Statistika, Vol. 10 No. 1, 13 – 23
Mei 2010

Pendugaan Selang Kepercayaan Persentil Bootstrap Nonparametrik untuk Parameter Regresi

MARZUKI, HIZIR SOFYAN, ASEP RUSYANA

Jurusan Matematika FMIPA Universitas Syiah Kuala
Jl. Syech Abdul Rauf No. 3 Darussalam, Banda Aceh.
E-mail: marz_ukie@yahoo.com

ABSTRAK

Persentil bootstrap merupakan salah satu metode pendugaan selang kepercayaan dengan menetapkan batas bawah dan atas selang berdasarkan persentase dari replikasi bootstrap. Penelitian ini bertujuan untuk menduga selang kepercayaan persentil bootstrap untuk parameter model regresi linier satu dan dua peubah bebas dengan melakukan beberapa variasi jumlah sampel bootstrap dan jumlah pengulangan pendugaan parameter. Data yang disimulasikan adalah data riil agar dapat dipastikan ada hubungan fungsionalnya antara peubah-peubah bebas dan peubah takbebas. Simulasi dilakukan untuk 9 kasus, yaitu masing-masing untuk kombinasi n = 50, 100, dan 200 serta B = 1000, 5000, dan 10000. Hasil penelitian menunjukkan bahwa banyaknya perulangan dalam pendugaan parameter regresi tidak mempengaruhi selang kepercayaan bootstrap nonparametrik. Namun jika jumlah sampel bootstrap yang diambil semakin besar maka selang yang dihasilkan makin pendek.

Kata Kunci: dugaan; parameter; selang kepercayaan; bootstrap nonparametrik.

1. PENDAHULUAN

Model regresi diperoleh dengan pendugaan parameter regresi. Pendugaan parameter ini didasarkan pada variansi masing-masing parameter. Hal ini dilakukan atas dasar asumsi-asumsi tertentu dalam model regresi. Jika dugaan ini berupa titik maka galat dugaannya lebih tinggi daripada dugaan nilai rata-ratanya.

Hasil dugaan yang memberikan galat yang tinggi, hasilnya tidak akan baik. Solusi yang dapat ditempuh adalah dengan cara mendugakan selang kepercayaan yang lebih panjang. Menduga dengan satu titik berarti menduga dengan hanya satu angka. Untuk kasus ini hampir dapat dipastikan dugaannya keliru karena hampir mustahil dapat menduga angka parameter dengan tepat. Kemudian dikenal pendugaan selang yang berupa selang nilai. Bukan menduga bahwa market share produk tertentu sebesar 30%, tetapi menyatakan dalam bentuk 28% - 32%. Kalau ternyata angka pastinya adalah 31%, maka dugaan titik yang 30% itu salah, sedangkan kalau memakai selang menjadi benar. Menduga dengan memakai selang ini memiliki kemungkinan benar lebih besar daripada menduga hanya menggunakan satu titik.

Pendugaan selang kepercayaan untuk nilai respon tunggal, di antaranya dapat dilakukan dengan melakukan simulasi bootstrap. Persentil bootstrap merupakan salah satu metode pendugaan selang kepercayaan dengan menetapkan batas bawah dan atas selang berdasarkan persentase dari replikasi bootstrap yang dilakukan. Model regresi yang digunakan dalam penelitian ini adalah model regresi linier sederhana dan model regresi linier dengan dua peubah bebas.

Penelitian ini bertujuan untuk menduga selang kepercayaan persentil bootstrap untuk parameter model regresi linier dan membandingkan selang kepercayaan persentil bootstrap untuk beberapa variasi jumlah sampel bootstrap dan jumlah perulangan pendugaan parameter. Adapun manfaat dari kajian tentang dugaan selang kepercayaan persentil bootstrap ini dapat dipakai dalam kasus model regresi yang data sampelnya tidak banyak.

2. SIMULASI BOOTSTRAP NONPARAMETRIK

Simulasi bootstrap sering digunakan dalam analisis statistik seperti menghitung galat dari suatu statistik, analisis regresi dan korelasi, analisis komponen utama, analisis deret waktu, analisis masalah dua sampel, pendugaan selang kepercayaan, dan masalah pengujian hipotesis (Efron dan Tibshirani, 1993).

Misalkan $x_1, x_2, \cdots, x_n$ adalah data pengamatan, biasanya dinyatakan oleh suatu vektor $x = (x_1, x_2, \cdots, x_n)$, kemudian dihitung suatu statistik dari pengamatan tersebut misalkan *s*(***x***). Suatu sampel bootstrap $x^* = (x_1^*, x_2^*, \cdots, x_n^*)$ diperoleh melalui sampling acak dengan pengembalian sebanyak *n* kali dari data pengamatan $x_1, x_2, \cdots, x_n$. Algoritma bootstrap diawali dengan membangkitkan *B* buah sampel bootstrap yang saling bebas, yaitu $x^{*1}, x^{*2}, \cdots, x^{*B}$, masing-masing berukuran *n*. Banyaknya sampel bootstrap *B*, berkisar dari 50 sampai dengan 200 untuk menaksir standar error dari suatu statistik. Berkaitan dengan sampel bootstrap adalah replikasi bootstrap dari suatu statistik, katakanlah $s(x^{*b})$ untuk *b* = 1, 2, …, *B*, nilai statistik *s* yang dievaluasi untuk x^{*b}, jika $s(x)$ adalah rata-rata sampel, maka $s(x^{*b})$ adalah rata-rata sampel bootstrap.

3. SELANG KEPERCAYAAN PERSENTIL BOOTSTRAP

Misalkan akan dibentuk selang kepercayaan $(1-\alpha)$ untuk parameter θ. Efron dan Tibshirani (1993) memberikan langkah-langkah untuk membentuk selang kepercayaan tersebut dengan menggunakan metode persentil bootstrap:

1) Nilai dugaan parameter θ, katakanlah $\hat{\theta}$, dihitung berdasarkan data sampel asli $x = (x_1, x_2, \cdots, x_n)$
2) Nilai dugaan parameter θ, katakanlah $\hat{\theta}$, dihitung berdasarkan data sampel asli $x = (x_1, x_2, \cdots, x_n)$
3) Sampel acak diambil dengan pengembalian dari data sampel asli untuk mendapatkan sampel bootstrap $x^* = (x_1^*, x_2^*, \cdots, x_n^*)$
4) Statistik yang sama dihitung seperti langkah 1 dengan menggunakan data sampel bootstrap untuk mendapatkan $\hat{\theta}^*$
5) Langkah 2 dan 3 diulang sebanyak *B* kali $(B \geq 1000)$, sehingga diperoleh replikasi bootstrap $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \cdots, \hat{\theta}^{*B}$
6) Replikasi bootstrap $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \cdots, \hat{\theta}^{*B}$ diurut dari nilai terkecil ke nilai terbesar.
7) Batas bawah selang adalah replikasi bootstrap urutan ke-$(B.\alpha/2)$ sedangkan batas atasnya adalah replikasi bootstrap urutan ke-$(B.(1-\alpha/2))$ dari replikasi bootstrap.

4. SELANG KEPERCAYAAN PARAMETER MODEL REGRESI LINIER

Model regresi linier sederhana adalah

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{1}$$

dengan dugaan modelnya

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 \tag{2}$$

Pendugaan parameter β diperoleh dengan rumus

$$\hat{\beta}_1 = \frac{s_{XY}}{s_{XX}} \; ; \; \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1 \tag{3}$$

dengan

$s_{YX} = \sum (X_{1i} - \overline{X}_1) Y_i$ dan $s_{XX} = \sum (X_{1i} - \overline{X}_1)^2$

Selang kepercayaan $(1-\alpha)100\%$ untuk masing-masing parameter β_1 dan β_0, dalam Myers (1990) adalah

$$\beta_1 \pm t_{\alpha/2,v} s \sqrt{1/s_{XX}} \tag{4}$$

$$\beta_0 \pm t_{\alpha/2,v} s \sqrt{\frac{1}{n} + \frac{\overline{X}}{s_{XX}}} \tag{5}$$

dengan $t_{\alpha/2}$ adalah titik $\alpha/2$ persen pada distribusi-t, v adalah derajat bebas dari pendugaan σ, dan s adalah standar deviasi sampel.

Sedangkan model regresi linier berganda dengan dua peubah adalah

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \tag{6}$$

dengan dugaan modelnya

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 \tag{7}$$

Pendugaan parameter β diperoleh dengan rumus

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = [\boldsymbol{X}'\boldsymbol{X}]^{-1}[\boldsymbol{X}'\boldsymbol{Y}] \; ; \; \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}_1 - \hat{\beta}_2 \overline{X}_2 \tag{8}$$

dengan

$$[\boldsymbol{X}'\boldsymbol{X}] = \begin{bmatrix} \Sigma(X_1 - \overline{X}_1)^2 & \Sigma(X_1 - \overline{X}_1)(X_2 - \overline{X}_2) \\ \Sigma(X_1 - \overline{X}_1)(X_2 - \overline{X}_2) & \Sigma(X_2 - \overline{X}_2)^2 \end{bmatrix} \tag{9}$$

$$[\boldsymbol{X}'\boldsymbol{Y}] = \begin{bmatrix} \Sigma(X_1 - \overline{X}_1)(Y - \overline{Y}) \\ \Sigma(X_2 - \overline{X}_2)(Y - \overline{Y}) \end{bmatrix} \tag{10}$$

Selang kepercayaan $(1-\alpha)100\%$ untuk masing-masing parameter β_i, dalam Ryan (1997) adalah

$$\beta_i \pm t_{\alpha/2,v} s \sqrt{c_{ii}} \tag{11}$$

dengan $i = 0, 1, 2$, $t_{\alpha/2}$ adalah titik $\alpha/2$ persen pada distribusi-t, v adalah derajat bebas dari pendugaan σ, s adalah standar deviasi sampel, dan c_{ii} merupakan unsur diagonal matriks $[\boldsymbol{X}'\boldsymbol{X}]^{-1}$.

5. DATA DAN METODE ANALISA

Data yang digunakan dalam penelitian tentang selang kepercayaan bootstrap untuk parameter regresi ini adalah dua kumpulan data riil, masing-masing untuk model regresi linier sederhana dan berganda. Data riil digunakan supaya data yang disimulasikan ini dapat dipastikan ada hubungan fungsionalnya antara peubah bebas atau peubah-peubah bebas dan peubah takbebas (model regresi linier sederhana dan berganda).

Data ongkos pemeliharaan dan umur dari 17 traktor diambil dari buku *Analisis Regresi Terapan Edisi Kedua* (Draper dan Smith, 1992). Data sampel ini disajikan dalam Tabel 1.

Sedangkan untuk model regresi berganda diambil data dari buku *A Second Course in Statistics: Regression Analysis Sixth Edition* (Mendenhall dan Sincich, 1996). Data ini adalah 15 sampel yang diambil acak dari hasil wawancara seribuan pedagang kaki lima di Kota Puebla, Meksiko untuk melihat faktor-faktor yang mempengaruhi pendapatan pedagang kaki lima. Peubah yang diperhatikan dalam penelitian ini adalah Umur (X_1), Jam kerja per hari (X_2) dan Penghasilan per tahun (Y). Data sampel ini disajikan dalam Tabel 2.

Tabel 1. Data Ongkos Pemeliharaan dan Umur Traktor

Nomor Sampel	Ongkos dalam 6 bulan ($)	Umur (tahun)
	Y	X_1
1	619	4,5
2	1049	4,5
3	1033	4,5
4	495	4,0
5	723	4,0
6	681	4,0
7	890	5,0
8	1522	5,0
9	987	5,5
10	1194	5,0
11	163	0,5
12	182	0,5
13	764	6,0
14	1373	6,0
15	978	1,0
16	466	1,0
17	549	1,0

Tabel 2. Data Sampel Pedagang Kaki Lima di Kota Puebla Tahun 1996

Nomor Sampel	Penghasilan per tahun (x$100)	Umur (tahun)	Jam kerja per hari (jam)
	Y	X_1	X_2
1	28,41	29	12
2	18,76	21	8
3	29,34	62	10
4	15,52	18	10
5	30,65	40	11
6	36,70	50	11
7	20,05	65	5
8	32,15	44	8
9	19,30	17	8
10	20,10	70	6
11	31,11	20	9
12	28,82	29	9
13	16,83	15	5
14	18,17	14	7
15	40,66	33	12

Kumpulan data riil pertama (Tabel 1) diduga model regresi linier sederhana dan selang kepercayaan masing-masing parameter regresinya. Selang kepercayaan persentil bootstrap untuk parameter model regresi linier untuk kumpulan data pertama dilakukan pendugaan berdasarkan langkah-langkah berikut:

1) Sampel bootstrap diambil sebanyak *n* buah berdasarkan data riil, lalu dicari nilai parameter-parameternya.
2) Pengambilan sampel ini diulang sebanyak *B* kali, sehingga diperoleh *B* nilai parameter untuk masing-masing parameter.
3) Nilai-nilai parameter ini diurutkan dari nilai yang terkecil sampai nilai yang terbesar. Hal ini dilakukan masing-masing terhadap dua parameter.
4) Selang kepercayaan untuk kedua parameter, diperoleh dari nilai-nilai parameter yang telah diperoleh di atas, dengan menggunakan teori selang kepercayaan bootstrap.
5) Langkah ketiga sampai keenam disimulasikan sebanyak 9 kali masing-masing untuk kombinasi *n* = 50, 100, dan 200 serta *B* = 1000, 5000, dan 10000. Kombinasi ini selengkapnya disajikan dalam Tabel 3.

Tabel 3. *Setting*-an Kasus Simulasi

Kasus	1	2	3	4	5	6	7	8	9
n	50	50	50	100	100	100	200	200	200
B	1000	5000	10000	1000	5000	10000	1000	5000	10000

6) Selang kepercayaan bootstrap yang telah diperoleh ini dibandingkan dengan selang kepercayaan asli.

Data pada Tabel 2 diduga model regresi linier berganda dan selang kepercayaan masing-masing parameter regresinya. Pendugaan selang kepercayaan persentil bootstrap untuk parameter-parameter model regresi ini juga dilakukan berdasarkan enam langkah di atas, namun parameter yang diperhatikan dalam kasus ini adalah sebanyak 3 buah.

6. MODEL DAN SELANG KEPERCAYAAN DATA RIIL

Data riil dijadikan dasar data analisis untuk mencapai tujuan penelitian. Data riil pada Tabel 1, diduga model regresi sederhana sehingga memberikan nilai $\hat{\beta}_0 = b_0 = 323{,}6223$, dan $\hat{\beta}_1 = b_1 = 131{,}7165$. Model regresi linier berganda yang diperoleh,

$$\hat{Y} = 323{,}6223 + 131{,}7165 X_1$$

Sedangkan selang kepercayaan 95% untuk kedua paramater atau koefisien regresi di atas dicantumkan dalam Tabel 4.

Tabel 4. Selang Kepercayaan Parameter Regresi untuk Data Riil

Parameter	Batas Bawah	Batas Atas
β_0	10,4294	636,8152
β_1	55,8217	207,6113

7. SAMPEL DAN SELANG KEPERCAYAAN BOOTSTRAP NONPARAMETRIK

Sampel bootstrap diambil secara nonparametrik dari 17 sampel data riil. Kasus pertama, ambilan sebanyak $n = 50$ sampel. Dari sampel ini diduga nilai parameter β_0 dan β_1. Misalkan nilai dugaan untuk pengambilan pertama ini masing-masing adalah $\hat{\beta}_0^{\,1}$ dan $\hat{\beta}_1^{\,1}$. Pengambilan masing-masing 50 sampel ini diulang sebanyak $B = 1000$ kali sehingga diperoleh 1000 nilai dugaan parameter regresi untuk masing-masing parameter, yaitu $\hat{\beta}_0^{\,1}, \hat{\beta}_0^{\,2}, \cdots, \hat{\beta}_0^{\,1000}$ dan $\hat{\beta}_1^{\,1}, \hat{\beta}_1^{\,2}, \cdots, \hat{\beta}_1^{\,1000}$.

Selang kepercayaan bootstrap untuk parameter β_0 diperoleh dengan mengurutkan nilai dugaan $\hat{\beta}_0^{\,1}, \hat{\beta}_0^{\,2}, \cdots, \hat{\beta}_0^{\,1000}$ dari nilai yang paling kecil sampai dengan nilai yang paling besar. Batas bawah dari selang ini adalah nilai dugaan yang ke- $B(\alpha/2)$. Jika tingkat kepercayaan yang diambil $\alpha = 0{,}05$ maka batas bawah dari selang bootstrap adalah nilai dugaan yang ke-25. Sedangkan batas atasnya adalah nilai dugaan yang ke- $B(1-\alpha/2)$, yaitu nilai dugaan yang ke-975. Dugaan ke-25 diperoleh 170,4176 dan dugaan ke-975 diperoleh 499,1868. Dengan demikian selang kepercayaan bootstrap parameter β_0 untuk kasus $n = 50$ dan $B = 1000$ adalah [170,4176; 499,1868].

Selanjutnya, dengan cara yang sama, pengurutan nilai dugaan $\hat{\beta}_1^{\,1}, \hat{\beta}_1^{\,2}, \cdots, \hat{\beta}_1^{\,1000}$ menghasilkan selang kepercayaan bootstrap paramater β_1 yaitu [91,3019; 167,5159]. Kasus kedua, untuk pengambilan sebanyak $n = 50$ sampel dan $B = 5000$. Selang kepercayaan bootstrap untuk parameter β_0 diperoleh dengan mengurutkan nilai dugaan $\hat{\beta}_0^{\,1}, \hat{\beta}_0^{\,2}, \cdots, \hat{\beta}_0^{\,5000}$ dari nilai yang paling kecil sampai dengan nilai yang paling besar. Batas bawah dari selang ini adalah nilai dugaan yang ke- $B(\alpha/2)$. Jika tingkat kepercayaan yang diambil $\alpha = 0{,}05$ maka batas bawah dari selang bootstrap adalah nilai dugaan yang ke-125. Sedangkan batas atasnya adalah nilai dugaan yang ke- $B(1-\alpha/2)$, yaitu nilai dugaan yang ke-4875. Selang kepercayaan bootstrap yang dihasilkan untuk β_0 dan β_1 masing-masing adalah [170,8309; 493,6922] dan [89,9474; 170,2340].

Sedangkan kasus ketiga, juga dengan pengambilan sampel bootstrap sebanyak $n = 50$ namun untuk $B = 10000$. Batas bawah selang kepercayaan bootstrap untuk kasus ini adalah dugaan parameter secara bootstrap yang ke-250 dan batas atasnya adalah nilai dugaan yang ke-9750. Untuk kasus ketiga ini, selang kepercayaan bootstrap masing-masing parameter, berturut-turut adalah [171,0484; 497,1565] dan [90,1025; 170,1622].

Tabel 5. Selang Kepercayaan Bootstrap untuk Parameter Regresi Sederhana

n	B	Parameter	
		β_0	β_1
50	1000	[170,4176; 499,1868]	[91,3019; 167,5159]
	5000	[170,8309; 493,6922]	[89,9474; 170,2340]
	10000	[171,0484; 497,1565]	[90,1025; 170,1622]
100	1000	[220,2671; 438,2440]	[103,3896; 159,2207]
	5000	[213,5346; 445,8541]	[102,0398; 159,0101]
	10000	[214,2004; 439,8448]	[103,7537; 158,8598]
200	1000	[246,0963; 407,8866]	[110,1981; 150,8733]
	5000	[246,1496; 406,0958]	[112,2884; 151,2543]
	10000	[244,8499; 405,6663]	[111,7471; 151,3524]

Kasus keempat, kelima, dan keenam adalah dengan mencari masing-masing $B = 1000$, $B = 5000$, dan $B = 10000$ dugaan parameter untuk masing-masing parameter regresi, sedangkan pengambilan sampel bootstrapnya adalah $n = 100$. Batas bawah selang kepercayaan bootstrap untuk kasus keempat, kelima, dan keenam masing-masing adalah nilai dugaan yang ke-25, 125, dan 250. Sedangkan batas atas selang bootstrap untuk ketiga kasus ini masing-masing adalah nilai dugaan yang ke-975, 4875, dan 9750.

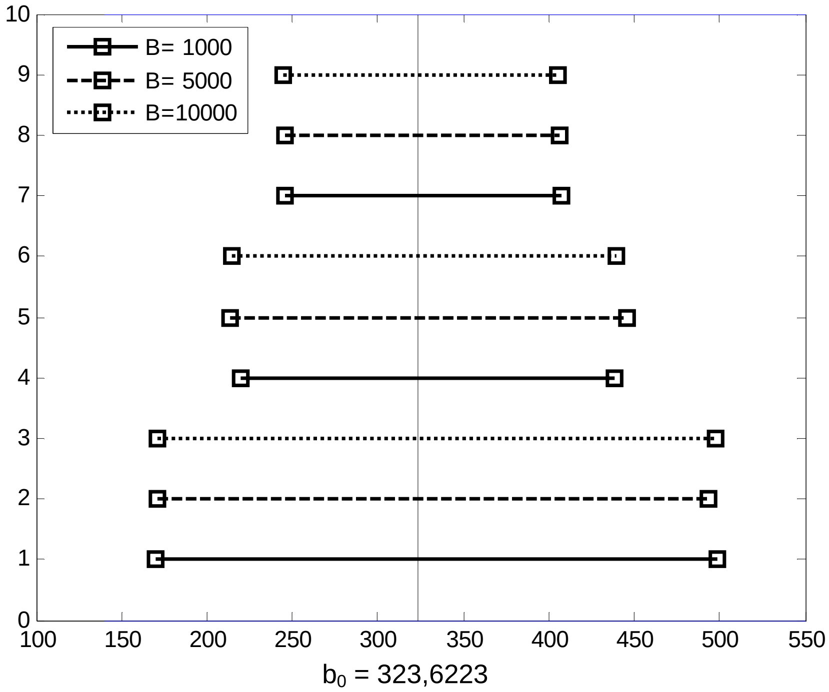
Tiga kasus terakhir, yaitu kasus ketujuh sampai sembilan adalah dengan mengambil 200 sampel bootstrap. Setelah semua nilai parameter ini diurutkan, maka cara pengambilan batas bawah dan batas atas selang kepercayaan bootstrap identik dengan cara untuk kasus pertama sampai ketiga atau kasus keempat sampai keenam.

Dugaan selang kepercayaan bootstrap dengan tingkat kepercayaan $\alpha = 0{,}05$ untuk kesembilan kasus tersebut di atas disajikan dalam Tabel 5.

8. PERBANDINGAN SELANG KEPERCAYAAN BOOTSTRAP NONPARAMETRIK

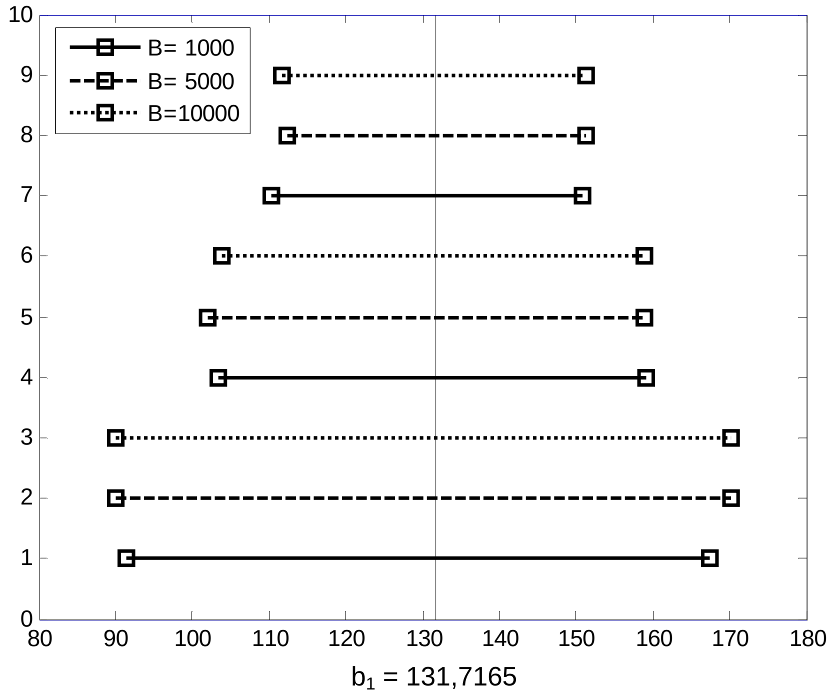
Selang kepercayaan bootstrap yang dihasilkan untuk pendugaan parameter β_0 cenderung sama untuk setiap jumlah perulangan B yang ditetapkan. Untuk jumlah sampel bootstrap yang sama, selang kepercayaannya cenderung sama untuk $B = 1000$, $B = 5000$, maupun $B = 10000$. Kecenderungan ini dapat dilihat pada Gambar 1. Tiga selang pertama dari bawah merupakan selang kepercayaan parameter β_0 untuk kasus $n = 50$.

Kecenderungan ini juga berlaku untuk kasus jumlah sampel bootstrap yang lain, yaitu $n = 100$ dan $n = 200$. Selang kepercayaan untuk kasus ini ditunjukkan pada Gambar 1, yaitu tiga selang yang di tengah (untuk $n = 100$) dan tiga selang teratas (untuk $n = 200$). Untuk kedua kasus ini pun, menunjukkan bahwa jumlah perulangan perhitungan parameter tidak begitu mempengaruhi terhadap selang yang dihasilkan. Jumlah perulangan perhitungan $B = 1000$, $B = 5000$, maupun $B = 10000$ cenderung menghasilkan dugaan selang yang relatif sama.



Gambar 1. Perbandingan Selang Kepercayaan Bootstrap untuk Parameter β_0

Namun, perbedaan selang kepercayaan yang dihasilkan terlihat pada jumlah sampel bootstrap. Semakin banyak sampel bootstrap yang diambil maka selang kepercayaan yang dihasilkan cenderung semakin sempit. Hal ini berlaku untuk setiap jumlah perulangan B yang ditetapkan. Semakin besar sampel bootstrap yang diambil (makin ke atas pada gambar), maka semakin sempit selang yang dihasilkan. Gambar ini menunjukkan bahwa selang kepercayaan yang dihasilkan semakin sempit, jika sampel bootstrap yang diambil semakin banyak. Analogi dengan penjelasan tentang pendugaan selang kepercayaan untuk parameter β_0 di atas, selang kepercayaan untuk paramater yang lain, yaitu β_1, juga menghasilkan hal yang sama. Kenyataan ini ditunjukkan oleh Gambar 2.



Gambar 2. Perbandingan Selang Kepercayaan Bootstrap untuk Parameter β_1

Garis vertikal di tengah gambar-gambar di atas menunjukkan nilai dugaan untuk masing-masing parameter regresi. Nilai dugaan parameter ini semuanya berada dalam selang kepercayaan bootstrap yang diperoleh. Jika selang kepercayaan dibandingkan antara data riil dan data sampel bootstrap, maka selang kepercayaan bootstrap jauh lebih pendek daripada selang kepercayaan tanpa bootstrap. Selang kepercayaan bootstrap 95% merupakan bagian dari selang untuk data sebelum di-bootstrap (Tabel 4).

9. SELANG KEPERCAYAAN UNTUK PARAMETER REGRESI BERGANDA

Data riil yang kedua memberikan nilai dugaan parameter regresi berganda $\hat{\beta}_0 = b_0 = -0{,}2035$, $\hat{\beta}_1 = b_1 = 0{,}1335$, dan $\hat{\beta}_2 = b_2 = 2{,}4371$. Sehingga model regresi linier berganda yang diperoleh, $\hat{Y} = -0{,}2035 + 0{,}1335X_1 + 2{,}4371X_2$, dan selang kepercayaan 95% untuk ketiga paramater di atas dicantumkan dalam Tabel 6.

Tabel 6. Selang Kepercayaan Parameter Regresi untuk Data Riil

Parameter	Batas Bawah	Batas Atas
β_0	-14,426	14,019
β_1	-0,034	0,301
β_2	1,053	3,821

Sampel bootstrap diambil secara nonparametrik dari 15 sampel data riil. Setingan kasus untuk pendugaan selang kepercayaan parameter regresi berganda, mengambil setingan untuk regresi sederhana. Kasus pertama, ambilan sebanyak $n = 50$ sampel. Dari sampel ini diduga nilai parameter β_0, β_1, dan β_2. Misalkan nilai dugaan untuk pengambilan pertama ini masing-masing adalah $\hat{\beta}_0^{\,1}$, $\hat{\beta}_1^{\,1}$, dan $\hat{\beta}_2^{\,1}$. Pengambilan masing-masing 50 sampel ini diulang sebanyak $B = 1000$ kali sehingga diperoleh 1000 nilai dugaan parameter regresi untuk masing-masing parameter, yaitu $\hat{\beta}_0^{\,1}, \hat{\beta}_0^{\,2}, \cdots, \hat{\beta}_0^{\,1000}$, $\hat{\beta}_1^{\,1}, \hat{\beta}_1^{\,2}, \cdots, \hat{\beta}_1^{\,1000}$, dan $\hat{\beta}_2^{\,1}, \hat{\beta}_2^{\,2}, \cdots, \hat{\beta}_2^{\,1000}$.

Dugaan selang kepercayaan bootstrap dengan tingkat kepercayaan $\alpha = 0{,}05$ untuk kesembilan kasus tersebut di atas disajikan dalam Tabel 7.

Tabel 7. Selang Kepercayaan Bootstrap untuk Parameter Regresi

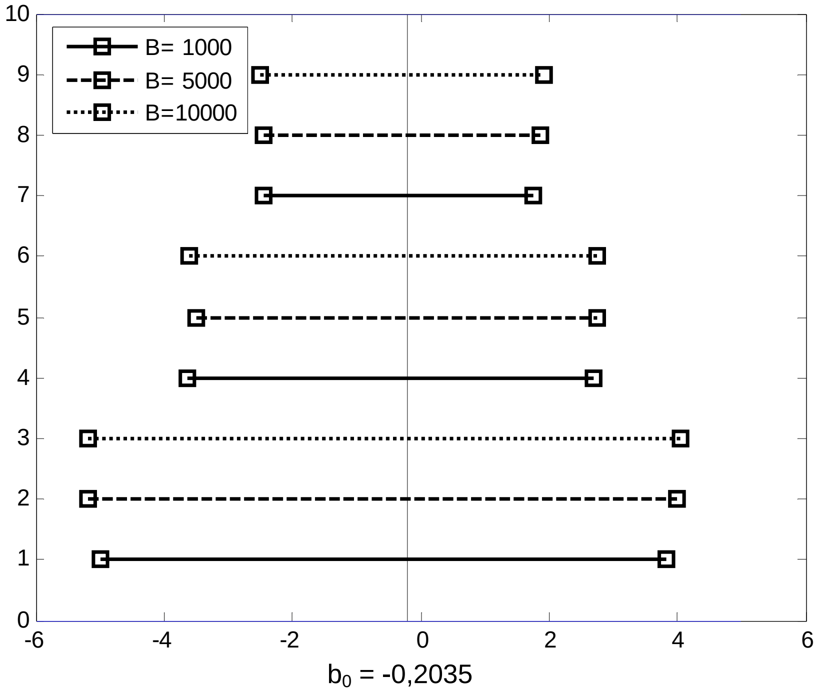
n	B	Parameter		
		β_0	β_1	β_2
50	1000	[-5,0016; 3,8403]	[0,0796; 0,2084]	[1,8503; 2,9638]
	5000	[-5,1919; 3,9990]	[0,0781; 0,2072]	[1,8842; 2,9550]
	10000	[-5,1837; 4,0552]	[0,0780; 0,2086]	[1,8822; 2,9675]
100	1000	[-3,6420; 2,6909]	[0,0927; 0,1818]	[2,0453; 2,7824]
	5000	[-3,4861; 2,7429]	[0,0937; 0,1822]	[2,0678; 2,7939]
	10000	[-3,6140; 2,7540]	[0,0940; 0,1822]	[2,0514; 2,8050]
200	1000	[-2,4482; 1,7445]	[0,1041; 0,1668]	[2,1967; 2,6718]
	5000	[-2,4584; 1,8724]	[0,1043; 0,1650]	[2,1879; 2,6922]
	10000	[-2,5032; 1,9109]	[0,1048; 0,1656]	[2,1794; 2,6894]

Selang kepercayaan bootstrap yang dihasilkan untuk pendugaan ketiga parameter regresi berganda cenderung sama untuk setiap jumlah perulangan B yang ditetapkan. Ini menunjukkan bahwa kecenderungan ini sama seperti pendugaan selang kepercayaan parameter model regresi sedehana. Untuk jumlah sampel bootstrap yang sama, selang kepercayaannya cenderung sama untuk $B = 1000$, $B = 5000$, maupun $B = 10000$. Kecenderungan ini bukan hanya untuk kasus $n = 50$, melainkan juga untuk kasus jumlah sampel bootstrap yang lain, yaitu $n = 100$ dan $n = 200$.

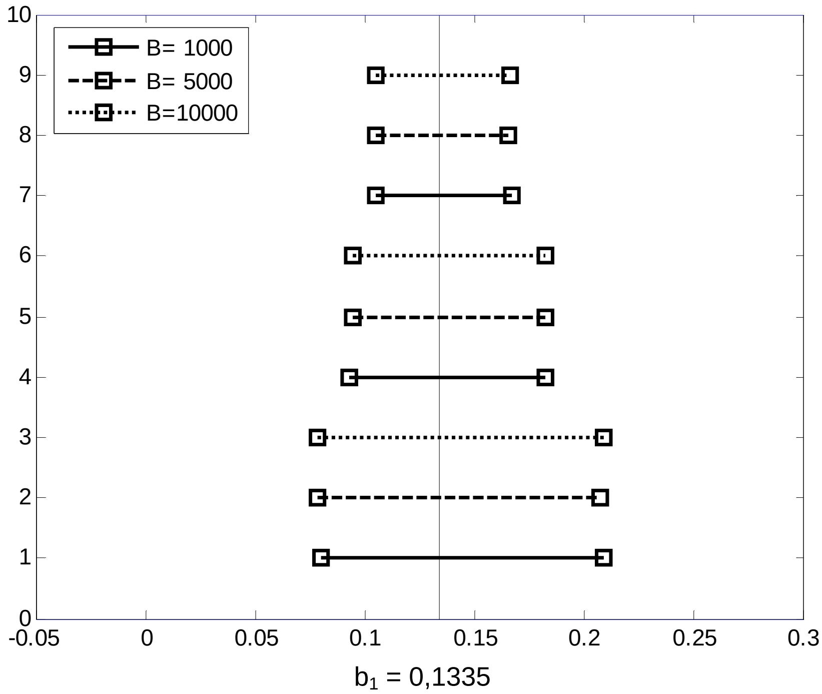
Jumlah perulangan perhitungan parameter tidak begitu mempengaruhi terhadap selang yang dihasilkan. Jumlah perulangan perhitungan $B = 1000$, $B = 5000$, maupun $B = 10000$ cenderung menghasilkan dugaan selang yang relatif sama. Ketiga gambar di bawah ini memperlihatkan kecenderungan ini.

Sama juga halnya dengan pendugaan selang pada model regresi sederhana, perbedaan selang kepercayaan yang dihasilkan terlihat pada jumlah sampel bootstrap. Semakin banyak sampel bootstrap yang diambil maka selang kepercayaan yang dihasilkan cenderung semakin sempit yang berlaku untuk setiap jumlah perulangan B yang ditetapkan.

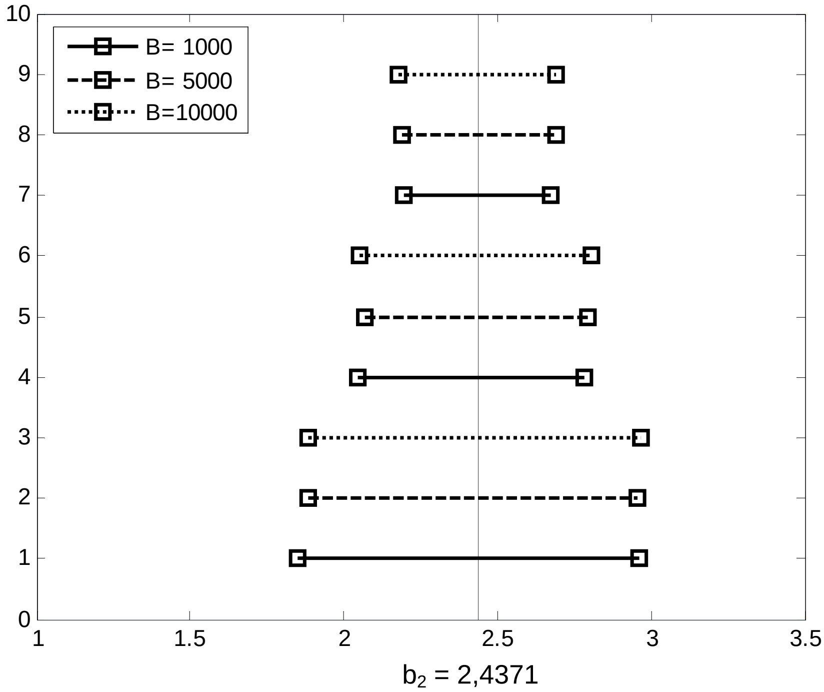
Jika selang kepercayaan dibandingkan antara data riil dan data sampel bootstrap, maka selang kepercayaan bootstrap jauh lebih pendek daripada selang kepercayaan tanpa bootstrap. Selang kepercayaan bootstrap 95% merupakan bagian dari selang untuk data sebelum di-bootstrap (Tabel 6).



Gambar 3. Perbandingan Selang Kepercayaan Bootstrap untuk Parameter β_0



Gambar 4. Perbandingan Selang Kepercayaan Bootstrap untuk Parameter β_1



Gambar 5. Perbandingan Selang Kepercayaan Bootstrap untuk Parameter β_2

10. KESIMPULAN

Kajian terhadap pendugaan selang kepercayaan dengan metode bootstrap nonparametrik dilakukan untuk sembilan kasus dengan berbagai variasi jumlah sampel bootstrap dan variasi jumlah perulangan pendugaan parameter. Pendugaan ini dilakukan masing-masing untuk parameter pada model regresi linier sederhana dan regresi linier berganda. Variasi tersebut masing-masing adalah untuk jumlah sampel bootstrap yang berbeda, yaitu $n = 50, 100, 200$ dan untuk jumlah perhitungan parameter yang berbeda, yaitu $B = 1000, 5000, 10000$. Ada beberapa hal yang dapat disimpulkan sehubungan dengan kajian tersebut, yaitu:

1) Selang kepercayaan bootstrap nonparametrik persentil, baik untuk parameter regresi linier sederhana maupun berganda, cenderung tidak dipengaruhi oleh banyaknya perulangan.
2) Selang kepercayaan bootstrap nonparametrik untuk setiap parameter regresi linier dan berganda, cenderung berbeda untuk setiap perubahan jumlah sampel bootstrap yang diambil. Selang yang dihasilkan makin pendek, jika jumlah sampel bootstrap yang diambil semakin besar.

DAFTAR PUSTAKA

[1] Draper, N. dan Smith, H. 1992. *Analisis Regresi Terapan Edisi Kedua*. P.T. Gramedia Pusaka Utama, Jakarta.
[2] Efron, B. dan Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York.
[3] Mendenhall, W. dan Sincich, T. 1996. *A Second Course in Statistics: Regression Analysis Sixth Edition*. Pearson Printice Hall, Florida.
[4] Myers. 1990. *Classical and Modern Regression with Applications Second Edition*, PWS-KENT Publishing Company, Boston.
[5] Ryan, T.P. 1997. *Modern Regression Methods*. John Wiley & Sons, New York.