

**Kesalahan Baku Penduga.** Kesalahan baku penduga disebut juga dengan *standard error estimation*, yaitu melihat seberapa jauh nilai penduga, yaitu  $b_1$  dan  $b_2$ , dari nilai sebenarnya, yaitu  $B_1$  dan  $B_2$  (parameter populasi). Oleh karena nilai ini menunjukkan besarnya penyimpangan atau *error*, maka semakin kecil nilainya dianggap akan lebih baik. Untuk menghitung kesalahan baku penduga dapat digunakan rumus sebagai berikut.

$$sb_1 = \frac{S_{Y.X_1X_2}}{\sqrt{(\sum X_1^2 - n\bar{X}_1^2)(1 - r_{X_1X_2}^2)}}$$

$$sb_2 = \frac{S_{Y.X_1X_2}}{\sqrt{(\sum X_2^2 - n\bar{X}_2^2)(1 - r_{X_1X_2}^2)}}$$

di mana:

- $sb_1$  : Kesalahan baku penduga  $b_1$
- $sb_2$  : Kesalahan baku penduga  $b_2$
- $S_{Y.X_1X_2}$  : Standar *error* variabel Y berdasarkan variabel X yang diketahui
- $\sum X_1^2$  : Jumlah  $X_1$  kuadrat
- $\sum X_2^2$  : Jumlah  $X_2$  kuadrat
- $r_{X_1X_2}$  : Korelasi sederhana antara  $X_1$  dan  $X_2$
- $\bar{X}_2^2$  : Jumlah kuadrat dari  $X_2$  rata-rata
- $\bar{X}_1^2$  : Jumlah kuadrat dari  $X_1$  rata-rata

#### CONTOH 16-6

Dengan menggunakan data pada Contoh 16-1, 16-3, dan 16-4, hitunglah kesalahan baku penduga untuk  $b_1$  dan  $b_2$ !

**Jawab:**

Dari soal 16-1 diketahui:  $\sum X_1^2 = 405$ ,  $\bar{X}_1 = 39,69$ ;  $\sum X_2^2 = 324$ ,  $\bar{X}_2 = 21,16$

Dari soal 16-4 diketahui:  $S_{Y.X_1X_2} = 0,72$

Dari soal 16-3 diketahui:  $r_{X_1X_2} = 0,901$

$$S_{b_1} = \frac{0,72}{\sqrt{(405 - (10)(39,69))(1 - 0,901^2)}} = 0,72/1,23 = 0,580$$

$$S_{b_2} = \frac{0,72}{\sqrt{(324 - (10)(21,16))(1 - 0,901^2)}} = 0,72/4,59 = 0,156$$

Kesalahan baku penduga  $b_1$  sebesar 0,580 menunjukkan bahwa nilai dugaan populasi ( $B_1$ ), yang diduga dengan sampel ( $b_1$ ) berada pada kisaran  $b_1 \pm 0,580$ , sehingga kisaran dugaan nilai koefisien parameter populasi dapat dinyatakan dalam interval, yaitu  $0,435 < B_1 < 1,95$ . Dengan cara yang sama didapatkan kesalahan baku penduga  $b_2$  sebesar 0,156, sehingga nilai dugaan populasinya berada pada kisaran  $-0,41 \pm 0,156$  atau pada interval  $-0,566 < B_2 < -0,254$ .

... output menyatakan sebagai ... rarian karian kolom koefisien. Pada kolom standar error, terlihat  $sb_1 = 0,580445685$  dan  $sb_2 = 0,155821874$ . Perhatikan gambar output berikut.



Menggunakan MS Excel untuk nilai kesalahan baku

The screenshot shows the following data from the Excel output:

Regression Statistics	
Multiple R	0.966030126
R Square	0.933214204
Adjusted R Square	0.918132549
Standard Error	0.715114241
Observations	20

ANOVA					
	df	SS	MS	F	Significance F
Regression	2	30.62028136	15.31014068	48.90633334	0.0000779
Residual	17	3.37971864	0.198807		
Total	19	34			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	15.0816687	1.01512876	14.85522	0.001818155	7.90927467	22.26305277	7.90927467	22.26305277
X Variable 1	-1.015480128	0.580445685	-1.74904	0.123766217	-2.357799917	0.357119261	-2.357799917	0.357119261
X Variable 2	0.41902098	0.155821874	2.687003	0.01368316	0.779362878	0.04247915	0.779362878	0.04247915

## 16.5 Pengujian Hipotesis pada Regresi Berganda

Pada subbab ini akan dibahas tentang pengujian hipotesis, baik pengujian secara bersama-sama maupun parsial. Pengujian hipotesis untuk melihat apakah variabel bebas mampu secara menyeluruh bersama-sama menjelaskan tingkah laku variabel terikat adalah dengan menggunakan Uji Global atau Uji F. Selain mengetahui kemampuan secara bersama-sama variabel bebas menjelaskan variabel terikat, juga perlu mengetahui apakah setiap variabel bebas juga berpengaruh terhadap variabel terikatnya, untuk pengujian ini digunakan Uji t.

**Pengujian hipotesis:** apakah pengaruhnya sama dengan nol?

### 16.5.1 Uji Global

Uji global disebut juga uji signifikansi serentak atau Uji F. Uji ini dimaksudkan untuk melihat kemampuan menyeluruh dari variabel bebas ( $X_1, X_2, \dots, X_k$ ) dapat atau mampu menjelaskan tingkah laku atau keragaman variabel terikat ( $Y$ ). Uji global juga dimaksudkan untuk mengetahui apakah semua variabel bebas memiliki koefisien regresi sama dengan nol.

**Uji global:** uji kemampuan variabel bebas terhadap variabel terikat secara bersama-sama

Untuk melakukan pengujian secara global, ada beberapa langkah yang diperlukan, yaitu:

1. **Menyusun hipotesis.** Hipotesis yang ingin diuji adalah kemampuan variabel bebas menjelaskan tingkah laku variabel terikat, apabila variabel bebas tidak dapat memengaruhi variabel terikat dapat dianggap nilai koefisien regresinya sama dengan nol, sehingga

berapapun nilai variabel terikat, tidak akan berpengaruh terhadap variabel bebas. Persamaan pada contoh satu, yaitu  $Y = 15,086 - 1,015X_1 - 0,41X_2$ , variabel bebas  $X_1$  dan  $X_2$  dikatakan mampu memengaruhi  $Y$  apabila nilai koefisien  $b_1$  dan  $b_2$  tidak sama dengan nol, apabila sama dengan nol, maka dikatakan tidak mampu memengaruhi variabel bebas  $Y$ .

Dalam menyusun hipotesis, selalu ada hipotesis nol dan hipotesis alternatif. Hipotesis nol selalu mengandung unsur kesamaan, sehingga dapat dirumuskan hipotesis nol adalah koefisien regresi sama dengan nol. Untuk hipotesis alternatifnya adalah koefisien regresi tidak sama dengan nol. Kedua hipotesis tersebut kemudian dirumuskan sebagai berikut.

$$H_0 : B_1 = B_2 = 0$$

$$H_1 : B_1 \neq B_2 \neq 0$$

- Menentukan daerah keputusan.** Penentuan daerah keputusan dilakukan dengan mencari nilai  $F$ . Untuk mencari nilai  $F$ -tabel perlu diketahui derajat bebas pembilang pada kolom, derajat bebas penyebut pada baris, dan taraf nyata. Umumnya ada 2 taraf nyata yang dipakai, yaitu 1% dan 5%, untuk ilmu pasti lebih baik menggunakan taraf nyata 1%, sedangkan ilmu sosial dapat menggunakan taraf nyata 5%. Untuk derajat pembilang, digunakan nilai  $k - 1$ , yaitu jumlah variabel dikurangi 1. Untuk derajat penyebut digunakan  $n - k$ , yaitu jumlah sampel dikurangi dengan jumlah variabel.

Berikut adalah contoh mencari nilai  $F$ -tabel dari tabel  $F$ . Pada contoh soal sebelumnya, diketahui ada tiga variabel, yaitu  $Y, X_1$ , dan  $X_2$ , jadi  $k = 3$ , sedangkan jumlah  $n = 10$ . Jadi derajat pembilang  $k - 1 = 3 - 1 = 2$ , sedangkan derajat penyebutnya adalah  $n - k = 10 - 3 = 7$  dengan taraf nyata 5%. Nilai  $F$ -tabel dengan derajat pembilang 2, penyebut 7 dan taraf nyata 5% adalah 4,74. Tabel selengkapnya disajikan pada lampiran.

		Derajat bebas pembilang							
		1	2	3	4	5	...	120	$\infty$
Derajat bebas penyebut	1	161	200	216	225	230	...	253	254
	2	18,5	19,0	19,2	19,2	19,3	...	19,5	19,5
	3	10,1	9,55	9,28	9,12	9,01	...	8,55	8,53
	4	7,71	6,94	6,59	6,39	6,26	...	5,66	5,63
	5	6,61	5,79	5,41	5,19	5,05	...	4,40	4,37
	6	5,99	5,14	4,76	4,53	4,39	...	3,70	3,67
	7	5,59	4,74	4,35	4,12	3,97	...	3,27	3,23
	...	...	...	...	...	...	...	...	...
	$\infty$	3,84	3,00	2,60	2,37	2,21	...	1,22	1,00

3. **Menentukan nilai F-hitung.** Nilai F-hitung ditentukan dengan rumus sebagai berikut.

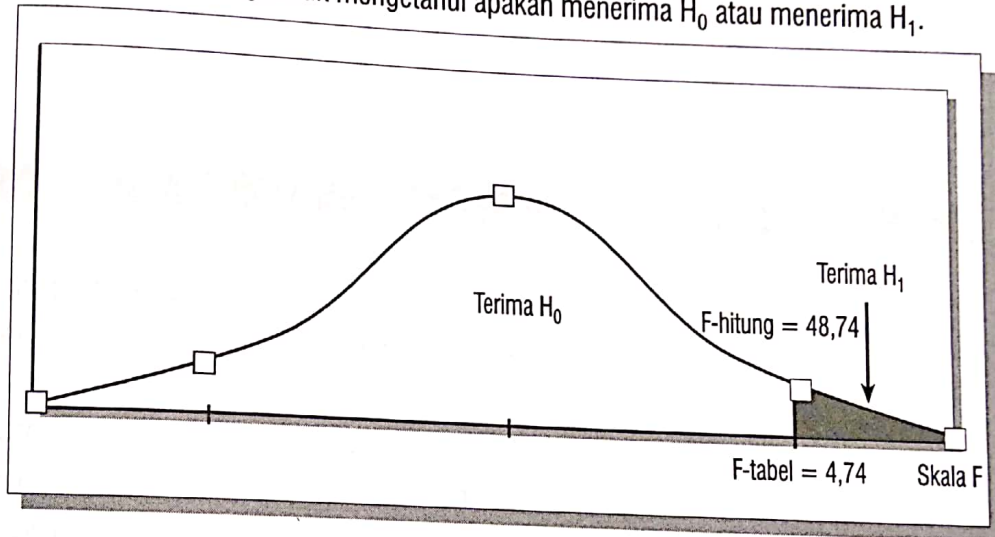
$$F = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - 3)}$$

Rumus: F-hitung

Dari soal diketahui bahwa  $R^2 = 0,933$  dan  $n = 10$ , sehingga nilai F-hitung adalah:

$$F = \frac{0,933 / (3 - 1)}{(1 - 0,933) / (10 - 3)} = 0,4665 / 0,0096 = 48,73881$$

4. **Menentukan daerah keputusan.** Menentukan wilayah  $H_0$  dan  $H_1$ , serta membandingkan dengan nilai F-hitung untuk mengetahui apakah menerima  $H_0$  atau menerima  $H_1$ .



Daerah keputusan menerima atau menolak hipotesis

5. **Memutuskan hipotesis.** Berdasarkan Gambar pada langkah 4, nilai F-hitung  $>$  dari F-tabel dan berada di daerah terima  $H_1$ . Ini menunjukkan bahwa terdapat cukup bukti untuk menolak  $H_0$  dan menerima  $H_1$ . Kesimpulan dari diterimanya  $H_1$  adalah nilai koefisien regresi tidak sama dengan nol, dengan demikian variabel bebas dapat menerangkan variabel terikat, atau dengan kata lain variabel bebas, yaitu  $X_1$  dan  $X_2$  secara bersama-sama berpengaruh nyata terhadap variabel terikatnya ( $Y$ ), yaitu permintaan minyak goreng.

Dalam praktik sehari-hari, Anda tidak harus dipusingkan dengan perhitungan dan pencarian yang rumit. Analisis komputer yang sudah dilakukan pada Soal 1 dapat langsung memberikan nilai F-hitungnya. Pada *output* terlihat bahwa nilai F-hitung adalah 48,91. Untuk memutuskan apakah akan menerima atau menolak  $H_0$ , kita harus membandingkan F-hitung dengan F-tabel. Apabila F-hitung  $>$  F-tabel, maka  $H_0$  ditolak dan apabila F-hitung  $<$  F-tabel, maka  $H_0$  diterima. Hasil perhitungan menunjukkan bahwa F-hitung 48,91 dan F-tabel 4,74 jadi F-hitung  $>$  F-tabel dan kesimpulannya adalah tolak  $H_0$  dan terima  $H_1$ .

## 16.5.2 Uji Signifikansi Parsial atau Individual

Uji signifikansi parsial atau individual digunakan untuk menguji apakah suatu variabel bebas berpengaruh atau tidak terhadap variabel terikat. Pada regresi berganda  $Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k$ , mungkin variabel  $X_1$  sampai  $X_k$  secara bersama-sama berpengaruh nyata. Namun demikian, belum tentu secara individu atau parsial seluruh variabel dari  $X_1$  sampai  $X_k$  berpengaruh nyata terhadap variabel terikatnya ( $Y$ ). Nyata atau tidaknya pengaruh suatu variabel bebas terhadap variabel terikatnya juga bergantung pada hubungan variabel tersebut dan kondisi sosial dan ekonomi masyarakat.

Untuk mengetahui apakah suatu variabel secara parsial berpengaruh nyata atau tidak digunakan uji  $t$  atau *t-student*. Untuk melakukan uji  $t$  ada beberapa langkah yang diperlukan seperti berikut.

1. **Menentukan Hipotesis.** Variabel bebas berpengaruh tidak nyata apabila nilai koefisiennya sama dengan nol, sedangkan variabel bebas akan berpengaruh nyata apabila nilai koefisiennya tidak sama dengan nol. Hipotesis lengkapnya adalah sebagai berikut.

$$H_0 : B_1 = 0 \quad H_1 : B_1 \neq 0$$

$$H_0 : B_2 = 0 \quad H_1 : B_2 \neq 0$$

2. **Menentukan daerah kritis.** Daerah kritis ditentukan oleh nilai t-tabel dengan derajat bebas  $n - k$ , dan taraf nyata  $\alpha$ . Untuk kasus ini derajat bebasnya adalah  $10 - 3 = 7$ , sedangkan taraf nyata  $\alpha$  dapat 1%, 5% atau yang lainnya. Untuk kasus ini digunakan taraf nyata 5%. Dengan melakukan uji dua arah, maka nilai t-hitung adalah 2,36. Berikut adalah cara penggunaan t-tabel. Perhatikan untuk taraf nyata 5%, karena menggunakan uji dua arah, maka ujungnya ada dua sehingga  $5\%/2 = 2,5\% = 0,025$  dan derajat bebasnya adalah 7.

		Taraf Nyata					
		0,25	0,10	0,05	0,025	0,01	0,005
Derajat Bebas	1	1,00	3,08	6,31	12,71	31,82	63,06
	2	0,82	1,89	2,92	4,30	6,98	9,92
	3	0,76	1,64	2,35	3,18	4,54	5,84
	4	0,74	1,53	2,13	2,78	3,75	4,60
	5	0,73	1,48	2,02	2,57	3,36	4,03
	6	0,72	1,44	1,94	2,45	3,14	3,71
	7	0,71	1,41	1,89	2,36	3,00	3,50
	...						
	$\alpha$	0,67	1,28	1,64	1,96	2,33	2,58

3. **Menentukan Nilai t-hitung.** Nilai t-hitung untuk koefisien  $b_1$  dan  $b_2$  dapat dirumuskan sebagai berikut.

$$t\text{-hitung} = \frac{b - B}{S_b}$$

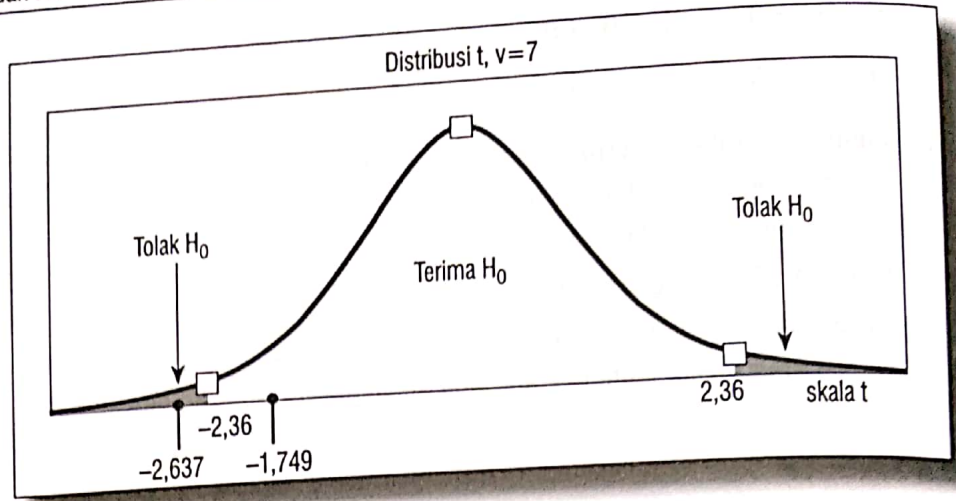
Nilai t-hitung untuk  $b_1$

$$t\text{-hitung} = \frac{b_1 - B_1}{S_{b_1}} = \frac{-1,015 - 0}{0,58} = -1,750$$

Nilai t-hitung untuk  $b_2$

$$t\text{-hitung} = \frac{b_2 - B_2}{S_{b_2}} = \frac{-0,410 - 0}{0,1558} = -2,637$$

4. **Menentukan daerah keputusan.** Daerah keputusan untuk menerima  $H_0$  atau menolak  $H_0$  dengan derajat bebas 7 dan taraf nyata 5% untuk uji dua arah adalah sebagai berikut.



- Menentukan Keputusan.** Nilai t-hitung untuk koefisien regresi  $b_1$  sebesar  $-1,750$  dan berada di daerah terima  $H_0$ . Ini menunjukkan bahwa koefisien regresi tidak berbeda dengan nol atau variabel  $X_1$  tidak berpengaruh nyata terhadap  $Y$ . Nilai t-hitung untuk koefisien regresi  $b_2$  sebesar  $-2,637$  dan berada di daerah tolak  $H_1$ . Ini menunjukkan bahwa koefisien regresi berbeda dengan nol, atau variabel  $X_2$  berpengaruh nyata terhadap  $Y$ . Dalam kehidupan sehari-hari, kita cukup memerhatikan variabel yang nyata saja sebagai dasar pengambilan keputusan.

Komputer mempermudah pekerjaan pengujian parsial ini. Pada *output* hasil regresi sudah tercantum secara otomatis nilai t-hitung. Kita hanya memerlukan nilai t-tabel sesuai dengan derajat bebas dan taraf nyatanya. Suatu variabel akan berpengaruh nyata apabila nilai t-hitung lebih besar dari t-tabel ( $t\text{-hitung} > 2,36$ ) atau lebih kecil dari negatif t-tabel ( $t\text{-hitung} < -2,36$ ). Dengan memperhatikan kondisi tersebut, maka dengan mudah dapat dilihat bahwa  $b_1$  tidak berpengaruh nyata terhadap  $Y$  dan  $b_2$  berpengaruh nyata terhadap  $Y$ .



Screen: penggunaan Excel untuk mencari t-hitung

Regresi dengan Excel - Excel

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0,96030126					
R Square	0,91214204					
Adjusted R Square	0,914132549					
Standard Error	0,715154241					
Observations	10					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	2	50,02028136	25,01014068	48,90453136	0,0000770	
Residual	7	8,37971864	0,911388377			
Total	9	58,4				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	15,04610647	3,015108769	4,970552188	0,001618163	7,90927467	22,26305827
X Variable 1	-1,015240128	0,580459685	-1,749042001	0,123766217	-2,387799917	0,357919261
X Variable 2	-0,410902696	0,155822874	-2,637003664	0,033569316	-0,779362878	-0,042442515

Hasil pengolahan data dengan bantuan komputer, memungkinkan Anda tidak perlu melihat t-tabel, karena *output* dari pengolahan dengan bantuan program Excel sudah menampilkan (*P-value*) dari t-hitung setiap koefisien. Dengan membandingkan nilai *P-value* dengan taraf pengujian (1% atau 5%), maka dapat diketahui apakah  $H_0$  diterima atau ditolak. Jika nilai *P-value* < taraf pengujian, maka kesimpulannya tolak  $H_0$  (terima  $H_1$ ) dan sebaliknya jika nilai *P-value* > taraf pengujian terima  $H_0$ . Dengan menggunakan contoh kasus sebelumnya, maka diperoleh nilai *P-value* untuk koefisien  $b_1$  sebesar 0,124 yang lebih besar dari taraf 5% sehingga terima  $H_0 : B_1 = 0$ ; dan *P-value* untuk koefisien  $b_2$  sebesar 0,034 (> 0,025) sehingga kesimpulannya adalah tolak  $H_0$  (terima  $H_1 : B_2 \neq 0$ ).

## 16.6 Asumsi dan Pelanggaran Asumsi pada Regresi Berganda

Sejauh ini telah dibahas mengenai hal-hal penting dalam regresi berganda, yaitu koefisien regresi, koefisien determinasi dan korelasi, kesalahan baku pendugaan, dan pengujian hipotesis. Pada subbab ini akan dibahas mengenai asumsi-asumsi pokok dalam regresi berganda, sehingga nilai koefisien regresi yang dihasilkan baik atau tidak bias. Beberapa asumsi dalam regresi berganda adalah sebagai berikut.

**Asumsi:** hasilnya tidak bias dan efisien

1. Variabel terikat dan variabel bebas memiliki hubungan yang linier atau hubungan garis lurus. Jadi hubungan Y dengan X harus linier, bagaimana kalau tidak linier? Masalah ini akan dibahas pada Bab 17. Untuk persamaan yang tidak linier, maka datanya harus ditransformasi terlebih dahulu menjadi linier, biasanya data di-log-kan terlebih dahulu sehingga menjadi linier.
2. Variabel terikat haruslah variabel yang bersifat kontinu dan paling tidak berskala selang. Variabel kontinu ini adalah variabel yang dapat menempati pada semua titik dan biasanya merupakan data dari proses pengukuran.
3. Nilai keragaman atau residu, yaitu selisih antara data pengamatan dan data dugaan hasil regresi harus sama untuk semua nilai Y. Asumsi ini menyatakan bahwa nilai residu bersifat konstan untuk semua data Y. Asumsi ini memperlihatkan kondisi HOMOSKEDASTISITAS, yaitu nilai residu yang sama untuk semua nilai Y, menyebar normal, dan mempunyai rata-rata 0.
4. Pengamatan-pengamatan untuk variabel terikat dari satu pengamatan ke pengamatan lain harus bebas atau tidak berkorelasi. Hal ini penting untuk data yang bersifat deret berkala.

**Asumsi-asumsi linier berganda**

Asumsi-asumsi tersebut harus dipenuhi dalam menyusun regresi berganda agar hasilnya tidak bias. Untuk itu, perlu dilakukan beberapa tes yang memungkinkan pendeteksian pelanggaran asumsi biasanya adalah *multikolinieritas*,