Jurnal Matematika UNAND
Vol. **5** No. **1** Hal. 125 – 130
ISSN : 2303–291X

ANALISIS ESTIMASI PARAMETER REGRESI KUANTIL DENGAN METODE *BOOTSTRAP*

MESI OKTAFIA, FERRA YANUAR, MAIYASTRI
Program Studi Magister Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas
Limau Manis Padang, Sumatera Barat, Indonesia,
email : mesioktafia@yahoo.co.id

Abstrak. Regresi kuantil merupakan salah satu metode regresi dengan pendekatan memisahkan atau membagi data menjadi kuantil-kuantil tertentu. Metode pendekatan yang dilakukan adalah dengan meminimumkan error mutlak berbobot yang tidak simetris dan menduga fungsi kuantil bersyarat pada suatu sebaran data. Pendugaan parameter regresi kuantil ini tidak membutuhkan asumsi parametertik. Parameter model yang dihasilkan kemudian diuji keakuratannya dengan menggunakan metode *Bootstrap*. Metode *Bootstrap* merupakan suatu teknik pendekatan nonparametrik untuk menaksir berbagai kuantitas statistik seperti *mean*, *standar error*, dan bias suatu estimasi atau untuk membentuk interval konfidensi dengan mengikuti algoritma tertentu. Pada kajian ini estimasi parameter regresi kuantil dengan metode *Bootstrap* diperoleh bahwa hasil estimasi parameter regresi kuantil sudah cukup akurat, karena nilai estimasi parameter regresi kuantil dengan metode *Bootstrap* hampir mendekati nilai estimasi regresi kuantil untuk data simulasi. Selanjutnya estimasi parameter regresi kuantil dengan metode *Bootstrap* diperoleh nilai *mean square error* (MSE) yang cukup kecil untuk setiap estimasi parameter pada setiap kuantilnya, ini mengindikasikan bahwa nilai estimasi parameter regresi kuantil dengan metode *Bootstrap* sudah cukup baik.

Kata Kunci: Regresi kuantil, metode Bootstrap

1. Pendahuluan

Analisis regresi adalah metode statistika yang paling sering digunakan dalam segala bidang ilmu pengetahuan, analisis ini bertujuan untuk memodelkan hubungan antara dua variabel yang terdiri dari variabel tak bebas (Y) dengan satu atau lebih variabel bebas (X) dalam suatu sistem. Hubungan antara variabel-variabel tersebut biasanya dinyatakan dalam suatu model yang biasanya diistilahkan sebagai model regresi yang secara umum dinyatakan sebagai $Y = f(x) + e$, e adalah komponen galat (error), model tersebut menghubungkan variabel bebas dan variabel tak bebas melalui suatu parameter yang dinamakan sebagai parameter regresi dinotasikan dengan β.

Metode regresi yang biasa digunakan oleh para analisis regresi dan merupakan dasar teknik regresi adalah metode kuadrat terkecil (MKT) atau metode regresi klasik, dimana prinsip kerjanya adalah meminimumkan jumlah kuadrat penyimpangan atau error nilai observasi terhadap rata-ratanya, dalam metode kuadrat terkecil ini disyaratkan memenuhi asumsi yang ada, namun metode ini dikenal peka terhadap penyimpangan asumsi pada data, misalnya jika data tidak memenuhi salah satu asumsi regresi maka penduga motode kuadrat terkecil tidak lagi baik digunakan, maka berkembanglah metode regresi median (*Median Regression*), metode regresi median dilakukan dengan mengganti pendekataan rataan (*mean*) pada MKT menjadi median. Hal ini dilakukan dengan mempertimbangkan apabila data berbentuk lonceng atau tidak simetris. Tetapi pada kenyataannya, pendekatan regresi median juga dianggap kurang tepat karena regresi ini hanya melihat pada dua kelompok data. Padahal ada kemungkinan data bisa terbagi menjadi lebih dari dua kelompok, sehingga berkembanglah metode regresi kuantil (*Quantile Regresion*). Regresi kuantil pertama kali diperkenalkan oleh Koenker dan Bassett [5], metode ini merupakan salah satu metode regresi dengan pendekatan memisahkan atau membagi data menjadi kuantil-kuantil tertentu, dengan meminimumkan error mutlak berbobot yang tidak simetris dan menduga fungsi kuantil bersyarat pada suatu sebaran data.

Pada penelitian ini dilakukan analisis pendugaan paremanter regresi kuantil dengan metode *Bootstrap*, karena dengan metode *Bootstrap* dapat menghitung tingkat keakuratan nilai estimasi setiap parameter regresi kuantil. Metode *Bootstrap* pertama kali diperkenalkan oleh Efron pada tahun 1979. Metode *Bootstrap* digunakan untuk mengestimasi koefisien dari suatu persamaan regresi dengan melakukan penyampelan ulang dari sampel yang sudah ada. Metode *Bootstrap* juga digunakan untuk data yang tidak memenuhi asumsi kenormalan. Metode *Bootstrap* dapat menghasilkan nilai-nilai statistik yang digunakan untuk menentukan selang kepercayaan setiap parameter model yang diestimasi. Metode kuantil *Bootstrap* tersebut, kemudian akan diilustrasikan pada simulasi data yang dibangkitkan.

2. Data dan Metode Penelitian

Data yang digunakan dalam penelitian terdiri dari dua variabel independen (X_1 dan X_2) dan satu variabel dependen (Y), dimana variabel independen (X_1 dan X_2) bersifat stokastik (tidak deterministik). Variabel independen (X_1 dan X_2) masing-masingnya menyebar menurut sebaran normal ($X1 \sim N(0,1)$ dan $X_2 \sim N(0,1)$), sedangkan variabel dependen (Y) ditetapkan nilai $Y = 2X_1 + 4X_2 + e$, dimana $e = 0.2 \times Z$ dengan $Z \sim N(0,1)$.

Analisis regresi adalah suatu teknik yang digunakan untuk mencari hubungan antara dua atau lebih variabel kuantitatif sehingga suatu variabel dapat diprediksikan dari satu atau beberapa variabel yang lain. Pada penelitian ini merupakan regresi linier berganda terdiri dari satu variabel dependen (Y) dan beberapa variabel independen (X). Pada analisis regresi klasik, pendugaan parameter dilakukan dengan metode kuadrat terkecil prinsipnya meminimumkan jumlah kuadrat error. Jumlah kuadrat error secara matematis dapat dituliskan sebagai berikut:

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \tag{2.1}$$

Dalam melakukan pendugaan parameter dengan menggunakan metode MKT terdapat beberapa asumsi yang harus dipenuhi [4]. Pada saat melakukan estimasi dengan metode MKT kemudian ada asumsi yang tidak terpenuhi, maka model yang diperoleh tidak efisien dan tidak dapat dipercaya akibatnya dapat menyesatkan kesimpulan yang diambil dari model regresi yang dibentuk. Sedangkan pada analisis regresi kuantil tidak membutuhkan asumsi parametrik [1]. Kuantil adalah suatu teknik pembagian sekelompok data kepada beberapa bagian yang sama, setelah data diurutkan dari yang paling kecil atau paling besar. Regresi Kuantil merupakan suatu pendekatan dalam analisis regresi yang dikenalkan oleh Koenker dan Bassett [5]. Pendekatan ini menduga berbagai fungsi kuantil dari suatu distribusi Y sebagai fungsi dari X. Penggunaan metode regresi ini dilakukan dengan pembagian atau pemisahan data menjadi beberapa kelompok yang dicurigai mempunyai perbedaan nilai dugaan pada kuantil-kuantil tersebut. Secara umum, regresi kuantil sangat bermanfaat ketika ingin menganalisis bagian tertentu dari suatu sebaran bersyarat. Model linier dari persamaan regresi kuantil yang terdiri dari satu variabel independen Y dan dua atau lebih variabel independen X, yaitu:

$$Y_i = X_{i1}\beta_1(\theta) + X_{i2}\beta_2(\theta) + \ldots + X_{ip}\beta_p(\theta) + e_i \qquad \text{untuk } i = 1, \ldots, n.$$

Dalam bentuk matriks, persamaan diatas ditulis sebagai berikut:

$$\mathbf{Y} = \mathbf{X}\beta(\theta) + \mathbf{e},$$

dengan:

$$\begin{aligned} \mathbf{Y} &= \text{vektor berukuran } n \times 1 \text{ dari variabel dependen} \\ \mathbf{X} &= \text{matriks berukuran } n \times p \text{ dari variabel independen} \\ \beta(\theta) &= \text{vektor berukuran } n \times 1 \text{ dari koefisien regresi kuantil yang bergantung} \\ &\quad\ \text{pada } \theta(0 < \theta < 1) \\ \mathbf{e} &= \text{vektor berukuran } n \times 1 \text{ dari error.} \end{aligned}$$

Estimasi dengan metode regresi kuantil ke-θ diperoleh dengan meminimumkan jumlah nilai mutlak dari error dengan pembobot θ untuk error positif dan pembobot $(1-\theta)$ untuk error negatif yaitu [5]:

$$\min_{\beta \in R^p} \sum_{i=1}^{n} \rho_p(y_i - Q_\theta(Y|X)), \tag{2.2}$$

dimana p menyatakan indeks kuantil $\in (0,1)$, ρ_p merupakan *loss function* yang asimetrik, dimana $0 < p < 1$, dan $\rho_p(\varepsilon)$ merupakan *Loss Function* dari regresi kuantil yaitu:

$$\rho_p(\varepsilon) = \begin{cases} p\varepsilon & , \text{ jika } \ \varepsilon \geq 0, \\ (p-1)\varepsilon & , \text{ jika } \ \varepsilon < 0. \end{cases} \tag{2.3}$$

Analisis estimasi regresi kuantil dilakukan dengan metode *Bootstrap*. Metode *Bootstrap* merupakan metode yang digunakan untuk mengestimasi suatu distribusi populasi yang tidak diketahui dimana distribusi empirisnya diperoleh dari proses penyampelan ulang [3]. Metode *Bootstrap* merupakan suatu teknik pendekatan nonparametrik untuk menaksir berbagai kuantitas statistik seperti *mean*, *standar error*, dan bias suatu estimasi atau untuk membentuk interval konfidensi dengan mengikuti algoritma tertentu. Berdasarkan [2] nilai rata-rata parameter *Bootstrap* yaitu:

$$\overline{\widehat{\beta}}(\theta) = \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}_b(\theta). \tag{2.4}$$

Sedangkan selang kepercayaan yang digunakan dalam penelitian ini yaitu selang kepercayaan persentil. Metode persentil didasarkan pada $(\widehat{\beta}(\theta_q)_{lo})$ persentil ke α dan $(\widehat{\beta}(\theta_q)_{up})$ persentil ke $(1-\alpha)$ fungsi distribusi kumulatif dari estimasi parameter vektor *Bootstrap* yaitu:

$$[\widehat{\beta}(\theta_q)_{lo}, \widehat{\beta}(\theta_q)_{up}] = [\widehat{F}(\alpha), \widehat{F}(1-\alpha)],$$

dimana *lo* dan *up* masing-masing merupakan batas bawah dan batas atas dari selang kepercayaan [2]. Selanjutnya melihat kebaikan penduga parameter berdasarkan nilai bias dan ragam secara bersamaan maka diperlukan *Mean Square Error*(MSE) [6].

$$MSE(\widehat{\beta}_j(\theta_q)) = Var(\widehat{\beta}_j(\theta_q)) + Bias(\widehat{\beta}_j(\theta_q))^2. \tag{2.5}$$

3. Hasil dan Pembahasan

Analisis terhadap regresi klasik dan regresi kuantil dengan metode *Bootstrap* kemudian akan diilustrasikan penerapannya dengan menggunakan data yang dibangkitkan dengan menggunakan *software R*. Estimasi parameter dengan Regresi Klasik diberikan pada Gambar 1.

	Estimasi Parameter Regresi Klasik	p-value	Standar error
β_0	0,00764	0,749	0,02382
β_1	2,04086 **	0	0,02756
β_2	4,00776 **	0	0,02686

(**) Signifikan pada taraf $\alpha = 0{,}05$

Gambar 1. Hasil Estimasi Parameter Regresi Klasik

Pada Gambar 1 dapat diketahui bahwa estimasi parameter regresi klasik untuk β_1 dan β_2 signifikan pada taraf $\alpha = 0.05$. Sedangkan estimasi koefisien regresi untuk β_0 tidak signifikan pada taraf $\alpha = 0.05$.

Pada saat melakukan estimasi dengan regresi klasik kemudian ada asumsi yang tidak terpenuhi, maka model yang diperoleh tidak efisien dan tidak dapat dipercaya akibatnya dapat menyesatkan kesimpulan yang diambil dari model regresi yang dibentuk. Sehingga dilakukanlah analisis regresi kuantil.

Pada Gambar 2 diberikan estimasi parameter dengan regresi kuantil. Dapat diketahui bahwa estimasi parameter regresi untuk β_1 dan β_2 signifikan pada taraf $\alpha = 0.05$. Sedangkan estimasi parameter regresi untuk β_0 tidak signifikan, yaitu pada kuantil ke-0.4 dan 0.5. Koefisien intersep pada kuantil ke 0.1 sampai 0.4 bernilai negatif, sementara untuk kuantil ke 0.5 sampai 0.9 bernilai positif. Sedangkan

Kuantil θ	Estimasi Parameter Regresi Kuantil					
	β_0 (se)	p-value	β_1(se)	p-value	β_2 (se)	p-value
0.1	-0,16523 ** (0,03498)	0,00001	2,06568 ** (0,04331)	0	4,01547 ** (0,03099)	0
0.2	-0,22227 ** (0,03616)	0	2,00673 ** (0,04464)	0	4,04865 ** (0,03067)	0
0.3	-0,09374 ** (0,03364)	0,00668	2,05088 ** (0,04065)	0	4,01692 ** (0,03814)	0
0.4	-0,04550 (0,03245)	0,16489	2,03395 ** (0,03481)	0	3,98515 ** (0,03389)	0
0.5	0,00056 (0,03065)	0,98545	2,02366 ** (0,04061)	0	3,98846 ** (0,02700)	0
0.6	0,06513 ** (0,03065)	0,03678	2,02745 ** (0,03577)	0	3,96080 ** (0,02932)	0
0.7	0,08566 ** (0,02804)	0,00309	2,03441 ** (0,03020)	0	3,96445 ** (0,03189)	0
0.8	0,15119 ** (0,04624)	0,00161	2,05935 ** (0,04654)	0	3,98025 ** (0,04432)	0
0.9	0,28915 ** (0,05833)	0,00001	2,06181 ** (0,07611)	0	4,03084 ** (0,05953)	0

(**) Signifikan pada taraf $\alpha = 0{,}05$ se = standar error

Gambar 2. Hasil Estimasi Parameter Regresi Kuantil

untuk koefisien variabel independen X_1 dan X_2 semuanya bernilai positif pada setiap kuantil yang dipilih artinya terjadi hubungan yang positif antara variabel X_1 dan X_2 dengan variabel Y, semakin naik nilai X_1 dan X_2 maka semakin meningkat nilai Y.

Selanjutnya analisis dilakukan dengan metode *Bootstrap* dengan jumlah replikasi yang digunakan yaitu replikasi sebanyak 25 kali, 50 kali, dan 100 kali. Artinya dilakukan 25 kali, 50 kali, dan 100 kali resampling dengan pengembalian. Diperoleh bahwa nilai estimasi parameter regresi kuantil dengan metode *Bootstrap* menghasilkan nilai yang hampir mendekati nilai estimasi regresi kuantil untuk setiap data simulasi yang dibangkitkan, kemudian hasil estimasi parameter regresi kuantil untuk semua parameter model terdapat didalam selang kepercayaan persentil *Bootstrap*, kemudian dianalisis kebaikan penduga parameter regresi kuantil dengan metode *Bootstrap* berdasarkan nilai bias dari metode *Bootstrap* dan variansi dari *Bootstrap* yaitu dengan diperoleh nilai *mean square error* (MSE) masing masing parameter cendrung mengecil pada posisi tengah kuantil.

4. Kesimpulan

Hasil estimasi parameter regresi kuantil dapat dilihat keakuratannya dengan membandingkan hasilnya dengan hasil estimasi parameter dengan metode *Bootstrap*. Pada kajian ini nilai estimasi rata-rata dari masing-masing replikasi dibandingkan dengan hasil estimasi parameter regresi kuantil untuk melihat kedekatan antara ke-dua nilai tersebut. Diperoleh bahwa nilai estimasi *Bootstrap* sudah mendekati nilai estimasi parameter regresi kuantil, karena nilai estimasi parameter regresi kuantil berada dalam selang kepercayaan persentil *Bootstrap*. Selanjutnya hasil analisis yang dilakukan berdasarkan nilai MSE parameter regresi kuantil dengan metode *Bootstrap* disimpulkan bahwa model dugaan yang lebih baik digunakan yaitu model dugaan pada posisi tengah kuantil.

Daftar Pustaka

[1] Buhai S. 2005. *Quantile Regression: Overview and Selected Applications*. ad-Astra

[2] Davino, Cristina., dkk. 2014.*Quantile Regression Theory and Applications*. Pondicherry. India.

[3] Efron, Bradley & Tibshirani, J.Robert. 1993. *An Introduction to the Bootstrap*. New York: Champman & Hall, Inc.

[4] Gujarati. 2003. *Essential Of Econometrics*. 4th Edition, The McGraw-Hill Companies. New Jersey.

[5] Koenker, R. dan Basset, G.Jr. 1978. Regression Quantiles. *Econometrica*. **46** : 33 – 50.

[6] Mendenhall, W., Sheaffer, R.L dan Wackerly, D.D. 1990. *Mathematical Statistics with Application Fourth Edition*. Wasworth. California.