INTERVAL KONFIDENSI SIMULTAN
PADA REGRESI GANDA

SKRIPSI



Oleh :

ERMA WINDU SUSANTI

No. Mahasiswa : 96 411 006

NIRM : 961051013206120012

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS ISLAM INDONESIA

YOGYAKARTA

2000

HALAMAN PENGESAHAN

Pengesahan Skripsi

Berjudul

INTERVAL KONFIDENSI SIMULTAN PADA REGRESI GANDA

Oleh :

ERMA WINDU SUSANTI

No. Mahasiswa : 96 411 006
NIRM : 961051013206120012

Telah disetujui dan disahkan pada tanggal: .....

Pembimbing 1	Pembimbing 2
(Dra. Sri Haryatmi M.Sc.)	(Akhmad Fauzy, M.Si.)

Mengetahui
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

(Dr. Chairil Anwar)

HALAMAN PENGESAHAN

Pengesahan Skripsi

Berjudul

INTERVAL KONFIDENSI SIMULTAN PADA REGRESI GANDA

Oleh :

ERMA WINDU SUSANTI

No. Mahasiswa : 96 411 006
NIRM : 961051013206120012

Telah dipertahankan dihadapan Panitia Penguji Skripsi
Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

Tanggal : 25 Nop 2000

Penguji	Tanda Tangan
1. Dra. Sri Haryatmi, M.Sc.	
2. Akhmad Fauzy, M.Si.	
3. Supriyono, M.Sc.	
4. Jaka Nugraha, M.Si.	

Mengetahui

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Islam Indonesia

(Dr. Chairil Anwar)

HALAMAN PERSEMBAHAN

Dengan keikhlasan dan ketulusan hati, karya tulis ini penulis persembahkan kepada :

- Bapak dan Mama tercinta yang telah memberi dukungan, doa serta kasih sayangnya.
- Mbah kakung dan putri, yang selalu mendoakan tanpa kenal lelah.
- Keluarga besar Bapak H. Kardjono yang telah mendoakan dan mendukungnya.
- Gurun Bayu Saputro, yang telah memberikan nasehat, kasih sayang dan bantuannya berupa tenaga maupun doa yang selalu dipanjatkannya.
- Adikku tercinta Susan dan Anto yang selalu mendukung dan mendoakannya serta Ika yang telah banyak membantu dengan nasehatnya.
- Semua saudaraku tercinta, kel. Lik Nono, kel. Lik Katam, Lik Dawud, Lik Ati, Lik Djumadi, Lik Enik, Lik Yani, Lik Utik (yang telah membelikan printer) yang telah memberi motifasi dan kasih sayangnya.
- Dosen-dosen Fakultas MIPA UII yang telah memberi ilmu serta petunjuknya.
- Temen seangkatanku Yuli, Retno, Wuryani, Sofwan, Budi, Etek, Agung Hidayat, mas Andang dan yang lainnya, yang sudi membantu dan meminjamkan bukunya.
- Adik angkatanku Puji Sulistiono, Akhmadi, Sulistiana Indah, Shofiani, Neneng yang telah banyak membantu.
- Sahabatku Yeyen, Evy, Ninik dan Dwi yang dengan setia selalu memberi nasehat serta hiburannya saat kesulitan datang.
- Semua teman kostku Anis, Inan, Aris, Kiki, Lian, Iis, Endah, Indah, Wiwid, Santi, Laila, Ery, Lifah, Risma, mba Lisa, Ratna, Rini, Candra, Ani, Rova, Erlin, Dewi dll yang telah banyak membantu dengan nasehat dan buku-buku yang dipinjamkannya.

KATA PENGANTAR

Bissmilahirrahmanirrahiim

Assalamu'alaikum Wr.Wb

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya kepada penulis, sehingga dapat menyelesaikan Laporan Tugas Akhir.

Laporan ini merupakan salah satu wujud nyata dari apa yang telah penulis dapatkan selama mengadakan penelitian literatur yang berjudul Interval Konfidensi Simultan pada Regresi Ganda.

Penulis sadar sepenuhnya bahwa dalam melaksanakan penelitian ini, penulis tidak bekerja sendiri, tetapi mendapat bantuan dari berbagai pihak. Untuk itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

- Bapak Dr. Chairil Anwar selaku Dekan Fakultas MIPA Universitas Islam Indonesia.
- Ibu Dra. Sri Haryatmi M.Sc. selaku Dosen Pembimbing pertama.
- Bapak Ahmad Fauzy M.Si selaku Dosen Pembimbing kedua.
- Bapak Joko Nugroho M.Si, selaku Ketua Jurusan Statistik Fakultas MIPA Universitas Islam Indonesia.
- Semua Dosen Jurusan Statistika Fakultas MIPA Universitas Islam Indonesia.
- Teman-teman yang tidak dapat penulis sebutkan satu persatu, khususnya dari Jurusan Statistika Fakultas MIPA Universitas Islam Indonesia.

Semoga Allah SWT menganugerahkan balasan yang lebih baik kepada semua pihak yang telah membantu penyelesaianTugas Akhir ini.

Akhir kata penulis mohon maaf apabila ada kekurangan maupun kesalahan dalam penyusunan laporan Tugas Akhir ini. Penulis berharap agar hasil dari Tugas Akhir ini dapat bermanfaat bagi ilmu pengetahuan khususnya dalam bidang statistika.

Wassalamu'alaikum Wr.Wb

Yogyakarta, November 2000

Penulis

DAFTAR ISI

HALAMAN PENGESAHAN i
HALAMAN PERSEMBAHAN iii
KATA PENGANTAR iv
DAFTAR ISI vi
INTISARI viii

BAB I : PENDAHULUAN
1.1. Latar Belakang 1
1.2. Perumusan Masalah 4
1.3. Batasan Masalah 4
1.4. Tujuan Penelitian 4
1.5. Manfaat penelitian 5

BAB II: LANDASAN TEORI
2.1. Tipe Data Statistik 6
2.1.1. Data Kualitatif (Data Kategorik) 6
2.1.2. Data kuantitatif (Data Kontinu) 7
2.2. Model Regresi Linier Secara Umum 8
2.2.1. Estimasi Koefisien Regresi 9
2.3. Interval Konfidensi 11
2.3.1. Interval Konfidensi Koefisien Regresi 11
2.3.2. Interval Konfidensi Simultan dari Koefisien Regresi 20
2.4. Analisis Kovariansi 23

BAB III: PEMBAHASAN
3.1. Metode untuk Menentukan Interval Konfidensi Simultan 26
3.1.1. Metode Bonferroni 26
3.1.2. Metode Scheffe 29
3.1.3. Metode Hybrid 31

3.2. Perbandingan dari Metode Scheffe dan Hybrid dalam Menentukan Interval Konfidensi Simultan 32
BAB IV: APLIKASI
4.1. Menduga Persamaan Regresi 40
4.2. Menentukan Interval Konfidensi Simultan 42
BAB V : KESIMPULAN DAN SARAN 46
DAFTAR PUSTAKA 47
LAMPIRAN 49

INTISARI

Metode Hybrid adalah salah satu metode yang digunakan untuk menentukan interval konfidensi pada koefisien kombinasi linier dalam regresi ganda, dengan fokus interval pada nilai tetap. Metode Hybrid adalah kombinasi dari pendekatan Bonferroni dan Scheffe. Pendekatan dari metode Hybrid ini ketika kedua prediktornya kontinu dan kategorik. Dari interval ini biasanya akan diperoleh interval yang lebih sempit dari metode Scheffe, walaupun tidak selalu.

Metode Hybrid ini belum sering dipakai untuk beberapa masalah yang berkenaan dengan penentuan interval konfidensi simultan. Kelebihan dari metode ini adalah dapat digunakan pada data kategorik dan kontinu.

Kata-kata kunci (*Key Words*) :
Metode Bonferroni; Regresi Ganda; Metode Scheffe; Metode Hybrid.

UNIVERSITAS ISLAM INDONESIA
PERPUSTAKAAN

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dalam ilmu statistika banyak metode yang dapat diterapkan dalam berbagai segi kehidupan. Selama ini banyak orang beranggapan kalau statistika hanya berupa tabel atau grafik dari suatu variabel semata, namun sebenarnya itu adalah sebagian kecil dari ilmu statistika. Analisis statistika itu sendiri mempunyai bagian yang sangat banyak seperti analisis regresi, analisis multivariabel, analisis kuantitatif, analisis kualitatif, analisis variansi dan masih banyak lagi.

Dalam penelitian ini akan dibahas masalah analisis regresi khususnya menentukan interval konfidensi simultan (keseluruhan) pada analisis regresi ganda. Pandang sebuah model regresi linier secara umum :

$$Y = X\beta + \varepsilon \qquad \ldots (1.1)$$

dimana

- Y adalah nilai pengamatan respon untuk setiap pengamatan dalam bentuk matrik dituliskan $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$

- X adalah vektor berdimensi p dari nilai prediktor yang diketahui, dalam bentuk matrik dituliskan $X = \begin{bmatrix} x_{10} & x_{11} & \cdots & x_{1,p-1} \\ x_{20} & x_{21} & \cdots & x_{2,p-1} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n0} & x_{n1} & \cdots & x_{n,p-1} \end{bmatrix}$

- β adalah vektor berdimensi p dari koefisien-koefisien yang belum diketahui, dalam bentuk matrik ditulis $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-i} \end{bmatrix}$

- $\varepsilon_i \approx N(0,\sigma^2)$ yang berarti suatu peubah acak normal, dengan nilai tengah nol dan variansi σ^2 yang belum diketahui, dalam bentuk matrik dituliskan $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Banyak dari ilmuwan atau khususnya statistisi yang ingin mengetahui batas suatu daerah konfidensi simultan untuk semua permukaan regresi, yaitu dengan mengobservasi n variabel independen dari nilai respon $y_1, y_2, \ldots . y_n$ dan vektor prediktor $x_1, x_2, \ldots . x_n$ yang bersesuaian. Selanjutnya dihitung batas-batas konfidensi mean $x'\hat{\beta}$ yang belum diketahui, sehingga bats konfidensi secara simultan memenuhi semua nilai-nilai x.

Dengan adanya perkembangan ilmu statistika telah banyak metode yang dapat digunakan untuk menghitung interval konfidensi simultan, untuk mencocokkan nilai maupun juga untuk kombinasi linier lainnya dari koefisien regresi. Dua prosedur yang

biasa digunakan dalam menentukan interval konfidensi simultan yaitu dengan menggunakan metode Scheffe dan Bonferroni. Metode ini telah diterangkan dalam banyak buku seperti Seber (1977), buku Simultaneous Inference oleh Miller (1981) dan buku Raymond H. Myers (1989).

Metode Scheffe menerangkan tentang konfidensi simultan untuk semua kombinasi linier yang mungkin dari koefisien. Metode ini biasanya digunakan untuk semua masalah regresi, akan tetapi penggunaannya akan lebih bermanfaat ketika prediktornya kontinu. Sedangkan metode Bonferroni menerangkan tentang konfidensi simultan bagi himpunan berhingga pada kombinasi-kombinasi linier. Metode Bonferroni biasanya kurang dapat digunakan pada saat prediktornya tanpa pembatasan. Metode Bonferroni ini akan dapat digunakan dan akan lebih bermanfaat apabila banyaknya nilai prediktor adalah suatu himpunan berhingga atau terbatas. Metode Bonferroni ini biasanya sering menghasilkan interval yang lebih ringkas dari metode Scheffe.

Dalam penelitian ini akan diperlihatkan metode baru, yaitu metode Hybrid. Metode ini secara teori akan menghasilkan suatu perhitungan interval yang lebih ringkas dari metode-metode sebelumnya. Metode ini akan bermanfaat ketika prediktornya kontinu maupun kategorik. Metode Hybrid merupakan metode yang secara tidak langsung memasukkan metode Bonferroni dan Scheffe sebagai kasus khusus, atau lebih mudah dicerna kalau metode Hybrid adalah gabungan dari dua metode Bonferroni dan Scheffe.

1.2. Perumusan Masalah

Permasalahan yang muncul dari penelitian ini adalah :

a. Membahas metode baru, yaitu metode Hybrid dalam menentukan interval konfidensi simultan.

b. Membicarakan kelemahan dan keunggulan dari Scheffe dan Hybrid dalam menentukan interval konfidensi simultan metode.

1.3. Batasan Masalah

a. Metode yang digunakan dalam menentukan interval konfidensi keseluruhan dengan menggunakan metode Scheffe, Bonferroni dan Hybrid.

b. Data yang diambil sebagai permasalahan bukan suatu data yang baku, artinya data tersebut hanya untuk mengaplikasikan teori agar permasalahan serta penjelasannya akan lebih mudah dimengerti.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

a. Memperkenalkan metode Hybrid dalam membentuk interval konfidensi simultan untuk setiap parameter pada koefisien regresi ganda.

b. Membandingkan metode Hybrid dengan metode Scheffe.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah :

a. Memberikan solusi bahwa dalam menentukan interval konfidensi keseluruhan, khususnya pada regresi ganda tidak hanya menggunakan metode Scheffe dan Bonferroni, akan tetapi ada metode lain yang dapat digunakan dalam berbagai data baik data kategorik maupun data kontinu, yaitu metode Hybrid.

b. Memperkenalkan metode Hybrid dalam menyelesaikan masalah interval konfidensi keseluruhan pada regresi ganda, khususnya untuk tipe data kategorik dan kontinu.



BAB II

LANDASAN TEORI

2.1 Tipe Data Statistik

Statistik dalam prakteknya tidak bisa dilepaskan dari data yang berupa angka, baik statistik deskriptif maupun statistik inferensi yang melakukan analisis terhadap data. Dalam data statistik berdasarkan tingkat pengukurannya dapat dibedakan dalam empat jenis.

2.1.1 Data Kualitatif (Data Kategorik)

Data kualitatif adalah data yang bukan berupa angka, yang mempunyai ciri tidak bisa dilakukan operasi matematika dan hanya sekedar untuk identitas saja.
Data kualitatif dibedakan menjadi dua:

a. Data Nominal

Suatu angka hasil pengukuran disebut berada pada tingkat nominal atau berskala nominal apabila angka tersebut berfungsi untuk identifikasi, yaitu membedakan antara satu subjek dengan subjek yang lainnya. Perbedaan angka hanya menunjukkan adanya objek atau subjek yang terpisah dan tidak sama. Sebagai contoh adalah nomer punggung pemain sepak bola. Nomor punggung ini digunakan hanyalah untuk identifikasi pemain, sama sekali bukan untuk menunjukkan bahwa pemain yang satu lebih berkualitas dari pemain yang lainnya.

Selain untuk identifikasi, angka dapat dikatakan berada pada level nominal apabila digunakan untuk klasifikasi atau kategorisasi. Sebagai contoh adalah kategori jenis kelamin dimana pria diberi angka 1 dan wanita diberi angka 0.

b. Data Ordinal

Suatu hasil pengukuran termasuk data ordinal apabila angkanya berfungsi sebagai petunjuk adanya penjenjangan kualitatif. Perbedaan angka yang dimiliki oleh objek yang satu dari yang lain tidak menunjukkan adanya perbedaan kuantitatif melainkan perbedaan jenjang kualitatif saja. Bila terdapat jenjang 1,2,dan 3, maka kita tidak dapat mengatakan bahwa jarak 3-2 adalah sama dengan jarak 2-1. Yang dapat dikatakan hanyalah bahwa 3 > 2, 2 > 1 dan 3 > 1. Jarak jenjang antara angka yang berurutan tidak selalu sama. Lebih lanjut angka 0 dalam skala ini tidak memiliki arti mutlak.

2.1.2 Data Kuantitatif (Data Kontinu)

Data kuantitatif adalah data berupa angka dalam arti sebenarnya, jadi bisa dilakukan operasi matematika. Data kuantitatif dibedakan menjadi dua:

a. Data Interval

Hasil pengukuran berskala interval adalah hasil pengukuran ordinal yang memiliki jarak antar jenjang yang tetap. Jadi dalam deretan angka 2,3,4,5,6,7,.. kita tentu saja dapat mengatakan bahwa jarak 6-4 adalah sama dengan jarak 5-3 dan sama pula dengan jarak antara 7-5. Sebagaimana sebagai ukuran ordinal, skala inipun tidak memiliki harga 0 mutlak, sehingga kita tidak dapat mengatakan bahwa 6 adalah dua

kali 3. Sebagai contoh data yang berada pada tingkat interval adalah hasil pengukuran suhu pada termometer.

b. Data Rasio

Pada level pengukuran yang tertinggi adalah hasil pengukuran yang berskala rasio. Data rasio pada dasarnya adalah level interval yang memiliki harga 0 mutlak, artinya harga 0 pada skala ini memang menunjukkan bahwa atribut yang diukur sama sekali tidak ada pada objek yang bersangkutan. Ukuran-ukuran berat, panjang, banyaknya benda dalam suatu ruangan dan sebagainya, semua merupakan contoh data berlevel rasio.

2.2 Model Regresi Linier Secara Umum

Dalam banyak penelitian kita sering ingin menyelidiki bagaimana perubahan-perubahan pada satu peubah mempengaruhi peubah lainnya, dengan bentuk umum regresi liniernya adalah :

$$Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_{p-1} x_{ip-1} + \varepsilon_i \qquad \ldots (2.1)$$

$$i = 1,2,\ldots,n$$

dalam lambang matrik persamaan 2.1 dapat dituliskan menjadi:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} x_{10} & x_{11} \cdots x_{1p-1} \\ x_{20} & x_{21} \cdots x_{2p-1} \\ \vdots & \vdots \cdots \vdots \\ x_{n0} & x_{n1} \cdots x_{np-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan demikian persamaan 2.1 dapat ditulis kembali menjadi $Y = X\beta + \varepsilon$ yang merupakam bentuk umum persamaan regresi dalam bentuk matrik, dengan :

- Y merupakan vektor respon, berukuran n x 1,
- X menyatakan matrik peubah bebas, berukuran n x p,
- β merupakan vektor parameter, berukuran p x 1,
- ε merupakan vektor galat, berukuran n x 1.

Agar penelitian ini dianggap tidak menyalahi aturan dalam persamaan regresi ganda, perlu memperhatikan spesifikasi mengenai error dan variabel X secara jelas. Untuk itu regresi ganda ini diasumsikan sebagai berikut:

1. linieritas pada koefisien artinya semua koefisiennya linier dan ini sangat berkaitan langsung dengan pengembangan uji F dan uji t,
2. nilai sisaan atau error independen, ini juga berkaitan dengan penggunaan uji F dan uji t,
3. variansinya konstan, dimana model regresi dianggap mempunyai distribusi kesalahan menyebar normal disekitar nilai tengah nol dan mempunyai variansi V_e,
4. nilai sisaan atau error berdistribusi normal,
5. variabel bebas bersifat tetap, yang berarti diukur tanpa kesalahan.

2.2.1 Estimasi Koefisien Regresi

Untuk mengetahui tafsiran dari koefisien regresi kita bisa menggunakan metode kuadrat terkecil (Ordinary Least Square).

Dengan persamaan regresi $Y = X\beta + \varepsilon$, kita ingin mencari $\hat{\beta}$ dengan meminimumkan jumlah kuadrat sesatan.

Lemma 2.1

Dengan menggunakan metode kuadrat terkecil akan diperoleh estimasi untuk β *adalah* $\hat{\beta} = (X'X)^{-1}X'Y$

Bukti :

Jumlah kuadrat kesalahan dengan menggunakan metode matrik adalah :

$$\begin{aligned} J = \varepsilon'\varepsilon &= (Y - X\beta)'(Y - X\beta) \\ &= Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \end{aligned}$$

dengan matrik $\beta'X'Y$ adalah matrik (1 x 1) atau matrik skalar. Transpose dari matrik tersebut adalah $(\beta'X'Y)' = Y'X\beta$. Untuk mencari estimasi kuadrat terkecil maka dilakukan penggantian parameter β dengan penaksir $\hat{\beta}$, dan diturunkan terhadap $\hat{\beta}$:

$$\left.\frac{\partial J}{\partial \beta}\right|_{\hat{\beta}} = -2X'Y + 2X'X\hat{\beta} = 0$$

$$\Leftrightarrow X'X\hat{\beta} = X'Y$$

dimana

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ \vdots & \vdots & \cdots & \vdots \\ x_{ip} & x_{2p} & \cdots & x_{np-1} \end{bmatrix}, \quad Y' = (y_1, y_2, \ldots, y_n), \quad \hat{\beta}' = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_{p-1})$$

maka persamaan normal $(X'X)\hat{\beta} = X'Y$. Dengan asumsi bahwa $(X'X)$ adalah matrik yang tidak singular maka matrik $(X'X)$ mempunyai invers, dengan

demikian persamaan diatas mempunyai penyelesaian tunggal $\hat{\beta} = (X'X)^{-1}X'Y$ ■

2.3. Interval Konfidensi

Nilai penaksir yang hanya memuat satu nilai saja disebut sebagai penaksir titik. Penaksir ini sifatnya eksak, tetapi penaksiran yang menyangkut parameter tidak dapat diketahui nilainya, jadi hampir mustahil dapat diperoleh nilai penaksir yang sama dengan nilai parameternya. Karena itu diperlukan adanya penaksiran dengan melihat probabilitas kesalahan itu akan dapat terjadi. Untuk menyatakan hal itu dibuatlah penaksiran interval, artinya nilai penaksir yang dipakai tidak dinyatakan dalam satu angka saja tetapi dalam angka interval. Karena besarnya parameter yang ditaksir belum diketahui, maka bukan kepastian yang dibuat dalam penaksiran itu, tetapi probabilitas.

Interval konfidensi yang biasa dilakukan adalah dengan menentukan interval konfidensi pada tiap parameter regresi atau sering dikenal dengan menentukan selang kepercayaan tiap parameter satu persatu. Selain itu ada juga interval konfidensi simultan atau keseluruhan yang sering dikenal dengan menentukan daerah kepercayaan.

2.3.1 Interval Konfidensi Koefisien Regresi

Interval Konfidensi biasanya digunakan sebagai tafsiran suatu parameter dan dapat pula dipandang sebagai pengujian hipotesis, yaitu apakah suatu parameter sama dengan suatu nilai tertentu.

Dengan persamaan regresi $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ didapat :

a. Interval Konfidensi untuk β_1

Lemma 2.2

$$E(\hat{\beta}) = \beta$$

$$var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Bukti :

Dalam bentuk matrik didapat,

$$E(\hat{\beta}) = E\left((X'X)^{-1}X'Y\right) = (X'X)^{-1}X'E(Y)$$
$$= (X'X)^{-1}X'X\beta = \beta$$

$$\text{var}(\hat{\beta}) = V\left[(X'X)^{-1}X'Y\right] = (X'X)^{-1}X'V(Y)X(X'X)^{-1}$$
$$\Leftrightarrow \sigma^2(X'X)^{-1}X'IX(X'X)^{-1} = \sigma^2(X'X)^{-1}$$

(Seber;48,1977) ■

Variansi koefisien $\hat{\beta}_1$ dapat ditulis dalam bentuk $\text{var}(\hat{\beta}_1) = \sigma^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2$ dan σ^2 biasanya tidak diketahui maka harus ditaksir dari data, yaitu

$$s^2 = \frac{Y'Y - Y'X(X'X)^{-1}X'Y}{n-p}$$

Lemma 2.3

$E(s^2) = \sigma^2$.

Bukti :

$$E(s^2)(n-p) = \sigma^2 \text{tr}\left[I - X(X'X)^{-1}X'\right] + (X\beta)'\left[I - X(X'X)^{-1}X'\right](X\beta)$$

$$= \sigma^2(n-p) + \beta'X'X\beta - \beta'X'X(X'X)^{-1}X'X\beta$$

$$= \sigma^2(n-p)$$

dimana $trX(X'X)^{-1}X' = tr(X'X)(X'X)^{-1} = p$

jadi dapat dituliskan :

$$E(s^2) = \sigma^2$$ *(Raymond H. Myers;462,1990)* ■

Untuk membentuk uji statistik mengenai β_1 ,maka dibentuk nisbah.

Teorema 2.4

$$(n-2)s^2/\sigma^2 \sim \chi^2_{n-2}$$

Bukti :

$$\frac{1}{\sigma^2}\sum_{i=2}^{n}(x_i-\mu)^2 = \sum_{i=2}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2 + \frac{n-1(\bar{x}-\mu)^2}{\sigma^2}$$

dalam sampel random menggunakan persamaan:

$$\frac{1}{\sigma^2}\sum_{i=2}^{n}(x_i-\mu)^2 = \frac{(n-2)s^2}{\sigma^2} + \frac{n-1(\bar{x}-\mu)^2}{\sigma^2}$$

misalkan :

$$V_1 = \frac{1}{\sigma^2}\sum_{i=2}^{n}(x_i-\mu)^2 \sim \chi^2_{(n)} \qquad V_2 = \frac{(n-2)s^2}{\sigma^2} \qquad V_3 = \frac{n-1(\bar{x}-\mu)^2}{\sigma^2} \sim \chi^2_{(2)}$$

dimana V_2 dan V_3 independan.

maka:

$$M_{V_1}(t) = M_{V_2}(t)M_{V_3}(t)$$

$$M_{V_2}(t) = \frac{M_{V_1}(t)}{M_{V_3}(t)} = \frac{(1-2t)^{-n/2}}{(1-2t)^{-2/2}} = (1-2t)^{-(n-2)/2}$$

sehingga $V_2 = \left[(n-2)s^2\right]/\sigma^2 \sim \chi^2_{(n-2)}$

(Lee. J. Bain, 1992) ■

Teorema 2.5

Jika $Z \sim N(0,1)$ dan $V \sim \chi^2_{(v)}$ dengan Z dan V independen, maka distribusi $\frac{Z}{\sqrt{V/v}}$ *adalah berdistribusi t dengan derajat bebas v.*

Bukti :

Gabungan dari densitas Z dan V adalah

$$f_{Z,V}(z,v) = \frac{v^{v/2-1}e^{-v/2}e^{-z^2/2}}{\sqrt{2\pi}\,\Gamma(v/2)2^{v/2}} \qquad 0 < v < \infty, -\infty < z < \infty$$

dengan mempertimbangkan transformasi $T = Z/\sqrt{V/v}$, $W = V$, yang mempunyai invers $v = w$, $z = t\sqrt{w/v}$. Jakobiannya adalah $J = \sqrt{w/v}$

dan $f_{T,W}(t,w) = \dfrac{(w/v)^{1/2}w^{v/2-1}e^{-w/2}e^{-t^2w/2v}}{\sqrt{2}\,\Gamma(v/2)2^{v/2}} \qquad -\infty < t < \infty,\ 0 < w < \infty$ i

probabilitas densitas fungsi (pdf) marjinalnya:

$$f(t;v) = \int_0^\infty f_{T,W}(t,w)dw = \int_0^\infty \frac{(w/v)^{1/2}w^{v/2-1}e^{-w/2}e^{-t^2w/2v}}{\sqrt{2\pi}\Gamma(v/2)2^{v/2}}dw$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi v}\,\Gamma(v/2)2^{v/2}}.w^{\frac{(v+1)}{2}-1}.\exp\left[\frac{-w}{2}(1+\frac{t^2}{v})\right]dw$$

misalkan : $z = \left[\dfrac{-w}{2}(1+\dfrac{t^2}{v})\right] = w(1+(t^2/v))/2$

$$f(t:v) = \int_0^{\infty} \frac{1}{\sqrt{2\pi v}\,\Gamma(v/2)2^{v/2}} \cdot \left(\frac{2z}{1+t^2/v}\right)^{\frac{(v+1)}{2}-1} .e^{-z}\left(\frac{2}{1+t^2/v}\right) dz$$

$$= \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v}} \left(1+\frac{t^2}{v}\right)^{-(v+1)/2}$$, yang merupakan pdf dari distribusi student- t. *(Hogg and Craig, 1978)* ■

Lemma 2.6

$\frac{\hat{\beta}_1 - \beta_1}{s(\hat{\beta}_1)}$ *berdistribusi t dengan derajat bebas n-2*

Bukti :

Tulis kembali persamaan $\frac{\hat{\beta}_1 - \beta_1}{s(\hat{\beta}_1)} = \frac{\frac{\hat{\beta}_1 - \beta_1}{\sigma(\hat{\beta}_1)}}{\frac{s(\hat{\beta}_1)}{\sigma(\hat{\beta}_1)}}$

dimana: pembilang $\frac{\hat{\beta}_1 - \beta_1}{\sigma(\hat{\beta}_1)} \sim N(0,1)$ namakan Z dan penyebut adalah:

$$\frac{s^2(\hat{\beta}_1)}{\sigma^2(\hat{\beta}_1)} = \frac{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n-2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}{\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n-2}{\sigma^2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sigma^2(n-2)} \sim \frac{\chi^2_{(n-2)}}{n-2}$$

sehingga : $\frac{\hat{\beta}_1 - \beta_1}{s(\hat{\beta}_1)} \sim \frac{Z}{\sqrt{\frac{\chi_{(n-2)}}{n-2}}}$, karena Z dan χ^2 saling bebas maka menurut

teorema 2.5 dapat disimpulkan bahwa $\frac{\hat{\beta}_1 - \beta_1}{s(\hat{\beta}_1)} \sim t_{(n-2)}$ ■

Nisbah dari β_1 adalah dengan mengurangi $\hat{\beta}_1$ dengan $E(\hat{\beta}_1)$ dan kemudian membagi hasilnya dengan simpangan baku dari $\hat{\beta}_1$, yaitu $s(\hat{\beta}_1)$. Di bawah asumsi normal, $\hat{\beta}_1$ akan berdistribusi normal sedangkan s^2 berdistribusi χ^2 dengan db = n-2 sehingga nisbah di atas berdistribusi – t dengan db = n-2 atau disingkat t(n-2). Dengan demikian bila nilai α diberikan maka kita dapat melihat tabel nilai t dengan db = n-2, sehingga interval konfidensi untuk β_1 dengan kepercayaan 1- α adalah

$$\hat{\beta}_1 - t(n-2;1-\alpha/2)s(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t(n-2;1-\alpha/2)s(\hat{\beta}_1)$$

b. Interval Konfidensi untuk β_0

Interval Konfidensi untuk β_0 dapat dikerjakan seperti untuk β_1. Sebenarnya pengujian β_0 tidaklah terlalu berarti karena tidak banyak yang dapat diambil keterangannya dari pengetahuan apakah β_0 nol atau bukan. Jika x=0, maka β_0 tidak punya tafsiran yang jelas. Jika $\beta_0 = 0$, maka model garis regresi melalui titik nol. Untuk mencari interval konfidensi dari β_0 adalah :

$$\hat{\beta}_0 - t(n-2;1-\alpha/2)s(\hat{\beta}_0) \le \beta_0 \le \hat{\beta}_0 + t(n-2;1-\alpha/2)s(\hat{\beta}_0)$$

Lemma 2.7

$$\text{var}(\hat{\beta}_0) = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

Bukti :

$$\text{var}(\hat{\beta}_0) = \text{var}(\bar{Y} - \hat{\beta}_1\bar{x}) = \text{var}(\bar{Y}) + \bar{x}^2\,\text{var}(\hat{\beta}_1)$$

$$= \frac{\sigma^2}{n} + \bar{x}^2\left(\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

$$= \sigma^2\left(\frac{1}{n} + \frac{\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

(Raymond H. Myers; 15, 1990) ■

c. Interval Konfidensi untuk $E(Y/x_0)$

Dengan asumsi berdistribusi normal untuk $x = x_0$, peubah acak $\hat{Y}$ berdistribusi normal akan dibentuk interval konfidensi untuk $\hat{Y}$.

Definisi 2.8

$$Var(X+Y) = kov(X+Y, X+Y)$$

$$= var(X) + var(Y) + 2kov(X,Y)$$

$$Kov(X,Y) = E(XY) - E(X)E(Y)$$

Lemma 2.9

$$Kov\,(\hat{\beta}_0,\hat{\beta}_1) = -\frac{\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

Bukti:

$$\begin{aligned}
\text{kov}(\hat{\beta}_0,\hat{\beta}_1) &= \text{kov}(\bar{y}-\hat{\beta}_1\bar{x},\hat{\beta}_1)\\
&= E\left[(\bar{y}-\hat{\beta}_1\bar{x})(\hat{\beta}_1)\right]-E(\bar{y}-\hat{\beta}_1\bar{x})E(\hat{\beta}_1)\\
&= E(\hat{\beta}_1\bar{y}-\hat{\beta}_1^2\bar{x})-E(\hat{\beta}_0)E(\hat{\beta}_1)\\
&= E(\hat{\beta}_1\bar{y})-E(\hat{\beta}_1^2\bar{x})-\beta_0\beta_1\\
&= E(\hat{\beta}_1\bar{y})-\bar{x}E(\hat{\beta}_1^2)-\beta_0\beta_1
\end{aligned}$$

dimana,

$$E(\hat{\beta}_1\bar{y}) = E((\hat{\beta}_1)(\hat{\beta}_0+\hat{\beta}_1\bar{x})) = \beta_0\beta_1+\beta_1^2\bar{x}$$

$$E(\hat{\beta}_1^2) = \text{var}(\hat{\beta}_1)+\left[E(\hat{\beta}_1)\right]^2 = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}+\beta_1^2$$

Jadi

$$\text{kov}(\hat{\beta}_0,\hat{\beta}_1) = \beta_0\beta_1+\beta_1^2\bar{x}-\bar{x}\frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}-\bar{x}\beta_1^2-\beta_0\beta_1$$

$$= -\bar{x}\frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

(LeeJ. Bain, 1992) ■

UNIVERSITAS ISLAM INDONESIA PERPUSTAKAAN

Lemma 2.10

1. $E(\hat{Y}) = E(Y/x_0)$

2. $var(\hat{Y}) = \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]$

Bukti:

Bagian 1.

$$\begin{aligned} E(\hat{Y}) &= E(\hat{\beta}_0 + \hat{\beta}_1 x_0) \\ &= E(\hat{\beta}_0) + E(\hat{\beta}_1 x_0) \\ &= E(\hat{\beta}_0) + x_0 E(\hat{\beta}_1) \\ &= \beta_0 + \beta_1 x_0 \\ &= E(Y|x_0) \end{aligned}$$

(R.K Sembiring) ■

bagian 2.

$$\begin{aligned} var(\hat{Y}) &= var(\hat{\beta}_0 + \hat{\beta}_1 x_0) \\ &= var(\hat{\beta}_0) + x_0^2\, var(\hat{\beta}_1) + 2x_0 kov(\hat{\beta}_0, \hat{\beta}_1) \\ &= \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] + x_0^2\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} + 2x_0\left[-\frac{\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] \\ &= \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} + \frac{x_0^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} - \frac{2x_0\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] \\ &= \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2 + x_0^2 - 2x_0\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] = \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right] \end{aligned}$$

(R.K Sembiring) ■

Jadi interval konfidensi untuk $E(Y/x_0)$ dengan tingkat kepercayaan 1 - α diberikan oleh:

$$\hat{y} - t(n-2;1-\alpha/2)s(\hat{y}) \le E(Y/x_0) \le \hat{y} + (n-2;1-\alpha/2)s(\hat{y})$$

bila $s(\hat{y}) = s\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$

dalam persamaan regresi $Y = X\beta + \varepsilon$, nilai dari $s(\hat{y})$ adalah:

$$s(\hat{y}) = s\sqrt{x_0'(X'X)^{-1}x_0}$$

2.3.2 Interval Konfidensi Simultan dari Koefisien Regresi

Selama ini untuk menentukan interval konfidensi biasanya dengan menyusun interval konfidensi pada tingkat kepercayaan tertentu untuk tiap parameter secara sendiri-sendiri. Kesulitannya adalah bahwa tiap interval konfidensi yang terbentuk untuk tiap parameter tidak akan memberikan kepercayaan pada tingkat tertentu bahwa kesimpulan bagi tiap parameter itu benar. Peluang tiap parameter itu benar adalah besarnya tingkat kepercayaan pangkat banyaknya parameter. Dalam kenyataannya tiap parameter itu tidak bebas sebab tiap parameter tersebut berasal dari contoh data yang sama, sehingga membuat perhitungan peluang tiap parameter itu benar menjadi jauh lebih sempit.

Dalam analisis data seringkali memerlukan serangkaian pendugaan dengan memperoleh jaminan tentang kebenaran keseluruhan nilai dugaan tersebut. Untuk itu perlu adanya pendugaan secara bersama atau simultan. Karena dengan menentukan

interval konfidensi secara simultan akan lebih menjamin keseluruhan dugaan adalah benar jika dibandingkan secara sendiri-sendiri.

Untuk interval konfidensi simultan kita menaksir beberapa parameter secara serentak yang sering kita kenal dengan daerah kepercayaan. Jika pada selang kepercayaan kita menggunakan uji-t, tetapi pada daerah kepercayaan kita menggunakan uji-F. Pada interval konfidensi simultan kita berhadapan dengan lebih dari satu parameter, sedangkan pada selang kepercayaan kita berhadapan hanya dengan satu parameter.

Teorema 2.11

Jika Z dan U adalah variabel independen berdistribusi Chi-square dengan derajat bebas v_1 dan v_2, distribusi dari

$$F = \frac{Z/v_1}{U/v_2}$$

Bukti :

Fungsi Probabilitas bersama Z dan U adalah

$$f(z,u) = \frac{1}{\Gamma(v_1/2)\Gamma(v_2/2)2^{(v_1+v_2)/2}} \cdot z^{v_1/2-1}u^{v_2/2-1}e^{-(z+u)/2} \quad 0 < z < \infty, 0 < u < \infty$$

Kita dapatkan persamaan baru, yaitu:

$F = \frac{Z/v_1}{U/v_2}$, pdf g(f) dari persamaan F adalah

$f = \frac{z/v_1}{u/v_2}$ dan $t = u$, dimana $z = (v_1/v_2)\,tf$, $u = t$ dan $|J| = (v_1/v_2)t$

pdf join g(f,t) variabel random F dan T = U adalah

$$g(f,t) = \frac{1}{\Gamma(v_1/2)\Gamma(v_2/2)2^{(v_1+v_2)/2}} \cdot \left(\frac{v_1 tf}{v_2}\right)^{v_1/2-1} \cdot t^{v_2/2-1} \cdot \exp\left[-\frac{t}{2}\left(\frac{v_1 f}{v_2}+1\right)\right]\frac{v_1 t}{v_2}$$

marjinal pdf g(f) :

$$g(f) = \int_{-\infty}^{\infty} g(f,t)\,dt$$

$$= \int_{-\infty}^{\infty} \frac{(v_1/v_2)^{v_1/2}(f)^{v_1/2-1}}{\Gamma(v_1/2)\Gamma(v_2/2)2^{(v_1+v_2)/2}} \cdot t^{(v_1+v_2)/2-1} \cdot \exp\left[-\frac{t}{2}\left(\frac{v_1}{v_2}+1\right)\right] dt$$

misalkan $y = \frac{t}{2}\left(\frac{v_1 f}{v}+1\right)$, maka :

$$g(f) = \int_{-\infty}^{\infty} \frac{(v_1/v_2)^{v_1/2}(f)^{v_1/2-1}}{\Gamma(v_1/2)\Gamma(v_2/2)2^{(v_1+v_2)/2}} \cdot \left(\frac{2y}{v_1 f/v_2+1}\right)^{(v_1+v_2)/2-1} \cdot e^{-y} \cdot \left(\frac{2}{v_1 f/v_2+1}\right) dt$$

Dengan demikian distribusi F mempunyai pdf untuk $f > 0$:

$$g(f; v_1; v_2) = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}\left(\frac{v_1}{v_2}\right)^{v_1/2} f^{(v_1/2)-1}\left(1+\frac{v_1}{v_2}f\right)^{-(v_1+v_2)/2}$$

yang mana merupakan densitas dari distribusi F dengan derajat bebas v_1 dan v_2, notasinya adalah $F \sim F(v_1, v_2)$. *(Hogg and Craig, 1978)* ■

Adapun untuk mendapatkan interval konfidensi simultan $100(1-\alpha)$ % untuk parameter β adalah :

$$(\beta - \hat{\beta})X'X(\beta - \hat{\beta}) \le ps^2 F(p, v, 1-\alpha)$$

dalam hal ini $F(p,v,1-\alpha)$ adalah persentil pada sebaran F(p, v) dan s^2 adalah sebagai nilai dugaan dari σ^2 diasumsikan bahwa modelnya benar. Bila modelnya tidak benar, s^2 biasanya menjadi terlalu besar sebab ia merupakan suatu peubah acak dengan nilai tengah lebih besar dari σ^2. Secara umum ini akan bermanfaat bila p kecil, misalnya 2,3,4, kecuali diambil langkah untuk menyajikan informasi tersebut dalam bentuk yang mudah dipahami.

2.4 Analisis Kovariansi

Dalam mempelajari analisis kovariansi harus memperhatikan efek pertama atau faktor utama untuk membuat variabel respon, misalkan Y, bisa juga berupa variabel lain atau beberapa variabel yang mendekati Y. Dalam analisis kovariansi seringkali tidak memungkinkan untuk mengontrol semua variabel pada tingkat konstan untuk seluruh eksperimen. Variabel Y ini bisa dijadikan ukuran sebagai variabel respon, yang mana sebagai penghubung variabel bersama dengan variabel respon Y. Dalam kelompok tersebut dapat memperkirakan pengaruh dari perlakuan pada Y. Salah satu percobaan yang pertama dengan menghilangkan efek dari variabel bersama X. Teknik penghilangan efek X (atau semua X) pada Y dan selanjutkan menganalisis sisaan untuk efek perlakuan pada Y yang dinamakan analisis kovariansi. Definisi lain dari analisis kovariansi itu sendiri adalah merupakan kombinasi antara analisis variansi dengan analisis regresi.

Manfaat dari analisis kovariansi adalah

1. membantu dalam menaksir data, terutama dengan memperhatikan sifat-sifat dasar pengaruh dari perlakuan,

2. memisahkan suatu total kovariansi di dalam bagian-bagiannya,
3. menyesuaikan mean perlakuan variabel dependen untuk himpunan harga-harga yang berbeda,
4. menduga data yang hilang.

Dalam hal ini variabel yang akan dianalisis adalah variabel dependen yang diberi notasi Y dan dipakai dalam mengontrol kesalahan dan mengontrol mean yang disesuaikan dengan variabel independen yang diberi notasi X. X adalah tertentu dan diukur tanpa kesalahan, sebab X adalah parameter yang dihubungkan dengan mean populasi Y. X juga dipengaruhi oleh perlakuan-perlakuan lain tapi X mempunyai pengaruh terhadap Y. Sebuah korelasi positif adalah dugaan antara X dan Y, dalam sebuah model regresi dapat dituliskan sebagai berikut :

$$Y_{ij} = \mu + \beta\left(X_{ij} - \overline{X}\right) + \varepsilon_{ij} \quad \ldots \quad (2.2)$$

dimana :

β adalah koefisien regresi linier (atau slope) antara X dan Y untuk semua data.

$\overline{X}$ adalah rata-tara harga X.

Dalam model di atas diasumsikan ukuran dari variabel X berasal dari beberapa unit dari ukuran Y yang bersesuaian. Untuk model analisis kovariansi satu faktor yaitu:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \ldots \quad (2.3)$$

dan model persamaan 2.2 dapat dikombinasikan menjadi model kovariansi yaitu :

$$Y_{ij} = \mu + \beta\left(X_{ij} - \overline{X}\right) + \tau_i + \varepsilon_{ij} \quad , i = 1,2,3.....I$$

$$j = 1,2,3,....J \qquad \ldots \quad (2.4)$$

dimana:

Y_{ij} = pengamatan ke-j dari group yang mendapat perlakuan ke-i,

μ = mean efek sebenarnya,

τ_i = efek perlakuan ke-i yang disesuaikan terhadap x,

ε_{ij} = kesalahan random.

BAB III

PEMBAHASAN

3.1 Metode untuk menentukan interval konfidensi simultan

Ada dua metode yang telah dikenal baik untuk menghitung interval konfidensi simultan dari koefisien regresi yaitu metode Scheffe dan Bonferroni. Serta untuk metode Hybrid merupakan metode yang masih baru.

3.1.1 Metode Bonferroni

Suatu prosedur yang menghasilkan koefisien konfidensi simultan seringkali sangat diingini sebab memungkinkan pengguna menyatukan hasil-hasil yang terpisah menjadi suatu kumpulan yang terpadu, dengan jaminan bahwa keseluruhan dugaan adalah benar. Metode Bonferroni salah satu metode untuk menyusun interval konfidensi simultan dengan tingkat kepercayaan tertentu. Metode Bonferroni menerangkan konfidensi simultan bagi himpunan berhingga dari kombinasi-kombinasi linier. Metode ini tidak bisa digunakan pada prediktor kontinu tanpa pembatasan banyaknya nilai prediktor yang mungkin ke suatu himpunan berhingga. Pada metode ini biasanya akan memuat prediktor yang kategorik. Dengan kata lain secara teori metode ini akan menghasilkan interval konfidensi sempit pada saat prediktornya kategorik.

Dalam persamaan regresi sederhana $Y = \beta_0 + \beta_1 X_1 + \varepsilon$, batas-batas kepercayaan untuk β_0 dan β_1 dengan koefisien kepercayaan masing-masing $1 - \alpha$ adalah :

$$\hat{\beta}_0 \pm t(1-\alpha/2;n-2)s(\hat{\beta}_0) \qquad \hat{\beta}_1 \pm t(1-\alpha;n-2)s(\hat{\beta}_1)$$

peluang bahwa kedua interval itu benar adalah dengan memisalkan A1 menyatakan kejadian bahwa interval konfidensi pertama tidak mencakup β_0 dan A_2 kejadian bahwa interval konfidensi kedua tidak mencakup β_1,

sehingga :

$P(A_1) = \alpha$ $\qquad$ $P(A_2) = \alpha$

dimana,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$
$$1 - P(A_1 \cup A_2) = 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2)$$

dapat dituliskan,

$$1 - P(A_1 \cup A_2) = P(\overline{A_1 \cup A_2}) = P(\overline{A_1} \cap \overline{A_2})$$

$P(\overline{A_1} \cap \overline{A_2})$ adalah peluang bahwa kedua interval konfidensi adalah benar, sehingga dari persamaan dapat dituliskan:

$$P(\overline{A_1} \cap \overline{A_2}) = 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2)$$

untuk $P(A_1 \cap A_2) \geq 0$ maka diperoleh pertidaksamaan Bonferroni :

$$P(A_1 \cap A_2) \geq 1 - P(A_1) - P(A_2)$$

untuk kasus di atas menjadi;

$$P(A_1 \cap A_2) \geq 1 - \alpha - \alpha = 1 - 2\alpha$$

jadi jika β_0 dan β_1 diduga secara sendiri-sendiri, misalkan diambil selang kepercayaan 95%, maka pertidaksamaan Bonferroni memberi tingkat kepercayaan bersama 90% bahwa kedua interval konfidensi yang didasarkan pada contoh yang sama adalah benar.

Perhitungan di atas menghasilkan batas interval Bonferroni $1-\alpha/2-\alpha/2=1-\alpha$. Jadi batas-batas kepercayaan bersama $1-\alpha$ bagi β_0 dan β_1 menurut prosedur Bonferroni adalah

$$\hat{\beta}_0 \pm Bs(\hat{\beta}_0) \qquad \text{dan} \qquad \hat{\beta}_1 \pm Bs(\hat{\beta}_1)$$

dimana $B = t(1-\alpha/4;n-2)$

Untuk q interval konfidensi simultan dengan tingkat kepercayaan $1-\alpha$, didapat:

$P\left[\bigcap_{i=1}^{q} \bar{A}_1\right] \geq 1-2q$, maka sudah memenuhi syarat jika setiap selang dibuat dengan tingkat kepercayaan sebesar $1-\alpha/2q$. Semakin besar harga q maka akan memperbesar harga B.

Untuk menyusun interval konfidensi bagi rataan respon dengan amatan baru pada nilai x_0 yang berbeda, dapat dihitung batas-batas kepercayaan bagi rataan respon amatan baru $E(Y/x_0)$ dengan q buah nilai x_0, dengan tingkat kepercayaan bersama 1-α, maka batas-batas kepercayaan untuk Bonferroni adalah:

$$\hat{Y} \pm Bs(\hat{Y})$$

dimana: $B = t(1-\alpha/2q;n-2) = \sqrt{F(1;n-p;1-\alpha/q)}$, karena

$$F = \frac{\hat{\beta}_1^{\,2}\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum(y_i-\hat{y}_i)^2/n-2} \quad \text{sedangkan} \quad s^2(\hat{\beta}_1) = \frac{\sum(y_i-\hat{y}_i)^2/n-2}{\sum_{i=1}^{n}(x_i-\bar{x})}$$

maka diperoleh:

$$F = \frac{\hat{\beta}_1^{\,2}}{s^2(\hat{\beta}_1)} = \left(\frac{\hat{\beta}_1}{s(\hat{\beta}_1)}\right)^2 \quad \text{dimana} \quad t = \frac{\hat{\beta}_1}{s[\hat{\beta}_1]}$$

sehingga dapat disimpulkan bahwa:

$$(t)^2 = \left[\frac{\hat{\beta}_1}{s(\hat{\beta}_1)}\right]^2 = F$$

Dalam bentuk matrik dapat dituliskan :

$$s(\hat{Y}) = s\sqrt{x_0'(X'X)^{-1}x_0}$$
$$\hat{Y} = x_0'\hat{\beta}$$

jadi interval konfidensi simultan dengan menggunakan metode Bonferroni adalah:

$$x_0'\hat{\beta} \pm s\sqrt{F(1;n-p;1-\alpha/q)} \cdot \sqrt{x_0'(X'X)^{-1}x_0} \quad \ldots \quad (2.5)$$

dengan,

x_0 = amatan baru,

p = banyaknya suku atau banyaknya parameter termasuk β_0, dan dapat berubah besarnya sesuai dengan bentuk persamaan regresinya,

q = banyaknya amatan baru atau besarnya kombinasi yang mungkin dari prediktor kategorik (kalau dalam kasus terdapat prediktor kategorik),

s = standar deviasi yang diduga dari data.

Dasar interval dalam metode di atas mengganti α menjadi α/q yang dinamakan interval Bonferroni.

3.1.2 Metode Scheffe

Secara teori metode Scheffe sama dengan metode Bonferroni tetapi pada metode Scheffe biasanya akan menghasilkan interval konfidensi simultan yang lebih sempit bila data yang digunakan adalah kontinu, walaupun dalam kenyataannya tidak

selalu. Jadi pada intinya metode Scheffe lebih efisien untuk digunakan apabila prediktornya kontinu. Pada metode ini menggunakan sebaran F, dalam persamaan regresi $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ akan didapat batas-batas konfidensi bagi p banyaknya amantan baru dengan tingkat kepercayaan 1-α adalah

$$\hat{Y} \pm Ss\left(\hat{Y} / x_0\right)$$

dimana, $S = pF(p; n-2; 1-\alpha)$

Misalkan $Y = (y_i, \ldots, y_n)$ vektor n independen, variabel random distribusi Normal dengan variansi biasa σ^2. Misalkan rata-rata vektor y didapat $Y = X\beta$ dimana vektor Y berukuran (n x 1) dan matrik X berukuran (n x p),

$$x = \begin{bmatrix} x_{11}, \ldots, x_{1p} \\ \ldots\ldots\ldots\ldots \\ x_{n1}, \ldots, x_{np} \end{bmatrix}$$

dimana Rank x = p ; (p < n)

$$\beta = (\beta_1, \ldots\ldots, \beta_p)$$

untuk matrik X variabel independen sudah diketahui, tapi untuk parameter regresi β, variansi error σ^2 belum diketahui, dengan maksimum likelihood β adalah $\hat{\beta} = (X^T X)^{-1} X^T Y$ dan estimasi variansi tidak bias $s^2 = (1/(n-p))\sum (y - \hat{y})^2$, akan didapatkan interval konfidensi simultan yang hampir sama dengan metode Bonferroni, hanya saja pada persentil F-nya saja yang membedakan. Dengan demikian diperoleh interval konfidensi simultan Scheffe sebagai beriku:

$$x_0' \hat{\beta} \pm s\sqrt{pF(1; n-p; 1-\alpha)} \cdot \sqrt{x_0'(X'X)^{-1} x_0} \qquad \ldots \quad (2.6)$$

dimana x_0 = amatan baru

p = banyaknya suku atau banyaknya parameter termasuk β_0, dan dapat berubah besarnya sesuai dengan bentuk persamaan regresinya.

3.1.3 Metode Hybrid

Metode Hybrid adalah metode baru dimana penggunaannya lebih luas dari metode yang lain, karena metode ini dapat digunakan pada saat prediktornya kontinu dan kategorik, metode Scheffe dan Bonferroni adalah kasus khusus dari metode Hybrid.

Seperti halnya pada metode Scheffe dan Bonferroni, metode Hybrid juga hampir sama. Misalkan Y adalah suatu vektor respon berdimensi n atau bisa dikatakan vektor Y berukuran (n x 1), dan misalkan X adalah matrik berukuran (n x p). Estimasi kuadrat dari vektor β dimana vektor yang belum diketahui adalah $\hat{\beta} = (X'X)^{-1}X'Y$. Dimisalkan juga q adalah banyaknya tingkat kombinasi yang mungkin dari prediktor kategorik dan p_0 adalah banyaknya suku yang berbeda dalam persamaan, maka dapat dituliskan interval konfidensi dari metode Hybrid adalah:

$$x_0'\hat{\beta} \pm s\sqrt{p_0 F(p_0; n-p; 1-\alpha/q)}\sqrt{x_0'(X'X)^{-1}x_0} \quad \ldots (2.7)$$

dengan probabilitas seperti pada metode Bonferroni yaitu $1-\alpha/q$ untuk harga kritik. Metode Hybrid itu sendiri secara tidak langsung merupakan penggabungan dari dua metode, yaitu metode Scheffe dan Bonferroni. Dengan kata lain metode Bonferroni adalah kasus khusus untuk metode Hybrid pada saat tidak ada prediktor kontinu, sehingga p_0 = 1 dan q adalah banyaknya tingkat kombinasi pada prediktor kategorik. Metode Scheffe merupakan kasus khusus juga bagi metode Hybrid pada saat tidak ada

prediktor kategorik, sehingga $p_0 = p$ dan $q = 1$, dan memperlakukannya sebagaimana semua prediktornya adalah kontinu.

3.2 Perbandingan Metode Scheffe dan Hybrid dalam Menentukan Interval Konfidensi Simultan

Dalam bab ini akan dijelaskan mengenai perbandingan dari dua metode dalam menentukan interval konfidensi simultan, yaitu metode Scheffe dan Hybrid.

Model regresi yang digunakan adalah model regresi kuadratik dengan tipe model sebagai berikut :

Dengan persamaan regresinya adalah

$$Y = \beta_{0ij} + \beta_{1ij}X_1 + \beta_{2ij}X_2 + \beta_3X_1X_2 + \beta_4X_1^2 + \beta_5X_2^2$$

Tabel 1. Jenis model dalam regresi kuadratik

Model	Tetap	Linier	Kuadratik	Jumlah suku (p)
1	A	C	C	I+J+4
2	I	C	C	IJ+5
3	A	A	C	3I+3J
4	I	A	C	IJ+2I+2J+1
5	I	I	C	3IJ+3

dimana:

Tetap (C) : $\beta_{0ij} = \gamma$ 1 derajat bebas,

Aditive (A) : $\beta_{0ij} = \gamma_i + \delta_J$ I + J – 1 derajat bebas,

Interaksi (I) : $\beta_{0ij} = \gamma_{ij}$ IJ derajat bebas.

Dengan cara yang sama, tiap-tiap β_{1ij} dan β_{2ij} bisa mempunyai sembarang bentuk dari tiga bentuk di atas dan mengkontribusikan banyaknya derajat bebas yang sama terhadap regresi.

Dalam banyak kasus, metode Hybrid menghasilkan interval yang lebih sempit dari metode Scheffe namun tidak selalu. Kedua metode menghitung setengah lebar sebagai perkalian dari tiga faktor yaitu s, fungsi X dan model matrik dan fungsi dari persentil F. Harga faktor yang ketiga yang membedakan antara dua metode. Oleh karena itu kita asumsikan perkalian dua suku pertama adalah 1, maka setengah lebar dari dua interval adalah :

$\sqrt{pF(p, n-p, 1-\alpha)}$ untuk interval Scheffe dan

$\sqrt{p_0 F(p_0, n-p, 1-\alpha/q)}$ untuk interval Hybrid

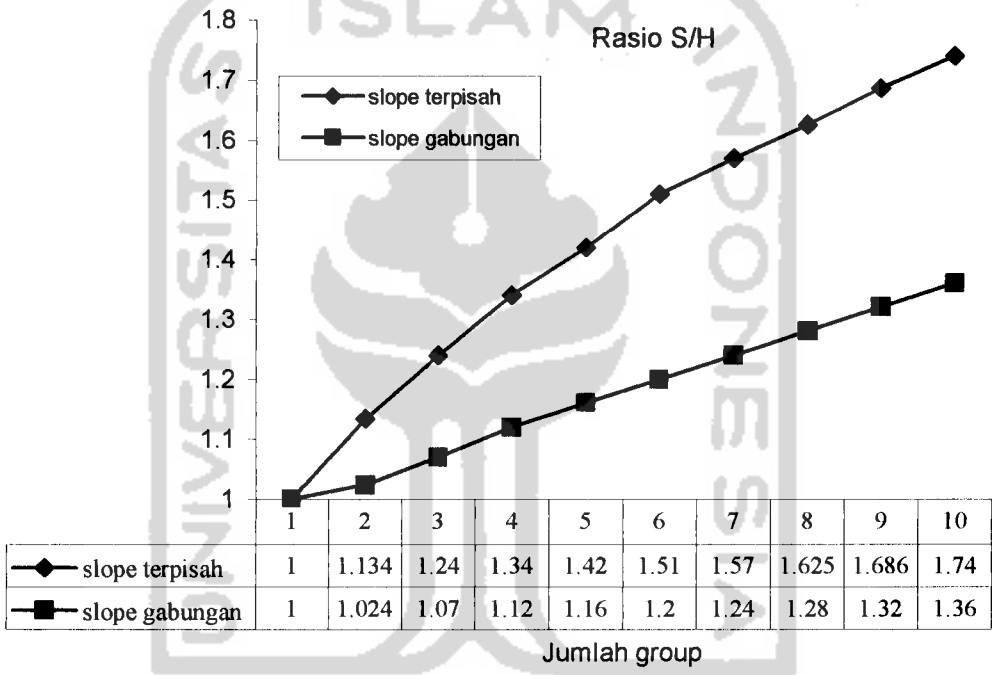
untuk membandingkan harga-harga pada metode regresi biasa. Pertama dipandang suatu model Anova satu arah keduanya dengan slope terpisah dan slope gabungan dimana:

slope terpisah : $E[Y/i, x] = \beta_{0i} + \beta_{1i} x$

slope gabungan : $E[Y/i, x] = \beta_{0i} + \beta_1 x$

Untuk lebih jelasnya dapat dilihat pada gambar 1, dimana memperlihatkan perbandingan antara persentil F Scheffe dengan persentil F Hybrid dengan $1 \leq i \leq I$, nilai metode Hybrid dari $\sqrt{p_0 F(p_0, n-p, 1-\alpha/q)}$ diambil I bervariasi dari 1 hingga 10, dan tiap-tiap masuk hitungan dengan 30 derajat bebas dan tingkat kepercayaan 95%, untuk semua nilai Hybrid parameter q = I dan $p_0 = 2$. Pendekatan Scheffe untuk model

slope terpisah diperoleh dengan mengeset p = 2I. Harga-harga untuk model gabungan diperoleh dengan mengeset p = I + 1. Harga n dipilih sedemikian hingga n – p = 30. Pandang model slope terpisah $E[Y/i,x] = \beta_{0i} + \beta_{1i}x$ jika hanya ada satu grup, maka akan tereduksi ke model regresi biasa sehingga metode Scheffe dan Hybrid akan indentik.



	1	2	3	4	5	6	7	8	9	10
slope terpisah	1	1.134	1.24	1.34	1.42	1.51	1.57	1.625	1.686	1.74
slope gabungan	1	1.024	1.07	1.12	1.16	1.2	1.24	1.28	1.32	1.36

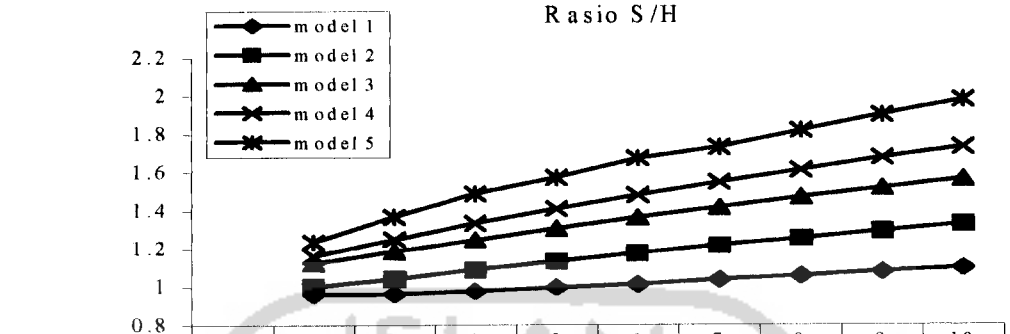
Gambar 1. Perbandingan panjang interval dari Hybrid dan Scheffe

Dari gambar 1 diatas dapat dijelaskan kalau banyaknya grup bertambah, lebar interval Scheffe meningkat lebih cepat daripada lebar interval Hybrid. Jika hanya ada tiga grup, interval Scheffe hampir 25% lebih lebar daripada interval Hybrid, namun pada slope terpisah interval Scheffe akan bertambah lebih cepat dibandingkan dengan interval Hybrid daripada dalam model slope gabungan. Untuk kedua model, interval Hybrid lebih kecil dari interval Scheffe jika banyaknya grup dua atau lebih . dan dapat

dikatakan juga dari gambar 1 menunjukkan interval Hybrid ini dapat namun tidak selalu lebih baik dari metode Scheffe.

Perbandingan interval Scheffe dan Hybrid yang lebih umum dapat dibuat dengan catatan bahwa sembarang model regresi $\sqrt{pF(p, n-p, 1-\alpha)}$ dan $\sqrt{p_0 F(p_0, n-p, 1-\alpha/q)}$ bergantung hanya pada p, p_0 , dan q (asumsikan bahwa α dan derajat bebas sesatan db = n – p tetap). Interval Scheffe dan Hybrid 95 % akan mempunyai panjang sama jika $\sqrt{pF(p, n-p, 0.95)} = \sqrt{p_0 (p_0, db, 1-0.05/q)}$ untuk p, p_0 dan db tetap. Metode Hybrid akan menghasilkan interval yang lebih pendek dan cenderung bekerja dengan baik untuk model dimana p besar dibandingkan p_0 dan mempunyai derajat bebas yang besar. Metode Scheffe akan cenderung menghasilkan interval yang pendek pada model-model dimana p mempunyai selisih yang kecil dengan p_0, dan mempunyai derajat bebas yang kecil.

Selanjutnya lihat gambar 2, dimana menunjukkan rasio dari persentil F Scheffe dengan persentil F Hybrid sebagai fungsi dari J, dengan I = 2. Dan diambil dari kelima model pada tabel 1 yang menunjukkan slope terpisah. Dalam tiap kasus, parameter-parameter interval Hybrid adalah p_o=6 dan q=IJ dan tiap-tiap perhitungan menggunakan 30 derajat bebas dan tingkat konfidensi 95 %.



	1	2	3	4	5	6	7	8	9	10
model 1		0.957	0.96	0.975	0.993	1.01	1.033	1.053	1.074	1.094
model 2		0.999	1.041	1.085	1.129	1.173	1.213	1.252	1.291	1.327
model 3		1.122	1.183	1.245	1.303	1.361	1.415	1.466	1.516	1.563
model 4		1.1603	1.249	1.33	1.407	1.479	1.547	1.61	1.673	1.73
model 5		1.232	1.366	1.486	1.57	1.672	1.728	1.817	1.899	1.978

besarnya J (pada data kategorik mulai dari tingkat ke-2)

Gambar 2. Perbandingan dari kelima model pada Hybrid dan Scheffe

Dari gambar 2 di atas dapat terlihat kalau metode Hybrid lebih baik dari metode Scheffe untuk 3 dari 5 model dengan 30 derajat bebas residual. Dan dari gambar tersebut menunjukkan kalau metode Hybrid ini akan bekerja dengan baik jika derajat bebasnya paling sedikit 6.

Perhitungan yang mendasar pada perbandingan interval Scheffe dan Hybrid tidak tergantung pada respon namun lebih difokuskan pada persentil F-nya saja. Dalam beberapa kasus terapan regresi, memungkinkan untuk menghitung dari kedua interval dan memilih salah satu yang mempunyai interval lebih sempit, tanpa memperhatikan tingkat konfidensi simultannya.

UNIVERSITAS ISLAM INDONESIA
PERPUSTAKAAN
F MIPA

BAB IV

APLIKASI

Aplikasi tentang penentuan interval konfidensi simultan dapat digambarkan dengan menggunakan contoh sebagai berikut:

Dalam hal ini diambil data yang memuat prediktor kategorik dan prediktor kontinu. Dengan menggunakan 30 lahan pertanian seluas 2 ha (1 ha = 100m x 100m) sebagai sampel, lahan tersebut ditanami tanaman padi dengan masing-masing lahan mempunyai luas yang sama dan mempunyai kusuburan tanah yang sama . Dari lahan tersebut akan diketahui hasil panen padi dari tiap-tiap lahan yang diberi pupuk dengan berat dan komposisi yang berbeda pula. Pupuk yang digunakan adalah pupuk alami (kandang dan kompos) dan pupuk buatan (Urea, ZA, dan NPK). Dari lahan satu ha dapat dihasilkan 3 ton gabah kering dengan pupuk ± 2 kwintal. Dengan banyaknya panen tiap lahan (dengan satuan kwintal) adalah sebagai berikut :

Tabel 2. banyaknya panen dalam kwintal pada 30 lahan

		Pupuk Buatan								
		Urea			ZA			NPK		
		Y	X_1	X_2	Y	X_1	X_2	Y	X_1	X_2
Pupuk Alami	Pupuk Kandang	85	2.15	1.85	73	2.60	1.90	88	2.80	1.00
		87	2.30	1.75	68	2.00	2.50	86	2.80	1.20
		75	1.50	2.58	78	1.80	2.00	90	3.00	1.50
		80	1.80	2.20	82	2.20	2.15	83	2.30	1.80
		82	2.20	1.80	85	2.60	1.60	76	2.00	2.30

		Pupuk Buatan								
		Urea			ZA			NPK		
		Y	X_1	X_2	Y	X_1	X_2	Y	X_1	X_2
Pupuk Alami	Pupuk Kompos	89	2.30	1.80	72	2.40	2.00	78	2.00	1.90
		92	2.80	1.60	78	2.30	1.75	68	1.80	2.40
		86	2.15	1.80	83	1.80	2.00	80	2.10	1.50
		93	2.40	1.40	80	1.60	2.40	66	1.60	2.40
		82	2.30	2.00	84	2.00	1.80	84	2.30	1.30

Untuk memperjelas dapat dilihat rincian komposisi banyaknya pupuk tiap lahan (dengan satuan Kwintal) adalah sebagai berikut :

Y : hasil panen

X_1 : pupuk buatan jenis urea

X_2 : pupuk alami jenis kandang

No	Y	X_1	X_2
1	85	2.15	1.85
2	87	2.30	1.75
3	75	1.50	2.58
4	80	1.80	2.20
5	82	2.20	1.80

Y : hasil panen

X_1 : pupuk buatan jenis NPK

X_2 : pupuk alami jenis kandang

No	Y	X_1	X_2
1	88	2.80	1.00
2	86	2.80	1.20
3	90	3.00	1.50

Y : hasil panen

X_1 : pupuk buatan jenis ZA

X_2 : pupuk alami jenis kandang

No	Y	X_1	X_2
1	73	2.60	1.90
2	68	2.00	2.50
3	78	1.80	2.00
4	82	2.20	2.15
5	85	2.60	1.60

Y : hasil panen

X_1 : pupuk buatan jenis urea

X_2 : pupuk alami jenis kompos

No	Y	X_1	X_2
1	89	2.30	1.80
2	92	2.80	1.60
3	86	2.15	1.80

4	93	2.30	1.80
5	76	2.00	2.30

4	93	2.40	1.40
5	82	2.30	2.00

Y : hasil panen

X_1 : pupuk buatan jenis ZA

X_2 : pupuk alami jenis kompos

Y : hasil panen

X_1 : pupuk buatan jenis NPK

X_2 : pupuk alami jenis kompos

No	Y	X_1	X_2
1	72	2.40	2.00
2	78	2.30	1.75
3	83	1.80	2.00
4	80	1.60	2.40
5	84	2.00	1.80

No	Y	X_1	X_2
1	78	2.00	1.90
2	68	1.80	2.40
3	80	2.10	1.50
4	66	1.60	2.40
5	84	2.30	1.30

Dari data di atas kita ingin membuat persamaan regresi yang bertujuan untuk meramalkan hasil panen. Dalam kasus di atas kita ingin melihat bahwa hasil panen hanya dipengaruhi oleh dua kelompok pupuk yaitu pupuk buatan dan pupuk alami. Dengan demikian kasus di atas dapat dijadikan satu data, dimana masing-masing jenis tiap kelompok sebagai data kategorik. Untuk kelompok pupuk alami ada dua kategori (I), untuk pupuk kelompok pupuk buatan tiga kategori (J). data di atas dapat disederhanakan menjadi :

No	Y	X_1	X_2	I	J
1	85	2.15	1.85	1	1
2	73	2.60	1.90	1	2
3	88	2.80	1.00	1	3
4	87	2.30	1.75	1	1
5	68	2.00	2.50	1	2

6	86	2.80	1.20	1	3
7	75	1.50	2.85	1	1
8	78	1.80	2.00	1	2
9	90	3.00	1.50	1	3
10	80	1.80	2.20	1	1
11	82	2.20	2.15	1	2
12	83	2.30	1.80	1	3
13	82	2.20	1.80	1	1
14	85	2.60	1.60	1	2
15	76	2.00	2.30	1	3
16	89	2.30	1.80	2	1
17	72	2.40	2.00	2	2
18	78	2.00	1.90	2	3
19	92	2.80	1.60	2	1
20	78	2.30	1.75	2	2
21	68	2.10	2.40	2	3
22	86	2.15	1.80	2	1
23	83	1.80	2.00	2	2
24	80	2.10	1.50	2	3
25	93	2.40	1.40	2	1
26	80	1.60	2.40	2	2
27	66	1.60	2.40	2	3
28	82	2.30	2.00	2	1
29	84	2.00	1.80	2	2
30	84	2.30	1.30	2	3

4.1 Menduga Persamaan Regresi

Data di atas akan diduga dengan menggunakan persamaan regresi model kuadratik. Model persamaan regresi kuadratik tersebut adalah sebagai berikut:

$$Y = \beta_{0ij} + \beta_{1ij}X_1 + \beta_{2ij}X_2 + \beta_3X_1X_2 + \beta_4X_1^2 + \beta_5X_2^2$$

untuk mencari koefisien-koefisien regresi model di atas, kita menggunakan metode matrik, karena koefisien regresi dari data di atas lebih dari satu dan mempunyai data yang kompleks, sehingga akan lebih mudah mencarinya dengan menggunakan matrik. Dalam lambang matrik dapat dituliskan sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \cdots x_{1p} \\ 1 & x_{21} & x_{22} \cdots x_{2p} \\ \vdots & \vdots & \vdots \quad \cdots \vdots \\ 1 & x_{n1} & x_{n2} \cdots x_{np-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{p-1} \end{bmatrix}$$

untuk menghitung koefisien regresinya dapat dihitung dengan menggunakan bantuan komputer untuk menyelesaikannya.

Harga dari koefisien $\hat{\beta}$ adalah :

$$X' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2.15 & 2.60 & \cdots & 2.30 \\ 1.85 & 1.90 & \cdots & 1.30 \\ 3.98 & 4.94 & \cdots & 2.99 \\ 4.64 & 6.76 & \cdots & 5.29 \\ 3.42 & 3.64 & \cdots & 1.69 \end{pmatrix}$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

$$= \left[\begin{pmatrix} 1 & 1 & \ldots & 1 \\ 2.15 & 2.60 & \ldots & 2.30 \\ 1.85 & 1.90 & \ldots & 1.30 \\ 3.98 & 4.94 & \ldots & 2.99 \\ 4.64 & 6.76 & \ldots & 5.29 \\ 3.42 & 3.64 & \ldots & 1.69 \end{pmatrix} \begin{pmatrix} 1 & 2.15 & 1.85 & 3.98 & 4.64 & 3.42 \\ 1 & 2.60 & 1.90 & 4.94 & 6.76 & 3.64 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2.30 & 1.30 & 2.99 & 5.29 & 1.69 \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 2.15 & 2.60 & \ldots & 2.30 \\ 1.85 & 1.90 & \ldots & 1.30 \\ 3.98 & 4.94 & \ldots & 2.99 \\ 4.64 & 6.76 & \ldots & 5.29 \\ 3.42 & 3.64 & \ldots & 1.69 \end{pmatrix} \begin{bmatrix} 85 \\ 73 \\ 88 \\ 87 \\ \vdots \\ 84 \end{bmatrix}$$

$$= \begin{pmatrix} 30.00 & 66.20 & 56.45 & 121.16 & 150.14 & 111.08 \\ 66.20 & 150.14 & 121.16 & 267.53 & 349.42 & 232.13 \\ 56.45 & 121.16 & 111.08 & 232.14 & 267.53 & 227.54 \\ 121.16 & 267.53 & 232.13 & 499.37 & 606.86 & 463.29 \\ 150.14 & 349.42 & 267.53 & 606.86 & 833.20 & 499.37 \\ 111.08 & 232.13 & 227.54 & 463.29 & 499.37 & 483.38 \end{pmatrix} \cdot \begin{pmatrix} 2433.0 \\ 5409.7 \\ 4519.4 \\ 9773.7 \\ 12359.0 \\ 8781.9 \end{pmatrix}$$

$$= \begin{pmatrix} -279.304 \\ 168.890 \\ 199.753 \\ -48.872 \\ -17.527 \\ -27.831 \end{pmatrix} = \hat{\beta}$$

Dari hasil koefisien regresi di atas maka didapat peramalan persamaan regresinya adalah :

$$\hat{Y} = -279 + 169\ X_1 + 200\ X_2 - 48.9\ X_1X_2 - 17.5\ X_1^2 - 27.8\ X_2^2$$

4.2 Menentukan Interval Konfidensi Simultan

Dalam menentukan interval konfidensi simultan kita harus meramalkan berapa besarnya Y peramal apabila nilai dari X diketahui:

$x_{01} = 2.7 \quad x_{02} = 1.8$

dengan demikian dapat dituliskan

$$\hat{Y} = \begin{bmatrix} 1 & x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = x'\hat{\beta}$$

$$\hat{Y} = \begin{bmatrix} 1 & 2.7 & 1.8 & 4.86 & 7.29 & 3.249 \end{bmatrix} \begin{bmatrix} -279.304 \\ 168.890 \\ 199.753 \\ -48.872 \\ -17.527 \\ -27.831 \end{bmatrix} = \begin{bmatrix} 80.7904 \end{bmatrix} = x'\hat{\beta}$$

menghitung besarnya ,

$$V(\hat{Y}) = x_o'(X'X)^{-1}x_o\sigma^2$$

dimana σ^2 diduga dengan $s^2 = \dfrac{Y'Y - b'X'Y}{n-p}$

$$s^2 = \frac{\begin{pmatrix} 85 & 73 & 88 & 87 & \cdots & 84 \end{pmatrix} \begin{bmatrix} 85 \\ 73 \\ 88 \\ 87 \\ \vdots \\ 84 \end{bmatrix}}{30-6} - \frac{\begin{pmatrix} -279.304 & 168.890 & 199.753 & -48.872 & -17.527 & -27.831 \end{pmatrix}}{30-6}$$

$$\Leftrightarrow \frac{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2.15 & 2.60 & \cdots & 2.30 \\ 1.85 & 1.90 & \cdots & 1.30 \\ 3.98 & 4.94 & \cdots & 2.99 \\ 4.64 & 6.76 & \cdots & 5.29 \\ 3.42 & 3.64 & \cdots & 1.69 \end{pmatrix} \begin{bmatrix} 85 \\ 73 \\ 88 \\ 87 \\ \vdots \\ 84 \end{bmatrix}}{30-6} = 23.59$$

dan harga dari

$$x_o'(X'X)^{-1}x_o = \begin{pmatrix}1 & 2.7 & 1.8 & 4.86 & 7.9 & 3.24\end{pmatrix}\left[\begin{pmatrix}1 & 1 & \dots & 1\\ 2.15 & 2.60 & \dots & 2.30\\ 1.85 & 1.90 & \cdots & 1.30\\ 3.98 & 4.94 & \dots & 2.99\\ 4.64 & 6.76 & \dots & 5.29\\ 3.42 & 3.64 & \dots & 1.69\end{pmatrix}\begin{pmatrix}1 & 2.15 & 1.85 & 3.98 & 4.62 & 3.42\\ 1 & 2.60 & 1.90 & 4.94 & 6.76 & 3.61\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & 2.30 & 1.30 & 2.99 & 5.29 & 1.69\end{pmatrix}\right]^{-1}$$

$$\Leftrightarrow \begin{pmatrix}1\\ 2.7\\ 1.8\\ 4.86\\ 7.9\\ 3.24\end{pmatrix} = 0.248876$$

$$V(\hat{Y}) = 0.248876 \cdot 23.59$$
$$= 5.87098$$

Untuk menentukan fungsi persentil F maka perlu dihitung:

a p dicari dengan rumus IJ +2I +2J + 1, karena kita menggunakan model regresi kuadratik pada model ke-4.

dimana I adalah banyaknya kombinasi dari pupuk alami

J adalah banyaknya kombinasi dari pupuk buatan

Sehingga harga dari p adalah dimana I = 2 dan J = 3adalah :

$P = 2 \cdot 3 + 2 \cdot 2 + 2 \cdot 3 + 1 = 17$

b p_0 untuk permasalahan diatas adalah 6

c q dengan rumus $IJ = 2 \cdot 3 = 6$

d n-p adalah derajat bebas 30 – 17 = 13

- Interval konfidensi simultan dengan metode Scheffe

  Rumus untuk setengah interval simultannya adalah

  $$x_o'\hat{\beta} \pm s\sqrt{pF(p,n-p,1-\alpha)}\sqrt{x_o'(X'X)^{-1}x_o}$$

  dengan menggunakan tingkat kepercayaan 90% maka :

  harga $F(p,n-p,1-\alpha) = F(17,13,0.1) = 2.031798$

  sehingga : $80.7904 \pm \sqrt{5.871} \cdot \sqrt{34.540} \cdot \sqrt{0.249}$

  jadi interval konfidensi simultan 90% untuk $x_0'\beta$ dengan menggunakan metode Sehiffe adalah

  $$80.7904 - 7.106 \leq x_0'\beta \leq 80.7904 + 7.106$$

  $$73.684 \leq x_0'\beta \leq 87.896$$

- Interval konfidensi simultan dengan metode Hybrid

  Rumus untuk setengah interval simultannya adalah

  $$x_o'\hat{\beta} \pm s\sqrt{p_0F(p_0,n-p,1-\alpha/q)}\sqrt{x_o'(X'X)^{-1}x_o}$$

  dengan menggunakan tingkat kepercayaan 90% maka :

  harga $F(p_0,n-p,1-\alpha/q) = F(6,13,0.1/6) = 4.0357$

  sehingga : $80.7904 \pm \sqrt{5.871} \cdot \sqrt{24.214} \cdot \sqrt{0.249}$

  jadi interval konfidensi simultan 90% untuk $x_0'\beta$ dengan menggunakan metode Hybrid adalah

  $$= 80.7904 - 5.949 \leq x_0'\beta \leq 80.7904 + 5.949$$

  $$74.841 \leq x_0'\beta \leq 86.739$$

BAB V

KESIMPULAN DAN SARAN

Kesimpulan

Beberapa hal yang dapat diambil dari penelitian ini berkaitan dengan penggunaan metode baru, yaitu metode Hybrid dalam menentukan interval konfidensi simultan, adalah :

1. Penggunaan metode Hybrid lebih luas dibandingkan dengan metode Bonferroni dan dalam beberapa kasus regresi, metode Hybrid dapat menghasilkan interval yang lebih sempit dari metode Scheffe.
2. Metode Hybrid akan menghasilkan interval konfidensi simultan yang sempit dari metode Scheffe apabila mempunyai derajat bebas yang cukup besar, dan harga dari p lebih besar dibandingkan dari p_o.

Saran

Dengan banyaknya bentuk data yang bermacam-macam dan karakteristik yang berbeda-beda, maka alangkah baiknya dalam menentukan interval konfidensi simultan menggunakan metode Hybrid. Karena selain penggunaannya lebih luas, metode Hybrid ini bisa diterapkan dalam data kategorik maupun kontinu.

DAFTAR PUSTAKA

Thomas P. Lane and William H. DuMoushel, 1994, *Simultaneous Confidence ntervals in Multiple Regression*, The American Statistican, 48, 315-321.

Hogg, R.V. and Craig, A. , 1978, *Introduction to Mathematical Statistics*, Fourth Edition, London: The Dryden Press.

John A. Rice, 1995, *Matematical Statistics and Data Analisis*, Second Edition, University of California, Berkelay: Duxbury Press.

Raymond H. Myers, 1990, *Classical and Modern Regression with Applications*, Second Edition, Boston: PWS.

Lee J. Bain and Max Engelhardt, 1992, *Introduction to Probability and Mathematical Statistics*, Second Edition, California: Duxbury Press.

N.R. Draper and H. Smith, 1966, *Applied Regression Analisis*, Second Edition, Toronto: University of Wisconsin.

Charles. Hicks, 1982, *Fundamental Concepts in the Design of experiments*, Third Edition, Chicago: The Dryden Press.

Ruperts G. Miller, Jr, and McGraw-Hill, 1982, *Simultaneous Statistical Inference*, Toronto : Stanford University.

G.A.F Seber, 1977, *Linier Regression Analisis*, New York: John Wiley.

RK. Sebiring,1977, *Analisis Regresi*, Bandung: penerbit ITB.

John, Neter, William Wasserman and Michael H. Kutner, 1990, *Model Linier Terapan Analisis Regresi Ganda dan Sederhana,* Diterjemahkan oleh Bambang Sumantri, Bogor: penerbit IPB

J. Supranto, M.A. , 1994, *Statistik Teori dan Aplikasi*, Edisi Kelima, Jakarta: Penerbit Erlangga.

Walpole, R.E and Reymond H. Myers, 1995, *Ilmu Peluang dan Statistika untuk Insiyur dan Ilmuwan,* Diterjemahkan oleh R.K. Sembiring, Edisi Kedua, Bandung: Penerbit ITB.

Prof. Drs. Sutrisno Hadi , M.A. , 1988, *Statistik Jilid II,* Edisi 18, Yogyakarta, Penerbit: Andi Offset

John Netter, Wiliam Wasserman and Michael H. Katner, 1990, *Applied Linier Statistical Models in Regression, Analisis of Variance, and Experimental Designs,* Trird Edition, Boston: Irwin.

LAMPIRAN

3.2. Perbandingan dari Metode Scheffe dan Hybrid dalam Menentukan Interval Konfidensi Simultan 32
BAB IV: APLIKASI
4.1. Menduga Persamaan Regresi 40
4.2. Menentukan Interval Konfidensi Simultan 42
BAB V : KESIMPULAN DAN SARAN 46
DAFTAR PUSTAKA 47
LAMPIRAN 49

UNIVERSITAS ISLAM INDONESIA
PERPUSTAKAAN
F MIPA

TABEL A.4 (*lanjutan*) Persentil Sebaran *F*

db Pe-		db Pembilang								
sebut	*p*	10	12	15	20	24	30	60	120	∞
1	.50	2.04	2.07	2.09	2.12	2.13	2.15	2.17	2.18	2.20
	.90	60.2	60.7	61.2	61.7	62.0	62.3	62.8	63.1	63.3
	.95	242	244	246	248	249	250	252	253	254
	.975	969	977	985	993	997	1001	1010	1014	1018
	.99	6056	6106	6157	6209	6235	6261	6313	6339	6366
	.995	24224	24426	24630	24836	24940	25044	25253	25359	25464
	.999	605620	610670	615760	620910	623500	626100	631340	633970	636620
2	.50	1.34	1.36	1.38	1.39	1.40	1.41	1.43	1.43	1.44
	.90	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.48	9.49
	.95	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
	.975	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5	39.5
	.99	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5
	.995	199	199	199	199	199	199	199	199	200
	.999	999.4	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5
3	.50	1.18	1.20	1.21	1.23	1.23	1.24	1.25	1.26	1.27
	.90	5.23	5.22	5.20	5.18	5.18	5.17	5.15	5.14	5.13
	.95	8.79	8.74	8.70	8.66	8.64	8.62	8.57	8.55	8.53
	.975	14.4	14.3	14.3	14.2	14.1	14.1	14.0	13.9	13.9
	.99	27.2	27.1	26.9	26.7	26.6	26.5	26.3	26.2	26.1
	.995	43.7	43.4	43.1	42.8	42.6	42.5	42.1	42.0	41.8
	.999	129.2	128.3	127.4	126.4	125.9	125.4	124.5	124.0	123.5
4	.50	1.11	1.13	1.14	1.15	1.16	1.16	1.18	1.18	1.19
	.90	3.92	3.90	3.87	3.84	3.83	3.82	3.79	3.78	3.76
	.95	5.96	5.91	5.86	5.80	5.77	5.75	5.69	5.66	5.63
	.975	8.84	8.75	8.66	8.56	8.51	8.46	8.36	8.31	8.26
	.99	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.6	13.5
	.995	21.0	20.7	20.4	20.2	20.0	19.9	19.6	19.5	19.3
	.999	48.1	47.4	46.8	46.1	45.8	45.4	44.7	44.4	44.1
5	.50	1.07	1.09	1.10	1.11	1.12	1.12	1.14	1.14	1.15
	.90	3.30	3.27	3.24	3.21	3.19	3.17	3.14	3.12	3.11
	.95	4.74	4.68	4.62	4.56	4.53	4.50	4.43	4.40	4.37
	.975	6.62	6.52	6.43	6.33	6.28	6.23	6.12	6.07	6.02
	.99	10.1	9.89	9.72	9.55	9.47	9.38	9.20	9.11	9.02
	.995	13.6	13.4	13.1	12.9	12.8	12.7	12.4	12.3	12.1
	.999	26.9	26.4	25.9	25.4	25.1	24.9	24.3	24.1	23.8
6	.50	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.12
	.90	2.94	2.90	2.87	2.84	2.82	2.80	2.76	2.74	2.72
	.95	4.06	4.00	3.94	3.87	3.84	3.81	3.74	3.70	3.67
	.975	5.46	5.37	5.27	5.17	5.12	5.07	4.96	4.90	4.85
	.99	7.87	7.72	7.56	7.40	7.31	7.23	7.06	6.97	6.88
	.995	10.2	10.0	9.81	9.59	9.47	9.36	9.12	9.00	8.88
	.999	18.4	18.0	17.6	17.1	16.9	16.7	16.2	16.0	15.7
7	.50	1.03	1.04	1.05	1.07	1.07	1.08	1.09	1.10	1.10
	.90	2.70	2.67	2.63	2.59	2.58	2.56	2.51	2.49	2.47
	.95	3.64	3.57	3.51	3.44	3.41	3.38	3.30	3.27	3.23
	.975	4.76	4.67	4.57	4.47	4.42	4.36	4.25	4.20	4.14
	.99	6.62	6.47	6.31	6.16	6.07	5.99	5.82	5.74	5.65
	.995	8.38	8.18	7.97	7.75	7.65	7.53	7.31	7.19	7.08
	.999	14.1	13.7	13.3	12.9	12.7	12.5	12.1	11.9	11.7

TABEL A.4 (*lanjutan*) Persentil Sebaran *F*

db Penyebut		db Pembilang								
		1	2	3	4	5	6	7	8	9
8	.50	0.499	0.757	0.860	0.915	0.948	0.971	0.988	1.00	1.01
	.90	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.95	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.975	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.99	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.995	14.7	11.0	9.60	8.81	8.30	7.95	7.69	7.50	7.34
	.999	25.4	18.5	15.8	14.4	13.5	12.9	12.4	12.0	11.8
9	.50	0.494	0.749	0.852	0.906	0.939	0.962	0.978	0.990	1.00
	.90	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.95	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.975	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.99	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.995	13.6	10.1	8.72	7.96	7.47	7.13	6.88	6.69	6.54
	.999	22.9	16.4	13.9	12.6	11.7	11.1	10.7	10.4	10.1
10	.50	0.490	0.743	0.845	0.899	0.932	0.954	0.971	0.983	0.992
	.90	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.95	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.975	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.99	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.995	12.8	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
	.999	21.0	14.9	12.6	11.3	10.5	9.93	9.52	9.20	8.96
12	.50	0.484	0.735	0.835	0.888	0.921	0.943	0.959	0.972	0.981
	.90	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.95	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.975	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.99	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.995	11.8	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
	.999	18.6	13.0	10.8	9.63	8.89	8.38	8.00	7.71	7.48
15	.50	0.478	0.726	0.826	0.878	0.911	0.933	0.949	0.960	0.970
	.90	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.95	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.975	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.99	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.995	10.8	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
	.999	16.6	11.3	9.34	8.25	7.57	7.09	6.74	6.47	6.26
20	.50	0.472	0.718	0.816	0.868	0.900	0.922	0.938	0.950	0.959
	.90	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.95	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.975	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.99	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.995	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96
	.999	14.8	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
24	.50	0.469	0.714	0.812	0.863	0.895	0.917	0.932	0.944	0.953
	.90	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.95	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.975	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.99	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.995	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
	.999	14.0	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80

TABEL A.4 (*lanjutan*) Persentil Sebaran *F*

db Penyebut	A	db Pembilang: 10	12	15	20	24	30	60	120	∞
8	.50	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.08	1.09
	.90	2.54	2.50	2.46	2.42	2.40	2.38	2.34	2.32	2.29
	.95	3.35	3.28	3.22	3.15	3.12	3.08	3.01	2.97	2.93
	.975	4.30	4.20	4.10	4.00	3.95	3.89	3.78	3.73	3.67
	.99	5.81	5.67	5.52	5.36	5.28	5.20	5.03	4.95	4.86
	.995	7.21	7.01	6.81	6.61	6.50	6.40	6.18	6.06	5.95
	.999	11.5	11.2	10.8	10.5	10.3	10.1	9.73	9.53	9.33
9	.50	1.01	1.02	1.03	1.04	1.05	1.05	1.07	1.07	1.08
	.90	2.42	2.38	2.34	2.30	2.28	2.25	2.21	2.18	2.16
	.95	3.14	3.07	3.01	2.94	2.90	2.86	2.79	2.75	2.71
	.975	3.96	3.87	3.77	3.67	3.61	3.56	3.45	3.39	3.33
	.99	5.26	5.11	4.96	4.81	4.73	4.65	4.48	4.40	4.31
	.995	6.42	6.23	6.03	5.83	5.73	5.62	5.41	5.30	5.19
	.999	9.89	9.57	9.24	8.90	8.72	8.55	8.19	8.00	7.81
10	.50	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.06	1.07
	.90	2.32	2.28	2.24	2.20	2.18	2.16	2.11	2.08	2.06
	.95	2.98	2.91	2.84	2.77	2.74	2.70	2.62	2.58	2.54
	.975	3.72	3.62	3.52	3.42	3.37	3.31	3.20	3.14	3.08
	.99	4.85	4.71	4.56	4.41	4.33	4.25	4.08	4.00	3.91
	.995	5.85	5.66	5.47	5.27	5.17	5.07	4.86	4.75	4.64
	.999	8.75	8.45	8.13	7.80	7.64	7.47	7.12	6.94	6.76
12	.50	0.989	1.00	1.01	1.02	1.03	1.03	1.05	1.05	1.06
	.90	2.19	2.15	2.10	2.06	2.04	2.01	1.96	1.93	1.90
	.95	2.75	2.69	2.62	2.54	2.51	2.47	2.38	2.34	2.30
	.975	3.37	3.28	3.18	3.07	3.02	2.96	2.85	2.79	2.72
	.99	4.30	4.16	4.01	3.86	3.78	3.70	3.54	3.45	3.36
	.995	5.09	4.91	4.72	4.53	4.43	4.33	4.12	4.01	3.90
	.999	7.29	7.00	6.71	6.40	6.25	6.09	5.76	5.59	5.42
15	.50	0.977	0.989	1.00	1.01	1.02	1.02	1.03	1.04	1.05
	.90	2.06	2.02	1.97	1.92	1.90	1.87	1.82	1.79	1.76
	.95	2.54	2.48	2.40	2.33	2.29	2.25	2.16	2.11	2.07
	.975	3.06	2.96	2.86	2.76	2.70	2.64	2.52	2.46	2.40
	.99	3.80	3.67	3.52	3.37	3.29	3.21	3.05	2.96	2.87
	.995	4.42	4.25	4.07	3.88	3.79	3.69	3.48	3.37	3.26
	.999	6.08	5.81	5.54	5.25	5.10	4.95	4.64	4.48	4.31
20	.50	0.966	0.977	0.989	1.00	1.01	1.01	1.02	1.03	1.03
	.90	1.94	1.89	1.84	1.79	1.77	1.74	1.68	1.64	1.61
	.95	2.35	2.28	2.20	2.12	2.08	2.04	1.95	1.90	1.84
	.975	2.77	2.68	2.57	2.46	2.41	2.35	2.22	2.16	2.09
	.99	3.37	3.23	3.09	2.94	2.86	2.78	2.61	2.52	2.42
	.995	3.85	3.68	3.50	3.32	3.22	3.12	2.92	2.81	2.69
	.999	5.08	4.82	4.56	4.29	4.15	4.00	3.70	3.54	3.38
24	.50	0.961	0.972	0.983	0.994	1.00	1.01	1.02	1.02	1.03
	.90	1.88	1.83	1.78	1.73	1.70	1.67	1.61	1.57	1.53
	.95	2.25	2.18	2.11	2.03	1.98	1.94	1.84	1.79	1.73
	.975	2.64	2.54	2.44	2.33	2.27	2.21	2.08	2.01	1.94
	.99	3.17	3.03	2.89	2.74	2.66	2.58	2.40	2.31	2.21
	.995	3.59	3.42	3.25	3.06	2.97	2.87	2.66	2.55	2.43
	.999	4.64	4.39	4.14	3.87	3.74	3.59	3.29	3.14	2.97

TABEL A.4 (*lanjutan*) Persentil Sebaran *F*

db Penyebut	A	db Pembilang 1	2	3	4	5	6	7	8	9
30	.50	0.466	0.709	0.807	0.858	0.890	0.912	0.927	0.939	0.948
	.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.975	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.995	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
	.999	13.3	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
60	.50	0.461	0.701	0.798	0.849	0.880	0.901	0.917	0.928	0.937
	.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.975	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.995	8.49	5.80	4.73	4.14	3.76	3.49	3.29	3.13	3.01
	.999	12.0	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
120	.50	0.458	0.697	0.793	0.844	0.875	0.896	0.912	0.923	0.932
	.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.95	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
	.975	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.995	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
	.999	11.4	7.32	5.78	4.95	4.42	4.04	3.77	3.55	3.38
∞	.50	0.455	0.693	0.789	0.839	0.870	0.891	0.907	0.918	0.927
	.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	.975	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.995	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62
	.999	10.8	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10

TABEL A.4 (*lanjutan*) Persentil Sebaran *F*

db Penyebut	A	db Pembilang								
		10	12	15	20	24	30	60	120	∞
30	.50	0.955	0.966	0.978	0.989	0.994	1.00	1.01	1.02	1.02
	.90	1.82	1.77	1.72	1.67	1.64	1.61	1.54	1.50	1.46
	.95	2.16	2.09	2.01	1.93	1.89	1.84	1.74	1.68	1.62
	.975	2.51	2.41	2.31	2.20	2.14	2.07	1.94	1.87	1.79
	.99	2.98	2.84	2.70	2.55	2.47	2.39	2.21	2.11	2.01
	.995	3.34	3.18	3.01	2.82	2.73	2.63	2.42	2.30	2.18
	.999	4.24	4.00	3.75	3.49	3.35	3.22	2.92	2.76	2.59
60	.50	0.945	0.956	0.967	0.978	0.983	0.989	1.00	1.01	1.01
	.90	1.71	1.66	1.60	1.54	1.51	1.48	1.40	1.35	1.29
	.95	1.99	1.92	1.84	1.75	1.70	1.65	1.53	1.47	1.39
	.975	2.27	2.17	2.06	1.94	1.88	1.82	1.67	1.58	1.48
	.99	2.63	2.50	2.35	2.20	2.12	2.03	1.84	1.73	1.60
	.995	2.90	2.74	2.57	2.39	2.29	2.19	1.96	1.83	1.69
	.999	3.54	3.32	3.08	2.83	2.69	2.55	2.25	2.08	1.89
120	.50	0.939	0.950	0.961	0.972	0.978	0.983	0.994	1.00	1.01
	.90	1.65	1.60	1.55	1.48	1.45	1.41	1.32	1.26	1.19
	.95	1.91	1.83	1.75	1.66	1.61	1.55	1.43	1.35	1.25
	.975	2.16	2.05	1.95	1.82	1.76	1.69	1.53	1.43	1.31
	.99	2.47	2.34	2.19	2.03	1.95	1.86	1.66	1.53	1.38
	.995	2.71	2.54	2.37	2.19	2.09	1.98	1.75	1.61	1.43
	.999	3.24	3.02	2.78	2.53	2.40	2.26	1.95	1.77	1.54
∞	.50	0.934	0.945	0.956	0.967	0.972	0.978	0.989	0.994	1.00
	.90	1.60	1.55	1.49	1.42	1.38	1.34	1.24	1.17	1.00
	.95	1.83	1.75	1.67	1.57	1.52	1.46	1.32	1.22	1.00
	.975	2.05	1.94	1.83	1.71	1.64	1.57	1.39	1.27	1.00
	.99	2.32	2.18	2.04	1.88	1.79	1.70	1.47	1.32	1.00
	.995	2.52	2.36	2.19	2.00	1.90	1.79	1.53	1.36	1.00
	.999	2.96	2.74	2.51	2.27	2.13	1.99	1.66	1.45	1.00

Sumber: Disalin dari Tabel 5 dalam Pearson dan Hartley, *Biometrika Tables for Statisticians*, Volume 2, 1972, yang diterbitkan oleh the Cambridge University Press, atas nama The Biometrika Society, atas izin penulis dan penerbit.