CONTOH 16-2

Contoh: menghitung koefisien determinasi

Hitunglah koefisien determinasi antara permintaan minyak goreng dengan harga minyak goreng dan pendapatan pada Contoh 16-1, yang menghasilkan persamaan regresi $Y = 15{,}086 - 1{,}015X_1 - 0{,}41X_2$!

Jawab:

Rumus:

$$R^2 = \frac{n(a.\sum Y + b_1.\sum YX_1 + b_2.\sum YX_2) - (\sum Y)^2}{n.\sum Y^2 - (\sum Y)^2}$$

Dari soal diketahui bahwa:

$\Sigma Y = 68$	$\Sigma YX_1 = 409$	$\Sigma YX_2 = 239$	$\Sigma Y^2 = 516$
$n = 10$	$a = 15{,}086$	$b_1 = -1{,}015$	$b_2 = -0{,}41$

$$R^2 = \frac{10\{(15{,}086.68) + (-1{,}015)(409) + (-0{,}41)(239)\} - 68^2}{(10).(516) - 68^2} = \frac{503{,}23}{536} = 0{,}939$$

Nilai koefisien determinasi (R^2) = 0,939 dapat diartikan bahwa variabel bebas, yaitu X_1 maupun X_2, mampu menerangkan atau menjelaskan 93,9% terhadap total varians permintaan minyak goreng. Nilai koefisien determinasi ini relatif besar dan menunjukkan bahwa variabel yang dipilih, yaitu harga dan pendapatan, dapat menjelaskan dengan baik terhadap variabel terikatnya, yaitu permintaan minyak goreng. Sedangkan sebanyak 6,10% (100 – 93,9%) dari varian permintaan minyak goreng dijelaskan oleh variabel yang tidak diukur atau diteliti, dalam hal ini disebut dengan residu. Kita mengetahui bahwa permintaan tidak hanya dipengaruhi oleh kedua variabel, tetapi juga dipengaruhi oleh musim, selera, harga barang, dan lain-lain, peran dari semua variabel yang tidak dimasukkan ini sebesar 6,10%.

Koefisien korelasi: hubungan variabel bebas dan terikat

Koefisien Korelasi digunakan untuk mengukur keeratan hubungan antara variabel terikat Y dengan variabel bebas X. Semakin besar nilai koefisien korelasi menunjukkan semakin eratnya hubungan dan sebaliknya. Koefisien korelasi merupakan akar kuadrat dari koefisien determinasi dan dirumuskan sebagai berikut.

$$R = \sqrt{R^2}$$

Kolerasi Berganda.
- semakin besar nilai koefesien kolerasi hubungan semakin erat.

Dari soal di atas, diketahui bahwa koefisien determinasi sebesar 0,939, maka nilai koefisien korelasinya adalah 0,939 = 0,969. Jadi, hubungan antara variabel terikat (permintaan minyak goreng) dengan variabel bebas (harga minyak dan pendapatan) sebesar 96,9%. Nilai ini menunjukkan hubungan yang kuat, karena lebih besar dari 0,5. Untuk regresi berganda, nilai koefisien korelasi relatif kurang penting dibandingkan dengan koefisien determinasi. Hal ini disebabkan karena regresi berganda tidak terfokus pada hubungannya, tetapi lebih berfokus pada seberapa besar variabel bebas dapat menerangkan dengan baik variabel terikatnya.

Korelasi Parsial. Korelasi parsial dalam regresi berganda digunakan untuk melihat besarnya hubungan antara dua variabel bebas dari variabel terikatnya. Pada Subbab 16.5 kita juga mengenal koefisien korelasi yang menghubungkan antara dua variabel seperti r_{YX1}, r_{YX2}, r_{X1X2}, yang menggambarkan hubungan antara Y dengan X_1, Y dengan X_2, dan X_1 dengan X_2. Koefisien korelasi yang demikian dikenal dengan koefisien korelasi tingkat nol. Pada regresi berganda kita juga mengenal koefisien korelasi yang dilambangkan R, koefisien ini menggambarkan hubungan antara Y dengan X_2 dan X_1 sekaligus. Korelasi parsial dilambangkan dengan $r_{YX1.X2}$ yang menyatakan hubungan antara y dengan X_1 di mana X_2 dianggap tetap, $r_{YX2.X1}$ yang menyatakan hubungan antara Y dengan X_2 di mana X_1 dianggap tetap, dan $r_{X1X2.Y}$ yang menyatakan hubungan antara X_1 dengan X_2 di mana Y dianggap tetap. Koefisien parsial ini memang khusus dimaksudkan untuk melihat hubungan dua variabel, dan terbebas dari pengaruh variabel lainnya dalam regresi berganda. Berikut adalah rumus-rumus untuk menghitung koefisien korelasi parsial, melalui koefisien korelasi sederhana atau korelasi tingkat nol.

Korelasi parsial: hubungan variabel bebas dan variabel terikat dengan variabel lain tetap

Koefisien Korelasi Sederhana

Korelasi sederhana

$$r_{YX1} = \frac{n\sum YX_1 - \sum Y \sum X_1}{\sqrt{\left[n\sum Y^2 - \left(\sum Y\right)^2\right]\left[n\sum X_1^2 - \left(\sum X_1\right)^2\right]}}$$

$$r_{YX2} = \frac{n\sum YX_2 - \sum Y \sum X_2}{\sqrt{\left[n\sum Y^2 - \left(\sum Y\right)^2\right]\left[n\sum X_2^2 - \left(\sum X_2\right)^2\right]}}$$

$$r_{X1X2} = \frac{n\sum X_1X_2 - \sum X_1 \sum X_2}{\sqrt{\left[n\sum X_1^2 - \left(\sum X_1\right)^2\right]\left[n\sum X_2^2 - \left(\sum X_2\right)^2\right]}}$$

Koefisien korelasi parsial diturunkan dari koefisien korelasi sederhana sebagai berikut.

Rumus: korelasi parsial

$$r_{YX1 \cdot X2} = \frac{r_{YX1} - r_{YX2} r_{X1X2}}{\sqrt{\left(1 - r_{YX2}^2\right)\left(1 - r_{X1X2}^2\right)}}$$

$$r_{YX2 \cdot X1} = \frac{r_{YX2} - r_{YX1} r_{X1X2}}{\sqrt{\left(1 - r_{YX1}^2\right)\left(1 - r_{X1X2}^2\right)}}$$

$$r_{X1 \cdot X2 \cdot Y} = \frac{r_{X1X2} - r_{YX2} r_{YX2}}{\sqrt{\left(1 - r_{YX1}^2\right)\left(1 - r_{YX2}^2\right)}}$$

CONTOH 16-3

Hitunglah koefisien korelasi parsial pada Contoh 16-1, dan apa artinya?

Contoh: menghitung koefisien korelasi

Jawab:

Dari Contoh 16-1 diketahui:

$\Sigma YX_1 = 409$ $\Sigma Y = 68$ $\Sigma X_1 = 63$ $\Sigma Y^2 = 516$ $\Sigma X_1^2 = 405$

$\Sigma YX_2 = 239$ $\Sigma X_2 = 46$ $\Sigma X_2^2 = 324$ $\Sigma X_1X_2 = 317$

Langkah 1 Menghitung Koefisien Korelasi Sederhana

$$r_{YX_1} = \frac{10.409-(68).(63)}{\sqrt{[10.(516)-68^2][10.(405)-63^2]}} = -\frac{194}{208} = -0,931$$

$$r_{YX_2} = \frac{10.239-(68).(46)}{\sqrt{[10.(516)-68^2][10.(324)-46^2]}} = -\frac{738}{776} = -0,951$$

$$r_{X_1X_2} = \frac{10.317-(63).(46)}{\sqrt{[10.(405)-63^2][10.(324)-46^2]}} = \frac{272}{301} = 0,901$$

Langkah 2 Menghitung Koefisien Korelasi Parsial

$$r_{YX_1 \cdot X_2} = \frac{-0,931-(-0,951).(0,901)}{\sqrt{\left(1-(-0,951)^2\right)\left(1-(0,901)^2\right)}} = -0,074/0,134 = -0,55$$

$$r_{YX_2 \cdot X_1} = \frac{-0,951-(-0,931).(0,901)}{\sqrt{\left(1-(-0,931)^2\right)\left(1-(0,901)^2\right)}} = -0,111/0,158 = -0,71$$

$$r_{X_1.X_2.Y} = \frac{0,901-(-0,931).(-0,951)}{\sqrt{\left(1-(-0,931)^2\right)\left(1-(-0,951)^2\right)}} = 0,016/0,113 = 0,143$$

Nilai koefisien korelasi parsial untuk Y dengan X_1, di mana X_2 tetap ($r_{YX1.X2}$), dan Y dengan X_2, di mana X_1 tetap ($r_{YX2.X1}$), lebih besar dari 0,5 yang berarti menunjukkan hubungan antara Y dengan X_1, maupun X_2 erat. Sedangkan nilai korelasi antara X_2 dengan X_1 lebih kecil dari 0,5 menunjukkan hubungan yang lemah. Dalam regresi berganda, kondisi demikian justru yang diharapkan, yaitu terjadinya hubungan Y dengan X yang kuat disertai hubungan antara X dengan X lainnya yang lemah. Kondisi ini diperlukan karena kalau hubungan antara X kuat, maka akan terjadi multikolinieritas (akan dijelaskan pada subbab pelanggaran asumsi) yang menyebabkan nilai koefisien determinasi turun atau melemah.

16.4 Kesalahan Baku dalam Regresi Berganda

Kesalahan baku: besarnya penyimpangan nilai dugaan dengan nilai sebenarnya

Kesalahan baku dalam regresi berganda adalah suatu ukuran untuk melihat ketepatan antara nilai dugaan ($\hat{Y}$) dengan nilai sebenarnya (Y) sebagaimana juga berlaku untuk regresi sederhana yang telah dibahas pada Bab 15. Apabila nilai dugaan semakin mendekati nilai sebenarnya, maka persamaan yang kita peroleh juga semakin baik. Apabila nilai dugaan semakin jauh dari nilai sebenarnya, maka persamaan yang kita gunakan juga tidak baik. Dari hasil regresi kita mendapatkan persamaan $\hat{Y} = 15,086 - 1,015X_1 - 0,41X_2$, sehingga kita dapat menduga berapa permintaan sesorang dengan pendapatan Rp5 juta per bulan dan harga minyak Rp7 ribu per liter. Nilai dugaannya adalah $\hat{Y} = 15,086 - 1,015(7) - 0,41(5) = 5,92$ liter per bulan. Pada kenyataannya, orang tersebut membeli 6 liter per bulan, sehingga ada selisih nilai sebesar $6 - 5,92 = 0,08$ liter per bulan. Perbedaan antara nilai dugaan dengan nilai sebenarnya ($\hat{Y} - Y$)

disebut dengan residu atau *error*. Berdasarkan pada kenyataan bahwa tidak mungkin untuk mendapatkan nilai dugaan dengan ketepatan 100%, maka diperlukan suatu ukuran seberapa besar ketidakakuratan pendugaan terjadi. Suatu ukuran yang mengukur ketidakakuratan nilai dugaan disebut dengan Kesalahan Baku Pendugaan atau Standar *Error*. Kesalahan baku pendugaan untuk regresi berganda dirumuskan sebagai berikut.

Rumus: kesalahan baku

$$S_{Y.X1.X2} = \sqrt{\frac{\Sigma\left(\hat{Y}-Y\right)^2}{n-(k+1)}}$$

di mana:

- $S_{Y.X1.X2}$: Kesalahan baku atau standar *error* pendugaan variabel Y berdasarkan variabel X_1 dan X_2
- $\hat{Y}$: Nilai dugaan dari Y di mana X_1 dan X_2 diketahui
- Y : Nilai pengamatan dari Y
- n : Jumlah sampel atau data
- k : Jumlah variabel bebas

CONTOH 16-4

Hitunglah kesalahan baku pendugaan dari persamaan pada Contoh 16-1!

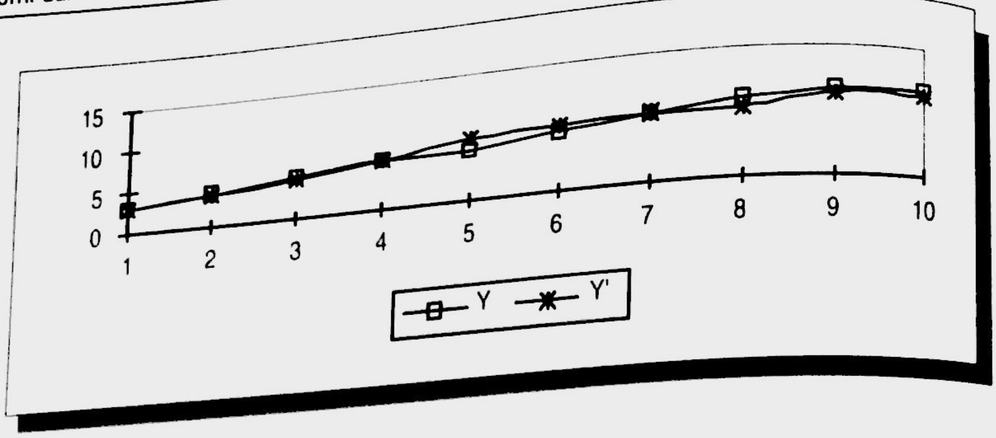
Contoh: menghitung kesalahan baku pendugaan

Jawab:

Untuk mendapatkan nilai kesalahan baku standar, nilai pendugaan harus diketahui. Nilai pendugaan dapat diperoleh melalui proses berikut.

Y	X_1	X_2	$\hat{Y} = 15{,}086 - 1{,}015X_1 - 0{,}41X_2$	$(Y-\hat{Y})$	$(Y-\hat{Y})^2$
3	8	10	2,86 = 15,086 – 1,015 (8) – 0,41(10)	0,14	0,02
4	7	10	3,87 = 15,086 – 1,015 (7) – 0,41(10)	0,13	0,02
5	7	8	4,69 = 15,086 – 1,015 (7) – 0,41 (8)	0,31	0,09
6	7	5	5,92 = 15,086 – 1,015 (7) – 0,41 (5)	0,08	0,01
6	6	4	7,35 = 15,086 – 1,015 (6) – 0,41 (4)	–1,35	1,83
7	6	3	7,76 = 15,086 – 1,015 (6) – 0,41 (3)	–0,76	0,58
8	6	2	8,17 = 15,086 – 1,015 (6) – 0,41 (2)	–0,17	0,03
9	6	2	8,17 = 15,086 – 1,015 (6) – 0,41 (2)	0,83	0,68
10	5	1	9,60 = 15,086 – 1,015 (5) – 0,41 (1)	0,40	0,16
10	5	1	9,60 = 15,086 – 1,015 (5) – 0,41 (1)	0,40	0,16
				$\Sigma\left(\hat{Y}-Y\right)^2$	3,58

Dari data di atas kita dapat membuat grafik hubungan Y dengan $\hat{Y}$, semakin dekat nilai Y dengan $\hat{Y}$, maka semakin kecil jarak nilai Y dengan $\hat{Y}$.



Nilai dugaan pada sampel 1 sampai 4 dan 7 relatif lebih baik dibandingkan dengan sampel 5 begitu pula untuk sampel 8, 9, dan 10. Nilai kesalahan baku $s_{Y.X1.X2}$ dapat dihitung sebagai berikut.

$$S_{Y.X1.X2} = \sqrt{\frac{\Sigma(\hat{Y}-Y)^2}{n-(k+1)}} = \sqrt{\frac{3,58}{10-(2+1)}} = 0,72$$

Nilai kesalahan baku sebesar 0,72; apa artinya nilai ini? Secara ekonomi dapat diartikan bahwa permintaan minyak goreng menyebar secara normal di sekitar garis regresi berganda $(\hat{Y})$ pada kisaran $\pm$ 0,72 liter/bulan atau $\pm\ s_{Y.X1.X2}$ sebanyak 68%, sedangkan 95,5% akan berada pada kisaran $\pm\ 2 \times 0{,}72$ liter/bulan atau $\pm\ 2 \times S_{Y.X1.X2}$, dan 99,7% akan berada pada kisaran $\pm\ 3 \times 0{,}72$ liter per bulan atau $\pm\ 3 \times s_{Y.X1.X2}$. Apabila sampel mempunyai pendapatan 5 juta per bulan dengan harga minyak 7 ribu, maka dapat diduga bahwa permintaannya adalah $\hat{Y} = 15{,}086 - 1{,}015$ (7) – 0,41 (5) 5,92 = 5,92. Dengan kesalahan baku 0,72, maka nilai sebenarnya (Y) sebanyak 68% akan berkisar antara 5,92 $\pm$ 0,72 atau antara 5,20 sampai 6,63 liter per bulan, demikian juga berlaku untuk probabilitas 95,5% dan 99,7%.

Kesalahan baku $s_{Y.X1.X2}$ yang diperoleh dengan cara menghitung $\hat{Y}$ dan selisih atau residu, yaitu $(\hat{Y} - Y)$, membutuhkan waktu yang relatif lama. Untuk kasus regresi berganda dengan 2 variabel independen (k=2) ada rumus lain yang dapat membantu mempercepat perhitungan, yaitu:

Rumus lain kesalahan baku

$$S_{Y \cdot X1 \cdot X2} = \sqrt{\frac{\sum Y^2 - a\sum Y - b_1\sum X_1Y - b_2\sum X_2Y}{n-3}}$$

Dari perhitungan sebelumnya, kita mengetahui $\sum Y^2 = 516$, a= 15,086, $\sum Y = 68$, $b_1 = -1{,}01524$, $\sum X_1Y = 409$, $b_2 = -0{,}41$, dan $\sum X_2Y = 239$, sehingga nilai kesalahan baku adalah:

$$S_{Y \cdot X1 \cdot X2} = \sqrt{\frac{\sum Y^2 - a\sum Y - b_1\sum X_1Y - b_2\sum X_2Y}{n-3}}$$

$$S_{Y \cdot X1 \cdot X2} = \sqrt{\frac{516-(15{,}086\times 68)-(-1{,}01524\times 409)-(-0{,}41\times 239)}{10-3}} = 0{,}72$$