kesesuaian model regresi Poisson dapat digunakan statistik uji menggunakan uji *likelihood ratio* (LR).

Hipotesis:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$; variabel prediktor $X_1, X_2, \dots, X_k$ tidak mempengaruhi variabel respon Y.

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

Statistik Uji:

$$G = -2ln\left[\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right] \quad (5)$$

dimana: $L(\widehat{\omega})$ = nilai *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor dan $L(\widehat{\Omega})$ = nilai *likelihood* untuk model yang lebih lengkap dengan melibatkan variabel prediktor.

Kriteria uji:

Pada kondisi H_0 statistik uji G mendekati distribusi *Chi-Square* (χ^2) dengan derajat bebas k dan taraf $\alpha = 0{,}05$. H_0 ditolak jika statistik uji $G > \chi^2{}_{(1-\alpha;k)}$ atau $p_{value} < \alpha$.

Apabila uji kesesuaian model menunjukan bahwa model regresi Poisson dapat digunakan, maka selanjutnya dilakukan uji signifikansi terhadap masing-masing koefisien regresi dengan uji *Wald*.

Hipotesis:

$H_0: \beta_j = 0$, untuk $j = 1,2,\dots,k$; variabel prediktor X_j tidak mempengaruhi variabel respon Y.

$H_1: \beta_j \neq 0$, untuk $j = 1,2,\dots,k$.

Statistik Uji:

$$W_j = \left(\frac{\widehat{\beta}_j}{\widehat{SE}(\beta_j)}\right)^2 \quad (6)$$

dimana: $\hat{\beta}_j$ = nilai taksiran parameter β_j dan $\widehat{SE}(\beta_j)$ = nilai taksiran *Standard Error* dari $\hat{\beta}_j$.

Kriteria Uji:

Pada kondisi H_0 statistik uji W_j mendekati distribusi Chi-Square (χ^2) dengan derajat bebas 1 dan taraf $\alpha = 0{,}05$. H_0 ditolak jika statistik uji $W_j > \chi^2{}_{(1-\alpha;1)}$ atau $p_{value} < \alpha$.

RCC untuk GLM diusulkan oleh Zheng dan Agresti (2000). RCC dapat digunakan untuk mengukur kekuatan prediksi suatu model karena memenuhi persyaratan yang diberikan yaitu:

a. *Interpretability* (interprebilitas): ukuran kekuatan prediksi harus masuk akal dan memiliki kekuatan diskriminan yang tinggi.
b. *Applicability* (penerapan): harus bisa diterapkan pada berbagai GLM.
c. *Consitency* (konsistensi): dapat dibandingkan diantara rangkaian data yang berbeda.
d. *Affinity* (afiniti): tidak bertentangan dengan ukuran kekuatan prediksi lainnya.

Berdasarkan Zheng dan Agresti (2000) mendefinisikan RCC sebagai berikut:

$$RCC(Y,X) = cor\big(Y, E(Y|X)\big) \quad (7)$$

Jika variabel respon Y memiliki korelasi kuat dengan harapan bersyarat $E(Y|X)$, maka nilai cenderung ke 1 dan menilai model tepat. Sebaliknya, jika nilai cenderung ke 0, maka menilai model tidak tepat.

Berdasarkan Persamaan (7), maka RCC dapat ditulis juga sebagai berikut:

$$RCC(Y,X) = cor\big(Y, E(Y|X)\big) = \sqrt{1 - \frac{E\big(Var(Y|X)\big)}{Var(Y)}} \quad (8)$$

RCC untuk model regresi Poisson diusulkan oleh Takahashi dan Kurosawa (2016), yang dikembangkan dari RCC untuk model GLM.

Misalkan $\boldsymbol{\Sigma}$ adalah matriks definit positif. Asumsikan bahwa $Y|X$ mengikuti model regresi Poisson dan vektor dari variabel prediktor X memiliki distribusi normal multivariat $N_k(\mu,\boldsymbol{\Sigma})$. Kemudian, $E(Y)$ dan

$Var(Y)$ didefinisikan sebagai berikut:

$$E(Y) = E[E(Y|X)]$$
$$= exp\left(\alpha + \boldsymbol{\mu}^T\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta}\right) \quad (9)$$
$$Var(Y) = E(Y)\{1 + E(Y)\ exp(\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta}) - 1\}$$

Dengan memasukkan Persamaan (9) ke dalam Persaman (8), maka RCC untuk model regresi Poisson adalah:

$$R = \sqrt{\frac{exp\left[\alpha+\boldsymbol{\mu}^T\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta}\right]\cdot(exp(\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta})-1)}{1+exp\left[\alpha+\boldsymbol{\mu}^T\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta}\right](exp(\boldsymbol{\beta}^T\boldsymbol{\Sigma}\boldsymbol{\beta})-1)}} \quad (10)$$

Karena RCC adalah nilai populasi, maka diusulkan penaksir RCC dengan menggunakan bentuk eksplisitnya, maka akan diperoleh penaksir $\hat{R}$ dari RCC untuk model regresi Poisson sebagai berikut:

$$\hat{R} = \sqrt{\frac{exp\left[\hat{\alpha}+\boldsymbol{\mu}^T\hat{\boldsymbol{\beta}}+\frac{1}{2}\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}}\right](exp(\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}})-1)}{1+exp\left[\hat{\alpha}+\boldsymbol{\mu}^T\hat{\boldsymbol{\beta}}+\frac{1}{2}\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}}\right](exp(\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}})-1)}} \quad (11)$$

Dimana μ adalah vektor rata-rata variabel prediktor X, ukuran $(p \times 1)$. $\boldsymbol{\Sigma}$ adalah matriks kovarians variabel prediktor X, ukuran $(p \times p)$. α adalah perpotongan dari model regresi skala, ukuran (1×1). $\boldsymbol{\beta}$ adalah vektor dari koefisien regresi, ukuran $(p \times 1)$.

Untuk mencari modifikasi RCC untuk model regresi Poisson, ada beberapa asumsi yang harus dipenuhi yaitu:

1) Uji Kecocokan Distribusi Poisson

Untuk melihat bentuk distribusi dari data respon (Y), maka dapat diketahui melalui uji kecocokan distribusi dengan menggunakan uji *Chi-Square* (χ^2).

$$\chi^2 = \sum_{k=1}^{l}\frac{(n_k-np_k)^2}{np_k} \quad (12)$$

2) Uji Normalitas untuk Masing-masing Variabel Prediktor

Untuk melihat kenormalan setiap data variabel prediktor, maka dilakukan uji normalitas menggunakan uji *Kolomogorov-Smirnov*.

$$D = max|F_0(y) - S_n(y)| \quad (13)$$

3) Uji Multikolinieritas

Untuk melihat multikolinieritas antar data variabel prediktor dapat dilihat dengan menggunakan korelasi.

$$r_{ij} = \frac{n\sum X_iX_j-(\sum X_i)(\sum X_j)}{\sqrt{\{n\sum X_i^2-(\sum X_i)^2\}\{n\sum X_j^2-(\sum X_j)^2\}}} \quad (14)$$

R_{adj}^2 dikembangkan dari koefisien determinan ganda (R^2). Menambahkan lebih banyak variabel prediktor ke dalam model, selalu tampak bahwa R^2 lebih tinggi. Namun, tidak berarti bahwa model berkinerja baik. Di sisi lain, hal tersebut dapat menyebabkan memiliki model dengan *Mean Square Error* (MSE) yang tinggi. Oleh karena itu, untuk mengurangi nilai bias, nilai R_{adj}^2 berkisar antara 0 dan 1 diusulkan sebagai berikut (Mittlbock dan Waldhor, 2000):

$$R^2{}_{Adj} = 1 - \frac{(n-1)}{(n-p-1)}\frac{\sum_{i-1}^{n}(Y_i-\hat{Y}_i)^2}{\sum_{i-1}^{n}(Y_i-\bar{Y}_i)^2}$$

Pedoman untuk modifikasi adalah dengan meningkatkan penduga RCC dari Takahashi dan Kurosawa (2016) dengan menggunakan istilah $(n-p-1)/(n-1)$. Berdasarkan Kaengthong dan Domthong (2017), penaksir baru modifikasi RCC untuk model regresi Poisson sebagai berikut:

$$\hat{R}^* = \frac{(n-p-1)}{(n-1)}\sqrt{\frac{exp\left[\hat{\alpha}+\boldsymbol{\mu}^T\hat{\boldsymbol{\beta}}+\frac{1}{2}\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}}\right](exp(\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}})-1)}{1+exp\left[\hat{\alpha}+\boldsymbol{\mu}^T\hat{\boldsymbol{\beta}}+\frac{1}{2}\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}}\right](exp(\hat{\boldsymbol{\beta}}^T\boldsymbol{\Sigma}\hat{\boldsymbol{\beta}})-1)}} \quad (15)$$

di mana n adalah ukuran sampel dan p adalah jumlah variabel prediktor.

Untuk menemukan nilai modifikasi RCC yang harus dilakukan sebelumnya yaitu menghitung terlebih dahulu vektor rata-rata variabel prediktor $(\boldsymbol{\mu})$ dan matriks varians variabel prediktor $(\boldsymbol{\Sigma})$. Jika $\boldsymbol{\mu}$ dan $\boldsymbol{\Sigma}$ tidak diketahui, maka untuk mencarinya menggunakan rata-rata sampel dan matriks varians dari sampel X. Sehingga taksiran untuk vektor rata-rata adalah rata-rata sampel X, yaitu:

$$\bar{X}_i = \sum_{i=1}^{n}\frac{X_{1i}+X_{2i}+\cdots+X_{ni}}{n} \quad (16)$$

Dan taksiran untuk matriks varians adalah matriks varians sampel S, yaitu:

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_i)^T(X_i - \bar{X}_i) \quad (17)$$

C. Hasil Penelitian dan Pembahasan

Pengujian Asumsi

1. Uji Kecocokan Distribusi Poisson

Hipotesis:

H_0: Jumlah kematian ibu berdistribusi Poisson

H_1: Jumlah kematian ibu tidak berdistribusi Poisson

Statistik uji yang digunakan untuk uji Chi-Square (χ^2) yaitu menggunakan Persamaan (12). Dengan bantuan software Minitab, diperoleh nilai p-value sebesar 0,163. Dengan menggunakan nilai taraf $\alpha = 0,05$, maka diputuskan untuk menerima H_0 karena p-value = 0,163 $> \alpha =$ 0,05, sehingga dapat disimpulkan bahwa data jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017 berdistribusi Poisson.

2. Uji Normalitas untuk Masing-masing Variabel Prediktor

Hipotesis:

H_0: Data variabel prediktor X_j; $j = 1,2,3,4$ berdistribusi Normal

H_1: Data variabel prediktor X_j; $j = 1,2,3,4$ tidak berdistribusi Normal

Statistik uji yang digunakan untuk uji *Kolmogorov-Smirnov* yaitu menggunakan Persamaan (13). Dengan bantuan *software* Minitab, diperoleh nilai *p-value* untuk setiap data variabel prediktor yang dapat dilihat pada Tabel 1.

Tabel 1. Nilai Statistik Uji *Kolmogorov-Smirnov* dan *P-value*

Variabel	**Statistik Uji**	***p-value***	**Keterangan**
X_1	0,113	>0,150	Terima H_0
X_2	0,200	0,133	Terima H_0
X_3	0,118	>0,150	Terima H_0
X_4	0,185	>0,150	Terima H_0

Berdasarkan Tabel 1, dengan menggunakan nilai taraf $\alpha = 0,05$, maka H_0 diterima karena *p-value* $> \alpha$, sehingga dapat disimpulkan bahwa data X_1, X_2, X_3 dan X_4berdistribusi Normal.

3. Uji Multikolinieritas

Hipotesis:

H_0: Tidak terdapat hubungan antara variabel prediktor

H_1: Terdapat hubungan antara variabel prediktor

Statistik uji yang digunakan untuk korelasi yaitu menggunakan Persamaan (14). Dengan bantuan *software* Minitab, diperoleh nilai statistik uji korelasi dan *p-value* antara variabel prediktor yang dapat dilihat pada Tabel 2.

Tabel 2. Nilai Korelasi dan *P-value* antara Variabel Prediktor

		$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	$\boldsymbol{X_3}$
$\boldsymbol{X_2}$	Corr	0,879		
	p-value	0,000		
$\boldsymbol{X_3}$	Corr	0,859	0,732	
	p-value	0,000	0,003	
$\boldsymbol{X_4}$	Corr	0,428	0,159	0,705
	p-value	0,126	0,588	0,005

Dengan menggunakan nilai taraf $\alpha = 0,05$. Berdasarkan Tabel 2, maka terdapat hubungan antara variabel prediktor X_1 dan X_2, X_1 dan X_3, X_2 dan X_3, serta X_3 dan X_4. Sehingga dapat disimpulkan bahwa data persentase ibu bersalin yang ditolong tenaga kesehatan (X_3) cenderung terdapat multikolinieritas.

Pemodelan Regresi Poisson

Penaksiran parameter untuk model regresi Poisson menggunakan metode MLE. Dengan bantuan *software* SAS 9.4, diperoleh hasil nilai taksiran parameter untuk model regresi

Poisson yang disajikan pada Tabel 3.

Tabel 3. Nilai Taksiran Parameter untuk Model Regresi Poisson

Para Meter	Taksiran	*p-value*	Keterangan
Intercept	1,5327	0,6341	
X_1	-0,0097	0,8597	Terima H_0
X_2	0,065	0,0199	Tolak H_0
X_3	-0,0634	0,0401	Tolak H_0
X_4	0,0149	0,0605	Terima H_0

Pengujian signifikasi Model ini menggunakan uji *Likelihood Ratio* (LR) dengan hipotesis:

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$; variabel prediktor X_1, X_2, X_3, X_4 tidak mempengaruhi variabel respon Y.

H_1: paling sedikit ada satu $\beta_j \neq 0$, dengan $j = 1,2,3,4$

Statistik uji yang digunakan adalah Persamaan (6). Dengan bantuan *software* SAS 9.4, diperoleh nilai statistik uji G adalah 11,5196. Berdasarkan tabel *Chi-Square* dengan taraf $\alpha =$ 0,05 dan derajat bebas 4, diperoleh nilai $\chi^2_{(0,95;4)}$= 9,49. H_0 ditolak karena nilai G = 11,5196 $> \chi^2_{(0,95;4)}$= 9,49, maka terdapat paling sedikit ada satu parameter $\beta_1, \beta_2, \beta_3, \beta_4$ yang signifikan.

Pengujian signifikansi parameter model regresi Poisson dilakukan dengan menggunakan uji *Wald* . Hipotesis untuk pengujian ini adalah:

$H_0: \beta_j = 0$, untuk $j = 1,2,3,4$; variabel prediktor X_j tidak mempengaruhi variabel respon Y.

$H_1: \beta_j \neq 0$, untuk $j = 1,2,3,4$; variabel prediktor X_j mempengaruhi variabel respon Y.

Statistik uji yang digunakan adalah Persamaan (7). Dengan menggunakan bantuan *software* SAS 9.4, diperoleh *p-value* untuk uji parsial pada Tabel 3. Hasilnya dapat dilihat bahwa parameter β_2 dan β_3 signifikan pada taraf $\alpha = 0{,}05$ karena *p-value* $< \alpha$.

Pemilihan Model Terbaik Menggunakan $\hat{R}^*$

Berdasarkan pengujian signifikansi model dan parameter, terdapat 2 variabel yang signifikan yaitu X_2 dan X_3. Dari kedua variabel prediktor tersebut, diperoleh 3 kandidat model yang akan dicari nilai modifikasi RCC untuk model regresi Poisson $\hat{R}^*$.

Untuk model regresi Poisson dengan 1 variabel, diperoleh 2 kandidat model. Dengan menggunakan bantuan *software* SAS 9.4, diperoleh hasil nilai taksiran parameter untuk model regresi Poisson dengan 1 variabel yang dapat dilihat pada Tabel 4.

Tabel 4. Nilai Taksiran Parameter untuk Model Regresi Poisson dengan 1 Variabel

Variabel X_2		Variabel X_3	
Parameter	**Taksiran**	**Parameter**	**Taksiran**
Intercept	-0,1644	Intercept	0,7764
X_2	0,0250	X_3	0,0145

Langkah selanjutnya yaitu mencari nilai rata-rata menggunakan Persamaan (16) dan varians menggunakan Persamaan (17). Dengan bantuan Excel 2013, diperoleh nilai rata-rata dan varians sebagai berikut:

Rata-rata dan varians untuk variabel K4 (X_2) yaitu:

$$\overline{\mathrm{X}} = (82{,}9071) \quad \mathrm{S} = (143{,}6007)$$

Rata-rata dan varians untuk variabel Nakes (X_4) yaitu:

$$\overline{\mathrm{X}} = (80{,}1071) \quad \mathrm{S} = (105{,}6546)$$

Langkah terakhir yaitu mencari nilai modifikasi RCC untuk model regresi Poisson menggunakan Persamaan (15), dimana diketahui untuk n = 23 dan p = 1. Dengan bantuan Excel 2013, diperoleh nilai $\hat{R}^*$ dari masing-masing kandidat model yang disajikan pada Tabel 4.5.

Tabel 5. Nilai $\hat{R}^*$ untuk Model Regresi Poisson dengan 1 Variabel

Model	$\hat{R}^*$
$log\ (\lambda) = \alpha + \beta_2 X_2$	0,5826
$log\ (\lambda) = \alpha + \beta_3 X_3$	0,3407

Berdasarkan Tabel 5 dapat dilihat bahwa model regresi Poisson dengan model $log\ (\lambda) = \alpha + \beta_2 X_2$ memiliki nilai $\hat{R}^*$ lebih besar dibandingkan dengan yang lain yaitu sebesar 0,5826 dan merupakan model terbaik dari model regresi Poisson dengan 1 variabel.

Untuk model regresi Poisson dengan 2 variabel, diperoleh 1 kandidat model. Dengan menggunakan bantuan *software* SAS 9.4, diperoleh hasil nilai taksiran parameter untuk model regresi Poisson dengan 2 variabel yang dapat dilihat pada Tabel 6.

Tabel 6. Nilai Taksiran Parameter untuk Model Regresi Poisson dengan 2 Variabel

Parameter	Taksiran
Intercept	0,0554
X_2	0,0367
X_3	-0,0150

Dengan bantuan Excel 2013, diperoleh nilai rata-rata dan varians sebagai berikut:

Rata-rata dan varians untuk variabel K4 (X_2) + Nakes (X_4) yaitu:

$$\overline{\mathbf{X}} = \begin{pmatrix} 82,9071 \\ 80,1071 \end{pmatrix} \mathbf{S} = \begin{pmatrix} 143,6007 & 90,1153 \\ 90,1153 & 105,6546 \end{pmatrix}$$

Langkah terakhir yaitu mencari nilai modifikasi RCC untuk model regresi Poisson menggunakan Persamaan (15), dimana diketahui untuk n = 23 dan p = 2. Dengan bantuan Excel 2013, diperoleh nilai $\hat{R}^*$ dari masing-masing kandidat model yang disajikan pada Tabel 7.

Tabel 7. Nilai $\hat{R}^*$ untuk Model Regresi Poisson dengan 2 Variabel

Model	$\hat{R}^*$
$log\ (\lambda) = \alpha + \beta_2 X_2 + \beta_3 X_3$	0,5798

Berdasarkan Tabel 7, dapat dilihat bahwa model regresi Poisson dengan model $log\ (\lambda) = \alpha + \beta_2 X_2 + \beta_3 X_3$ memiliki nilai $\hat{R}^*$ yaitu sebesar 0,5798.

Nilai $\hat{R}^*$ terbesar dari setiap model regresi Poisson dengan 1 variabel dan 2 variabel disajikan pada tabel 8.

Tabel 8. Nilai $\hat{R}^*$ Terbesar

Model	$\hat{R}^*$
$log\ (\lambda) = \alpha + \beta_2 X_2$	0,5826
$log\ (\lambda) = \alpha + \beta_3 X_3$	0,3407
$log\ (\lambda) = \alpha + \beta_2 X_2 + \beta_3 X_3$	0,5798

Berdasarkan Tabel 8, dapat dilihat bahwa model regresi Poisson dengan model $log\ (\lambda) = \alpha + \beta_2 X_2$ memiliki nilai $\hat{R}^*$ lebih besar dibandingkan dengan yang lain yaitu sebesar 0,5826.

Pemodelan Regresi Poisson Terbaik

Penaksiran parameter untuk model regresi Poisson terbaik. Dengan menggunakan *software* SAS 9.4, diperoleh hasil nilai taksiran parameter untuk model regresi Poisson terbaik yang disajikan pada Tabel 9.

Tabel 9. Nilai Taksiran Parameter untuk Model Regresi Poisson Terbaik

Parameter	Taksiran	*p-value*	Keterangan
Intercept	-0,1644	0,8471	Terima H_0
X_2	0,0250	0,0113	Tolak H_0

Pengujian signifikansi model terbaik.

Hipotesis:

$H_0: \beta_2 = 0$ dan $H_1: \beta_2 \neq 0$

Statistik uji yang digunakan

yaitu menggunakan Persamaan (6). Dengan bantuan *software* SAS 9.4, diperoleh nilai statistik uji G adalah 6,9844.

Berdasarkan tabel *Chi-Square* dengan taraf $\alpha =$ 0,05 dan derajat bebas 1, diperoleh nilai $\chi^2{}_{(0,95;1)}=$ 3,84. H_0 ditolak karena nilai G = 6,9844 $> \chi^2{}_{(0,95;1)}=$ 3,84, maka parameter β_2 signifikan.

Hipotesis:

$H_0 : \beta_2 = 0$ dan $H_1 : \beta_2 \neq 0$

Statistik uji yang digunakan yaitu menggunakan Persamaan (7). Dengan menggunakan bantuan *software* SAS 9.4, diperoleh nilai *p-value* untuk uji parsial pada Tabel 4.9. Hasilnya dapat dilihat bahwa parameter β_2 signifikan pada taraf $\alpha =$ 0,05 karena *p-value* $< \alpha$. Artinya, persentase ibu hamil melaksanakan program K4 (X_2) mempengaruhi jumlah kematian ibu (Y).

Berdasarkan hasil nilai taksiran dari Tabel 9, maka model regresi Poisson terbaik adalah:

$\lambda = \exp(0{,}0250X_2)$

dimana: $X_2 =$ persentase ibu hamil melaksanakan program K4.

Interpretasi parameter $\beta_2$sebagai berikut:

b_2 = 0,0250

Setiap kenaikan 1% persentase ibu hamil melaksanakan program K4 (X_2), maka rata-rata jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017 (Y) adalah sebesar exp(0,0250) = 1,0253 kali.

Kesimpulan

Berdasarkan pembahasan pada bab sebelumnya, diperoleh kesimpulan bahwa modifikasi RCC untuk model regresi Poisson yang diaplikasikan untuk data jumlah kematian ibu di Provinsi Kalimantan Barat tahun 2017 dapat digunakan sebagai ukuran kekuatan prediksi alternatif dalam pemilihan model terbaik. Hasilnya diperoleh model terbaik dengan nilai $\hat{R}^*$ sebesar 0,5826 dengan model regresi Poisson yaitu:

$\ell og\ (\lambda) = 0{,}0250X_2$

D. Saran

Berdasarkan hasil dari penelitian yang dilakukan, saran yang dapat dikemukakan adalah penelitian berikutnya dapat melakukan penelitian dengan menerapkan modifikasi RCC untuk berbagai model GLM yang lainnya. Pengujian signifil

Daftar Pustaka

Achmad, A.I. dan Hajarisman, N. (2011). *Analisis Regresi Lanjut.* Seri buku Ajar Universitas Islam Bandung.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York: John Wiley & Sons.

Dinas Kesehatan Aceh. (2018*). Profil Kesehatan Aceh Tahun 2017*. Aceh: Dinas Kesehatan Aceh.

Dobson, A.J. and Barnett, A.G. (2008). *An Introduction to Generalized Linear Models*. In: Statistical Science Series, Chapman and Hall/CRC Texts.

Hajarisman Nusar. (2008). *Analisis Data Kategorik.* Seri buku Ajar Universitas Islam Bandung.

Johson, R.A. and Wichern, D.W. *Applied Multivariate Statistical Analysis* (3rd Ed).New Jersey: Prentice Hall, Englewood Cliffs.

Kaengthong, N dan Domthong, U. (2017). *Modified Regression Correlation Coefficient for a Poisson Regression Model*. Journal of Physics: Conf Series 890.

Kartaningrum, E.D. (2017). *Faktor yang Mempengaruhi Angka Kematian Ibu*. Surakarta: CV Kekata Group.

McCullagh, P. and J.A. Nelder. (1983).

Generalized Linear Models. (2nd Ed). New York: Chapman & Hall/CRC.

Mittlbock, M. and Waldhor, T. (2000). *Adjustments for R^2-measures for Poisson Regression Mode*l. Computational Statistics and Data Analysis 34 p 461-472.

Sembiring. (2003). *Analisis Regresi*. Bandung: Institut Teknologi Bandung.

Sudjana. (2005). *Metode Statistika.* Bandung: PT. Tarsito.

Takahashi, A. dan Kurosawa, T. (2015). *Regression Correlation Coefficient for a Poisson Regression Model*. Computational Statistics and Data Analysis Journal : Elsevier B.V.

Zheng, B. dan Agresti, A. (2000). *Summarizing the Predictive Power of a Generalized Linear Model*. USA: John Wiley & Sons.