

**ANALISIS KETEPATAN *HETEROSKEDASTICITY-CONSISTENT*
COVARIANCE MATRIX ESTIMATOR (HCCME) TERHADAP
PELANGGARAN ASUMSI HOMOSKEDASTISITAS MELALUI SIMULASI
MONTE CARLO**

Skripsi

Oleh

**WINNE PASCILIA NAPITU
NPM. 2217031176**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

ABSTRACT

ANALYSIS OF THE ACCURACY OF HETEROSKEDASTICITY-CONSISTENT COVARIANCE MATRIX ESTIMATOR (HCCME) IN RESPONSE TO HOMOSKEDASTICITY ASSUMPTION VIOLATION THROUGH MONTE CARLO SIMULATION

By

Winne Pascilia Napitu

Violation of the homoskedasticity assumption in a linear regression model can cause the OLS covariance estimator to produce inaccurate standard errors. This condition may affect statistical inference, especially hypothesis testing and confidence interval construction. This study aims to analyze the accuracy of the Heteroskedasticity-Consistent Covariance Matrix Estimator (HCCME), namely HC0, HC1, HC2, and HC3, and to compare their performance with the OLS covariance estimator under heteroskedasticity. This study uses a Monte Carlo simulation based on a simple linear regression model. The explanatory variable is generated from a normal distribution, while the error term is formed with nonconstant variance. The simulation is conducted for sample sizes of 30, 50, 100, and 250 with 1000 replications for each scenario. Estimator performance is evaluated based on bias, Mean Squared Error (MSE), and coverage probability. The results show that the OLS covariance estimator tends to estimate the variance of regression parameters lower than the true variance. This is indicated by negative bias values and coverage probability values that are far below the nominal level of 95%. In contrast, the HCCME estimators produce covariance estimates that are closer to the Monte Carlo variance. The bias and MSE values of the HCCME estimators tend to decrease as the sample size increases. Based on coverage probability, HC3 provides the best performance because it produces values closest to the nominal level of 95%. Therefore, HCCME is more appropriate than the OLS covariance estimator when the homoskedasticity assumption is violated.

Keywords: HCCME, Heteroskedasticity, OLS, Monte Carlo simulation.

ABSTRAK

ANALISIS KETEPATAN *HETEROSKEDASTICITY-CONSISTENT COVARIANCE MATRIX ESTIMATOR* (HCCME) TERHADAP PELANGGARAN ASUMSI HOMOSKEDASTISITAS MELALUI SIMULASI MONTE CARLO

Oleh

Winne Pascilia Napitu

Pelanggaran asumsi homoskedastisitas dalam model regresi linear dapat menyebabkan estimator kovarians OLS menghasilkan standard error yang tidak akurat. Kondisi ini dapat memengaruhi hasil inferensi statistik, terutama dalam pengujian hipotesis dan pembentukan interval kepercayaan. Penelitian ini bertujuan untuk menganalisis ketepatan estimator *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME), yaitu HC0, HC1, HC2, dan HC3, serta membandingkan kinerjanya dengan estimator kovarians OLS pada kondisi heteroskedastisitas. Penelitian ini menggunakan simulasi Monte Carlo dengan model linear sederhana. Peubah penjelas dibangkitkan dari distribusi normal, sedangkan galat dibentuk dengan varians tidak konstan. Simulasi dilakukan pada ukuran sampel 30, 50, 100, dan 250 dengan 1000 replikasi untuk setiap skenario. Kinerja estimator dievaluasi berdasarkan bias, *Mean Squared Error* (MSE), dan *coverage probability*. Hasil penelitian menunjukkan bahwa estimator kovarians OLS cenderung mengestimasi varians parameter regresi lebih kecil daripada varians sebenarnya. Hal ini terlihat dari nilai bias negatif dan *coverage probability* yang berada jauh di bawah tingkat nominal 95%. Sebaliknya, estimator HCCME menghasilkan estimasi kovarians yang lebih mendekati varians Monte Carlo. Nilai bias dan MSE estimator HCCME cenderung menurun seiring bertambahnya ukuran sampel. Berdasarkan *coverage probability*, HC3 memberikan kinerja paling baik karena menghasilkan nilai yang paling mendekati tingkat nominal 95%. Dengan demikian, HCCME lebih tepat digunakan daripada estimator kovarians OLS ketika terjadi pelanggaran asumsi homoskedastisitas.

Kata kunci: HCCME, Heteroskedastisitas, OLS, Simulasi Monte Carlo.

**ANALISIS KETEPATAN *HETEROSKEDASTICITY-CONSISTENT*
COVARIANCE MATRIX ESTIMATOR (HCCME) TERHADAP
PELANGGARAN ASUMSI HOMOSKEDASTISITAS MELALUI SIMULASI
MONTE CARLO**

WINNE PASCILIA NAPITU

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

Judul Skripsi : **ANALISIS KETEPATAN
HETEROSKEDASTICITY-CONSISTENT
COVARIANCE MATRIX ESTIMATOR
(HCCME) TERHADAP PELANGGARAN
ASUMSI HOMOSKEDASTISITAS
MELALUI SIMULASI MONTE CARLO**

Nama Mahasiswa : **Winne Pascilia Napitu**

Nomor Pokok Mahasiswa : **2217031176**

Program Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing

A handwritten signature in black ink, appearing to read 'Mustaf'.

Prof. Drs. Mustofa Usman, M.A., Ph.D.
NIP 195701011984031020

A handwritten signature in black ink, appearing to read 'Bernadhita'.

Dr. Bernadhita Herindri S Utami, S.Si., M.Sc.
NIP 199206302023212034

2. Ketua Jurusan Matematika

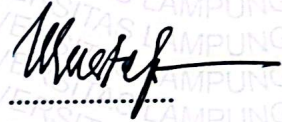
A handwritten signature in black ink, appearing to read 'Aang Nuryaman'.

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua : Prof. Drs. Mustofa Usman, M.A.,
Ph.D.



Sekretaris : Dr. Bernadhita Herindri S Utami,
S.Si., M.Sc.



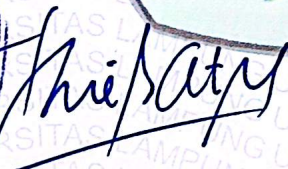
Penguji
Bukan Pembimbing : Widiarti, S.Si., M.Si.



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002



Tanggal Lulus Ujian Skripsi: 18 Juni 2026

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Winne Pascilia Napitu
Nomor Pokok Mahasiswa : 2217031176
Jurusan : Matematika
Judul Skripsi : Analisis Ketepatan *Heteroskedasticity - Consistent Covariance Matrix Estimator* (HCCME) Terhadap Pelanggaran Asumsi Homoskedastisitas Melalui Simulasi Monte Carlo

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 18 Juni 2026

Penulis,



Winne Pascilia Napitu

RIWAYAT HIDUP

Penulis memiliki nama lengkap Winne Pascilia Napitu, lahir di Kota Batam pada tanggal 03 Mei 2004. Penulis merupakan anak ketiga dari lima bersaudara, putri dari pasangan Bapak Bagaolan Damianus Napitu dan Ibu Rusti Br Tanjung.

Penulis menempuh pendidikan dimulai dari SD Kristen Immanuel Batam, Kepulauan Riau pada tahun 2010 sampai tahun 2016. Kemudian, penulis melanjutkan pendidikan sekolah menengah pertama di SMP Negeri 31 Batam, Kepulauan Riau pada tahun 2016 sampai tahun 2019. Kemudian, penulis melanjutkan pendidikan sekolah menengah atas di SMA Negeri 21 Batam, Kepulauan Riau pada tahun 2019 sampai tahun 2022.

Pada tahun 2022, penulis diterima sebagai mahasiswi di program studi S1 Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN). Penulis juga merupakan penerima Bantuan Beasiswa KIP-Kuliah. Selama masa perkuliahan, penulis juga aktif mengikuti Unit Kegiatan Mahasiswa (UKM) Saintek Universitas Lampung dan aktif juga mengikuti kepanitiaan sebagai koordinator bidang pdd.

Selama menempuh pendidikan, penulis menyelesaikan Kerja Praktik (KP) di Pemerintah Provinsi Lampung selama 40 hari dari Desember 2024 sampai dengan Februari 2025. Selain itu, penulis juga melaksanakan program Merdeka Belajar Kampus Merdeka (MBKM) Magang Berdampak selama 4 bulan pada bulan Maret-Juni 2025 di PT PLN BATAM, Kota Batam, Kepulauan Riau.

KATA INSPIRASI

”Karena masa depan sungguh ada, dan harapanmu tidak akan hilang.”
(Amsal 23:18)

”Sebab Aku ini mengetahui rancangan-rancangan apa yang ada pada-Ku mengenai kamu, demikianlah firman TUHAN, yaitu rancangan damai sejahtera dan bukan rancangan kecelakaan, untuk memberikan kepadamu hari depan yang penuh harapan.”
(Yeremia 29:11)

”if you can’t beat the fear, just do it scared. just do it. ”
(Glennon Doyle)

*”You’re on your own, kid. Yeah, you can face this.
You’re on your own, kid. You always have been ”*
(Taylor Swift)

PERSEMBAHAN

Dengan mengucapkan puji dan syukur atas kehadiran Tuhan Yang Maha Esa karena limpahan rahmat dan karunia-Nya sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya.

Dengan rasa syukur dan Bahagia, saya persembahkan rasa terimakasih saya kepada:

Bapak dan Ibuku Tercinta

Terimakasih kepada orang tuaku atas segala pengorbanan, motivasi, doa serta dukungannya selama masa perkuliahan sampai selesainya masa skripsi ini. Terimakasih telah memberikan pelajaran berharga tentang makna perjalanan hidup yang sebenarnya. Terimakasih sudah menjadi rumah dan tempat berpulang dari segala keluh kesah akan hal-hal pelik yang datang.

Kakak Tercinta

Terimakasih kepada kak Sari Hanna Angelina yang selalu memberikan dukungan secara moral dan material selama penulis menempuh pendidikan. Terimakasih atas kebaikan yang sudah diberikan kepada penulis.

Dosen Pembimbing dan Pembahas

Terimakasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan arahan serta ilmu yang berharga.

Sahabat-sahabatku

Terimakasih kepada semua orang-orang baik yang telah memberikan pengalaman, semangat, motivasinya, serta doa-doanya dan senantiasa memberikan dukungan dalam hal apapun.

Almamater Tercinta

Universitas Lampung

SANWACANA

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "Analisis Ketepatan *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) Terhadap Pelanggaran Asumsi Homoskedastisitas melalui Simulasi Monte Carlo" dengan baik dan lancar serta tepat pada waktu yang telah ditentukan.

Dalam proses penyusunan skripsi ini, banyak pihak yang telah membantu memberikan bimbingan, dukungan, arahan, motivasi serta saran sehingga skripsi ini dapat terselesaikan. Oleh karena itu, dalam kesempatan ini penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D. selaku Pembimbing 1 yang telah banyak meluangkan waktunya untuk memberikan arahan, bimbingan, motivasi, saran serta dukungan kepada penulis sepanjang proses penyusunan skripsi ini.
2. Ibu Bernadhita Herindri Samodera Utami, M.Sc. selaku Pembimbing 2 yang telah memberikan arahan, serta dukungan sehingga penulis dapat menyelesaikan skripsi ini.
3. Ibu Widiarti, S.Si., M.Si. selaku Penguji yang telah bersedia memberikan saran, kritik, serta evaluasi yang membangun sehingga penulis dapat menyelesaikan skripsi ini.
4. Ibu Dina Eka Nurvazly, S.Pd., M.Si. selaku dosen Pembimbing Akademik.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh dosen, staff dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

7. Keluarga tercinta, yaitu Bapak, Ibu, kedua kakak penulis yaitu Kak Yanti dan Kak angel, dan kedua adik penulis yaitu May darnita dan Jordan yang senantiasa menjadi sumber doa, kasih sayang, dan motivasi. Tanpa dukungan, kesabaran, dan cinta yang tak pernah berhenti dari kalian, penulis tidak akan mampu menyelesaikan skripsi ini. Terima kasih atas segala pengorbanan dan perhatian yang selalu menguatkan langkah penulis hingga hari ini.
8. Teman seperjuangan sirkel Pomp op Asep, diantaranya Yolan, Ayu, Lusi, Gracia, Mega, Elisabeth, Rafael, Agustino dan Niluh yang telah menemani penulis sejak semester pertama hingga saat ini. Terima kasih atas kebersamaan, tawa, dan dukungan yang saling menguatkan yang kalian bagikan selama masa perkuliahan.
9. Teman seperjuangan dari awal Kelas D, diantaranya Puan dan Vira. Terima kasih atas, dukungan, kerja sama, canda tawa, serta berbagai pengalaman yang telah dilalui bersama selama masa perkuliahan.
10. Teman seiman penulis, diantaranya Betesdha dan Ike Juliana. Terimakasih karena sudah menjadi teman berbagi cerita, tempat saling menguatkan, dan memberi semangat selama penulis menyelesaikan skripsi ini.
11. Teman-teman Batam penulis, diantaranya Geby, Shonalia, Debora, Selin, Manda, dan Priska. Terima kasih sudah menjadi teman yang baik selama ini.
12. Rekan seperjuangan skripsi sepembimbingan Pak Mustofa, terkhusus Nada dan Tri Harsono yang selalu membantu dan saling memberikan dukungan satu sama lain. Terimakasih sudah bersama-sama menyelesaikan ini.
13. Teman-teman seperjuangan Jurusan Matematika angkatan 2022.

Semoga skripsi ini dapat bermanfaat bagi kita semua. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, sehingga penulis mengharapkan kritik dan saran yang membangun untuk menjadikan skripsi ini lebih baik lagi.

Bandar Lampung, 18 Juni 2026

Winne Pascilia Napitu

DAFTAR ISI

DAFTAR ISI	xiii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	3
II TINJAUAN PUSTAKA	4
2.1 Model Linear	4
2.2 Estimator <i>Ordinary Least Squares</i> (OLS)	5
2.2.1 Sifat-Sifat <i>Estimator Ordinary Least Squares</i> (OLS)	7
2.2.2 Varians dan Kovarians <i>Ordinary Least Squares</i> (OLS)	8
2.3 Teorema Gauss Markov	10
2.4 Pelanggaran Asumsi Homoskedastisitas	11
2.4.1 Distribusi Normal pada Pembangkitan Galat dalam Model Linear	13
2.4.2 Pembentukan Galat Heteroskedastik dalam Simulasi	14
2.5 <i>Heteroskedasticity-Consistent Covariance Matrix Estimator</i> (HCCME)	16
2.5.1 Jenis–Jenis Estimator HCCME	17
2.6 Simulasi Monte Carlo	21
III METODOLOGI PENELITIAN	23
3.1 Waktu dan Tempat Penelitian	23
3.2 Data Penelitian	23
3.3 Metode Penelitian	24
IV HASIL DAN PEMBAHASAN	30
4.1 Pembentukan Data Simulasi berdasarkan Matriks Desain dan Galat Heteroskedastik	30
4.2 Indikasi Pola Heteroskedastisitas pada Data Simulasi	33
4.3 Estimasi Parameter OLS dan Pembentukan Estimator Kovarians HCCME	36

4.4	Varians Monte Carlo	39
4.5	Analisis Kinerja Estimator berdasarkan Simulasi Monte Carlo . . .	41
4.5.1	Bias Estimator Kovarians OLS dan HCCME	41
4.5.2	<i>Mean Squared Error</i> Estimator Kovarians OLS dan HCCME	43
4.5.3	<i>Coverage Probability</i> Estimator Kovarians	45
V	KESIMPULAN DAN SARAN	48
5.1	Kesimpulan	48
5.2	Saran	49
	DAFTAR PUSTAKA	50

DAFTAR TABEL

1	Bentuk ω_i dari setiap estimator HC	18
1	Hasil Estimasi Parameter pada Seluruh Replikasi	37
2	Hasil Rata-Rata Estimator Kovarians $\widehat{Var}(\hat{\beta}_1)$ pada seluruh Replikasi	38
3	Varians Monte Carlo dari estimator $\hat{\beta}_1$ untuk setiap ukuran sampel .	40
4	Nilai Bias Estimator Kovarians berdasarkan Ukuran Sampel	42
5	<i>Mean Squared Error</i> (MSE) Estimator Kovarians OLS dan HCCME	44
6	Nilai <i>Coverage Probability</i> Estimator Kovarians berdasarkan Ukuran Sampel	46

DAFTAR GAMBAR

1	Kondisi Varians Homoskedastisitas dan Heteroskedastisitas	12
2	Diagram Alur Penelitian	29
3	Pola Galat terhadap Peubah Penjelas	34
4	Pola Heteroskedastisitas Residual	34
5	Pola Heteroskedastisitas Residual Kuadrat	35
6	Grafik Bias Estimator Kovarians terhadap Ukuran Sampel	43
7	Grafik MSE Estimator Kovarians terhadap Ukuran Sampel	45
8	Grafik <i>Coverage Probability</i> pada Berbagai Ukuran Sampel	46

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam analisis regresi linear, asumsi mengenai struktur galat merupakan bagian yang sangat penting untuk menjamin validitas inferensi statistik. Salah satu asumsi klasik yang harus dipenuhi agar estimator *Ordinary Least Square* (OLS) memiliki sifat *Best Linear Unbiased Estimator* (BLUE) berdasarkan teorema Gauss-Markov adalah asumsi homoskedastisitas, yaitu varian galat yang konstan pada seluruh pengamatan. Sebagaimana dijelaskan Rancher *and* Schaalje (2008), ketidakstabilan varian galat menyebabkan penduga varian OLS menjadi tidak akurat sehingga mengganggu proses inferensi.

Namun, dalam banyak aplikasi nyata asumsi homoskedastisitas sering kali tidak terpenuhi. Varians galat yang berubah terhadap nilai variabel penjelas atau faktor lain dikenal sebagai heteroskedastisitas. Kondisi ini dapat muncul akibat berbagai faktor seperti perubahan skala pengukuran, perbedaan karakteristik unit observasi, hingga model yang disalah spesifikasi. Ketika heteroskedastisitas terjadi, OLS tetap memberikan penduga koefisien yang tidak bias, tetapi penduga varians dari koefisien menjadi tidak konsisten, sehingga standard error dan uji statistik yang dihasilkan menjadi bias (Hayes *and* Cai , 2007). Dengan kata lain, masalah utama dari heteroskedastisitas bukan pada nilai koefisien regresi, tetapi pada standard error yang menjadi dasar pengujian hipotesis dan interval kepercayaan.

Sebagai Solusi, White (1980) memperkenalkan pendekatan yang kini dikenal sebagai *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME). Pendekatan ini merupakan metode penduga matriks kovarians parameter regresi yang tetap konsisten meskipun varians galat tidak konstan. Secara umum, HCCME bekerja dengan menggantikan matriks varians galat yang tidak diketahui dengan fungsi residual kuadrat hasil estimasi OLS, sehingga diperoleh penduga kovarians parameter yang robust terhadap heteroskedastisitas.

Seiring perkembangan penelitian, pendekatan HCCME tidak hanya terbatas pada bentuk awal yang diperkenalkan oleh White (1980). MacKinnon and White (1985), serta peneliti lain seperti Davidson and MacKinnon (1993) mengembangkan beberapa modifikasi terhadap estimator tersebut untuk memperbaiki performa pada kondisi sampel kecil dan adanya pengaruh *leverage*. Pengembangan ini menghasilkan beberapa varian estimator yang masih berada dalam kerangka HCCME, yaitu HC0, HC1, HC2 dan HC3.

Keempat estimator tersebut pada dasarnya merupakan bentuk khusus dari HCCME yang dibedakan oleh cara menduga matriks varians galat menggunakan residual OLS dan faktor koreksi tertentu. HC0 merupakan bentuk dasar HCCME yang secara langsung menggunakan residual kuadrat sebagai penduga varians galat. HC1 merupakan modifikasi HC0 dengan penyesuaian derajat bebas untuk mengurangi bias sampel pada sampel kecil. Selanjutnya, HC2 dan HC3 dikembangkan dengan mempertimbangkan nilai *leverage* pada setiap observasi, sehingga mampu memberikan koreksi yang lebih baik terhadap pengaruh observasi berpengaruh tinggi. Dengan demikian, dapat dikatakan bahwa HC0, HC1, HC2 dan HC3 merupakan anggota dalam keluarga estimator HCCME yang memiliki tujuan sama, yaitu menghasilkan penduga kovarians parameter regresi yang konsisten dalam kondisi heteroskedastisitas, namun dengan karakteristik dan tingkat koreksi yang berbeda.

Perbedaan karakteristik masing-masing estimator tersebut menunjukkan bahwa, tidak terdapat satu estimator HCCME yang selalu unggul di semua kondisi. Efektivitas setiap estimator bergantung pada ukuran sampel, struktur matriks desain, serta tingkat dan pola heteroskedastisitas yang terjadi. Oleh karena itu, diperlukan kajian empiris untuk membandingkan performa masing-masing estimator dalam berbagai kondisi data. Salah satu pendekatan yang umum digunakan dalam penelitian metodologis adalah simulasi Monte Carlo. Melalui simulasi ini, peneliti dapat membangkitkan data sesuai model tertentu, menetapkan bentuk heteroskedastisitas yang jelas, menerapkan berbagai estimator kovarians, dan mengukur performanya menggunakan ukuran seperti Bias, *Mean Squared Error* (MSE), serta *standard error* rata-rata (Flachaire, 2005).

Heteroskedastisitas yang diabaikan dapat mengakibatkan kesimpulan yang menyesatkan, terutama pada penelitian sosial ekonomi dan eksperimental yang sering melibatkan data dengan variabilitas tinggi antar unit observasi. Oleh karena itu, penilaian terhadap performa dan ketepatan berbagai estimator HCCME menjadi hal yang krusial untuk memastikan keandalan hasil analisis.

Beberapa penelitian terdahulu seperti Hayes *and* Cai (2007), Nwangburuka *et al.*, (2023), Simsek *and* Orhan (2016), Kranz (2024) dan Majid *et al.*, (2017) telah mengkaji performa estimator HCCME melalui pendekatan simulasi maupun aplikasi pada data nyata. Hasil penelitian tersebut menunjukkan bahwa setiap varian estimator HCCME memiliki tingkat bias dan *Root Mean Squared Error* (RMSE) yang berbeda pada berbagai pola heteroskedastisitas dan ukuran sampel.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis ketepatan beberapa estimator dalam keluarga HCCME, yaitu HC0, HC1, HC2, dan HC3 dalam mengatasi pelanggaran asumsi homoskedastisitas. Analisis ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang perilaku setiap estimator dalam kondisi heteroskedastisitas.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Menganalisis ketepatan estimator HCCME (HC0, HC1, HC2, HC3) dalam kondisi heteroskedastisitas.
2. Membandingkan performa estimator HCCME dengan estimator kovarians OLS.
3. Menilai pengaruh ukuran sampel terhadap performance estimator berdasarkan hasil simulasi Monte Carlo.

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut.

1. Memberikan pemahaman mengenai bagaimana pelanggaran asumsi homoskedastisitas memengaruhi ketepatan pendugaan varians koefisien regresi.
2. Menyediakan informasi mengenai kinerja estimator HCCME (HC0, HC1, HC2, HC3) sebagai alternatif penduga varians ketika terjadi heteroskedastisitas.

BAB II

TINJAUAN PUSTAKA

2.1 Model Linear

Model linear merupakan kerangka dasar untuk menganalisis hubungan antara variabel respon dan satu atau lebih variabel penjelas. Model ini berfungsi untuk mengukur pengaruh variabel bebas terhadap variabel terikat dalam bentuk hubungan yang linier terhadap parameter, bukan terhadap variabelnya. Artinya, meskipun variabel-variabel dapat berbentuk transformasi nonlinear seperti logaritma atau kuadrat, model tetap disebut linear selama parameter β muncul secara linear di dalam persamaan (Gujarati *and* Porter, 2009).

Secara umum, model linear dapat dinyatakan sebagai:

$$Y = X\beta + \varepsilon \quad (\text{II.1})$$

dengan:

Y = vektor variabel respon

X = matriks yang memuat variabel penjelas

β = vektor parameter

ε = vektor galat acak .

Dalam bentuk matriks, model linear dapat dituliskan sebagai berikut:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}. \quad (\text{II.2})$$

Model linear mencakup baik regresi linear sederhana (satu variabel bebas) maupun regresi linear berganda (lebih dari satu variabel bebas). Menurut Roustaei (2024), tujuan utama dari regresi linear adalah *fitting* persamaan linear ke data yang diamati sehingga hubungan antara variabel bebas dan variabel terikat dapat diprediksi dan diinterpretasikan secara statistik. Hal ini memungkinkan peneliti untuk mengukur pengaruh perubahan pada variabel bebas terhadap besaran rata-rata variabel terikat.

2.2 Estimator *Ordinary Least Squares* (OLS)

Estimator *Ordinary Least Squares* (OLS) merupakan metode pendugaan parameter yang bertujuan meminimalkan jumlah kuadrat residual. Metode ini digunakan untuk memperoleh penduga parameter $\hat{\beta}$ yang menghasilkan selisih sekecil mungkin antara nilai pengamatan dan nilai prediksi model. Estimator yang dihasilkan diharapkan mampu menjelaskan variasi data secara optimal.

Secara konseptual, suatu estimator yang baik harus mampu memisahkan komponen sistematis yang berasal dari $X\beta$ dengan komponen stokastik yang berasal dari galat (Lee and Han, 2024). Apabila komponen galat lebih dominan, maka informasi yang terkandung pada variabel penjelas mengenai variabel respon menjadi semakin kecil.

Dengan demikian, pendekatan yang digunakan harus mampu menentukan parameter β sedemikian sehingga selisih antara nilai observasi dan nilai prediksi menjadi sekecil mungkin. Selisih antara nilai pengamatan dan nilai prediksi model dinyatakan sebagai residual. Vektor residual didefinisikan sebagai berikut:

$$\varepsilon = Y - X\beta \quad (\text{II.3})$$

Residual ini merepresentasikan bagian dari variabel respon yang tidak dapat dijelaskan oleh model regresi, yaitu selisih antara nilai pengamatan dan nilai yang diprediksi oleh model. Semakin kecil nilai residual yang dihasilkan suatu estimator, semakin besar proporsi variasi data yang dapat dijelaskan oleh model (Ayanlowo *et al.*, 2024).

Metode OLS meminimumkan jumlah kuadrat residual yang dikenal sebagai *Residual Sum of Squares* (RSS). RSS merupakan fungsi objektif utama dalam metode OLS yang digunakan untuk menentukan parameter optimal (Susiana *et al.*, 2024).

Fungsi RSS dinyatakan sebagai:

$$S(\beta) = \epsilon' \epsilon = (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) \quad (\text{II.4})$$

karena hasil perkalian tersebut berupa skalar, nilainya tetap sama ketika ditranspos. Dengan melakukan ekspansi aljabar, diperoleh:

$$\begin{aligned} S(\beta) &= (\mathbf{Y}' - \mathbf{X}'\beta')(\mathbf{Y} - \mathbf{X}\beta) \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\beta - \beta'\mathbf{X}'\mathbf{Y} + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{Y}'\mathbf{Y} - (\mathbf{Y}'\mathbf{X}\beta)' - \beta'\mathbf{X}'\mathbf{Y} + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{Y}'\mathbf{Y} - \beta'\mathbf{X}'\mathbf{Y} - \beta'\mathbf{X}'\mathbf{Y} + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{Y}'\mathbf{Y} - 2\beta'\mathbf{X}'\mathbf{Y} + \beta'\mathbf{X}'\mathbf{X}\beta. \end{aligned}$$

Untuk memperoleh penduga parameter yang meminimumkan fungsi RSS, diferensialkan $S(\beta)$ terhadap β sehingga diperoleh turunan pertama fungsi tersebut sebagai berikut:

$$\frac{\partial S(\beta)}{\partial \beta} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\beta. \quad (\text{II.5})$$

Nilai minimum dicapai ketika turunan pertama sama dengan nol. Dengan mengambil $\frac{\partial S(\beta)}{\partial \beta} = 0$, maka diperoleh parameter $\hat{\beta}$ sehingga diperoleh persamaan:

$$\begin{aligned} -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\hat{\beta} &= 0 \\ 2\mathbf{X}'\mathbf{X}\hat{\beta} &= 2\mathbf{X}'\mathbf{Y} \\ \mathbf{X}'\mathbf{X}\hat{\beta} &= \mathbf{X}'\mathbf{Y} \\ \hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \end{aligned} \quad (2.6)$$

Persamaan (2.6) merupakan bentuk umum penduga parameter menggunakan metode *Ordinary Least Squares* (OLS).

Menurut Rancher *and* Schaalje (2008), metode ini memiliki sifat-sifat optimal di bawah asumsi Gauss Markov sehingga estimator yang dihasilkan bersifat *best linear unbiased estimator* (BLUE). Oleh karena itu, OLS menjadi dasar dalam berbagai pengembangan metode estimasi yang lebih kompleks, termasuk metode robust yang dikembangkan untuk mengatasi pelanggaran asumsi klasik (Susiana *et al.*, 2024).

2.2.1 Sifat-Sifat *Estimator Ordinary Least Squares* (OLS)

Estimator OLS memiliki beberapa sifat penting yang menjadi dasar penggunaannya dalam analisis regresi. Salah satu sifat utama estimator OLS adalah bersifat linear terhadap pengamatan (Rancher *and* Schaalje, 2008).

Suatu estimator dikatakan linear apabila dapat dinyatakan sebagai kombinasi linear dari observasi Y . Estimator OLS dinyatakan sebagai $\hat{\beta} = (X'X)^{-1}X'Y$, sehingga $\hat{\beta}$ dapat ditulis sebagai $\hat{\beta} = AY$. Hal ini dapat ditunjukkan oleh bentuk:

$$\hat{\beta} = (X'X)^{-1}X'Y = AY \quad (\text{II.7})$$

dengan $A = (X'X)^{-1}X'$.

Karena matriks A hanya bergantung pada matriks X dan tidak bergantung pada Y , maka $\hat{\beta}$ merupakan kombinasi linear dari elemen-elemen vektor Y . Dengan demikian, estimator OLS merupakan estimator yang bersifat linear.

Selain bersifat linear, estimator OLS juga memiliki sifat takbias di bawah asumsi model regresi linear klasik. Sifat ini dinyatakan dalam teorema berikut.

Teorema 2.1 Menurut Clarke (2008), jika model linear memenuhi asumsi $E(\varepsilon) = 0$ dan $\text{Var}(\varepsilon) = \sigma^2 \mathbf{I}_n$, maka estimator OLS bersifat penduga takbias terhadap β , yaitu:

$$E(\hat{\beta}) = \beta. \quad (\text{II.8})$$

Bukti: akan ditunjukkan sifat pendugaan estimator OLS takbias, maka

$$\begin{aligned} E(\hat{\beta}) &= E((X'X)^{-1}X'Y) \\ &= (X'X)^{-1}X'E(Y) \\ &= (X'X)^{-1}X'X\beta \\ &= \beta \end{aligned} \quad (2.9)$$

sehingga terbukti bahwa $\hat{\beta}$ sebagai estimator OLS merupakan penduga takbias terhadap parameter β .

2.2.2 Varians dan Kovarians *Ordinary Least Squares* (OLS)

Analisis pendugaan varian dan kovarians dari estimator parameter penting karena varian dari $\hat{\beta}$ digunakan dalam pengujian hipotesis, pembentukan interval kepercayaan, serta menjadi dasar bagi penyesuaian ketika asumsi homoskedastisitas tidak terpenuhi.

Pada persamaan (2.6) diberikan estimator OLS untuk parameter β . Dengan mensubstitusikan model linear $Y = X\beta + \varepsilon$ ke dalam persamaan tersebut, maka diperoleh

$$\hat{\beta} = \beta + (X'X)^{-1}X'\varepsilon \quad (\text{II.10})$$

Persamaan (2.10) menunjukkan bahwa estimator $\hat{\beta}$ terdiri atas parameter sebenarnya β dan komponen galat. Untuk memperoleh varians dari $\hat{\beta}$ digunakan sifat dasar varians vektor acak dari Persamaan (2.10), sehingga diperoleh:

$$\text{Var}(\hat{\beta}) = \text{Var}[(X'X)^{-1}X'\varepsilon] \quad (\text{II.11})$$

menggunakan sifat kovarians linear $\text{Var}(A\varepsilon) = A \text{Var}(\varepsilon) A'$, dengan $A = (X'X)^{-1}X'$, maka diperoleh:

$$\text{Var}(\hat{\beta}) = (X'X)^{-1}X' \text{Var}(\varepsilon) X (X'X)^{-1} \quad (\text{II.12})$$

bentuk ini merupakan bentuk umum matriks varians-kovarians estimator OLS tanpa asumsi tambahan mengenai struktur varians galat.

Menurut Mustofa Usman dan Warsono (2001) jika estimator OLS $\hat{\beta}$ merupakan penduga takbias terhadap parameter β dan asumsi klasik regresi terpenuhi, yaitu galat bersifat homoskedastisitas dan tidak mengandung autokorelasi, maka struktur varians galat dinyatakan sebagai:

$$\text{Var}(\varepsilon) = \sigma^2 \mathbf{I}_n \quad (\text{II.13})$$

dengan mensubstitusikan varians galat tersebut ke dalam persamaan (2.12), maka diperoleh:

$$\text{Var}(\hat{\beta}) = (X'X)^{-1}X'(\sigma^2 \mathbf{I}_n)X(X'X)^{-1} \quad (\text{II.14})$$

karena σ^2 merupakan skalar, maka persamaan tersebut dapat disederhanakan menjadi:

$$\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} \quad (\text{II.15})$$

Varians ini menunjukkan tingkat ketelitian estimator OLS dan menjadi dasar dalam perhitungan *standard error*. Semakin besar variasi pada variabel bebas, semakin kecil varians $\hat{\beta}$, yang berarti semakin presisi estimasi parameter yang dihasilkan.

Dalam praktik, nilai sebenarnya dari σ^2 tidak diketahui, sehingga digunakan penduga takbias berdasarkan residual $\hat{\varepsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta}$ yang didefinisikan sebagai:

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n - k} \quad (\text{II.16})$$

dengan,

- $\hat{\sigma}^2$ = penduga varians galat
- $\hat{\varepsilon}$ = vektor residual
- n = jumlah observasi
- k = jumlah parameter yang diestimasi.

Nilai $\hat{\varepsilon}'\hat{\varepsilon}$ merupakan jumlah kuadrat residual yang menggambarkan seberapa jauh hasil prediksi model menyimpang dari data aktual $\mathbf{Y}$. Penduga ini juga dikenal sebagai *Mean Squared Error* (MSE) dari model OLS. Penggunaan pembagi $(n - k)$ bertujuan untuk memperhitungkan derajat kebebasan yang hilang akibat pendugaan k parameter dalam β , sehingga $\hat{\sigma}^2$ menjadi penduga takbias bagi σ^2 (Clarke, 2008). Sifat ketakbiasan penduga varians galat tersebut dinyatakan dalam teorema berikut.

Teorema 2.2 Menurut Clarke (2008), jika $E(\varepsilon) = 0$ dan $Var(\varepsilon) = \sigma^2 \mathbf{I}_n$, maka $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k}$ merupakan penduga tak bias bagi σ^2 , sehingga:

$$E(\hat{\sigma}^2) = \sigma^2.$$

Bukti: Akan ditunjukkan $\hat{\sigma}^2$ merupakan penduga tak bias bagi σ^2 , residual OLS dapat ditulis sebagai $\hat{\varepsilon} = \mathbf{M}\varepsilon$, dengan $\mathbf{M} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Karena $\mathbf{M}$ simetris dan idempoten serta $rank(\mathbf{M}) = n - k$, maka $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon'\mathbf{M}\varepsilon$. Menggunakan sifat umum bentuk kuadrat, $E(\varepsilon'\mathbf{A}\varepsilon) = \sigma^2 tr(\mathbf{A})$ untuk setiap matriks simetris $\mathbf{A}$. Ambil $\mathbf{A} = \mathbf{M}$, maka:

$$E(\hat{\varepsilon}'\hat{\varepsilon}) = \sigma^2 tr(\mathbf{M}) = \sigma^2(n - k) \quad (\text{II.17})$$

sehingga,

$$E(\hat{\sigma}^2) = \frac{1}{n-k} \sigma^2 (n-k) \quad (\text{II.18})$$

$$E(\hat{\sigma}^2) = \sigma^2. \quad (\text{II.19})$$

Terbukti.

Dengan demikian, diperoleh penduga empiris matriks kovarians estimator $\hat{\beta}$ adalah:

$$\widehat{Cov}(\hat{\beta}) = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}. \quad (\text{II.20})$$

Persamaan (2.20) menunjukkan bahwa bentuk matriks varians–kovarians penduga OLS diperoleh dengan mengasumsikan bahwa galat memiliki varians konstan atau bersifat homoskedastik. Asumsi ini merupakan bagian dari asumsi model linear klasik yang sering digunakan dalam analisis regresi.

Namun, dalam praktiknya khususnya pada data ekonomi dan sosial, asumsi homoskedastisitas sering kali tidak terpenuhi, sehingga varians galat dapat berbeda antar pengamatan. Ketika kondisi tersebut terjadi, penduga OLS tetap bersifat tak bias, tetapi matriks varians–kovarians klasik tidak lagi valid.

2.3 Teorema Gauss Markov

Teorema Gauss–Markov merupakan dasar utama dalam teori model linear. Teorema ini menyatakan bahwa misalkan model linear dinyatakan dalam bentuk matriks $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$, dengan $\mathbf{X}$ memiliki derajat kolom penuh.

Diasumsikan bahwa:

- Model bersifat linear dalam parameter β .
- $E(\varepsilon) = 0$.
- $\text{Var}(\varepsilon) = \sigma^2 \mathbf{I}_n$, yaitu galat bersifat homoskedastis dan tidak mengandung autokorelasi.
- Tidak terdapat multikolinearitas sempurna pada matriks $\mathbf{X}$.

Dengan demikian, Teorema Gauss–Markov menyatakan bahwa di bawah asumsi klasik model linear, estimator OLS merupakan *Best Linear Unbiased Estimator*

(BLUE). Artinya, estimator OLS bersifat linear, takbias, dan memiliki varians paling minimum dibandingkan estimator linear takbias lainnya (Hallin , 2013).

Dalam perspektif yang lebih modern, efisiensi OLS tidak hanya berlaku pada model linear sederhana tetapi juga dalam bentuk model yang lebih umum selama masih memenuhi asumsi ekspektasi nol dan varians konstan (Hansen , 2022).

2.4 Pelanggaran Asumsi Homoskedastisitas

Dalam model linear klasik, salah satu asumsi dasar dari Teorema Gauss–Markov adalah bahwa *error* ε_i memiliki varians yang sama untuk setiap observasi. Asumsi ini dikenal dengan istilah homoskedastisitas. Asumsi homoskedastisitas menyatakan bahwa varians *error* bersifat konstan untuk setiap observasi, yaitu

$$\text{Var}(\varepsilon_i) = \sigma^2, \quad i = 1, 2, \dots, n$$

dengan σ^2 merupakan varians konstan dan $\mathbf{I}_n$ matriks identitas berukuran $n \times n$. Artinya, seluruh komponen error ε_i memiliki varians yang sama sebesar σ^2 dan tidak saling berkorelasi (Wooldridge , 2016).

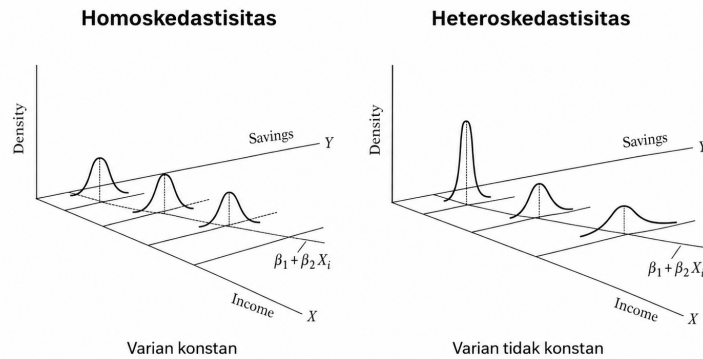
Jika varians *error* tidak konstan dan bergantung pada nilai variabel bebas, maka kondisi ini disebut heteroskedastisitas, yaitu

$$\text{Var}(\varepsilon_i) = \sigma_i^2, \quad \sigma_i^2 \neq \sigma^2$$

yang menunjukkan bahwa setiap observasi memiliki varians galat yang berbeda. Dalam bentuk matriks, varians–kovarians galat dapat dinyatakan sebagai $\text{Var}(\varepsilon) = \Omega$ dengan Ω merupakan matriks diagonal dengan elemen diagonal yang tidak sama, yaitu

$$\Omega = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{pmatrix} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2) \quad (\text{II.21})$$

Dengan demikian, model linear yang semula mengasumsikan varians *error* homogen $\sigma^2 \mathbf{I}_n$ kini berubah menjadi bentuk diagonal dengan nilai varians yang berbeda pada setiap elemen diagonalnya. Kondisi ini dapat digambarkan sebagai berikut:



Gambar 1 Kondisi Varians Homoskedastisitas dan Heteroskedastisitas

Menurut Greene (2018), jika dalam model $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ berlaku asumsi $E(\varepsilon) = 0$ dan $\text{Var}(\varepsilon) = \Omega$ dengan Ω merupakan matriks diagonal positif definit, maka estimator OLS $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ tetap merupakan estimator takbias terhadap parameter β , namun tidak lagi efisien. Selain itu, varians sebenarnya dari estimator OLS tidak lagi mengikuti bentuk klasik melainkan menjadi,

$$\text{Cov}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Omega\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \quad (\text{II.22})$$

Hal ini menunjukkan bahwa meskipun estimator OLS tetap tidak bias, matriks varians kovarians klasik yang digunakan untuk inferensi tidak lagi valid ketika heteroskedastisitas terjadi. Akibatnya, standar *error* yang diperoleh dari model OLS menjadi tidak akurat dan pengujian hipotesis dapat menghasilkan kesimpulan yang menyesatkan.

Dalam praktik analisis data, pelanggaran terhadap asumsi homoskedastisitas merupakan permasalahan yang sering dijumpai. Oleh karena itu, diperlukan suatu pendekatan yang mampu menghasilkan penduga matriks varians-kovarians yang tetap konsisten meskipun terjadi heteroskedastisitas. Salah satu pendekatan yang sering digunakan adalah penggunaan *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) seperti yang diperkenalkan oleh White (1980).

2.4.1 Distribusi Normal pada Pembangkitan Galat dalam Model Linear

Dalam teori model linear, peubah penjelas berperan dalam membentuk matriks desain yang digunakan untuk menjelaskan struktur hubungan linear antara vektor respon, parameter, dan galat. Menurut Rigollet (2015), dalam pembahasan model linear desain dapat dipandang sebagai desain tetap (*fixed design*) atau sebagai desain acak (*random design*), yaitu ketika titik-titik desain merupakan realisasi dari peubah acak. Pada penelitian ini, peubah penjelas x_i dibangkitkan sebagai desain tetap (*fixed design*), yaitu nilai-nilai x_i yang ditentukan secara deterministik dan tidak berubah-ubah antar replikasi simulasi, sebagaimana dijelaskan lebih lanjut pada Bab III. Adapun unsur acak dalam model diperoleh dari pembangkitan galat, yang dikonstruksi melalui peubah acak baku z_i yang mengikuti distribusi normal.

Salah satu distribusi peluang kontinu yang sering digunakan dalam pembangkitan peubah acak adalah distribusi normal. Menurut Walpole *et al.*, (2012) distribusi normal memiliki bentuk kurva yang simetris di sekitar rata-ratanya. Secara umum, suatu peubah acak X dikatakan mengikuti distribusi normal dengan rata-rata μ dan varians σ^2 , yang dituliskan sebagai $X \sim N(\mu, \sigma^2)$. Parameter μ menyatakan pusat sebaran data, sedangkan σ^2 menyatakan besar kecilnya penyebaran data terhadap rata-ratanya. Fungsi kepadatan peluang dari distribusi normal dengan rata-rata μ dan varians σ^2 dinyatakan sebagai:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty. \quad (\text{II.23})$$

dengan:

$f(x)$ = fungsi kepadatan peluang dari peubah acak X

x = nilai realisasi dari peubah acak X

μ = rata-rata dari distribusi normal

σ = simpangan baku dari distribusi normal, dengan $\sigma > 0$.

Menurut Montgomery *and* Runger (2014), parameter μ pada distribusi normal menunjukkan lokasi pusat distribusi, sedangkan parameter σ menunjukkan ukuran penyebaran distribusi. Apabila nilai μ berubah, maka kurva distribusi akan bergeser ke kiri atau ke kanan. Sementara itu, apabila nilai σ^2 semakin besar, maka kurva distribusi akan menjadi lebih lebar karena nilai-nilai X lebih menyebar dari rata-ratanya. Sebaliknya, apabila nilai σ^2 semakin kecil, maka nilai-nilai X akan lebih terkonsentrasi di sekitar rata-ratanya.

Distribusi normal baku merupakan bentuk khusus dari distribusi normal. Distribusi normal baku atau distribusi normal standar diperoleh ketika parameter distribusi normal memiliki $\mu = 0$ dan $\sigma = 1$. Karena $\sigma = 1$, maka variansnya adalah $\sigma^2 = 1$. Dengan demikian, distribusi normal baku dapat dituliskan sebagai $X \sim N(0, 1)$, yaitu distribusi normal dengan rata-rata 0 dan varians 1. Oleh karena itu, distribusi normal baku merupakan kasus khusus dari distribusi normal umum $X \sim N(\mu, \sigma^2)$, yaitu ketika $\mu = 0$ dan $\sigma^2 = 1$.

Penggunaan distribusi normal baku dalam membangkitkan z_i , dibandingkan membangkitkan galat ε_i secara langsung dari distribusi normal dengan varians yang berbeda-beda, memberikan keuntungan teknis karena memungkinkan pemisahan yang jelas antara komponen acak (z_i) dan komponen skala yang membawa struktur heteroskedastisitas (σ_i). Dengan pendekatan ini, distribusi peluang z_i tetap konsisten $N(0, 1)$ untuk seluruh observasi, sementara variasi varians galat antar observasi diatur sepenuhnya melalui σ_i sebagai fungsi dari x_i . Pendekatan ini umum digunakan dalam studi simulasi Monte Carlo mengenai heteroskedastisitas (Long and Ervin, 2000; MacKinnon and White, 1985) karena memberikan kendali penuh kepada peneliti terhadap bentuk fungsional varians galat yang dirancang.

2.4.2 Pembentukan Galat Heteroskedastik dalam Simulasi

Menurut Wooldridge (2016), heteroskedastisitas terjadi ketika varians galat tidak konstan untuk seluruh observasi melainkan dapat berubah mengikuti nilai peubah penjelas atau karakteristik tertentu dari data. Dalam kerangka model linear, kondisi tersebut dapat dinyatakan dengan $\text{Var}(\varepsilon_i | x_i) = \sigma_i^2$, yaitu varians galat pada observasi ke- i tidak harus sama dengan varians galat pada observasi lainnya. Dengan demikian, pembentukan galat heteroskedastik dalam simulasi perlu dilakukan dengan menentukan suatu fungsi varians yang bergantung pada nilai peubah penjelas.

Menurut Greene(2018), pada kondisi heteroskedastisitas matriks varians-kovarians galat tidak lagi berbentuk $\sigma^2 \mathbf{I}$, tetapi dapat berbentuk matriks diagonal dengan elemen diagonal yang berbeda-beda antarobservasi. Artinya, setiap observasi dapat memiliki varians galat yang berbeda, meskipun galat antar observasi tetap diasumsikan tidak berkorelasi. Dalam simulasi, kondisi ini dapat dibentuk dengan menyatakan galat sebagai hasil perkalian antara simpangan baku galat dan suatu peubah acak standar, yaitu

$$\varepsilon_i = \sigma_i z_i \quad (\text{II.24})$$

dengan,

σ_i simpangan baku galat pada observasi ke- i
 z_i = komponen acak standar dalam pembentukan galat.

Menurut Walpole *et al.*, (2012), peubah acak normal standar adalah peubah acak yang mengikuti distribusi normal dengan rata-rata nol dan varians satu. Dalam penelitian ini, z_i didefinisikan sebagai peubah acak normal standar, sehingga $z_i \sim N(0, 1)$. Dengan demikian, z_i memiliki nilai harapan $E(z_i) = 0$ dan varians $\text{Var}(z_i) = 1$. Peubah acak z_i berfungsi sebagai sumber keacakan dasar dalam pembentukan galat, sedangkan besar kecilnya varians galat dikendalikan oleh σ_i^2 . Dengan kata lain, z_i memberikan unsur acak pada galat, tetapi bentuk heteroskedastisitasnya ditentukan oleh fungsi varians yang dipilih.

Dalam pembentukan galat heteroskedastik, bentuk umum $\varepsilon_i = \sigma_i z_i$ digunakan agar galat tetap memiliki nilai harapan nol tetapi variansnya dapat berubah antarobservasi. Jika $z_i \sim N(0, 1)$ dan diasumsikan bebas terhadap peubah penjelas x_i , maka nilai harapan galat bersyarat terhadap x_i adalah

$$E(\varepsilon_i | x_i) = E(\sigma_i z_i | x_i) = \sigma_i E(z_i) = 0. \quad (\text{II.25})$$

karena σ_i dianggap tetap ketika dikondisikan pada x_i , maka varians galat bersyaratnya adalah

$$\text{Var}(\varepsilon_i | x_i) = \text{Var}(\sigma_i z_i | x_i) = \sigma_i^2 \text{Var}(z_i) = \sigma_i^2. \quad (\text{II.26})$$

Hasil tersebut menunjukkan bahwa penggunaan $z_i \sim N(0, 1)$ membuat nilai harapan galat tetap sama dengan nol, sedangkan varians galat dapat diatur melalui pemilihan σ_i^2 . Oleh karena itu, z_i tidak digunakan untuk membentuk pola heteroskedastisitas secara langsung, tetapi digunakan sebagai komponen acak standar agar galat yang dibangkitkan memiliki struktur probabilistik yang terkontrol.

Pada penelitian ini, fungsi varians galat ditentukan sebagai $\sigma_i^2 = 4 + 12x_i^2$. Karena σ_i merupakan simpangan baku galat, maka $\sigma_i = \sqrt{4 + 12x_i^2}$. Dengan demikian, galat pada observasi ke- i dibangkitkan dengan bentuk:

$$\varepsilon_i = \sqrt{4 + 12x_i^2} z_i, \quad z_i \sim N(0, 1). \quad (\text{II.27})$$

Berdasarkan bentuk tersebut, varians galat bersyarat terhadap x_i menjadi:

$$\text{Var}(\varepsilon_i | x_i) = \text{Var} \left(\sqrt{4 + 12x_i^2} z_i | x_i \right) = (4 + 12x_i^2) \text{Var}(z_i) = 4 + 12x_i^2. \quad (\text{II.28})$$

Persamaan tersebut menunjukkan bahwa varians galat tidak konstan, tetapi bergantung pada nilai peubah penjelas. Jika nilai x_i mendekati nol, maka varians galat mendekati satu. Sebaliknya, jika nilai $|x_i|$ semakin besar, maka nilai x_i^2 juga semakin besar, sehingga varians galat $1 + x_i^2$ ikut meningkat. Dengan demikian, bentuk $\varepsilon_i = \sqrt{4 + 12x_i^2} z_i$ menghasilkan pola heteroskedastisitas yang meningkat ketika nilai peubah penjelas semakin jauh dari nol.

2.5 Heteroskedasticity-Consistent Covariance Matrix Estimator (HCCME)

Heteroskedasticity-Consistent Covariance Matrix Estimator (HCCME) pertama kali diperkenalkan oleh White (1980) sebagai suatu pendekatan untuk memperoleh penduga matriks kovarians yang tetap konsisten terhadap heteroskedastisitas. Menurut Kranz (2024), penggunaan HCCME secara teoritis bertujuan menghindari dampak negatif heteroskedastisitas terhadap pengujian hipotesis tanpa memerlukan informasi mengenai bentuk atau pola ketidaksemaan varians galat.

Dalam kondisi heteroskedastisitas, varians–kovarians sebenarnya dari penduga *Ordinary Least Squares* (OLS) pada Persamaan (2.20) berlaku $\text{Var}(\varepsilon) = \Omega$. Apabila galat antar observasi tidak berkorelasi, maka matriks varians–kovarians galat memiliki bentuk diagonal sebagai:

$$\Omega = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2) \quad (\text{II.29})$$

dengan,

$$\begin{aligned} \sigma_n^2 &= E(\varepsilon_n^2) = \text{variens galat pada observasi ke-}n \\ \Omega &= \text{matriks varians–kovarians galat berukuran } n \times n. \end{aligned}$$

Struktur ini menunjukkan bahwa varians galat dapat berbeda pada setiap observasi sehingga tidak lagi memenuhi asumsi homoskedastisitas. Masalah utama pada kasus heteroskedastisitas adalah bahwa matriks Ω tidak diketahui dan tidak dapat diasumsikan konstan. Oleh karena itu diperlukan penduga terhadap matriks varians–kovarians galat. White (1980) mengusulkan untuk mengganti matriks Ω

dengan matriks diagonal yang dibentuk dari kuadrat residual hasil estimasi LS, yaitu:

$$\hat{\Omega} = \text{diag}(\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, \dots, \hat{\varepsilon}_n^2) \quad (\text{II.30})$$

dengan $\hat{\varepsilon}_i$ merupakan residual hasil estimasi LS, yaitu:

$$\hat{\varepsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta} \quad (\text{II.31})$$

sehingga diperoleh bentuk umum penduga kovarians konsisten terhadap heteroskedastisitas:

$$\hat{V}_{HC} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \quad (\text{II.32})$$

Penduga $\hat{V}_{HC}$ sering disebut sebagai *White's robust covariance estimator*.

Menurut White (1980), dalam model linear $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ dengan $E(\varepsilon) = 0$ dan $\text{Var}(\varepsilon) = \Omega$, penduga kovarians *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) yang didefinisikan pada Persamaan (2.26) merupakan penduga yang konsisten secara asimtotik terhadap matriks kovarians sebenarnya $\text{Var}(\hat{\beta})$. Hal ini menunjukkan bahwa meskipun terjadi heteroskedastisitas, penduga kovarians yang diperoleh melalui pendekatan HCCME tetap mendekati kovarians sebenarnya ketika ukuran sampel semakin besar.

Selain itu, penggunaan HCCME tidak mengubah bentuk estimator parameter OLS, sehingga mempertahankan linearitas penduga LS. Perbedaannya terletak pada penyesuaian matriks kovarians yang digunakan untuk menghitung standar *error*. Dengan menggunakan matriks kovarians yang robust terhadap heteroskedastisitas, pengujian hipotesis seperti uji-t dan uji-F tetap valid secara asimtotik meskipun varians *error* tidak homogen (Davidson and MacKinnon, 1993).

2.5.1 Jenis-Jenis Estimator HCCME

Estimator kovarians konsisten terhadap heteroskedastisitas (*Heteroskedasticity-Consistent Covariance Matrix Estimator*) yang diperkenalkan oleh White (1980) menjadi dasar bagi pengembangan berbagai bentuk koreksi lanjutan. White memperkenalkan bentuk dasar estimator yang dikenal sebagai HC0, sedangkan MacKinnon and White (1985) serta Davidson and MacKinnon (1993) mengembangkan beberapa modifikasi bentuk, yaitu HC1, HC2 dan HC3 untuk memperbaiki sifat performa pada ukuran sampel terbatas serta mengoreksi pengaruh *lverage* dari setiap observasi.

Secara umum, *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) dapat dibedakan ke dalam empat bentuk utama sebagai berikut:

$$\hat{V}_{HC_i} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}_{HC_i}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \quad i = 0, 1, 2, 3 \quad (\text{II.33})$$

dengan,

i = tipe estimator HC

$\hat{\Omega}_{HC_i}$ matriks diagonal berukuran $n \times n$

dengan elemen diagonalnya didefinisikan sebagai berikut:

$$\hat{\Omega}_{HC_i} = \text{diag}(\omega_i) \quad (\text{II.34})$$

dengan ω_i menyatakan elemen diagonal ke- i yang merepresentasikan estimasi varians galat pada observasi ke- i . Nilai ω_i bergantung pada jenis estimator HCCME yang digunakan. Nilai diagonal ω_i didefinisikan pada tabel berikut:

Tabel 1. Bentuk ω_i dari setiap estimator HC

Estimator	Bentuk ω_i	Keterangan
HC0	$\hat{\varepsilon}_i^2$	Estimator dasar <i>White's estimator</i>
HC1	$\frac{n}{n-k}\hat{\varepsilon}_i^2$	Koreksi bias untuk ukuran sampel kecil
HC2	$\frac{\hat{\varepsilon}_i^2}{1-h_{ii}}$	Penyesuaian terhadap <i>leverage</i> observasi
HC3	$\frac{\hat{\varepsilon}_i^2}{(1-h_{ii})^2}$	Penyesuaian lebih konservatif terhadap <i>leverage</i> ekstrem

Keempat bentuk tersebut memiliki tujuan yang sama, yaitu memberikan penduga kovarians yang konsisten terhadap heteroskedastisitas, tetapi dengan mekanisme koreksi yang berbeda.

Definisi 2.1 (Estimator HC0) Menurut White (1980), diberikan matriks $\mathbf{X}$ berukuran $n \times k$ dan residual *Least Squares* $\hat{\varepsilon} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n)'$, estimator HC0 sebagai bentuk dasar *White's heteroskedasticity-consistent estimator* menggunakan kuadrat residual LS sebagai penduga langsung dari varians *error*, yaitu:

$$\hat{\Omega}_{HC0} = \text{diag}(\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, \dots, \hat{\varepsilon}_n^2)$$

sehingga diperoleh kovarians robustnya adalah:

$$\hat{V}_{HC0} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}_{HC0}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad (\text{II.35})$$

Estimator ini tidak melakukan koreksi terhadap ukuran sampel maupun *leverage*. Secara asimtotik, $\hat{V}_{HC0}$ merupakan penduga yang konsisten terhadap $\text{Var}(\hat{\beta})$, namun pada ukuran sampel kecil varians yang dihasilkan sering kali *underestimated*.

Definisi 2.2 (Estimator HC1) Menurut MacKinnon *and* White (1985), diberikan matriks $\mathbf{X}$ berukuran $n \times k$ dan residual OLS $\hat{\varepsilon}$, variasi estimator ini didefinisikan dengan koreksi derajat kebebasan sebagai berikut:

$$\hat{\Omega}_{HC1} = \frac{n}{n-k} \text{diag}(\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, \dots, \hat{\varepsilon}_n^2)$$

dengan, n menyatakan jumlah observasi dan k menyatakan jumlah parameter yang diestimasi dalam model regresi. sehingga kovarians robustnya adalah:

$$\hat{V}_{HC1} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}_{HC1}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad (\text{II.36})$$

Faktor koreksi $\frac{n}{n-k}$ merupakan bentuk analog dengan koreksi *unbiased estimator* pada varians OLS. Koreksi ini tidak mempengaruhi konsistensi estimator, tetapi dapat meningkatkan performa pada ukuran sampel kecil.

Selanjutnya, MacKinnon *and* White(1985) memperkenalkan koreksi terhadap pengaruh *leverage* dari setiap observasi. Pengembangan estimator HC2 dan HC3 didasarkan pada pengaruh *leverage* setiap observasi dalam model regresi.

Dalam model linear, nilai penduga OLS dapat dituliskan sebagai $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Estimasi nilai dugaan $\hat{\mathbf{Y}}$ diperoleh melalui proyeksi orthogonal dari $\mathbf{Y}$ ke ruang kolom $C(\mathbf{X})$. Proyeksi tersebut diberikan oleh matriks

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'. \quad (\text{II.37})$$

Matriks $\mathbf{H}$ disebut *hat matrix* karena menghasilkan nilai dugaan $\hat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}$.

Misalkan $\mathbf{H}$ merupakan matriks berukuran $n \times n$, elemen ke- i dari diagonal *hat matrix* $\mathbf{H}$ disebut nilai *leverage* (h_{ii}) dari pengamatan ke- i . Jadi, nilai *leverage* diperoleh dari elemen diagonal *hat matrix*, bukan dari seluruh elemen matriks $\mathbf{H}$.

Jika x'_i menyatakan baris ke- i dari matriks $\mathbf{X}$, maka nilai *leverage* dapat dituliskan sebagai:

$$h_{ii} = x'_i(\mathbf{X}'\mathbf{X})^{-1}x_i. \quad (\text{II.38})$$

Nilai *leverage* berada pada interval $0 \leq h_{ii} \leq 1$. Nilai *leverage* mengukur sejauh mana posisi pengamatan ke- i dalam ruang desain mempengaruhi nilai dugaan $\hat{\mathbf{Y}}$ (Rancher and Schaalje, 2008).

Besarnya nilai *leverage* memengaruhi perilaku residual OLS. Karena vektor residual OLS didefinisikan sebagai $\hat{\boldsymbol{\varepsilon}} = (\mathbf{I}_n - \mathbf{H})\mathbf{Y}$ dan $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, maka varians residual dapat dituliskan sebagai $\hat{\boldsymbol{\varepsilon}} = (\mathbf{I}_n - \mathbf{H})\boldsymbol{\varepsilon}$. Dengan demikian, matriks varians-kovarians residual dapat dituliskan sebagai berikut.

$$\text{Var}(\hat{\boldsymbol{\varepsilon}}) = (\mathbf{I}_n - \mathbf{H})\boldsymbol{\Omega}(\mathbf{I}_n - \mathbf{H})'. \quad (\text{II.39})$$

Persamaan (2.38) menunjukkan bahwa varians residual tidak hanya dipengaruhi varians galat $\boldsymbol{\Omega}$, tetapi juga oleh matriks *hat* $\mathbf{H}$. Semakin besar nilai *leverage* (h_{ii}) semakin besar pengaruh observasi tersebut terhadap nilai dugaan regresi.

Dalam kasus homoskedastisitas, hubungan antara *leverage* dan varians residual dapat dilihat lebih jelas, karena berlaku $\text{Var}(\hat{\varepsilon}_i) = \sigma^2(1 - h_{ii})$. Bentuk tersebut menunjukkan bahwa semakin besar nilai h_{ii} , semakin kecil nilai $1 - h_{ii}$, sehingga varians residual pada observasi tersebut menjadi lebih kecil. Dengan kata lain, residual OLS pada observasi dengan *leverage* tinggi cenderung mengecil. Akibatnya, penggunaan residual kuadrat $\hat{\varepsilon}_i^2$ secara langsung dapat menyebabkan pendugaan varians galat menjadi terlalu kecil.

Untuk mengoreksi hal tersebut, MacKinnon and White(1985) memperkenalkan faktor berbasis $(1 - h_{ii})^{-1}$ pada estimator HC2 dan $(1 - h_{ii})^{-2}$ pada estimator HC3. Koreksi ini digunakan agar residual dari observasi dengan *leverage* tinggi tidak memberikan pendugaan varians yang terlalu kecil dalam perhitungan estimator kovarians robust.

Definisi 2.3 (Estimator HC2) Menurut MacKinnon and White (1985) koreksi pertama terhadap pengaruh *leverage* dilakukan dengan menyesuaikan varian residual setiap observasi dengan faktor $(1 - h_{ii})^{-1}$, menghasilkan

$$\hat{\boldsymbol{\Omega}}_{HC2} = \text{diag}\left(\frac{\hat{\varepsilon}_1^2}{1 - h_{11}}, \frac{\hat{\varepsilon}_2^2}{1 - h_{22}}, \dots, \frac{\hat{\varepsilon}_i^2}{1 - h_{ii}}\right)$$

maka kovarians robust-nya adalah:

$$\hat{V}_{HC2} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}_{HC2}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad (\text{II.40})$$

Koreksi ini memperbesar varian residual bagi observasi dengan *leverage* tinggi, sehingga penduga $\hat{V}_{HC2}$ lebih stabil dibandingkan HC0 dan HC1 pada data dengan pengaruh observasi tidak seragam.

Definisi 2.4 (Estimator HC3) Davidson *and* MacKinnon (1993) mengusulkan koreksi yang lebih konservatif dengan menggunakan faktor kuadrat dari $(1 - h_{ii})$ terutama untuk ukuran sampel kecil

$$\hat{\Omega}_{HC3} = \text{diag}\left(\frac{\hat{\varepsilon}_1^2}{(1 - h_{11})^2}, \frac{\hat{\varepsilon}_2^2}{(1 - h_{22})^2}, \dots, \frac{\hat{\varepsilon}_i^2}{(1 - h_{ii})^2}\right)$$

sehingga,

$$\hat{V}_{HC3} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}_{HC3}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad (\text{II.41})$$

Estimator HC3 memberikan koreksi yang lebih kuat terhadap pengaruh *leverage* ekstrem sehingga sering dianggap lebih stabil dan lebih andal pada ukuran sampel kecil.

Secara keseluruhan, keempat estimator HCCME tersebut dikembangkan untuk menghasilkan penduga kovarians yang tetap konsisten dalam kondisi heteroskedastisitas, dengan perbedaan utama terletak pada bentuk penyesuaian terhadap ukuran sampel dan pengaruh *leverage* masing-masing observasi.

2.6 Simulasi Monte Carlo

Simulasi Monte Carlo merupakan metode komputasi yang menggunakan pembangkitan bilangan acak secara berulang untuk mempelajari perilaku suatu model atau estimator statistik. Menurut Robert *and* Casella (2004), simulasi Monte Carlo digunakan untuk memperoleh pendekatan numerik terhadap suatu permasalahan statistik melalui pengulangan sampel acak dari distribusi tertentu. Metode ini banyak digunakan ketika bentuk analitik suatu estimator sulit diperoleh atau ketika peneliti ingin mengevaluasi performa estimator dalam berbagai kondisi data yang dikendalikan.

Dalam simulasi Monte Carlo, data dibangkitkan berdasarkan model probabilistik yang telah ditentukan, kemudian proses estimasi dilakukan secara berulang sebanyak sejumlah replikasi tertentu. Hasil dari seluruh replikasi tersebut digunakan untuk membentuk distribusi empiris dari estimator yang diamati. Dengan demikian, simulasi Monte Carlo memungkinkan peneliti untuk mempelajari sifat statistik suatu estimator, seperti bias, varians, dan kestabilan estimator pada berbagai kondisi sampel.

Menurut Robert *and* Casella (2004), salah satu keunggulan simulasi Monte Carlo adalah kemampuannya dalam mengevaluasi metode statistik pada kondisi yang sulit dianalisis secara teoritis. Dalam konteks model linear, simulasi Monte Carlo dapat digunakan untuk membangkitkan peubah penjelas, galat, dan peubah respon berdasarkan struktur model tertentu, kemudian membandingkan performa estimator yang digunakan pada model tersebut.

Pada penelitian ini, simulasi Monte Carlo digunakan untuk mengevaluasi ketepatan estimator kovarians *Ordinary Least Squares* (OLS) dan *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) pada kondisi heteroskedastisitas. Data simulasi dibangkitkan berdasarkan model linear sederhana dengan struktur galat yang memiliki varians tidak konstan. Selanjutnya, proses estimasi dilakukan secara berulang untuk setiap ukuran sampel dan setiap kombinasi parameter yang digunakan dalam penelitian.

Hasil dari seluruh replikasi simulasi digunakan untuk menghitung ukuran kinerja estimator, seperti bias, *Mean Squared Error* (MSE), dan *coverage probability*. Bias digunakan untuk mengukur kecenderungan estimator dalam mengestimasi varians lebih besar atau lebih kecil dari nilai acuan. *Mean Squared Error* (MSE) digunakan untuk mengukur besar kesalahan estimasi secara kuadrat, sedangkan *coverage probability* digunakan untuk menilai kemampuan selang kepercayaan dalam memuat nilai parameter sebenarnya. Dengan demikian, simulasi Monte Carlo memberikan dasar empiris untuk membandingkan kinerja estimator kovarians OLS, HC0, HC1, HC2, dan HC3 dalam kondisi heteroskedastisitas.

BAB III

METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun ajaran 2025/2026 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang beralamatkan di Jalan Prof. Dr. Ir. Soemantri Brojonegoro, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung, Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data simulasi yang dibentuk secara tetap menggunakan perangkat komputasi SAS IML. Data ini tidak berasal dari pengamatan empiris, melainkan dirancang dengan struktur yang menyerupai struktur model linear sederhana di bawah kondisi pelanggaran asumsi homoskedastisitas. Data yang dibangkitkan terdiri dari peubah penjelas, galat, peubah respon, dan matriks desain yang digunakan dalam proses estimasi parameter serta perhitungan estimator kovarians. Simulasi dilakukan untuk 4 ukuran sampel, yaitu 30, 50, 100, dan 250 observasi.

Untuk setiap ukuran sampel, peubah penjelas dibentuk sebagai desain tetap dengan $x_i = -1 + \frac{2(i-1)}{n-1}$, untuk $i = 1, 2, \dots, n$. Dengan demikian, nilai peubah penjelas berada pada rentang -1 sampai 1. Penggunaan desain tetap dilakukan agar pola heteroskedastisitas dapat terlihat lebih jelas berdasarkan nilai peubah penjelas dan agar perbandingan antar ukuran sampel tidak dipengaruhi oleh perubahan skala peubah penjelas.

Untuk ukuran sampel 30, nilai peubah penjelas yang dibentuk adalah $x_1 = -1 + \frac{2(1-1)}{30-1}$, $x_2 = -1 + \frac{2(2-1)}{30-1}$, $\dots$, $x_{30} = -1 + \frac{2(30-1)}{30-1}$. Untuk ukuran sampel 50, nilai peubah penjelas yang dibentuk adalah $x_1 = -1 + \frac{2(1-1)}{50-1}$, $x_2 = -1 +$

$\frac{2(2-1)}{50-1}, \dots, x_{50} = -1 + \frac{2(50-1)}{50-1}$. Untuk ukuran sampel 100, nilai peubah penjelas yang dibentuk adalah $x_1 = -1 + \frac{2(1-1)}{100-1}, x_2 = -1 + \frac{2(2-1)}{100-1}, \dots, x_{100} = -1 + \frac{2(100-1)}{100-1}$. Selanjutnya, untuk ukuran sampel 250, nilai peubah penjelas yang dibentuk adalah $x_1 = -1 + \frac{2(1-1)}{250-1}, x_2 = -1 + \frac{2(2-1)}{250-1}, \dots, x_{250} = -1 + \frac{2(250-1)}{250-1}$. Nilai x_i menyatakan nilai peubah penjelas pada observasi ke- i . Nilai-nilai $x_1, x_2, \dots, x_n$ kemudian digunakan untuk membentuk matriks desain $\mathbf{X}$ pada model linear. Seluruh proses pembentukan data dilakukan dengan jumlah replikasi $R = 1000$ untuk setiap skenario simulasi.

Galat heteroskedastik dalam penelitian ini dibentuk dengan menggunakan struktur heteroskedastisitas yang bergantung pada nilai peubah penjelas. Galat dibangkitkan sebagai $\varepsilon_i = \sigma_i \cdot z_i$ dengan $\sigma_i = \sqrt{4 + 12x_i^2}$ dan $z_i \sim N(0, 1)$. Dengan demikian, galat dapat dinyatakan sebagai $\varepsilon_i = \sqrt{4 + 12x_i^2} z_i$. Berdasarkan struktur tersebut, varians galat bersifat tidak konstan dan dinyatakan sebagai $\text{Var}(\varepsilon_i | x_i) = 4 + 12x_i^2$. Struktur ini menghasilkan varians yang minimum pada saat x_i berada di sekitar nol, dan meningkat ketika nilai x_i semakin jauh dari nol hingga mencapai nilai maksimum pada titik ekstrem, yaitu mendekati -1 atau 1. Dengan demikian, data simulasi dibentuk dalam kondisi heteroskedastisitas yang bergantung secara kontinu pada peubah penjelas.

3.3 Metode Penelitian

Penelitian ini dilakukan menggunakan pendekatan simulasi Monte Carlo untuk mengevaluasi ketepatan estimator kovarians OLS dan HCCME di bawah kondisi heteroskedastisitas. Setiap langkah dilakukan secara terstruktur agar dapat mengamati performa estimator secara empiris melalui pengulangan data acak yang terkontrol. Adapun tahapan penelitian yang dilakukan dapat dijelaskan sebagai berikut:

1. Menentukan rancangan simulasi. Menetapkan ukuran sampel yang akan digunakan $n = 30, 50, 100, 250$. Menetapkan model linear sederhana yang digunakan, yaitu $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, dengan $\beta_0 = -1, 0, 1$ dan $\beta_1 = -1, 1$. Sehingga kombinasi parameternya, yaitu $(-1, -1), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 1)$. Untuk indeks i menunjukkan observasi. Nilai ini ditetapkan konstan untuk seluruh replikasi simulasi.

2. Membentuk peubah penjelas. Nilai peubah penjelas ditentukan dengan $x_i = -1 + \frac{2(i-1)}{n-1}$, untuk $i = 1, 2, \dots, n$. Dengan bentuk tersebut, nilai x_i berada pada rentang 0 sampai 1. Penggunaan desain tetap ini bertujuan agar pola perubahan varians galat dapat terlihat lebih jelas terhadap nilai peubah penjelas.
3. Membentuk matriks desain. Setelah nilai peubah penjelas diperoleh, nilai-nilai tersebut disusun ke dalam matriks desain. Nilai tersebut dinotasikan sebagai x_i . Karena model yang digunakan merupakan model linear sederhana dengan intersep, maka matriks desain terdiri atas dua kolom. Kolom pertama berisi konstanta satu, sedangkan kolom kedua berisi nilai peubah penjelas hasil simulasi. Dengan demikian, setelah peubah acak X_i dibangkitkan, diperoleh nilai-nilai $x_1, x_2, \dots, x_{30}$ untuk ukuran sampel 30, dalam bentuk matriks desain dinyatakan sebagai:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{30} \end{bmatrix}$$

untuk ukuran sampel 50, 100 dan 250 matriks dibentuk dengan struktur yang sama dan jumlah baris yang mengikuti banyaknya observasi.

4. Membangkitkan komponen acak galat. Komponen acak galat dinotasikan dengan z_i , dengan $z_i \sim N(0, 1)$. Untuk ukuran sampel 30, dibangkitkan $z_1, z_2, \dots, z_{30}$. Untuk ukuran sampel 50, dibangkitkan $z_1, z_2, \dots, z_{50}$. Untuk ukuran sampel 100, dibangkitkan $z_1, z_2, \dots, z_{100}$. Untuk ukuran sampel 250 dibangkitkan $z_1, z_2, \dots, z_{250}$.
5. Membentuk galat heteroskedastik. Galat heteroskedastik dalam penelitian ini dibentuk berdasarkan fungsi yang bergantung pada nilai peubah penjelas. Galat dibangkitkan dengan mengalikan komponen acak berdistribusi normal standar $z_i \sim N(0, 1)$ dengan fungsi skala yang merupakan fungsi kuadrat dari x_i yaitu, $\sqrt{4 + 12x_i^2}$. Sehingga, galat dapat dituliskan sebagai $\varepsilon_i = \sqrt{4 + 12x_i^2} z_i$. Dengan demikian, varians galat dinyatakan sebagai $\text{Var}(\varepsilon_i | x_i) = 4 + 12x_i^2$. Varians galat tidak konstan untuk seluruh observasi karena bergantung pada nilai x_i sehingga semakin besar nilai absolut x_i , semakin besar pula varians galat. Dengan demikian, data simulasi memiliki struktur heteroskedastisitas kontinu terhadap peubah penjelas.
6. Membentuk peubah respon. Peubah respon dibentuk berdasarkan persamaan

model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. untuk setiap kombinasi parameter β_0 dan β_1 , nilai y_i dibentuk menggunakan nilai x_i dan ε_i yang telah dibangkitkan. Dengan demikian, proses ini menghasilkan pasangan data (x_i, y_i) untuk setiap ukuran sampel dan setiap kombinasi parameter.

7. Melakukan estimasi parameter. Parameter β diestimasi menggunakan metode OLS: $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Dari hasil estimasi tersebut diperoleh nilai $\hat{\beta}_0$ dan $\hat{\beta}_1$. Dalam penelitian ini, perhatian utama diberikan pada $\hat{\beta}_1$ karena estimator kovarians yang dibandingkan dinilai berdasarkan varians estimator koefisien slope tersebut.
8. Menghitung residual. Nilai dugaan respon diperoleh dari hasil perkalian antara matriks desain dan estimator parameter OLS, yaitu $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$. Pada persamaan tersebut $\hat{\mathbf{Y}}$ merupakan vektor nilai dugaan respon, $\mathbf{X}$ merupakan matriks desain, dan $\hat{\beta}$ merupakan vektor estimator parameter OLS. Setelah nilai dugaan respon diperoleh, residual dihitung sebagai selisih antara vektor respon aktual dan vektor nilai dugaan respon, yaitu $\hat{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$. Dalam bentuk komponen, residual pada observasi ke- i adalah $\hat{\varepsilon}_i = y_i - \hat{y}_i$ atau dapat dituliskan sebagai $\hat{\varepsilon}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$. Residual ini menunjukkan selisih antara nilai respon hasil simulasi dengan nilai dugaan respon yang diperoleh dari model OLS. Untuk setiap ukuran sampel yang digunakan residual dihitung untuk seluruh observasi. Dengan demikian, untuk ukuran sampel 30 diperoleh residual $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_{30}$. Untuk ukuran sampel 50 diperoleh residual $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_{50}$. Untuk ukuran sampel 100 diperoleh residual $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_{100}$. Untuk ukuran sampel 250 diperoleh residual $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_{250}$. Residual dihitung dengan cara yang sama sesuai jumlah observasi pada masing-masing ukuran sampel.
9. Menghitung estimator kovarians OLS dan HCCME. Kovarians parameter dihitung menggunakan dua pendekatan sebagai berikut:

a. Kovarians OLS homoskedastik

$$\widehat{\text{Var}}(\hat{\beta}) = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$$

dengan, $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k}$. Pada model linear sederhana dengan intersep, jumlah parameter yang diestimasi adalah $k = 2$, yaitu β_0 dan β_1 . Nilai n menyatakan ukuran sampel yang digunakan.

b. Kovarians HCCME

$$\hat{V}_{HCCME}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}(\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, \dots, \hat{\varepsilon}_n^2)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\hat{V}_{HC1}(\hat{\beta}) = \frac{n}{n-k}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}(\hat{\varepsilon}_1^2, \hat{\varepsilon}_2^2, \dots, \hat{\varepsilon}_n^2)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\hat{V}_{HC2}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}\left(\frac{\hat{\varepsilon}_1^2}{1-h_{11}}, \frac{\hat{\varepsilon}_2^2}{1-h_{22}}, \dots, \frac{\hat{\varepsilon}_n^2}{1-h_{nn}}\right)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\hat{V}_{HC3}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}\left(\frac{\hat{\varepsilon}_1^2}{(1-h_{11})^2}, \frac{\hat{\varepsilon}_2^2}{(1-h_{22})^2}, \dots, \frac{\hat{\varepsilon}_n^2}{(1-h_{nn})^2}\right)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

10. Seluruh proses pembangkitan data dan estimasi diulang sebanyak $R = 1000$ kali untuk setiap ukuran sampel $n = 30, 50, 100$, dan 250 . Setiap hasil estimasi disimpan untuk perhitungan ukuran kerja estimator.
11. Varian sebenarnya dari estimator $\hat{\beta}_1$ dihitung secara empirik dari distribusi hasil simulasi Monte Carlo sebagai:

$$\text{Var}_{MC}(\hat{\beta}_1) = \frac{1}{R-1} \sum_{r=1}^R \left(\hat{\beta}_1 - \bar{\beta}_1 \right)^2$$

dengan, $\bar{\beta}_1 = \frac{1}{R} \sum_{r=1}^R \hat{\beta}_1$ pada setiap replikasi. Varians Monte Carlo digunakan sebagai nilai acuan untuk menilai ketepatan estimator kovarians OLS, HC0, HC1, HC2 dan HC3.

12. Menghitung ukuran kinerja estimator, dari hasil simulasi diperoleh ukuran:
 - a. Bias estimator varians

$$\text{Bias}_{OLS}(\hat{\beta}_1) = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{OLS}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)$$

$$\text{Bias}_{HC0}(\hat{\beta}_1) = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC0}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)$$

$$\text{Bias}_{HC1}(\hat{\beta}_1) = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC1}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)$$

$$\text{Bias}_{HC2}(\hat{\beta}_1) = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC2}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)$$

$$\text{Bias}_{HC3}(\hat{\beta}_1) = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC3}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)$$

- b. Mean Squared Error (MSE)

$$\text{MSE}_{OLS} = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{OLS}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)^2$$

$$\text{MSE}_{HC0} = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC0}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)^2$$

$$\text{MSE}_{HC1} = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC1}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)^2$$

$$\text{MSE}_{HC2} = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC2}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)^2$$

$$\text{MSE}_{HC3} = \frac{1}{R} \sum_{r=1}^R \left(\widehat{\text{Var}}_{HC3}(\hat{\beta}_1) - \text{Var}_{MC}(\hat{\beta}_1) \right)^2$$

nilai bias dan MSE dibandingkan untuk menilai ketepatan estimator kovarians.

13. Menghitung *Coverage Probability*

Selang kepercayaan untuk parameter $\hat{\beta}_1$ dibentuk sebagai:

$$CI = \hat{\beta}_1 \pm t_{\alpha/2, n-k} \sqrt{\widehat{\text{Var}}(\hat{\beta}_1)}$$

Coverage Probability dihitung sebagai proporsi selang kepercayaan yang mengandung nilai parameter sebenarnya β_1 , yaitu:

$$\text{Coverage}_{OLS} = \frac{1}{R} \sum_{r=1}^R I(\beta_1 \in CI_{OLS})$$

$$\text{Coverage}_{HC0} = \frac{1}{R} \sum_{r=1}^R I(\beta_1 \in CI_{HC0})$$

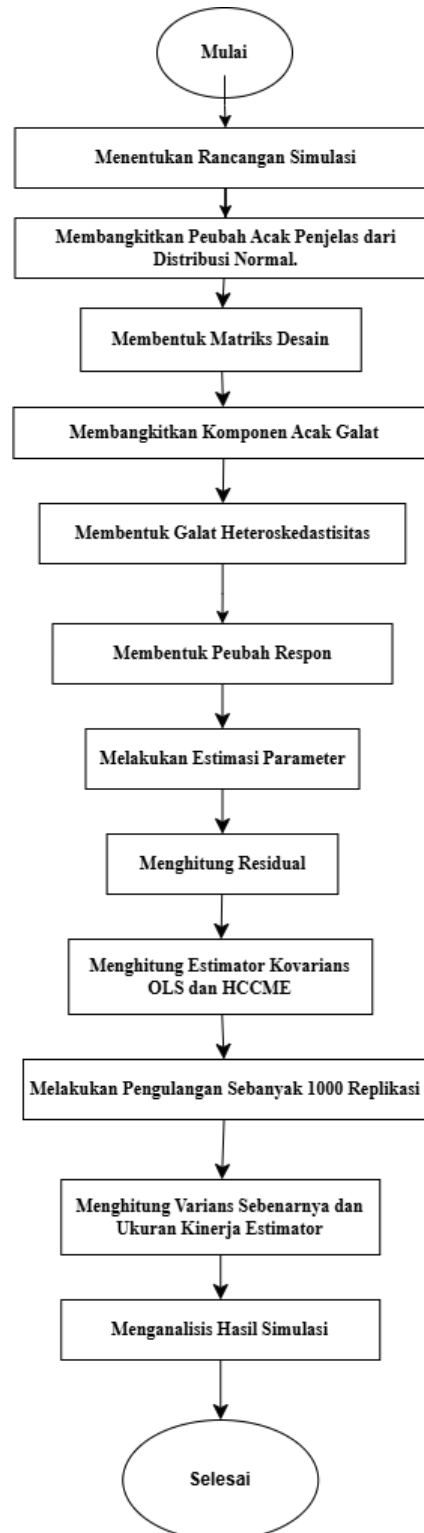
$$\text{Coverage}_{HC1} = \frac{1}{R} \sum_{r=1}^R I(\beta_1 \in CI_{HC1})$$

$$\text{Coverage}_{HC2} = \frac{1}{R} \sum_{r=1}^R I(\beta_1 \in CI_{HC2})$$

$$\text{Coverage}_{HC3} = \frac{1}{R} \sum_{r=1}^R I(\beta_1 \in CI_{HC3})$$

14. Menganalisis hasil simulasi, dari setiap replikasi dihitung nilai bias dan *Mean Squared Error* (MSE) dari estimasi kovarians $\hat{\beta}_1$. Sementara itu, estimator dengan *coverage probability* yang lebih mendekati 95% dianggap lebih baik dalam membentuk selang kepercayaan yang memuat nilai parameter sebenarnya. Kemudian dibandingkan untuk menilai ketepatan masing-masing estimator.

Berikut ini merupakan diagram alur tahapan analisis ketepatan HCCME dalam mengatasi pelanggaran asumsi homoskedastisitas:



Gambar 2 Diagram Alur Penelitian

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil simulasi Monte Carlo yang telah dilakukan untuk mengevaluasi kinerja estimator kovarians *Ordinary Least Squares* (OLS) dan *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) dalam kondisi heteroskedastisitas, dapat diperoleh beberapa kesimpulan sebagai berikut:

1. Estimator kovarians OLS menunjukkan kinerja yang kurang baik dalam kondisi heteroskedastisitas, karena secara konsisten menghasilkan estimasi varians yang lebih kecil dibandingkan varians sebenarnya. Hal ini terlihat dari nilai bias yang relatif besar dan bernilai negatif pada seluruh ukuran sampel, yang menunjukkan bahwa estimator OLS cenderung mengunderestimasi varians parameter regresi.
2. Estimator *Heteroskedasticity-Consistent Covariance Matrix Estimator* (HCCME) yang terdiri dari HC0, HC1, HC2, dan HC3 mampu memberikan estimasi varians yang lebih mendekati varians sebenarnya dibandingkan estimator OLS. Hal ini ditunjukkan oleh nilai bias yang lebih kecil serta nilai *Mean Squared Error* (MSE) yang relatif lebih rendah pada sebagian besar kondisi simulasi. Dengan demikian, estimator HCCME dapat menjadi alternatif yang lebih robust dalam mengatasi pelanggaran asumsi homoskedastisitas.
3. Ukuran sampel memiliki pengaruh yang signifikan terhadap kinerja estimator. Seiring dengan bertambahnya ukuran sampel, nilai bias dan *Mean Squared Error* (MSE) dari seluruh estimator cenderung menurun, yang menunjukkan bahwa estimasi varians menjadi semakin stabil dan akurat. Pada ukuran sampel yang lebih besar, perbedaan kinerja antar estimator juga semakin kecil.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan adalah, Penelitian selanjutnya dapat mengembangkan simulasi dengan mempertimbangkan model regresi yang lebih kompleks, seperti regresi linear berganda, sehingga dapat memberikan gambaran yang lebih luas mengenai kinerja estimator HCCME dalam berbagai kondisi.

Penelitian selanjutnya juga dapat mempertimbangkan variasi bentuk heteroskedastisitas atau distribusi galat yang berbeda, sehingga dapat mengevaluasi ketahanan estimator HCCME dalam kondisi yang lebih beragam.

DAFTAR PUSTAKA

- Ayanlowo, E. A., Oladapo, D. I., Odeyemi, A. S., and Obadina, G. O. 2024. Enhanced weighted least squares regression: A robust approach for managing outliers and heteroscedasticity. *International Journal of Science and Technology Research Archive*. 7(2): 97–106.
- Clarke, B. R. 2008. *Linear Models: The Theory and Application of Analysis of Variance*. John Wiley & Sons, Perth.
- Davidson, R., and MacKinnon, J. G. 1993. *Estimation and Inference in Econometrics*. Oxford University Press, Oxford.
- Davidson, R., and MacKinnon, J. G. 2004. *Econometric Theory and Methods*. Oxford University Press, Oxford.
- Flachaire, E. 2005. Bootstrapping heteroskedastic regression models: Wild bootstrap vs. pairs bootstrap. *Computational Statistics & Data Analysis*. 49(2): 361–376.
- Greene, W. H. 2018. *Econometric Analysis* (8th ed.). Pearson Education.
- Gujarati, D. N., and Porter, D. C. 2009. *Basic Econometrics* (5th ed.). McGraw-Hill Education, New York.
- Hallin, M. 2013. Gauss–Markov theorem. In F. Ruggeri, R. Kenett, and F. Faltin (Eds.), *Encyclopedia of Statistics in Quality and Reliability*. Wiley.
- Hansen, B. E. 2022. A modern Gauss–Markov theorem. *Econometrica*. 90(2): 667–702.
- Hayes, A. F., and Cai, L. 2007. Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*. 39(4): 709–722.
- Kranz, S. 2024. *From Replications to Revelations: Heteroskedasticity-Robust Inference*. Ulm University.

- Lee, M.-J., and Han, C. 2024. Ordinary least squares and instrumental-variables estimators for any outcome and heterogeneity. *The Stata Journal*. 24(1): 72–92.
- Long, J. S., and Ervin, L. H. 2000. Using heteroskedasticity-consistent standard errors in the linear regression model. *The American Statistician*. 54(3): 217–224.
- MacKinnon, J. G., and White, H. 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*. 29(3): 305–325.
- Majid, A., Khan, M. A., and Shah, S. Z. A. 2017. Performance of heteroskedasticity-consistent covariance matrix estimators in linear regression model. *Journal of Statistical Computation and Simulation*. 87(14): 2789–2806.
- Montgomery, D. C., and Runger, G. C. 2014. *Applied Statistics and Probability for Engineers*. 6th ed. John Wiley & Sons.
- Mustofa Usman, dan Warsono. 2001. *Teori Model Linear dan Aplikasinya*.
- Nwangburuka, C. C., Ijomah, M. A., and Nwakuya, M. T. 2023. Heteroskedasticity of unknown form: A comparison of five heteroskedasticity-consistent covariance matrix estimators. *Asian Journal of Probability and Statistics*. 22(1): 1–13.
- Rancher, A. C., and Schaalje, G. B. 2008. *Linear Models in Statistics*. 2nd ed. John Wiley and Sons.
- Rigollet, P. 2015. *High Dimensional Statistics: Chapter 2 Linear Regression Model*. MIT OpenCourseWare.
- Robert, C. P., and Casella, G. 2004. *Monte Carlo Statistical Methods*. 2nd ed. Springer, New York.
- Roustaei, N. 2024. Application and interpretation of linear regression analysis. *Medical Hypothesis, Discovery and Innovation in Ophthalmology*. 13(3): 151–159.
- Simsek, A., and Orhan, M. 2016. Heteroskedasticity-consistent covariance matrix estimators in small samples with high leverage points. *Open Journal of Statistics*. 6(4): 680–692.
- Susiana, S., Chairunisah, and Refisis, N. R. 2024. Comparison of OLS regression and robust regression in overcoming outlier problems. *Zero: Jurnal Sains, Matematika, dan Terapan*. 8(2): 108–118.

- Walpole, R. E., Myers, R. H., Myers, S. L., and Ye, K. 2012. *Probability & Statistics for Engineers & Scientists*. 9th ed. Pearson Education.
- White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*. 48(4): 817–838.
- Wooldridge, J. M. 2016. *Introductory Econometrics: A Modern Approach*. 6th ed. Cengage Learning.
- Zimmerman, D. W. 2020. *Linear Model Theory: With Examples and Exercises*. Springer.