

PERBANDINGAN KEMAMPUAN REGRESI KUANTIL MEDIAN DAN TRANSFORMASI BOX-COX DALAM MENANGANI HETEROSKEDASTISITAS

Febria Pradita Prima Andani¹⁾, Edy Widodo²⁾

^{1), 2)} Program Studi Statistika FMIPA Universitas Islam Indonesia febriapradita@gmail.com,
edykafifa@gmail.com

Abstrak

Analisis regresi merupakan salah satu analisis statistik yang memanfaatkan hubungan antara dua atau lebih variabel, dengan tujuan mengetahui pengaruh variabel prediktor (X) terhadap variabel respon (Y). Dengan menggunakan analisis regresi dapat dilakukan pemodelan dan estimasi. Metode pendekatan standar untuk mendapatkan nilai dugaan parameter adalah Metode Kuadrat Terkecil (MKT). Pendugaan parameter dengan menggunakan MKT memiliki beberapa asumsi yang harus dipenuhi supaya mendapatkan penduga yang bersifat Best Linier Unbiased Estimation (BLUE). Salah satu asumsi tersebut adalah homoskedastisitas, yang berarti nilai varians residual adalah konstan (identik). Apabila asumsi nilai varians residual tidak konstan maka dikatakan terjadi heteroskedastisitas, yang dapat menyebabkan penduga parameter tidak memenuhi sifat BLUE sehingga model yang diperoleh tidak efisien dan tidak dapat dipercaya. Dalam penelitian ini dibandingkan metode regresi kuantil median dan transformasi Box-Cox dalam menangani masalah heteroskedastisitas. Didapatkan hasil perbandingan bahwa nilai R^2 MKT setelah dilakukan transformasi Box-Cox lebih besar dari nilai R^2 regresi kuantil median. Hal ini menunjukkan bahwa transformasi Box-Cox lebih baik daripada regresi kuantil median dalam menangani masalah heteroskedastisitas.

Kata Kunci: BLUE; Heteroskedastisitas; MKT; Regresi Kuantil; Transformasi Box-Cox

1. PENDAHULUAN

Analisis regresi adalah suatu analisis statistik yang memanfaatkan hubungan antara dua variabel atau lebih (Soejoeti, 1986). Tujuan dari analisis regresi adalah mengetahui hubungan antara dua atau lebih variabel prediktor (X) terhadap variabel respon (Y). Dengan menggunakan analisis regresi dapat dilakukan pemodelan dan estimasi. Menurut Qudratullah (2013), salah satu metode estimasi parameter dalam analisis regresi yang biasa digunakan adalah Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Square (OLS)*. Prinsip dari metode kuadrat terkecil yaitu dengan meminimumkan jumlah kuadrat residual. Menurut Uthami (2013), MKT memiliki beberapa asumsi yang harus dipenuhi untuk mendapatkan penduga yang memiliki sifat *Best Linier Unbiased Estimator (BLUE)*. Asumsi-asumsi yang harus dipenuhi yaitu, data berdistribusi normal, homoskedastisitas, tidak terdapat multikolinieritas, dan tidak terdapat autokorelasi. Apabila asumsi tersebut tidak terpenuhi, maka penduga tidak memenuhi sifat *BLUE*.

Apabila asumsi homoskedastisitas, yaitu residual memiliki variansi yang konstan, tidak terpenuhi, berarti varians residual tersebut tidak konstan. Hal ini disebut dengan heteroskedastisitas (Mendenhall, 1996). Menurut Mokosolang (2015), heteroskedastisitas dapat dideteksi dengan menggunakan

beberapa metode statistika, yaitu dengan metode grafik dan uji statistik. Uji statistik yang dapat digunakan dalam pendeteksian heteroskedastisitas antara lain Uji Glejser, Uji Park, Uji White, Uji Korelasi Rank Spearman, dan Uji Goldfeld-Quandt.

Heteroskedastisitas menyebabkan model yang akan diestimasi menjadi terganggu, karena model tersebut menjadi tidak efisien baik dalam sampel kecil maupun sampel besar, meskipun penduga menggambarkan populasi. Oleh karena itu, penanganan heteroskedastisitas penting untuk dilakukan supaya mendapatkan model yang dapat dipercaya dan tidak menyesatkan kesimpulan (Maziyya, 2015). Terjadinya kesalahan dalam menyimpulkan hasil penelitian akan menimbulkan permasalahan baru, seperti kesalahan dalam pengambilan kebijakan sehingga dapat memperburuk permasalahan atau kondisi yang diteliti. Maka sangatlah penting untuk menangani masalah heteroskedastisitas tersebut.

Penanganan heteroskedastisitas dapat dilakukan dengan beberapa metode, antara lain dengan Regresi Kuantil Median dan Transformasi Box-Cox (Cahyani, 2015). Menurut Uthami, dkk (2013), regresi kuantil bertujuan untuk memperluas ide-ide dalam estimasi model fungsi kuantil bersyarat, dimana distribusi kuantil bersyarat dari peubah respon dinyatakan sebagai fungsi dari kovariat yang diamati. Sedangkan transformasi Box-Cox bertujuan untuk menormalkan data, melinierkan model regresi, dan menghomogenkan varians (Cahyani, 2015).

Berdasarkan uraian di atas, maka penulis ingin mengetahui bagaimanakah kemampuan regresi kuantil dan transformasi Box-Cox dalam mengatasi heteroskedastisitas dalam melakukan estimasi parameter dalam analisis regresi. Hal ini sangat penting karena dengan mengetahui perbedaan hasil analisis dengan kedua metode tersebut, dapat diketahui metode manakah yang paling baik digunakan dalam penanganan heteroskedastisitas. Sehingga penanganan pada data heteroskedastik benar-benar menggunakan metode yang tepat dan didapatkan hasil analisis yang baik, yang mampu mendapatkan model yang baik, dapat dipercaya dan tidak menyesatkan peneliti dalam mengambil kesimpulan.

Penelitian sebelumnya mengenai regresi kuantil untuk mengatasi heteroskedastisitas telah dilakukan oleh Rahmawati (2011) dan Uthami (2013), dengan hasil bahwa regresi kuantil mampu untuk mengatasi masalah heteroskedastisitas. Sedangkan mengenai perbandingan regresi kuantil dan transformasi Box-Cox dalam mengatasi heteroskedastisitas, pernah juga dilakukan oleh Cahyani (2015).

1.1 Metode Kuadrat Terkecil

MKT adalah metode estimasi parameter yang sering digunakan dalam analisis regresi. MKT akan menemukan nilai-nilai estimasi dengan cara meminimumkan jumlah kuadrat residual (Draper dan Smith, 1992). Persamaan estimasi kuadrat terkecil adalah

dengan :

: Matriks penduga

: Inverse matriks dari tranpose matriks X dikalikan matriks X

: Tranpose matriks X dikalikan matriks Y

MKT memiliki asumsi-asumsi yang harus dipenuhi supaya dapat menilai kebaikan suatu persamaan regresi (Sembiring, 2003). Model regresi yang didapatkan dari MKT merupakan model regresi yang menghasilkan estimasi dengan sifat *BLUE* apabila asumsi-asumsinya terpenuhi. Asumsi-asumsi tersebut disebut dengan asumsi klasik, yaitu normalitas, homoskedastisitas, non-multikolinieritas, dan non-autokorelasi (Algifari, 1997).

1.2 Heteroskedastisitas

Heteroskedastisitas adalah kondisi dimana variansi dari variabel independen mempunyai nilai yang tidak konstan untuk setiap nilai variabel independen tersebut. Pemeriksaan heteroskedastisitas dapat dilakukan dengan pengujian Glesjer. Adanya heteroskedastisitas menyebabkan hal-hal berikut (Qudratullah, 2013).

- a. Penaksiran tetap *unbiased*, karena asumsi heteroskedastisitas tidak digunakan dalam membuktikan sifat *unbiased* parameter. Jadi, hal ini bisa dimaklumi.
- b. Variansi estimator akan salah, karena variansi estimator tersebut akan berubah-ubah nilainya atau tidak tetap.
- c. Estimator-estimator menjadi tidak efisien. Meskipun menggunakan MKT, estimator yang dihasilkan bersifat *unbiased* tetapi memiliki variansi yang lebih besar dari estimator lain yang bersifat bias.
- d. Prediksi terhadap koefisien populasi tidak akan tepat atau keliru.

1.3 Regresi Kuantil Median

Regresi kuantil median adalah salah satu pendekatan dalam analisis regresi dengan melakukan estimasi pada kuantil median, yaitu 0.5. Regresi kuantil median ini diperkenalkan oleh Roger Koenker dan Basset pada tahun 1978. Regresi kuantil baik dimanfaatkan ketika residual data tidak memenuhi asumsi homoskedastisitas, karena regresi kuantil merupakan metode yang tidak akan terpengaruh pada masalah heteroskedastisitas. Berikut adalah persamaan untuk meminimumkan kuantil tak bersyarat dengan estimasi regresi kuantil median. Misalkan data (Y_i, X_i) untuk $i = 1, 2, \dots, n$ dimana X_i berukuran $k \times 1$. Maka model linier persamaan regresi kuantil adalah sebagai berikut.

Dengan y_i adalah kuantil ke- i dari y dengan regresi kuantil ke- i yang diperoleh dengan meminimumkan jumlah nilai mutlak residual dengan pembobot w_i untuk residual. Selanjutnya untuk penduga regresi kuantil digunakan fungsi berikut (Koenker dan Basset, 1978)

$$\sum$$

dengan :

- : Penduga kuantil ke- i
- : *Loss function*
- : Variabel respon ke- i
- : Variabel bebas ke- i
- : Parameter regresi

1.4 Transformasi Box-Cox

Metode transformasi Box-Cox merupakan salah satu metode untuk melakukan transformasi pada variabel respon dengan melakukan pemangkatan (Y^λ) (Weisberg, 2005). Transformasi Box-Cox bertujuan untuk menormalkan data, menghomogenkan varians dan melinierkan model regresi. Nilai λ yang digunakan adalah nilai λ yang optimal. Misalkan dalam analisis regresi, nilai yang optimal tersebut didapatkan dari pencarian nilai λ yang meminimalkan jumlah kuarat residual. Berikut adalah model transformasi untuk masing-masing nilai λ .

Tabel 1. Model Transformasi untuk Masing-Masing Nilai λ
(Sumber: Kutner, dkk, 2004)

Nilai λ	Model Transformasi
-2	—
-1	-
-0.5	$\sqrt{-}$
0	
0.5	$\sqrt{-}$
1	
2	

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data penderita Demam Berdarah Dengue (DBD) (Y), pengguna jaminan kesehatan dari Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) (X_1), dan jumlah

penduduk (X_2) pada setiap kabupaten di Provinsi Jawa Tengah tahun 2014 yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Tengah.

Langkah pertama dalam penelitian ini adalah melakukan pengujian heteroskedastisitas pada data dengan menggunakan Uji Glejser. Selanjutnya melakukan estimasi regresi kuantil median pada dengan menggunakan *software R* i386 3.2.3. Setelah itu melakukan transformasi Box-Cox dengan menggunakan *software Minitab* 14. Kemudian melakukan estimasi dengan MKT. Langkah selanjutnya dilakukan pengujian heteroskedastisitas untuk melihat apakah masih terdapat masalah heteroskedastisitas pada model yang didapatkan. Kemudian membandingkan R^2 hasil regresi kuantil median dan MKT setelah dilakukan transformasi Box-Cox.

3. HASIL PENELITIAN DAN PEMBAHASAN

Berikut adalah hasil penelitian dan pembahasan mengenai penggunaan metode regresi kuantil median dan transformasi Box-Cox dalam analisis pengaruh pengguna jaminan kesehatan dari Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) dan jumlah penduduk terhadap penderita Demam Berdarah Dengue (DBD) pada setiap kabupaten di Provinsi Jawa Tengah. Data tersebut didapatkan dari *website* Badan Pusat Statistik (BPS) Jawa Tengah, yaitu jateng.bps.go.id.

a. Pengujian Heteroskedastisitas

Heteroskedastisitas adalah keadaan dimana residual variansi tidak konstan (identik). Adanya heteroskedastisitas menyebabkan model yang akan diestimasi menjadi terganggu karena model tersebut menjadi tidak efisien. Maka hal ini perlu ditangani. Untuk mendeteksi adanya heteroskedastisitas dapat dilakukan dengan Uji Glejser. Uji Glejser ini dilakukan dengan meregresikan nilai *absolute* residual terhadap variabel prediktornya. Kemudian lihat pada nilai estimasi β , apabila berpengaruh secara signifikan maka dikatakan terjadi heteroskedastisitas.

Hasil pendeteksian heteroskedastisitas pada data menunjukkan bahwa nilai *P-value* yang didapatkan adalah 0.004 untuk X_1 dan 0.000 untuk X_2 , nilai tersebut lebih kecil dari nilai tingkat signifikansi $\alpha=0.01$. Maka dapat disimpulkan bahwa terjadi heteroskedastisitas pada data.

b. Estimasi dengan Regresi Kuantil Median

Regresi kuantil median adalah salah satu metode analisis regresi yang tidak memerlukan asumsi homoskedastisitas, sehingga dapat digunakan untuk menyelesaikan masalah heteroskedastisitas. Hasil estimasi dengan menggunakan kuantil median, mendapatkan model sebagai berikut.

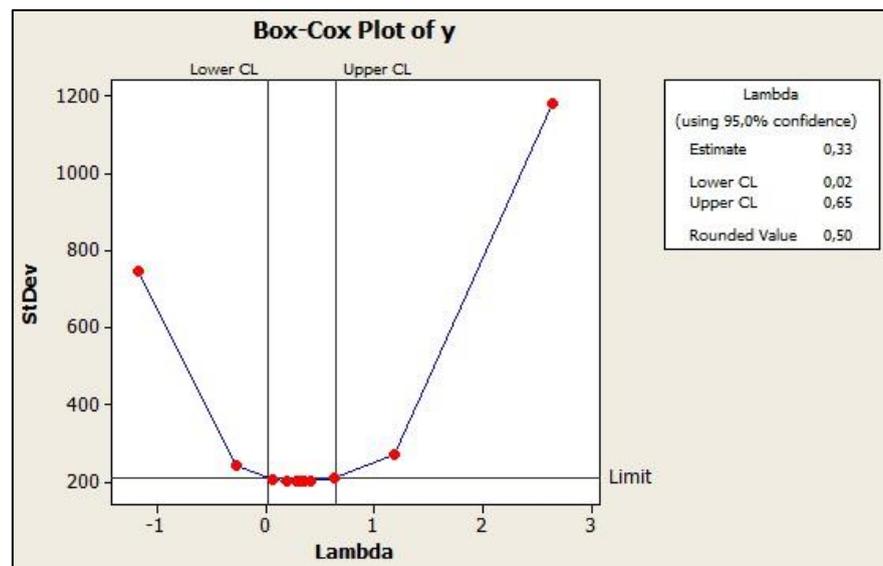
Nilai R^2 yang didapatkan adalah sebesar 36.05%. Hal ini menunjukkan bahwa variabel prediktor yang digunakan dalam model mampu menjelaskan variabel respon sebesar 36.05%, sedangkan sisanya dijelaskan oleh variabel lain di luar model.

Selanjutnya dilakukan pendeteksian heteroskedastisitas pada hasil estimasi tersebut. Hasil dari pengujian dengan menggunakan Uji Glejser didapatkan nilai p -value untuk X_1 adalah 0.001 dan untuk X_2 adalah 0.000. Karena p -value X_1 dan X_2 kurang dari tingkat signifikansi $\alpha=0.01$, maka terlihat bahwa masih terdapat masalah heteroskedastisitas.

c. Penggunaan Transformasi Box-Cox

Transformasi Box-Cox adalah metode transformasi pangkat atau kuasa yang digunakan pada data yang bernilai positif pada variabel Y . Dalam transformasi Box-Cox digunakan nilai λ optimal. Nilai λ optimal didapatkan dengan cara mencari nilai λ yang dapat meminimalkan jumlah kuadrat residual.

Dalam penelitian ini, transformasi Box-Cox dilakukan dengan bantuan *software* Minitab 14. Berikut adalah hasil transformasi Box-Cox pada data.



Gambar 1. Hasil Keluaran Minitab 14 berupa Plot Data Transformasi Box-Cox

Gambar tersebut menunjukkan bahwa nilai λ berada pada interval $-0.02 < \lambda < 0.65$. Namun nilai λ optimal ada pada 0.50, maka model transformasi yang tepat adalah transformasi $\sqrt{}$. Apabila transformasi telah dilakukan, maka selanjutnya adalah melakukan estimasi dengan MKT. Berikut adalah hasil dari estimasi MKT.

Nilai R^2 yang didapatkan adalah sebesar 47.22%. Hal ini menunjukkan bahwa variabel prediktor yang digunakan dalam model mampu menjelaskan variabel respon sebesar 47.22%, sedangkan sisanya dijelaskan oleh variabel lain di luar model.

Selanjutnya dilakukan pendeteksian heteroskedastisitas pada hasil estimasi tersebut. Hasil dari pengujian dengan menggunakan Uji Glejser didapatkan nilai p -value untuk X_1 adalah 0.096 dan untuk X_2 adalah 0.037. Karena p -value X_1 dan X_2 lebih besar dari tingkat signifikansi yang digunakan, yaitu $\alpha=0.01$, berarti tidak terdapat masalah heteroskedastisitas.

d. Perbandingan Regresi Kuantil Median dan Transformasi Box-Cox

Model yang baik dalam analisis regresi adalah model yang menghasilkan nilai R^2 yang besar. Dalam mengatasi heteroskedastisitas pada analisis pengaruh pengguna jaminan kesehatan dari Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) dan jumlah penduduk terhadap penderita Demam Berdarah Dengue (DBD) pada setiap kabupaten di Provinsi Jawa Tengah dengan menggunakan regresi kuantil median dan transformasi Box-Cox, didapatkan nilai R^2 sebagai berikut.

Tabel 2. Tabel Perbandingan R^2

R^2	
Regresi Kuantil Median	MKT
36,05%	47,22%

Tabel perbandingan tersebut menunjukkan bahwa nilai R^2 MKT setelah dilakukan transformasi Box-Cox lebih besar daripada nilai R^2 regresi kuantil median. Hal ini menunjukkan bahwa transformasi Box-Cox lebih baik dalam mengatasi heteroskedastisitas dibandingkan regresi kuantil median.

4. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa analisis regresi dengan MKT setelah dilakukan transformasi Box-Cox lebih baik dalam mengatasi heteroskedastisitas dibandingkan dengan menggunakan regresi kuantil median pada analisis pengaruh pengguna jaminan kesehatan dari Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) dan jumlah penduduk terhadap penderita Demam Berdarah Dengue (DBD) di setiap kabupaten di Provinsi Jawa Tengah.

5. DAFTAR PUSTAKA

Algifari. (1997). *Analisis Regresi Teori, Kasus dan Solusi*. Yogyakarta: BPFE.

- Cahyani, N. W. Y., Srinadi, I. G. A. M. & Susilawati, M. (2015). Perbandingan Transformasi Box-Cox dan Regresi Kuantil Median dalam Mengatasi Heteroskedastisitas. *E-Jurnal Matematika*, Vol. 4, No. 1: 8-13. Diakses dari <http://ojs.unud.ac.id/index.php/mtk/article/view/12283>
- Draper, N. R. & Smith, H. (1992). *Analisis Regresi Terapan*. Jakarta: PT. Gramedia Pustaka Utama.
- Koenker, R. & Basset, G. (1978). *Regression Quantile*. New York: Cambridge University Press.
- Kutner, M. H., Nachtsheim, C. J., Neter, J. & Li, W. (2004). *Applied Linear Statistical Models Fifth Edition*. New York: McGraw-Hill.
- Maziyya, P. A., Sukarsa, I. K. G. & Asih, N. M. (2015). Mengatasi Heteroskedastisitas pada Regresi dengan Menggunakan Weighted Least Square. *E-Jurnal Matematika*, Vol. 4, No. 1: 20-25. Diakses dari <http://ojs.unud.ac.id/index.php/mtk/article/view/12285>
- Mendenhall, W., & Sincich, T. (1996). *A Second Course in Statistics: Regression Analysis, Seventh Edition*. United States of America: Pearson Education, Inc.
- Mokosolang, C. A., Prang, J. D. & Mananohas, M. L. (2015). Analisis Heteroskedastisitas pada Data Cross Section dengan White Heteroscedasticity Test dan Weighted Least Squares. *JdC*, Vol. 4, No. 2. Diakses dari <http://ejournal.unsrat.ac.id/index.php/decartesian/article/view/9056>
- Qudratullah, M. F. (2013). *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS*. Yogyakarta: Penerbit ANDI.
- Rahmawati, R., Widiarti & Novianti, P. (2011). Regresi Kuantil (Studi Kasus pada Data Suhu Harian). *Peran dan Implementasi Statistika dalam Analisis Finansial dan Pengambilan Keputusan Bisnis: Prosiding Seminar Nasional Statistika*. Diselenggarakan oleh Program Studi Statistika, Undip, 21 MEI 2011. Semarang: Universitas Diponegoro. Diakses dari <http://eprints.undip.ac.id/39118/1/4.kuantil.pdf>
- Sembiring, R. K. (2003). *Analisis Regresi Edisi Kedua*. Bandung: Penerbit ITB.
- Soejoeti, Z. (1986). *Buku Materi Pokok Metode Statistika II*. Jakarta: Karunika Jakarta.
- Uthami, I. A. P., Sukarsa, I. K. G. & Kencana, I. P. E. N. (2013). Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas pada Analisis Regresi. *E-Jurnal Matematika*, Vol. 2, No. 1. Diakses dari <http://ojs.unud.ac.id/index.php/mtk/article/view/4911>
- Weisberg, S. (2005). *Applied Linear Regression*. USA: John Wiley & Sons, Inc.