

**Model Regresi Logistik Spasial
(Studi Kasus: Penyebaran Penyakit Tuberkulosis Paru di setiap Kelurahan di Kota Samarinda pada Tahun 2013)**

Spatial Logistic Regression Model (Case Study: The Spread of Pulmonary Tuberculosis disease in every village in Samarinda City in the Year 2013)

Tiara Nurul Ma'ala¹, Desi Yuniarti², dan Memi Nor Hayati³

¹Laboratorium Statistika Komputasi FMIPA Universitas Mulawarman

^{2,3}Program Studi Statistika FMIPA Universitas Mulawarman

Email: tiaranurulmaalalah@gmail.com¹, desy_yunt@yahoo.com², memiorhayati@yahoo.co.id³

Abstract

Logistic regression modeling procedure is applied to model the response variable (Y) which is based on one or more categorical explanatory variable (X) which is categorical or continuous. In the application of logistic regression is often found that there are spatial influences that affect the model. The existence of spatial relationships between regions that cause necessary to accommodate the spatial diversity into the model, so that the analysis used logistic regression spatial. First law of geography says that everything is related to everything else, but near things are more related than distant things. Then, when a region becomes a major cause of the spread of a disease is suspected, the region will provide the spread of a disease to the new area adjacent to it. The way to find out the adjacent area with the same characteristics can be done with spatial logistic regression method. The spread of TB disease in Samarinda City is quite high. TB is a chronic disease which has been known by the public and feared of its infection. This study's aim is to determine the appropriate model to estimate the spread of TB disease. The results showed that the correct model to predict the spread of TB by using the backwards elimination method is $g(X) = -3,803 + 0,336 X_4 + 0,124 X_7$. From this model it is known that the factors that influence the number of people with TB disease in every village in Samarinda City in the year 2013 are the number of primary school in every village and the spatial effect. This means that there is the influence of spatial factors to the spread of TB disease in every village in Samarinda City in the Year 2013.

Keywords: Tuberculosis (TB), Logistic regression spatial.

Pendahuluan

Regresi logistik sama halnya dengan regresi linier, yang membedakan model regresi logistik dengan model regresi linier bahwa variabel respon dalam regresi logistik adalah bersifat biner atau *dichotomus*, sedangkan variabel respon dalam regresi linier adalah kontinu (Hosmer & Lemeshow, 2000).

Penerapan analisis regresi logistik sudah sangat luas, namun pada penerapannya seringkali ditemukan bahwa terdapat pengaruh spasial yang mempengaruhi model. Adanya hubungan spasial antar wilayah menyebabkan perlu mengakomodir keragaman spasial ke dalam model, sehingga analisis yang digunakan menjadi analisis regresi logistik spasial.

Hukum I Geografi berbunyi, *'everything is related to everything else, but near things are more related than distant things'* (Lee & Wong, 2001). Maksud lainnya adalah segala sesuatu berhubungan satu sama lain, dan sesuatu yang berada lebih dekat mempunyai hubungan yang lebih erat dibandingkan dengan yang berada lebih jauh. Oleh karena itu, apabila suatu wilayah menjadi penyebab utama penyebaran suatu penyakit maka diduga wilayah tersebut akan memberikan penyebaran suatu penyakit yang baru bagi wilayah yang berdekatan dengannya. Cara

untuk mengetahui daerah yang berdekatan akan memiliki karakteristik yang sama dapat dilakukan dengan metode regresi logistik spasial (Lee & Wong, 2001).

Penyakit TB (*Tuberculosis*) merupakan penyakit kronis atau menahun yang telah lama dikenal oleh masyarakat luas dan ditakuti karena penyakit TB dapat menular. Menurut Robert Kock (1882), secara meyakinkan telah dapat memberikan bukti bahwa TB adalah suatu penyakit infeksi yang disebabkan oleh bakteri yang bernama *Mycobacterium tuberculosis*. Pada Tahun 2013 penemuan kasus TB Paru BTA positif (+) di Kalimantan Timur mencapai 1.969. Penelitian ini didasarkan pada permasalahan penyebaran penyakit TB di Kota Samarinda yang cukup tinggi.

Adapun batasan masalah pada penelitian ini adalah penyelesaian masalah untuk mengetahui pengaruh penyebaran penyakit TB dengan menggunakan metode regresi logistik spasial dengan pendekatan matriks *contiguity*. Data yang digunakan untuk mengetahui pengaruh penyebaran penyakit TB adalah data jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013.

Tujuan yang ingin dicapai dari penelitian ini adalah yang pertama untuk Menentukan model

regresi logistik spasial yang tepat untuk menduga penyebaran penyakit TB tersebut dan untuk mengetahui pengaruh dari faktor spasial terhadap penyebaran penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013.

Analisis Regresi Logistik

Analisis Regresi logistik memiliki tujuan untuk mendapatkan model terbaik dan sederhana yang menggambarkan hubungan antara variabel respon dengan variabel-variabel penjelas. Syarat utama dalam regresi logistik adalah variabel responnya berupa variabel biner yaitu variabel diskrit dengan dua nilai. Misalnya diambil ilustrasi, bila respon tidak terjadi diberi nilai 0 dan bila respon terjadi diberi nilai 1, sedangkan variabel penjelasnya dapat berupa variabel kuantitatif dan kualitatif (Hosmer dan Lemeshow, 2001).

Untuk mencari persamaan logistiknya maka model yang dipakai adalah:

$$f(X) = \frac{e^{s_0 + \sum_{j=1}^k s_j X_j}}{1 + e^{s_0 + \sum_{j=1}^k s_j X_j}} \tag{1}$$

Penaksiran Parameter Model Regresi Logistik

Penaksiran parameter pada model regresi logistik yang mempunyai variabel respon berskala biner adalah menggunakan metode MLE. Pada dasarnya metode MLE menetapkan asumsi distribusi Bernoulli dan obyek pengamatan yang saling bebas atau memberikan nilai taksiran parameter dengan memaksimumkan fungsi likelihood (*likelihood function*) (Agresti, 2002).

Pengujian Parameter Model Regresi Logistik

Pengujian signifikansi parameter merupakan pemeriksaan untuk menentukan apakah variabel penjelas dalam model signifikan atau berpengaruh secara nyata terhadap variabel respon. Pengujian signifikansi parameter terdiri dari dua uji yaitu sebagai berikut:

1) Pengujian Signifikansi Parameter Secara Simultan

Pengujian secara simultan juga merupakan uji model *chi-square* yang digunakan untuk menguji parameter hasil estimasi secara bersama (Hosmer dan Lemeshow, 2000).

Hipotesis:

H_0 : $\beta_1 = \beta_2 = \dots = \beta_j$ (secara simultan variabel penjelas tidak berpengaruh terhadap variabel respon)

H_1 : Minimal terdapat satu $\beta_j \neq 0$, dimana $j = 0, 1, 2, 3, \dots, k$ (secara simultan variabel penjelas berpengaruh terhadap variabel respon)

Statistik Uji:

Statistik Uji yang digunakan adalah *G likelihood ratio*:

$$G = -2 \ln \left(\frac{\frac{n_1^{n_1} n_0^{n_0}}{n^n}}{\prod_{i=1}^n f_i^{Y_i} (1-f_i)^{(1-Y_i)}} \right) = -2 \ln \frac{L_0}{L_k} \tag{2}$$

Daerah Penolakan H_0 :

H_0 ditolak apabila $G > \chi^2_{(r,k)}$, dengan nilai $\chi^2_{(r,k)}$ dapat diperoleh dari tabel *chi-square* atau tolak H_0 bila *p-value* < α .

2) Pengujian Signifikansi Parameter Secara Parsial

Dalam uji parsial ini, pengujian dilakukan setiap variabel penjelas secara individual akan menunjukkan apakah suatu variabel penjelas layak untuk masuk dalam model atau tidak (Hosmer dan Lemeshow, 2000):

Hipotesis

H_0 : $\beta_j = 0$, dimana $j = 0, 1, 2, 3, \dots, k$ (tidak ada pengaruh variabel penjelas ke- j terhadap variabel respon).

H_1 : $\beta_j \neq 0$, dimana $j = 0, 1, 2, 3, \dots, k$ (ada pengaruh variabel penjelas ke- j terhadap variabel respon).

Statistik uji

$$W = \frac{\hat{s}_j}{SE(\hat{s}_j)}; j = 1, 2, \dots, k \tag{3}$$

Daerah Penolakan H_0 :

H_0 ditolak apabila $|W| > Z_{(\alpha/2)}$, dengan nilai $Z_{(\alpha/2)}$ dapat diperoleh dari tabel distribusi normal atau tolak H_0 bila *p-value* < α .

Pengujian Kesesuaian Model (Goodness of Fit)

Uji ini digunakan untuk menilai kecocokan model dengan membandingkan hasil pengamatan dengan nilai dugaan.

Hipotesis:

H_0 : Model sesuai (tidak ada perbedaan antara hasil pengamatan dengan nilai dugaan)

H_1 : Model tidak sesuai (ada perbedaan antara hasil pengamatan dengan nilai dugaan)

Statistik Uji:

Statistik Uji yang digunakan adalah Uji Hosmer dan Lemeshow (\hat{C}):

$$\hat{C} = \sum_{j=1}^k \frac{(O_j - n_j \hat{f}_j)^2}{n_j \hat{f}_j (1 - \hat{f}_j)} \tag{4}$$

Daerah Penolakan H_0 :

H_0 ditolak apabila $\hat{C} > t^2_{(k-2)}$, dengan nilai $t^2_{(k-2)}$ dapat diperoleh dari tabel *chi-square* atau H_0 ditolak bila $p\text{-value} < \alpha$.
(Hosmer dan Lemeshow, 2000)

Penafsiran Koefisien Model Regresi Logistik

Dalam menginterpretasikan atau menafsirkan koefisien β_j pada regresi logistik, hal yang harus selalu diperhatikan adalah jenis variabel penjelasnya, berupa dikotom (terdiri dua kategori), polikotomus (variabelnya memiliki lebih dari dua kategori), atau kontinyu. Sedangkan untuk menginterpretasikan koefisien parameter digunakan *odds ratio* (OR).

Odds ratio adalah nilai perbandingan *odds*. *Odds ratio* merupakan perbandingan *odds* untuk $X = 1$ terhadap *odds* untuk $X = 0$, dilambangkan dengan OR dan dinyatakan dalam persamaan:

$$OR = \frac{f(1)/(1-f(1))}{f(0)/(1-f(0))} = \frac{f(1) \times (1-f(0))}{f(0) \times (1-f(1))} \tag{5}$$

Bila suatu model logistik memiliki variabel penjelas yang kontinyu, maka interpretasi dari koefisien variabel tersebut akan tergantung pada bagaimana variabel tersebut diperlukan di dalam model (Hosmer dan Lemeshow, 2000).

Nagelkerke R²

Nilai Nagelkerke R² ini akan menunjukkan seberapa besar variabel-variabel penjelas dalam penelitian ini menjelaskan variabel responnya. Nilai Nagelkerke R² biasanya dibentuk dalam persen agar dapat mengetahui dengan pasti seberapa jauh penjelasan variabel-variabel penjelas terhadap variabel respon (Yatani, 2014).

Analisis Spasial

Analisis spasial merupakan analisis yang memasukkan pengaruh spasial atau ruang ke dalam analisisnya. Selain memperhatikan temporal atau waktu, juga ketinggian atau variabel utama lainnya seperti kelembaban masuk di dalam variabel yang harus dipastikan. Apabila batasan ruang lebih bersifat *non made* seperti halnya tata ruang, maka istilah spasial lebih *concern* kepada ekosistem. Spasial mempunyai arti sesuatu yang dibatasi oleh ruang. Komunikasi dan atau transformasi, data spasial menunjukkan posisi, ukuran dan kemungkinan hubungan topologis (bentuk dan tata letak) dari obyek di muka bumi ini. Pada analisis spasial selalu ada korelasi antar ruang yang biasa disebut korelasi spasial (Ward & Gleditsch, 2008).

Matriks Contiguity

Matriks *contiguity* adalah matriks yang menggambarkan hubungan kedekatan antar daerah. Kedekatan suatu daerah dihitung berdasarkan *Queen criterion*. *Queen criterion* merupakan gerakan. Pemberian nilai 1 diberikan

jika daerah- i bertetangga langsung dengan daerah- j sedangkan nilai 0 diberikan jika daerah- i tidak bertetangga dengan daerah- j (Lee & Wang, 2001).

Matriks *contiguity* sering disebut dengan *binary matrix* dan juga *connectivity matrix*, yang dinotasikan dengan:

$$c_i = \sum_{j=1}^n c_{ij}; j = 1, 2, \dots, n \tag{6}$$

Matriks Pembobot Spasial

Matriks pembobot spasial pada dasarnya merupakan matriks *contiguity* yang distandarisasi. Pada matriks *contiguity*, nilai 1 menunjukkan daerah yang bertetangga satu sama lain. Untuk dapat melihat seberapa besar pengaruh masing-masing tetangga terhadap suatu daerah dapat dihitung dari rasio antara nilai pada daerah tertentu dengan total nilai daerah tetangganya.

Hasilnya merupakan nilai pembobot (w_{ij}) untuk setiap kebertetangaan. Sesuai dengan persamaan:

$$w_{ij} = \frac{c_{ij}}{c_i} \tag{7}$$

Nilai w_{ij} menggambarkan pengaruh alami yang diberikan wilayah ke- j untuk wilayah ke- i . Sehingga matriks pembobot spasial dapat dikatakan juga sebagai matriks yang menggambarkan kekuatan interaksi antar lokasi. Bentuk umum matrik spasial (W) adalah:

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \ddots & \dots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \tag{8}$$

(Lee dan Wong, 2001)

Regresi Logistik Spasial

Menurut Xie *et al.* (2005) dalam Thaib (2008) menjelaskan bahwa regresi dapat dikatakan sebagai proses untuk ekstraksi koefisien-koefisien hubungan empirik dari observasi. Pendekatan regresi yang biasa digunakan adalah regresi linear, regresi log linear, dan regresi logistik. Variabel respon pada regresi logistik merupakan bilangan biner atau kategorik sedangkan variabel penjelas pada regresi logistik dapat berupa gabungan dari variabel kotinu maupun kategorik.

Analisis regresi logistik dapat diekspresikan variabel respon (Y) yang ditentukan oleh variabel penjelas (X) adalah:

$$Y = f(X) + v, 0 \leq Y \leq 1 \tag{9}$$

dengan v adalah nilai galatnya yang bernilai $[-f(X), 1-f(X)]$.

Jika Y memiliki hubungan spasial, dan dipengaruhi oleh X_1, X_2, \dots, X_k , maka model yang biasa dibentuk adalah:

$$y = X_i\beta + \rho W_i Y \quad (10)$$

Pemilihan Model Terbaik

Pada eliminasi langkah mundur model awal yang dibentuk memuat semua variabel penjelas yang dianggap peneliti berpengaruh terhadap variabel respon. Eliminasi dilakukan pada variabel yang memiliki nilai statistik uji Wald terbesar (Drapper dan Smith, 1998).

Tuberkulosis (TB)

Kock (1882) secara meyakinkan telah dapat memberikan bukti bahwa tuberkulosis adalah suatu penyakit infeksi yang disebabkan oleh bakteri yang diberi nama *Mycobacterium tuberculosis* berbentuk batang lurus tidak berspora dan juga tidak berkapsul. Bakteri ini berukuran lebar 0,3-0,6 mm dan panjang 1-4 mm. Orang yang pertama kali dapat membuktikan bahwa TB adalah suatu penyakit yang dapat ditularkan yaitu Villenim yang hidup pada tahun (1827-1894). Menurut Robbins (1957) TB adalah penyakit infeksi kronis yang disebabkan oleh *Mycobacterium tuberculosis*. Penyakit TB biasa terdapat pada paru-paru, tetapi mungkin juga pada organ lain seperti kelenjar getah bening (*nodus lymphaticus*) (Misnadiarly, 2006).

Terdapat berbagai faktor resiko yang bisa menyebabkan tertularnya penyakit TB, beberapa faktor resiko mempunyai pengaruh yang sangat kuat dan yang lainnya lemah, meskipun jelas berhubungan dengan penyakit TB tetapi biasanya hubungan ini masih bersifat tumpang tindih dengan faktor lainnya. Suatu gambaran penting dari faktor resiko adalah mereka saling menguatkan secara bersama-sama dan tidak hanya merupakan jumlah pertambahan dari masing-masing faktor resiko (Widaningrum, 2014). Namun yang dibahas dalam bab ini hanya variabel-variabel yang digunakan dalam penelitian terdiri dari 5 faktor sesuai dengan penelitian yang dilakukan Mahtumah (2011), yaitu sebagai berikut:

1) Luas Wilayah

Wilayah dapat diartikan sebagai suatu ruang pada permukaan bumi yang menunjuk pada tempat atau lokasi yang dilihat secara horizontal dan vertikal, termasuk yang ada di bawah permukaan bumi. Wilayah merupakan satu kesatuan ruang yang mempunyai tempat tertentu tanpa terlalu memperhatikan soal batas dan kondisinya. Satuan luas wilayah yang umumnya digunakan adalah Km^2 .

2) Kepadatan Penduduk

Kepadatan Penduduk adalah perbandingan dari jumlah penduduk dibagi dengan luas wilayahnya.

3) Jenis Kelamin

Jenis kelamin adalah pengenalan pada seseorang untuk menunjukkan dirinya, jenis kelamin hanya mencakup dua yaitu laki-laki dan perempuan. TB paru lebih banyak terjadi pada laki-laki dibandingkan dengan wanita karena laki-laki sebagian besar mempunyai kebiasaan merokok sehingga memudahkan terjangkitnya TB paru.

4) Tingkat Pendidikan

Tingkat pendidikan juga menjadi salah satu faktor resiko penularan penyakit TB. Tingkat pendidikan seseorang akan mempengaruhi terhadap pengetahuan seseorang diantaranya mengenai rumah yang memenuhi syarat kesehatan dan pengetahuan penyakit TB Paru, sehingga dengan pengetahuan yang cukup maka seseorang akan mencoba untuk mempunyai perilaku hidup bersih dan sehat.

5) Fasilitas Kesehatan

Fasilitas kesehatan adalah tempat yang digunakan untuk menyelenggarakan upaya kesehatan. Dalam melaksanakan kegiatan medis banyak sarana kesehatan yang dapat dijadikan sebagai tempat rujukan, baik pemerintah maupun swasta telah menyediakan berbagai bentuk dan jenis sarana kesehatan. Menurut Azwar (2008), adapun sarana-sarana kesehatan tersebut meliputi Balai pengobatan, Pusat Kesehatan Masyarakat, Rumah Sakit Umum, Rumah Sakit Khusus, Praktek Dokter, Praktek Bidan, Toko Obat, Apotek, Pedagang Besar Farmasi, Pabrik Obat dan Bahan Obat, Laboratorium, Sekolah dan Akademi Kesehatan, Balai Pelatihan Kesehatan dan Sarana Kesehatan Lainnya.

Metodologi Penelitian

Adapun variabel respon yang digunakan dalam penelitian ini adalah jumlah penderita penyakit TB (Y) di setiap kelurahan di Kota Samarinda pada Tahun 2013, dengan kategori sebagai berikut :

$Y = 0$: Jumlah penderita penyakit TB kurang dari rataannya

$Y = 1$: Jumlah penderita penyakit TB lebih dari atau sama dengan rataannya

variabel penjelasnya adalah variabel luas wilayah tiap kelurahan (X_1), kepadatan penduduk tiap kelurahan (X_2), rasio penduduk menurut jenis kelamin (X_3), banyaknya sekolah dasar tiap kelurahan (X_4), banyaknya puskesmas/puskesmas pembantu tiap kelurahan (X_5), banyaknya rumah sakit tiap kelurahan (X_6) dan pengaruh spasial (X_7).

Adapun langkah-langkah pengujian regresi logistik spasial ini adalah sebagai berikut:

1. Menentukan variabel pengaruh spasial pengaruh spasial yang dinyatakan sebagai variabel penjelas X_7 ditentukan dengan langkah-langkah sebagai berikut:

- a. Membuat peta persebaran jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda untuk melihat lokasi yang bersisian atau titik sudutnya bertemu dengan lokasi lain yang nantinya akan dibutuhkan untuk membuat matriks *contiguity*
 - b. Menentukan kedekatan antar kelurahan dengan membuat matriks *contiguity* antar kelurahan yang mengacu pada *queen contiguity* sesuai dengan persinggungan lokasi berdasarkan peta persebaran
 - c. Menghitung persentase tetangga yang menderita penyakit TB dari matriks *contiguity* menjadi matriks pembobot spasial yang kemudian dijadikan variabel penjelas spasial
 - d. Matriks pembobot spasial dari matriks *contiguity* dikalikan dengan variabel *Y*
 - e. Kemudian variabel penjelas baru dimasukkan ke dalam model regresi logistik baru yang dinamakan dengan “*X spasial*”
2. Melakukan pengujian regresi logistik spasial pada data jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 dengan tujuan untuk mencari probabilitas dengan langkah-langkah sebagai berikut:
 - a. Memodelkan regresi logistik dengan penambahan satu variabel penjelas baru sebagai variabel penjelas spasial
 - b. Melakukan pengujian *likelihood ratio* untuk melihat pengaruh spasial lainnya pada model regresi logistik spasial tersebut
 - c. Melakukan pengujian signifikansi parameter secara parsial dengan menggunakan uji Wald
 - d. Melakukan pengujian kesesuaian model
 - e. Menentukan model regresi logistik spasial terbaik dengan menggunakan metode eliminasi langkah mundur
 - f. Penafsiran atau interpretasi koefisien model regresi logistik spasial terbaik dengan menggunakan *Odds Ratio*.

Hasil dan Pembahasan
Analisis Statistika Deskriptif

Statistika deskriptif atau statistika deduktif adalah bagian dari ilmu statistik yang mempelajari cara pengumpulan dan penyajian data sehingga mudah dipahami. Statistika deskriptif hanya berhubungan dengan menguraikan atau memberikan keterangan-keterangan mengenai suatu data keadaan atau fenomena (Hasan, 2006).

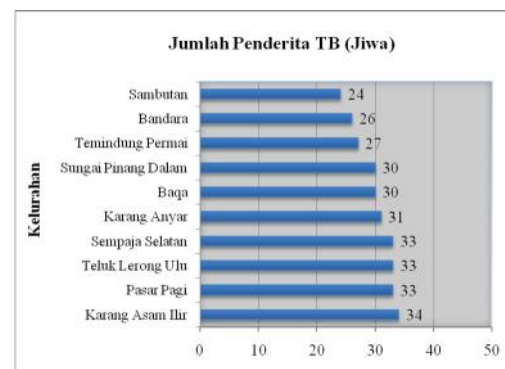
Analisis statistika deskriptif yang digunakan untuk jumlah penderita penyakit TB di setiap

kelurahan di Kota Samarinda pada Tahun 2013 dilakukan dengan membuat tabulasi penyajian karakteristik seperti nilai minimum, nilai maksimum, *mean*, standar deviasi untuk variabel *Y*. Karakteristik jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 dapat dilihat pada Tabel 1.

Tabel 1. Statistika Deskriptif Variabel *Y*

Y	N	Min	Maks	Total	Mean	Std. Dev
	53	0	34	732	14	10

Berdasarkan Tabel 4.1 dapat diketahui bahwa banyaknya data yang dianalisis adalah 53 data. Jumlah penderita penyakit TB yang tertinggi adalah sebanyak 34 jiwa, sedangkan yang terendah adalah sebanyak 0 jiwa. Rata-rata jumlah penderita penyakit TB yaitu 14 jiwa. Ukuran standar penyimpangan dari rata-rata penderita penyakit TB adalah sebesar 10 jiwa. Kemudian untuk total seluruh penderita penyakit TB adalah 732 jiwa. Gambaran tentang jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 yang terlihat pada Gambar 1.

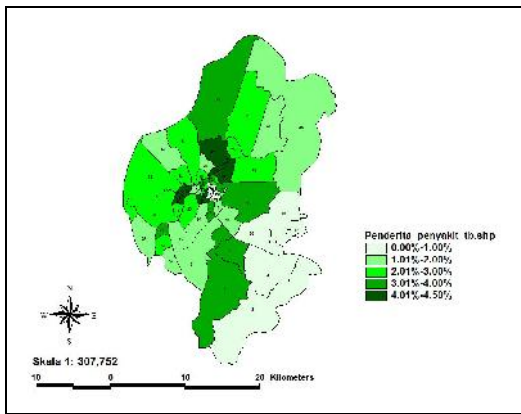


Gambar 1. Diagram Batang Sepuluh Kelurahan yang Memiliki Jumlah Penderita Penyakit TB Terbesar di Kota Samarinda pada Tahun 2013

Berdasarkan Gambar 1 dapat dinyatakan bahwa dari sepuluh kelurahan yang memiliki jumlah penderita penyakit TB terbesar di Kota Samarinda pada Tahun 2013, maka penderita penyakit TB di Kota Samarinda yang paling banyak adalah di Kelurahan Karang Asam Ilir yaitu sebanyak 34 jiwa. Sedangkan yang paling sedikit di Kelurahan Sambutan yaitu sebanyak 24 jiwa.

Peta Persebaran

Peta dibuat untuk melihat lokasi yang bersisian atau titik sudutnya bertemu dengan lokasi lain yang akan dibutuhkan untuk pembuatan matriks *contiguity*. Berikut adalah gambaran peta persebaran jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda:



Gambar 2. Peta Persebaran untuk Jumlah Penderita Penyakit TB di setiap Kelurahan di Kota Samarinda pada Tahun 2013

Warna pada setiap lokasi menunjukkan persentase dari jumlah penderita penyakit TB dibagi dengan total jumlah penderita penyakit TB pada setiap kelurahan. Daerah kelurahan yang ditandai dengan warna hijau pekat menunjukkan lokasi yang memiliki persentase penderita penyakit TB tertinggi yaitu sebesar 4,01% sampai dengan 4,50%, warna hijau tua menunjukkan lokasi penderita penyakit TB dengan persentase yaitu sebesar 3,01% sampai dengan 4,00%, warna hijau muda menunjukkan lokasi penderita penyakit TB dengan persentase yaitu sebesar 2,01% sampai dengan 3,00%, warna hijau muda kebiruan menunjukkan lokasi penderita penyakit TB dengan persentase yaitu sebesar 1,01% sampai dengan 2,00%, dan warna hijau muda keputihan menunjukkan lokasi yang memiliki persentase penyakit TB terendah yaitu sebesar 0,00% sampai dengan 1,00%. Sebagai contoh, kode 1 merupakan Kelurahan Simpang Pasir dengan nilai persentase jumlah penderita penyakit TB adalah sebesar 1,093 sehingga ditandai dengan warna hijau muda kebiruan yang menunjukkan bahwa Kelurahan Simpang Pasir memiliki persentase penderita penyakit TB yaitu sebesar 1,01% sampai dengan 2,00%.

Matriks Matriks Contiguity dan Matriks Pembobot Spasial

Matriks *contiguity* adalah matriks yang menggambarkan hubungan kedekatan antar daerah. Pemberian nilai 1 diberikan jika daerah-*i* bertetangga langsung dengan daerah-*j*, sedangkan nilai 0 diberikan jika daerah-*i* tidak bertetangga dengan daerah-*j*. Berikut ini matriks **C** dari 53 kelurahan di Kota Samarinda berdasarkan hubungan kedekatannya pada Gambar 2:

$$C = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \dots & \dots & 53 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix}$$

Matriks pembobot spasial dapat ditentukan dengan berbagai metode. Salah satu metode penentuan matriks pembobot spasial yang digunakan dalam penelitian ini adalah persentuhan sisi wilayah satu dengan wilayah yang lain (*Queen Contiguity*). Matriks pembobot w_{ij} yang berukuran $n \times n$, dimana lokasi yang bersisian atau titik sudutnya bertemu dengan lokasi yang menjadi perhatian diberi pembobotan $w_{ij} = 1$, sedangkan untuk lokasi lainnya adalah $w_{ij} = 0$. Jika dalam bentuk matriks maka akan menjadi matriks **W** adalah sebagai berikut:

$$W = \begin{bmatrix} 0 & 1/7 & 0 & 0 & 0 & 0 & 1/7 & 0 & 0 & 0 & \dots & \dots & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1/4 & 1/4 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & \dots & \dots & 0 \\ 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 1/4 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 1/4 & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 1/6 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix}$$

Pengaruh spasial yang dinyatakan sebagai variabel penjelas X_7 diperoleh dengan mengalikan matriks **W** terhadap jumlah penderita penyakit TB (Y) berdasarkan persamaan (10). Sehingga diperoleh variabel penjelas X_7 seperti pada Tabel 2 adalah sebagai berikut:

Tabel 2. Variabel Pengaruh Spasial (X_7)

Kelurahan	X_7
Simpang Pasir	9,714
Handil Bhakti	6,400
Bantuas	6,500
Bukuan	8,250
Rawa Makmur	4,250
Pelita	17,500
Selili	8,500
Sungai Dama	9,750
Sidodamai	10,000
Sidomulyo	12,333
Bugis	16,000
Pasar Pagi	11,750
Pelabuhan	15,800
⋮	⋮
⋮	⋮
Gunung Lingai	21,667

Langkah selanjutnya adalah pengujian regresi logistik pada data jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 adalah sebagai berikut:

Model Awal Regresi Logistik Spasial

Model awal regresi logistik spasial adalah:
 $g(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$

dimana:

- $g(X)$ = Persamaan bentuk logit
- β_0 = Konstanta
- β_j = Koefisien regresi dari variabel ke- j ; $j = 1, 2, \dots, 7$
- Y = Jumlah penderita penyakit TB di setiap kelurahan di kota Samarinda, dengan kategori adalah sebagai berikut:
 $Y = 0$: Jumlah penderita penyakit TB kurang dari rataannya
 $Y = 1$: Jumlah penderita penyakit TB lebih dari atau sama dengan rataannya
- X_1 = Luas wilayah tiap kelurahan
- X_2 = Kepadatan penduduk tiap kelurahan
- X_3 = Rasio penduduk menurut jenis kelamin
- X_4 = Banyaknya sekolah dasar tiap kelurahan
- X_5 = Banyaknya puskesmas/puskesmas pembantu tiap kelurahan
- X_6 = Banyaknya rumah sakit tiap kelurahan
- X_7 = Pengaruh spasial

Estimasi Parameter Model Regresi Logistik Spasial

Estimasi parameter model regresi logistik spasial seperti pada Tabel 3.

Tabel 3. Nilai Koefisien Model Regresi Logistik Spasial

Variabel	S
Konstanta	-4,099
X_1	-0,017
X_2	0,000
X_3	0,002
X_4	0,404
X_5	0,209
X_6	-1,143
X_7	0,167

Berdasarkan Tabel 3 diperoleh persamaan regresi logistik spasial dengan variabel respon (Y) adalah jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 dan variabel penjelasnya adalah luas wilayah tiap kelurahan (X_1), kepadatan penduduk tiap kelurahan (X_2), rasio penduduk menurut jenis kelamin (X_3), banyaknya sekolah dasar tiap kelurahan (X_4), banyaknya puskesmas/puskesmas pembantu tiap kelurahan (X_5), banyaknya rumah sakit tiap kelurahan (X_6) dan pengaruh spasial (X_7) adalah sebagai berikut:

$$\hat{g}(X) = -4,099 - 0,017 X_1 + 0,000 X_2 + 0,002 X_3 + 0,404 X_4 + 0,209 X_5 - 1,143 X_6 + 0,167 X_7$$

Pengujian Signifikansi Secara Simultan

Pengujian secara simultan juga merupakan uji model *chi-square* yang digunakan untuk menguji parameter hasil estimasi secara bersama.

Hipotesis

- H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$
 (Secara simultan variabel penjelas ke- j tidak ada yang berpengaruh terhadap variabel respon)
- H_1 : Minimal terdapat satu $\beta_j \neq 0$, $j = 1, 2, 3, 4, 5, 6, 7$ (Secara simultan variabel penjelas ke- j ada yang berpengaruh terhadap variabel respon)

Taraf Signifikansi

$$\alpha = 0,05$$

Statistik Uji

Berdasarkan perhitungan diperoleh nilai $-2 L_0 = 71,938$ dan $-2 L_k = 56,125$. Sehingga diperoleh nilai $L_0 = -35,969$ dan $L_k = -28,063$.

$$G = -2 (L_0 - L_k) = -2 ((-35,969) - (-28,063)) = 15,812$$

Keputusan

Dapat dilihat bahwa nilai $G = 15,812$ dimana nilai $G > t^2_{(r,k)} = 14,067$, maka H_0 ditolak

Kesimpulan

Secara simultan variabel penjelas ke- j ada yang berpengaruh terhadap variabel respon.

Pengujian Signifikansi Secara Parsial

Pengujian ini dilakukan untuk mengetahui pengaruh dari masing-masing variabel penjelas terhadap variabel respon. Berdasarkan perhitungan diperoleh nilai W untuk variabel penjelas ke- j dan diperoleh nilai $Z_{(0,05/2)} = 1,960$, dapat diambil keputusan seperti pada Tabel 4 sebagai berikut:

Tabel 4. Hasil Pengujian Signifikansi Parameter Secara Parsial

Variabel	W	p -value	Keputusan
Konstanta	5,921	0,015	H_0 ditolak
X_1	0,281	0,596	H_0 diterima
X_2	0,154	0,694	H_0 diterima
X_3	0,348	0,555	H_0 diterima
X_4	4,168	0,041	H_0 ditolak
X_5	0,205	0,650	H_0 diterima
X_6	1,944	0,163	H_0 diterima
X_7	4,420	0,036	H_0 ditolak

Berdasarkan pengujian signifikansi parameter secara parsial dengan menggunakan uji wald, dapat disimpulkan bahwa ada pengaruh variabel banyaknya sekolah dasar tiap kelurahan dan variabel pengaruh spasial terhadap jumlah penderita penyakit TB.

Pengujian Kesesuaian Model (Goodness of Fit)

Uji Hosmer-Lemeshow berfungsi untuk menilai kesesuaian model regresi logistik spasial dengan cara membandingkan hasil pengamatan dengan nilai dugaan.

Hipotesis

- H_0 : Model sesuai (tidak ada perbedaan antara hasil pengamatan dengan nilai dugaan)
- H_1 : Model tidak sesuai (ada perbedaan antara hasil pengamatan dengan nilai dugaan)

Statistik Uji

Berdasarkan tabel *chi-square*, diperoleh nilai $t^2_{(0,05,6)} = 12,592$. Diketahui nilai $\hat{C} = 7,899 < t^2_{(0,05,6)} = 12,592$ atau nilai p -value = 0,443 > α (0,05), maka H_0 diterima

Kesimpulan

Model sesuai, sehingga tidak ada perbedaan antara hasil pengamatan dengan nilai dugaan.

Pemilihan Model Regresi Logistik Terbaik

Berdasarkan perhitungan dengan menggunakan eliminasi langkah mundur, didapatkan model regresi logistik spasial terbaik adalah model 6 dengan nilai AIC terkecil yaitu variabel respon $Y = f(X)$ adalah jumlah penderita penyakit TB dengan variabel penjelas yaitu variabel banyaknya sekolah dasar tiap kelurahan (X_4) dan variabel pengaruh spasial (X_7) adalah sebagai berikut:

$$f(X) = \frac{e^{-3,803+0,336X_4+0,124X_7}}{1 + e^{-3,803+0,336X_4+0,124X_7}}$$

Sehingga di dapat model regresi logistik spasial terbaik sebagai berikut:

$$g(X) = -3,803 + 0,336 X_4 + 0,124 X_7 \quad (11)$$

Interpretasi persamaan (11) adalah:

1. Jika diperoleh hasil tetap tanpa adanya pengaruh banyaknya sekolah dasar tiap kelurahan dan pengaruh spasial tidak berpengaruh maka besarnya jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 yaitu sebesar 4 jiwa.
2. Jika ada penambahan tiga unit banyaknya sekolah dasar tiap kelurahan maka akan menaikkan jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 yaitu sebesar 1 jiwa.
3. Jika ada penambahan sepuluh lag spasial maka akan menaikkan jumlah penderita penyakit TB di setiap kelurahan di Kota Samarinda pada Tahun 2013 yaitu sebesar 1 jiwa.

Penafsiran Koefisien Model Regresi Logistik Spasial

Setelah didapatkan model regresi logistik spasial terbaik pada persamaan (11), maka dilakukan penafsiran koefisien dari model regresi dengan melihat nilai *Odds Ratio* (\mathbb{E}). Berdasarkan nilai *Odds Ratio* untuk variabel penjelas yang signifikan dalam model adalah variabel banyaknya sekolah dasar tiap kelurahan (X_4) dan variabel pengaruh spasial (X_7). Interpretasi dari nilai *Odds Ratio* dapat dilihat dari Tabel 5.

Tabel 5. Penafsiran Koefisien Model Regresi Logistik Spasial

Variabel	\mathbb{E} (<i>Odds Ratio</i>)
X_4	1,400
X_7	1,132

1. Untuk variabel banyaknya sekolah dasar tiap kelurahan

$$\mathbb{E} = e^{(s_0+s_1)-s_0} = 1,400$$

Jadi, daerah yang memiliki kenaikan satu unit banyaknya sekolah dasar tiap kelurahan akan menyebabkan kemungkinan suatu daerah tersebut memiliki persentase penderita penyakit TB lebih besar dari nilai rataannya meningkat sebesar 1,400 kali.

2. Untuk variabel pengaruh spasial

$$\mathbb{E} = e^{(s_0+s_1)-s_0} = 1,132$$

Jadi, daerah yang memiliki kenaikan satu lag spasial akan menyebabkan kemungkinan suatu daerah memiliki persentase penderita penyakit TB lebih besar dari nilai rataannya meningkat sebesar 1,132 kali.

Nagelkerke R²

Berikut adalah nilai nagelkerke R² seperti pada Tabel 6.

Step	Nagelkerke R ²
1	0,347

Dari Tabel 6 diperoleh nilai Nagelkerke R² adalah 0,347. Hal ini menunjukkan bahwa variabel banyaknya sekolah dasar tiap kelurahan (X_4) dan pengaruh spasial (X_7) mampu menjelaskan variabel jumlah penderita penyakit TB di setiap Kelurahan di Kota Samarinda pada Tahun 2013 yaitu sebesar 34,7% dan sisanya sebesar 65,3% dijelaskan oleh variabel-variabel lain di luar kedua variabel yang terdapat dalam model.

Kesimpulan

Berdasarkan hasil analisis dan pembahasan di atas diperoleh kesimpulan sebagai berikut:

1. Model regresi logistik spasial yang paling tepat untuk menduga penyebaran penyakit TB adalah sebagai berikut:

$$g(X) = -3,803 + 0,336 X_4 + 0,124 X_7$$
 Berdasarkan model regresi logistik spasial terbaik di atas dapat disimpulkan bahwa banyaknya sekolah dasar tiap kelurahan dan pengaruh spasial dapat mempengaruhi penyebaran penyakit TB di setiap Kelurahan di Kota Samarinda pada Tahun 2013.
2. Ada pengaruh dari faktor spasial terhadap penyebaran penyakit TB di setiap Kelurahan Kota Samarinda pada Tahun 2013.

Daftar Pustaka

- Agresti, A. 2002. *Categorical Data Analysis*. New York: John Wiley & Sons, Inc.
- Azwar, Azrul. 1996. *Pengantar Administrasi Kesehatan Edisi Ketiga*. Jakarta: Binarupa Aksara.
- Draper, N.R & H., Smith. 1998. *Applied Regression Analysis, Third Edition*. Canada: John Wiley & Sons.
- Hasan, Iqbal. 2006. *Analisis Data Penelitian dengan Statistik*. Jakarta: Bumi Aksara.
- Hosmer, D. W., dan S., Lemeshow. 2000. *Applied Logistic Regression*. New York: John Wiley and Sons.
- Lee J, dan DWS. Wong. 2001. *Statistical Analysis ArcView GIS*. New York: John Wiley & Sons, Inc.
- Mahtumah, U. 2011. Penerapan Model Regresi Logistik Spasial (Studi Kasus: Penyebaran

Penyakit Demam Berdarah *Dengue* di Kota Bogor Tahun 2008). Skripsi. Bogor: Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.

Misnadiarly. 2006. *Penyakit Infeksi TB Paru dan Ekstra Paru: Mengenal, Mencegah, Menanggulangi TBC Paru, Ekstra Paru, Anak, dan pada Kehamilan*. Jakarta: Pustaka Populer Obor.

Thaib, Z. 2008. *Pemodelan Regresi Logistik Spasial Dengan Pendekatan Matriks Contiguity*. Skripsi. Bogor: Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.

Ward, MD., dan KS Gleditsch. 2008. *Spatial Regression Models*. Los Angeles: Sage Publications, Inc.

Widaningrum, Christina. 2014. *Pedoman Nasional Pengendalian Tuberkulosis*. Kementerian Kesehatan RI: Jakarta.

Xie. Chengin, Bo Huang, Christophe Claramunt dan Magesh Chandramouli. 2005. *Spatial Logistic Regression and GIS to Model Rural-Urban Land Conversion*. http://www.civ.utoronto.ca/sect/traeng/ilute/processus2005/PaperSession/Paper10_Xie-et-al_SpatialLogisticRegression_CD.pdf Diakses pada tanggal 21 Maret 2015

Yatani, Koji. 2014. *Logistic Regression*. <http://yatani.jp/teaching/doku.php?id=hcistats:logisticregression> Diakses pada tanggal 16 Agustus 2015

