

Uji Asumsi Klasik untuk Menghindari Pelanggaran Asumsi Klasik pada Regresi Linier Ordinary Least Squares (OLS) dalam Ekonometrika

Fauziah Allysa Sabrina*¹

Wanda Vidya Nandita²

Dinda Dewi Maharani³

^{1,2,3} Pogram Studi Manajemen, Fakultas Ekonomi Bisnis, Universitas Muhammadiyah Surakarta, Indonesia

*e-mail: b100210239@students.ums.ac.id¹, b100210242@students.ums.ac.id²,
b100210244@students.ums.ac.id³

Abstrak

Regresi linier merupakan salah satu bagian dari ekonometrika, yang dilakukan untuk memprediksi hubungan antara variabel terikat dan variabel bebas dalam ekonomi. Hasil analisis regresi linier dapat berguna untuk menganalisis struktur dari suatu variabel atau fenomena ekonomi, meramalkan variabel atau fenomena ekonomi, dan merencanakan serta mengevaluasi kebijakan. Akan tetapi, agar hasil dari analisis regresi bersifat BLUE best, linear, unbiased, dan estimator maka harus dipastikan bahwa seluruh asumsi klasik pada regresi linier OLS terpenuhi dan tidak terjadi pelanggaran asumsi klasik apapun. Berdasarkan pertimbangan tersebut, menggunakan metode kajian pustaka. artikel ini berusaha mengulas mengenai asumsi klasik dan pengujiannya dalam regresi linier OLS ekonometrika.

Kata kunci: Pelanggaran Asumsi Klasik, Regresi Linier, OLS

Abstract

Linear regression in econometrics is conducted to predict the relation between a dependent variable and independent variables in economics. The results of linear regression analysis can be used for analyzing the structure of an economic variable or phenomenon, predicting economic variables or phenomena, and planning and evaluating economic policies. Nevertheless, in order for the regression result to be BLUE (best, linear, unbiased, and estimator) it must be ensured that all classical linear regression assumptions in OLS are met and that no violations of occur. Based on these considerations, using the literature review method, this article attempts to review classical assumptions and the testing of the classical OLS linear regression assumptions.

Keywords: Violations of Classical Linear Regression Assumptions, Linear Regression, OLS

PENDAHULUAN

Konsep regresi atau *regression* pertama kali dikemukakan oleh Franssisco Galton para paruh akhir abad 19 melalui *paper* penelitiannya yang berusaha megkaji mengenai tinggi badan anak. Dalam penelitian tersebut, diperoleh hasil bahwa orang tua berbadan tinggi akan mewariskan tinggi badan ke anaknya, sehingga anak juga akan berbadan tinggi. Sementara itu, orang tua berbadan pendek akan mewariskan tinggi badan ke anaknya sehingga anaknya juga akan berbadan pendek. Kemudian, adanya kecenderungan rata-rata tinggi badan anak untuk “regress” ke rata-rata tinggi badan dari suatu populasi. Konsep regresi, oleh Galton digunakan untuk memprediksi, meramal, atau menebak satu variabel dengan memanfaatkan variabel lainnya (Galton, 1886). Di awal abad ke-20, teori regresi tersebut kemudian dicoba dan dikonfirmasi oleh Karl Pearson dengan objek penelitian yang sama, yakni tinggi badan anak. Dalam penelitiannya, Pearson menyatakan bahwa dari sejumlah seribu total data tinggi badan keluarga, rata-rata dari tinggi badan anak yang ayahnya berbadan tinggi justru di bawah rata-rata tinggi badan ayahnya. Sebaliknya, rata-rata tinggi badan dari anak yang ayahnya berbadan pendek justru berada di atas rata-rata tinggi badan ayahnya. Pearson menyatakan bahwa tinggi badan anak mengalami *regress* ke tinggi badan rata-rata dari keseluruhan data yang diambil (*regression to mediocrity*) (Pearson & Lee, 1903). Konsep regresi yang dibangun dan dikukuhkan oleh kedua peneliti tersebut memiliki nilai guna dan kebermanfaatan bahkan hingga sekarang.

Dalam berbagai penelitian maupun analisis yang melibatkan statistik data, banyak digunakan konsep regresi. Namun, konsep tersebut telah berkembang sedemikian rupa menjadi lebih kompleks di era ilmu pengetahuan modern saat ini. Analisis regresi berhubungan dengan suatu telaah mengenai ketergantungan dari satu variabel terikat pada satu atau lebih variabel penjelasnya untuk mengetahui perkiraan atau prediksi nilai rata-rata dari variabel terikat dalam ketentuan nilai yang diketahui atau tetap (dalam pengambilan sampel berulang) dari nilai penjelasnya. Dalam ilmu ekonometrika, analisis regresi memiliki kedudukan yang esensial. Ekonometrika menjadi salah satu bentuk pengembangan regresi, yakni sebuah ilmu pengetahuan yang merupakan gabungan dari teori-teori ekonomi, statistika, dan matematika yang berusaha memberikan gambaran empiris berupa data angka dari model-model yang dibangun berdasarkan teori ekonomi (Damodar, *et al.*, 2004). Dengan kata lain, ekonometrika berusaha menggabungkan teori ekonomi dengan data faktual melalui teori dan teknik statistika (Setiawan & Kusri, 2010). Ekonometrika di era modern ini menjadi salah satu ilmu yang penting karena sifat fungsional dan praktikalnya. Ekonometrika berperan besar dalam analisis struktural dari suatu variabel atau fenomena ekonomi, peramalan dari variabel atau fenomena ekonomi, dan perencanaan serta evaluasi kebijakan (Ghozali, 2014).

Sebagai bidang ilmu, ekonometrika tentunya juga memiliki metodologi ilmiah yang khas, yang digunakan untuk menjawab permasalahan-permasalahan ekonometrik. Untuk dapat mengkaji suatu permasalahan ekonomi, secara berurutan dilakukan (a) proses spesifikasi model matematis dan teorinya, (b) pelaksanaan estimasi model ekonometrika, (c) pengevaluasian hasil estimasi, (d) pengeluaian daya prediksi atau ramalan (Setiawan & Kusri, 2010), (Ghozali, 2014) Saat evaluasi daya prediksi, diperlukan analisis regresi. Pembuatan model regresi ditujukan sebagai peramalan atau prediksi. Jika nilai dari variabel penjelas atau eksplanatori telah diketahui, maka nilai variabel terikat dapat diprediksi dengan nilai tersebut. Regresi linier dapat dibedakan menjadi dua jenis. Apabila variabel penjelas atau eksplanatorinya berjumlah satu, dan hendak digunakan untuk memprediksi satu variabel bebas, maka digunakan analisis regresi sederhana. (Rencher, 2008) Sementara itu, apabila variabel penjelas atau eksplanatorinya lebih dari satu, maka digunakan model analisis regresi berganda (Rupper, 2003). Variabel-variabel penjelas tersebut disebut juga sebagai variabel predictor karena fungsinya sebagai dasar untuk memprediksi variabel terikat. Meski demikian, hubungan ketergantungan antarvariabel yang dihasilkan melalui analisis regresi tidak harus dan tidak selalu menyiratkan hubungan sebab-akibat atau causation (Damodar, 2004).

Model regresi linier memiliki persamaan yang mengekspresikan hubungan linier antara variabel terikat yang diberi simbol Y , dan salah satu atau lebih variabel prediktor yang diberi simbol X . Jika variabel prediktornya lebih dari satu, maka digunakan simbol $X_1, X_2, \dots, \dots, X_k$ jika terdapat lebih dari satu variabel prediktor. Koefisien regresi apabila hanya terdapat satu prediktor ataupun koefisien regresi parsial apabila terdapat variabel prediktor lebih dari satu disimbolkan dengan β . Berikut adalah bentuk model persamaannya,

- Untuk regresi linier sederhana:
 $Y = a + \beta_1 X_1$
- Untuk regresi linier berganda:
 $Y = a + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$

Sementara itu, untuk model regresi, di mana Y adalah nilai sebenarnya dari hasil observasi, maka persamaan menyertakan kesalahan (error term / residual) sehingga menjadi sebagai berikut,

- Untuk regresi linier sederhana
 $Y = a + \beta_1 X_1 + e$
- Untuk regresi linier berganda
 $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e$

Dalam rangkaian metodologi ekonometrika, sebagaimana disebutkan sebelumnya, analisis regresi dilakukan setelah evaluasi hasil estimasi atau disebut juga statistik inferensi. Misal, untuk tujuan memprediksi variabel ekonomi dan perencanaan kebijakan ekonomi, untuk koefisien marginal propensity to consume (MPC) yang disimbolkan dengan β_1 diperoleh nilai 0,7. Untuk memperoleh nilai pendapatan yang diperlukan untuk mencapai target pengeluaran konsumsi senilai 4000 miliar rupiah, dapat digunakan fungsi konsumsi dan dihitung seperti di bawah ini,

$$\begin{aligned} C &= \beta_0 + \beta_1 Y \\ 4000 &= -231,8 + 0,7194 X \\ X &= 5882 \end{aligned}$$

Dengan tingkat pendapatan 5882 Miliar Rupiah dan MPC sebesar 0,7 akan menghasilkan pengeluaran konsumsi sebesar 4000 Miliar Rupiah. Perhitungan ini kemudian Berdasarkan perhitungan ini, diperoleh model estimasi yang dapat digunakan sebagai dasar perencanaan kebijakan ekonomi. Dapat pula dilakukan manipulasi melalui pengontrolan variabel X guna menghasilkan variabel tertentu. (Ruppert, 2003)

Namun, meskipun rangkaian metodologi ekonometrika berhenti di evaluasi daya prediksi, masih ada tahap penting lainnya untuk memastikan bahwa pemodelan regresi telah tepat. Pada suatu analisis regresi, dibutuhkan pendekatan untuk mengarahkan indikator menuju hasil yang best, linear, unbiased, dan estimator (BLUE). Dari berbagai pendekatan dalam analisis regresi, metode ordinary least square (OLS) atau dalam Bahasa Indonesia disebut dengan metode kuadrat terkecil (MKT) merupakan yang paling sering digunakan dan dianggap paling efektif untuk mencapai BLUE. (Gujarti, 2007) Untuk itu, terlebih dulu harus dipenuhi asumsi-asumsi klasik dalam regresi linier dengan metode OLS. Pasalnya, sesuai konsep awalnya, analisis regresi pada dasarnya merupakan prediksi atau ramalan. Koefisiennya pun adalah 'estimasi' dari parameter atau indikator populasi yang sebenarnya. Untuk itu, agar estimasinya tepat sasaran dan gap antara nilai estimasi dan nilai sebenarnya dapat diminimalisasi diperlukan pemenuhan terhadap asumsi-asumsi klasik yang menjadi premis dasar dari regresi linier OLS (Basuki & Purwoto, 2016). Dengan demikian, baru ketika asumsi klasik tersebut dipenuhi dalam analisis regresi linier, koefisien yang dihasilkan melalui metode OLS menjadi paling mendekati nilai sebenarnya. Sebaliknya, ketika asumsi-asumsi klasik ini tidak diperhatikan atau malah tidak dipenuhi, maka semakin besar kemungkinan model regresi tidak memenuhi sifat BLUE. Hal ini terutama menjadi lebih esensial bagi regresi linier berganda yang variabel prediktornya berjumlah lebih dari satu sehingga lebih kompleks. Permasalahannya adalah, masih banyak sekali model regresi linier yang tidak memenuhi seluruh asumsi klasik melainkan hanya melakukan uji t dan uji f (Mardiatmoko, 2020). Dari sini, dapat dikatakan bahwa asumsi klasik menjadi signifikan dalam regresi linier terutama dengan pendekatan OLS. Berdasarkan uraian yang telah dijabarkan di bagian ini, tulisan ini berusaha mengulas mengenai asumsi klasik dan pelanggaran asumsi klasik dalam regresi serta metode perbaikannya.

METODE

Penelitian ini merupakan penelitian kualitatif yang disusun menggunakan metode studi pustaka dengan sumber data sekunder. Studi pustaka atau disebut juga kajian pustaka merupakan keseluruhan upaya dalam proses penghimpunan informasi yang berhubungan dengan suatu topik permasalahan memanfaatkan buku, hasil penelitian, skripsi, tesis, dan berbagai sumber ilmiah

lainnya dilanjutkan dengan proses pengkajian dari sumber tersebut sebagai bahan atau landasan menganalisis masalah penelitian. Cooper dikutip dalam Creswell menjelaskan esensi dari proses ini adalah agar suatu penelitian yang sedang berlangsung dapat berhubungan dengan yang telah dilakukan sebelumnya dan dapat mengisi kekosongan celah yang ada (Creswell, 2012). Dengan menggunakan literasi yang telah ada, artikel ini berusaha mengulas dan menganalisis mengenai asumsi klasik dan pelanggaran asumsi klasik regresi linier dengan metode OLS.

HASIL DAN PEMBAHASAN

Regresi Linier

Regresi linear merupakan sebuah metode statistika yang digunakan untuk membentuk model hubungan antara variabel terikat atau dependen (Y) dengan satu atau lebih variabel bebas sebagai variabel eksplanatori atau prediktor (X). Jika variabel prediktornya berjumlah satu, maka dikategorikan sebagai regresi linier sederhana dan jika variabel prediktornya berjumlah lebih dari satu maka dikategorikan sebagai regresi linier berganda. Perolehan data untuk variabel X pada regresi linier dapat bersumber dari hasil observasi (observational data) yang sebelumnya tidak ditentukan oleh peneliti, contohnya melalui suatu kuisioner atau juga bersumber dari data yang sebelumnya telah ditentukan (experimental atau fixed data) contohnya melalui percobaan laboratorium. Yang membedakan antara data tersebut adalah bahwa dengan fixed data, perolehan informasi menjadi lebih kuat dalam memberikan penjelasan hubungan kausalitas atau causation antara variabel X dan variabel Y. Sementara itu, observational data tidak memberikan kepastian akan terlihatnya hubungan kausalitas antarvariabel. Meskipun, sebagaimana disebutkan dalam [3] bahwa tidak semua hasil harus menunjukkan hubungan causation antarvariabel. Penyebab perbedaan hasil tersebut adalah karena pada fixed data, peneliti dapat mengontrol terlebih dahulu variabel-variabel X yang akan digunakan, dibandingkan dengan menggunakan observational data yang sifatnya tidak pasti, cenderung menyesuaikan situasi pada saat dilangsungkannya penelitian. Variabel X yang diamati bisa berapa saja, tergantung keadaan di lapangan (Kurniawan, 2008).

Pada sebuah model regresi, akan ditemukan koefisien yang menjadi nilai duga parameter atau indikator di dalam model regresi terhadap keadaan yang sebenarnya atau aktual. Secara sederhana, konsep ini dapat dipersamakan dengan konsep rata-rata (mean) pada statistika tingkat dasar. Yang membedakan adalah bahwa pada sebuah mode regresi, koefisiennya adalah nilai rata-rata yang diprediksi muncul jika nilai X diberikan terhadap variabel Y (variabel terikat). Koefisien regresi dikelompokkan menjadi dua jenis,[14]

1. Intersep

Dalam teori matematika, intersep atau intercept adalah titik potong antara suatu garis dengan sumbu Y pada diagram/sumbu kartesius saat nilai $X = 0$. Dalam teori statistika, intersep merupakan nilai rata-rata pada variabel Y apabila nilai pada variabel X bernilai 0. Maksudnya adalah, jika X tidak berkontribusi apapun, maka dilihat dari nilai rata-ratanya, nilai variabel Y adalah sama dengan intersep. Intersep adalah konstanta yang membuka peluang atau kemungkinan untuk koefisien lain dapat muncul dalam suatu model regresi. Akan tetapi, tidak selalu atau tidak harus dilakukan interpretasi terhadap intersep itu sendiri. Jika data pengamatan pada variabel X tidak menyentuh nilai 0 ataupun mendekati 0, maka intersep tidak harus dilakukan interpretasi terhadap intersep karena kemungkinan besar tidak akan ditemui makna tertentu.

2. Slope

Dalam teori matematika, slope digunakan untuk mendefinisikan ukuran kemiringan dari suatu garis. Slope adalah koefisien regresi untuk variabel X. Dalam teori statistika, slope berperan dalam memperlihatkan besarnya kontribusi atau pengaruh dari suatu variabel X terhadap variabel Y. Nilai slope dapat pula diartikan sebagai rata-rata nilai tambah ataupun nilai kurang variabel y terhadap setiap bertambahnya variabel X.

Apabila di bab pendahuluan telah disebutkan mengenai bentuk dasar dari persamaan regresi linier sederhana dan berganda, berikut adalah model regresi dengan memperhitungkan slope, intersep, dan error. Berikut adalah model persamaannya,

$$Y = 9.2 + 0.6 * X + \varepsilon$$

Dalam persamaan tersebut, angka 9.2 merupakan intersep, 0.7 merupakan slope, sedangkan ε merupakan error. Error dalam regresi linier ekonometrika ini memberikan ruang bagi kemungkinan-kemungkinan lainnya yang dapat memberikan pengaruh pada variabel y yang berada di luar ruang lingkup penelitian.

Asumsi Klasik & Uji Asumsi Klasik

Pada umumnya, untuk melakukan uji asumsi klasik regresi linier diperlukan data residual. Seluruh uji asumsi klasik pada regresi linier, kecuali uji asumsi multikolinieritas, menggunakan data residual, yakni perkiraan sampel kesalahan untuk per satu pengamatan. Residual diperoleh dari mengurangi nilai yang diamati (yang ingin dicari) dengan nilai yang dipasangkan. Ada perangkat lunak statistika saat ini yang digunakan untuk melakukan pengujian asumsi klasik regresi linier, banyak yang telah dilengkapi teknologi untuk melakukan pengujian asumsi klasik secara otomatis berdasarkan masukkan data, tanpa meminta peneliti atau pengguna aplikasi untuk secara terpisah memasukkan data residual. Hal ini, pada status quo, menyebabkan banyak peneliti yang tidak memahami data residualnya. [14] Penyebabnya adalah karena sejak awal, tidak ada tuntutan untuk memilah perolehan data. Padahal, ketika menganalisis regresi menggunakan metode OLS, residual menjadi sangat penting karena berkaitan dengan asumsi klasik.

Terdapat tujuh asumsi klasik yang harus dipenuhi regresi linier menggunakan metode OLS dalam ekonometrika. Dari tujuh asumsi berikut, enamnya secara langsung menentukan kualitas hasil akhir model regresi, dan satu sisanya adalah bagian dari asumsi klasik yang harus dipenuhi. Berikut adalah tujuh asumsi klasik tersebut (Gujarti, 2003).

1. Model regresi linier harus spesifik dan memenuhi pola linier

Model regresi linier harus fungsional, tidak bisa hanya berisikan variabel X dan Y, karena model regresi baru akan memenuhi pola linearitas ketika keseluruhan istilah adalah konstanta parametermultiplied oleh variabel x. Meski demikian, terkhusus untuk asumsi klasik ini tidak terdapat uji statistika yang dapat dilakukan, karena untuk masing-masing model regresi telah disesuaikan dengan konsep teoritis penelitian atau tujuan yang hendak dicapai dari analisis regresi linier dalam ekonometrika. Dalam model regresi, harus terdapat parameter dan kemungkinan error atau kesalahan. Meski demikian, tetap dapat dimasukkan variabel non-linear diikuti dengan mengatur fungsi eksponensial dan nantinya model linier dapat membentuk kelengkungan. Berikut adalah contoh model persamaannya,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

2. Nilai rata-rata error adalah nol

Model permasalahan regresi harus memiliki nilai rata-rata error yang mendekati atau bahkan menyentuh nol agar hasil analisisnya tidak bias atau memiliki gap yang jauh antara nilai sebenarnya dan nilai prediksi. Nilai rata-rata nol tersebut dapat diperoleh dengan memasukkan konstanta dalam model regresi. Dengan kata lain, dijabarkan dalam persamaan berikut:

$$\varepsilon = Y - Y' \text{ dan } \varepsilon = 0 \text{ jika } Y = Y'$$

3. Error tidak berkorelasi dengan variabel X

Variabel X tidak boleh berkorelasi dengan ragam error (disebut juga asumsi eksogenitas) karena dapat menyebabkan estimasi koefisien menjadi bias. Alasannya, korelasi error dengan variabel X akan menunjukkan hubungan ramalan. Maksudnya, hal tersebut berarti variabel X dapat memprediksi error sehingga sifat error sebagai kesalahan acak yang tak dapat diprediksi menjadi tidak lagi terpenuhi. Apabila X sampai berkorelasi dengan error,

maka telah terjadi pelanggaran asumsi klasik yang disebut dengan endogenitas. Endogenitas dapat terjadi ketika variabel X dan variabel Y simultan, terjadi kesalahan saat mengukur variabel X, atau dihilangkannya bias variabel.

4. Error tidak berautokorelasi satu sama lainnya

Satu error tidak seharusnya berpengaruh atau meramalkan error pada pengamatan berikutnya. Jika ditemukan autokorelasi pada error, hal tersebut dapat menjadi indikasi bahwa terdapat satu atau lebih variabel yang berpengaruh terhadap variabel Y dan tidak dimasukkan ke dalam model regresi. Autokorelasi sering pula muncul pada kasus dimana data yang digunakan memasukkan unsur waktu (data time-series). Hasil yang saling berkorelasi akan mengurangi ketepatan hasil regresi linier OLS dalam ekonometrika.

5. Ragam error sifatnya homogen (homoskedastisitas)

Maksud dari ragam bersifat homogen adalah bahwa error memiliki nilai ragam yang sama antara error ke-i dan error ke-j. Untuk memastikan, dapat dilihat data residual dari model. Setiap data residual seharusnya menunjukkan nilai ragam yang serupa atau mirip sehingga jika residualnya sudah homogen, maka pada dasarnya error akan turut homogen. Dengan demikian, ragam dari error harus selalu konsisten untuk keseluruhan pengamatan atau homoskedastisitas (pencar yang sama). Pelanggaran terhadap asumsi ini disebut juga heteroskedastisitas (sebaran berbeda). Menguji homogenitas ragam error dapat dilakukan dengan membuat plot nilai residual versus nilai pas (fitted value). Jika memang ragam errornya tidak homogen, maka akan muncul bentuk kerucut dengan residu yang menyebar secara meningkat searah. Akibatnya adalah, heteroskedastisitas akan mengurangi akurasi analisis regresi linier dalam ekonometrika.

6. Tidak adanya variabel Y menunjukkan fungsi linier yang sempurna dari variabel X lainnya

Korelasi dikatakan sempurna jika dua variabel memiliki koefisien korelasi Pearson sebesar +1 atau -1. Ketika salah satu variabel berubah, variabel lainnya juga berubah dengan proporsi yang sama persis. Jika terjadi korelasi sempurna, artinya dua variabel tersebut sebenarnya adalah bentuk berbeda dari variabel yang sama. Regresi linier OLS tidak memiliki kemampuan untuk membedakan satu variabel dari yang lain ketika mereka berkorelasi secara sempurna dan hasil akhirnya analisisnya kemungkinan besar akan menghasilkan error atau kesalahan lainnya. Jika hal ini terjadi, maka salah satu variabel dari model persamaan harus dihapus. Terkhusus untuk regresi linier berganda, apabila dua atau lebih variabel X berkorelasi erat maka disebut sebagai multikolinieritas.

7. Ragam error menyebar dengan normal

Sebenarnya, regresi linier OLS tidak mengharuskan error untuk berdistribusi secara normal agar hasil prediksi regresi linier tidak 'bias' atau cenderung ke satu arah yang diekspektasikan. Akan tetapi, dengan memastikan bahwa ragam error berdistribusi dengan normal maka pengujian hipotesisnya sangat mungkin untuk menghasilkan interval reliabilitas yang dapat diandalkan. Untuk memastikan ragam error berdistribusi normal, dapat dilakukan uji normalitas Kolmogorov-Smirnov, yakni dengan mengomparasikan distribusi data yang dihipotesiskan dan distribusi data yang teramati. Jika hasilnya serupa atau mendekati, maka dapat disimpulkan bahwa data pengamatan berdistribusi secara normal.

Asumsi-asumsi klasik yang disebutkan di atas, harus diuji terdahulu sebelum dilakukan analisis regresi linier lebih lanjut. Meskipun telah disinggung secara singkat pada penjabaran mengenai tujuh asumsi klasik regresi linier menggunakan OLS, di paragraf ini akan diuraikan lagi mengenai pengujian asumsi klasik. Pertama, untuk menguji normalitas distribusi dari variabel X dan variabel Y suatu model regresi, maka dilakukan uji normalitas. Tujuannya adalah mencegah degradasi kualitas analisis regresi jika nantinya variabel tidak berdistribusi dengan normal. Uji normalitas data dapat dilakukan dengan menggunakan uji One Sample Kolmogorov Smirnov dengan ketentuan apabila signifikansi bernilai lebih dari 5% atau 0,05 maka data berdistribusi normal, sementara apabila hasil ujinya menunjukkan nilai signifikan kurang dari 5% atau 0,05

maka distribusi data tidak normal. Kedua, untuk mengetahui ada atau tidaknya korelasi dari tiap variabel X, maka dilakukan uji multikolinearitas. Tujuannya adalah untuk mencegah standar error besar dalam analisis regresi linier. Pengujian dapat dilakukan dengan melihat nilai toleransi dan nilai variance inflation factor (VIF). Nilai Tolerance mengukur variabilitas dari variabel bebas yang terpilih yang tidak dapat dijelaskan oleh variabel bebas lainnya. Jadi nilai tolerance rendah sama dengan nilai VIF tinggi, dikarenakan $VIF = 1/\text{tolerance}$, dan menunjukkan terdapat kolinearitas yang tinggi. Nilai cut off yang digunakan adalah untuk nilai tolerance 0,10 atau nilai VIF diatas angka 10 (Ghozali, 2016).

Untuk mendeteksi ketidaknyamanan varian dari residual dalam satu pengamatan ke pengamatan lainnya maka dilakukan uji heteroskedastisitas. Uji heteroskedastisitas dapat dilakukan dengan menggunakan grafik scatterplot atau dari nilai prediksi variabel terikat yaitu SRESID dengan residual error yaitu ZPRED. Apabila tidak terdapat pola tertentu dan tidak menyebar diatas maupun dibawah angka nol pada sumbu y, maka dapat disimpulkan tidak terjadi heteroskedastisitas. Terakhir, jika sepanjang pengamatan dilakukan observasi secara urut sepanjang waktu, maka dapat terjadi autokorelasi. Penyebabnya adalah karena dari satu observasi ke observasi lain, data residual tidak bergerak dengan bebas. Padahal, model regresi linier OLS seharusnya tidak menunjukkan autokorelasi. Untuk menguji autokorelasi, dapat dilakukan Run Test yang menguji residual. Jika antarresidual menunjukkan menunjukkan relasi yang acak, maka tidak ada hubungan korelasi atau autokorelasi. Dengan hipotesis sebagai dasar pengambilan keputusan apabila nilai Asymp. Sig. (2-tailed) kurang dari 5% atau 0,05, maka untuk H_0 ditolak dan H_a diterima. Hal tersebut berarti data residual terjadi secara tidak acak (sistematis). Apabila nilai Asymp. Sig. (2-tailed) lebih dari 5% atau 0,05, maka untuk H_0 diterima dan H_a ditolak. Hal tersebut berarti data residual terjadi secara acak (random).

KESIMPULAN

Dalam ilmu ekonometrika, analisis regresi menjadi sangat penting karena berperan dalam memprediksi atau meramalkan hubungan antara variabel- variabel ekonomi. Dalam regresi linier, terdapat variabel penjelas, eksplanatori, atau prediktor yang disebut sebagai variabel X, yang akan memprediksi variabel terikatnya yaitu variabel Y. peramalan atau prediksi. Jika nilai dari variabel penjelas atau eksplanatori Dalam model regresi, apabila variabel penjelas atau eksplanatorinya berjumlah satu, dan hendak digunakan untuk memprediksi satu variabel bebas, maka digunakan analisis regresi sederhana. Sementara itu, apabila variabel penjelas atau eksplanatorinya lebih dari satu, maka digunakan model analisis regresi berganda. Dari seluruh pendekatan, OLS adalah pendekatan yang paling banyak digunakan karena memiliki akurasi paling tinggi dan paling efektif untuk mencapai sifat BLUE (best, linear, unbiased, dan estimator).

Dalam setiap analisis regresi linier dengan OLS dalam ekonometrika, tujuan akhir yang ingin dicapai adalah memperoleh hasil analisis struktural dari suatu variabel atau fenomena ekonomi, peramalan dari variabel atau fenomena ekonomi, dan perencanaan serta evaluasi kebijakan. Agar tujuan tersebut dapat tercapai melalui regresi linier OLS, maka tujuh asumsi klasik yang berhubungan dengan model persamaan, distribusi dan korelasi variabel, serta error harus terpenuhi. Pengabaian atau tidak terpenuhinya asumsi klasik disebut sebagai pelanggaran asumsi klasik. Untuk mencegah terjadinya pelanggaran asumsi klasik, maka harus dilakukan serangkaian uji asumsi klasik yang meliputi uji normalitas, uji multikolinearits, uji heteroskedastisitas, dan uji autokorelasi. Serangkaian pengujian tersebut dilakukan sebelum dilakukan analisis regresi linier yang lebih lanjut, dengan tujuan meminimalisasi error dan memastikan hasil analisis akan mencapai sifat BLUE. Dengan demikian, tujuan dari dilaksanakan analisis regresi OLS dalam ekonometrika akan tercapai.

DAFTAR PUSTAKA

- F. Galton, "Family Likeness in Stature," *Proc. R. Soc. London*, vol. 40, no. 242–245, pp. 42–73, 1886, doi: <https://doi.org/10.1098/rspl.1886.0009>.
- K. Pearson and A. Lee, "On the Laws of Inheritance in Man: I. Inheritance of Physical Characters," *Biometrika*, vol. 2, no. 4, pp. 357–462, 1903, doi: <https://doi.org/10.2307/2331507>.
- G. Damodar, *Basic Econometrics*, Fourth Edi. New York: McGraw-Hill, 2004.
- T. Haavelmo, "The Probability Approach in Econometrics," *Econometrica*, vol. 12, pp. 1–115, 1944, doi: <https://doi.org/10.2307/1906935>.
- Setiawan and D. I. Kusriani, *Ekonometrika*. Yogyakarta: Andi Offset, 2010.
- I. Ghozali, *Ekonometrika*. Semarang: Badan Penerbit Universitas Diponegoro, 2014.
- A. C. Rencher, *Linear Models in Statistics*. New Jersey: John Wiley and Sons, 2008.
- D. Ruppert, M. P. Wand, and R. J. Carroll, *Semiparametric Regression*. United Kingdom: Cambridge University Press, 2003.
- R. Mubarak, *Pengantar Ekonometrika*. Pamekasan: Duta Media Publishing, 2021.
- D. Gujarati, *Dasar-dasar Ekonometrika*, Third Edit. Jakarta: Erlangga, 2007.
- A. T. Basuki and N. Purwoto, *Analisis Regresi dalam Penelitian Ekonomi dan Bisnis: Dilengkapi Aplikasi SPSS dan Eviews*. Jakarta: Rajawali Pers, 2016.
- G. Mardiatmoko, "PENTINGNYA UJI ASUMSI KLASIK PADA ANALISIS REGRESI LINIER BERGANDA (STUDI KASUS PENYUSUNAN PERSAMAAN ALLOMETRIK KENARI MUDA [CANARIUM INDICUM L.]) The Importance of the Classical Assumption Test in Multiple Linear Regression Analysis (A Case Study of," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 14, no. 3, pp. 333–342, 2020, doi: <https://doi.org/10.30598/barekengvol14iss3pp333-342>.
- J. W. Creswell, *Educational Research Planning, Conducting and Evaluating Quantitative and Qualitative Research*. Boston: Pearson Education, 2012.
- D. Kurniawan, *Regresi Linier (Linear Regression)*. Vienna: R Foundation for Statistical Computing, 2008.
- D. Gujarati, *Basic Econometrics*. New York: McGraw-Hill, 2003.
- I. Ghozali, *Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23*. Semarang: Badan Penerbit Universitas Diponegoro, 2016.