

Penerapan Uji Asumsi Klasik Untuk Mendeteksi Kesalahan Pada Data Sebagai Upaya Menghindari Pelanggaran Pada Asumsi Klasik

Wilma Arum Nurcahya¹, Nadia Prasista Arisanti², Audrey Nabilla Hanandhika³

^{1,2,3}Department of Management, Universitas Muhammadiyah Surakarta, Indonesia

Email: b100210324@student.ums.ac.id¹, b100210330@student.ums.ac.id²,
b100210360@student.ums.ac.id³

Abstract

In this research, the focus of discussion is on solving problems from 4 tests, including the normality test, multicollinearity test, heteroscedasticity test, and autocorrelation test. The aim is to find out the efforts that researchers can make in overcoming problems in testing classical assumptions so that research data is valid and can be accounted for. This research was carried out with a literature review adopting a comparative descriptive approach which was carried out to examine problems theoretically in comparing modern econometrics with various methods of violating classical assumption tests, namely the normality test, multicollinearity test, heteroscedasticity test and autocorrelation test. The normality test is a test to find out whether research data has a normal distribution or not. The purpose of the multicollinearity test is to detect whether there is a relationship between the independent variables. The heteroscedasticity test is to test whether the error variance in each time period is different. And the last one is the autocorrelation test which is used to conceptualize the correlation between variables and time periods.

Keywords: Violation of classical assumption tests, Error variance, Application of assumption tests.

Abstrak

Pada penelitian ini, fokus pembahasan pada pemecahan permasalahan dari 4 uji yang diantaranya ialah pada uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Tujuannya untuk mengetahui upaya yang dapat dilakukan oleh peneliti dalam mengatasi permasalahan pada pengujian asumsi klasik agar data penelitian valid dan dapat dipertanggungjawabkan. Penelitian ini dilakukan dengan kajian pustaka dengan mengadopsi pendekatan deskriptif komparatif yang dilakukan untuk mengkaji permasalahan secara teoritis dalam membandingkan ekonometrika modern dengan berbagai metode pelanggaran uji asumsi klasik, yaitu dengan uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Uji normalitas ialah uji untuk mengetahui apakah data penelitian memiliki distribusi normal atau tidak. Uji multikolinearitas ditujukan untuk mendeteksi apakah ada hubungan antar variable bebas. Uji heteroskedastisitas ialah untuk menguji apakah varians error ditiap periode waktunya berbeda. Serta yang terakhir ialah uji autokorelasi yang dipergunakan untuk mengkonsep korelasi hubungan antara variable dengan periode waktu.

Kata kunci: Pelanggaran uji asumsi klasik, Varians error, Penerapan uji asumsi .

Article Info

Received date: 10 December 2021

Revised date: 20 December 2023

Accepted date: 27 December 2023

PENDAHULUAN

Uji signifikansi statistik biasanya digunakan untuk mengevaluasi validitas atau signifikansi temuan penelitian dalam ilmu biologi, medis, dan sosial. Uji signifikansi menjamin tingkat validitas temuan ilmiah dan seharusnya menurunkan kuantitas informasi palsu yang masuk ke dalam literatur.

Hal ini bahkan dianggap oleh Brunner dan Austin (2009) sebagai "fungsi utama pengujian hipotesis statistik dalam wacana ilmu pengetahuan." Namun, apakah asumsi model dilanggar atau tidak dapat menentukan apakah uji signifikansi parametrik valid atau tidak (Gelman & Hill, 2007; Zuur et al., 2009).

Para peneliti telah menyuarakan kekhawatiran tentang hasil yang tidak dapat dipertanggungjawabkan dalam sejumlah penelitian (Camerer et al., 2018; Ebersole et al., 2016; Open Science Collaboration, 2015; Silberzahn et al., 2018). Telah disebutkan bahwa salah satu penyebab utama hasil yang tidak dapat dipertanggungjawabkan adalah penggunaan statistik yang tidak tepat (Forstmeier et al., 2017). Namun, para ilmuwan sering kali tidak yakin dengan prosedur statistik mana

yang memungkinkan mereka untuk menjawab pertanyaan ilmiah mereka dengan benar dan mana yang mungkin dianggap rawan kesalahan.

Oleh karena itu, sebelum data hasil penelitian dipublikasi maka peneliti sebaiknya melakukan beberapa pengujian pada asumsi klasik. Pengujian ini bertujuan untuk memastikan bahwa persamaan regresi yang diperoleh memiliki ketepatan dalam estimasi, tidak memihak, dan konsisten. Model regresi linear OLS harus memenuhi persyaratan tertentu agar valid sebagai alat penduga. Persyaratan yang harus dipenuhi ialah asumsi klasik.

Asumsi klasik sendiri meliputi beberapa pengujian diantaranya, uji normalitas, uji multikolinearitas, uji heteroskedastisitas dan uji autokorelasi. Uji normalitas ialah uji untuk mengetahui apakah data penelitian memiliki distribusi normal atau tidak. Uji multikolinearitas ditujukan untuk mendeteksi apakah ada hubungan antar variable bebas. Uji heteroskedastisitas ialah untuk menguji apakah varians error di tiap periode waktunya berbeda. Serta yang terakhir ialah uji autokorelasi yang dipergunakan untuk mengkonsep korelasi hubungan antara variable dengan periode waktu.

Pada penelitian ini, fokus pembahasan ialah pada pemecahan permasalahan dari 4 uji tersebut yang diantaranya ialah pada uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Tujuannya ialah untuk mengetahui upaya apa saja yang dapat dilakukan oleh peneliti dalam mengatasi permasalahan pada pengujian asumsi klasik agar data penelitian valid dan dapat dipertanggungjawabkan.

LITERATUR REVIEW

pada dasarnya sebuah penelitian tidaklah selalu berjalan dengan hasil yang kita inginkan namun adakalanya hasil dari sebuah penelitian menunjukkan hasil yang tidak sesuai dengan dugaan awal maupun hipotesis awal sehingga mengharuskan para peneliti melakukan tindakan untuk dapat memperbaiki model penghitungan yang dipakai pada penelitian yang sedang dilakukannya. Hal ini dilakukan agar hasil dari penelitian valid serta dapat dipertanggung jawabkan. Oleh sebab itu peneliti harus dapat memahami, mengevaluasi tindakan perbaikan seperti apa yang dapat dipakai dalam model perhitungan yang digunakan dalam penelitiannya.

Keresahan dapat terjadi jika terdapat ketidaksesuaian pada hasil uji asumsi, reaksi wajar termasuk untuk mencoba menggunakan model uji alternatif yang lebih sesuai dengan data, mengubah data untuk sesuai dengan model yang diinginkan, atau sampai memanipulasi data supaya terlihat sesuai dengan asumsi yang diinginkan. Namun sangat disayangkan hal tersebut sering dilakukan tanpa memahami permasalahan yang sedang dibahas.

Selama pengelolaan data, peneliti selalu berkaitan dengan menetapkan keterkaitan antar dua mapun lebih peubah (Atmira et al, 2017). Pada analisis regresi linear dipakai untuk menyelidiki keterkaitan antar dua atau lebih variable, utamanya untuk menemukan gambaran maupun pola hubungan pada model yang belum diketahui sepenuhnya maupun untuk medeteksi bagaimana variable bebas dari beberapa variasi.

Asumsi dasar seperti homoskedastisitas serta nomalitas, harus dipenuhi dalam uji klasik parametrik supaya hasil yang didapat akurat. Hasilnya tidak dapat akurat jika metode klasik parametrik digunakan dengan asumsi melanggar. Berdasarkan hal tersebut penting bagi para peneliti untuk dapat memahami metode pengukuran yang dipakai serta bagaimana cara yang dapat dilakukan pada saat hasil dugaan awal tidak sesuai atau mengalami pelanggaran uji asumsi klasik, supaya mendapatkan hasil pada penelitian yang valid serta dapat dipertanggungjawabkan.

METODE

Penelitian ini dilakukan dengan kajian pustaka dengan mengadopsi pendekatan deskriptif komparatif yang dilakukan untuk mengkaji permasalahan secara teoritis dalam membandingkan ekonometrika modern dengan berbagai metode pelanggaran uji asumsi klasik, yaitu dengan uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Dalam mengkaji hal ini akan dilakukan pengkajian permasalahan secara teoritis, yang dimaksud dengan teoritis adalah suatu model yang menjelaskan bagaimana suatu hubungan teori dengan menganalisis faktor yang ada didalamnya yang sudah diketahui dalam masalah tertentu hingga menghasilkan cara dalam mengatasi permasalahan tersebut.

Langkah awal penelitian akan terfokus pada pemahaman konsep dasar asumsi klasik dan akan memberikan dasar pengetahuan yang kuat mengenai kerangka kerja teoritis yang mendasari metode deteksi pelanggaran yang akan dibahas. Selain itu juga dilakukan studi pustaka yang digunakan dalam mengumpulkan data yang diperlukan pada saat penelitian. Studi pustaka dilakukan dengan pengumpulan berbagai sumber data pustaka dan akan dikaji, setelah itu sumber pustaka akan menjadi landasan dalam mengatasi permasalahan dengan analisis uji asumsi klasik.

HASIL DAN PEMBAHASAN

Asumsi Klasik

Untuk mendapatkan nilai taksiran parameter model regresi linier dapat diestimasi dengan metode Ordinary Least Square atau OLS. Meskipun estimasi menggunakan metode OLS telah melalui prosedur yang sesuai dengan kaidah persyaratan dalam metodologi ekonometrika, namun seringkali kita mendapatkan nilai-nilai taksiran tidak seperti yang diharapkan. Oleh karena itu supaya diperoleh nilai taksiran parameter yang sesuai dengan nilai sebenarnya, maka hasil estimasi menggunakan OLS perlu diuji asumsi klasik.

Model regresi yang menggunakan metode estimasi OLS harus memenuhi asumsi klasik tertentu untuk menghasilkan apa yang dikenal sebagai BLUE (Best Linear Unbiased Estimator), yang merupakan singkatan dari karakteristik tidak bias, konsisten, dan efisien. Yang pertama best adalah hasil regresi menghasilkan suatu garis regresi dengan nilai error terkecil. Kedua ialah linier, linear dalam model ekonometrika berarti bahwa model linear dalam parameter. Selanjutnya unbiased atau tidak bias, suatu estimator dikatakan tidak bias jika nilai harapan dari estimator beta memiliki nilai yang benar atau mendekati nilai aktualnya. Artinya nilai rata-rata beta estimated sama dengan beta aktual dan dikatakan bias apabila rata-rata beta estimated tidak sama dengan beta aktual.

Asumsi yang harus dipenuhi mengacu pada Gauss Markov theorem atau the gaussian theorem dan sering dikenal dengan classical linear regression model (CLMR) meliputi :

- Asumsi 1: Hubungan antara variabel bebas dan Variabel terikat bersifat linear
- Asumsi 2: Tidak terdapat korelasi sempurna diantara variabel bebas
- Asumsi 3: Pada penyampelan berulang, akan diperoleh nilai taksiran parameter yang tetap
- Asumsi 4: Nilai harapan dari disturbance term atau error term adalah bernilai nol
- Asumsi 5: Homoskedastisitas atau varian dari disturbance term atau error term bersifat konstan
- Asumsi 6: Tidak terdapat korelasi antar disturbance term untuk periode berbeda
- Asumsi 7: Jumlah observasi atau sampel yang digunakan dalam model harus lebih besar dibandingkan dengan jumlah parameter
- Asumsi 8: Terdapat variabilitas dalam data pada variabel bebas, misal x_1 memiliki nilai yang berbeda atau memiliki variasi
- Asumsi 9: Model memiliki bentuk fungsional yang benar

Untuk menghasilkan nilai taksiran parameter yang bersifat best linear unbiased estimator atau BLUE, maka model regresi harus memenuhi persyaratan asumsi klasik. Pengujian asumsi klasik dilakukan untuk mengantisipasi kemungkinan terjadinya pelanggaran-pelanggaran asumsi klasik, namun perlu dipahami bahwa tidak semua asumsi perlu dilakukan pengujian. Asumsi klasik yang perlu dilakukan pengujian diantaranya:

1. Pengujian terhadap asumsi 4: Nilai harapan dari disturbance term atau error term adalah bernilai nol. $E(\epsilon_i) = 0$, untuk setiap $i = 1, 2, \dots, n$
Nilai error term umumnya tidak bernilai sama dengan nol, sedangkan asumsi keempat menyatakan harus bernilai nol.

$$\epsilon^* = \epsilon_i - \mu$$

$$y_i = \alpha^* + \beta X_i + \epsilon_i^* \text{ dimana } \alpha^* = \alpha + \mu$$

Koefisien determinasi (R^2) adalah ukuran goodness of fit yang menjelaskan apakah garis linear sesuai dengan data observasi. Sebagian besar model regresi memberikan nilai koefisien determinasi lebih kecil dari 1 ($R^2 < 1$)

2. Pengujian terhadap asumsi 5: Homoskedastisitas atau varian dari disturbance term atau error term bersifat konstan. Jika komponen pengganggu memiliki variasi konstan disebut homoskedastisitas, sebaliknya jika variasi tidak konstan disebut heteroskedastisitas. Perlu dilakukan uji heteroskedastisitas.

3. Pengujian terhadap asumsi 6: tidak terdapat korelasi antar disturbance term untuk periode berbeda.
 - a. Korelasi dengan variable bebas mengakibatkan terjadinya spurious correlation (korelasi lancing) yang membuat bilai taksiran menjadi bias. Perlu dilakukan uji multikolinearitas.
 - b. Korelasi dengan periode waktu berbeda: adanya korelasi antar komponen pengganggu disebut serial korelasi. Perlu dilakukan uji autokorelasi.
4. Pengujian terhadap asumsi 7: jumlah observasi atau sampel yang digunakan dalam model harus lebih besar dibandingkan dengan jumlah parameter (koefisien dalam model).
 Dalam suatu model ekonometrika yang menggunakan data kuantitatif (data skala rasio dan interval) umumnya diasumsikan berdistribusi normal. Perlu dilakukan uji normalitas.

Uji Normalitas

Menguji apakah dalam model regresi, variable residual memiliki distribusi normal merupakan tujuan dari dilakukannya uji normalitas. Uji normalitas ini dilakukan karena belum tentu data yang lebih dari 30 berdistribusi normal, sebaliknya data yang kurang dari 30 belum tentu tidak berdistribusi normal. Menurut Ghazali (2001) uji t dan uji f mengasumsikan bahwa nilai residual harus mengikuti distribusi normal. Sedangkan menurut Widarjono (2010) uji t melihat signifikansi variable independent terhadap variable dependen tidak bisa diaplikasikan jika residual tidak mempunyai distribusi normal.

Iqbal (2016) menyatakan bahwa seluruh data harus menggunakan distribusi normal pada beberapa metode, namun tidak diperlukan untuk metode lainnya. Salah satu yang menjadi keharusan sebuah data (residualnya bukan variabelnya) terdistribusi normal adalah dalam regresi linear (dengan OLS). Dalam regresi linier, salah satu asumsi klasik menyatakan bahwa residual model harus memiliki distribusi normal.

Cara mendeteksi residual berdistribusi normal atau tidak dapat melalui cara dibawah ini:

1. Analisis grafik

Menurut Ghazali (2001) dasar pengambilan keputusan meliputi

 - Model regresi memenuhi asumsi normalitas jika data menyebar disekitar garis diagonal atau jika grafik histogramnya menunjukkan pola distribusi normal.
 - Model regresi tidak memenuhi asumsi normalitas jika data menyebar jauh dari diagonal, tidak mengikuti arah garis diagonal, atau tidak menampilkan pola distribusi normal pada grafik histogram.
2. Analisis statistika
 - Chi-Square

menerapkan metode penjumlahan untuk menghitung deviasi data yang diamati setiap kelas dari nilai yang diprediksi. Uji normalitas data ditampilkan dalam kelompok. Data yang digunakan adalah data nominal atau ordinal.
 - Kolmogorov Smirnov

Pengujian ini digunakan untuk sampel data yang berukuran lebih dari 2000 sampel. Dalam pengujian, suatu data dikatakan berdistribusi normal apabila nilai signifikansi lebih dari 0,05.
 - Lilliefors

Untuk menguji apakah data berdistribusi normal atau tidak. Terdapat persyaratan dalam menggunakan metode ini diantaranya:

 1. Data berbentuk kuantitatif
 2. Data tunggal atau belum dikelompokkan pada table distribusi frekuensi.
 3. Dapat untuk n besar maupun n kecil.
 - Shapiro Wilk

Uji yang dilakukan untuk mengetahui sebaran data acak suatu sampel kecil kurang dari 50 sampel. Suatu data dikatakan berdistribusi normal apabila nilai signifikansi lebih dari 0.05.
 - Jarque Bera

Salah satu uji normalitas jenis goodness of fit test yang mana mengukur apakah skewness dan kurtosis sample sesuai dengan distribusi normal.

Multikolinearitas

Multikolinearitas dijelaskan sebagai suatu keadaan dimana terjadi korelasi linear yang sempurna diantara Sebagian atau seluruh variable bebas dalam sebuah model regresi, sehingga variable bebas dan variabel terikatnya sulit untuk diidentifikasi (Gujarati dan Porter, 2012). Masalah multikolinearitas pertama kali diperkenalkan oleh Ragnar Fisch (1934). Sebenarnya dalam suatu model, adanya korelasi merupakan hal biasa, karena dapat memberikan hasil taksiran yang valid. Namun dalam hal ini, yang dimaksud korelasi bukanlah korelasi yang seperti disebutkan di atas, namun yang dimaksud ialah korelasi yang bersifat linear, karena korelasi yang bersifat linear dapat menggagalkan proses estimasi. Dalam upaya menghindari korelasi linear, maka dilakukan uji multikolinearitas yang mana untuk mendeteksi apakah ada hubungan antar variable bebas. Disebutkan bahwa adanya masalah kolinearitas pada model jika dari hasil uji tersebut terdapat korelasi. Padahal yang seharusnya sebuah model regresi dikatakan baik jika tidak terjadi korelasi antar variable bebasnya.

Terjadinya multikolinearitas pada suatu model dipengaruhi oleh beberapa hal yaitu:

1. Terbatasnya ukuran sampel serta cara dan proses pengumpulan data.
Multikolinearitas bisa disebabkan oleh beberapa hal salah satunya yaitu kesalahan dalam proses, cara pengambilan data, serta ukuran sampel yang terlalu kecil. Ukuran sampel yang terlalu kecil dapat mempengaruhi atau meperjelas hubungan antar data, maka dari itu dalam penelitian kuantitatif disarankan untuk menggunakan sampel besar.
2. Adanya pola pergerakan data yang relative sama atau saling terkait.
Data yang tersedia dari berbagai sumber memiliki variable yang berbeda, namun terkadang memiliki pola pergerakan yang sama dan terikat satu sama lain, sehingga dapat menyebabkan multikolinearitas.
3. Spesifikasi model
Spesifikasi model ini biasanya dikaitkan dengan proses pembentukan model yang dilandaskan teori dan penelitian terdahulu agar mengurangi kemungkinan terjadinya multikolinearitas. Data yang spesifik juga mempengaruhi.
4. Adanya variable kelambanan (lag) dalam variable bebas pada model regresi.
5. Model yang overdetermined

Konsekuensi bila terjadi multikolinearitas

Multikolinearitas disebabkan karena adanya hubungan linear sempurna diantara variable bebas dengan tidak mengubah sifat parameter yang dihasilkan dari metode OLS sebagai BLUE. Karena nilai variansnya paling kecil, maka parameter turunan tetap menjadi yang terbaik di antara penduga linier dan masih valid untuk mencerminkan kondisi populasi suatu model. Multikolinearitas memiliki dampak negative yaitu varian parameter dapat bernilai besar sebagai konsekuensi dari jumlah sampel yang digunakan relative sedikit. Akibatnya, temuan estimasi menghasilkan kemampuan menolak hipotesis yang rendah dan tingkat presisi yang relatif rendah. Konsekuensi lanjutan dari kemungkinan terjadinya nilai-nilai varian parameter yang menjadi lebih besar adalah standar eror atau yang merupakan akar dari varian juga semakin besar. Jika standar eror parameter membesar maka dapat menurunkan nilai t-statistiknya, hal ini karena nilai t-statistik sangat sensitive terhadap perubahan standar eror. Jika nilai t-statistik semakin kecil maka jika nilai tersebut kemudian dibandingkan dengan t-tabel, juga akan memiliki nilai yang lebih rendah. Kecenderungan koefisien determinasi yang besar merupakan efek lain dari multikolinearitas yang menyulitkan peneliti untuk menilai seberapa besar kontribusi variabel independen terhadap variabel dependen.

Beberapa cara untuk mendeteksi Multikolinearitas diantaranya:

1. Memahami pengertian operasional dan pengukuran variable
2. Menganalisis matriks korelasi parsial
Korelasi antar dua variable bebas melebihi 0,8 maka multikolinearitas serius. 2 cara membuktikan multikolinearitas:
 - a. Melakukan regresi parsial antar variable bebas.
Terjadi karena adanya korelasi yang tinggi diantara variable bebas.
 - b. Membandingkan nilai R^2 model parsial dengan nilai R^2 model utama R^2 model parsial $> R^2$ model utama, maka terdapat multikolinearitas.
3. Mengecek keberadaan nilai koefisien determinasi (R^2)
Multikolinearitas terjadi ketika R^2 cukup tinggi, namun hanya sedikit variable bebas yang signifikan secara statistik.

4. Melihat nilai variance inflation factor (VIF)

Variance inflation factor adalah factor inflasi penyimpangan baku kuadrat untuk mengukur keeratan hubungan antar variable bebas dimana VIF merupakan fungsi dari nilai R^2 .

$$VIF = 1/(1-R^2)$$

Semakin besar R^2 , maka VIF semakin tinggi.

Uji Heteroskedastisitas

Dapat dikatakan bahwa uji heteroskedastisitas digunakan dalam model regresi untuk mengidentifikasi ketidaksamaan dalam variasi residual dari pengamatan satu ke pengamatan lainnya. Sementara itu jika hasil uji menunjukkan variasi residual dari pengamatan satu ke pengamatan lain tetap sehingga bisa disebut dengan homoskedastisitas, sebaliknya heteroskedastisitas tatkala tidak.

Dan bisa dikatakan pula bahwa homoskedastisitas merupakan regresi model yang baik. Pada data crossection biasanya memuat keadaan heteroskedastisitas dikarenakan pada data ini menghubungkan data dari berbagai ukuran, ukuran besar, sedang, maupun kecil. Dimana masalah heteroskedastisitas menyebabkannya estimasi yang diperoleh memiliki karakteristik antara lain:

1. Estimasi pada metode OLS masih tak bias
2. Estimasi pada metode OLS masih linear
3. Dikarenakan adanya heteroskedastisitas maka estimator pada metode OLS tidak memiliki lagi varian minimum, padahal asumsi yang seharusnya dipenuhi ketika memakai OLS yaitu varian dari variable gangguan diharuskan minim.

Oleh sebab itu terdapatnya heteroskedastisitas estimasi OLS tak lagi mendapatkan estimasi BLUE atau best liner unbiased estimator.

Menurut Gujarati, beberapa alasan yang menjelaskan mengenai keberadaan heteroskedastisitas dalam suatu model antara lain:

1. Adanya situasi error learning models
Seiring berjalannya waktu tingkat kesalahan akan semakin menurun. Contohnya dengan adanya teknologi yang sudah terprogram dalam suatu produksi, maka akan mengurangi kesalahan yang terjadi dalam proses produksi dan berkemungkinan juga meningkatkan jumlah serta kualitas produk.
2. Peningkatan pendapatan
Adanya peningkatan pendapatan individu yang tercatat pada data, mengakibatkan peningkatan pada data pendapatan dan jenis pendapatan, serta memiliki variasi yang lebih tinggi,
3. Pengambilan data
Dengan adanya teknologi pendukung yang diikuti dengan berkembangnya teknik pengambilan data, memberikan dampak positif terhadap data yang digunakan. Seperti sekarang ini yang mana sudah banyak sekali metode pengumpulan data yang didukung oleh perkembangan teknologi. Jika dilihat, sejauh ini memberikan kemudahan, keefektifitasan dan juga dapat dengan mudah memperluas jangkauan pengambilan data tau sampel.
4. Keberadaan outlier
Kualitas data yang menurun diakibatkan dari kondisi data yang menyimpang dari rata-rata hitung data. Data yang menyimpang atau tidak sesuai dengan rata-rata hitung tentunya akan mempengaruhi kualitas suatu data. Data dikatakan baik apabila data yang dihasilkan dari rata-rata hitung sesuai dan menghasilkan data yang valid.
5. Kesalahan spesifikasi
Model menjadi tidak valid karena kesalahan dalam pemilihan variable. Pemilihan variable sangat berpengaruh penting dalam proses pengolahan data, dikarenakan dengan variable yang tepat sebuah data akan terolah dengan baik dan akan menghasilkan data yang valid, sedangkan jika variable yang digunakan salah, maka otomatis dalam proses pengolahannya juga mengalami kesalahan yang pada akhirnya data yang dihasilkanpun akan salah atau tidak valid

Konsekuensi yang terjadi bila estimator yang dimiliki tidak memiliki varian yang minimum lagi, antara lain:

1. Dikarenakan tidak minimum nya varian menjadikan perhitungan standar eror dalam metode OLS itu tidak lagi dapat dipercaya, dikarenakan varian yang tidak minimum.

2. Dikarenakan tidak minimum nya varian hingga interval estimasi dan juga hipotesis uji yang berdasarkan kepada distribusi t dan f tidak lagi bisa diandalkan dalam uraian evaluasi regresi, hingga akan menghasilkan model regresi yang kurang baik.

Terperoleh beragam cara dalam mengetahui masalah heteroskedastisitas didalam model regresi.

1. Model Informal

Pada model ini merupakan cara yang tercepat yang bisa dipakai dalam menguji permasalahan heteroskedastisitas yaitu melalui melihat residual pola dari model regresi yang diperoleh biasanya melalui grafik. Jika residual memiliki varians yang sama lantas dari residual tidak memiliki pola pasti. Tetapi jika residual memiliki ciri heteroskedastisitas memiliki pola tersendiri.

2. Metode Park

Pada metode ini diasumsikan bahwa heteroskedastisitas muncul dikarenakan residual yang tergantung pada variable independent yang terdapat pada model. Jadi dapat disimpulkan bahwa ada atau tidak heteroskedastisitas yang berdasar uji statistic pada estimasi β pada persamaan, apabila melewati t uji β tidak signifikan maka bisa diputuskan heteroskedastisitas tidak ada.

Langkah dalam melakukan uji park antara lain :

- Menjalankan regresi terhadap model seperti regresi liner biasa yang kemudian kita akan mendapatkan residual nya.
- Melakukan regresi terhadap residual.
- Menghitung nilai statistic pada t hitung, apabila nilai lebih kecil t hitung dari nilai table atau pada nilai kritis maka tidak ada permasalahan pada heteroskedastisitas.

3. Metode korelasi Spearman

Pada metode ini dapat dijelaskan melalui regresi sederhana sebagai berikut

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Langkah-langkah uji menggunakan metode spearman adalah sebagai berikut:

- Yang pertama dapat melakukan regresi pada persamaan regresi yang terdapat diatas yang kemudian mendapatkan residualnya.
- Yang kedua yaitu mencari nilai absolut yang kemudian dirangking dimulai dari nilai yang terbesar maupun dari nilai yang terkecil, setelah merangking residual kemudian merangking variable independent, setelahnya dapat mencari korelasi spearman.

4. Metode uji Glejser

Pada metode ini dilakukan dengan regresi antar variable bebas melalui nilai absolut residualnya. Dapat dikatan jika nilai signifikan $> 0,05$ bahwa regresi model tidak terdapat heteroskodasitas. Apabila nilai $< 0,05$ bahwa model pada regresi terdapat heteroskodastisitas

Uji Autokorelasi

Autokorelasi merupakan suatu korelasi antar variabel gangguan satu melalui gangguan variable yang lain. Sementara itu salah satu dugaan penting yang terdapat pada OLS metode yaitu berhubungan melalui variable gangguan yang tidak ada kaitannya antar variable gangguan dengan variable gangguan yang lainnya. Jika ditulis dalam sebuah persamaan seperti berikut ini,

$E(\epsilon_i, \epsilon_j) = 0$, $i \neq j$, jadi dari variable gangguan ke i dengan variable gangguan ke j tidak ada atau diartikan nol, yang dimana i tidak sama dengan observasi j, jadi harus tidak ada autokorelasi yang dimana pada nilai eksptasi nya nol. Biasanya masalah autokorelasi ini timbul pada waktu runtut data, sementara itu pada data crossection jarang ditemukan adanya permasalahan autokorelasi.

Dikarenakan autokorelasi ini berhubungan dengan adanya trend dan waktu yang diamana waktu berhubungan dengan time series waktu beruntut. Maka jika terdapat autokorelasi antar variable gangguan dapat dinyatakan didalam pernyataan berikut, $E(\epsilon_i, \epsilon_j) \neq 0$, $i \neq j$, jika pada saat tidak ada autokorelasi maka nilai estimsinya adalah nol, namun pada saat terdapat autokorelasi dapat memiliki nilai tidak sama dengan nol.

Konsekuensi bila Autokorelasi terjadi. Sama halnya dengan adanya heteroskedastisitas, jika semua asumsi klasik didalam regresi model terpenuhi, kecuali terjadi autokorelasi, maka dapat dikatakan penduga OLS tetap masih tak bias dan konsisten, tetapi tidak efisien (varian meningkat),

dampaknya ialah tidak validnya uji parameter regresi dengan uji statistic t. Jadi terdapatnya autokorelasi estimasi OLS yang dimana tidak mendatangkan estimasi yang BLUE.

Konsekuensi bila estimasi yang dimiliki tidak memiliki varian minimum antara lain:

1. Perhitungan pada standar eror dari OLS metode tidak dipercaya kembali, tidak terpercaya kebenarannya.
2. Estimasi interval ataupun hipotesis uji yang berdasarkan kepada persebaran t ataupun f tidak dipercaya lagi untuk evaluasi pada regresi hasil. Jika begitu maka hasil regresi yang dimiliki akan diragukan kebenarannya.

Terperoleh beragam metode yang dapat dipakai dalam mendeteksi autokorelasi antara lain:

1. Metode Durbin-Watson (d)²

Pada metode ini ialah salah satu metode populer yang dipakai untuk mendapati permasalahan autokorelasi, dalam metode ini terdapat beberapa tahapan diantaranya yaitu:

- Menentukan hipotesis

$$H_0 : \rho = 0$$

$$H_a : \rho \neq 0$$

Jika $\rho = 0$ maka ini menyatakan bahwa tidak terdapat autokorelasi antara variable gangguan. Jika $\rho \neq 0$ maka menyatakan adanya autokorelasi. Pada hipotesis alternatif menyatakan jika adanya autokorelasi positif (>0), sedangkan jika korelasi negative (<0).

- Menghitung nilai rho yang kemudian setelahnya melakukan uji statistic apakah signifikan atau tidak.

2. Pengujian Tanda

Pada pengujian ini yaitu teknik non parametrik statistic yang dimana bersifat umum, soal ini menguji tanda dari residual. Jadi tahap awal dalam pengujian ini ialah regresi antar Y dan X menggunakan metode OLS. Yang kemudian menjalankan uji dengan residualnya.

3. Metode Breusch Godfrey

Pada metode sebelumnya yaitu uji DW memiliki kelemahan antara lain yaitu:

- Pada uji DW sekedar berlaku bila variable independennya berkarakter acak atau stokastik yaitu dimana bila pada uji ini memuat variable independent dimana bersifat non stokastik sebagaimana memasukkan variable kelayakan atau lag dari variable dependen sebagai variable independent yang disebut uji model autoregresi, oleh sebab itu pengujian tidak dapat dipakai.

- Pengujian tersebut tidak bisa dilakukan kepada model autoregresi yang relatif tinggi, seperti lag 2, lag 3 dan selanjutnya. Jika dilihat dalam lag 1 koefisien autokorelasi paling besar dibanding koefisien lag 2, lag 3. Namun apabila koefisien pada lag 1 signifikan, maka perlu dilakukannya pengujian koefisien lag 2 serta lag lainnya. Dan untuk menguji hal tersebut dipergunakan uji BG.

Mengatasi permasalahan

Tentunya diperlukan suatu cara untuk memecahkan permasalahan permasalahan dalam pengujian asumsi klasik agar tidak terjadi pelanggaran asumsi klasik. Pemecahan masalah tersebut meliputi masing masing Teknik pengujian yaitu uji normalitas, uji multikolinearitas, uji heteroskedastisitas dan uji autokorelasi.

Untuk mengatasi permasalahan uji multikolinearitas dapat dilakukan dengan:

1. Informasi mengenai keterkaitan antar variable.

Dalam membentuk suatu model, seorang peneliti akan berusaha mempelajari berbagai Pustaka baik dari teori maupun studi empiris terdahulu yang berkaitan dengan topik penelitian. Dari situ diperoleh temuan mengenai keterkaitan antar variable yang digunakan dalam model.

2. Menambah data

Banyak data yang digunakan akan memberikan hasil yang baik, oleh karena itu jika terdapat masalah multikolinearitas dapat diatasi dengan menambah jumlah data yang digunakan.

3. Menggabungkan data time series dan cross section

Multikolinearitas dapat terjadi pada kedua model tersebut, untuk memperbaikinya maka dapat dilakukan dengan cara menggabungkan kedua jenis data dalam model yang sama. Dengan melakukan penggabungan tersebut maka sifat dan struktur data menjadi berubah dan berpengaruh pada rendahnya multikolinearitas.

4. Mengeluarkan variable yang terdeteksi multikolinearitas
Mengeluarkan satu atau lebih dari variable bebas yang terdeteksi memiliki multikolinearitas yang tinggi. Contohnya pada perhitungan matriks korelasi didapati bahwa korelasi antara X_1 dan X_2 sebesar 0,85 maka nilai tersebut mengidentifikasi bahwa kedua variable berhubungan sangat kuat, maka jika kedua variable dimasukkan dalam model akan ada kemungkinan terjadi multikolinearitas.
5. Transformasi variable
Transformasi variable dari bentuk level menjadi bentuk logaritma natural dapat meminimalisir terjadinya multikolinearitas.

Untuk memecahkan permasalahan heteroskedastisitas dapat dilakukan dengan dua cara dibawah ini meliputi:

1. Melakukan transformasi data
Maksudnya ialah dengan mentransformasikan dengan cara log atau di lern kan. Cara ini biasanya yang paling umum digunakan.
2. Membiarkan datanya
Yang dimaksud membiarkan ialah dengan menggunakan robust standard error, yang mana standard eror yang kebal akan heteroskedastisitas, meskipun ada heteroskedastisitas dia tidak akan terpengaruh. Biasanya memakai huber white.

Salah suatu cara untuk menangani masalah pada autokorelasi ialah melalui memakai kuadrat terkecil umum GLS atau Generalized Least Square.

1. Jika ρ diketahui
Jika ρ diketahui maka dapat menggunakan regresi sederhana yang kemudian diregresikan dengan autoregresif, yang kemudian mencari eror atau variable gangguan, kemudian setelah itu mendapatkan persamaan baru atau barulah mendapatkan variable gangguannya.
2. Jika ρ tidak diketahui
Maka dapat melakukan estimasi dengan memakai beberapa cara yaitu menggunakan statistic Durbin Watson d . Dalam cara ini menggunakan aturan bahwa nilai ρ harus terletak antara -1 sampai dengan 1. Jika sudah mendapatkan nilai ρ kemudian disimpulkan, jika nilai pada ρ 0 maka berarti tidak ada autokorelasi. Sedangkan jika pada nilai ρ ± 1 maka berarti terdapat autokorelasi pada model yang dimiliki bisa itu korelasi negative maupun positive.

SIMPULAN

Dapat disimpulkan berdasarkan dari pembahasan yang telah diberikan, bahwa asumsi klasik merupakan suatu persyaratan yang harus dipenuhi pada model regresi yang menggunakan metode estimasi OLS supaya menghasilkan karakteristik tidak bias, konsisten dan efisien atau yang biasa disingkat dengan BLUE. Terdapat beberapa asumsi yang harus dilakukan pengujian yaitu uji normalitas menguji apakah dalam model regresi, variable residual memiliki distribusi normal atau tidak. Kemudian uji multikolinearitas keadaan dimana terjadi korelasi linear yang sempurna diantara sebagian atau seluruh variable bebas. Bila terjadi multikolinearitas maka dapat ditangani dengan menggali informasi mengenai keterkaitan antar variabel, menambah data, menggabungkan data time series dan cross section, dan mengeluarkan variabel yang terdeteksi multikolinearitas.

Pada uji heteroskedastisitas digunakan untuk mengidentifikasi ketidaksamaan dalam variasi residual dari suatu pengamatan. Untuk membenahi dapat dilakukan dengan melakukan transformasi data atau dapat dengan membiarkan datanya. Adapun uji autokorelasi untuk mendeteksi kuatnya hubungan antar periode waktu. Jika ρ diketahui maka dapat menggunakan regresi sederhana yang kemudian diregresikan dengan autoregresif, yang kemudian mencari eror atau variable gangguan, kemudian setelah itu mendapatkan persamaan baru atau barulah mendapatkan variable gangguannya.

REFERENSI

- [1] W. Forstmeier, E. Wagenmakers, and T. H. Parker, “Detecting and avoiding likely false-positive findings – a practical guide,” vol. 1954, pp. 1941–1968, 2017, doi: 10.1111/brv.12315.
- [2] P. C. Austin, “regression when independent are measured with error Inflation of Type I error rate in multiple variables,” vol. 37, no. 1, pp. 33–46, 2009, doi: 10.1002/cjs.
- [3] B. Arendt and D. M. Benjamin, “ODU Digital Commons Systematizing Confidence in Open Research and Evidence (SCORE),” 2021.
- [4] M. S. Isroah, “hitung > tabel,” vol. VII, no. 1, 2018.
- [5] A. Q. Sari, Y. L. Sukestiyarno, and A. Agoestanto, “Unnes Journal of Mathematics,” vol. 6, no. 2, pp. 168–177, 2017.
- [6] A. H. Rahmah, “The Study of In-Migration between Provinces in Indonesia by Panel Data Regression,” pp. 124–136.
- [7] U. Knief, “Violating the normality assumption may be the lesser of two evils,” pp. 2576–2590, 2021.
- [8] M. S. Reza Mubarak, *Pengantar ekonometrika*, Edisi Pert. Duta Media Publishing, 2021.
- [9] J. Salim, “Perbanas Review Volume 1, Nomor 1, November 2015 Page 19,” vol. 1, no. November, pp. 19–34, 2015.
- [10] J. O. B. Relevant and I. Jri, “PENGARUH PARTISIPASI PENGANGGARAN , DAN VOLATILITAS LINGKUNGAN PADA PERUSAHAAN MANUFAKTUR,” vol. 5, no. 2, pp. 101–130.