PEMODELAN *EIGENVECTOR SPATIAL FILTERING* DAN *RANDOM EFFECT EIGENVECTOR SPATIAL FILTERING* PADA DATA TINGKAT PENGANGGURAN TERBUKA DI SULAWESI SELATAN

AYUNI SIHOMBING
H051201043

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR

2024

PEMODELAN *EIGENVECTOR SPATIAL FILTERING* DAN *RANDOM EFFECT EIGENVECTOR SPATIAL FILTERING* PADA DATA TINGKAT PENGANGGURAN TERBUKA DI SULAWESI SELATAN

AYUNI SIHOMBING
H051201043

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

2024

PEMODELAN *EIGENVECTOR SPATIAL FILTERING* DAN *RANDOM EFFECT EIGENVECTOR SPATIAL FILTERING* PADA DATA TINGKAT PENGANGGURAN TERBUKA DI SULAWESI SELATAN

AYUNI SIHOMBING
H051201043

Skripsi

sebagai salah satu syarat untuk mencapai gelar sarjana

Program Studi Statistika

pada

DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

SKRIPSI

PEMODELAN *EIGENVECTOR SPATIAL FILTERING* DAN *RANDOM EFFECT EIGENVECTOR SPATIAL FILTERING* PADA DATA TINGKAT PENGANGGURAN TERBUKA DI SULAWESI SELATAN

AYUNI SIHOMBING
H051201043

Skripsi,

telah dipertahankan di depan Panitia Ujian Sarjana pada tanggal 6 Novembre 2024 dan dinyatakan telah memenuhi syarat kelulusan pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:	Mengetahui:
Pembimbing tugas akhir,	Ketua Program Studi,
Siswanto, S.Si., M.Si.	Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19920107 201903 1 012	NIP. 19770808 200501 2 002

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS HASANUDDIN
DEPARTEMEN STATISTIKA
FAKULTAS

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Eigenvector Spatial Filtering* dan *Random Effect Eigenvector Spatial Filtering* pada data Tingkat Pengangguran Terbuka di Sulawesi Selatan" adalah benar karya saya dengan arahan dari pembimbing Siswanto, S.Si., M.Si. sebagai Pembimbing Utama. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas pembuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 02 Desember 2024

Ayuni Sihombing
NIM. H051201043

UCAPAN TERIMAKASIH

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala limpahan cinta-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "**Pemodelan *Eigenvector Spatial Filtering* dan *Random Effect Eigenvector Spatial Filtering* pada data Tingkat Pengangguran Terbuka di Sulawesi Selatan**". Penulis menyampaikan ucapan terima kasih kepada Bapak **Siswanto, S.Si., M.Si.** selaku Pembimbing Utama yang dengan penuh kesabaran telah meluangkan waktu, pemikiran, semangat, dan bimbingan kepada penulis selama proses penulisan tugas akhir ini. Terima kasih kepada Ibu **Dr. Anna Islamiyati, S.Si., M.Si.** dan Ibu **Sitti Sahriman, S.Si., M.Si**. selaku Tim Penguji yang senantiasa memberikan saran dan kritikan yang membangun dalam penulisan tugas akhir ini. Terima kasih kepada **Pimpinan Universitas Hasanuddin, Departemen Statistika, Jajaran Dosen,** dan **Staf Departemen Statistika** yang telah memfasilitasi, memberikan ilmu bermanfaat, dan membantu penulis selama menempuh studi.

Penghargaan dan ucapan terima kasih yang tulus juga penulis ucapkan kepada kedua orang tua penulis yang tercinta, Ayahanda **Paian Sihombing** dan Ibunda **Meldina Datubara** yang telah memberikan dukungan penuh, pengorbanan luar biasa, limpahan cinta dan kasih sayang, kesabaran hati, serta dengan ikhlas telah menemani setiap langkah penulis dengan doa dan restu mulianya. Terima kasih kepada saudara-saudari tersayang penulis **Irvan**, **Irna**, **Tisza** dan **Levi** yang telah menasehati, memberikan doa, dukungan, semangat dan kebahagiaan kepada penulis. Terima kasih kepada **Oma** dan **Opa Pailang** yang sudah memberikan kenyamanan dan menjadi orang tua penulis selama tinggal di Makassar.

Terima kasih kepada sahabat penulis **Divia** dan **Salsabila** yang memberikan kenyamanan serta menjadi guru terbaik penulis selama perkuliahan. Terima kasih kepada teman seperantauan penulis **Evina** dan **Yelni** yang selalu memberikan dukungan dalam bentuk canda dan tawa. Terima kasih kepada teman-teman **Ridwan, Fadlan, Isti, Faldi, Rani, Hakam** dan **Gading** yang sudah menjadi tempat bertanya penulis serta garda terdepan dalam proses pembelajaran dan penyelesaian tugas akhir ini. Terima kasih kepada **Izzul, Reza, Mukhlis, Fahmi, Nahdi, Isra, Edward, Heri, Kur, Razy, Theo, Febi, Azalia, Naje, Ayu Afri, Aliyah, Uci, Liza, Linda, Alisha dan** teman-teman **Statistika'20** yang namanya tidak bisa disebut satu persatu yang telah menemani penulis menghadapi suka duka dan menjadi teman seperjuangan yang menyenangkan selama perkuliahan. Terima kasih kepada keluarga besar **KMK FMIPA Unhas** dan **Himastat FMIPA Unhas** atas kebersamaannya selama ini.

Makassar, 4 November 2024

Penulis,

Ayuni Sihombing

ABSTRAK

Ayuni Sihombing. **Pemodelan *Eigenvector Spatial Filtering* dan *Random Effect Eigenvector Spatial Filtering* pada data Tingkat Pengangguran Terbuka di Sulawesi Selatan** (dibimbing oleh Siswanto).

Latar Belakang. Autokorelasi spasial merupakan keterkaitan atau ketergantungan antara suatu lokasi dengan lokasi-lokasi disekitarnya. Masalah autokorelasi spasial dalam analisis regresi spasial dapat di atasi dengan pendekatan *spatial filtering* yang memanfaatkan eigenvektor dari matriks pemobobot spasial. Pendekatan *spatial filtering* dapat digunakan dengan dua metode yaitu *Eigenvector Spatial Filtering* (ESF) dan *Random Effect Eigenvector Spatial Filtering* (RE-ESF)*.* Eigenvektor yang digunakan dalam model regresi spasial dapat mengidentifikasi pola spasial yang dominan dalam data. Oleh karena itu, dilakukan pemodelan ESF dan RE-ESF untuk mengidentifikasi faktor-faktor yang mempengaruhi Tingkat Pengangguran Terbuka (TPT) di Provinsi Sulawesi Selatan pada Tahun 2022. **Tujuan**. Penelitian ini bertujuan untuk mendapatkan model terbaik antara *Eigenvector Spatial Filtering* dan *Random Effect Eigenvector Spatial Filtering* pada data Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan pada Tahun 2022. **Metode**. Penelitian ini dilakukan dengan dua metode estimasi yang berbeda yaitu *Maximum Likelihood Estimation* (MLE) untuk ESF dan *Restricted Maximum Likelihood Estimation* (REML) untuk RE-ESF. **Hasil**. Model ESF dengan metode estimasi MLE dapat menjelaskan Tingkat Pengangguran Terbuka di Sulawesi Selatan pada Tahun 2022 sebesar 80,84%, dengan nilai AIC sebesar 78,1573. Sementara itu, model RE-ESF yang menggunakan metode estimasi REML mampu menjelaskan Tingkat Pengangguran Terbuka di Sulawesi Selatan pada Tahun 2022 sebesar 81,37%, dengan nilai AIC sebesar 76,2670. **Kesimpulan**. Model terbaik untuk mengidentifikasi faktor-faktor yang mempengaruhi Tingkat Pengangguran Terbuka di Sulawesi Selatan pada Tahun 2022 adalah *Random Effect Eigenvector Spatial Filtering* dengan metode estimasi *Restricted Maximum Likelihood Estimation.*

Kata kunci: *Spatial Filtering*, *Eigenvector Spatial Filtering, Random Effect Eigenvector Spatial Filtering, Maximum Likelihood Estimation*, *Restricted Maximum Likelihood Estimation*, Tingkat Pengangguran Terbuka.

ABSTRACT

Ayuni Sihombing. **Modeling Eigenvector Spatial Filtering and Random Effect Eigenvector Spatial Filtering on Open Unemployment Rate data in South Sulawesi** (supervised by Siswanto).

Background. Spatial autocorrelation is the relationship or dependence between a location and its surrounding locations. The problem of spatial autocorrelation in spatial regression analysis can be overcome with a spatial filtering approach that utilizes the eigenvectors of the spatial weighting matrix. The spatial filtering approach can be used with two methods, namely Eigenvector Spatial Filtering (ESF) and Random Effect Eigenvector Spatial Filtering (RE-ESF). Eigenvectors used in spatial regression models can identify dominant spatial patterns in the data. Therefore, ESF and RE-ESF modeling was conducted to identify factors that affect the Open Unemployment Rate in South Sulawesi Province in 2022. **Aim**. This research aims to get the best model between Eigenvector Spatial Filtering and Random Effect Eigenvector Spatial Filtering on Open Unemployment Rate data in South Sulawesi Province in 2022. **Methods.** This research was conducted with two different estimation methods, namely Maximum Likelihood Estimation (MLE) for ESF and Restricted Maximum Likelihood Estimation (REML) for RE-ESF. **Results**. The ESF model with the MLE estimation method can explain the Open Unemployment Rate in South Sulawesi in 2022 by 80.84%, with an AIC value of 78.1573. Meanwhile, the RE-ESF model using the REML estimation method was able to explain the Open Unemployment Rate in South Sulawesi in 2022 by 81.37%, with an AIC value of 76.2670. **Conclusion**. The best model to identify factors affecting the Open Unemployment Rate in South Sulawesi in 2022 is Random Effect Eigenvector Spatial Filtering with Restricted Maximum Likelihood Estimation method.

Keywords: Spatial Filtering, Eigenvector Spatial Filtering, Random Effect Eigenvector Spatial Filtering, Maximum Likelihood Estimation, Restricted Maximum Likelihood Estimation, Open Unemployment Rate.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Spasial Laten	Proses atau variabel yang tidak teramati secara langsung tetapi memiliki pengaruh pada pola spasial dalam data.
Bandwith	Ukuran jarak dalam fungsi pembobot yang menunjukkan hasil kalibrasi lokal yang dihaluskan.
Cross Validation	Teknik untuk mengevaluasi kinerja model dalam mendapatkan nilai bandwidth yang sangat berguna untuk mendapatkan nilai pembobotan.
Linear Mixed Model	Model regresi yang digunakan untuk menganalisis data di mana terdapat efek tetap (*fixed effects*) dan efek acak (*random effects*).
Fixed Effect	Parameter yang mewakili populasi secara keseluruhan dan diasumsikan konstan di seluruh pengamatan.
Random Effect	Variabel yang mewakili variasi yang berasal dari kelompok atau unit yang berbeda dalam data
Eigenvalue	Himpunan skalar khusus yang terkait dengan sistem persamaan linear.
Eigenvector	Vektor (bukan nol) yang tidak mengubah arah saat transformasi linier diterapkan.
Log-likelihood	Logaritma dari fungsi *likelihood* yang sering digunakan dalam estimasi parameter model statistic.
Iterasi	Proses pengulangan langkah-langkah dalam suatu algoritma atau prosedur sampai kondisi tertentu tercapai.
Gradien	Vektor yang berisi turunan pertama dari fungsi, menunjukkan arah laju perubahan tercepat.
Hessian	Matriks yang berisi turunan keduadari fungsi.
Fungsi Kepadatan Peluang	Fungsi yang menggambarkan probabilitas relatif dari variabel acak kontinu, dikenal sebagai fungsi kepadatan probabilitas (PDF).
Kontinu	Mengacu pada data atau variabel yang dapat mengambil nilai apa pun dalam suatu rentang tertentu, tidak terbatas pada nilai diskrit.
Konvergen	Kondisi ketika suatu urutan atau seri mendekati nilai tertentu seiring berjalannya waktu atau iterasi.
Euclidean	Konsep matematika yang digunakan untuk mengukur jarak antara dua titik dalam ruang dua atau tiga dimensi.
Variabel Acak	Variabel yang hasilnya merupakan nilai acak dan ditentukan oleh probabilitas.

DAFTAR SINGKATAN DAN LAMBANG

LAMBANG/SINGKATAN	**Arti dan Penjelasan**
Y	Variabel respon
X	Variabel prediktor
β	Parameter independen
ε	Nilai galat
$\boldsymbol{W}$	Matriks pembobot spasial
h	Lebar *bandwith*
H	Hessian
g	Gradien
I	Indeks *moran's I*
$\boldsymbol{\Lambda}$	Matriks diagonal n *eigenvalue*
$\boldsymbol{E}^*$	Matriks eigenvektor
$\boldsymbol{M}$	Matriks pemusatan
$\boldsymbol{\mathcal{D}}$	Matriks varians kovarians dari γ
$\boldsymbol{\mathcal{R}}$	Matriks varians kovarians dari ε
σ^2	Parameter varians
$\boldsymbol{\gamma}$	Vektor parameter eigenvektor
λ	*Eigenvalue*
R^2	Koefisien determinasi
E	Eigenvektor
b_k	Penduga koefisien variabel ke-k
e_n	Penduga koefisien eigenvektor ke$-n$
$\boldsymbol{I}$	Matriks identitas
$\mathbf{1}$	Vektor satu
ESF	*Eigenvektor Spatial Filtering*
RE-ESF	*Random Effect Eigenvector Spatial Filtering*
MLE	*Maximum Likelihood Estimation*
REML	*Restricted Maximum Likelihood*
OLS	*Ordinary Least Squares*
IID	*Independent and Identically Distributed*
BLUE	*Best Linear Unbiased Estimator*
AIC	*Akaike's Information Criterion*

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi spasial merupakan metode statistik yang digunakan untuk membangun model regresi guna menggambarkan hubungan antara variabel prediktor dan variabel respon, dengan mempertimbangkan efek lokasi atau *spatial effect* (Yasin dkk., 2020). Menurut Anselin (1988) efek lokasi terdiri dari dua jenis yaitu dependensi (autokorelasi) spasial dan heterogenitas spasial. Autokorelasi spasial mengacu pada keterkaitan atau ketergantungan antara suatu lokasi dengan lokasi-lokasi di sekitarnya. Hal ini menggambarkan kondisi nilai pengamatan di wilayah i dipengaruhi oleh nilai pengamatan di wilayah sekitarnya, misalnya wilayah j $(i \neq j)$. Jika terdapat pola sistematik di dalam penyebaran sebuah variabel, maka terdapat autokorelasi spasial.

Autokorelasi spasial menunjukkan bahwa nilai suatu variabel di suatu wilayah dipengaruhi oleh nilai variabel yang sama di wilayah-wilayah tetangga yang berdekatan. Autokorelasi spasial positif mengindikasikan lokasi yang berdekatan mempunyai nilai yang mirip dan cenderung berkelompok, autokorelasi spasial negatif mengindikasikan lokasi yang berdekatan mempunyai nilai yang berbeda dan cenderung menyebar, dan tidak terdapat autokorelasi spasial mengindikasikan pola lokasi acak (Mailanda dan Kusnandar, 2022). Salah satu pendekatan analisis regresi spasial untuk mengatasi autokorelasi spasial adalah penyaringan spasial atau *spatial filtering.*

Spatial Filtering adalah salah satu teknik untuk menangani masalah autokorelasi spasial dengan menyaring variabel dalam model menjadi dua tipe variabel yaitu variabel nonspasial dan variabel spasial. Salah satu metode *spatial filtering* untuk mengatasi masalah autokorelasi spasial dengan memanfaatkan eigenvektor dari matriks pembobot spasial adalah *Eigenvector Spasial Filtering* (Griffith, 1996). Pendekatan *Eigenvector Spasial Filtering* (ESF) merupakan suatu metode untuk mengatasi masalah autokorelasi spasial dengan cara menambahkan kombinasi linier dari eigenvektor matriks pembobot spasial ke dalam spesifikasi model regresi (Griffith, 2002). Menurut Thayn dan Simanis (2012), penggunaan ESF dapat mengurangi kesalahan spesifikasi spasial, meningkatkan kesesuaian model, serta memperbaiki normalitas dan homoskedastisitas sisaan.

Murakami dan Griffith (2013) mengembangkan metode ESF dengan memperhitungkan efek acak (*random effect*) yang dikenal dengan *Random Effect Eigenvector Spatial Filtering* (RE-ESF). Metode RE-ESF bermanfaat untuk menganalisis isu dependensi spasial dengan memasukkan komponen proses spasial acak dalam modelnya. Metode RE-ESF mampu meminimalkan nilai ragam akibat faktor acak, yang terjadi antara variabel prediktor dan proses spasial laten.

Secara matematis, model dasar dari ESF identik dengan model regresi linier biasa. Sementara itu, model linier dengan RE-ESF identik dengan model linier campuran (*mixed model*). Model linier campuran merupakan perluasan dari model linier dengan penambahan komponen efek acak dalam modelnya. Proses pendugaan parameter untuk model linier campuran dapat dilakukan menggunakan metode *Maximum Likelihood* (MLE) atau metode *Restricted Maximum Likelihood* (REML).

Zhang dkk (2018) pertama kali menggunakan metode ESF untuk memodelkan konsentrasi PM 2.5 dan membandingkannya dengan model *Ordinary Least Squares* (OLS). Pada penelitian tersebut, diperoleh bahwa model ESF secara efektif mengatasi autokorelasi spasial dalam residual OLS dan menghasilkan peningkatan dalam kebaikan model serta pengurangan dalam *standard errors* dan kesalahan *cross-validation errors*, dibandingkan dengan model OLS klasik. Penelitian tentang model ESF juga pernah dilakukan oleh Sahputri (2022) dengan melakukan analisis regresi spasial yang mengandung efek autokorelasi spasial di variabel respon atau sering disebut model *spatial autoregressive* (SAR) dengan pendekatan ESF. Wijayanto (2019) juga melakukan pemodelan multilevel intersep acak dengan metode ESF untuk memodelkan nilai UN di Kota Kendari dan didapatkan metode ESF berhasil menurunkan tingkat *overestimate* (bias) pada pendugaan ragam level-2.

Metode yang memperhitungkan aspek spasial kini semakin penting dalam analisis fenomena sosial-ekonomi yang dipengaruhi oleh lokasi geografis. Fenomena seperti distribusi penduduk, kemiskinan, hingga tingkat pengangguran terbuka sering kali menunjukkan adanya pola spasial, dengan kondisi di suatu wilayah dapat dipengaruhi oleh kondisi wilayah sekitarnya. Tingkat Pengangguran Terbuka (TPT) adalah salah satu bentuk indikator yang digunakan untuk mengukur besarnya angkatan kerja yang tidak terserap oleh pasar kerja dan menjadi potret akan kurang termanfaatkannya tenaga kerja (BPS, 2022). Pada Tahun 2022 Badan Pusat Statistik (BPS) melalui publikasinya menyatakan bahwa TPT di Provinsi Sulawesi Selatan berada pada angka 4,51% yang mana telah mengalami penurunan dari Tahun sebelumnya yang berada dalam angka 5,72%. Dari publikasi tersebut diketahui bahwa Kota Makassar memiliki tingkat pengangguran tertinggi yaitu sebesar 11,82% sedangkan yang terendah adalah Kabupaten Enrekang sebesar 0,58%.

TPT di suatu daerah diperkirakan dipengaruhi oleh TPT di daerah sekitarnya. Hal ini mungkin terjadi karena adanya faktor kedekatan atau ketetanggaan antar daerah (Wuryandari dkk., 2014). Adanya hubungan spasial dalam variabel respon akan menyebabkan pendugaan menjadi tidak tepat karena asumsi keacakan galat dilanggar. Sehingga untuk mengatasinya diperlukan suatu model regresi yang memasukkan hubungan spasial antar lokasi ke dalam model, yaitu dengan model regresi spasial (Sanusi dkk., 2018). Berdasarkan penelitian yang telah dipaparkan, maka peneliti ingin melakukan sebuah penelitian dengan **Pemodelan *Eigenvector Spatial Filtering* dan *Random Effect Eigenvector Spatial Filtering* pada data Tingkat Pengangguran Terbuka di Sulawesi Selatan.**

1.2 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data dari 24 Kabupaten/Kota di Provinsi Sulawesi Selatan pada Tahun 2022.
2. Estimasi parameter yang digunakan pada pemodelan ESF adalah *Maximum Likelihood Estimation.*
3. Estimasi parameter yang digunakan pada pemodelan RE-ESF adalah *Restricted Maximum Likelihood.*

1.3 Tujuan Penelitian

Tujuan dari penelitian adalah:

1. Memperoleh model ESF dan RE-ESF untuk memodelkan hubungan antara variabel prediktor dan variabel respon yang mengandung autokorelasi spasial.
2. Memperoleh model terbaik antara ESF dan RE-ESF pada data Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan Tahun 2022.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan berbagai manfaat sebagai berikut:

1. Memberikan pengetahuan baru bagi pembaca dalam mengembangkan analisis regresi spasial dengan menggunakan pendekatan *spatial filtering.*
2. Menambah informasi mengenai model *Eigenvector Spatial Filtering* (ESF) dan *Random Effect Eigenvector Spatial Filtering* (RE-ESF).
3. Memberikan informasi kepada Pemerintah Daerah di Provinsi Sulawesi Selatan mengenai faktor-faktor yang mempengaruhi Tingkat Pengangguran Terbuka guna membantu untuk menurunkan angka pengangguran di Provinsi Sulawesi Selatan.

1.5 Teori

1.5.1 Model Regresi Linier

Analisis regresi merupakan suatu metode yang digunakan untuk menganalisis hubungan antara variabel prediktor dan variabel respon. Model regresi yang memuat satu variabel respon dan satu variabel prediktor disebut dengan regesi linier sederhana, sedangkan jika menggunakan beberapa variabel prediktor maka dinamakan regresi linier berganda. Bentuk umum dari regresi linier berganda dinyatakan pada Persamaan (1) (Draper & Smith, 1998).

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

dengan:

y_i : nilai pengamatan ke-i variabel respon, $i = 1, 2, \dots, n$

x_{ki} : nilai pengamatan ke-i variabel prediktor k

β_0 : nilai intersep model regresi

β_k : koefisien parameter regresi untuk variabel prediktor k

ε_i : sisan pada pengamatan ke-i dengan asumsi $\varepsilon_i \sim NIID(0, \sigma^2)$

n : banyaknya unit pengamatan

k : banyaknya variabel prediktor, $k = 1, 2, \ldots, p$

Model regresi linier berganda pada Persamaan (1) dapat ditulis dalam notasi matriks sebagai berikut:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2}$$

dengan :

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & \ldots & X_{1p} \\ 1 & X_{21} & \ldots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \ldots & X_{np} \end{bmatrix} dan\ \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan:

$\boldsymbol{Y}$: variabel respon berukuran $n \times 1$

$\boldsymbol{\beta}$: parameter regresi beukuran $(p + 1) \times 1$

$\boldsymbol{X}$: matriks variabel prediktor berukuran $n \times (p + 1)$

$\boldsymbol{\varepsilon}$: vektor galat berukuran $n \times 1$

n : banyak pengamatan

Nilai parameter regresi $\boldsymbol{\beta}$ diduga dengan Persamaan (3).

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{3}$$

Asumsi yang melandasi penduga parameter regresi $\boldsymbol{\beta}$ pada Persamaan (3) yaitu :

a. $E(\varepsilon_i) = 0$
b. $Var(\varepsilon_i) = \sigma^2$
c. $cov(\varepsilon_i, \varepsilon_j) = 0, untuk\ i \neq j$

Berdasarkan Persamaan (3), nilai ε_i (galat) untuk melakukan uji autokorelasi spasial dapat dihitung dengan Persamaan (4).

$$\varepsilon_i = y_i - (\beta_0 + \sum_{k=1}^{p} \beta_k x_{ki}) \tag{4}$$

1.5.2 Matriks Pembobot Spasial

Matriks pembobot spasial atau matriks $\boldsymbol{W}$ merupakan matriks yang menggambarkan suatu hubungan kedekatan antar wilayah/lokasi pengamatan yang berukuran $n \times n$. Untuk setiap lokasi unit pengamatan ke-i, elemen matriks w_{ij} menentukan lokasi j yang dapat mempengaruhi nilai variabel di lokasi i (Lesage, 1999). Dalam matriks $\boldsymbol{W}$, setiap satuan luas digambarkan sebagai baris dan kolom. Sebagai contoh, nilai $w_{1,2}$ atau $w_{2,1}$ menunjukkan bobot antar lokasi pertama dan kedua. Lokasi yang dekat dengan lokasi tertentu akan cenderung memiliki w_{ij} besar, sedangkan lokasi yang berjauhan akan cenderung memiliki w_{ij} kecil (Suryoto dan Faradila, 2018). Matriks $\boldsymbol{W}$ memiliki beberapa asumsi:

a. Tidak adanya pengaruh pada dirinya sendiri, atau lokasi i tidak mempengaruhi sifat pada lokasi i sendiri, maka $w_{ii} = 0$, artinya matriks $\boldsymbol{W}$ merupakan matriks yang bernilai nol pada diagonal utamanya (Arbia dan Baltagi, 2008).
b. Merupakan matriks simetris dengan $w_{ij} = w_{ji}$ yang pada dasarnya menggambarkan keterkaitan hubungan spasial.
c. Matriks baris menunjukkan bagaimana suatu lokasi secara spasial terkait dengan wilayah lain. Oleh karena itu, nilai total pada baris i adalah total tetangga yang dimiliki oleh lokasi i (Suryoto dan Faradila, 2018).

Bentuk umum matriks pembobot spasial dapat dilihat pada Persamaan (5).

$$\boldsymbol{W}_{ij} = \begin{bmatrix} W_{11} & \cdots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{n1} & \cdots & W_{nn} \end{bmatrix} \tag{5}$$

Matriks W_{ij} dibentuk berdasarkan pendekatan titik dan area. Dalam pendekatan titik adalah dengan memperhatikan jarak antar wilayah satu dengan wilayah yang lain kemudian dihitung berdasarkan titik koordinat garis lintang dan bujur. Dalam pendekatan area, matriks pembobot spasial dapat didefinisikan menggunakan konsep ketetanggaan dan jarak. Penentuan pembobot spasial yang digunakan sangat penting karena pembobot dapat menjelaskan besarnya pengaruh spasial antar suatu daerah. Salah satu jenis matriks pembobot jarak adalah jarak eksponensial (Murakami dan Griffith, 2015).

$$\boldsymbol{W}_{ij} = \exp\left[-\left(\frac{d_{ij}}{h}\right)\right] \tag{6}$$

Keterangan:

$\boldsymbol{W}_{ij}$: Nilai pembobot pada lokasi ke-i dan lokasi ke-j

h : Lebar *bandwith*

d_{ij} : Jarak antara lokasi ke-i dan lokasi ke-j

Bandwith (h) adalah ukuran jarak dalam fungsi pembobot yang menunjukkan hasil kalibrasi lokal yang dihaluskan. Pemilihan *bandwith* optimum menjadi hal yang penting karena dapat mempengaruhi hasil pendugaan dari model (Haryanto dkk., 2019). Salah satu metode dalam mendapatkan nilai *bandwith* optimum adalah metode *Cross Validation* (CV) melalui Persamaan (7).

$$CV(h) = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(h)\right)^2 \tag{7}$$

dengan:

n : Banyaknya pengamatan

y_i : Penggamatan ke-i

$\hat{y}_{\neq i}$: Nilai duga pengamatan ke-i yang didapat tanpa pengamatan ke$-i$

Nilai *bandwith* optimum didapatkan dari nilai h yang menghasilkan nilai CV minimum. d_{ij} merepresentasikan jarak antara titik ke-i dan titik ke-j. Perhitungan jarak menggunakan jarak *Euclidean* melaluin Persamaan (8).

$$d_{ij}^{s} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{8}$$

1.5.3 Uji Autokorelasi Spasial

Dependensi spasial (ketergantungan spasial) atau disebut autokorelasi spasial muncul karena berdasarkan hukum Tobler I (1979) yang berisi bahwa segala sesuatu saling berhubungan dengan hal lain tetapi sesuatu yang lebih dekat mempunyai pengaruh yang lebih besar (Musfika *cit* Nurul, 2017).

Autokorelasi spasial merupakan pengujian untuk melihat pengamatan di suatu lokasi berpengaruh terhadap lokasi lain yang letaknya berdekatan atau bertetangga. Autokorelasi spasial merupakan unit pengamatan pada suatu lokasi i dengan unit pengamatan pada lokasi j dengan $j \neq i$ tidak saling bebas (Lesage, 1999). Uji yang digunakan untuk mengetahui autokorelasi spasial adalah dengan menggunakan indeks *Moran's I* (Anselin, 1988).

Indeks Moran bertujuan untuk melihat nilai autokorelasi spasial secara global. Indeks Moran mengukur hubungan satu variabel prediktor $X\ (x_i\ dan\ x_j)$ dengan $i \neq j, i = 1,2,3, \dots, n$ dan $j = 1,2,3, \dots, n$ (Rahmadeni, 2020). Nilai rentang indeks *moran's I* dalam matriks pembobot spasial yang telah distandarisasi berkisar antara $-1 \leq I \leq 1$. Jika nilai $-1 \leq I \leq 0$ menunjukkan terjadinya autokorelasi spasial negatif, sedangkan $0 \leq I \leq 1$ menunjukkan terjadinya autokorelasi spasial positif (Safitri dan Ruliana, 2022).

a) Hipotesis

H_0 : $I = 0$ (tidak terdapat autokorelasi spasial antar lokasi)

H_1 : $I \neq 0$ (terdapat autokorelasi spasial antar lokasi)

b) Statistik Uji

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\, (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{j=1}^{n} (x_i - \bar{x})^2} \tag{9}$$

dengan:

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

I : Indeks *moran's I*

n : Banyaknya lokasi pengamatan

x_i : Nilai pada lokasi ke-i

x_j : Nilai pada lokasi ke-j

$\bar{x}$: Nilai rata-rata x_i dari lokasi ke-n

w_{ij} : Elemen pada pembobot yang telah distandarisasi antara lokasi i dan j.

c) Kriteria uji

$$Z(I) = \frac{I - E(I)}{\sqrt{var(I)}} \tag{10}$$

dengan:

$E(I) = -\frac{1}{n-1}$ dan $var(I) = \frac{n^2 S_1 - nS_2 + 3S_0^2}{(n^2-1)S_0^2} - [E(I)]^2$

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

$S_1 = \frac{1}{2} S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} (w_{ij} + w_{ji})^2$

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

Jika $|Z_{hitung}| > Z_{\frac{\alpha}{2}}$ maka H_0 ditolak.

Jika $|Z_{hitung}| < Z_{\frac{\alpha}{2}}$ maka H_0 diterima.

1.5.4 Analisis Regresi Spasial

Regresi spasial adalah analisis yang mengevaluasi hubungan antara satu variabel dengan beberapa variabel lain dengan memberikan efek spasial pada beberapa lokasi yang menjadi pusat pengamatan. Model regresi spasial dikembangkan oleh Anselin pada tahun 1988 menggunakan data *control-section*, model ini disebut dengan *General Spatial Model*. Menurut Anselin (1988), bentuk umul model regresi spasial ditunjukkan pada Persamaan (11).

$$\begin{aligned} \boldsymbol{y} &= \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{u} \\ \boldsymbol{u} &= \lambda \boldsymbol{W} \boldsymbol{u} + \boldsymbol{\varepsilon} \\ & \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \end{aligned} \tag{11}$$

Keterangan:

y	: Vektor variabel respon berukuran $n \times 1$
ρ	: Koefisien autoregresi lag spasial
W	: Matriks pembobot spasial berukuran $n \times n$
X	: Matriks variabel prediktor berukuran $n \times (p + 1)$
β	: Vektor koefisien parameter regresi berukuran $(p + 1) \times 1$
u	: Vektor galat, diasumsikan mengandung autokorelasi berukuran $n \times 1$
λ	: Koefisien autoregresi galat spasial
ε	: Vektor galat yang bebas autokorelasi berukuran $n \times 1$
I	: Matriks indentitas berukuran $n \times n$
n	: Banyaknya amatan/lokasi

Jika $\rho \neq 0$ dan $\lambda = 0$ maka model tersebut merupakan model *Spatial Autoregressive Regression* (SAR). Hal ini berarti bahwa model memiliki variabel respon yang berkorelasi spasial. Model regresi spasialnya menjadi (Fatati dkk., 2017):

$$
\begin{aligned}
\boldsymbol{y} &= \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\
\boldsymbol{\varepsilon} &= \boldsymbol{y} - \rho\boldsymbol{W}\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta} \\
&= (\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}
\end{aligned}
$$

1.5.5 *Spatial Filtering*

Spatial Filtering adalah salah satu teknik untuk menangani masalah autokorelasi spasial dengan menyaring variabel dalam model menjadi dua tipe variabel yaitu variabel nonspasial dan variabel spasial. Menurut Griffith (2003) proses filter menggunakan matriks pembobot yang dapat menangkap kovariat dari satu atau lebih variabel acak yang menggambarkan karakteristik dari satu unit area. *Spatial Filtering* atau penyaringan spasial adalah metode statistik yang bertujuan untuk memperoleh hasil yang lebih baik dan andal dalam analisis data spasial.

Spatial Filtering berfokus pada dekomposisi data spasial menjadi 3 komponen utama, yaitu:

a. Tren, komponen ini mewakili pola umum atau tren jangka panjang dalam data spasial.
b. Komponen acak yang terstruktur secara spasial adalah komponen yang menunjukkan hubungan spasial yang lebih halus atau terlokalisasi dalam data.
c. *Random Noise* adalah komponen acak yang tidak memiliki struktur spasial yang jelas dalam data.

Tujuan dari *spatial filtering* adalah untuk memisahkan ketiga komponen tersebut sehingga komponen yang terstruktur secara parsial dapat dianalisis dengan lebih tepat, sementara tren dan *random noise* dapat diabaikan atau ditangani secara terpisah. Tiefelsdorf dan Griffith (2007) menjelaskan bahwa proses pemisahan ini menggunakan operator matematika yang lebih baik, menghasilkan *eigenvalue* dan eigenvektor.

Berbagai metode penyaringan spasial yang ada saat ini masih terbatas, di antaranya adalah operator linier autoregresif, spesifikasi berbasis Gi dari Getis, eigenfungsi dari matriks jarak antar titik, serta eigenfungsi dari matriks bobot spasial. Metode-metode ini menggunakan operator matematika yang berbeda untuk mendekomposisi variabel spasial menjadi tiga komponen.

1. Operator linier autoregresif digunakan untuk menghilangkan autokorelasi pada variabel dependen agar menghasilkan kesalahan acak yang independen dan terdistribusi identik (Tobler, 1975).
2. Metode berbasis Gi dari Getis mengonversi setiap variabel yang terkorelasi secara spasial menjadi dua variat, satu yang menangkap SA dan yang lainnya yang mengandung efek sistematis dan acak nonspasial (Getis, 1990).
3. Eigenfungsi dari matriks jarak antar titik (*principal coordinates of neighbor matrices*) yang memanfaatkan eigenfungsi dari matriks $n \times n$ dari jarak geografis yang terpotong antar lokasi [pemotongan dapat dilakukan pada

jangkauan efektif autokorelasi spasial yang diidentifikasi, misalnya, oleh semivariogram].

4. Metode *eigenvector spatial filtering* (ESF) memanfaatkan eigenvektor dari matriks bobot spasial $\boldsymbol{W}$ (Griffith, 2003).

1.5.6 *Eigenvector Spatial Filtering*

Eigenvector Spatial Fitering (ESF) merupakan pendekatan regesi spasial berbasis eigenvektor moran. ESF memanfaatkan eigenvektor dari matriks pembobot spasial untuk menangkap autokorelasi spasial (Tiefelsdorf dan Griffith, 2007). Matriks pembobot spasial menggambarkan hubungan spasial antar lokasi dan eigenvektor yang dihasilkan mencerminkan pola-pola spasial yang ada dalam data. Masing-masing eigenvektor mewakili tingkat autokorelasi spasial dengan syarat eigenvektor tersebut saling ortogonal ($\boldsymbol{E}'\boldsymbol{E} = \boldsymbol{I}$ dan $\boldsymbol{E}\boldsymbol{1}' = \boldsymbol{0}$) dan tidak berkorelasi ($\boldsymbol{E_i}$ dan $\boldsymbol{E_j} = 0$ ketika $i \neq j$). *Eigenvalue* menunjukkan besarnya variabilitas yang dijelaskan oleh eigenvektor tersebut (Griffith, 2000).

Metode ESF memiliki dua komponen utama yaitu mengekstraksi eigenvektor (E_n) dan mengidentifikasi subset dari eigevektor (E_l). ESF menggunakan dekomposisi spektral dari matriks pembobot spasial yang ditransformasikan ($\boldsymbol{W}$). Dekomposisi matriks sprektal $\boldsymbol{MWM}$ menghasilkan satu set n *eigenvalue* dan eigenvektor yang sesuai. Matriks $\boldsymbol{MWM}$ dapat ditunjukkan pada Persamaan (12).

$$\boldsymbol{MWM} = \boldsymbol{E}^{*}\boldsymbol{\Lambda}\boldsymbol{E}^{*^{-1}} = \boldsymbol{E}^{*}\boldsymbol{\Lambda}\boldsymbol{E}^{*^{T}} \tag{12}$$

dengan:

$\boldsymbol{\Lambda}$: Matriks diagonal n *eigenvalue* $(\lambda_1, \lambda_2, \dots, \lambda_n)$ berukuran $n \times n$

$\boldsymbol{E}^{*}$ **:** Matriks eigenvektor $(\boldsymbol{e}_1, \boldsymbol{e}, \dots, \boldsymbol{e_n})$ berukuran $n \times n$

$\boldsymbol{W}$: Matriks pembobot spasial berukuran $n \times n$

$\boldsymbol{M}$: Matriks pemusatan $\left(\boldsymbol{I} - \frac{\boldsymbol{11}'}{n}\right)$ berukuran $n \times n$

Identifikasi eigenvektor yang signifikan (E_l) dapat dilakukan dengan metode *stepwise selection* yang kemudian eigenvektor yang signifikan tersebut digunakan pada model ESF. Spesifikasi model ESF dalam regresi linier dapat dinyatakan pada Persamaan (13).

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{E\gamma} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2\mathbf{I}) \tag{13}$$

Keterangan:

$\boldsymbol{y}$: Vektor variabel dependen berukuran $n \times 1$

$\boldsymbol{X}$: Matriks variabel independen berukuran $n \times k$

$\boldsymbol{E}$: Matriks eigenvektor yang signifikan berukuran $n \times L$

$\boldsymbol{\varepsilon}$: Vektor galat berukuran $n \times 1$

$\boldsymbol{\beta}$: Vektor parameter variabel independen berukuran $k \times 1$

$\boldsymbol{\gamma}$: Vektor parameter eigenvektor berukuran $L \times 1$

σ^2 : Parameter varians

$\boldsymbol{0}$: Vektor nol berukuran $n \times 1$

1.5.7 *Random Effect Eigenvector Spatial Filtering*

Pengembangan metode *eigenvector spatial filtering* (ESF) dengan memperhitungkan efek acak (*random effect*) telah menarik perhatian, karena pendekatan RE-ESF ini bermanfaat untuk menganalisis isu dependensi spasial dengan memasukkan komponen proses spasial acak dalam modelnya (Murakami dan Griffith, 2015). Hughes dan Haran (2013) telah membahas metode *random effect eigenvector spatial filtering* (RE-ESF) dan mendemonstrasikan kegunaan pendekatan ini untuk menganalisis dependensi spasial. Metode RE-ESF mampu meminimalkan nilai ragam akibat faktor acak, yang terjadi antara variabel prediktor dan proses spasial laten. Selanjutnya, Murakami dan Griffith (2015) menunjukkan bahwa penggunaan ESF dengan efek acak dapat meningkatkan akurasi estimasi koefisien regresi, dengan waktu komputasi yang lebih singkat dibandingkan pendekatan lain. Secara matematis, model dasar dari ESF identik dengan model regresi linier biasa. Sementara itu, model linier dengan RE-ESF serupa dengan model linier campuran (*mixed model*). Model linier campuran merupakan perluasan dari model linier, di mana terdapat penambahan komponen efek acak dalam modelnya.

Berdasarkan Hughes dan Haran (2013), model linier RE-ESF ditunjukkan pada Persamaan (14).

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{E\gamma} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\gamma} \sim N\left(\boldsymbol{0}, \sigma_\gamma^2 \boldsymbol{\Lambda}\right), \qquad \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{14}$$

dengan σ_γ^2 adalah parameter ragam dan $\boldsymbol{\Lambda}$ adalah matriks diagonal berukuran $L \times L$ yang elemennya adalah L pertama *eigenvalue* dari $\boldsymbol{MWM}$, $\lambda_1, \dots, \lambda_L$, sesuai dengan L eigenvektor di $\boldsymbol{E}$. Persamaan (14) identik dengan model campuran linier (*linear mixed model*). Model campuran linier merupakan perluasan dari model linier yaitu dengan menambahkan efek acak.

1.5.8 *Maximum Likelihood Estimation*

Maximum Likelihood Estimation (MLE) merupakan metode parameter yang digunakan untuk menaksir parameter suatu model yang diketahui fungsi probabilitasnya. Metode MLE bekerja dengan memanfaatkan fungsi peluang dari parameter yang ditaksir. Misalkan $Y_1, Y_2, \dots, Y_n$ adalah variabel acak dengan fungsi peluang $f(y_i|\boldsymbol{\theta})$ dengan $\boldsymbol{\theta}$ yang merupakan parameter yang tidak diketahui. Karena $Y_1, Y_2, \dots, Y_n$ saling bebas, maka fungsi probabilitas bersama dari $Y_1, Y_2, \dots, Y_n$ dapat dinyatakan sebagai berikut.

$$f(Y_1, Y_2, \dots, Y_n|\boldsymbol{\theta}) = f(y_1|\boldsymbol{\theta}) \cdot f(y_2|\boldsymbol{\theta}) \cdot \dots \cdot f(y_n|\boldsymbol{\theta}) \tag{15}$$

Persamaan (15) dapat dituliskan ke dalam bentuk fungsi *likelihood* yang dinyatakan sebagai (Hogg dkk., 2012):

$$\begin{aligned} L(\boldsymbol{\theta}|y_i) &= f\left(y_1, y_2, \dots, y_n\middle|\boldsymbol{\theta}\right) \\ L(\boldsymbol{\theta}|y_i) &= f\left(y_1\middle|\boldsymbol{\theta}\right). f\left(y_2\middle|\boldsymbol{\theta}\right). \dots . f\left(y_n\middle|\boldsymbol{\theta}\right) \\ L(\boldsymbol{\theta}|y_i) &= \textstyle\prod_{i=1}^{n} f\left(y_i\middle|\boldsymbol{\theta}\right) \end{aligned} \tag{16}$$

Metode estimasi MLE dilakukan dengan memaksimalkan fungsi *likelihood* $L(\boldsymbol{\theta}|y_i)$. Fungsi *likelihood* dikerjakan dengan logaritma natural yang fungsinya merupakan fungsi monoton naik, sehingga nilai yang memaksimumkan fungsi $l(\boldsymbol{\theta}|y_i)$ sama dengan memaksimumkan fungsi $L(\boldsymbol{\theta}|y_i)$ atau dapat dituliskan seperti pada Persamaan (17).

$$l(\boldsymbol{\theta}|y_i) = \ln L(\boldsymbol{\theta}|y)$$
$$l(\boldsymbol{\theta}|y_i) = \ln\{\textstyle\prod_{i=1}^{n} f(y_i|\boldsymbol{\theta})\}$$
$$l(\boldsymbol{\theta}|y_i) = \textstyle\sum_{i=1}^{n} \ln f(y_i|\boldsymbol{\theta}) \tag{17}$$

Nilai parameter $\widehat{\boldsymbol{\theta}}$ dapat diperoleh dengan menurunkan $l(\boldsymbol{\theta}|y_i)$ terhadap $\boldsymbol{\theta}$ yang hasilnya disamakan dengan 0 yang ditulis pada Persamaan (18):

$$l'(\boldsymbol{\theta}|y_i) = \frac{\partial l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta}} = \boldsymbol{0} \tag{18}$$

1.5.9 *Restricted Maximum Likelihood*

Estimasi varians menggunakan MLE diketahui bersifat bias. Sebagai contoh, model linear regresi $\boldsymbol{y}\sim N\,(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I})$ dengan menggunakan MLE maka estimasi varians adalah $\hat{\sigma}^2 = SS/n$, dengan SS merupakan jumlah kuadrat residual dan n adalah jumlah observasi pengamatan. Untuk estimator tak bias (*unbiased*) suatu varians adalah $\frac{SS}{n-p}$ dimana p adalah jumlah parameter regresi yang digunakan dalam pemodelan. Oleh karena itu, untuk memperkecil bias dari suatu varians diperkenalkan metode estimasi yaitu *restricted maximum likelihood* (REML).

Contoh penggunaan REML ini untuk model linear umum $\boldsymbol{y}\sim N\,(\boldsymbol{X\beta}, \boldsymbol{V})$ dimana $\boldsymbol{V}$ merupakan matriks kovarian berukuran $n \times n$ dan dependen terhadap parameter θ. REML digunakan untuk memaksimumkan fungsi *log-likelihood* vektor residual, $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$, dengan $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$, dan untuk $\hat{\boldsymbol{e}}$ dan $\widehat{\boldsymbol{\beta}}$ saling independen. Untuk itu, fungsi *likelihood* untuk $\boldsymbol{y}$ adalah perkalian fungsi *likelihood* $\hat{\boldsymbol{e}}$ dan $\widehat{\boldsymbol{\beta}}$. Karena $\widehat{\boldsymbol{\beta}}\sim N\left(\boldsymbol{\beta}, ((\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1})\right)$ dan dapat disimpulkan bahwa fungsi *likelihood* untuk $\boldsymbol{y}$ adalah $L(\boldsymbol{y}, \boldsymbol{\theta}) = L(\,\hat{\boldsymbol{e}}, \boldsymbol{\theta})L\big(\widehat{\boldsymbol{\beta}}, \boldsymbol{\theta}\big)$.

Fungsi *log-likelihood* untuk REML menggunakan residual yaitu:

$$\log L(\,\hat{\boldsymbol{e}}, \boldsymbol{\theta}) = \log L(\boldsymbol{y}, \boldsymbol{\theta}) - \log L\big(\widehat{\boldsymbol{\beta}}, \boldsymbol{\theta}\big) \tag{19}$$

Namun, Laird dan Ware (1982) menggunakan cara alternatif untuk esimasi REML ini dengan menggunakan pendekatan Bayesian, yaitu:

$$L_R(\boldsymbol{\beta}, \sigma^2) = \int L(\boldsymbol{\beta}, \sigma^2)\, d\beta \tag{20}$$

dengan $L(\boldsymbol{\beta}, \sigma^2)$ merupakan fungsi *likelihood* MLE.

Sehingga fungsi *likelihood* REML untuk regresi linear adalah (Pangaribuan, 2014).

$$L_R(\boldsymbol{\beta}, \sigma^2) = (2\pi)^{-(n-p)}|\boldsymbol{V}|^{-\frac{1}{2}}|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta})\right) \tag{21}$$

1.5.10 Sifat Distribusi Normal Bivariat

Sepasang variabel acak kontinu X dan Y dikatakan memiliki Distribusi Normal Bivariat jika variabel acak kontinu X dan Y memiliki fungsi densitas gabungan dari bentuk sebagai berikut:

$$f(x,y) = \frac{1}{\sigma_x\sigma_y 2\pi(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\left(\frac{1}{2(1-\rho^2)}\right)\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\} \quad (22)$$

Untuk $-\infty < x < \infty$ dan $-\infty < y < \infty$ dengan $-1 < \rho < 1$

Dimana μ_x dan σ_x^2 adalah nilai harapan dan variansi dari variabel acak X, μ_y dan σ_y^2 adalah nilai harapan dan variansi dari variabel acak Y, dengan r adalah koefisien korelasi antara variaabel X dan Y. Fungsi Distribusi Normal Bivariat mempunyai lima parameter yaitu $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ dan ρ.

Teorema. Jika sepasang variabel acak kontinu Xdan Y adalah Distribusi Normal Bivariat dengan parameter $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2$ dan ρ maka $X \sim N\ (\mu_x, \sigma_x^2)$ dan $Y \sim (\mu_y, \sigma_y^2)$ dengan r adalah koefisien korelasi antara X dan Y.

Misalkan $g(x)$ adalah distribusi marginal X dan $h(y)$ adalah distribusi marginal Y. Dibuktikan bahwa $X \sim N\ (\mu_x, \sigma_x^2)$ dan $Y \sim (\mu_y, \sigma_y^2)$.

Bukti:

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$g(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\left(\frac{1}{2(1-\rho^2)}\right)\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\} dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\left(\frac{1}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right)\left[(x-\mu_x)^2\sigma_y^2 - 2\rho(x-\mu_x)(y-\mu_y)\sigma_x\sigma_y + (y-\mu_y)^2\sigma_x^2\right]\right\} dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2\sigma_y^2(1-\rho^2)}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\exp\left\{-\frac{\left[(y-\mu_y)\sigma_x - (x-\mu_x)\sigma_y\rho\right]^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right\}\right\} dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2\sigma_y^2(1-\rho^2)}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right\}\int_{-\infty}^{\infty}\exp\left\{-\frac{\left[(y-\mu_y)\sigma_x - (x-\mu_x)\sigma_y\rho\right]^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right\} dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{\left[(y-\mu_y)\sigma_x-(x-\mu_x)\sigma_y\rho\right]^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right\} dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{\left[y-\left(\mu_y+(x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho\right)\right]^2}{2(1-\rho^2)\sigma_y^2}\right\} dy$$

$$= \frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \int_{-\infty}^{\infty} \frac{1}{\sigma_y(2\pi)^{\frac{1}{2}}(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{\left[y-\left(\mu_y+(x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho\right)\right]^2}{2(1-\rho^2)\sigma_y^2}\right\} dy$$

$$= \frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \times 1$$

$$= \frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \qquad (23)$$

Maka $g(x)$ adalah distribusi normal dengan nilai harapan μ_x dan variansi σ_x^2.

Dengan Langkah yang sama didapatkan bahwa:

$$h(y) \quad = \frac{1}{\sigma_y(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}\right\} \qquad (24)$$

Maka $h(y)$ adalah distribusi normal dengan nilai harapan μ_y dan variansi σ_y^2.
Terbukti bahwa $X \sim N\ (\mu_x, \sigma_x^2)$ dan $Y \sim \left(\mu_y, \sigma_y^2\right)$.

Sepasang variabel acak X dan Y pada distribusi normal bivariat mempunyai dua hubungan yaitu variabel ack X dan Y yang independen dan dua variabel acak X dan Y yang tidak independent. Selanjutnya dikaji distribusi normal bivariat yang mempunyai variabel acak X dan Y yang independen.

$$f(x,y) \quad = \frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \times \frac{1}{\sigma_y(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}\right\}$$

$$= \frac{1}{\sigma_x\sigma_y 2\pi} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \exp\left\{-\frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}\right\}$$

$$= \frac{1}{\sigma_x\sigma_y 2\pi} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}\right\} \qquad (25)$$

Persamaan (25) diatas adalah fungsi distribusi normal bivariat yang mempunyai $\rho = 0$. Dengan demikian fungsi distribusi normal bivariat dengan X dan Y yang independent terbentuk dari $g(x) \sim N\ (\mu_x, \sigma_x^2)$ dan $h(y) \sim \left(\mu_y, \sigma_y^2\right)$.

Jika fungsi distribusi normal bivariat mempunyai $\rho \neq 0$, maka variable acak X dan Y tidak independent. Fungsi distribusi normal bivariat dengan $\rho \neq 0$ mempunyai fungsi bersyarat yang harus dipenuhi agar hasil kali dari $X \sim N\ (\mu_x, \sigma_x^2)$ dengan $Y \sim \left(\mu_y, \sigma_y^2\right)$ menghasilkan fungsi distribusi normal bivariat yang mempunyai $\rho \neq 0$. Untuk itu dikaji fungsi distribusi bersyarat Y dengan syarat $X =_X$.
Fungsi distribusi bersyarat Y dengan syarat $X =_X$ adalah sebagai berikut:

$$f(y|x) = \frac{f(x,y)}{g(x)}$$

$$f(y|x) = \frac{\frac{1}{\sigma_x\sigma_y 2\pi(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\left(\frac{1}{2(1-\rho^2)}\right)\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}}{\frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}}$$

$$= \frac{\frac{1}{\sigma_x\sigma_y 2\pi(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\left(\frac{(x-\mu_x)^2\sigma_y^2(1-\rho^2)}{2(1-\rho^2)\sigma_x^2\sigma_y^2} - \frac{\left[(y-\mu_y)\sigma_x(x-\mu_x)\sigma_y(1-\rho^2)\right]^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\right)\right\}}{\frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}}$$

$$= \frac{\frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{\left[y-\left(\mu_y+(x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho\right)\right]^2}{2(1-\rho^2)\sigma_y^2}\right\}}{\frac{1}{\sigma_x(2\pi)^{\frac{1}{2}}}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}}$$

$$= \frac{\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{\left[y-\left(\mu_y+(x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho\right)\right]^2}{2(1-\rho^2)\sigma_y^2}\right\}}{\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}(2\pi)^{\frac{1}{2}}(2\pi)^{\frac{1}{2}}} \times \frac{\sigma_x(2\pi)^{\frac{1}{2}}}{\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}}$$

$$= \frac{1}{\sigma_y\left((1-\rho^2)(2\pi)^{\frac{1}{2}}\right)^{\frac{1}{2}}}\exp\left\{-\frac{\left[y-\left(\mu_y+(x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho\right)\right]^2}{2(1-\rho^2)\sigma_y^2}\right\} \quad (26)$$

Dengan langkah sama didapatkan fungsi distribusi bersyarat X dengan syarat $Y =_Y$ (Turyadi dkk., 2013).

$$f(x|y) = \frac{1}{\sigma_x\left((1-\rho^2)(2\pi)^{\frac{1}{2}}\right)^{\frac{1}{2}}}\exp\left\{-\frac{\left[x-\left(\mu_x+(y-\mu_y)\frac{\sigma_x}{\sigma_y}\rho\right)\right]^2}{2(1-\rho^2)\sigma_x^2}\right\} \quad (27)$$

Dengan demikian $(y|x) \sim N\left[\mu_y + (x-\mu_x)\frac{\sigma_y}{\sigma_x}\rho, (1-\rho^2)\sigma_y^2\right]$ dan $(x|y) \sim N\left[\mu_x + (y-\mu_y)\frac{\sigma_x}{\sigma_y}\rho, (1-\rho^2)\sigma_x^2\right]$.

1.5.11 Metode Iterasi *Newton Raphson*

Metode *newton raphson* merupakan metode iterasi untuk menyelesaikan persamaan nonlinier. Metode ini membutuhkan matriks Hessian yang merupakan matriks dengan elemen-elemennya adalah turunan parsial orde dua dari fungsi *likelihood* terhadap masing-masing kombinasi parameter yang digunakan.

Persamaan iterasi *newton raphson* untuk parameter $\boldsymbol{\theta}$ dapat dinyatakan melalui Persamaan (28) (Agresti, 2012).

$$\widehat{\boldsymbol{\theta}}^{(s+1)} = \widehat{\boldsymbol{\theta}}^{(s)} - \boldsymbol{H}\left(\boldsymbol{\theta}^{(s)}\right)^{-1} \boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right), s = 0,1,2, \dots, q \quad (28)$$

Matriks Hessian parameter $\boldsymbol{\theta}$ pada iterasi ke-s dilambangkan dengan $\boldsymbol{H}(\boldsymbol{\theta}^{(s)})$ pada Persamaan (30).

$$\boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right) = \frac{\partial l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta}} \quad (29)$$

$$\boldsymbol{H}\left(\boldsymbol{\theta}^{(S)}\right) = \frac{\partial^2 l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \quad (30)$$

dengan:

$\widehat{\boldsymbol{\theta}}^{(s)}$	:	parameter yang diestimasi pada iterasi ke-s sebanyak q iterasi
$\boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right)$	:	Vektor gradien parameter $\boldsymbol{\theta}$ pada iterasi ke-s
$\boldsymbol{H}\left(\boldsymbol{\theta}^{(s)}\right)$	:	Matriks hessian parameter $\boldsymbol{\theta}$ pada iterasi ke-s

Adapun langkah-langkah dalam algoritma *Newton Raphson* adalah sebagai berikut:

1. Menentukan taksiran awal parameter $\widehat{\boldsymbol{\theta}}^{(0)}$ yang diperoleh dari estimasi model regresi *Poisson Log-Normal*.
2. Menghitung vektor gradien $\boldsymbol{g}(\boldsymbol{\theta})$.
3. Menghitung matriks *Hessian* $\boldsymbol{H}(\boldsymbol{\theta})$.

Melakukan iterasi menggunakan Persamaan (28) sampai diperoleh nilai yang konvergen, yaitu ketika $\left|\widehat{\boldsymbol{\theta}}^{(s+1)} - \widehat{\boldsymbol{\theta}}^{(s)}\right| \leq \varepsilon$, dengan ε adalah sebesar 0,001.

1.5.12 Uji Signifikansi Parameter

Pengujian signifikansi parameter dilakukan untuk mengetahui variabel independen yang berpengaruh secara signifikan terhadap variabel dependen (Azizah, 2023).

1. Hipotesis

H_0 : $\beta_k = 0$ (koefisien variabel independen ke- k tidak signifikan)
H_1 : $\beta_k \neq 0$ untuk $k = 1,2, \dots, p$ (keofisien variabel independen ke-k signifikan)

2. Statistik uji

$$t = \frac{b_k}{s(b_k)} \quad (31)$$

dengan:

b_k : Penduga koefisien variabel ke-k
$s(b_k)$: Standar deviasi untuk penduga koefisien variabel ke-k

3. Kriteria uji

Jika $|t| > t_{(\alpha,n-p)}$ maka H_0 ditolak.

Jika $|t| < t_{(\alpha,n-p)}$ maka H_0 diterima.

1.5.13 Pemilihan Model Terbaik

a. *Akaike's Information Criterion*

AIC merupakan kriteria yang dapat mengukur kesesuaian model dalam menduga parameter model secara statistik dan mengetahui seberapa dekat parameter yang diduga dengan nilai populasi yang sebenarnya. Suatu model regresi dikatakan model terbaik jika memiliki nilai AIC terkecil. Rumus untuk menghitung nilai AIC adalah sebagai berikut (Anselin, 1988).

$$AIC = -2\ln L(\boldsymbol{\beta}) + 2k \tag{32}$$

dengan:

k : Jumlah parameter yang digunakan

$L(\boldsymbol{\beta})$: Maksimum logaritma *likelihood*

b. *Adjusted-R^2*

Adjusted-R^2 digunakan untuk mengetahui besarnya nilai dari variabel dependen yang dapat dijelaskan oleh variasi variabel independen dan sisanya merupakan bagian nilai dari varibel lain yang tidak termasuk di dalam model. Nilai *Adjusted*-R^2 adalah 0 sampai 1, jika nilai mendekati 1 artinya variabel indepen mempu memberikan hamper semua informasi yang dibutuhkan untuk memprediksi variabel dependen dan sebaliknya jika nilai mendekati 0 maka kemampuan variabel independen untuk mempresikdi variabel dependen sangat terbatas (Ghozali, 2018). Rumus R^2_{adj} adalah sebagai berikut:

$$R^2_{adj} = 1 - \left[(1 - R^2)\left(\frac{n-1}{n-p-1}\right)\right] \tag{33}$$

1.5.14 Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka (TPT) adalah nilai yang menunjukkan persentase jumlah pengangguran terbuka terhadap jumlah angkatan kerja. TPT menggambarkan proporsi angkatan kerja yang tidak memiliki pekerjaan dan secara aktif mencari dan bersedia untuk bekerja. Pengangguran terbuka adalah orang yang termasuk ke dalam kategori seperti mereka yang tidak memiliki pekerjaan dan mencari pekerjaan kegiatan, mereka yang tidak memiliki pekerjaan dan mempersiapkan usaha, mereka yang tidak memiliki pekerjaan dan tidak mencari pekerjaan, karena merasa tidak mungkin mendapatkan pekerjaan, dan mereka yang sudah memiliki pekerjaan, tetapi belum mulai bekerja.

Menurut Sukirno (2011) pengangguran terbuka adalah pengangguran yang tercipta akibat adanya pertambahan lowongan pekerjaan yang lebih rendah daripada pertambahan tenaga kerja. Efek jangka panjang yang ditimbulkan adalah masyarakat tidak melakukan suatu pekerjaan atau menganggur secara nyata dan sepenuh waktu. Pengangguran terbuka juga dapat tercipta akibat adanya penurunan kegiatan ekonomi, kemajuan teknologi atau tenaga manusia digantikan oleh tenaga mesin, serta kemunduran perkembangan industri.

Berdasarkan data dari BPS, “mencari pekerjaan” adalah kegiatan seseorang yang sedang mencari pekerjaan pada saat dilakukan survei, yakni orang yang belum pernah bekerja dan sedang berusaha mendapatkan pekerjaan, orang yang sudah pernah bekerja namun karena suat hal, orang tersebut berhenti atau diberhentikan dan sedang berusaha untuk mendapatkan pekerjaan. Kegiatan “mempersiapkan usaha” adalah suatu kegiatan yang dilakukan oleh seseorang dalam rangka mempersiapkan suatu pekerjaan baru, yang bertujuan untuk memperoleh keuntungan atas resiko sendiri, baik dengan dan/atau tanpa mempekerjakan karyawan dibayar maupun tidak dibayar.

Nanga (2001) menjelaskan bahwa Edgar Edwards membedakan jenis pengangguran, khususnya yang sering terjadi di negara berkembang ke dalam bentuk sebagai berikut:

1. Setengah Pengangguran adalah para pekerja yang jumlah jam kerjanya lebih sedikit dari yang sebenarnya mereka inginkan (sebagian besar hanya bekerja harian, mingguan atau musiman).
2. Pengangguran Terbuka atau *open unemployement* adalah mereka yang benar-benar sedang tidak bekerja baik secara sukarela maupun karena terpaksa dan secara aktif sedang mencari pekerjaan (Prasanti, 2015).

Pertumbuhan ekonomi juga merupakan salah satu faktor yang mempengaruh pengangguran terbuka. Pertumbuhan ekonomi yang tinggi akan berdampak pada penyerapan tenaga kerja yang artinya jumlah pengangguran akan berkurang. Pertumbuhan ekonomi dapat dihitung berdasarkan perubahan jumlah Produk Domestik Regional Bruto. PDRB adalah jumlah nilai tambah atas barang dan jasa yang dihasilkan oleh berbagai unit produksi pada suatu wilayah dalam jangka waktu tertentu (Kasanah dkk., 2018).

Simanjuntak dalam (Prakoso, 2020) menjelaskan adanya hubungan antara pendidikan dan pengangguran, bahwa peningkatan tingkat pendidikan di suatu daerah akan memberikan kontribusi positif terhadap peningkatan kualitas sumber daya manusia yang berada di wilayah tersebut, yang pada gilirannya meningkatkan peluang mendapatkan pekerjaan. Selain itu, Muslim (2014) menambahkan bahwa semakin banyak waktu yang diinvestasikan oleh masyarakat dalam mengejar pendidikan, maka semakin tinggi pula kualitas atau reputasi pekerjaan yang bisa mereka dapatkan, dan juga semakin minim kemungkinan mereka mengalami masalah pengangguran. Orang-orang yang memiliki pendidikan tinggi umumnya memiliki berbagai keterampilan dan keahlian yang beragam, sehingga peluang mendapatkan pekerjaan meningkat, dan masalah pengangguran dapat ditekan (Al-faridzi dkk., 2023).

BAB II
METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari situs resmi Badan Pusat Statistik Sulawesi Selatan www.sulsel.bps.go.id berupa data Tingkat Pengangguran Terbuka, PDRB, Rata-rata Lama Sekolah dan Tingkat Setengah Pengangguran di 24 Kabupaten/Kota Provinsi Sulawesi Selatan Tahun 2022.

2.2 Variabel Penelitian

Variabel yang digunakan pada penelitian ini merupakan faktor pertumbuhan ekonomi dan tenaga kerja yang diduga berpengaruh terhadap Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan. Variabel penelitian ini terbagi dalam dua variabel yaitu, variabel respon (Y) dan variabel prediktor (X). Adapun rincian variabel dalam penelitian ini disajikan dalam Tabel 1.

Tabel 1. Variabel Penelitian

No	Variabel	Keterangan	Deskripsi
1	Y	Tingkat Pengangguran Terbuka	Persentase pengangguran terhadap jumlah angkatan kerja. Angkatan kerja adalah penduduk usia 15-65.
2	X_1	Produk Domestik Regional Bruto (PDRB)	Jumlah nilai tambah atas barang dan jasa yang dihasilkan oleh berbagai unit produksi pada suatu wilayah dalam jangka waktu tertentu.
3	X_2	Rata-rata Lama Sekolah	Indikator yang digunakan untuk mengukur rata-rata jumlah tahun yang dihabiskan oleh penduduk usia 15 tahun ke atas untuk menempuh pendidikan formal.
4	X_3	Tingkat Setengah Pengangguran	Penduduk yang bekerja di bawah jam kerja normal (kurang dari 35 jam seminggu), dan masih mencari pekerjaan atau masih bersedia menerima pekerjaan (setengah pengangguran terpaksa).

2.3 Prosedur Penelitian

Adapun tahapan analisis yang dilakukan untuk data Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan pada Tahun 2022 adalah sebagai berikut:

a. Eksplorasi Data

Melakukan eksplorasi data menggunakan peta tematik dari sudut kewilayahannya untuk melihat sebaran data Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan Tahun 2022.

b. Pemodelan Regresi Linier Berganda

1. Melakukan estimasi parameter regresi linier berganda pada Persamaan (3).
1. Menentukan model regresi linier berganda.
2. Uji signifikansi parameter model regresi linier berganda secara parsial.

c. Matriks Pembobot Spasial

Analisis spasial diperlukan pembobotan untuk menunjukkan ketetanggaan antar lokasi. Matriks pembobot spasial yang digunakan dalam penelitian ini adalah matriks pembobot fungsi *Exponential* pada Persamaan (6).

d. Uji Autokorelasi Spasial

Menguji efek autokorelasi spasial pada data dengan Indeks Moran. Menghitung *Moran's I* dapat menggunakan persamaan (13) dan untuk pengujian statistiknya menggunakan Persamaan (9). Jika $p-value < \alpha$ maka terdapat autokorelasi spasial antar lokasi.

e. Pemodelan ESF

1. Dekomposisi eigenvektor $(\boldsymbol{MWM})$ berdasarkan matriks pembobot spasial pada Persamaan (12).
2. Menentukan L eigenvektor dari matriks pembobot spasial dengan menghilangkan eigenvektor yang *eigenvalue* kecil.
3. Melakukan estimasi parameter model ESF dengan metode MLE pada Persamaan (37) dan Persamaan (38).
4. Menentukan model linier dasar dari ESF.
5. Uji signifikansi parameter model ESF secara parsial pada Persamaan (31).

f. Pemodelan RE-ESF

1. Menentukan L eigenvektor dari matriks pembobot spasial dengan menghilangkan eigenvektor yang *eigenvalue* kecil.
2. Melakukan estimasi parameter model RE-ESF dengan metode REML pada Persamaan (50) dan Persamaan (61).
3. Menetukan model linier dasar RE-ESF.
4. Uji signifikansi parameter model RE-ESF secara parsial pada Persamaan (31).

g. Pemilihan Model Terbaik

Melakukan evaluasi pendugaan parameter model ESF dan RE-ESF berdasarkan nilai *Akaike's Information Criterion* (AIC) terkecil atau *Adjusted*-R^2 berdasarkan nilai terbesar.