STATISTIKA DAN PROBABILTAS

JURUSAN TEKNIK INFORMATIKA

DARMAJAYA

2022

Hipotesis merupakan suatu anggapan yang mungkin benar atau tidak mengenai suatu populasi atau lebih. Penolakan suatu hipotesis berarti menyimpulkan bahwa hipotesis itu tidak benar, sedangkan penerimaan hipotesis menunjukkan bahwa tidak cukup petunjuk untuk mempercayai hal yang sebaliknya. Perhatikan tabel berikut :

Peluang	**Kenyataan H_0 benar**	**Kenyataan H_0 salah**
menolak H_0	**α Galat tipe I (taraf keberartian)**	**1- α**
menerima H_0	**1-β**	**β Galat tipe II (kuasa uji)**

Peluang menolak H_0 padahal kenyataannya H_0 benar adalah . Memperkecil galat jenis II akan menaikkan peluang melakukan galat jenis I atau . Akan tetapi peluang melakukan kedua jenis galat dapat diperkecil dengan memperbesar ukuran sampel.

Langkah Pengujian Hipotesis

1. Rumuskan hipotesis nol dan hipotesis tandingannya ;
2. Pilih taraf keberartian atau α ;
3. Pilih uji statistik yang sesuai dan cari daerah kritisnya ;
4. Hitunglah nilai statistik dari sampel acak ukuran n.

5. Kesimpulan : tolak H_0 bila statisik tersebut mempunyai nilai dalam daerah kritis (daerah penolakan H_0); jika tidak terima H_0.

Uji Rataan

Perhatikan contoh tentang uji rataan berikut.

Contoh :

Misalkan rata-rata berat mahasiswa pria di suatu Perguruan Tinggi berdistribusi normal dengan simpangan baku populasi 3,6 kg. Uji bahwa rata-rata berat mahasiswa pria tersebut 68 kg lawan rata-rata berat mahasiswa tersebut tidak sama dengan 68 kg. Jika diambil sampel berukuran 36 dan dihitung ternyata dengan rata-rata sampel 67 kg. Apa kesimpulan anda ? Pilih taraf keberartian : α = 5%.

Penyelesaian :

- Akan diuji H_0 : μ = 68 (μ0) vs H1 : μ ≠ 68 .
- Dibawah H_0 : $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

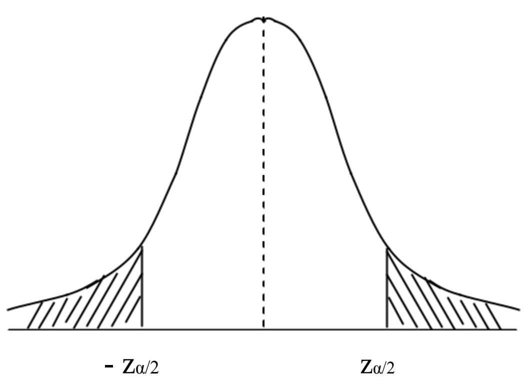
  Jika dipilih α = 5%, maka berarti :

  $$P(Z \quad z_{\alpha/2} \mid H_0 benar) \qquad P(Z \quad z_{\alpha/2} \mid H_0 benar)$$

- Dari tabel : $z_{\alpha/2} = z_{0,025} = 1,96$.
- z hitung = $(\bar{x} - \mu_0) / (\sigma/\sqrt{n})$

  = (67 - 68) / (3,6 / 6) = 1,67.

- Karena z hitung < $z_{\alpha/2}$, maka H_0 diterima. z hitung masuk dalam daerah penerimaan yaitu daerah diantara - $z_{\alpha/2}$ dan $z_{\alpha/2}$.



Contoh tadi merupakan uji dua arah karena ada dua daerah penolakan yaitu $Z > z_{\alpha/2}$ untuk $\mu>\mu_0$ (kanan) dan $Z < - z_{\alpha/2}$ untuk $\mu<\mu_0$ (kiri). Sedangkan uji satu arah mengenai rataan :

(i) $H_0 : \mu=\mu_0$ vs $H_1 : \mu>\mu_0$

(ii) $H_0 : \mu=\mu_0$ vs $H_1 : \mu<\mu_0$.

Contoh :

Rata-rata waktu yang diperlukan siswa untuk mendaftar pada permulaan kuliah baru di suatu PT pada waktu lalu adalah 50 menit dengan simpangan baku 10 menit. Suatu cara pendaftaran baru dengan menggunakan komputer yang sedang dicobakan. Bila sampel acak dengan 12 mahasiswa membutuhkan rata-rata mendaftarkan diri 42 menit dengan simpangan baku 11,9 menit menggunakan cara baru, ujilah hipotesis bahwa rataan populasi sekarang lebih kecil dari 50 dengan menggunakan taraf keberartian 0,05 dan 0,01. Anggap populasi waktu mendaftar berdistribusi normal.

Penyelesaian :

Uji : $H_0 : \mu = 50$ menit vs $H_1 : \mu< 50$ menit .

Pilih (1)   = 0,05 dan (2) 0,01, sehingga daerah kritis : (1) $T < -1{,}796$;

(2) $T < -2{,}718$.

Nilai t hitung = $\frac{\bar{x} \quad _0}{s\sqrt{\ n}} = \frac{42 \quad 50}{11{,}9/\sqrt{12}} = - 2{,}33.$

Kesimpulan : tolak H_0 pada taraf keberartian 0,05 tapi tidak pada taraf 0,01.

Ini berarti bahwa rataan sesungguhnya kemungkinan besar kurang dari 50 menit tapi perbedaannya tidaklah begitu besar sehingga penggunaan komputer dengan biaya yang begitu besar tidaklah menguntungkan.

Uji Proporsi

Contoh :

Suatu pabrik mengeluarkan suatu pernyataan bahwa 90% dari barang produksinya tidak cacat. Suatu peningkatan proses sedang dicobakan dan menurut mereka akan menurunkan proporsi yang cacat di bawah 10% yang sekarang. Dalam suatu percobaan dengan 100 barang yang dihasilkan dengan proses baru tersebut ternyata ada 5 yang cacat. Apakah kenyataan ini cukup untuk menyimpulkan bahwa telah ada peningkatan proses? Gunakan taraf keberartian 0,05.

Penyelesaian :

Uji : H_0 : p = 0,9 vs H_1 : p > 0,9.

= 0,05 , sehingga daerah kritis : (1) Z > 1,645.

Nilai z hitung = $\frac{x \;\; np_0}{\sqrt{np_0\,(1 \;\; p_0)}} = \frac{95 \;\; 90}{\sqrt{100(0,9)(0,1)}}$ = 1,67.

Kesimpulan : tolak H_0 dan simpulkan bahwa perbaikan telah menurunkan proporsi yang cacat.

Uji Normalitas

Uji kenormalan data nilai ujian statistika dasar 80 orang mahasiswa yang telah dibahas pada pertemuan kedua dengan uji khi-kuadrat dan uji K-S. Rumusan hipotesis yang akan diuji :

H_0 : Data berdistribusi normal vs

H_1 : Data tidak berdistribusi normal.

Rumusan hipotesis tersebut ekuivalen dengan :

$H_0 : F(x) = F^*(x)$ untuk semua x

$H_1 : F(x) \neq F^*(x)$ untuk paling sedikit satu x

Fungsi distribusi normal untuk v. a. X :

$$F^*(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt$$

Pengujian kenormalan dengan uji khi-kuadrat :

Statistik Uji (*Test Statistic*) : $T = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$

Di bawah H_0 , T berdistribusi Khi-Kuadrat dengan derajat kebebasan :

dk = banyaknya sel – banyaknya besaran yang diperoleh dari data amatan yang diperlukan dalam perhitungan frekuensi harapan.

Dari data diperoleh : mean (rata-rata) = 76,10 dan simpangan baku = 13,818.

Kemudian buatlah tabel berikut.

Kelas	**Batas Kelas**	**z untuk batas kelas**	**Cdf**	**Luas tiap kelas Interval**	**frekuensi amatan (O_i)**	**frekuensi Harapan (E_i)**	$\frac{(O_i - E_i)^2}{E_i}$
31-50	30.5	-3.30004	0.000483				
				0.031483	5	2.519	2.443573
51-60	50.5	-1.85266	0.031966				
				0.097491	5	7.799	1.004539
61-70	60.5	-1.12896	0.129457				
				0.213183	14	17.055	0.547231
71-80	70.5	-0.40527	0.34264				
				0.282279	24	22.582	0.089041
81-90	80.5	0.31843	0.624919				
				0.226403	20	18.112	0.196806
91-100	90.5	1.04212	0.851322				
				0.109964	12	8.797	1.166217
	100.5	1.76581	0.961286				
						Jumlah	5.447407

Misalkan dipilih α = 5% , karena t hitung = 5,4398 $\leq$ 7,815 = $\chi^2_{0,95}$,

dengan dk=6-3=3 , maka H_0 diterima. Jadi dapat disimpulkan data berdistribusi normal.

Pengujian kenormalan dengan uji Kolmogorov-Smirnov (K-S) :

Asumsi : Sampelnya adalah sampel acak

Statistik Uji : $T = \sup_{x} \left| F^*(x) - S(x) \right|$

dengan S(x) = fungsi distribusi empiris.

Tolak H_0 jika pada tingkat kepercayaan α , T $\geq w_{1-\alpha}$ (Conover, 1986).

Dapat ditunjukkan dengan menghitung statistik ujinya untuk K-S :

T hitung < $w_{1-\alpha}$ (Conover, hal. 462), maka H0 diterima.

Uji Kesamaan Dua Varians

Contoh :

Ada dua macam pengukuran kelembaban suatu zat. Cara pertama dilakukan 10 kali yang menghasilkan s_1^2 = 24,7 sedangkan cara kedua dilakukan 13 kali yang menghasilkan s_2^2 = 37,2. Dengan taraf keberartian 10%, tentukan apakah kedua cara pengukuran mempunyai varians yang homogen? Anggaplah kedua sampel berasal dari populasi yang normal.

Penyelesaian :

Uji : H_0 : $_1$ = $_2$ vs H_1 : $_1 \neq {}_2$.

Pilih = 0,10 , sehingga daerah kritis : $F > f_{\frac{\ }{2}; v_1, v_2} = f_{0,05;\ 9,12} = 2,80$

dan $F < f_{1-\frac{\ }{2}; v_1, v_2} = \dfrac{1}{f_{\frac{\alpha}{2}; v_2, v_1}} = \dfrac{1}{3,07} = 0,328$ dengan $v_1 = n_1 - 1 = 9$ dan $v_2 = n_2 - 1 = 12$. Karena nilai f hitung = 24,7 / 37,2 = 0,664 tidak masuk ke dalam daerah kritis maka H_0 diterima, sehingga disimpulkan kedua varians homogen.

Uji Selisih Dua Rataan

Contoh :

Suatu percobaan dilakukan untuk membandingkan keausan karena gosokan dua bahan yang dilapisi. Dua belas potong bahan 1 diuji dengan memasukkan tiap potong bahan ke dalam mesin pengukur aus. Sepuluh potong bahan 2 diuji dengan cara yang sama dan diamati. Sampel bahan 1 memberikan rata-rata keausan (setelah disandi) sebanyak 85 satuan dengan simpangan baku 4. Sedang bahan 2 rata-ratanya 81 dan simpangan baku 5. Uji hipotesis bahwa kedua jenis bahan memberikan rata-rata keausan yang sama pada taraf keberartian 0,10. Anggap kedua populasi hampir normal dengan variansi sama.

Penyelesaian :

Uji : H_0 : $_1 = {}_2$ atau $_1 - {}_2 = 0$

vs H_1 : $_1 \neq {}_2$ atau $_1 - {}_2 \neq 0$.

= 0,10 , daerah kritis : T < -1,725 dan T > 1,725 karena dk = 20.

Nilai t hitung : $t = \dfrac{\bar{x}_1 - \bar{x}_2 - d_0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = \dfrac{(85 - 81) - 0}{4{,}478\sqrt{\dfrac{1}{12} + \dfrac{1}{10}}} = 2{,}07$

dengan simpangan baku gabungan sampel : $s_p = \sqrt{\dfrac{11.16 + 9.25}{12 + 10 - 2}} = 4{,}478$.

Kesimpulan : tolak H_0 dan simpulkan bahwa kedua jenis bahan tidak menunjukkan keausan yang sama karena gosokan.

Uji Kesamaan Lebih Dari Dua Varians (Uji Bartlett)

Misalkan k sampel acak diambil masing-masing dari k populasi yang dianggap saling bebas dan berdistribusi normal dengan rataan $_1$, $_2$, ..., $_k$ dan variansi σ_1^2, σ_2^2,...,σ_k^2 . Akan diuji :

H_0 :
$\sigma_1^2 \quad \sigma_2^2 \quad \ldots \quad \sigma_k^2$ vs H_1 : tidak semua variansi sama.

Statistik uji : b = 2,3026 $\dfrac{q}{h}$ dengan $q \ (N \ k) \log s^2_p (n_i \quad 1) \log s_i^2$; $\scriptstyle k \atop i\,1$

$$s^2_p \quad \frac{\overset{k}{\underset{i\,1}{}} (n_i \quad 1) s_i^2}{N \quad k}$$ dan $\quad \frac{1}{3(k \ 1)} \ \overset{k}{\underset{i\,1}{}} \frac{1}{n_i \ 1} \quad \frac{1}{N \ k}$; b merupakan peubah acak yang berdistribusi khi-kuadrat dengan dk=k-1.

Gunakan uji Bartlett untuk menguji kesamaan variansi ketiga populasi sampel ciptaan berikut:

	Sampel			
	A	B	C	
	4	5	8	
	7	1	6	
	6	3	8	
	6	5	9	
		3	5	
		4		
Jumlah	23	21	36	40

Penyelesaian :

Akan diuji : H_0 : $\sigma_1^2 \quad \sigma^2_2 \quad \sigma^2_3$ vs H_1 : tidak semua variansi sama.

Pilih = 0,05 , sehingga daerah kritisnya : B > 5,991 karena dk = k-1 = 2.

Tunjukkan bahwa : $s_1^2 = 1{,}583$, $s_2^2 = 2{,}300$, $s_3^2 = 2{,}700$ sehingga $s^2_p =$ 2,254 ; q = 0,1034 dan h = 1,1167. Jadi b = = 0,213.

Kesimpulan : terima H_0 dan simpulkan bahwa variansi ketiga populasi sama.

Uji Kesamaan Lebih Dari Dua Rataan

Sampel acak ukuran n diambil masing-masing dari k populasi yang dianggap saling bebas dan berdistribusi normal dengan rataan $_1$, $_2$, ..., $_k$ dan variansi ^{2} yang sama. Akan diuji :

H_0 : $_1$ = $_2$ = ... = $_k$

H_1 : paling sedikit dua di antara rataan tidak sama.

Tiap pengamatan dapat ditulis dalam bentuk : y_{ij} $_i$ $_{ij}$, dengan $_{ij}$ (galat acak) menyatakan penyimpangan pengamatan ke j sampel ke i dari rataan perlakuan padanannya.

	Perlakuan				
	1	2	...	k	
	y_{11}	y_{21}	…	y_{k1}	
	y_{12}	y_{22}	…	y_{k2}	
	y_{1n}	y_{2n}	…	y_{kn}	
Jumlah	$T_{1.}$	$T_{2.}$	…	$T_{k.}$	$T_{..}$
Rataan	$\bar{y}_{1.}$	$\bar{y}_{2.}$		$\bar{y}_{k.}$	$\bar{y}_{..}$

$$\text{Hitung : } JKT = \sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}^2 - \frac{T^2}{nk} \; ; \; JKA = \frac{\sum_{i=1}^{n} T_{i.}^2}{n} - \frac{T^2}{nk} \; ; JKG = JKT - JKA.$$

Kemudian buatlah tabel Analisis Variansi Ekaarah berikut :

Sumber Variasi	Jumlah kuadrat	Derajat kebebasan	Rataan kuadrat	f hitung
Perlakuan	*JKA*	*k - 1*	$s_1^2 = \frac{JKA}{k-1}$	s_1^2 / s^2
Galat	*JKG*	*k (n - 1)*	$s^2 = \frac{JKG}{n(k-1)}$	
Jumlah	*JKT*	*n k - 1*		

Jika H_0 benar, rasio $f = \frac{s_1^2}{s^2}$ merupakan peubah acak F yang berdistribusi F dengan derajat kebebasan k-1 dan k(n-1). Hipotesis nol ditolak pada taraf keberartian jika f hitung > $f_{;\ k\ 1, k(n\ 1)}$.

Perhatikan contoh berikut:

Misalkan seorang insinyur ingin menyelidiki bagaimana rataan pengisapan uap air dalam beton berubah antara lima adukan beton yang berbeda. Bahan dibiarkan kena uap selama 48 jam. Dari tiap adukan diambil 6 contoh untuk diuji, sehingga seluruhnya diperlukan 30 contoh. Data selengkapnya disajikan pada tabel berikut:

	Adukan (berat%)					
	1	**2**	**3**	**4**	**5**	
	551	595	639	417	563	
	457	580	615	449	631	
	450	508	511	517	522	
	731	583	573	438	613	
	499	633	648	415	656	
	632	517	677	555	679	
Jumlah	3320	3416	3663	2791	3664	16854
Rataan	553,33	569,33	610,50	465,17	610,67	561,80

Penyelesaian :

Akan diuji : H_0 : ${}_1 = {}_2 = \ldots = {}_6$

lawan H_1 : paling sedikit dua di antara rataan adukan tidak sama.

Pilih = 0,05 , sehingga daerah kritisnya : F > 2,26 dengan v_1=4 dan v_2=25.

Hitung jumlah kolom dan rataan masing-masing adukan, seperti pada tabel.

Total variasi dalam adukan dibagi menjadi dua bagian :

1. variasi antara adukan, yang mengukur variasi sistematik dan acak;
2. variasi dalam adukan, yang hanya mengukur variasi acak.

Perhitungan masalah analisis variansi diringkas dalam tabel berikut:

Tabel Analisis Variansi untuk Klasifikasi Ekaarah

Sumber Variasi	Jumlah kuadrat	Derajat kebebasan	Rataan Kuadrat	f Hitung
Perlakuan	85356	4	21339	4,30
Galat	124021	25	4961	
Jumlah	209377	29		

Karena f hitung = 4,30 > 2,26 tolak H_0 dan simpulkan bahwa kelima adukan tidak mempunyai rataan yang sama.

Dalam penelitian biasanya digunakan suatu model atau hubungan fungsional antara peubah. Dengan model kita berusaha memahami, menerangkan, mengendalikan dan kemudian memprediksikan kelakuan sistem yang diteliti. Model juga menolong peneliti dalam menentukan hubungan kausal. Rumusan hubungan tersebut yang dinyatakan dalam bentuk hipotesis dan diuji berdasarkan data yang dikumpulkan kemudian.

Misalkan X adalah peubah bebas (prediktor)dan Y peubah tak bebas yang bergantung pada Y (respons). Y (respon) tidak dikontrol dalam percobaan. Nilainya (y) bergantung pada satu atau lebih peubah bebas, misalnya (nilainya) x_1, $x_2,\ldots,x_k$, yang galat pengukurannya dapat diabaikan dan sesungguhnya sering peubah tersebut dikendalikan dalam percobaan. Jadi peubah bebas tersebut bukanlah peubah acak tapi k besaran yang ditentukan sebelumnya oleh peneliti dan tidak mempunyai sifat-sifat distribusi. Yang akan dibahas adalah regresi linear yang menyangkut hanya satu peubah saja.

Nyatakan sampel acak ukuran n dengan himpunan :$\{(x_i,y_i);i=1,2,\ldots,n\}$.
y_i merupakan nilai dari peubah acak Y_i selanjutnya akan ditulis Y|x "peubah acak yang berkaitan dengan nilai tetap x". Rataan Y|x berkaitan linear dengan x dalam bentuk persamaan : $Y \;|x\; x$
dengan α dan β adalah dua parameter yang akan ditaksir dari data sampel .

Bila semua rataan terletak pada satu garis lurus maka :

$$Y_i \qquad x_i \quad E_i$$

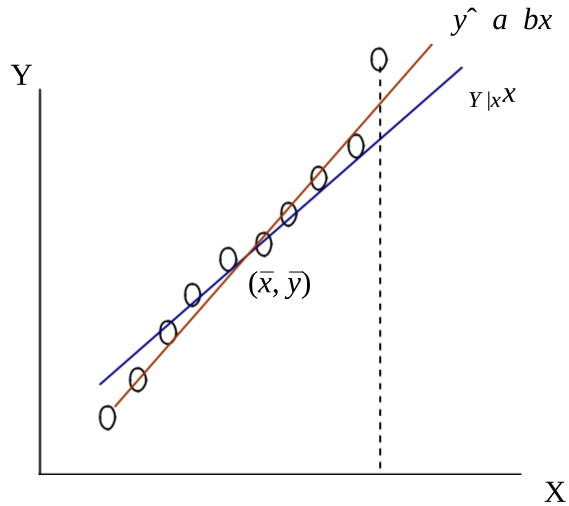
dengan asumsi : E_i galat yang bersifat acak dan rataannya = 0 dan variansinya konstan.

Setiap pengamatan (x_i,y_i) dalam sampel memenuhi :
dengan ε_i adalah nilai yang dicapai E_i bila Y_i berharga y_i
.

$$y_i \; x_i \quad i$$

Jika $\hat{\alpha} = a$ dan $\hat{\beta} = b$ maka setiap pengamatan dalam sampel memenuhi :

$$y_i = a + bx_i + e_i \; ; \; e_i \text{ disebut sisa}$$



Cara peminimuman untuk menaksir parameter dinamakan metode kuadrat terkecil (*least square method*), yaitu a dan b dicari sehingga :

$$JKG = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - a - bx_i\right)^2 \text{ minimum.}$$

Turunkan JKG terhadap a dan b maka diperoleh

$$\frac{\partial JKG}{\partial a} = -2 \sum_{i=1}^{n} \left(y_i - a - bx_i\right)$$

$$\frac{\partial JKG}{\partial b} = -2 \sum_{i=1}^{n} \left(y_i - a - bx_i\right) x_i$$

Samakan persamaan tsb dengan nol maka diperoleh *persamaan normal* :

$$na + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

Sehingga diperoleh :

$$b = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$a = \bar{y} - b\bar{x}$$

Di samping anggapan bahwa galat E_i dalam model $Y_i = \alpha + \beta x_i + E_i$ merupakan peubah acak dengan rataan nol, misalkan selanjutnya bahwa E_i berdistribusi normal dengan variansi sama σ^2 , dan $E_1, E_2, \ldots, E_n$ saling bebas dari suatu pengamatan ke pengamatan berikutnya dalam percobaan. Dengan asumsi kenormalan tersebut dapat dicari rataan dan variansi untuk penaksir α dan β.

Selang Kepercayaan dan Uji Keberartian

Uji $H_0 : \beta = 0$ (model tak linear) lawan $H_1 : \beta \neq 0$ (model linear) dan pilih taraf keberartian α=5%. Statistik ujinya : $T = \dfrac{B}{S/\sqrt{J_{xx}}} \sim t_{n-2}$

tolak jika $T < -t_{\alpha/2}$ atau $T > t_{\alpha/2}$.

Juga harus diuji : $H_0 : \alpha = 0$ (garis melalui titik asal) lawan $H_1 : \alpha \neq 0$ (garis tidak melalui titik asal) dan pilih taraf keberartian α=5%

Statistik ujinya : $T = \dfrac{A}{S\sqrt{\sum_{i=1}^{n} x_i^2 / nJ_{xx}}} \sim t_{n-2}$

tolak jika $T < -t_{\alpha/2}$ atau $T > t_{\alpha/2}$.

Pendekatan Analisis Variansi

Pengujian keberartian model selain dengan uji t juga dapat menggunakan uji F atau pendekatan analisis variansi dengan tabel berikut :

Sumber Variasi	JK(Jumlah Kuadrat)	dk(derajat kebebasan)	RK(Rataan Kuadrat)	f hitung
Regresi	JKR = b Jxy	1	RKR = JKR/1	JKR/s^2
Sisa	JKS (JKG) = JKT - JKR	n-2	RKS s^2=JKS/(n-2)	
Total	JKT = Jyy	n-1		

di mana : $J_{xx} = \sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$; $J_{yy} = \sum_{i=1}^{n} y_i^2 - \dfrac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$

dan $J_{xx} = \sum_{i=1}^{n} x_i y_i - \dfrac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$.

Tolak H_0 jika $F > F_{1,n-2}$ atau tolak H_0 jika f hitung > f tabel (dk_1=1,dk_2=n-2).

Uji t dan F yang digunakan bersifat kekar, yang berarti bahwa anggapan kenormalan dan kesamaan variansi tidak perlu dipenuhi dengan ketat tapi cukup agak kasar.

Selanjutnya harus dilakukan pemeriksaan sisa, yaitu :

1. Apakah sisa telah berpola acak ;
2. Apakah anggapan kenormalan tidak dilanggar ;
3. Apakah variansi dapat dianggap tidak berubah ;
4. Apakah ada data yang tidak mengikuti pola umum (pencilan) ;
5. Apakah peubah yang masuk dalam model mungkin bukan berbentuk Linear ;
6. Apakah peubah yang berpengaruh telah masuk ke dalam model.