PENGENDALIAN KOEFISIEN REGRESI LEAST *ABSOLUTE DEVIATION* PADA RENTANG BERMAKNA MENGGUNAKAN PROGRAM LINIER

Controlling Least Absolute Deviation Regression Coefficient on the Meaningful Range Using Linear Programming

Setyono[1], Agus Mohamad Soleh[2], Nur Rochman[1]

[1]*Jurusan Agroteknologi, Universitas Djuanda Bogor, Jl. Tol Ciawi 1 Bogor 16720, Indonesia*
[2]*Departemen Statistika, Institut Pertanian Bogor, Jl. Meranti, Kampus IPB Dramaga, Bogor 16680, Indonesia*
Telp. (0251)8240773 Fax. (0251)8240985
E-mail: setyono@unida.ac.id, agusms@apps.ipb.ac.id, nurdhr88@gmail.com

(Makalah diterima, 30 Januari 2018 – Disetujui, 04 Juni 2018)

ABSTRAK

Selama ini analisis regresi digunakan untuk memodelkan nilai tengah respons sebagai fungsi dari beberapa peubah bebas, menggunakan metode kuadrat terkecil (*least squares* - LS). Secara umum metode LS mampu menggambarkan ukuran pemusatan dengan baik, tetapi tidak kekar terhadap pencilan. Oleh sebab itu, pada kasus tertentu dibutuhkan regresi yang meminimumkan jumlah sisaan mutlak (*least absolute deviation* - LAD), yang dikenal lebih kekar. Sejauh ini, nilai koefisien regresi tidak pernah dimodelkan dan bergantung sepenuhnya pada data yang diolah. Pada beberapa kasus, tanda dan nilai koefisien regresi perlu dikendalikan agar berada pada rentang yang bermakna. Hasil penelitian ini menunjukkan modifikasi kendala pada regresi LAD mampu mengendalikan koefisien regresi untuk berada pada rentang yang bermakna. Hasil *bootstrap* menunjukkan sebaran koefisien regresi yang dikendalikan berbeda dengan yang tidak dikendalikan.

Kata kunci: program linier, sisaan mutlak, kendala, LAD, median, regresi

ABSTRACT

So far, regression analysis is used to model the mean of response variable as a function of some independent variables, using the least squares (LS) method. In general, the LS method is able to describe well the measure of central tendency, however it is not robust against outliers. Therefore, in certain cases, a regression analysis that minimizes the sum of absolute residuals (least absolute deviation - LAD) is required, which is more robust against outliers. So far, the value of the regression coefficient is not modeled and only depends entirely on the data processed. In some cases, the sign and the value of regression coefficients need to be controlled, in order to be in the meaningful range. The results of this study showed that the modification of the constraints on the LAD regression able to control the regression coefficients to be in the meaningful range. The results of bootstrap showed that distribution of controlled regression coefficients were different from distribution of uncontrolled regression coefficients.

Key words: *linear programming, absolute residual, constraints, LAD, median, regression*

PENDAHULUAN

Analisis regresi digunakan untuk memodelkan nilai tengah respons sebagai fungsi dari beberapa peubah bebas. Metode yang biasa digunakan untuk menduga koefisien regresi adalah metode yang meminimumkan jumlah kuadrat sisaan (*least squares* - LS). Secara umum metode LS mampu menggambarkan ukuran pemusatan dengan baik, tetapi tidak kekar terhadap pencilan. Oleh sebab itu, pada kasus tertentu dibutuhkan regresi yang meminimumkan jumlah sisaan mutlak (*least absolute deviation* - LAD), yang lebih kekar. Penduga LAD untuk ukuran pemusatan tidak khas, salah satunya adalah median (Hao dan Naiman, 2007). Median adalah salah satu ukuran kuantil, dan regresi LAD sudah dikembangkan menjadi regresi kuantil yang dirintis oleh Koenker dan Bassett (1978) dan dikembangkan antara lain oleh Koenker dan Hallock (2001).

Pada model linier Y = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p + \varepsilon$, ada beberapa hal yang dapat dimodelkan, tetapi yang sudah dimodelkan adalah sebaran peubah respon Y, kovariat X_j yang terlibat, dan hubungan fungsional peubah Y terhadap X_j. Tanda dan besar nilai β_j diserahkan sepenuhnya kepada data dan hubungan fungsional yang dimodelkan. Dengan demikian, sampai saat ini rentang nilai β_j belum pernah dimodelkan dan belum pernah dikendalikan.

Kenyataannya, nilai β_j tidak tidak terbatas karena ada kisaran nilai yang layak untuk β_j. Sebagai contoh pada bidang pertanian, produksi tanaman (Y) dimodelkan sebagai fungsi dari nutrisi (X) dalam bentuk sederhana $Y = \beta_0 + \beta_1 X + \varepsilon$ dengan kondisi $\beta_0 \geq 0$. Di bidang ekonomi, penawaran (Q) dimodelkan sebagai fungsi dari harga (P) dalam bentuk $Q = \beta_0 + \beta_1 P + \varepsilon$ dengan kondisi $\beta_1 > 0$ (Hansen, 2013). Jadi model linier seharusnya juga mengendalikan kisaran nilai koefisien regresi.

Koefisien regresi yang tidak wajar sering dijumpai pada regresi berganda. Sebagai contoh, bobot buah per tanaman merupakan fungsi dari tinggi tanaman, jumlah cabang, luas daun, jumlah buah, panjang buah, diameter buah, dan sebagainya, yang berperan sebagai peubah bebas (kovariat). Di antara peubah bebas tersebut terjadi korelasi positif, dan antara masing-masing peubah bebas dengan bobot buah per tanaman juga terjadi korelasi positif (Murniati *et al*., 2012). Dengan demikian, kalau dibuat regresi linier sederhana antara bobot buah per tanaman dengan masing-masing peubah bebas akan diperoleh koefisien regresi yang bernilai positif. Namun, jika dibuat regresi linier berganda bobot buah per tanaman terhadap semua peubah bebas mungkin diperoleh koefisien regresi yang tidak semuanya positif. Akibatnya, tanda koefisien regresi menjadi tidak masuk akal, karena meskipun semakin besar X semakin besar Y tetapi koefisien regresinya negatif.

Konsekuensi dari korelasi di antara peubah bebas (kolinear ganda) adalah galat baku koefisien regresi menjadi besar, bahkan kalau korelasinya sempurna maka koefisien regresi menjadi tidak khas atau tidak dapat diduga karena matriks X'X singular (determinan nol). Solusi yang dapat ditempuh adalah melakukan reduksi peubah, menerapkan regresi gulud (*ridge regression*), atau regresi komponen utama (Draper dan Smith, 1992). Penerapan reduksi peubah dapat berdampak pada kesalahan interpretasi, yaitu seolah Y tidak dipengaruhi oleh peubah yang dikeluarkan (tidak tercantum dalam model), padahal jika dibuat regresi linier sederhana dengan peubah yang dikeluarkan tadi maka koefisien regresinya signifikan.

Regresi lebih sering digunakan untuk menganalisis data survei dibandingkan dengan data percobaan. Pada data survei (pengambilan contoh), meskipun tidak terjadi kolinear ganda, tanda koefisien regresi yang tidak sesuai teori atau tidak logik masih mungkin terjadi akibat *sampling error*. Atau bisa saja tanda koefisien regresi masih sesuai tetapi nilainya melenceng dari koefisien regresi yang dihasilkan dari keseluruhan data (populasi) hasil sensus.

Koefisien regresi dengan tanda yang masuk akal merupakan hal penting karena regresi tidak hanya berfungsi untuk menduga nilai peubah respon (tak bebas) berdasarkan nilai-nilai peubah bebas, melainkan juga berfungsi menggambarkan hubungan fungsional peubah tak bebas sebagai fungsi dari peubah bebas. Tanpa pengendalian, pada pendugaan nilai respon tidak ada masalah, namun koefisien regresi menjadi tidak bermakna atau memberi makna yang salah. Dengan demikian, masalah yang dihadapi adalah bagaimana mengendalikan tanda dan rentang nilai koefisien regresi.

Berdasarkan uraian tersebut dapat dirumuskan bagaimana mendapatkan metode regresi yang mampu mengendalikan tanda dan besaran koefisien regresi agar berada pada rentang yang bermakna. Pendugaan regresi untuk masalah tersebut dapat dikerjakan dengan metode LAD dengan pertimbangan dapat dikerjakan dengan program linier sehingga berpotensi dapat dimodifikasi kendala untuk mencapai tujuan yang diinginkan. Pengembangan regresi melalui program linier dapat menggunakan program R, karena sudah diproduksi beberapa library untuk pemrograman linier seperti boot, linprog, dan Rglpk (Venables *et al*., 2016). Sebagai panduan komputasi dapat merujuk pada Rizzo (2008), Givens dan Hoeting (2005), serta Venables dan Ripley (2002).

Tujuan penelitian ini adalah mendapatkan metode pendugaan koefisien regresi yang dapat mengendalikan tanda dan nilai koefisien regresi agar berada pada rentang yang bermakna.

BAHAN DAN METODE

Program Linier untuk Regresi LAD

Pemrograman linier adalah memaksimumkan atau meminimumkan fungsi tujuan berupa kombinasi linier p peubah, yaitu $z=c_1x_1+c_2x_2+...+c_px_p$, dengan kendala sekumpulan persamaan atau pertidaksamaan. Pada program linier sederhana solusi dapat diperoleh melalui pendekatan grafis, sedangkan pada program linier yang melibatkan banyak kendala dan banyak peubah solusi dapat diperoleh menggunakan metode simplex. Metode simplex dikenalkan tahun 1947 oleh George B Danzig (1914-2005) dalam bukunya berjudul *Linear Programming and Extensions* yang dipublikasi pada tahun 1963 (Cottle *et al*., 2007).

Pada model regresi $y_i=\beta_0+\beta_1 x_i+\varepsilon_i$, misalnya b_0 adalah penduga β_0 dan b_1 adalah penduga β_1, sehingga model dugaannya adalah $y_i=b_0+b_1x_i+e_i$. Regresi LAD dapat dituliskan dalam argumen $min\left\{\sum_{i=1}^{n}|e_i|\right\}$.. Meminimumkan $\sum_{i=1}^{n}|e_i|$ tidak termasuk permasalahan program linier, tetapi dapat ditransformasi ke dalam masalah program linier dengan mengganti e_i dengan selisih dua peubah non negatif, yaitu e_i^+ dan e_i^-, sehingga $e_i=e_i^+-e_i^-$ (Ash, 2000).

Misalnya $e_i^+\geq 0$ adalah sisaan positif dan $e_i^-\geq 0$ adalah nilai mutlak dari sisaan negatif. Ketika sisaan bernilai negatif berlaku

$$y_i - \mathbf{x}_i'\mathbf{b} \leq 0 \Leftrightarrow e_i \leq 0 \Leftrightarrow e_i^+ = 0 \text{ dan } e_i^- \geq 0 \Leftrightarrow \mathbf{x}_i'\mathbf{b} - e_i^- = y_i$$

Sementara ketika sisaan bernilai positif berlaku

$$y_i - \mathbf{x}_i'\mathbf{b} \geq 0 \Leftrightarrow e_i \geq 0 \Leftrightarrow e_i^+ \geq 0 \text{ dan } e_i^- = 0 \Leftrightarrow \mathbf{x}_i'\mathbf{b} + e_i^+ = y_i$$

Ketika sisaan bernilai positif maka $e_i^+>0$ dan $e_i^-=0$ dan ketika sisaan bernilai negatif maka $e_i^->0$ dan $e_i^+=0$, sehingga $|e_i|=|e_i^+-e_i^-|=|e_i^+|+|e_i^-|=e_i^++e_i^-$. Akibatnya, dua persamaan di atas dapat diagregasi menjadi:

$$\mathbf{x}_i'\mathbf{b} + e_i^+ + e_i^- = y_i$$

yang berfungsi sebagai kendala untuk fungsi obyektif berikut:

$$\min \sum_{i=1}^{n} (e_i^+ + e_i^-)$$

Setiap pengamatan memiliki nilai e_i^+ dan e_i^- masing-masing, sehingga untuk n pengamatan dengan p parameter dibutuhkan (2n+p) variabel pada program liniernya. Solusi secara pemrograman linier menjadi rumit dan membutuhkan waktu lama terutama untuk n besar. Untuk mempelajari pemrograman linier lebih detil dapat merujuk pada McCarl dan Spreen (2003) atau Winston dan Goldberg (2004), sedangkan untuk mengimplementasikannya pada bahasa R dapat merujuk pada Rizzo (2008).

Program Linier untuk Pengendalian Rentang Nilai Koefisien Regresi LAD

Rentang nilai koefisien regresi dikendalikan dengan cara memberi tambahan kendala pada program linier. Kendala yang diberikan bergantung pada koefisien peubah yang dikendalikan dan pada rentang berapa koefisien tersebut dikendalikan. Misalnya model regresi yang digunakan adalah $Y=\beta_0+\beta_1X1+\beta_2\ X_2+\varepsilon$ dengan kendala $\beta_1+\beta_2$=0,3, sedangkan banyaknya data ada n pengamatan. Regresi LAD dilakukan dengan fungsi obyektif

$$\min \sum_{i=1}^{n} (e_i^+ + e_i^-)$$

dengan kendala $\mathbf{x}_i'\mathbf{b} + e_i^+ + e_i^- = y_i$ and b_1+b_2=0,3

Data untuk Aplikasi Pengendalian

Data yang digunakan dalam aplikasi pengendalian koefisien regresi adalah hasil pengambilan contoh secara acak berupa 30 tanaman cabai merah (*Capsicum annuum* L.) varietas TM-999 di Kebun Percobaan Pusat Pelatihan Pertanian dan Pedesaan Swadaya (P4S) Antanan di Kampung Tarikolot, Desa Cimande, Kecamatan Caringin, Bogor, Jawa Barat. Pengamatan dan pengukuran dilaksanakan oleh Murniati *et al*. (2012).

HASIL DAN PEMBAHASAN

Komputasi Regresi LAD Menggunakan Program Linier

Penduga LAD untuk koefisien regresi dapat diperoleh melalui program linier. Paket program linier pada umumnya hanya menyediakan solusi non-negatif, sedangkan koefisien regresi membutuhkan solusi bilangan nyata, sehingga koefisien regresi dinyatakan sebagai $b=b^+-b^-$, dalam hal ini b^+ dan b^- adalah peubah non-negatif. Sebagai ilustrasi, misalnya pasangan (x,y) yang akan digunakan untuk regresi dengan model dugaan y=a+bx+e adalah {(2,2), (4,3), (6,5), (8,7), (10,11)}. Struktur matriks tujuan dan kendala disajikan pada Tabel 1.

Pada pendugaan ukuran pemusatan sudah ditunjukkan secara grafis bahwa penduga LAD tidak khas, salah satunya adalah median (Setyono *et al*., 2015). Regresi median merupakan salah satu regresi kuantil, sehingga

Tabel 1. Matriks kendala pada program linier untuk regresi LAD

Tujuan	a^+	a^-	b^+	b^-	e_1^+	e_1^-	e_2^+	e_2^-	e_3^+	e_3^-	e_4^+	e_4^-	e_5^+	e_5^-	Tanda	Y
	0	0	0	0	1	1	1	1	1	1	1	1	1	1		
Kendala	1	-1	2	-2	1	-1	0	0	0	0	0	0	0	0	=	2
	1	-1	4	-4	0	0	1	-1	0	0	0	0	0	0	=	3
	1	-1	6	-6	0	0	0	0	1	-1	0	0	0	0	=	5
	1	-1	8	-8	0	0	0	0	0	0	1	-1	0	0	=	7
	1	-1	10	-10	0	0	0	0	0	0	0	0	1	-1	=	11
Hasil	0	1	1	0	1	0	0	0	0	0	0	0	2	0		

Koefisien regresi yang dihasilkan adalah $a=a^+-a^-=0-1=-1$ *dan* $b=b^+-b^-=1-0=1$, sedangkan sisaannya dihitung dengan cara serupa sehingga $e_1=1-0=1$, $e_2=0-0=0$, $e_3=0-0=0$, $e_4=0-0=0$, *dan* $e_5=2-0=2$.

```
> x=matrix(cbind(2,4,6,8,10))
> y=matrix(cbind(2,3,5,7,11))
> library(quantreg)
> hasil=rq(y~x,0.5)
> b=matrix(hasil$coef)
> w=matrix(cbind(1,x),5,2)
> e=y-w%*%b
> b
      [,1]
[1,]    -1
[2,]     1
> e
      [,1]
[1,]     1
[2,]     0
[3,]     0
[4,]     0
[5,]     2
```

Gambar 1. Keluaran program R untuk regresi LAD

regresi LAD dapat dibuat menggunakan regresi kuantil-0,5. Regresi kuantil sudah tersedia dalam paket program R, sehingga untuk data di atas dapat dikerjakan dengan koding seperti Gambar 1.

Tampak koefisien regresi yang diperoleh melalui program linier sama dengan yang diperoleh melalui paket program regresi kuantil. Meskipun paket program regresi kuantil dapat digunakan untuk mengerjakan regresi LAD standar, namun tidak dapat digunakan untuk menjalankan regresi LAD yang disertai dengan pengendalian koefisien regresi. Oleh sebab itu penggunaan program linier menjadi pilihan utama.

Tampak koefisien regresi yang diperoleh melalui program linier sama dengan yang diperoleh melalui paket program regresi kuantil. Meskipun paket program regresi kuantil dapat digunakan untuk mengerjakan regresi LAD

standar, namun tidak dapat digunakan untuk menjalankan regresi LAD yang disertai dengan pengendalian koefisien regresi. Oleh sebab itu penggunaan program linier menjadi pilihan utama.

Pengendalian Koefisien Regresi pada Data Cabai Merah

Pada 30 tanaman contoh dilakukan pengamatan terhadap 12 peubah, yaitu tinggi tanaman (cm), tinggi dikotom (cm), jumlah daun (helai), diameter batang (mm), lebar tajuk (cm), jumlah cabang, luas daun (cm^2), jumlah buah, rata-rata panjang buah (cm), rata-rata diameter buah (cm), rata-rata bobot buah (gram), dan bobot buah per tanaman (g). Pada analisis regresi ini, bobot buah total per tanaman berperan sebagai peubah tak bebas, sedangkan peubah yang lain sebagai peubah bebas. Untuk memodelkan bobot buah per tanaman sebagai fungsi dari peubah-peubah yang lain, pertama-tama dilakukan korelasi antara peubah tersebut dengan bobot buah total per tanaman. Peubah yang memiliki korelasi nyata dengan bobot buah per tanaman dijadikan sebagai kovariat pada regresi berganda untuk memodelkan bobot buah per tanaman.

Terdapat enam peubah yang berkorelasi nyata dengan bobot buah per tanaman (Y), yaitu diameter batang (X_1), lebar tajuk(X_2), jumlah buah (X_3), rata-rata panjang buah (X_4), rata-rata diameter buah (X_5), dan rata-rata bobot buah (X_6). Keenam peubah tersebut memiliki korelasi positif dengan bobot buah per tanaman, sehingga regresi linier sederhana dapat dibuat dengan model:

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

menggunakan LS sehingga nilai b_1 positif (Setyono *et al*., 2014).

Selanjutnya keenam peubah tersebut dilibatkan dalam regresi berganda untuk memodelkan bobot buah per tanaman dengan model:

$$Y_i=\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\beta_3 X_{3i}+\beta_4 X_{4i}+\beta_5 X_{5i}+\beta_6 X_{6i}+\varepsilon_i$$

menggunakan regresi LS, LAD, dan MLAD. Hasilnya disajikan pada Tabel 2. Baik regresi LS dan LAD maupun MLAD tidak dapat mempertahankan tanda koefisien regresi seperti regresi linier sederhana. Sesuai dengan namanya, metode LS unggul dalam jumlah kuadrat sisaan (paling kecil), metode LAD unggul dalam jumlah sisaan mutlak, dan metode MLAD unggul dalam maksimum sisaan mutlak.

Prosedur pendugaan koefisien regresi yang diajukan agar tanda koefisien regresi dapat dipertahankan seperti regresi linier sederhana adalah sebagai berikut:

1. Dihitung korelasi antara setiap peubah bebas dengan peubah respon.
2. Jika korelasi bertanda positif dan nyata maka ditambahkan kendala koefisien regresi untuk peubah bebas tersebut agar nilainya lebih atau sama dengan nol. Jika korelasi negatif dan nyata maka ditambahkan kendala koefisien regresi untuk peubah tersebut agar nilainya kurang atau sama dengan nol. Bilamana korelasi tidak nyata maka peubah bebas tersebut tidak dilibatkan dalam regresi (koefisien regresi sama dengan nol).
3. Dilakukan regresi dengan metode LAD dengan tambahan kendala hasil langkah 2.
4. Dilakukan *bootstrap* sisaan sebanyak 1.000 kali berdasarkan sisaan hasil langkah 3, untuk mendapatkan nilai rata-rata dan simpangan baku dari koefisien regresi yang dihasilkan.

Melalui prosedur tersebut, semua koefisien regresi

Tabel 2. Koefisien regresi LS, LAD, dan MLAD untuk data cabai merah

Koefisien regresi	LS	LAD	MLAD
Intersep	-221,39	-317,84	-83,07
Diameter batang	-6,93	-1,06	-10,80
Lebartajuk	1,69	0,74	1,92
Jumlahbuah	3,11	3,53	3,02
Rata-rata panjang buah	3,80	8,50	9,21
Rata-rata diameter buah	1,51	-0,11	-39,64
Rata-rata bobot buah	57,47	60,01	82,97
Maksimum sisaan mutlak	72,47	96,85	50,52
Jumlah sisaan mutlak	668,40	602,46	842,84
Jumlah kuadrat sisaan	23.753,43	38.767,88	34.672,01

setiap peubah bebas sudah tidak negatif seperti disajikan pada Tabel 3.

Hasil tersebut menunjukkan modifikasi kendala pada regresi LAD dapat mempertahankan tanda dan kisaran koefisien regresi seperti yang diinginkan. Hal serupa juga dapat dilakukan menggunakan regresi MLAD (Setyono *et al*., 2016). Namun, dengan pengendalian ini, sebaran statistik (penduga parameter) yang dihasilkan mungkin berubah sehingga perlu dikaji lebih lanjut.

Sebaran Koefisien Regresi Hasil *Bootstrap*

Setelah semua asumsi analisis regresi dipenuhi, koefisien regresi **b** dengan metode kuadrat terkecil (LS) menyebar normal dengan $E(\mathbf{b})=\boldsymbol{\beta}$ dan $Var(\mathbf{b})=(X'X)^{-1}\sigma^2$. Pada metode LAD, nilai **b** tidak dapat dinyatakan secara eksplisit, sehingga $Var(\mathbf{b})$ juga tidak dapat dinyatakan secara eksplisit. Dengan demikian, karakternya tidak dapat dipelajari secara analitik. Padahal sebaran koefisien regresi merupakan hal penting untuk pembuatan selang kepercayaan dan pengujian hipotesis. Jika dihadapkan pada satu set data maka salah satu penduga $Var(\mathbf{b})$ dapat diperoleh melalui *bootstrap*. Melalui *bootstrap* dapat diperoleh dugaan sebaran $\boldsymbol{b}$ jika hanya menggunakan satu set data.

Sebaran **b** hasil *bootstrap* metode LAD pada data cabai yang dicontohkan juga mendekati normal kalau koefisien regresi dibiarkan bebas. Namun, jika koefisien regresi dikendalikan agar non-negatif maka sebaran koefisien regresi bergantung pada tanda koefisien regresi pada saat tidak dikendalikan. Untuk koefisien regresi yang bernilai negatif jika tidak dikendalikan, sebarannya menjadi tidak normal atau seperti terpancung. Sebagai ilustrasi dapat dibandingkan sebaran koefisien regresi hasil *bootstrap* jika dibiarkan bebas dengan sebaran koefisien regresi dikendalikan agar non-negatif.

Gambar 2 menyajikan sebaran b_0 hasil bootstrap jika b_1, b_2, ..., b_6 dibiarkan bebas (a) dan sebaran b_0 hasil bootstrap jika b_1, b_2, ..., b_6 dikendalikan non-negatif (b). Pada kedua bootstrap tersebut, nilai b_0 tidak dikendalikan. Hasil *bootstrap* menunjukkan sebaran b_0 jika b_1, b_2, ..., b_6 tidak dikendalikan terlihat lebih setangkup dibandingkan dengan sebaran b_0 jika b_1, b_2, ..., b_6 dikendalikan untuk non-negatif.

Koefisien regresi LAD untuk peubah diameter batang (b_1) cabai adalah negatif. Setelah dilakukan *bootstrap* sisaan sebaran koefisien regresi b_1 relatif setangkup (Gambar 3a). Namun setelah koefisien b_1 dikendalikan agar non-negatif, sebaran koefisien regresi terlihat tidak memiliki ekor kiri (Gambar 3b).

Koefisien regresi LAD untuk peubah lebar tajuk (b_2) cabai adalah positif tetapi kecil (0,74). Sebaran b_2 hasil *bootstrap* setelah dibiarkan bebas relatif setangkup. Jika b_2 dikendalikan agar non-negatif, maka sebaran b_2 hasil *bootstrap* terpancung di sebelah kiri (Gambar 4). Hal ini terjadi karena nilai b2 hanya 0,74 sehingga setelah dilakukan *bootstrap* dimungkinkan diperoleh $b_2<0$, sehingga jika dikendalikan agar b_2 non-negatif sebarannya seperti terpancung.

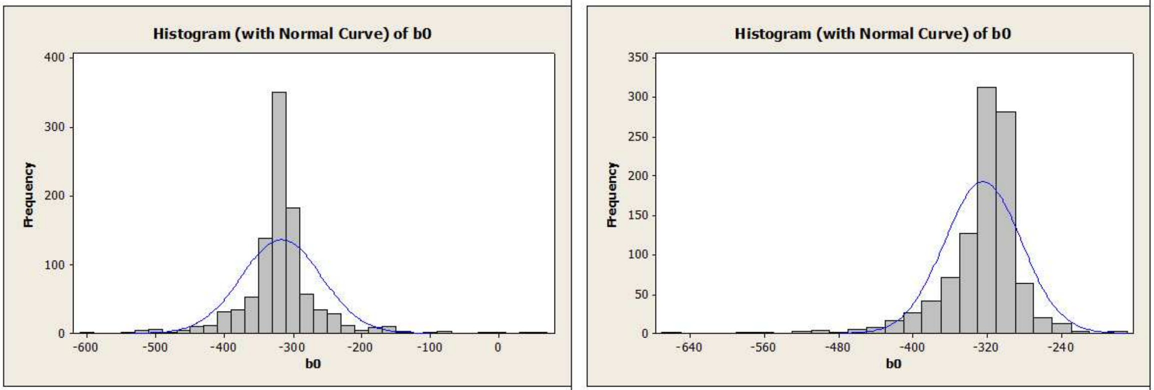
Koefisien regresi LAD untuk peubah jumlah buah (b_3) cabai bertanda positif (Tabel 2) dengan galat baku relatif kecil (Tabel 3). Nilai b_3 menggunakan LAD adalah 3,53 sehingga hasil *bootstrap* 1.000 kali hampir semua b_3 bernilai positif. Oleh sebab itu, dikendalikan atau tidak, sebaran b_3 hampir sama (Gambar 5).

Koefisien regresi LAD untuk rata-rata panjang buah (X_4) cabai relatif besar, yaitu 8,50. Sebaran b_4 yang dikendalikan non-negatif maupun yang dibiarkan

Tabel 3 Koefisien regresi LAD dengan kendala untuk data cabai merah.

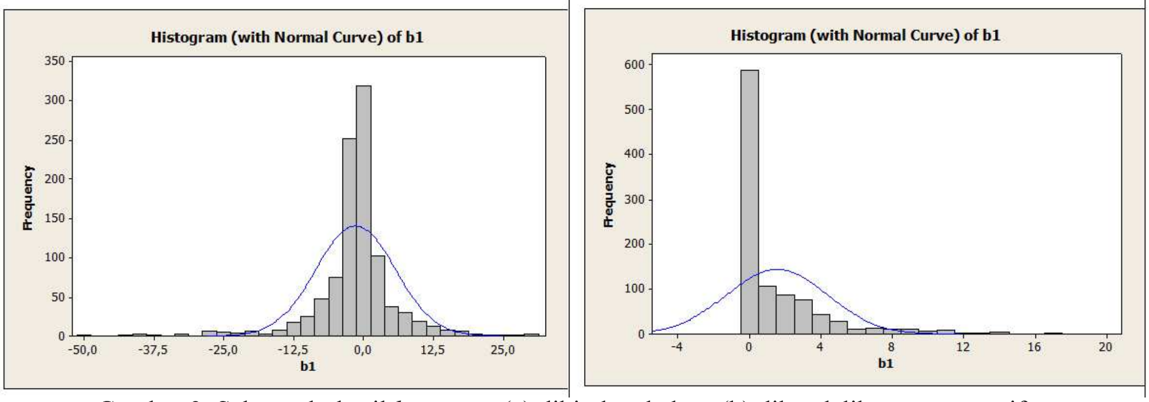
Peubah	Koefisien	Galat baku
Intersep	-327,49	41.72
Diameter batang	1,47	2.75
Lebar tajuk	0,64	0.61
Jumlah buah	3,48	0.12
Rata-rata panjang buah	7,63	3.52
Rata-rata diameter buah	3,62	7.13
Rata-rata bobot buah	56,51	8.50
Maksimum sisaan mutlak	92,25	
Jumlah sisaan mutlak	617,56	
Jumlah kuadrat sisaan	35.522,26	

(a)	(b)



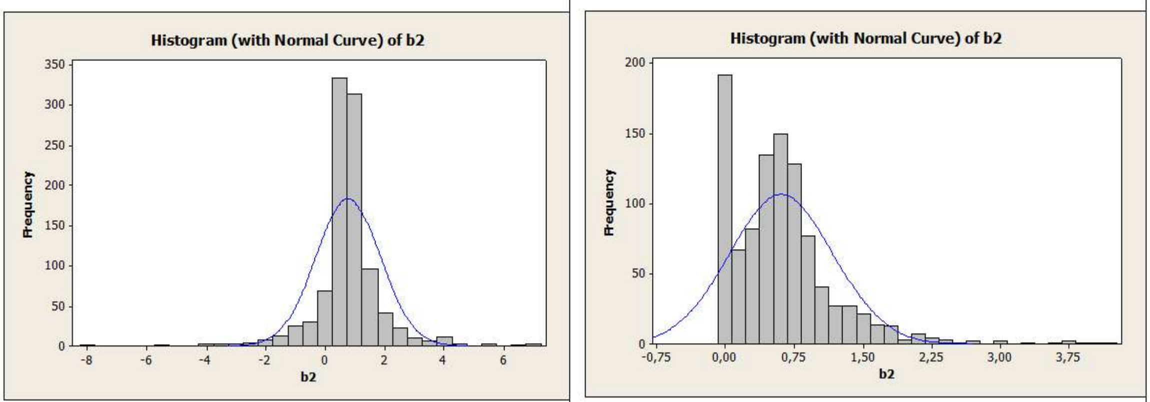
Gambar 2. Sebaran b_0 hasil *bootstrap* (a) $b_1, b_2, ..., b_6$ dibiarkan bebas, (b) $b_1, b_2, ..., b_6$ dikendalikan non-negatif

(a)	(b)

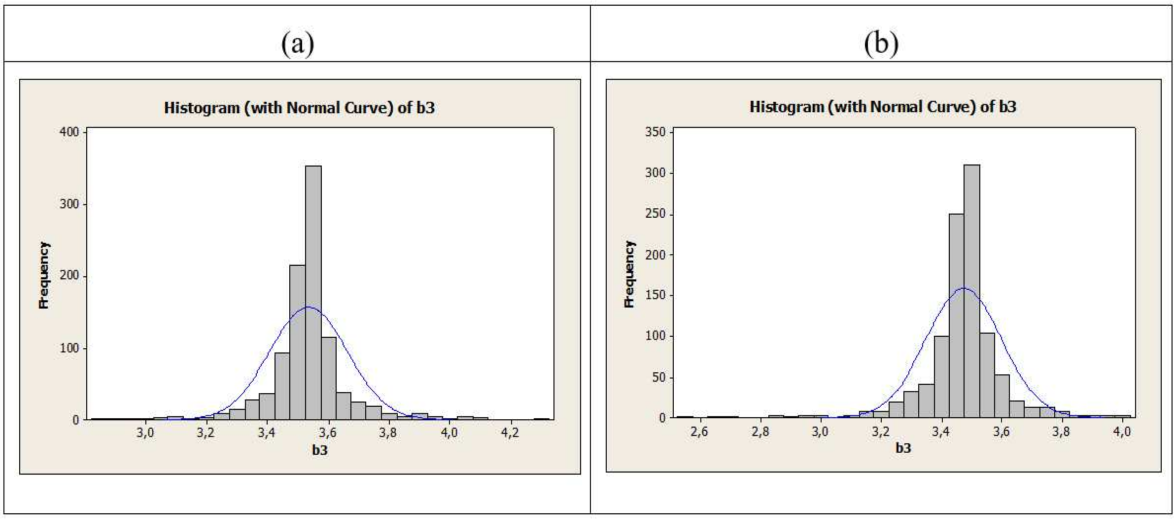


Gambar 3. Sebaran b_1 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif

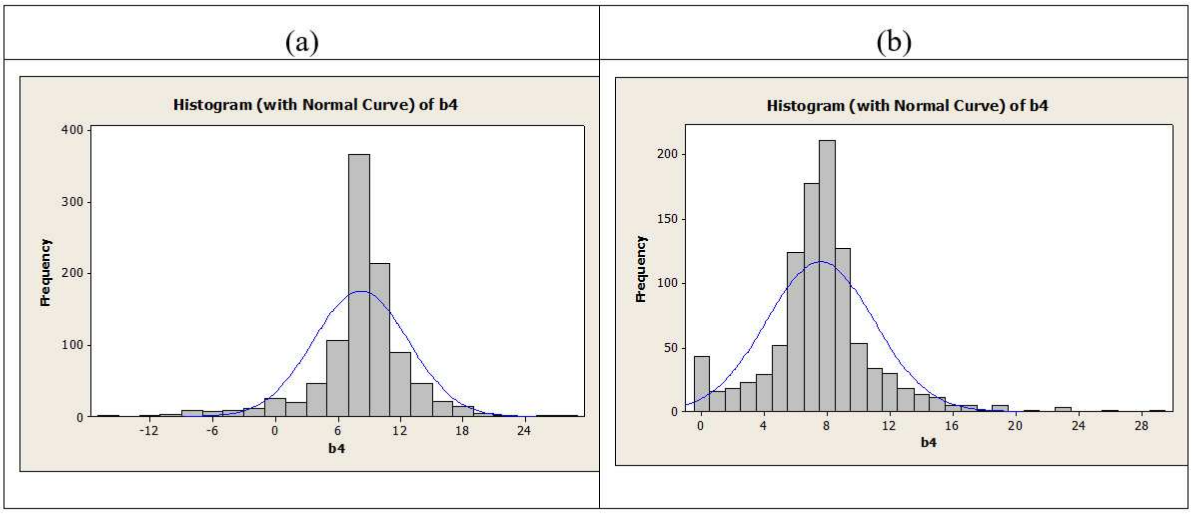
(a)	(b)



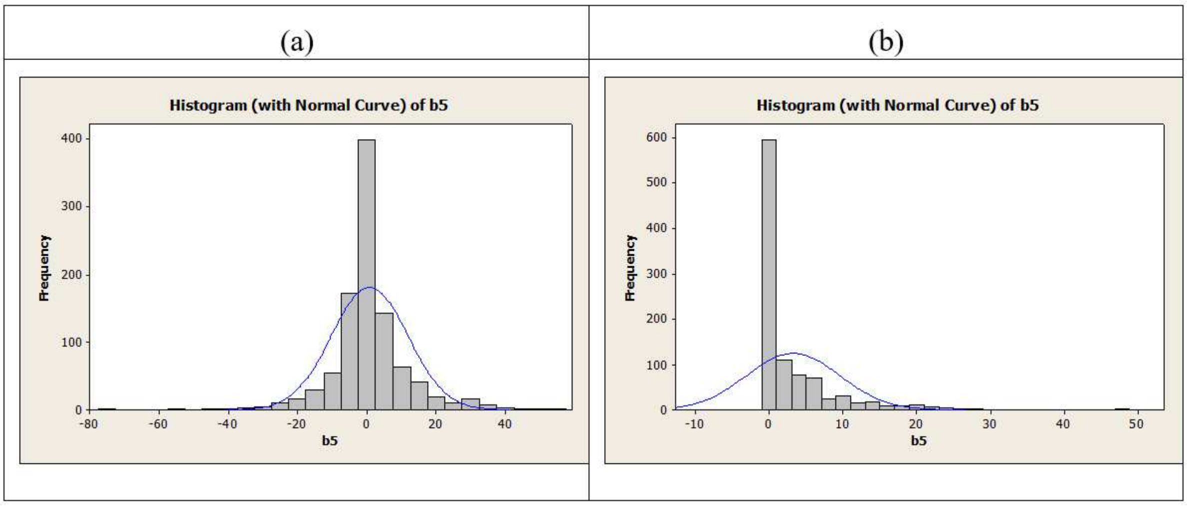
Gambar 4.Sebaran b_2 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif



Gambar 5. Sebaran b_3 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif



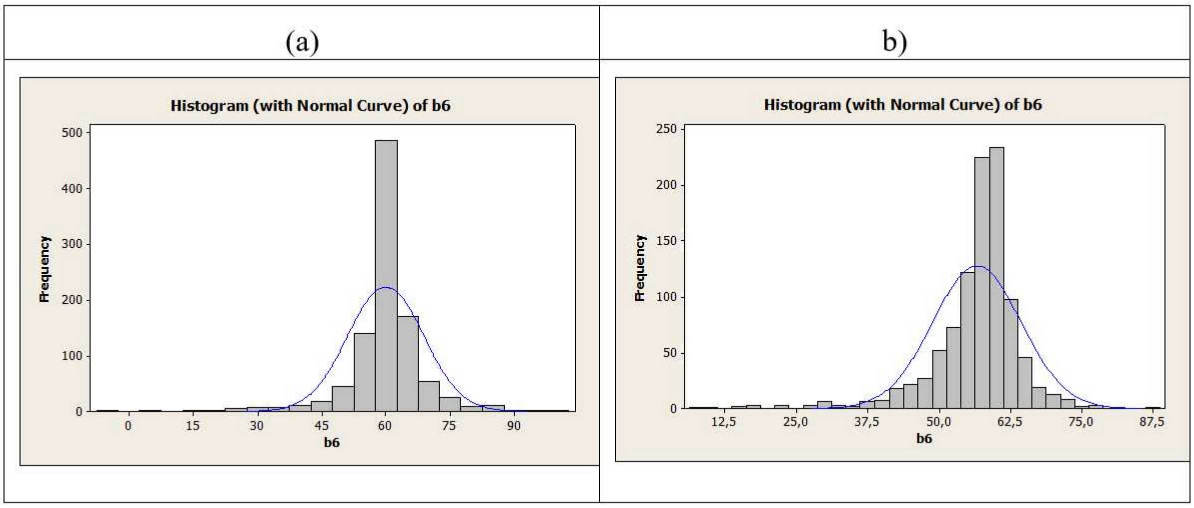
Gambar 6.Sebaran b_4 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif



Gambar 7.Sebaran b_5 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif

bebas seharusnya hampir sama. Namun karena galat baku b_4 cukup besar (Tabel 3), bisa saja hasil *bootstrap* menghasilkan b_4 yang negatif. Jika b_4 dikendalikan non-negatif, sebaran b_4 menjadi terpancung di sebelah kiri (Gambar 6).

Koefisien regresi LAD untuk rata-rata diameter buah (X_5) kalau dibiarkan bebas adalah negatif (Tabel 2). Sebaran b_5 hasil *bootstrap* jika dibiarkan bebas relatif



Gambar 8. Sebaran b_6 hasil *bootstrap* (a) dibiarkan bebas, (b) dikendalikan non-negatif

setangkup (Gambar 7a). Jika diberi kendala agar b_5 non-negatif, sebaran b_5 hasil bootsrap tampak terpancung di sebelah kiri (Gambar 7b).

Koefisien regresi LAD untuk rata-rata bobot buah (X_6) adalah positif dan cukup besar. Oleh sebab itu, tanpa dikendalikan untuk non-negatif pun sebaran b_6 sudah non-negatif. Dengan demikian, baik dibiarkan bebas maupun diberi kendala agar non- negatif, sebaran b_6 hasil *bootstrap* sama-sama setangkup dan tidak terpancung (Gambar 8).

Secara umum data cabai di atas tampak bahwa penambahan kendala pada program linier berhasil mengendalikan koefisien regresi untuk berada pada rentang tertentu yang wajar. Hasil *bootstrap* sisaan sebanyak 1.000 kali menunjukkan koefisien regresi yang kalau dibiarkan bernilai positif, jika dikendalikan agar non-negatif maka sebarannya relatif tidak berubah. Sebaliknya, koefisien regresi yang kalau dibiarkan bernilai negatif, maka jika dikendalikan agar non-negatif akan berdampak terhadap sebarannya yang terpancung di sebelah kiri. Sebaran statistik menjadi dasar dalam pengujian hipotesis. Oleh sebab itu, perubahan sebaran statistik akan mempengaruhi prosedur pengujian hipotesis yang digunakan.

KESIMPULAN

Penduga LAD untuk koefisien regresi dapat diperoleh melalui program linier, sehingga penambahan kendala atau modifikasi kendala dapat dilakukan. Dalam pengkajian ini, penambahan kendala pada program linier berhasil mengendalikan koefisien regresi untuk berada pada rentang tertentu yang wajar.

Memasukkan kendala batas bawah pada model regresi berakibat sebaran koefisien regresi tidak memiliki ekor kiri. Sebaliknya, memasukkan kendala batas atas pada model regresi berakibat sebaran koefisien regresi tidak memiliki ekor kanan. Jika ada informasi rentang nilai koefisien regresi, maka memasukkannya sebagai kendala pada model regresi akan memperkecil keragaman koefisien regresi yang dihasilkan.

Sebaran koefisien regresi merupakan hal penting dalam pembuatan selang kepercayaan dan pengujian hipotesis. Sebaran koefisien regresi LAD beserta galat bakunya dapat dipelajari melalui studi simulasi. Jika hanya menggunakan satu set data, maka galat baku regresi LAD dapat diperoleh melalui *bootstrap*.

UCAPAN TERIMA KASIH

Penelitian ini dibiayai oleh Kementerian Riset, Teknologi, dan Pendidikan Tinggi melalui Proyek Hibah Penelitian Produk Terapan.

DAFTAR PUSTAKA

Ash, R.B. 2000. Probability and Measure Theory. Second Edition. Harcourt Academic Press, California. 516 pages.

Cottle, R., E. Johnson,and R. Wets. 2007. George B Danzig (1914-2005). Notices of the AMS 54(3): 344-362.

Draper, N. and H. Smith. 1992. Analisis Regresi Terapan Edisi Kedua. Gramedia Pustaka Utama. Jakarta. 671 halaman.

Givens, G.H. and J. A. Hoeting. 2005.Computational Statistics. John Wiley and Sons, New York. 418 pages.

Hansen, B.E. 2013. Econometrics. University of Wisconsin, Madison. 346 pages.

Hao, L. And D. Q. Naiman. 2007. Quantile Regression. Sage Publications, Inc. California. 126 pages.

Koenker, R. and G. Bassett. 1978. Regression quantiles. Econometrica 46(1): 33-50.

Koenker, R. and K. F. Hallock. 2001. Quantile regression. Journal of Economic Perspectives 15(4): 143–156.

McCarl, B.A. and T.H. Spreen, 2003. Applied Mathematical Programming Using Algebraic Systems. Texas A&M University. 567 pages.

Murniati, N.S., Setyono, dan S. A. Adimihardja. 2012. Analisis korelasi dan sidik lintas peubah pertumbuhan terhadap produksi cabai merah (*Capsicum annuum* L.). JurnalPertanian3(2): 97-107.

Rizzo, M.L. 2008. Statistical Computing with R. Chapman and Hall, London.Jersey. 399 pages.

Setyono., I.M. Sumertajaya, A. Kurnia, dan A.A. Mattjik. 2014. Pengendalian koefisien regresi pada rentang yang bermakna melalui metode yang meminimumkan maksimum sisaan mutlak. Jurnal Pertanian 5(1): 52–57.

Setyono., I.M. Sumertajaya, A. Kurnia, dan A.A. Mattjik. 2015. Keragaan Galat pada Berbagai Metode Optimasi Sisaan. Jurnal Informatika Pertanian 24(2):19–204.

Setyono., I. M. Sumertajaya, A. Kurnia, and A. A. Mattjik. 2016. The addition of constraints on the MLAD regression coefficients to improve the meaningfulness. Global Journal of Pure and Applied Mathematics 12(6): 4827-4839.

Venables, W.N. and B. D. Ripley. 2002. Modern Applied Statistics with S. Springer, New York. 495 pages.

Venables, W.N., D.M. Smith, and the R Core Team. 2016. An Introduction to R. R Core Team. 99 pages.

Winston, W.L and Goldberg JB. 2004. Operations Research Applications and Algorithms. Belmont:Brooks/Cole—Thomson Learning. 1418 pages.