

**PEMODELAN REGRESI SPASIAL
DENGAN PEMBOBOT *CUSTOMIZE* BERDASARKAN
ANALISIS KLASSTER *AVERAGE LINKAGE* PADA
KASUS TUBERKULOSIS DI SULAWESI SELATAN**

SKRIPSI



ZAKIAH FITRI

H051171014

**PROGRAM STUDI STATISIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

UNIVERSITAS HASANUDDIN

MAKASSAR

DESEMBER 2023

**PEMODELAN REGRESI SPASIAL
DENGAN PEMBOBOT *CUSTOMIZE* BERDASARKAN
ANALISIS KLASTER *AVERAGE LINKAGE* PADA
KASUS TUBERKULOSIS DI SULAWESI SELATAN**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains
pada Program Studi Statistika Departemen Statistika Fakultas Matematika
dan Ilmu Pengetahuan Alam Universitas Hasanuddin**

ZAKIAH FITRI

H051171014

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

MAKASSAR

DESEMBER 2023

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**Pemodelan Regresi Spasial
dengan Pembobot *Customize* Berdasarkan Analisis Klaster *Average Linkage*
pada Kasus Tuberkulosis di Sulawesi Selatan**

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 14 Desember 2023



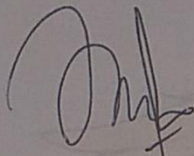
ZAKIAH FITRI

NIM. H051171014

PEMODELAN REGRESI SPASIAL
DENGAN PEMBOBOT *CUSTOMIZE* BERDASARKAN
ANALISIS KLASTER *AVERAGE LINKAGE* PADA KASUS
TUBERKULOSIS DI SULAWESI SELATAN

Disetujui Oleh:

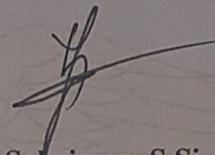
Pembimbing Utama



Anisa, S.Si., M.Si.

NIP. 19730227 199802 2 001

Pembimbing Pendamping



Sitti Sahriman, S.Si., M.Si.

NIP. 19881018 201504 2 002

Ketua Program Studi



Dr. Anna Islamiyati, S.Si., M.Si.

NIP. 19770806 200501 2 002

Pada 14 Desember 2023

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Zakiah Fitri
NIM : H051171702
Program Studi : Statistika
Judul Skripsi : Pemodelan Regresi Spasial dengan Pembobot
Customize Berdasarkan Analisis Klaster *Average Linkage* pada Kasus Tuberkulosis di Sulawesi Selatan

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Anisa, S.Si., M.Si. (.....)
2. Sekretaris : Sitti Sahriman, S.Si., M.Si. (.....)
3. Anggota : Drs. Raupong, M.Si (.....)
4. Anggota : Siswanto, S.Si., M.Si (.....)

Ditetapkan di : Makassar

Tanggal : 14 Desember 2023

KATA PENGANTAR

Alhamdulillah Robbil'alamin, Segala Puji senantiasa disenandungkan untuk menyampaikan rasa syukur penulis kepada **Allah Subhanahu Wata'ala** atas segala limpahan rahmat, nikmat, dan hidayah yang diberikan kepada penulis sehingga dapat menyelesaikan penulisan skripsi dengan judul "**Pemodelan Regresi Spasial dengan Pembobot *Customize* Berdasarkan Analisis Klaster *Average Linkage* pada Kasus Tuberkulosis di Sulawesi Selatan**" sebagai salah satu syarat untuk mendapat gelar Sarjana Sains pada Program Studi Statistika Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin Makassar.

Penulis menyadari dalam penulisan skripsi ini tidak lepas dari hambatan dan masalah namun dapat terselesaikan berkat dukungan dan bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan rasa terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Mama tersayang **Marhawati**, Bapak tercinta **Amir Mahmud**, dan Nenek tersayang **Indo Hawaedah** yang telah membesarkan dan mendidik dengan penuh kesabaran, memberikan cinta dan limpahan kasih sayang, dukungan dan doa yang tulus tanpa henti kepada penulis. Rasa terima kasih juga kepada kakak-kakak tercinta **Zulfadli**, **Ayusnawati** dan **Masnaeni** serta adik-adik tersayang **Syifa Ur Rahmah** dan **Syifa Ur Rahmi** atas doa, dukungan, semangat, dan bantuannya kepada penulis. Terima kasih pula kepada keluarga besar penulis yang selalu memberikan dukungan dalam segala hal. Tak lupa pula penulis ucapkan terima kasih yang sebesar-besarnya kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.** selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf Departemen Statistika yang telah membekali ilmu dan kemudahan-kemudahan kepada Penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.

4. **Ibu Anisa, S.Si., M.Si.** selaku Pembimbing utama dan Penasehat akademik yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan ditengah kesibukan beliau serta menjadi tempat berkeluh kesah untuk penulis.
5. **Ibu Sitti Sahriman, S.Si., M.Si.** selaku Pembimbing Pertama atas saran, nasehat, motivasi, ilmu dan dukungan yang telah diberikan kepada penulis selama menjadi mahasiswa dan meluangkan waktu untuk membimbing dan memberikan masukan dalam penulisan skripsi ini.
6. **Bapak Drs.Raupong, M.Si.** dan **Bapak Siswanto, S.Si., M.Si.** selaku Tim Penguji atas saran dan kritikan yang membangun dalam penyempurnaan penyusunan skripsi ini serta waktu yang telah diberikan kepada Penulis.
7. Sahabat-sahabat kesayangan **Nahdiatul Rabia Tusadiah, Nirmalasari** dan **Meliana Muliadi** yang telah menjadi saudara terbaik penulis, memberikan bantuan, motivasi dan selalu ada untuk penulis selama hampir 10 tahun ini.
8. Sahabat terbaik, **Nurhidayatullah, Sri Ulan S.S Abdullah** dan **Kak Muhammad Fadil** yang selalu membersamai penulis, memberikan dorongan dan dukungan serta tidak bosan mendengarkan segala keluh kesah penulis dalam setiap proses hingga skripsi ini terselesaikan.
9. Sahabat **17IMON, A. Restina, Fauziah, Astri, Alm.Dandi** dan semuanya yang tidak bisa penulis sebutkan satu persatu. Terima kasih karena tetap menjadi andalan penulis disegala situasi dan tetap menjadi saudara terbaik penulis.
10. Sahabat seperjuangan, **Aqilah, Zulfa, Siti Ihza, Aprilia, Riska, Kamalia, Risna, Izza** dan **Nahli** yang telah menjadi teman terbaik semasa perkuliahan, memberikan bantuan, dorongan, semangat, dan motivasi dalam setiap keadaan.
11. Teman-teman **Statistika 2017**, terkhusus **Iklil, Uti, Mirna, Ata, Alya, Haura, Fakhriyyah** dan **Fitri**. Terima kasih atas kebersamaan, suka, dan duka selama menjalani pendidikan di Departemen Statistika. Penulis senang mengenal kalian semua.
12. Keluarga besar **DISKRIT 2017** dan **MIPA 2017** terima kasih telah memberikan pelajaran yang berharga dan arti kebersamaan selama ini kepada

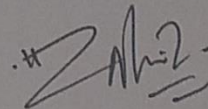
penulis. Pengalaman yang berharga telah penulis dapatkan dari teman-teman selama berproses.

13. **Keluarga Mahasiswa FMIPA Unhas** terkhusus anggota keluarga **Himatika FMIPA Unhas** dan **Himastat FMIPA Unhas**, terima kasih atas ilmu yang mungkin tidak bisa didapatkan pada proses perkuliahan dan telah menjadi keluarga selama penulis kuliah di Universitas Hasanuddin.
14. **IMPS Koperti UNHAS** khususnya **Kak Wahyu, Suardi, Indah** dan **Citra** yang telah memberikan banyak kenangan selama ber-IMPS
15. **Mujahidah 2018 UKM LDK MPM UNHAS**, *uktifillahku* di **BPA** dan **Pengurus Mushallah Istiqamah** yang telah banyak memberikan ilmu yang bermanfaat kepada penulis.
16. Semua pihak yang telah banyak berpartisipasi, baik secara langsung maupun tidak langsung yang tak sempat penulis sebutkan satu per satu. Terima kasih untuk segala bantuan dan dukungannya.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga skripsi ini dapat memberikan manfaat bagi pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar, 14 Desember 2023



Zakiah Fitri

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama : Zakiah Fitri
NIM : H051171014
Program Studi : Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty- Free Right*)** atas tugas akhir saya yang berjudul:

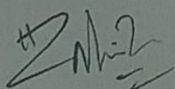
**“Pemodelan Regresi Spasial
dengan Pembobot *Customize* Berdasarkan Analisis Kluster *Average Linkage*
pada Kasus Tuberkulosis di Sulawesi Selatan”**

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 14 Desember 2023.

Yang menyatakan



(Zakiah Fitri)

ABSTRAK

Regresi spasial merupakan analisis regresi yang dapat menjelaskan hubungan antara variabel prediktor dan variabel respon dengan memberikan efek spasial pada lokasi yang menjadi pusat pengamatan. Dalam regresi spasial, diperlukan suatu matriks pembobot. Pemilihan pembobot spasial dikembangkan oleh beberapa peneliti dengan pemilihan pembobot yang didasarkan pada karakteristik masalah yang sedang diteliti atau biasa disebut dengan pembobot *customize*. Salah satu cara dalam mendapatkan pembobot *customize* yaitu dengan mengklasterkan daerah-daerah yang menjadi pengamatan. Analisis kluster yang dapat digunakan yaitu analisis kluster *average linkage*. Penelitian ini melakukan pemodelan regresi spasial pada kasus tuberkulosis di Sulawesi Selatan tahun 2021 dengan pembobot *customize* berdasarkan hasil analisis kluster metode *average linkage* pada faktor-faktor yang diduga memengaruhi tuberkulosis. Analisis kluster metode *average linkage* menghasilkan jumlah kluster optimum yaitu sebanyak lima kluster. Pemodelan regresi spasial dengan pembobot berdasarkan hasil analisis kluster diperoleh bahwa model terbaik yaitu model SARMA dengan nilai R^2 sebesar 90.85% dan nilai AIC sebesar 375.238 dengan faktor-faktor yang berpengaruh terhadap jumlah kasus tuberkulosis yaitu persentase penduduk miskin, persentase usia produktif dan ratio puskesmas.

Kata kunci: Pembobot *customize*, Analisis Kluster, *Average Linkage*, Regresi Spasial, SARMA, Tuberkulosis

ABSTRACT

Spatial regression is a regression analysis that can explain the relationship between predictor variables and response variables by providing spatial effects on the location that is the center of observation. In spatial regression, a weight matrix is required. The selection of spatial weights was developed by several researchers with the selection of weights based on the characteristics of the problem being studied or commonly referred to as customized weights. One way to get customized weights is by clustering the areas of observation. Cluster analysis that can be used is average linkage cluster analysis. This study conducted spatial regression modeling of tuberculosis cases in South Sulawesi in 2021 with customized weights based on the results of cluster analysis of the average linkage method on factors that are thought to affect tuberculosis. The cluster analysis of the average linkage method resulted in an optimum number of clusters, namely five clusters. Spatial regression modeling with weighting based on the results of cluster analysis obtained that the best model is the SARMA model with an R^2 value of 90.85% and an AIC value of 375.238 with factors that affect the number of tuberculosis cases, namely the percentage of poor people, the percentage of productive age and the ratio of health centers.

Keywords: *Customized weights, Cluster Analysis, Average Linkage, Spatial Regression, SARMA, Tuberculosis*

DAFTAR ISI

HALAMAN JUDUL.....	ii
LEMBAR PERNYATAAN KEOTENTIKAN	ii
LEMBAR PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR	v
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	viii
ABSTRAK.....	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Analisis Klaster	5
2.2 Analisis Klaster <i>Average Linkage</i>	6
2.3 Penentuan Jumlah Klaster Optimum	7
2.4 Regresi Linier Berganda.....	7
2.5 Analisis Spasial	8
2.6 Model Regresi Spasial.....	9
2.7 Pembobot Spasial	10
2.8 Dependensi Spasial	11
2.9 Uji Asumsi Residual.....	13
2.10 Uji Signifikansi Parameter	14

2.11	Pemilihan Model Terbaik	15
2.12	Tuberkulosis	15
BAB III METODOLOGI PENELITIAN		18
3.1	Sumber Data	18
3.2	Identifikasi Variabel	18
3.3	Metode Analisis	19
BAB IV HASIL DAN PEMBAHASAN		21
4.1	Deskripsi Data	21
4.2	Pembobot <i>Customize</i> Berdasarkan Analisis Kluster <i>Average Linkage</i>	22
4.3	Uji Indeks Moran	28
4.4	Uji <i>Lagrange Multiplier</i>	30
4.5	Model <i>Spatial Autoregressive Model</i>	34
4.6	Model <i>Spatial Error Model</i>	35
4.7	Model <i>Spatial Autoregressive Moving Average</i>	36
4.8	Pengujian Asumsi Residual	38
4.9	Pemilihan Metode Terbaik	39
4.10	Interpretasi Model	40
BAB V KESIMPULAN DAN SARAN		41
5.1	Kesimpulan	41
5.2	Saran	42
DAFTAR PUSTAKA		43
LAMPIRAN.....		46

DAFTAR GAMBAR

Gambar 4.1 Peta Sebaran Kasus Tuberkulosis di Provinsi Sulawesi Selatan.....	22
Gambar 4.2 Jumlah Klaster Optimum.....	23
Gambar 4.3 Dendogram dengan jumlah klaster $k = 5$	26
Gambar 4.4 Hasil Analisis Klaster pada Peta.....	27

DAFTAR TABEL

Tabel 3.1 Definisi Operasional Variabel.....	18
Tabel 4.1 Analisis Deskriptif	21
Tabel 4.2 Nilai VIF.....	23
Tabel 4.3 Pengujian Parameter Model SAR.....	34
Tabel 4.4 Pendugaan Parameter Model SAR	34
Tabel 4.5 Pengujian Parameter Model SEM.....	35
Tabel 4.6 Pendugaan Parameter Model SEM.....	35
Tabel 4.7 Pengujian Parameter Model SARMA.....	37
Tabel 4.8 Pendugaan Parameter Model SARMA.....	37
Tabel 4.9 Hasil Uji Distribusi Normal.....	38
Tabel 4.10 Hasil Uji Asumsi Identik.....	39
Tabel 4.11 Hasil Uji Independen.....	39
Tabel 4.12 Pemilihan Model Terbaik.....	39

DAFTAR LAMPIRAN

Lampiran 1. Data Tuberkulosis di Provinsi Sulawesi Selatan Tahun 2021.....	46
Lampiran 2. Jarak <i>Euclidean</i>	47
Lampiran 3. Matriks Pembobot <i>customize</i>	48
Lampiran 4. Nilai ϵ	49
Lampiran 5. Uji Indeks Moran.....	50
Lampiran 6. Uji <i>Lagrange Multiplier</i>	51
Lampiran 7. Estimasi Parameter dengan Metode Kuadrat Terkecil.....	52
Lampiran 8. Estimasi Parameter Model SAR.....	53
Lampiran 9. Estimasi Parameter Model SEM.....	54
Lampiran 10. Estimasi Parameter Model SARMA.....	55
Lampiran 11. Nilai AIC dan R^2	56
Lampiran 12. Pengujian Asumsi Residual.....	57

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tuberkulosis atau biasa disingkat TBC atau TB merupakan penyakit menular yang disebabkan oleh bakteri *Mycobacterium tuberculosis* (Kementerian Kesehatan RI, 2019). Tuberkulosis menjadi salah satu dari sepuluh penyakit penyebab kematian terbanyak di dunia (WHO, 2020). Kementerian Kesehatan Republik Indonesia menargetkan kasus penyakit TB turun 80% pada tahun 2030 dan tereliminasi pada tahun 2050. Upaya pemerintah tersebut salah satunya dengan memfokuskan penemuan penyakit TB di daerah padat penduduk, salah satunya Provinsi Sulawesi Selatan (Kementerian Kesehatan RI, 2019). Dalam rangka menyukseskan upaya pemerintah untuk mengeliminasi penyakit TB, maka perlu diketahui faktor-faktor yang memengaruhi terjadinya penyakit TB.

Beberapa penelitian telah dilakukan guna mengkaji faktor-faktor penyebab terjadinya TB paru. Rizky, dkk (2020) mengatakan bahwa kepadatan penduduk dan jumlah rumah sehat memengaruhi jumlah penderita TB. Selanjutnya, Pangaribuan (2020) mengatakan bahwa umur, jenis kelamin, daerah tempat tinggal, pendidikan dan kawasan merupakan faktor terjadinya TB. Penelitian lain oleh Wulandari (2012) mengatakan bahwa terdapat hubungan antara kepadatan penduduk, jumlah sarana kesehatan, dan jumlah tenaga kesehatan dengan jumlah kasus TB. Salah satu metode yang umum digunakan dalam menganalisis faktor-faktor penyebab terjadinya kasus TB yaitu regresi linear berganda.

Regresi linear berganda merupakan analisis regresi yang dapat menjelaskan hubungan antara variabel respon dengan lebih dari satu variabel prediktor (Maziyah, 2016). Metode yang dapat digunakan dalam mengestimasi parameter model regresi berganda adalah *ordinary least square* (OLS) atau metode kuadrat terkecil (Basri, *Pemodelan Regresi Berganda untuk Data dalam Studi kecerdasan Emosional*, 2018). Terdapat asumsi klasik yang harus dipenuhi dalam menganalisis data untuk mendapatkan model regresi yang baik yaitu distribusi datanya normal, tidak terjadi multikolinearitas, heteroskedasitas dan

autokorelasi. Jika terdapat pelanggaran asumsi non-heteroskedastisitas atau non-autokorelasi maka kemungkinan terdapat efek spasial sehingga OLS kurang baik untuk digunakan (Pertiwi dkk, 2020). Berdasarkan hal tersebut, maka analisis yang cocok digunakan adalah analisis regresi spasial.

Regresi spasial merupakan analisis yang mempertimbangkan hubungan antara suatu variabel dengan variabel lain dengan memberikan efek spasial pada lokasi yang menjadi pusat pengamatan (Kurnianto dkk, 2021). Anselin (1988) menyatakan bahwa terdapat dua efek spasial pada model regresi spasial yaitu *spatial dependence* (ketergantungan spasial) dan *spatial heterogeneity* (keragaman spasial). Ketergantungan spasial menunjukkan bahwa pengamatan pada sebuah lokasi dipengaruhi oleh pengamatan di lokasi yang lain. Pendekatan yang dapat dilakukan untuk menampilkan hubungan spasial antar lokasi yaitu dengan konsep persinggungan (*contiguity*). Metode yang digunakan untuk mendefinisikan persinggungan dalam membentuk matriks pembobot spasial berbasis geografis salah satunya adalah *queen contiguity*.

Beberapa peneliti mengembangkan konsep pemilihan matriks pembobot yang tidak hanya berdasarkan pada persinggungan secara geografis tetapi juga berdasar pada karakteristik masalah yang sedang diteliti. Hal tersebut dapat berupa aksesibilitas maupun kesamaan karakteristik sosial ekonomi (Sulistiyono, 2016). Pembobot dengan konsep tersebut biasa disebut dengan *customized contiguity*. Penelitian sebelumnya oleh Sulistiyono (2016) menggunakan pembobot *customized contiguity* dengan pendekatan lapangan usaha utama pada setiap desa. Penelitian lain oleh Fajri, dkk (2023) menggunakan pembobot *customized contiguity* berdasarkan pembagian wilayah bakorwil. Dengan demikian maka penelitian ini menggunakan pembobot *customized contiguity* berdasarkan faktor-faktor yang diduga memengaruhi TB yang dapat ditentukan menggunakan hasil dari analisis klaster.

Analisis klaster merupakan salah satu teknik statistik yang bertujuan untuk mengelompokkan suatu objek sehingga dalam suatu kelompok memiliki ciri yang relatif sama atau homogen sedangkan antar kelompok mempunyai ciri yang berbeda atau heterogen. Proses pembentukan klaster dapat dilakukan dengan dua cara yaitu metode hirarki dan non-hirarki (Damayanti & Wijayanto,

2021). Pembobot *customized contiguity* berdasarkan hasil analisis kluster dibentuk dengan memberikan nilai 1 pada daerah yang berada dalam kluster yang sama dan nilai 0 untuk daerah lainnya.

Rizky dkk (2020) melakukan pemodelan regresi spasial dengan pembobot *customized contiguity* dari hasil analisis kluster non-hirarki metode *K-means*. Penelitian tersebut membandingkan pembobot *queen contiguity* dengan pembobot dari hasil analisis kluster *K-means* dalam pemodelan regresi spasial pada kasus TB. Hasil yang diperoleh menunjukkan bahwa pembobot *queen contiguity* tidak signifikan sehingga pembobot yang digunakan dalam model regresi spasial adalah pembobot dari hasil analisis kluster metode *K-means*.

Adapun pada penelitian ini menggunakan pembobot *customized contiguity* berdasarkan hasil analisis kluster yang ditentukan dengan metode *average linkage*. Penentuan pembobot tersebut menggunakan data faktor-faktor yang diduga memengaruhi TB untuk pemodelan regresi spasial pada kasus TB di Sulawesi Selatan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka rumusan masalah yang akan dikaji pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil analisis kluster *average linkage* berdasarkan faktor yang diduga memengaruhi terjadinya tuberkulosis di Sulawesi Selatan?
2. Bagaimana dugaan model regresi spasial dengan pembobot dari hasil analisis kluster *average linkage* pada kasus tuberkulosis di Sulawesi Selatan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan dari penelitian ini antara lain sebagai berikut:

1. Mendapatkan hasil analisis kluster *average linkage* berdasarkan faktor yang diduga memengaruhi terjadinya tuberkulosis di Sulawesi Selatan.
2. Mendapatkan dugaan model regresi spasial dengan pembobot dari hasil analisis kluster *average linkage* pada kasus tuberkulosis di Sulawesi Selatan.

1.4 Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu:

1. Data yang digunakan memiliki efek dependensi spasial untuk pemodelan regresi spasial pada kasus tuberkulosis di Sulawesi Selatan tahun 2021.
2. Matriks pembobot yang digunakan adalah matriks pembobot yang distandarisasi

1.5 Manfaat Penelitian

Manfaat yang diharapkan dapat diambil dari penelitian ini bagi penulis yaitu sebagai bentuk penerapan dan pengembangan pada bidang statistika yang berkaitan dengan pemodelan regresi spasial menggunakan pembobot dari hasil analisis kluster *average linkage*. Adapun manfaat yang diharapkan bagi pembaca yaitu menambah wawasan dan pengetahuan mengenai hal-hal yang berkaitan dengan penelitian ini serta diharapkan dapat dijadikan sebagai salah satu acuan dalam pengambilan kebijakan untuk menangani kasus tuberkulosis di Sulawesi Selatan oleh pemerintah atau pihak terkait.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Klaster

Analisis klaster merupakan teknik analisa data untuk mengelompokkan suatu individu atau objek menjadi beberapa kelompok yang memiliki sifat yang relatif berbeda antar kelompok, sedangkan individu atau objek dalam satu kelompok memiliki sifat yang relatif sama atau homogen (Talakua dkk, 2017). Analisis klaster terdiri dari analisis klaster hirarki dan non-hirarki. Perbedaan antara analisis klaster hirarki dan non-hirarki yaitu, pada metode hirarki jumlah klaster tidak ditentukan sebelumnya, sedangkan pada metode non-hirarki jumlah klaster ditentukan sebelum melakukan pengelompokkan (Damayanti & Wijayanto, 2021).

Asumsi yang harus dipenuhi dalam analisis klaster yaitu data merupakan populasi atau sampel yang harus representatif. Selain itu, data yang digunakan pada analisis klaster sebaiknya tidak mengalami multikolinearitas (Fithriyyah, 2017). Multikolinearitas adalah adanya hubungan linear sempurna atau pasti diantara beberapa atau semua variabel. Salah satu cara untuk mengetahui terjadinya multikolinearitas adalah dengan menghitung nilai *variance inflation factor* (VIF) dengan persamaan berikut:

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

dengan,

VIF_j : nilai *variance inflation factor* variabel ke- j

R_j^2 : koefisien determinasi variabel ke- j

Apabila nilai $(VIF_j) \geq 10$ maka dapat diindikasikan bahwa terjadi multikolinearitas (Silvi, 2018).

Berdasarkan tujuan dari analisis klaster untuk mengelompokkan objek yang mirip dalam klaster yang sama, maka diperlukan sebuah ukuran untuk mengetahui seberapa mirip objek-objek yang akan dikelompokkan. Ukuran kemiripan tersebut dapat ditentukan dari kedekatan jarak. Ukuran jarak antar objek menunjukkan bahwa semakin kecil jarak antar objek maka akan semakin

mirip objek tersebut. Salah satu ukuran jarak yang dapat digunakan adalah jarak *Euclidean* (Prabowo dkk, 2020). Jarak *Euclidean* dirumuskan sebagai berikut:

$$d(i, j) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (2.2)$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, n; k = 1, 2, \dots, p$$

dengan,

$d(i, j)$: jarak antar objek ke- i dengan objek ke- j

x_{ik} : nilai dari objek ke- i pada variabel ke- k

x_{jk} : nilai dari objek ke- j pada variabel ke- k

n : jumlah objek

p : jumlah variabel kluster

2.2 Analisis Kluster *Average Linkage*

Average linkage adalah metode pengklasteran yang didasarkan pada jarak rata-rata antar objeknya. Jarak antara dua kluster dapat diasumsikan sebagai jarak rata-rata dari semua anggota dalam sebuah kluster dengan semua anggota dalam kluster lainnya. Tujuan dari metode ini adalah untuk meminimumkan rata-rata jarak semua pasangan observasi dan dua kluster yang digabungkan. Pengklasteran dengan metode *average linkage* dimulai dengan mendefinisikan matriks D untuk memperoleh objek-objek terdekat, misalnya objek U dan V. Kedua objek tersebut kemudian digabung menjadi kluster (UV). Adapun jarak antara kluster (UV) dan kluster lainnya (W) adalah sebagai berikut (Johnson & Wichern, 2007):

$$d_{UV(W)} = \frac{\sum_{i=1}^n \sum_{k=1}^n d_{ik}}{N_{(UV)}N_W} \quad (2.3)$$

dengan,

d_{ik} : jarak antar objek i pada kluster (UV) dan objek k pada kluster W

$n_{(UV)}$: jumlah objek pada kluster (UV)

n_W : jumlah objek pada kluster W

2.3 Penentuan Jumlah Kluster Optimum

Penentuan jumlah *cluster* optimum dilakukan dengan pendekatan indeks validitas *silhouette*. Indeks validitas *silhouette* merupakan suatu ukuran statistik yang digunakan untuk menyeleksi permasalahan penentuan jumlah kluster optimal yang dapat merepresentasikan grafis singkat seberapa baik setiap objek terletak dalam kluster.

Asumsikan data sudah dikelompokkan ke dalam kluster. Untuk setiap objek i , misalkan $a(i)$ adalah rata-rata jarak objek i ke semua objek dalam *cluster* yang sama dan $b(i)$ adalah rata-rata jarak minimum objek i ke semua objek pada suatu kluster serta i bukan anggota kluster. Dari penjelasan yang telah dipaparkan indeks validitas *silhouette* dapat ditulis dengan persamaan sebagai berikut:

$$(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.4)$$

Rata-rata (i) dari seluruh objek dalam suatu kluster menunjukkan seberapa dekat kemiripan objek dalam suatu kluster yang juga menunjukkan seberapa tepat objek telah dikelompokkan. Semakin dekat (i) kepada 1, maka semakin baik pengelompokan objek. Sebaliknya, semakin dekat (i) kepada -1, maka semakin buruk pengelompokan objek. Jumlah kluster k yang optimal merupakan estimasi dari harga k yang paling memaksimalkan nilai rata-rata (i) dan jika terdapat satu kluster yang anggotanya terdiri dari satu objek maka nilai rata-rata (i) adalah 0 (Rousseeuw, 1987).

2.4 Regresi Linier Berganda

Regresi linear berganda merupakan analisis regresi yang dapat menjelaskan hubungan antara variabel respon dengan lebih dari satu variabel prediktor (Maziyah, 2016). Model regresi linear untuk p variabel prediktor secara umum dapat ditulis:

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i$$

$$i = 1, 2, \dots, n; k = 1, 2, \dots, p$$

dengan,

y_i : variabel respon pengamatan ke- i

x_{ki} : variabel prediktor pengamatan ke- i

β_0 : intersep dari model

β_k : koefisien regresi

ε_i : residual pengamatan ke- i

Dari persamaan (2.6) di atas, jika diuraikan akan menjadi persamaan-persamaan berikut:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1p} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2p} + \varepsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{np} + \varepsilon_n \end{aligned}$$

Berdasarkan persamaan-persamaan di atas, diperoleh bentuk matriks sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dapat ditulis dengan:

$$Y = X\beta + \varepsilon$$

Metode yang dapat digunakan dalam mengestimasi parameter model regresi berganda adalah metode kuadrat terkecil atau *ordinary least square*/OLS (Basri, Pemodelan Regresi Berganda untuk Data dalam Studi kecerdasan Emosional, 2018), yaitu dengan meminimumkan jumlah kuadrat residual berdasarkan persamaan (2.7):

$$\varepsilon^T \varepsilon = (Y - X\beta)^T (Y - X\beta) = Y^T Y - 2Y^T X\beta + \beta^T X^T X\beta$$

Sehingga diperoleh bentuk estimasi sebagai berikut:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{2.5}$$

Berdasarkan persamaan (2.8), maka:

$$\hat{Y} = X\hat{\beta} \tag{2.6}$$

2.5 Analisis Spasial

Secara umum, analisis spasial membutuhkan data yang berdasarkan lokasi dan memuat karakteristik dari lokasi tersebut. Regresi spasial merupakan

analisis yang mempertimbangkan hubungan antara suatu variabel dengan variabel lain dengan memberikan efek spasial pada lokasi yang menjadi pusat pengamatan (Kurnianto, dkk, 2021). Anselin (1988) mengatakan bahwa terdapat dua efek spasial pada model regresi spasial yaitu *spatial dependence* (ketergantungan spasial) dan *spatial heterogeneity* (keragaman spasial). Pengujian efek ketergantungan spasial dapat menggunakan indeks Moran/*Moran's I* dan uji *lagrange multiplier*/uji LM (Weku, Eksplorasi Efektifitas Model Spasial untuk Menjelaskan Hubungan Antara Penduduk dan Infrastruktur Terhadap Kesejahteraan Masyarakat Kota Manado, 2020). Sedangkan untuk menguji efek keragaman spasial dengan uji *Breush-Pagan* (uji BP). Salah satu hal yang penting dalam analisis spasial yaitu adanya pembobot atau yang biasa disebut sebagai matriks pembobot spasial.

2.6 Model Regresi Spasial

Model umum regresi spasial adalah sebagai berikut (LeSage, 1999):

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

Beberapa model yang dapat terbentuk dari model umum regresi spasial adalah sebagai berikut:

1. Jika $\rho = 0$ dan $\lambda = 0$ maka persamaan menjadi model *ordinary least square* (OLS)

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

2. Jika $\rho \neq 0$ dan $\lambda = 0$ maka persamaan menjadi model *spasial autoregressive* (SAR)

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Adapun bentuk penaksir parameter dari model SAR yaitu:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \rho \mathbf{W}) \mathbf{y} \quad (2.7)$$

3. Jika $\rho = 0$ dan $\lambda \neq 0$ maka menjadi model *spatial error model* (SEM)

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon}$$

Adapun bentuk penaksir parameter dari model SEM yaitu:

$$\hat{\beta} = [(X - \lambda WX)^T(X - \lambda WX)]^{-1}(X - \lambda WX)^T(I - \lambda W)y \quad (2.8)$$

4. Jika $\rho \neq 0$ dan $\lambda \neq 0$ maka persamaan menjadi gabungan dari model SAR dan SEM atau biasa disebut *spatial autoregressive moving average* (SARMA)

$$y = \rho Wy + X\beta + u$$

$$u = \lambda Wu + \varepsilon$$

Adapun bentuk penaksir parameter dari model SARMA yaitu:

$$\hat{\beta} = [(X - \lambda WX)^T(X - \lambda WX)]^{-1}(X - \lambda WX)^T(I - \lambda W - \rho W)y \quad (2.9)$$

dengan,

y : vektor variabel respon ($n \times 1$)

X : matriks variabel prediktor ($n \times (p + 1)$)

β : vektor koefisien parameter regresi ($(p + 1) \times 1$)

λ : koefisien parameter spasial *error*

ρ : koefisien parameter spasial *lag* dari variabel respon

W : matriks pembobot spasial ($n \times n$)

u, ε : vektor residual ($n \times 1$)

I : matriks identitas ($n \times n$)

2.7 Pembobot Spasial

Matriks pembobot spasial adalah matriks yang menggambarkan hubungan suatu lokasi dengan lokasi yang lain. Penentuan matriks pembobot spasial W dapat menggunakan konsep *contiguity* (ketersinggungan spasial). Adapun bentuk umum matriks spasial W adalah sebagai berikut:

$$W = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2n} \\ w_{31} & w_{32} & w_{33} & \dots & w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{nn} \end{bmatrix}$$

Beberapa peneliti mengembangkan konsep pemilihan matriks pembobot yang tidak hanya berdasarkan pada persinggungan secara geografis tetapi juga berdasar pada karakteristik masalah yang sedang diteliti. Hal tersebut dapat berupa aksesibilitas maupun kesamaan karakteristik sosial ekonomi. Pembobot

dengan konsep tersebut biasa disebut dengan *customized contiguity* (Sulistiyono, 2016). Pada penelitian ini akan menggunakan pembobot *customized contiguity* yaitu pembobot dari hasil analisis kluster *average linkage*. Kabupaten/kota yang berada dalam kluster yang sama akan diberi nilai 1 dan nilai 0 untuk daerah pada kluster lainnya atau dapat ditulis sebagai berikut

$$w = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ berada dalam satu kluster} \\ 0, & \text{lainnya} \end{cases}$$

2.8 Dependensi Spasial

Dependensi spasial menunjukkan adanya ketergantungan antarlokasi. Misalnya lokasi i tergantung atau terhubung dengan lokasi j , dimana $i \neq j$. Pengujian efek dependensi spasial dapat menggunakan indeks moran (*Moran's I*) dan uji *lagrange multiplier* (uji LM).

Indeks Moran merupakan teknik dalam analisis spasial untuk mengetahui kebergantungan atau autokorelasi spasial antardaerah. Perhitungan autokorelasi spasial dengan *Moran's I* dapat dilakukan dengan cara berikut (Wuryandari, dkk, 2014):

Indeks Moran dengan matriks pembobot terstandarisasi \mathbf{W}

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_j - \bar{x}) (x_i - \bar{x})}{S_0 \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.10)$$

dengan,

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

I : Indeks Moran

n : banyaknya lokasi kejadian

x_i : nilai pada lokasi i

x_j : nilai pada lokasi j

\bar{x} : rata-rata dari jumlah variabel atau nilai

w_{ij} : elemen pada pembobot terstandarisasi antara daerah i dan j

Untuk mengidentifikasi adanya autokorelasi spasial dapat dilakukan dengan uji signifikansi Indeks Moran. Uji hipotesis untuk Indeks Moran adalah sebagai berikut:

i. Hipotesis H_0 : Tidak terdapat autokorelasi spasial

H_1 : Terdapat autokorelasi spasial

- ii. Tingkat signifikansi: α
 iii. Statistik uji: $Z(I) = \frac{I-E(I)}{\sqrt{Var(I)}} \approx N(0,1)$

dengan,

$$E(I) = I_0 - \frac{1}{n-1}$$

$$Var(I) = \frac{n^2.S_1 - n.S_2 + 3.S_0^2}{(n^2-1)S_0^2} - [E(I)]^2$$

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2$$

$$S_2 = \sum_{i=1}^n (\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji})^2$$

kriteria uji:

Tolak H_0 pada taraf signifikansi α jika $|Z(I)| > Z_{1-\alpha}$

Adapun uji *lagrange multiplier* atau *LM test* digunakan sebagai dasar dalam memilih model regresi spasial yang sesuai. Hipotesis yang digunakan pada uji LM ada tiga yaitu (Sanusi, dkk, 2018):

- i. Hipotesis yang digunakan:
- Hipotesis untuk SAR
 $H_0: \rho = 0$ (tidak ada dependensi spasial *lag*)
 $H_1: \rho \neq 0$ (ada dependensi spasial *lag*)
 - Hipotesis untuk SEM
 $H_0: \lambda = 0$ (tidak ada dependensi spasial *error*)
 $H_1: \lambda \neq 0$ (ada dependensi spasial *error*)
 - Hipotesis untuk SARMA
 $H_0: \rho, \lambda = 0$ (tidak ada dependensi spasial *lag* dan *error*)
 $H_1: \rho, \lambda \neq 0$ (ada dependensi spasial *lag* dan *error*)

ii. Statistik Uji

- Statistik uji untuk *lagrange multiplier error* (LM_{error})

$$LM_{error} = \frac{\left(\frac{\boldsymbol{\varepsilon}'\mathbf{W}\boldsymbol{\varepsilon}}{S^2}\right)^2}{T} \tag{2.11}$$

- Statistik uji untuk *lagrange multiplier lag* (LM_{lag})

$$LM_{lag} = \frac{\left(\frac{\boldsymbol{\varepsilon}'\mathbf{W}\mathbf{y}}{S^2}\right)}{\left(\frac{(\mathbf{W}\mathbf{X}\boldsymbol{\beta})'\mathbf{M}(\mathbf{W}\mathbf{X}\boldsymbol{\beta}) + TS^2}{S^2}\right)} \quad (2.12)$$

- Statistik uji untuk *lagrange multiplier lag* (LM_{SARMA})

$$LM_{SARMA} = \frac{\left(\frac{\boldsymbol{\varepsilon}'\mathbf{W}\mathbf{Y}}{S^2} - \frac{\boldsymbol{\varepsilon}'\mathbf{W}\boldsymbol{\varepsilon}}{S^2}\right)^2}{\left(\frac{(\mathbf{W}\mathbf{X}\boldsymbol{\beta})'\mathbf{M}\mathbf{W}\mathbf{X}\boldsymbol{\beta}}{S^2} + T\right) - T} + \frac{\left(\frac{\boldsymbol{\varepsilon}'\mathbf{W}\boldsymbol{\varepsilon}}{S^2}\right)^2}{T} \quad (2.13)$$

dengan,

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$T = tr((\mathbf{W}' + \mathbf{W})\mathbf{W})$$

$$S^2 = \frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{n}$$

\mathbf{X} : Matriks variabel prediktor ($n \times k$)

\mathbf{W} : Matriks pembobot spasial ($n \times n$)

$\boldsymbol{\varepsilon}$: Vektor residual ($n \times 1$)

\mathbf{y} : Vektor variabel respon ($n \times 1$)

kriteria uji:

Tolak H_0 jika nilai $LM > \chi^2_{(\alpha,1)}$ atau $p - value < \alpha$.

2.9 Uji Asumsi Residual

Beberapa asumsi yang harus dipenuhi dalam analisis regresi yaitu berdistribusi normal ($0, \sigma^2$), ragam residual homogen (identik) dan tidak ada ketergantungan antar residual (independen) atau tidak terjadi autokorelasi antar residual (Dewi & Winahju, 2014).

1. Uji Distribusi Normal

Pengujian asumsi residual berdistribusi normal dapat menggunakan uji *Kolmogorov-Smirnov* berikut:

i. Hipotesis H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

ii. Statistik uji: $D = \sup |S(x) - F_0(x)|$

dengan,

$S(x)$: sebaran kumulatif sampel

$F_0(x)$: sebaran kumulatif normal

iii. Kriteria uji: tolak H_0 apabila $D > D_{(\alpha,n)}$ atau nilai $p - value < \alpha$

2. Uji Asumsi Identik

Pengujian asumsi residual identik dapat menggunakan uji *Breusch-Pagan* berikut:

i. Hipotesis H_0 : residual identik

H_1 : residual tidak identik

ii. Statistik uji: $LM = \frac{NT}{2(t-1)} \sum_{i=1}^N \left[\frac{[\sum_{t=1}^T \varepsilon_{it}]^2}{\sum_{t=1}^T \varepsilon_{it}^2} - 1 \right]^2$

iii. Kriteria uji: tolak H_0 apabila $LM > X^2_{\alpha;N-1}$ atau $p - value < \alpha$

3. Uji Independen

Pengujian asumsi residual independen dapat menggunakan uji *Durbin-Watson* berikut:

i. Hipotesis H_0 : residual independen

H_1 : residual tidak independen

ii. Statistik uji: $d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n (e_i)^2}$

iii. Kriteria uji: gagal tolak H_0 apabila nilai $d_U < d < 4 - d_U$

2.10 Uji Signifikansi Parameter

Pengujian signifikansi pada model spasial dapat menggunakan *Likelihood Ratio Test*. Pengujian ini dilakukan untuk melihat ada atau tidaknya pengaruh yang signifikan antara variabel bebas dan variabel terikat. Pengujian hipotesis uji *Likelihood Ratio* pada model spasial sebagai berikut (Dewi & Ekaria, 2022)

$H_0: \beta = 0$ (parameter tidak signifikan)

$H_1: \beta \neq 0$ (parameter signifikan)

Statistik uji :

$$G = -2 \ln \left(\frac{L_0}{L_i} \right) \quad (2.14)$$

dengan L_0 merupakan log-*Likelihood* dari model tanpa variabel prediktor dan L_1 merupakan log-*Likelihood* dari model dengan p variabel prediktor.

Kriteria pengambilan keputusan yaitu tolak H_0 jika $G > \chi^2_{(\alpha,1)}$ atau $p - value < \alpha$

2.11 Pemilihan Model Terbaik

Model terbaik dalam pemodelan regresi dapat diperoleh dengan membandingkan hasil dari persamaan regresi metode-metode yang digunakan. Metode terbaik dapat ditentukan dengan koefisien determinasi R^2 dan *akaike information criterion* (AIC). Jika nilai R^2 semakin mendekati satu maka pengaruh variabel prediktor terhadap variabel respon semakin tinggi. Hal tersebut berarti bahwa kecocokan model dengan data semakin baik. Adapun menurut metode AIC, model dengan nilai AIC terkecil adalah model regresi terbaik (Sanusi, dkk, 2018). Koefisien determinasi R^2 dinotasikan sebagai berikut:

$$R^2 = \frac{SSR}{SST} \quad (2.15)$$

dengan SSR : *sum square regression* (jumlah kuadrat regresi)

SST : *sum square total* (jumlah kuadrat total)

Sedangkan AIC dinotasikan sebagai berikut:

$$AIC = -2Lm + m \quad (2.16)$$

dengan, Lm : maksimum log – *likelihood*

m : jumlah parameter dalam model

2.12 Tuberkulosis

Penyakit tuberkulosis/TB disebabkan oleh bakteri *Mycobacterium tuberculosis*. TB sebagian besar menyerang paru yang dikenal dengan istilah TB paru. Bakteri yang ada pada paru dapat mengikuti aliran darah sehingga dapat menyerang hampir semua organ tubuh. Beberapa contoh organ tubuh lain yang dapat terserang TB yaitu usus, otak, kelenjar, tulang maupun kulit (Kementerian Kesehatan RI, 2019). Kejadian penyakit tuberkulosis berdasarkan perspektif epimologi disebabkan oleh hasil interaksi antara pejamu (*host*), penyebab (*agent*) dan lingkungan (*environment*). Pada sisi pejamu, daya tahan tubuh seseorang sangat berpengaruh terhadap kerentanan infeksi dari bakteri

Mycobacterium tuberculosis (Kementerian Kesehatan RI, 2018). Adapun faktor risiko TB berdasarkan faktor pejamu dan lingkungan diantaranya sebagai berikut:

1) Usia

Usia merupakan salah satu hal yang berpengaruh terhadap kejadian TB paru. Beberapa penelitian telah membuktikan bahwa adanya perbedaan usia menyebabkan terjadinya frekuensi suatu penyakit berbeda pula. Hal tersebut berarti, variabel usia perlu diperhitungkan dalam meneliti sebuah penyakit. Pangaribuan (2020) menyatakan bahwa kelompok usia tua (55 tahun ke atas) lebih berisiko 1,73 kali mengalami penyakit TB dibanding usia 15-34 tahun. Hartanto (2019) menyatakan bahwa distribusi kasus TB paru lebih banyak pada kelompok usia 46-55 tahun.

2) Keluarga miskin

Penduduk miskin merupakan penduduk dengan rata-rata pengeluaran per kapita per bulan berada di bawah garis kemiskinan (BPS, 2020). Kemiskinan dapat menjadikan seseorang lebih rentan terhadap suatu penyakit termasuk TB paru. Hal tersebut terjadi karena semakin rendah pendapatan seseorang maka akan semakin sulit untuk memenuhi syarat kesehatan yang layak. Dariatno (2021) menyatakan bahwa terdapat hubungan positif dan signifikan antara kemiskinan terhadap angka TB di Indonesia. Rajak (2021) menyatakan bahwa jumlah penduduk miskin berpengaruh terhadap terjadinya kasus TB.

3) Fasilitas pelayanan kesehatan

Menurut Undang-Undang Nomor 36 Tahun 2009 tentang kesehatan, fasilitas pelayanan kesehatan merupakan suatu alat dan/atau tempat yang digunakan untuk menyelenggarakan pelayanan kesehatan, baik promotif, preventif, kuratif, maupun rehabilitatif yang dilakukan oleh pemerintah pusat, pemerintah daerah, dan/atau masyarakat. Keberadaan sarana kesehatan berpengaruh terhadap derajat kesehatan masyarakat suatu negara (Kementerian Kesehatan RI, 2020). Penelitian yang dilakukan Wulandari (2012) menyimpulkan bahwa analisis spasial menunjukkan adanya hubungan antara jumlah sarana kesehatan terhadap jumlah kasus baru dan insidens TB paru BTA (+). Salah satu sarana kesehatan yaitu puskesmas (Pusat Kesehatan Masyarakat).

4) Kebiasaan merokok

Kebiasaan merokok merupakan salah satu faktor yang dapat memicu seseorang lebih berisiko terkena TB paru. Hal tersebut karena ketika merokok seseorang menghisap racun sehingga dapat merusak kesehatan yang membuat mudah terinfeksi penyakit termasuk TB (Mathofani & Febriyanti, Faktor-faktor yang Berhubungan dengan Kejadian Tuberkulosis (TB) Paru di Wilayah Kerja Puskesmas Serang Kota Tahun 2019, 2020). Penelitian sebelumnya telah dilakukan oleh Hapsari, dkk (2013) menyatakan bahwa terdapat hubungan antara riwayat merokok dengan pasien TB paru.

5) Jenis kelamin

Jenis kelamin merupakan salah satu faktor yang berpengaruh terhadap terjadinya sebuah penyakit, termasuk TB. Berdasarkan laporan dari Kemenkes RI, jumlah kasus TB pada laki-laki lebih banyak jika dibandingkan pada perempuan baik secara nasional maupun pada setiap provinsi. Menurut Hartanto dkk (2020), laki-laki dua kali lebih berisiko mengalami TB paru karena mobilitasnya cenderung lebih tinggi sehingga kemungkinan terpapar penyakit TB paru lebih besar. Penelitian lain mengenai pengaruh jenis kelamin terhadap penularan penyakit TB dilakukan oleh Pangaribuan (2020) yang menyatakan bahwa laki-laki 2,07 kali lebih berisiko mengalami TB. Nafsi (2019) menyatakan bahwa kasus TB Paru lebih banyak terjadi pada laki-laki sebanyak 65,22% jika dibandingkan dengan perempuan. Dewi, dkk (2020) menyatakan bahwa variabel yang berpengaruh terhadap terjadinya kasus TB adalah jumlah penduduk laki-laki.