

II. TINJAUAN PUSTAKA

2.1 Standar SEM

Di dalam perilaku, pendidikan, kesehatan, dan sains sosial, teori substantif biasanya melibatkan dua jenis variable dengan nama variabel teramati dan variabel laten. Variabel teramati adalah variabel yang dapat diamati secara langsung seperti nilai ujian, penghasilan, tekanan darah sistolik/diastolik dan berat badan seseorang. Variabel laten adalah variabel yang tidak dapat diamati secara langsung seperti kecerdasan, kepribadian, kondisi kesehatan, kemampuan dan tekanan darah.

Karakteristik dari sebuah variabel laten dapat diamati secara parsial oleh kombinasi linier dari beberapa variabel teramati. Contohnya, kemampuan kuantitatif dari siswa sekolah tercermin dari nilai-nilai matematika, fisika dan kimia siswa tersebut. Di dalam penelitian itu penting untuk membuktikan sebuah model yang patut untuk mengevaluasi sebuah hipotesis tentang dampak variabel laten dan variabel teramati terhadap variabel lainnya dan menulis galat-galat pengamatan ke dalam catatan. *Structural Equation Modeling* (SEM) merupakan sesuatu yang diakui sebagai metode statistik yang sangat penting untuk menyajikan tujuan di atas. SEM dapat diterapkan di dalam banyak kalangan seperti diterapkannya untuk penelitian pasar, investigasi kesehatan dan hal-hal yang memiliki hubungan sebab-akibat (Lee, 2007).

2.2 SEM (*Structural Equation Modeling*)

SEM adalah salah satu teknik statistik yang digunakan untuk melakukan pengujian terhadap suatu model sebab akibat dengan menggunakan kombinasi teori yang ada. Dalam penggunaannya, SEM memiliki asumsi-asumsi yang mendasarinya. SEM memiliki variabel-variabel dan model-model di dalamnya, variabel-variabel dan model-model dalam SEM beserta asumsi yang mendasarinya adalah sebagai berikut:

2.2.1 Variabel dalam SEM

a. Variabel Laten

Variabel kunci yang menjadi perhatian di dalam SEM adalah variabel laten, dimana variabel laten merupakan konsep abstrak, seperti perilaku orang, sikap, perasaan, dan motivasi. Variabel laten dapat diamati secara tidak langsung dan tidak sempurna melalui efeknya pada variabel teramati. SEM mempunyai 2 jenis variabel laten, yaitu eksogen dan endogen. SEM membedakan kedua jenis variabel ini berdasarkan keikutsertaan variabel sebagai variabel terkait pada persamaan-persamaan dalam model. Variabel laten eksogen sebagai variabel bebas pada persamaan yang ada dalam model. Sedangkan variabel endogen merupakan variabel terikat pada persamaan yang ada dalam model.

b. Variabel Teramati (*Indicator Variable*)

Variabel teramati adalah variabel yang dapat diamati atau dapat diukur secara empiris dan sering disebut indikator atau variabel *manifest*. Variabel teramati merupakan efek atau ukuran variabel laten. Pada metode survey dengan

menggunakan kuesioner, setiap pernyataan pada kuesioner mewakili sebuah variabel teramati.

2.2.2 Model dalam SEM

SEM memiliki model-model antara lain model struktural dan model pengukuran, berikut ini gambaran kedua model.

a. Model struktural

Model struktural menggambarkan hubungan-hubungan yang ada diantara variabel-variabel laten. Variabel-variabel laten dibagi menjadi dua kelas, yaitu variabel eksogen dan variabel endogen. Hubungan-hubungan ini umumnya linear meskipun perluasan SEM memungkinkan untuk mengikutsertakan hubungan yang non-linear. Model struktural dapat dibuat dalam notasi sederhana

$$Y = YB + Z$$

Dimana variabel Y merupakan matriks variabel laten, baik variabel eksogen maupun endogen. Nilai rentang error Z diasumsikan menjadi terpusat dengan $E[Z] = 0$.

b. Model Pengukuran

Model pengukuran memodelkan hubungan antara variabel laten dengan variabel indikator. Hubungan tersebut bersifat reflektif dari variabel laten terkait. Pada diagram PLS, 1 variabel indikator hanya dapat dihubungkan pada 1 variabel laten. Seluruh indikator yang terhubung dengan satu variabel laten disebut blok. Jadi, masing-masing variabel laten memiliki bloknya sendiri. Blok tersebut dapat berhubungan secara formatif dan reflektif.

2.2.3 Asumsi-Asumsi SEM

Asumsi-asumsi yang mendasari penggunaan *SEM* adalah sebagai berikut:

a. Distribusi Normal Multivariat.

Masing-masing indikator mempunyai nilai yang berdistribusi normal terhadap masing-masing indikator lainnya. Karena permulaan yang kecil normalitas multivariat dapat menuntun kearah perbedaan yang besar dalam pengujian chi-square, dengan demikian akan melemahkan kegunaannya. Secara umum, pelanggaran asumsi ini menaikkan chi-square sekalipun demikian didalam kondisi tertentu akan menurunkannya.

Selanjutnya penggunaan pengukuran ordinal atau nominal akan menyebabkan adanya pelanggaran normalitas multivariat. Perlu diperhatikan bahwa normalitas multivariat diperlukan untuk MLE, yang merupakan metode dominan dalam SEM yang akan digunakan untuk membuat estimasi koefisien-koefisien jalur struktur. Khususnya, MLE membutuhkan variabel-variabel endogen yang berdistribusi normal. Secara umum, sebagaimana ditunjukkan dalam suatu studi simulasi menunjukkan bahwa dalam kondisi data yang sangat tidak normal, pendugaan parameter SEM seperti misalnya estimasi jalur, masih dianggap akurat tetapi koefisien-koefisien signifikansi yang bersangkutan akan menjadi terlalu tinggi, sehingga nilai chi-square akan meningkat. Perlu diingat bahwa untuk uji keselarasan chi-square dalam model keseluruhan, nilai chi-square tidak harus signifikan jika ada keselarasan model yang baik. Semakin tinggi nilai chi-square, semakin besar perbedaan model yang diestimasi dan matriks kovarian sesungguhnya. Namun, keselarasan model akan semakin tidak baik.

Chi-square yang tinggi dapat mengarahkan peneliti berpikir bahwa model-model yang sudah dibuat memerlukan modifikasi. Kurangnya normalitas multivariat biasanya menaikkan statistik chi-square. Misalnya, statistik keselarasan chi-square secara keseluruhan untuk model yang bersangkutan akan bias ke arah kesalahan tipe 1, yaitu menolak suatu model yang seharusnya diterima. Pelanggaran terhadap normalitas multivariat juga cenderung menurunkan (*deflate*) kesalahan-kesalahan standar mulai dari menengah sampai ke tingkat tinggi. Kesalahan-kesalahan yang lebih kecil dari yang seharusnya terjadi mempunyai makna jalur-jalur regresi dan kovarian-kovarian faktor / kesalahan didapati akan menjadi signifikan secara statistik dibandingkan dengan seharusnya yang terjadi.

b. Linearitas

SEM mempunyai asumsi adanya hubungan linear antara variabel-variabel indikator dan variabel-variabel laten, serta antara variabel-variabel laten sendiri. Sekalipun demikian, sebagaimana halnya dengan regresi, peneliti dimungkinkan untuk menambah transformasi eksponensial, logaritma, atau non-linear lainnya dari suatu variabel asli ke dalam model yang dimaksud.

c. Pengukuran Tidak Langsung (*Indirect measurement*)

Secara tipikal, semua variabel dalam model merupakan variabel-variabel laten.

d. Indikator Jamak

Beberapa indikator harus digunakan untuk mengukur masing-masing variabel laten dalam model. Regresi dapat dikatakan sebagai kasus khusus dalam SEM

dimana hanya ada satu indikator di setiap variabel laten. Kesalahan pemodelan dalam SEM membutuhkan adanya lebih dari satu pengukuran untuk masing-masing variabel laten.

e. Rekursifitas

Suatu model disebut rekursif jika semua anak panah menuju satu arah, tidak ada factor pengulangan (*feedback looping*), dan faktor gangguan (*disturbance terms*) atau kesalahan residual untuk variabel-variabel endogenous yang tidak dikorelasikan. Dengan kata lain, model-model rekursif merupakan model-model dimana semua anak panah mempunyai satu arah tanpa putaran umpan balik dan peneliti dapat membuat asumsi kovarian–kovarian gangguan kesalahan semua 0. Dapat diartikan bahwa semua variabel yang tidak diukur yang merupakan determinan dari variabel-variabel endogenous tidak dikorelasikan satu dengan lainnya sehingga tidak membentuk *feedback loops*. Model–model dengan gangguan kesalahan yang berkorelasi dapat diperlakukan sebagai model rekursif hanya jika tidak ada pengaruh-pengaruh langsung diantara variabel-variabel endogen.

f. Ketepatan yang Tinggi

Apakah data berupa data interval atau ordinal, data-data tersebut harus mempunyai jumlah nilai yang besar. Jika variabel–variabel mempunyai jumlah nilai yang sangat kecil, maka masalah-masalah metodologi akan muncul pada saat peneliti membandingkan varian dan kovarian, yang merupakan masalah sentral dalam SEM.

g. Residual-Residual Acak dan Kecil

Rata-rata residual–residual atau kovarian hasil pengitungan yang diestimasi minus harus sebesar 0, sebagaimana dalam regresi. Suatu model yang sesuai akan hanya mempunyai residual – residual kecil. Residual–residual besar menunjukkan kesalahan spesifikasi model, sebagai contoh, beberapa jalur mungkin diperlukan untuk ditambahkan ke dalam model tersebut.

h. Gangguan Kesalahan yang Tidak Berkorelasi (*Uncorrelated Error Terms*)

Seperti di dalam regresi, maka gangguan kesalahan diasumsikan saja. Sekalipun demikian, jika memang ada dan dispesifikasi secara eksplisit dalam model oleh peneliti, maka kesalahan yang berkorelasi (*correlated error*) dapat diestimasi dan dibuat modelnya dalam SEM.

i. Kesalahan Residual yang Tidak Berkorelasi (*Uncorrelated Residual Error*)

Kovarian nilai–nilai variabel tergantung yang diprediksi dan residual–residual harus sebesar 0.

j. Multikolinearitas yang Lengkap

Multikolinearitas diasumsikan tidak ada, tetapi korelasi antara semua variabel bebas dapat dibuat model secara eksplisit dalam SEM. Multikolinearitas yang lengkap akan menghasilkan matriks kovarian tunggal, yang mana peneliti tidak dapat melakukan penghitungan tertentu, misalnya inversi matriks karena pembagian dengan 0 akan terjadi.

k. Ukuran Sampel

Ukuran sampel tidak boleh kecil, karena SEM bergantung pada pengujian-pengujian yang sensitif terhadap ukuran sampel dan magnitudo perbedaan-perbedaan matrices kovarian. Secara teori, untuk ukuran sampelnya berkisar antara 200 - 400 untuk model-model yang mempunyai indikator antara 10 - 15. Satu survei terhadap 72 penelitian yang menggunakan SEM didapatkan median ukuran sampel sebanyak 198. Sampel di bawah 100 akan kurang baik hasilnya jika menggunakan SEM (Sarwono, 2007).

2.3 Distribusi Normal Multipeubah

Pembahasan distribusi normal multipeubah diawali dengan mengemukakan konsep distribusi normal univariat dan chi kuadrat. Peubah tunggal Z didefinisikan mempunyai distribusi normal univariat baku jika dan hanya jika fungsi densitas peluang atau *probability density function* (pdf):

$$n(z; 0, 1) = \left(\frac{1}{\sqrt{2\pi}} \right) \exp \left[-\frac{z^2}{2} \right], \quad -\infty < z < \infty$$

dengan median = 0 & varians = 1.

Fungsi distribusi kumulatif dan fungsi pembangkit momen peubah acak Z yang mempunyai distribusi normal univariat baku masing-masing diberikan berikut

$$N(z; 0, 1) = \int_{-\infty}^z n(u; 0, 1) du, -\infty < z < \infty \text{ dan } M_z(t) = e^{\frac{t^2}{2}}$$

Bukti

$$M_z(t) = E_z[e^{tZ}]$$

$$\begin{aligned}
&= \int \frac{1}{\sqrt{2\pi}} e^{tz} e^{-\frac{z^2}{2}} dz \\
&= \int (2\pi)^{-\frac{1}{2}} e^{tz} e^{-\frac{z^2}{2}} dz \\
&= e^{\frac{t^2}{2}} \int (2\pi)^{-\frac{1}{2}} e^{-\frac{(z-t)^2}{2}} dz \\
&= e^{\frac{t^2}{2}} \blacksquare
\end{aligned}$$

Terbukti.

Misalkan peubah acak berikut ini Z_1, Z_2, \dots, Z_n independen dan identik berdistribusi normal dengan nilai tengah 0 dan varians 1. Distribusi marjinal Z_1 adalah

$$f_{Z_i}(z) = (2\pi)^{-\frac{1}{2}} e^{-\frac{z_i^2}{2}}, \quad -\infty < z_i < \infty$$

untuk $i = 1, 2, \dots, n$. Karena Z_i adalah peubah acak independen dan identik maka distribusi peluang gabungan vektor acak $n \times 1: Z = (Z_1, Z_2, \dots, Z_n)'$ adalah

$$F_Z(z) = (2\pi)^{-\frac{n}{2}} e^{-\frac{z' \Sigma z}{2}}, \quad -\infty < z_i < \infty$$

Bentuk terakhir ini dilambangkan dengan

$$N(Z; 0, I) = (2\pi)^{-\frac{n}{2}} e^{-\frac{z' \Sigma z}{2}}, Z \in E_n$$

dengan simbol $N(Z; 0, I)$ melambangkan distribusi normal multipeubah baku (Mustofa dan Warsono, 2009).

2.4 Metode Pendugaan PLS

Metode pendugaan yang digunakan pada penelitian ini adalah metode kuadrat terkecil parsial atau *Partial Least Square* (PLS).

PLS adalah teknik statistika multivariat yang melakukan perbandingan antara variabel dependen berganda dan variabel independen berganda. PLS adalah salah satu metode statistika SEM berbasis varian yang secara simultan dapat melakukan pengujian model pengukuran sekaligus pengujian model struktural. Model pengukuran digunakan untuk uji validitas dan reliabilitas, sedangkan model struktural digunakan untuk uji kausalitas (pengujian hipotesis dengan model prediksi). Perbedaan mendasar PLS yang merupakan SEM berbasis varian dengan LISREL atau AMOS yang berbasis kovarian adalah tujuan penggunaannya. SEM berbasis kovarian bertujuan untuk mengestimasi model untuk pengujian atau konfirmasi teori, sedangkan SEM berbasis varian bertujuan untuk memprediksi model untuk pengembangan teori.

Sebagai alat untuk model prediksi, untuk menghindari masalah *intedeminancy*, yaitu skor faktor yang berbeda dihitung dari model faktor tunggal yang dihasilkan. PLS mengasumsikan bahwa semua ukuran varian adalah varian yang dijelaskan sehingga pendekatan estimasi variabel laten dianggap sebagai kombinasi linear dari indikator. Dalam menggunakan metode PLS, ada beberapa langkah-langkah yang harus dilaksanakan seperti berikut ini

1. Merancang Model Struktural (*inner model*)

Pada SEM perancangan model adalah berbasis teori, akan tetapi pada PLS dapat berupa:

- a. Teori
 - b. Hasil penelitian empiris
 - c. Analogi, hubungan antar variabel pada bidang ilmu lain
 - d. Normatif, misal peraturan pemerintah, undang-undang, dan lain sebagainya
 - e. Rasional (PLS: bisa eksplorasi hubungan antar variabel)
2. Merancang Model Pengukuran (*outer model*)
 Pada SEM semua bersifat refleksif, model pengukuran tidak penting.
 Namun pada PLS perancangan *outer model* sangat penting yaitu reflektif atau formatif.
 3. Kontruksi diagram jalur.
 4. Konversi diagram jalur ke bentuk persamaan.
 5. Estimasi parameter.
 6. Estimasi Jalur yang menghubungkan antar variabel laten (koefisien jalur) dan antara variabel laten dengan indikatornya (*loading*).
 7. Evaluasi kecocokan model.
 8. *Outer Model* refleksif.
 Untuk model penelitian yang menggunakan *outer model* refleksif dievaluasi berdasarkan *convergent*, *discriminant validity*, *composite reliability*. Nilai *convergent* dilihat dari nilai *loading*, nilai tersebut dianggap cukup antara 0.5 sampai 0.6 untuk jumlah variabel laten antara 3 sampai 7. Nilai *discriminant validity* dilihat berdasarkan nilai AVE, nilai AVE tersebut > 0.5. Nilai *composite reliability* yang masih dapat diterima adalah ≥ 0.7

9. *Outer Model* formatif

Untuk model penelitian yang menggunakan outer model formatif dievaluasi berdasarkan pada *substantive content*-nya yaitu dengan melihat signifikansi dan *weight*.

10. Inner Model *GOF*

Diukur menggunakan *Q-square predictive relevance*.

Rumus Q-Square:

$$Q^2 = 1 - (1 - R_1^2)(1 - R_2^2) \dots (1 - R_p^2)$$

Dimana $R_1^2, R_2^2 \dots R_p^2$ adalah R square variabel endogen dalam model.

Interpretasi Q^2 sama dengan koefisien determinasi total dalam analisis jalur (mirip dengan R^2 pada regresi).

11. Uji Hipotesis

Hipotesis statistik untuk outer model:

$$H_0: \gamma_i = 0, \text{ vs } H_1: \gamma_i \neq 0$$

Hipotesis statistik untuk *inner model*: Variabel eksogen terhadap endogen:

$$H_0: \gamma_i = 0, \text{ vs } H_1: \gamma_i \neq 0$$

Hipotesis statistik untuk *inner model*: Variabel endogen terhadap endogen:

$$H_0: \beta_i = 0, \text{ vs } H_1: \beta_i \neq 0$$

12. Statistik uji

t-test; p-value $\leq 0,05$ (alpha 5%); signifikan

Outer model signifikan: indikator bersifat valid

Inner model signifikan: terdapat pengaruh signifikan

PLS tidak mengasumsikan data berdistribusi normal: menggunakan teknik resampling dengan metode *bootstrap*.

PLS sebagai model prediksi tidak mengasumsikan distribusi tertentu untuk mengestimasi parameter dan memprediksi hubungan kausalitas. Oleh karena itu, teknik parametrik untuk menguji signifikansi parameter tidak diperlukan dan model evaluasi untuk prediksi bersifat non-parametrik. Evaluasi model PLS dilakukan dengan mengevaluasi outer model dan inner model. Outer model merupakan model pengukuran untuk menilai validitas dan reliabilitas model. Melalui proses iterasi algoritma, parameter model pengukuran (validitas konvergen, validitas diskriminan, *composite reliability* dan *crombach's alpha*) diperoleh, termasuk nilai R² sebagai parameter ketepatan model prediksi. Inner model merupakan model struktural untuk memprediksi hubungan kausalitas antar variabel laten. Melalui proses *bootstrapping*, parameter uji T statistik diperoleh untuk memprediksi adanya hubungan kausalitas (Jogiyanto dan Abdillah, 2009).

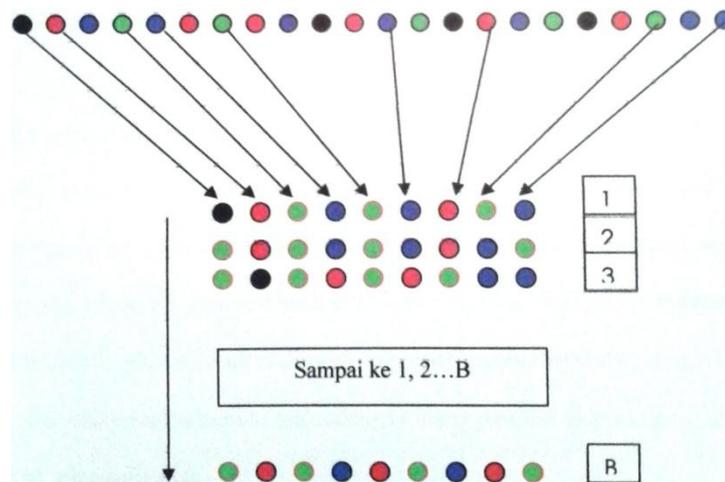
2.5 Uji Kecocokan

Setelah melakukan estimasi yang menghasilkan nilai parameter, perlu dilakukan pemeriksaan tingkat kecocokan. Pada tahap ini kita akan memeriksa tingkat kecocokan antara data dengan model, validitas dan reliabilitas model pengukuran, dan signifikansi koefisien-koefisien dari model struktural (Wijanto, 2008).

2.6 Metode *Bootstrap*

Bootstrap diperkenalkan oleh Bradley Efron pada tahun 1979. Istilah *bootstrap* berasal dari "*pull oneself up by one's bootstrap*", yang berarti berpijak di atas kaki sendiri, berusaha dengan sumber daya minimal. Dalam sudut pandang statistika

sumber daya minimal adalah data yang sedikit, data yang menyimpang dari asumsi tertentu, atau data yang tidak mempunyai asumsi apapun tentang distribusi populasinya. Teknik ini mampu menciptakan ukuran-ukuran dari ketakpastian dan bias, khususnya pada estimasi parameter dari variabel-variabel yang independen dan berdistribusi identik.



Gambar 2.1 Teorema B. Efron

Bootstrap adalah teknik resampling yang bertujuan untuk menaksir galat baku dan selang kepercayaan parameter populasi, seperti mean, median, proporsi, koefisien korelasi, dan regresi dengan tidak selalu memperhatikan asumsi distribusi.

Menurut Shao dan Tu (1995) dan jugamenurut Davison dan Hinkley (1997), distribusi sampling dianggap sebagai suatu model dengan sifat-sifat probabilitas yang diketahui. Seperti asumsi distribusi yang memerlukan formula analitis berdasarkan pada model untuk mengestimasi secara analitis parameter dalam distribusi samplingnya. Dalam prakteknya, distribusi sampling tidak selalu memenuhi distribusi normal dan kadang-kadang memerlukan penurunan

formulasi analitis yang sulit dilakukan sehingga dimungkinkan akurasi penduganya tidak valid.

Bootstrap memungkinkan seseorang untuk melakukan inferensi statistik tanpa membuat asumsi distribusi yang kuat dan tidak memerlukan formulasi analitis untuk distribusi sampling suatu penduga. Sebagai pengganti, *bootstrap* menggunakan distribusi empiris untuk mengestimasi distribusi sampling. Jadi jika penyelesaian analitik tidak mungkin dilakukan dimana anggapan (suatu distribusi, misalnya kenormalan data) tidak dipenuhi maka dengan menggunakan *bootstrap* masih dapat dilakukan suatu inferensi.

Dasar pendekatan *bootstrap* adalah dengan memperlakukan sampel sebagai populasi dan dengan menggunakan sampling Monte Carlo untuk membangkitkan dan mengkonstruksi penduga empiris dari distribusi sampling statistik. Distribusi sampling dapat dipandang sebagai harga-harga statistik yang dihitung dari sejumlah tak terhingga sampel acak berukuran n dari suatu populasi yang diberikan. Sampling Monte Carlo mengambil konsep ini untuk membangun distribusi sampling suatu penduga dengan mengambil sejumlah besar sampel berukuran n secara acak dari populasi dan menghitung statistik tersebut dari harga-harga distribusi sampling tersebut.

Metode *bootstrap* adalah melakukan *resampling* terhadap sampel awal x (berukuran n) secara satu per satu dengan pengembalian. Dengan prosedur ini didapat sampel baru.

$$x^* = x_1^*, x_2^*, \dots, x_n^*$$

Prosedur *resampling* tersebut diulang sebanyak B kali. Sehingga didapat sampel-sampel *bootstrap* sebanyak B berikut

$$\begin{aligned}x^{*1} &= (x_{11}^*, x_{21}^*, \dots, x_{n1}^*) \\x^{*2} &= (x_{12}^*, x_{22}^*, \dots, x_{n2}^*) \\&\vdots \\x^{*B} &= (x_{1B}^*, x_{2B}^*, \dots, x_{nB}^*)\end{aligned}$$

Selanjutnya dari tiap-tiap sampel *bootstrap* dihitung penduganya untuk mendapatkan $t(x)$, maka diperoleh penduga-penduga *bootstrapnya* (tx^*), $i = 1, 2, \dots, B$. Dengan distribusi *bootstrap* untuk sampel nilai tengah oleh *resampling* dari setiap sampel.

Tahapan/Prosedur *bootstrap*:

1. Sampel asal, dari eksperimen atau simulasi, berukuran n .
2. Resampel dengan pengembalian, didapatkan resampel ke- i ($i = 1, 2, \dots, \text{REP}$)
3. Perhitungan penaksir setiap hasil resampel, didapatkan: $\hat{\theta}_{b,1}, \hat{\theta}_{b,2}, \dots, \hat{\theta}_{b,\text{REP}}$; bentuk umumnya $\hat{\theta}_{b,i}$.
4. Perhitungan penaksir *bootstrap*:

$$\hat{\theta}_{bootstrap} = \hat{\theta}_b = \frac{1}{\text{REP}} \sum_{i=1}^{\text{REP}} \hat{\theta}_{b,i}$$

Keterangan: $\hat{\theta}_{b,i}$ atau $\hat{\theta}_{j,i}$ suatu penaksir parameter yang pada penerapannya dapat berupa: rata-rata, ragam, galat baku, korelasi, dan koefisien regresi.