

**PEMODELAN JUMLAH ANAK PUTUS SEKOLAH  
DI PROVINSI BALI DENGAN PENDEKATAN  
*SEMI-PARAMETRIC GEOGRAPHICALLY  
WEIGHTED POISSON REGRESSION***

**GUSTI AYU RATIH ASTARI<sup>1</sup>, I GUSTI AYU MADE SRINADI<sup>2</sup>,  
MADE SUSILAWATI<sup>3</sup>**

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail: <sup>1</sup>ayuratih.astari@yahoo.co.id, <sup>2</sup>srinadiigustiayumade@yahoo.co.id,  
<sup>3</sup>susilawati.made@gmail.com

***Abstract***

*Dropout number is one of the important indicators to measure the human progress resources in education sector. This research uses the approaches of Semi-parametric Geographically Weighted Poisson Regression to get the best model and to determine the influencing factors of dropout number for primary education in Bali. The analysis results show that there are no significant differences between the Poisson regression model with GWPR and Semi-parametric GWPR. Factors which significantly influence the dropout number for primary education in Bali are the ratio of students to school, ratio of students to teachers, the number of families with the latest educational fathers is elementary or junior high school, illiteracy rates, and the average number of family members.*

**Keywords:** GWPR, Semi-parametric GWPR, dropout

## **1. Pendahuluan**

Putus sekolah merupakan salah satu indikator yang berguna untuk mengukur kemajuan sumber daya manusia pada bidang pendidikan pada tiap wilayah. Oleh karena itu, masalah anak yang putus sekolah perlu mendapatkan perhatian. Untuk menekan laju pertambahan jumlah anak putus sekolah tersebut dapat dilakukan dengan cara mengetahui faktor-faktor yang berpengaruh terhadap jumlah anak putus sekolah dan berpotensi dalam meningkatkan laju pertumbuhan anak yang putus sekolah.

Faktor-faktor yang memengaruhi jumlah putus sekolah pada tiap wilayah berbeda-beda tergantung pada karakteristik dari masing-masing daerah tersebut. Oleh karena itu, diperlukan suatu analisis statistika yang memperhitungkan faktor geografis pada tiap pengamatannya, salah satunya adalah *Semi-parametric Geographically Weighted Poisson Regression*. Pada penelitian ini akan dilakukan analisis dengan menggunakan pendekatan *Semi-parametric Geographically Weighted Poisson Regression* untuk mendapatkan model terbaik dan mengetahui faktor-faktor yang memengaruhi jumlah anak putus sekolah usia pendidikan dasar pada tiap kecamatan di Provinsi Bali tahun 2010.

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2,3</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

### Regresi Poisson

Regresi Poisson adalah model regresi yang digunakan untuk memodelkan hubungan antara variabel respon ( $Y$ ) yang berupa data *count* (jumlah) dengan variabel bebas ( $X$ ). Regresi Poisson merupakan model yang berdasarkan pada penggunaan distribusi Poisson. Fungsi peluang dari distribusi Poisson pada suatu variabel random  $y$  didefinisikan seperti pada [4], yaitu:

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

Model regresi Poisson dapat dinyatakan:

$$y_i \sim \text{Poisson} \left[ \exp\left(\sum_{j=0}^k \beta_j x_{ij}\right) \right] \quad (2)$$

Metode regresi Poisson biasanya diterapkan pada penelitian kesehatan masyarakat, biologi, dan teknik dengan variabel responnya ( $y$ ) berupa cacahan objek yang merupakan fungsi dari sejumlah karakteristik tertentu ( $x$ ), sesuai dengan yang dijelaskan pada [1]. Penaksiran parameter model regresi Poisson dilakukan dengan menggunakan metode *Maximum Likelihood Estimator* (MLE).

### *Geographically Weighted Poisson Regression* (GWPR)

*Geographically Weighted Poisson Regression* (GWPR) merupakan model regresi yang memperhitungkan faktor lokasi pengamatan. Pada model GWPR, variabel respon  $Y$  diprediksi dengan menggunakan variabel bebas  $X$  yang masing-masing koefisien regresinya bergantung pada lokasi dimana data tersebut diamati. Penjelasan lebih lanjut mengenai model GWPR diberikan pada [5]. Model GWPR dapat ditulis sebagai berikut:

$$y_i \sim \text{poisson} \left[ \exp\left(\sum_{j=0}^k \beta_j(u_i, v_i) x_{ij}\right) \right] \quad (3)$$

dengan:

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

$x_{ij}$  : nilai observasi variabel bebas ke- $j$  pada pengamatan lokasi  $(u_i, v_i)$

$\beta_j(u_i, v_i)$  : koefisien regresi untuk setiap variabel  $x$  pada lokasi  $(u_i, v_i)$

### *Semi-parametric Geographically Weighted Poisson Regression*

*Semi-parametric Geographically Weighted Poisson Regression* adalah suatu model regresi yang memperhitungkan faktor lokasi pada tiap titik pengamatan yang merupakan pengembangan dari model GWPR. Pada model *Semi-parametric* GWPR, variabel respon  $y$  diprediksi dengan menggunakan variabel bebas  $x$  yang masing-masing koefisien regresinya ( $\beta_j(u_i, v_i)$ ) bergantung pada lokasi dimana data diamati dan koefisien regresi  $\gamma_m$  yang bersifat konstan. Model *Semi-parametric* GWPR dapat dinyatakan sebagai berikut:

$$y_i \sim \text{poisson} \left[ \exp\left(\sum_{j=0}^{k^*} \beta_j^*(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^k \gamma_m x_{im}\right) \right] \quad (4)$$

dengan:

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

$x_{ij}$  : nilai observasi variabel prediktor  $j$  pada pengamatan lokasi  $(u_i, v_i)$

$\beta_j^*(u_i, v_i)$  : koefisien regresi untuk setiap variabel  $x$  pada lokasi  $(u_i, v_i)$

$\gamma_m$  : koefisien regresi yang bersifat konstan  
 $x_{im}$  : nilai observasi variabel prediktor  $m$  dengan koefisien regresi yang bersifat konstan

Penaksiran parameter model *Semi-parametric* GWPR dilakukan dengan menggunakan metode *Maximum Likelihood Estimator* (MLE) dan kemudian diselesaikan dengan metode iterasi numerik yaitu iterasi Newton-Raphson.

## 2. Metode Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Bali. Variabel respon pada penelitian ini adalah jumlah anak putus sekolah usia pendidikan dasar pada tiap kecamatan di Provinsi Bali berdasarkan hasil Sensus Penduduk Provinsi Bali tahun 2010. Sedangkan variabel bebas yang digunakan terdiri dari rasio siswa terhadap sekolah ( $X_1$ ), rasio siswa terhadap guru ( $X_2$ ), jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP ( $X_3$ ), angka buta huruf ( $X_4$ ), dan rata-rata jumlah anggota keluarga ( $X_5$ ).

Langkah-langkah dalam penelitian ini adalah sebagai berikut: (1) Pemeriksaan multikolinieritas antar variabel bebas, (2) Menganalisis model regresi Poisson, (3) Menganalisis model GWPR, (4) Menganalisis model *Semi-parametric* GWPR, dan (5) Mendapatkan model terbaik.

## 3. Hasil dan Pembahasan

Sebelum dilakukan pembentukan suatu model, langkah awal yang dilakukan adalah pengujian multikolinieritas antar variabel bebas. Salah satu kriteria yang digunakan untuk melihat adanya multikolinieritas adalah dengan menggunakan nilai VIF (*Variance Inflation Factors*). Berdasarkan nilai VIF dari masing-masing variabel bebas diperoleh hasil bahwa di antara variabel bebas tidak mengandung multikolinieritas, sehingga variabel-variabel bebas tersebut dapat digunakan dalam pembentukan model selanjutnya, yaitu model regresi Poisson.

Berdasarkan analisis dengan menggunakan regresi Poisson, maka model regresi Poisson yang terbentuk untuk data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali adalah:

$$\hat{\mu} = \exp(3,840824 + 0,003199X_1 + 0,077056X_2 - 0,000087X_3 + 0,030462X_4 + 0,071964X_5)$$

Berdasarkan model regresi Poisson yang terbentuk maka dapat disimpulkan bahwa rasio siswa terhadap sekolah ( $X_1$ ), rasio siswa terhadap guru ( $X_2$ ), angka buta huruf ( $X_4$ ), dan rata-rata jumlah anggota keluarga ( $X_5$ ) dapat meningkatkan jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali jika variabel-variabel tersebut bertambah. Sedangkan untuk variabel jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP ( $X_3$ ) dapat menurunkan jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali jika variabel tersebut bertambah.

Selanjutnya dilakukan pemodelan dengan menggunakan model *Geographically Weighted Poisson Regression* (GWPR). Langkah-langkah yang dilakukan pada pembentukan model GWPR adalah menghitung jarak *Euclidean* antar lokasi pengamatan, menentukan nilai *bandwidth* optimum, dan menghitung matriks pembobot dari masing-masing titik lokasi. Pembobot spasial yang digunakan pada model GWPR untuk menghitung matriks pembobot tersebut adalah pembobot fungsi kernel *Gaussian*. Fungsi kernel *Gaussian* dapat ditulis seperti pada [2], yaitu:

$$w_j(u_i, v_i) = \exp\left(-\frac{1}{2}(d_{ij}/G)^2\right) \quad (5)$$

Pada model GWPR dilakukan pengujian kesesuaian model GWPR dengan model regresi Poisson dan pengujian parameter model GWPR secara parsial. Pengujian kesesuaian model dilakukan dengan menggunakan uji *F* untuk mengetahui apakah model GWPR lebih sesuai digunakan dibandingkan dengan model global (regresi Poisson). Berikut ini adalah tabel uji kesesuaian model GWPR:

Tabel 1. Uji Kesesuaian Model GWPR

Model	Devians	Df	Devians/df	$F_{hitung}$	$F_{(0.05,51,40)}$
Poisson	3533,758	51,000	69,289	1,38514	1,649979
GWPR	2062,090	41,223	50,023		

Sumber : Data diolah tahun 2013

Berdasarkan Tabel 1 dapat diambil keputusan gagal tolak  $H_0$  yang berarti bahwa tidak ada perbedaan yang signifikan antara model GWPR dengan model regresi Poisson. Tabel 1 juga menunjukkan bahwa pada model regresi Poisson dan GWPR mengandung kasus overdispersi karena nilai devians dibagi derajat bebasnya lebih dari 1. Variabel-variabel yang memengaruhi jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali dapat diketahui dengan melakukan pengujian parameter model GWPR secara parsial. Berdasarkan hasil analisis GWPR diperoleh hasil bahwa terdapat variabel yang berpengaruh signifikan di semua titik lokasi, yaitu variabel  $X_1$ (rasio siswa terhadap sekolah),  $X_2$ (rasio siswa terhadap guru),  $X_3$ (jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP), dan  $X_4$ (angka buta huruf). Variabel-variabel tersebut diduga berpengaruh secara global pada model *Semi-parametric* GWPR, sedangkan variabel  $X_5$  (rata-rata jumlah anggota keluarga) yang berpengaruh signifikan hanya di beberapa kecamatan diduga secara lokal pada model *Semi-parametric* GWPR, seperti yang dijelaskan pada [3].

Kemudian dibentuk model *Semi-parametric Geographically Weighted Poisson Regression* dengan langkah-langkah yaitu menghitung jarak *Euclidean* antar lokasi pengamatan, menentukan nilai *bandwidth* optimum, dan menghitung matriks pembobot masing-masing titik lokasi dengan menggunakan pembobot spasial fungsi kernel *Gaussian*. Pada pembentukan model *Semi-parametric Geographically Weighted Poisson Regression* dilakukan

pengujian kesesuaian model dan pengujian parameter model secara parsial. Berikut ini adalah tabel uji kesesuaian model untuk *Semi-parametric* GWPR dengan menggunakan uji  $F$ .

Tabel 2. Uji Kesesuaian Model Semi-parametric GWPR

Model	Devians	Df	Devians/df	$F_{hitung}$	$F_{(0.05,51,47)}$
Poisson	3533,758	51,000	69,289	1,16715	1,612235
S-GWPR	2835,148	47,757	59,366		

Sumber : Data diolah tahun 2013

Tabel 2 menunjukkan bahwa nilai  $F_{hitung} < F_{(0.05,51,47)}$ , sehingga keputusannya adalah gagal tolak  $H_0$  yang berarti bahwa tidak ada perbedaan yang signifikan antara model *Semi-parametric* GWPR dengan model regresi Poisson. Selain itu juga terlihat bahwa pada model *Semi-parametric* GWPR mengandung kasus overdispersi. Berdasarkan pengujian parameter model *Semi-parametric* GWPR secara parsial dengan menggunakan uji  $t$  diperoleh hasil bahwa variabel  $X_1, X_2, X_3$ , dan  $X_4$  berpengaruh signifikan di semua kecamatan, sedangkan variabel  $X_5$  berpengaruh signifikan hanya pada beberapa kecamatan di Provinsi Bali.

#### 4. Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang dilakukan maka diperoleh keputusan tidak ada perbedaan yang signifikan antara model regresi Poisson dengan model GWPR dan model *Semi-parametric* GWPR, sehingga dapat disimpulkan bahwa faktor spasial/lokasi tidak berpengaruh terhadap data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010. Dari model yang dihasilkan menunjukkan bahwa pada data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 mengandung kasus overdispersi, oleh karena itu model terbaik untuk data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 adalah model regresi yang dapat mengatasi adanya kasus overdispersi, misalnya model regresi Binomial Negatif dan regresi *Generalized Poisson*.

Variabel-variabel yang berpengaruh signifikan terhadap jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 berdasarkan model regresi Poisson adalah variabel rasio siswa terhadap sekolah pada tiap kecamatan ( $X_1$ ), rasio siswa terhadap guru pada tiap kecamatan ( $X_2$ ), jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP ( $X_3$ ), angka buta huruf ( $X_4$ ), dan rata-rata jumlah anggota keluarga ( $X_5$ ) dengan variabel rasio siswa terhadap guru pada tiap kecamatan ( $X_2$ ) merupakan variabel yang memiliki pengaruh paling besar dibandingkan dengan keempat variabel signifikan yang lainnya.

**Daftar Pustaka**

- [1] Baharuddin. 2005. Ukuran  $R^2$  dalam Model Regresi Poisson. *Integral*, vol.10, no.3, November, pp. 114-121.
- [2] Charlton M., and Fotheringham, A.S. 2009. Geographically Weighted Regression White Paper. *Jurnal*. National University of Ireland Maynooth.
- [3] Erdkhadifa, R. 2012. Perbandingan Geographically Weighted Poisson Regression, Geographically Weighted Poisson Regression Semiparametric (Studi Kasus : Kematian Demam Berdarah Dengue di Jawa Timur). *Skripsi*. Surabaya : Institut Teknologi Sepuluh November.
- [4] Myers, R.H. 1990. *Classical and Modern Regression with Application, second edition*. Boston : PWS-KENT Publishing Company.
- [5] Nakaya,T., Fotheringham, A.S., Brunson, C. and Charlton, M. 2005. Geographically Weighted Poisson Regression for Disease Association Mapping. *Statistics in Medicine*, vol. 24, no. 17, pp. 2695-2717.