

**MODUL AJAR**

**ANALISIS KUANTITATIF**



**Penulis:**  
**M. Rondhi, SP, MP, Ph.D**

**PROGRAM STUDI AGRIBISNIS**  
**JURUSAN SOSIAL EKONOMI FAKULTAS PERTANIAN**  
**UNIVERSITAS JEMBER**  
**SEMESTER GENAP, 2016/2017**

## **PENGANTAR**

Alhamdulillah penulisan modul ajar Analisis Kuantitatif dapat terselesaikan dengan ridlo Allah SWT. Penulisan modul ajar ini ditujukan bagi mahasiswa Fakultas Pertanian Program Studi Agribisnis yang menempuh mata kuliah Analisis kuantitatif pada semester genap.

Modul ini membahas beberapa model antara lain model regresi linear berganda, analisis regresi dengan variabel dependen yang bersifat terbatas (dengan kategori nominal dan ordinal), dan model dengan persamaan simultan. Modul ini juga dilengkapi dengan step-by-step aplikasi soft ekonometrik E-Views yang akan memberikan pengalaman bagi pembaca dalam mengaplikasikan software tersebut. Selain itu, modul ini juga dilengkapi dengan interpretasi hasil analisis yang dilakukan dengan menggunakan soft-ware SAS.

Namun demikian, materi dalam modul ini merupakan sebagian dari keseluruhan materi dalam analisis kuantitatif sehingga memiliki kekurangan. Penulisan beberapa materi akan disusun pada modul-modul mendatang.

Penulis

M. Rondhi

## DAFTAR ISI

Cover .....	i
Kata Pengantar .....	ii
Daftar Isi .....	iii
1. Analisis Regresi Linear Berganda .....	1
2. Uji Asumsi Klasik dalam Model Ordinary Least Square .....	10
3. Model Regresi Probabilitas Linear .....	24
4. Model Fungsi Distribusi Kumulatif .....	31
5. Model Logistik .....	46
6. Pendugaan Parameter pada Model Simultan .....	57
Daftar Bacaan	

**I.**  
**ANALISIS REGRESI LINEAR BERGANDA**

Kemampuan akhir yang diharapkan	:	Mahasiswa menganalisis analisis regresi dan menginterpretasikan
Kompetensi dasar	:	1. Mahasiswa memahami konsep dasar ekonometrik (regresi linear). 2. Mahasiswa mampu menggunakan regresi linear di bidang ekonomi pertanian 3. Mahasiswa mampu mengaplikasikan e-views
Metode Pembelajaran	:	- Metode SCL (Student Centered Learning) - Pendekatan: Case Study, Discovery Learning

### **1.1 Pendahuluan**

Sebagaimana dipahami bersama bahwa analisis regresi adalah analisis yang digunakan untuk mengetahui besarnya pengaruh perubahan variabel independent terhadap variabel dependen. Sebelum melakukan penggunaan EViews yang perlu dipahami adalah teori dari analisis regresi itu sendiri seperti uji F, uji-t, tingkat signifikansi, dan uji asumsi klasik. EViews mempermudah dalam mencari hasil-hasil tersebut. Selain itu, pengguna harus mempelajari arti secara ekonomi atau menurut kebutuhannya. Berikut dijelaskan sedikit gambaran analisis regresi berganda dan uji asumsi klasik.

### **2.2 Definisi**

- Analisis Regresi Linear Berganda : analisis regresi yang jumlah variable independennya dua atau lebih.

- Analisis regresi dilakukan untuk mengetahui pengaruh perubahan variable independen terhadap variable dependen jika variable independennya lebih dari 2.
- Persamaan umum regresi linear berganda adalah  

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n + e$$
Dimana:  
Y = Variabel dependen  
a = konstanta  
 $b_1, b_2, b_3, b_n$  = parameter untuk variabel  $X_1, X_2, X_3, X_n$   
 $X_1$  = variabel  $X_1$   
 $X_2$  = variabel  $X_2$   
 $X_3$  = variabel  $X_3$   
e = error, kesalahan
- Analisis ini digunakan untuk meramalkan tentang variable independen pada masa yang akan datang atau untuk mengetahui gejala umum yang terjadi pada sebuah peristiwa.
- Contoh.  
Untuk mengetahui faktor-faktor yang berpengaruh terhadap konvergensi atau divergensi di Jawa Timur dari tahun 1997 – 2007.

### 1.3 Hal-hal Penting yang Ada Dalam Analisis Regresi Linear Berganda

#### 1) Koefisien determinasi ( $R^2$ )

Nilai ini digunakan untuk mengetahui ketepatan model yang digunakan yang dinyatakan dengan berapa persen variasi variabel dependen dijelaskan oleh variasi variabel independen yang dimasukkan ke dalam model regresi. Model dianggap baik

apabila koefisien determinasi sama dengan satu atau mendekati satu (Gujarati, 1997).

Koefisien determinasi diformulasikan sebagai berikut :

$$R^2 = (ESS/TSS) = 1 - (RSS/TSS)$$

Keterangan :  $R^2$  = nilai koefisien determinasi,

ESS = explained sum of squares,

TSS = total sum of squares,

RSS = residual sum of squares

## 2) Uji F (*over all test*) →

Uji F digunakan untuk mengetahui tingkat pengaruh semua variabel independen secara bersama-sama terhadap variabel dependen.

Adapun rumus  $F_{hitung}$  adalah

$$F_{hitung} = \frac{[(R^2) / (k - 1)]}{[(1 - R^2) (n - k)]}$$

$$F_{tabel} = [ (k - 1) ; (n - k) ; \alpha ]$$

Dimana :

$R^2$  = koefisien determinasi

k = banyaknya koefisien regresi (termasuk intersep)

n = banyaknya sampel

Formula hipotesis yang akan diuji adalah:

$$H_0 : \beta_i = 0$$

$$H_a : \beta_i \neq 0$$

Jika  $F_{hitung} > F_{tabel}$  maka  $H_0$  ditolak pada tingkat kesalahan tertentu, artinya bahwa variabel independen yang diuji secara bersama-sama berpengaruh nyata terhadap variasi variabel dependen

### 3). Uji terhadap penduga parameter (*t test*)

Uji t digunakan untuk mengetahui pengaruh masing-masing variabel independent terhadap variabel dependen

Adapun rumus  $t_{hitung}$  adalah

$$T_{hitung} = \frac{\beta_i}{Se(\beta_i)}$$

$$T_{tabel} = (n - k ; \alpha / 2)$$

Dimana :  $\beta_i$  = koefisien regresi yang diestimasi

$Se$  = *Standard error* koefisien yang diestimasi

Sedangkan formula hipotesis yang diuji adalah sebagai berikut:

$H_0$  :  $\beta_i = 0$ , artinya tidak ada pengaruh variabel independen ke-i ( $X_i$ ) terhadap variable dependen.

$H_a$  :  $\beta_i \neq 0$ , artinya ada pengaruh variabel independen ke-i ( $X_i$ ) terhadap variable dependen.

Jika  $t_{hitung} > t_{tabel}$ , maka  $H_0$  ditolak pada tingkat kesalahan tertentu, artinya bahwa variabel independen ke-i berpengaruh nyata terhadap variabel dependen.

- Nilai a (konstanta) dan b (nilai parameter) masing-masing variabel independen.
- Tingkat signifikansi/tingkat probabilitas: Probabilitas dimana kesalahan dapat diterima. Tingkat probabilitas umum yang dapat diterima adalah : 1%, 5%, 10%

#### 1.4 Aplikasi.

Sebuah penelitian untuk mengetahui faktor-faktor yang berpengaruh terhadap tingkat konvergensi di Kabupaten Sidoarjo, dengan data time series dari tahun 1999 – 2007 seperti terdapat dalam tabel 1 di atas.

Model yang terbentuk adalah

$$\text{Konvergensi} = a + b_1 \text{ IPM} + b_2 \text{ PDRB} + b_3 \text{ Manufaktur}$$

Data yang tersedia adalah sebagai berikut.

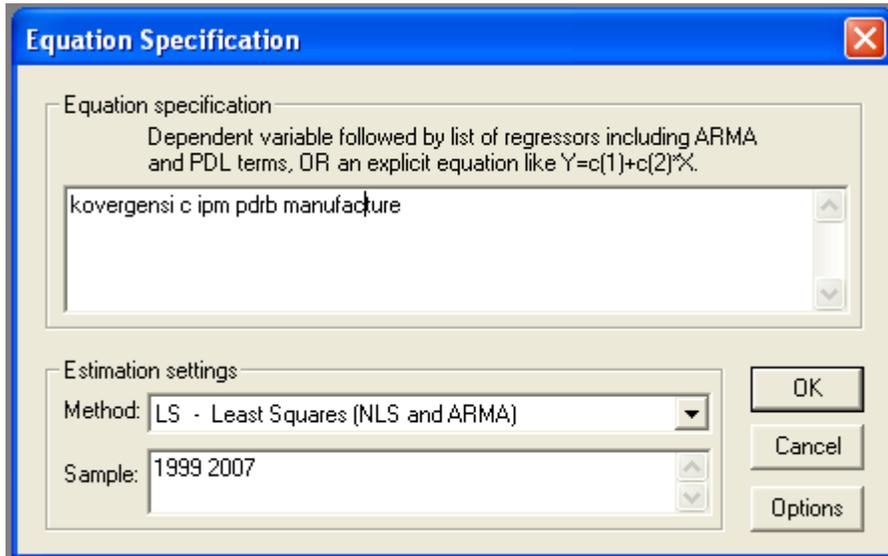
**Contoh.**

Tabel 1.1 Data Indeks Pembangunan Manusia, PDRB sumbangan Sektor Manufaktur dan Tingkat Konvergensi

Tahun	IPM	PDRB	MANUFACTURE	Konvergensi
1999	62	8049850	14.04	0
2000	63.97	9714722	14.01	0.303
2001	63.88	8438443	16.74	-0.303
2002	69.3	10400370	16.28	-0.009
2003	69.38	10791298	16.26	0.0004
2004	70.25	11309427	16.21	0.0011
2005	68.55	11971715	16.2	-0.0002
2006	71.55	12655286	15.78	0.003
2007	66.92	13380947	15.66	0.303

Proses analisis:

Masukkan data ke perintah regresi dengan perintah Quick – Estimate Equation. Selanjutnya pada jendela Equation specification cantumkan nama-nama variabel dalam model. Variabel dependent harus ditulis paling depan diikuti variabel independen termasuk konstanta. Penulisan dilakukan seperti contoh berikut.



Gambar 2.1 Memasukkan rumus

Pada kotak Method pilih LS – Least Squares, karena metode yang digunakan adalah ordinary Least. Klik OK untuk menampilkan hasilnya. Jika proses anda benar, maka hasil yang didapatkan adalah sebagai berikut.

Dependent Variable: KONVERGENSI  
Method: Least Squares  
Date: 05/01/10 Time: 09:39  
Sample: 1999 2007  
Included observations: 9

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.963868	0.880771	2.229715	0.0762
IPM	-0.010756	0.019535	-0.550610	0.6056
PDRB	8.37E-08	3.03E-08	2.764672	0.0396
MANUFACTURE	-0.134231	0.046526	-2.885088	0.0344

R-squared	0.804174	Mean dependent var	0.033144
Adjusted R-squared	0.686679	S.D. dependent var	0.182219
S.E. of regression	0.101997	Akaike info criterion	-1.426635
Sum squared resid	0.052017	Schwarz criterion	-1.338980
Log likelihood	10.41986	F-statistic	6.844311
Durbin-Watson stat	2.130369	Prob(F-statistic)	0.032052

Gambar 1.1. Hasil analisis dengan Metode OLS

Berikan nama untuk hasil analisis anda. Untuk keseragaman tuliskan hasil1.

Berdasarkan hasil analisis tersebut, terdapat beberapa hal penting terkait dengan hasil analisis tersebut.

### Nilai $R^2$

Nilai  $R^2$  sebesar 0,804, artinya pengaruh perubahan variabel variabel penjelas (IPM, PDRB dan sektor manufaktur) terhadap variabel konvergensi sebesar 80,4%, sedangkan sisanya sebesar 19,6% dipengaruhi oleh variabel lainnya di luar model.

### **Nilai F-hitung**

Berdasarkan hasil analisis didapatkan nilai F-hitung sebesar 6,844 dengan f-tabel tingkat signifikansi 5% sebesar 2,84 dengan demikian diketahui bahwa secara bersama-sama variabel independent berpengaruh terhadap variabel konvergensi di Kabupaten Pasuruan.

### Nilai t-hitung

Nilai t-statistik untuk masing-masing variabel independent (IPM, PDRB dan manufacture) masing-masing berturut-turut adalah -0,55; 2,76; dan -2,88 (diabsolutkan untuk uji dua-sisi). Nilai tersebut dibandingkan dengan t-tabel pada tingkat signifikansi 5% sebesar 1,76. Berdasarkan hasil analisis tersebut diketahui bahwa variabel PDRB dan manufaktur berpengaruh terhadap konvergensi di Kabupaten Jember, sedangkan variabel IPM tidak berpengaruh terhadap konvergensi di Kabupaten Jember.

### **Arah Pengaruh**

Untuk mengetahui arah pengaruh digunakan tanda dari koefisien regresi. Berdasarkan hasil analisis diketahui bahwa arah koefisien regresi masing-masing variabel (IPM, PDRB, manufaktur) adalah sebagai berikut -0,01; 0,00; -0,13. Arah positif berarti semakin tinggi perubahan variabel independen akan meningkatkan tingkat konvergensi, sebaliknya nilai koefisien regresi negative berarti semakin tinggi

perubahan variabel independen akan menurunkan tingkat konvergensi di Kabupaten Jember.

### **Tugas**

Berikut adalah data faktor-faktor yang berpengaruh terhadap permintaan karet alam Indonesia

Tabel 2.2 Faktor-faktor Permintaan Karet Alam Indonesia

NO	DNR	DSR	SNR	PNR
1	1400	2200	1310	2585.98245
2	1410	2230	1240	2588.76644
3	1400	2150	1570	2221.93342
4	1440	2200	1590	2463.13598
5	1470	2320	1460	2535.32379
6	1490	2390	1330	2275.47128
7	1460	2230	1510	2264.22688
8	1490	2300	1560	2308.12946
9	1540	2430	1610	2279.29313
10	1500	2410	1380	2086.42873
11	1530	2370	1650	2237.15562
12	1560	2390	1700	2347.00168
13	1650	2450	1680	2317.12357
14	1650	2500	1400	2269.42925
15	1620	2490	1640	2544.90957

Di mana:

DNR : Permintaan karet alam Indonesia

DSR : Permintaan karet sintesis dunia

SNR : Penawaran karet alam Indonesia

PNR : Harga karet alam Indonesia

1. Lakukan analisis regresi linear berganda.
2. Lakukan interpretasi dari hasil analisis tersebut.

## II. UJI ASUMSI KLASIK DALAM MODEL ORDINARY LEAST SQUARE

Standar Kompetensi	:	Mahasiswa dapat mempraktekkan Aplikasi E-Views untuk Pengujian Asumsi Klasik dalam OLS ( <i>Ordinary Least Square</i> )
Kompetensi dasar	:	<ol style="list-style-type: none"> <li>1. Mahasiswa memahami konsep dasar ekonometrik (uji asumsi klasik).</li> <li>2. Mahasiswa mampu menggunakan regresi linear di bidang ekonomi pertanian dan mampu melakukan uji asumsi klasik.</li> <li>3. Mahasiswa mampu mengaplikasikan e-views untuk uji asumsi klasik.</li> </ol>
Metode Pembelajaran	:	<ul style="list-style-type: none"> <li>- Metode SCL (Student Centered Learning)</li> <li>- Pendekatan: Case Study, Discovery Learning</li> </ul>

### 2.1 Pendahuluan

Untuk memperoleh validitas hasil pengujian ekonometrik dengan model Ordinary Least Square, maka perlu dilakukan pendeteksian penyimpangan dari asumsi-asumsi klasik dan terhadap kesesuaian model (Pindyck and Rubinfeld, 1991; Maddala, 1992; Green, 1993; Gujarati, 1997). Pengujian terhadap asumsi klasik ditujukan untuk mengetahui apakah koefisien regresi estimasi merupakan penaksir tak bias yang terbaik (Best Linear Unbiased Estimator BLUE).

#### a. Uji Asumsi *Multikolinieritas*

Uji *multikolinieritas* dimaksudkan untuk mengetahui apakah terdapat suatu hubungan linear yang kuat antar variabel bebas yang terdapat dalam suatu model. Dengan demikian antar variabel bebas tidak ortogonal artinya variabel tersebut memiliki korelasi yang tidak sama dengan nol (Arief, 1992).

Jika terdapat korelasi erat antar sesama variabel bebas, maka konsekuensinya:

1. koefisien regresi menjadi tidak dapat ditaksir, atau koefisien regresi menjadi tidak signifikan.
2. Adanya multikoleniaritas dapat menyebabkan tanda koefisien regresi menjadi berlawanan dengan yang diramalkan secara teoritis.

### **Pendeteksian**

Untuk mendeteksi adanya hubungan linear tersebut digunakan uji korelasi dengan rumus sebagai berikut.

$$r = \frac{\left\{ n \sum_{i=1}^n (X_1 \cdot X_2) \right\} - \left\{ \left( \sum_{i=1}^n X_1 \right) \left( \sum_{i=1}^n X_2 \right) \right\}}{\left\{ \left( n \sum_{i=1}^n X_1^2 \right) - \left( \sum_{i=1}^n X_1 \right)^2 \right\} \left\{ \left( n \sum_{i=1}^n X_2^2 \right) - \left( \sum_{i=1}^n X_2 \right)^2 \right\}}$$

$r$  = r-hitung/korelasi

$X_1$  = variabel  $X_1$

$X_2$  = variabel  $X_2$

Kriteria Pengambilan Keputusan:

Jika nilai korelasi  $\geq 0,8$  maka berarti terdapat *multikolinearitas* yang serius antar dua variabel tersebut.

### **Cara mengatasi Masalah Multikolinearitas.**

#### **1. Metode Koutsoyiannis (1977)**

Metode ini menjelaskan cara untuk memperbaiki model yang terdapat gejala multikolinearitas. Cara yang digunakan adalah melakukan regresi variabel independen terhadap variabel dependen yang secara apriori berpengaruh. Selanjutnya, memasukkan variabel-variabel independen satu-persatu ke dalam persamaan. Hasil analisis tersebut kita amati koefisien regresinya, standart error dan  $R^2$ . Variabel bebas yang baru dimasukkan dalam persamaan dapat diklasifikasikan sebagai variabel

bebas yang berguna (*useful*), tidak perlu (*superflous*), dan merusak hasil (*detrimental*).

1. *Useful*. Jika variabel baru yang dimasukkan dalam model menyebabkan nilai  $R^2$  menjadi lebih bagus dan tanda signifikansi koefisien regresi sesuai dengan teori.

2. *Superflous*. Jika variabel baru yang dimasukkan dalam model menyebabkan nilai  $R^2$  tidak menjadi lebih bagus dan tanda signifikansi koefisien regresi tidak sesuai dengan teori.

3. *Detriemental*. Jika variabel baru yang dimasukkan dalam model menyebabkan tanda koefisien regresi tidak sesuai dengan teori.

## **2. Mentransformasikan Variabel-variabel**

Mentransformasikan variabel berarti merubah bentuk variabel ke dalam bentuk tertentu. Cara yang digunakan dapat dengan menggunakan bentuk logaritma atau dengan first difference.

### **Cara bentuk logaritma.**

Nilai variabel independen sebesar 100, maka bentuk logartimanya sama dengan  $\log(100) = 2$

### **Cara first difference (lebih banyak digunakan pada data time series)**

1. 200

2. 300

Maka first differennya menjadi  $(300 - 200) = 100$

## **3. memperbanyak Data**

Cara yang lainnya adalah dengan memperbanyak data. Dengan memperbesar jumlah data, standart error cenderung turun yang akan memungkinkan kita dapat menaksir koefisien regresi secara lebih tepat.

### 3.3 Uji Asumsi *Heteroskedastisitas*

Uji *Heteroskedastisitas* apakah model yang dibangun memiliki unsur pengganggu (*disturbance*) atau *error term* ( $\mu$ ) yang konstan (*homoskedastis*), yang dilambangkan dengan  $E(\mu^2) = \sigma^2$ . Jika *error term* ( $\mu$ ) tidak konstan, dalam hal ini dilambangkan dengan  $E(\mu^2) = \sigma_i^2$ , berarti telah terjadi pelanggaran terhadap salah satu asumsi klasik pada OLS yaitu model mengandung *heteroskedastisitas*.

Kebanyakan data cross-section mengandung situasi heteroskedastic karena data ini menghimpun data yang mewakili berbagai ukuran (kecil, sedang dan besar). Contoh pengeluaran keluarga yang berpenghasilan tinggi akan lebih bervariasi dibandingkan dengan keluarga yang berpendapatan rendah.

Situasi heteroskedastik akan menyebabkan penaksiran koefisien-koefisien regresi menjadi tidak efisien. Hasil taksiran menjadi kurang semestinya, lebih dari semestinya atau menyesatkan.

#### Cara Mendeteksi

##### 1. Metode Park

Park (1966) mengemukakan metode berikut. Asumsikan bahwa  $\sigma_i^2$  merupakan fungsi dari variabel bebas  $\sigma_i^2 = \alpha X_i^{\beta}$ . Selanjutnya persamaan tersebut dibuat dalam bentuk persamaan linear sebagai berikut:

$$\ln \sigma_i^2 = \alpha + \beta \ln X_i + v_i$$

Oleh karena  $\sigma_i^2$  umumnya tidak diketahui, maka ini dapat ditaksir dengan  $e_i^2$ .

Metode Park mengandung prosedur dua tahap:

1. Melakukan regresi terhadap model tanpa mempersoalkan apakah situasi heteroskedastisitas ada atau tidak.

$$\text{Contoh: } Y_i = \beta_0 + \beta_1 X_i + \mu_i$$

$$\text{Dari regresi tersebut diperoleh } (\hat{Y} - Y_i) = e_i$$

2. Selanjutnya melakukan regresi  $X_i$  terhadap  $e_i^2$  seperti persamaan berikut.

$$\ln e_i^2 = \alpha_0 + \beta_1 \ln X_i + v_i$$

Kriteria pengambilan keputusan:

Jika secara statistik  $X_i$  signifikan terhadap  $\ln e_i^2$ , maka disimpulkan bahwa terdapat gejala heteroskedastisitas dalam persamaan atau model yang terbentuk

## 2. Metode Glesjer

Pada dasarnya metode Glesjer hampir sama dengan metode Park. Akan tetapi yang diregresikan untuk tahap kedua adalah dengan menggunakan nilai absolut dari  $e$ .

Menurut Arif, 1993, menyebutkan bahwa Glesjer membuat model alternatif sebagai berikut.

$$(1) |e_i| = \alpha_1 + \alpha_2 X_i + v_i$$

$$(2) |e_i| = \alpha_1 + \alpha_2 \sqrt{X_i} + v_i$$

$$(3) |e_i| = \alpha_1 + \alpha_2 1/X_i + v_i$$

$$(4) |e_i| = \alpha_1 + \alpha_2 1/\sqrt{X_i} + v_i$$

$$(5) |e_i| = \sqrt{\alpha_1 + \alpha_2 X_i} + v_i$$

$$(6) |e_i| = \sqrt{\alpha_1 + \alpha_2 X_i^2} + v_i$$

Model yang sering digunakan adalah model 1.

Kriteria pengambilan keputusan:

Jika secara statistik  $X_i$  signifikan terhadap  $\ln e_i^2$ , maka disimpulkan bahwa terdapat gejala heteroskedastisitas dalam persamaan atau model yang terbentuk

## 3. Metode Spearman Rank Correlation

Langkah yang dilakukan:

1. Dari hasil regresi suatu model regresi, perolehlah nilai-nilai residual ( $e_i$ )
2. Hitung koefisien korelasi masing-masing variabel bebas dengan  $e_i$  yang dihasilkan dalam langkah 1. Korelasi yang digunakan adalah korelasi Spearman Rank dengan formulasi sebagai berikut.

$$r' = 1 - \frac{6 \sum Di^2}{N(N^2 - 1)}$$

Di mana:

Di = perbedaan ranking residual dengan ranking variabel bebas

N = jumlah observasi dalam sampel.

Kriteria pengambilan keputusan:

Nilai  $r'$  yang tinggi menunjukkan adanya gejala heteroskedastisitas dalam model tersebut.

### **Cara Mengatasi Masalah Heteroskedastisitas**

1. Melakukan transformasi dalam bentuk membagi model regresi asal dengan salah satu variabel bebas yang digunakan dalam model.

Contoh:

Model Asal:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \mu_i$$

Mentrasnformasikan bentuk awal dengan membagi dengan variable  $X_i^2$

Menjadi persamaan berikut.

$$Y_i/X_{1i} = \beta_0/X_{1i} + \beta_1 + \dots + \beta_k X_{ki} / X_{1i} + \mu_i/X_{1i}$$

2. Melakukan transformasi logaritma.

Model asal:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \mu_i$$

Model transformasi logaritma:

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \dots + \beta_k \ln X_{ki} + \mu_i$$

Transformasi logartma akan mengirangi situasi heteroskedastisitas karena transformasi logaritma memperkecil skala ukuran variabel.

### 3.4 Uji Sumsi Korelasi Serial antar Error Terms

Korelasi sereal antar error term berarti adanya korelasi yang serius antara error pengamatan satu dengan error pengamatan yang lain.

Penyebab serial korelasi:

1. Adanya momentum dari situasi yang menyebabkan data mengalami kecenderungan naik atau kecenderungan turun.

Contoh : adanya krisis ekonomi (resesi ekonomi)

2. Tidak memasukkan variabel bebas tertentu yang sebetulnya mempengaruhi dependent variable.

Contoh:

$$Y_i = \beta_0 + \beta_1 P_1 + \beta_2 C_t + \beta_3 P_2 + e$$

Ditulis dengan

$$Y_i = \beta_0 + \beta_1 P_1 + \beta_2 C_t + e$$

3. Bentuk model yang tidak tepat.

Bentuk persamaan marginal cost adalah:

$$MC = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + e$$

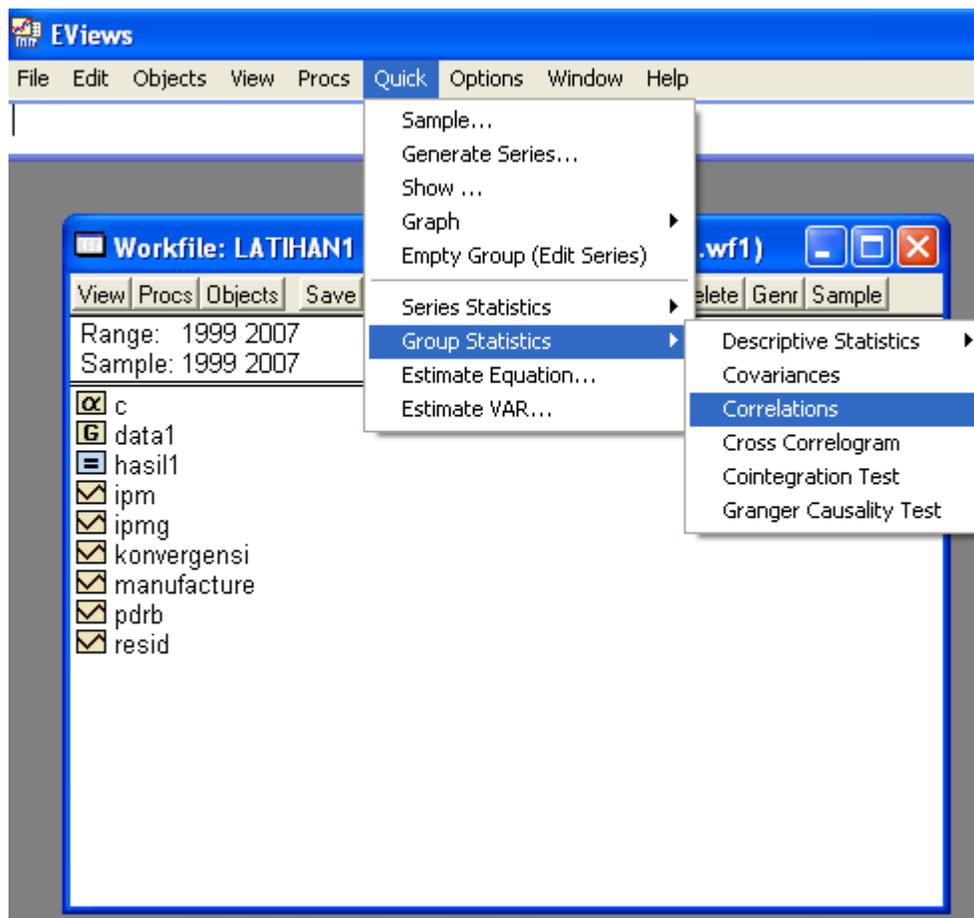
Akan tetapi kita membuat model

$$MC = \beta_0 + \beta_1 X_1 + e$$

Aplikasi

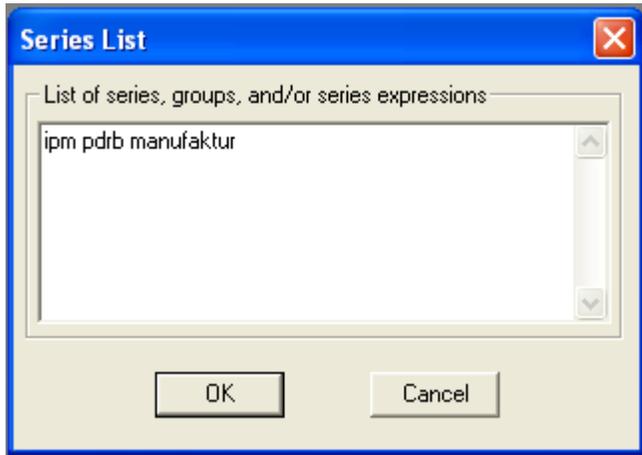
## 1. Uji Multikolinearitas

Untuk menguji multikolinearitas, langkah yang digunakan adalah sebagai berikut. Dari menu Quick → Group Statitic → correlation, sebagaimana dalam gambar berikut.



Gambar 3.1. Uji Multikolinearitas

Setelah tuliskan variabel independen dalam model yang anda bentuk seperti dalam gambar berikut.



Gambar 3.2. Memasukkan variabel dalam uji Multikolinearitas

Jika proses anda benar, maka hasil anda sebagai berikut.

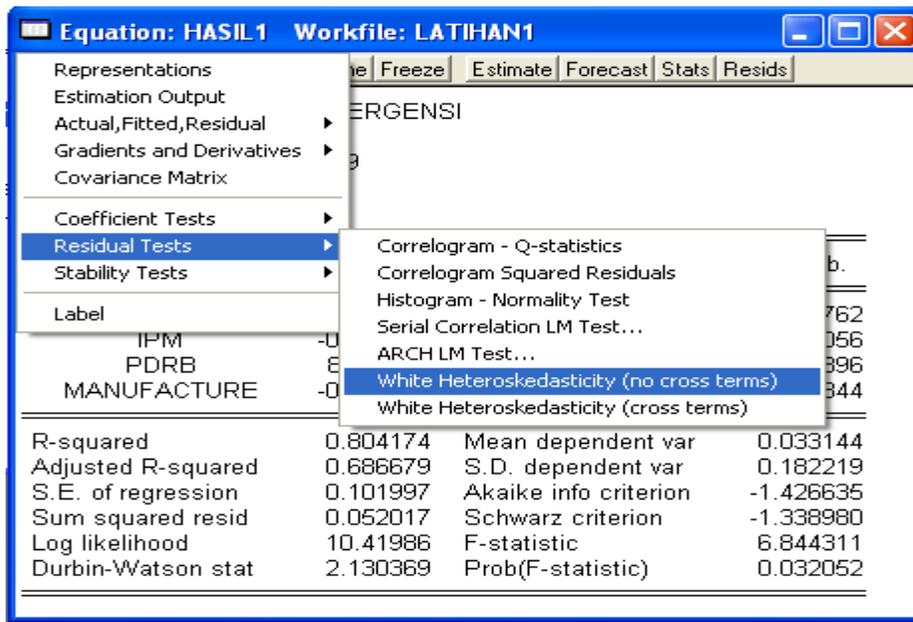
	IPM	PDRB	MANUFAC...
IPM	1.000000	0.737256	0.597312
PDRB	0.737256	1.000000	0.320007
MANUFAC...	0.597312	0.320007	1.000000

Gambar 3.3. Hasil uji multikolinearitas

Kesimpulan: karena nilai korelasi antar variabel indenpenen dibawah 0,8, maka disimpulkan bahwa dalam model tersebut tidak terdapat gejala multokolinearitas.

## 2. Heteroskedastisitas

Langkah untuk menguji heteroskedastisitas adalah sebagai berikut. Tampilkan hasil analisis seperti pada gambar 14. Klik View → Residual Test → White Heteroskedasticity (no cross term) seperti pada gambar berikut.



Gambar 3.4. Langkah uji Heteroskedastisitas

Jika proses anda benar, maka hasilnya sebagai berikut.

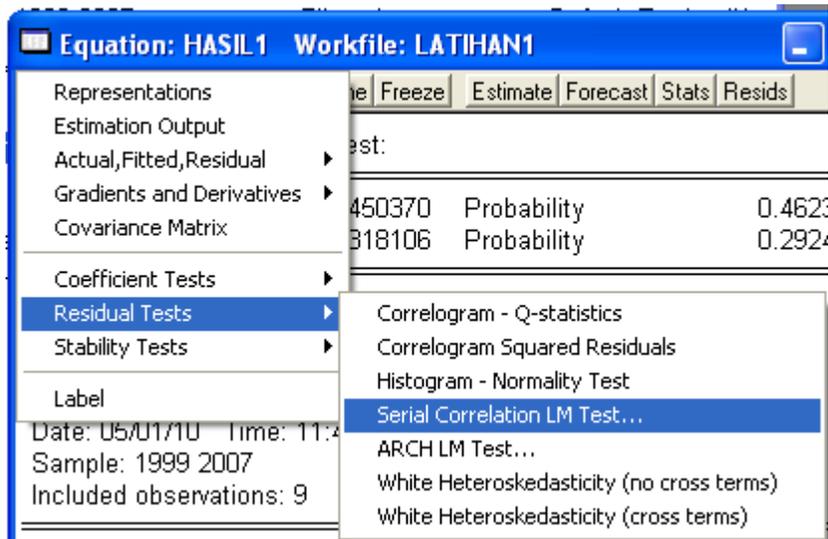
EViews - [Equation: HASIL1 Workfile: LATIHAN1]				
File Edit Objects View Procs Quick Options Window Help				
View Procs Objects Print Name Freeze Estimate Forecast Stats Resids				
White Heteroskedasticity Test:				
F-statistic	1.450370	Probability	0.462389	
Obs*R-squared	7.318106	Probability	0.292428	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 05/01/10 Time: 11:47				
Sample: 1999 2007				
Included observations: 9				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.716594	3.044164	0.563897	0.6296
IPM	0.269867	0.220614	1.223258	0.3458
IPM^2	-0.001891	0.001570	-1.204969	0.3514
PDRB	-1.97E-07	1.34E-07	-1.468852	0.2796
PDRB^2	8.93E-15	6.03E-15	1.479933	0.2770
MANUFACTURE	-1.316726	0.857620	-1.535325	0.2645
MANUFACTURE^2	0.042227	0.027568	1.531745	0.2653
R-squared	0.813123	Mean dependent var	0.005780	
Adjusted R-squared	0.252491	S.D. dependent var	0.005307	
S.E. of regression	0.004588	Akaike info criterion	-7.879203	
Sum squared resid	4.21E-05	Schwarz criterion	-7.725806	
Log likelihood	42.45642	F-statistic	1.450370	
Durbin-Watson stat	2.752569	Prob(F-statistic)	0.462389	

Gambar 3.5. Uji Heteroskedastisitas

Berdasarkan hasil analisis diketahui bahwa tidak ada variabel x (independen) yang berpengaruh terhadap error. Dengan demikian tidak ada gejala heteroskedastisitas dalam model penelitian ini.

### 3. Autokorelasi

Langkah untuk menguji heteroskedastistas adalah sebagai berikut. Tampilkan hasil analisis seperti pada gambar 14. Klik View → Residual Test → Serial Correlation test seperti pada gambar berikut.



Masukkan lag 2 seperti berikut.



Gambar 3.6. Langkah uji Autokorelasi

Jika proses anda benar, maka hasil yang anda dapatkan adalah sebagai berikut.

EViews - [Equation: HASIL1 Workfile: LATIHAN1]

File Edit Objects View Procs Quick Options Window Help

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.128525	Probability	0.883986
Obs*R-squared	0.710290	Probability	0.701072

Test Equation:  
 Dependent Variable: RESID  
 Method: Least Squares  
 Date: 05/01/10 Time: 11:52  
 Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.386122	1.719909	-0.224502	0.8368
IPM	0.002881	0.035046	0.082206	0.9397
PDRB	-1.81E-08	5.97E-08	-0.303569	0.7813
MANUFACTURE	0.024477	0.076810	0.318664	0.7709
RESID(-1)	-0.496234	1.031963	-0.480864	0.6635
RESID(-2)	0.062962	0.987593	0.063753	0.9532

R-squared	0.078921	Mean dependent var	-1.76E-16
Adjusted R-squared	-1.456210	S.D. dependent var	0.080636
S.E. of regression	0.126375	Akaike info criterion	-1.064401
Sum squared resid	0.047912	Schwarz criterion	-0.932918
Log likelihood	10.78980	F-statistic	0.051410
Durbin-Watson stat	2.078083	Prob(F-statistic)	0.996537

Gambar 3.7. Hasil uji Autokorelasi

Kesimpulan: karena nilai Obs\*R-squared < nilai Chi-square dan nilai signifikasni dari Obs\*R-squared > 0,05, disimpulkan bahwa tidak ada gejala autokorelasi dalam model.

## Tugas

Tabel 3.1 Berikut adalah data faktor-faktor yang berpengaruh terhadap karet alam Indonesia

NO	DNR	DSR	SNR	PNR
1	1400	2200	1310	2585.98245
2	1410	2230	1240	2588.76644
3	1400	2150	1570	2221.93342
4	1440	2200	1590	2463.13598
5	1470	2320	1460	2535.32379
6	1490	2390	1330	2275.47128
7	1460	2230	1510	2264.22688
8	1490	2300	1560	2308.12946
9	1540	2430	1610	2279.29313
10	1500	2410	1380	2086.42873
11	1530	2370	1650	2237.15562
12	1560	2390	1700	2347.00168
13	1650	2450	1680	2317.12357
14	1650	2500	1400	2269.42925
15	1620	2490	1640	2544.90957

Di mana:

DNR : Permintaan karet alam Indonesia

DSR : Permintaan karet sintesis dunia

SNR : Penawaran karet alam Indonesia

PNR : Harga karet alam Indonesia

## Pertanyaan

1. Lakukan Uji multikolinearitas, heteroskedastisitas dan autokorelasi.
2. Jika ada gejala seperti di atas, bagaimana cara memperbaiki model tersebut.

**III.**  
**MODEL REGRESI PROBABILITAS LINEAR**  
**(LINEAR PROBABILITY MODEL)**

Standar Kompetensi	:	Mahasiswa dapat mempraktekkan Aplikasi E-Views untuk menganalisis model Probabilitas Linear
Kompetensi dasar	:	<ol style="list-style-type: none"> <li>1. Mahasiswa memahami konsep dasar ekonometrik model regresi probabilitas Linear.</li> <li>2. Mahasiswa mampu menggunakan model regresi probabilitas linear di bidang ekonomi pertanian.</li> <li>3. Mahasiswa mampu mengaplikasikan e-views untuk pengujian model regresi probabilitas linear.</li> </ol>
Metode Pembelajaran	:	<ul style="list-style-type: none"> <li>- Metode SCL (Student Centered Learning)</li> <li>- Pendekatan: Case Study, Discovery Learning</li> </ul>

**4.1 Pendahuluan**

Dalam penjelasan sebelumnya model yang digunakan adalah model dengan variabel dependent (Y) dan independent (X) berupa data rasio dan juga kadang data kontinyu. Akan tetapi dalam kenyataannya banyak penelitian yang variabel dependennya berupa data diskrit dan terkadang juga nominal. Sebagai konsekuensinya, penggunaan analisis regresi linear tidaklah tepat.

Oleh karena itu, beberapa ahli ekonometrik (Gujarati, 2001; Green, 2006; Kutsosianis, 2001) memperkenalkan model dengan variabel dependent berbentuk kategori yang dikenal dengan istilah model probabilitas linear. Perdefinitif, model ini berarti model dengan probabilitas linear. Secara detail, berikut dibahas tentang model, contoh dan penyelesaian model tersebut.

## 4.2 Model

Menurut Arief (1992), model probabilitas linear dapat dijelaskan dengan contoh permodelan sebagai berikut.

$$Y_i = \alpha + \beta X_i + \varepsilon \dots\dots\dots(1)$$

Variabel  $X_i$  adalah variabel independen

Variabel  $Y_i$  adalah variabel dependen

$\alpha$  = konstanta

$\beta$  = koefisien parameter

$i$  = data ke- $i$

$\varepsilon$  = error

Berdasarkan sifat kualitatif dari dependent variable, maka dapat dinyatakan bahwa:

$$\begin{aligned} E(Y_i) &= 1 \times P_i + 0 \times (1 - P_i) \dots\dots\dots(2) \\ &= P_i \end{aligned}$$

Di mana  $P_i$  adalah probabilitas variabel  $Y$  dengan kemunculan angka 1. 0 adalah kemunculan data untuk variabel  $Y = 0$

Dari persamaan 1 dan 2 didapatkan persamaan berikut.

$$\begin{aligned} E(Y_i) &= \alpha + \beta X_i + \varepsilon \dots\dots\dots(3) \\ &= P_i \end{aligned}$$

Variabel  $Y$  yang berbentuk kualitatif berimplikasi pada bentuk penyebaran errornya (*error term distribution*). Arief (1992) menjelaskan bahwa untuk suatu nilai  $X$  tertentu ( $X_i$ ), jika  $Y_i=1$ , maka  $\varepsilon_i = 1 - \alpha - \beta X_i$ , dan jika  $Y_i = 0$ , maka  $\varepsilon_i = 0 - \alpha - \beta X_i$  atau  $\varepsilon_i = -\alpha - \beta X_i$ . Probabilitas untuk masing-masing nilai  $Y_i$  adalah  $P_i$  dan  $(1 - P_i)$ .

Oleh karena model ini diasumsikan linear maka  $E(\varepsilon_i) = 0$ , maka diperoleh:

$$(1 - \alpha - \beta X_i)P_i + (-\alpha - \beta X_i)(1 - P_i) = 0$$

varian error term dari  $\epsilon_i$  adalah kuadrat dari error yang dihasilkan. Karena nilai yang dihasilkan dari perhitungan ini relative kecil, maka dimungkinkan terjadi gejala heteroskedastisitas. Selain itu, nilai yang dihasilkan untuk  $Y_i$  yang diharapkan ( $E(Y_i)$ ) tidak sesuai dengan kondisi realitas di mana  $Y=0$  dan  $Y=1$ .

### Contoh

Sebuah penelitian tentang ingin mengetahui apakah faktor pendapatan dan lamanya pendidikan pengusaha kecil berpengaruh terhadap pilihan pengusaha dalam bermitra dengan perusahaan X. Diambil 15 sampel (diasumsikan memenuhi) dari pengusaha dan didapatkan hasil sebagai berikut.

### Data

No	Pilihan	Pendapatan (Rp000000/bulan)	Pengalaman (tahun)
	Y	X1	X2
1	1	4	6
2	0	1	4
3	1	3	7
4	1	3	7
5	1	4	8
6	1	3	6
7	1	3	6
8	0	1	4
9	0	1	3
10	0	1	3
11	1	4	7
12	1	4	7
13	1	3	7
14	0	1	2
15	0	1	2

**Ket:**

1: bermitra

0 : tidak bermitra

**Pertanyaan:**

1. Bagaimana model yang terbentuk?
2. Apakah model yang terbentuk sesuai dengan kondisi riil atau tidak?
3. Berapa nilai  $R^2$  dan F-hitung?

**Penyelesaian**

1. Gunakan metode least square untuk menghitung nilai  $\alpha$ ,  $\beta_1$  dan  $\beta_2$

$$\sum Y = na + b_1\sum X_1 + b_2\sum X_2 \dots\dots\dots (a)$$

$$\sum X_1Y = a\sum X_1 + b_1\sum X_1^2 + b_2\sum X_1 \sum X_2 \dots\dots\dots (b)$$

$$\sum X_2Y = a\sum X_2 + b_1\sum X_1 \sum X_2 + b_2\sum X_2^2 \dots\dots\dots (c)$$

Dari tabel soal di atas dapat dibuat tabel berikut.

No	Pilihan	Pendapatan (Rp000.000/ bulan)	Pengalaman (tahun)	YX <sub>1</sub>	YX <sub>2</sub>	X <sup>2</sup>	X <sub>2</sub> <sup>2</sup>	X <sub>1</sub> .X <sub>2</sub>
	Y	X <sub>1</sub>	X <sub>2</sub>					
1	1	4	6	4	6	16	36	24
2	0	1	4	0	0	1	16	4
3	1	3	7	3	7	9	49	21
4	1	3	7	3	7	9	49	21
5	1	4	8	4	8	16	64	32
6	1	3	6	3	6	9	36	18
7	1	3	6	3	6	9	36	18
8	0	1	4	0	0	1	16	4
9	0	1	3	0	0	1	9	3
10	0	1	3	0	0	1	9	3
11	1	4	7	4	7	16	49	28
12	1	4	7	4	7	16	49	28
13	1	3	7	3	7	9	49	21
14	0	1	2	0	0	1	4	2
15	0	1	2	0	0	1	4	2

Jumlah	9	37	79	31	61	115	475	229
--------	---	----	----	----	----	-----	-----	-----

$X'X :$

	15	37	79
	37	115	2923
	79	2923	475

$X'Y :$

9  
31  
61

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

$(X'X)^{-1} :$

	0.076055612	-0.001911329	-0.00089
	-0.001911329	-7.91975E-06	0.000367
	-0.000887532	0.00036662	-3.2E-06

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

$$\alpha \quad 0.571109872$$

$$\square \square \quad 0.004916337$$

$$\square \square \quad 0.003182879$$

Dengan demikian model yang terbentuk adalah

$$Y = 0,5711 + 0,0049X_1 + 0,0031X_2$$

Jika nilai  $X_1$  dan  $X_2$  dimasukkan dalam model maka didapatkan nilai  $Y$  sebagai berikut.

Resp.	Pilihan	Pendapatan (Rp000000/bulan))	Pengalaman (tahun)	Y expected	Error
	Y	X1	X2		
1.	1	4	6	0.61	0.39
2.	0	1	4	0.02	-0.02
3.	1	3	7	0.04	0.96
4.	1	3	7	0.04	0.96
5.	1	4	8	0.05	0.95
6.	1	3	6	0.03	0.97
7.	1	3	6	0.03	0.97
8.	0	1	4	0.02	-0.02
9.	0	1	3	0.01	-0.01
10.	0	1	3	0.01	-0.01
11.	1	4	7	0.04	0.96
12.	1	4	7	0.04	0.96
13.	1	3	7	0.04	0.96
14.	0	1	2	0.01	-0.01
15.	0	1	2	0.01	-0.01

Jika dicermati lebih detail nilai *Y expected* sangat berbeda dengan nilai *Y riel* . Sebagai salah satu contoh pada *Y riel* responden 1 adalah 1 yang berarti bahwa peternak memilih program bermitra. Hal ini sedikit berbeda jika dilihat dari nilai *Y Expected* yang didapatkan dari hasil perhitungan sebesar 0,6. Akan tetapi pada responden 3 *Y riel* sebesar 1 sedangkan *Y expected* sebesar 0,04 (jauh dari 1). Hal ini mengindikasikan bahwa terdapat penyimpangan dari model.

Sebagai bahan review bahwa adanya kondisi tersebut memunculkan gejala heteroskedastisitas (salah satu penyimpangan model linear), di mana varians dari error cenderung sama.

Selanjutnya untuk menghitung nilai  $R^2$  digunakan formulasi sebagai berikut.

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2}$$

Sehingga didapatkan nilai  $R^2$  hasil sebesar = 0,93

Untuk menghitung nilai F-hitung digunakan formulasi sebagai berikut.

$$F\text{-hitung} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \text{ sebesar } 181,2$$

### 4.3 Penutup

Karena nilai  $Y$  expected jauh dari nilai  $Y$  riel, maka model ini secara ekonometrik tidak dapat dipertanggungjawabkan secara sepenuhnya.

**IV.**  
**MODEL FUNGSI DISTRIBUSI KOMULATIF**  
**(CUMMULATIVE DISTRIBUTION FUNCTION)**

Standar Kompetensi	:	Mahasiswa dapat mempraktekkan Aplikasi E-Views untuk menganalisis model Fungsi Distribusi Komulatif
Kompetensi dasar	:	1. Mahasiswa memahami konsep dasar ekonometrik model fungsi distribusi komulatif. 2. Mahasiswa mampu menggunakan model fungsi distribusi komulatif di bidang ekonomi pertanian. 3. Mahasiswa mampu mengaplikasikan e-views untuk pengujian model fungsi distribusi komulatif.
Metode Pembelajaran	:	- Metode SCL ( <i>Student Centered Learning</i> ) - Pendekatan: Case Study, Discovery Learning

**5. 1 Pendahuluan**

Model Probabilitas linear memiliki beberapa kelemahan antara lain

1. Nilai  $Y$  expected tidak mampu menjelaskan kondisi riil yang pada prinsipnya adalah 0 dan 1.
2. Variabel gangguan tidak berdistribusi normal

$$Y_i = \beta_0 + \beta_1 X_i + e$$

$$e = Y_i - \beta_0 - \beta_1 X_i$$

Jika  $Y_i = 1$ , maka  $e = 1 - \beta_0 - \beta_1 X_i$  maka probabilitasnya  $P_i$

Jika  $Y_i = 0$ , maka  $e = 0 - \beta_0 - \beta_1 X_i$  maka probabilitasnya  $1 - P_i$

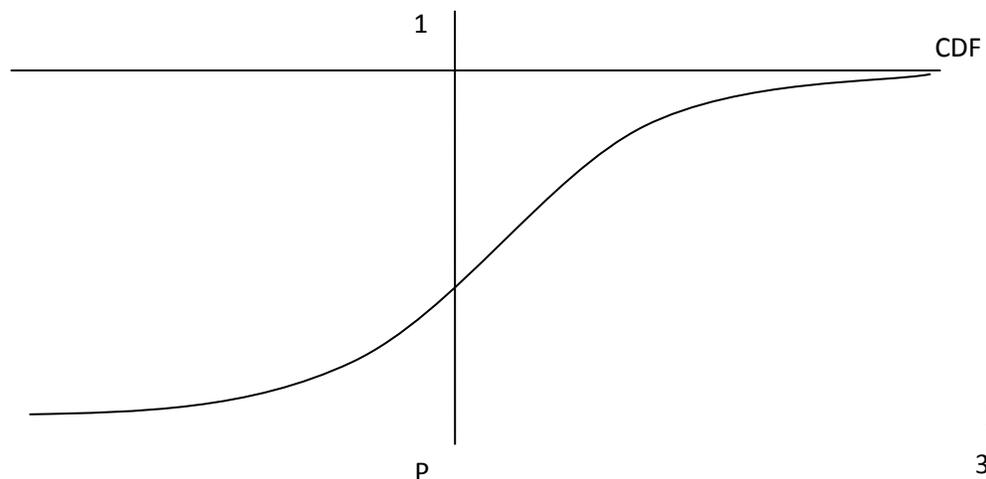
Menurut Widarjono (2007) menjelaskan bahwa gangguan yang tidak memiliki distribusi normal tidak menimbulkan masalah jika tujuan OLS hanya sekedar

estimasi bukan untuk inferensi (prediksi) karena akan menghasilkan estimator yang BLUE (*Best Linear Unbiased Estimator*).

3. varian dari variabel pengganggu mengandung unsur heteroskedastisitas
4. Nilai  $R^2$  diragukan kebenarannya.

Karena keterbatasan-keterbatasan tersebut, maka model LPM kurang bagus digunakan untuk prediksi. Seperti dalam contoh jika pendapatan naik satu rupiah, maka pengrajin akan cenderung untuk bermitra. Jika dikaitkan dengan konsep linear akan dapat dijelaskan jika pendapatan naik dua rupiah, maka akan terjadi kenaikan sebesar dua kali. Kondisi ini tidak sesuai dengan kondisi riil.

Oleh karenanya, untuk menjawab pertanyaan tersebut dikenalkan sebuah model probabilitas yang mampu menjamin bahwa nilainya terletak antara 0 dan 1, sehingga dapat menjamin bahwa nilainya terletak antara 0 dan 1. Model tersebut dikenal dengan istilah *cumulative distribution functions* (CDF). CDF mempunyai dua sifat yaitu (1) ketika  $X_i$  naik maka  $\Pr(Y_i=1 | X_i)$  akan naik akan tetapi tidak pernah melebihi interval 0 – 1; (2) hubungan antara  $P_i$  dan  $X_i$  adalah nonlinear sehingga tingkat perubahannya tetapi kenaikannya semakin besar dan kemudian mengecil.





Bermitra                      Jika  $Z_i \geq Z_i^*$

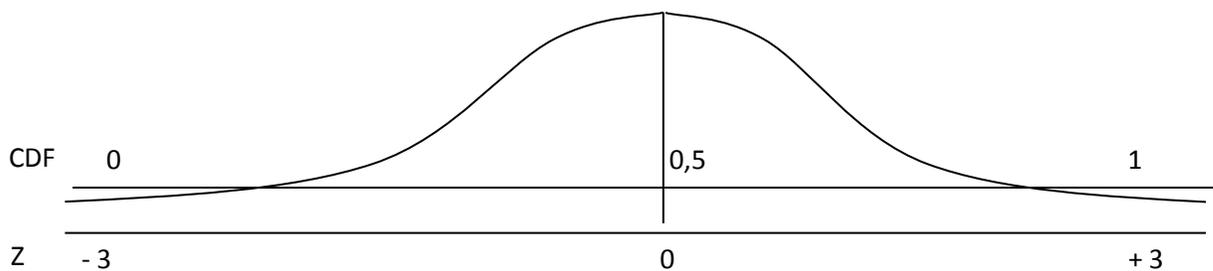
Tidak bermitra              Jika  $Z_i < Z_i^*$

Karena model ini mengikuti distribusi normal, probabilitas  $Z_i^*$  yang kurang atau sama dengan  $Z_i$  dapat dihitung melalui distribusi normal dari CDF.

$$P_i = f(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{-z^2/2} dz$$

$$P_i = f(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\beta_0 + \beta_1 X_i)} e^{-z^2/2} dz$$

Perlu dipahami bahwa probabilitas kumulatif dari kurva normal adalah 0 – 1 digambarkan sebagai berikut.



Dalam model probit ini terdapat dua jenis pendekatan analisis yaitu pendekatan pada data kelompok dan data individual. Pendekatan data kelompok artinya data yang ada digunakan adalah proporsi dari data yang telah dikelompokkan. Alat analisis yang digunakan adalah *ordinary least square* seperti biasa. Sedangkan jika data yang digunakan adalah data individu, maka alat analisis yang digunakan adalah *maximum likelihood estimation (MLE)*. Berikut dibahas satu persatu.

### 1. Model Probit Data Kelompok

Yang dimaksud data kelompok adalah data yang tiap tampilan datanya mewakili data kelompok. Sebagai contoh penelitian yang dilakukan pada 15 desa di Kecamatan

tertentu. Di setiap desa diambil sampel sebesar menurut proporsinya masing-masing. Adapun datanya sebagai berikut.

Tabel 4.1. Pendapatan, pengalaman jumlah yang bermitra dan jumlah sampel pada masing-masing desa.

No	Jumlah Sampel masing-masing	Jumlah yang bermitra	Pendapatan (Rp000000/bulan)	Pengalaman (tahun)
	N	n	$X_1$	$X_2$
1	30	5	4	6
2	25	6	1	4
3	20	7	3	7
4	15	7	3	7
5	16	10	4	8
6	20	10	3	6
7	25	12	3	6
8	20	11	1	4
9	30	15	1	3
10	45	20	1	3
11	30	15	4	7
12	35	15	4	7
13	40	20	3	7
14	45	40	1	2
15	30	25	1	2

Pada contoh di atas, data yang tersedia adalah data kelompok. Contoh, data no.1 jumlah sampel sebanyak 30 dan yang bermitra sebanyak 5 orang. Artinya probabilitas untuk bermitra sebesar 0,17 (5/30). Probabilitas ini merupakan dasar untuk merubah dari data nominal ke data pengukuran dengan bantuan tabel distribusi normal. Dasar yang digunakan jika probabilitas < 0,5 maka tanda normalitasnya negative, jika sebesar 0,5 maka besar normalitasnya adalah 0 dan jika > 0,5 maka normalitasnya positif. Sebagai contoh probabilitas 0,17 dirubah ke *commulative distributive function* sebesar  $1-0,17 = 0,83$  dan setara dengan nilai  $z = 0,96$ , karena nilai probabilitasnya <

0,5, maka tandanya negative. Dengan demikian nilai cdf untuk probabilitas 0,17 sebesar -0,96.

Hasil selengkapnya adalah sebagai berikut.

Tabel 4.2. Pendapatan, pengalaman jumlah yang bermitra yang sudah dinormalkan dan jumlah sampel pada masing-masing desa.

No	Jumlah sampel	Jumlah yang bermitra	Probabilitas	Zi	Zi+5	Pendapatan	pengalaman
1	30	5	0.17	-0.96	4.04	4	6
2	25	6	0.24	-0.71	4.29	1	4
3	20	7	0.35	-0.39	4.61	3	7
4	15	7	0.47	-0.08	4.92	3	7
5	16	10	0.63	0.33	5.33	4	8
6	20	10	0.50	0.00	5.00	3	6
7	25	12	0.48	-0.05	4.95	3	6
8	20	11	0.55	0.13	5.13	1	4
9	30	15	0.50	0.00	5.00	1	3
10	45	20	0.44	-0.15	4.85	1	3
11	30	15	0.50	0.00	5.00	4	7
12	35	15	0.43	-0.18	4.82	4	7
13	40	20	0.50	0.00	5.00	3	7
14	45	40	0.89	1.23	6.23	1	2
15	30	25	0.83	0.96	5.96	1	2

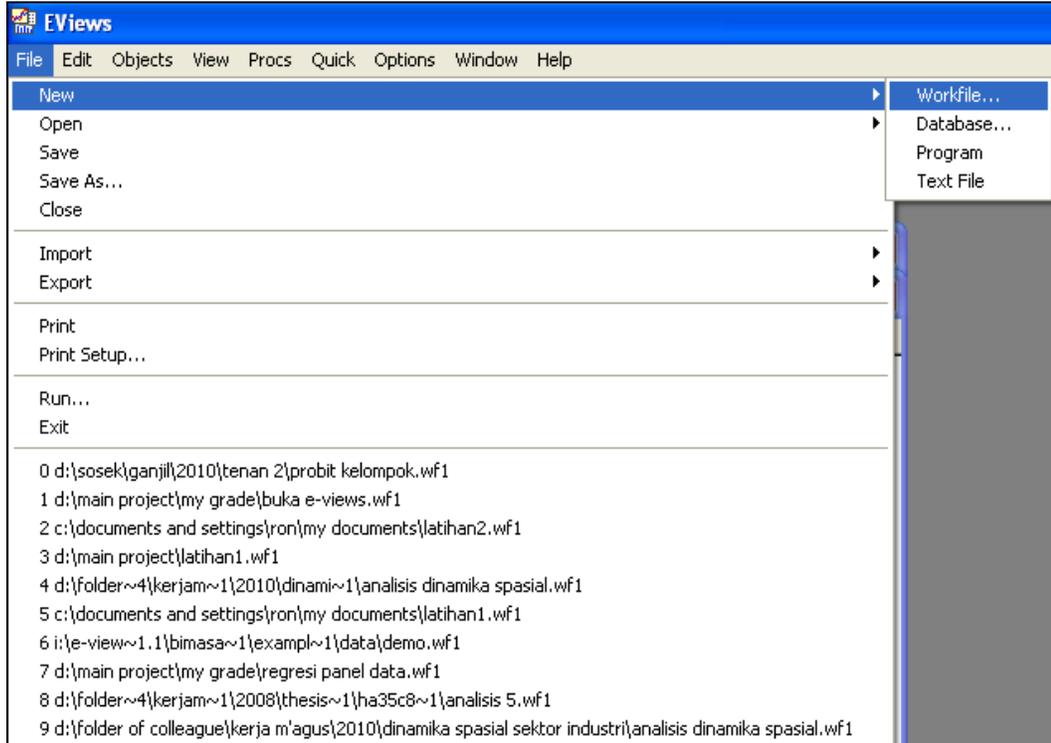
Sumber: Data diolah

Karena analisis yang digunakan adalah OLS, maka  $Z_i$  perlu ditambah dengan 5 (nilai perkiraan tertinggi dari nilai  $Z$ ). Selanjutnya, setelah didapatkan hasil tersebut, maka langkah selanjutnya adalah mengolah data berdasarkan analisis OLS.

Karena alat analisis yang digunakan adalah OLS, maka hal penting yang harus ditemukan adalah  $R^2$ , F-hitung, nilai estimasi (arah dan besarnya koefisien regresi), t-hitung.

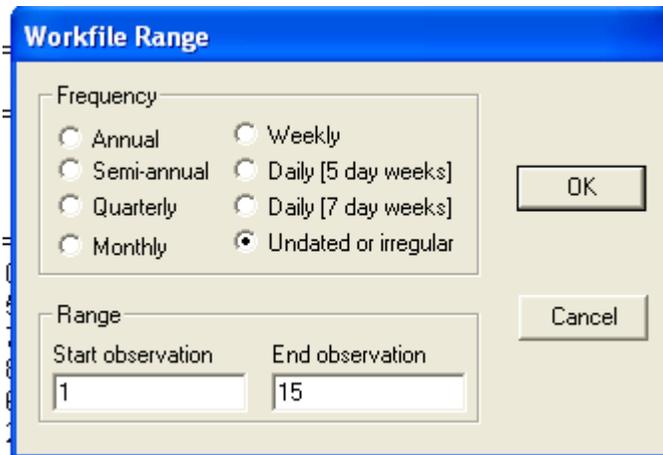
Penyelesaian model analisis yang digunakan sama seperti pada bab sebelumnya, sehingga pada bab ini digunakan penyelesaian dengan *soft ware* e-views. Adapun langkah yang digunakan adalah sebagai berikut.

1. Membuka file baru dalam e-views



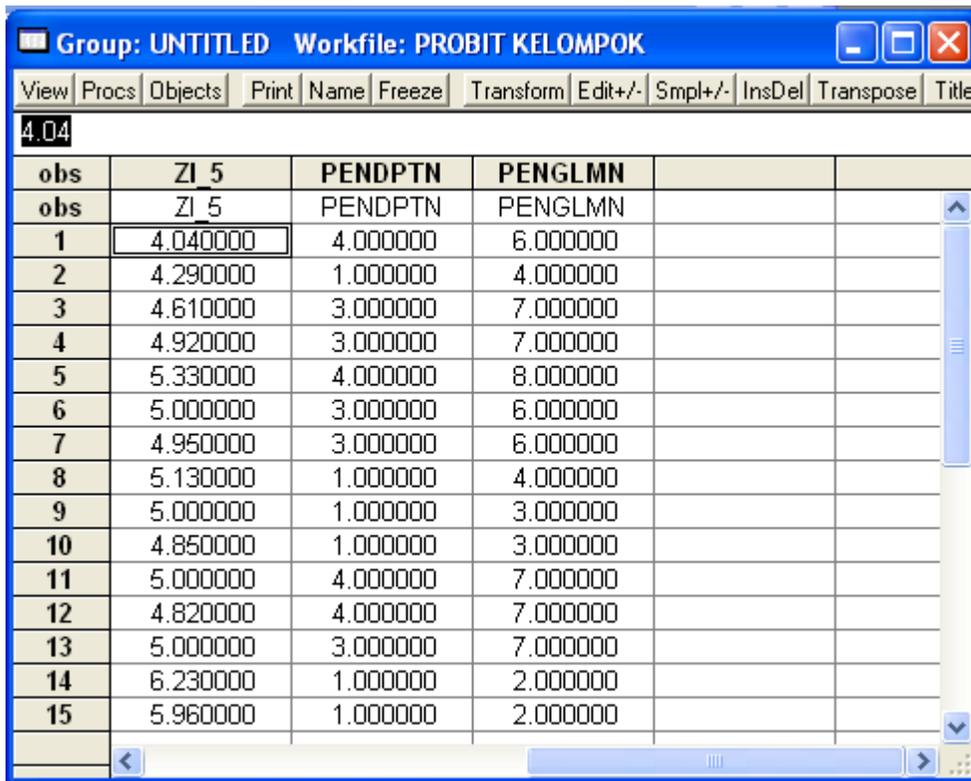
Gambar 4.1 Tampilan Awal E-Views Saat Membuka File Baru

Karena jenis data yang digunakan adalah cross-section, maka dipilih undated atau irregular.



Gambar 4.2 Tampilan E-Views Saat Menentukan Jenis Data dan Jumlah Data

2. Memasukkan data dari ms.excell



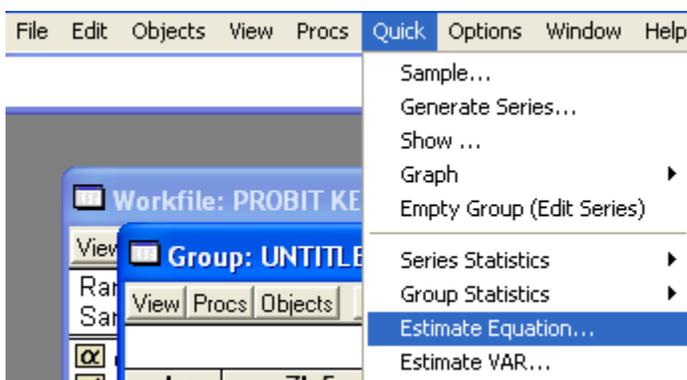
The screenshot shows the EViews software interface with a data table. The window title is "Group: UNTITLED Workfile: PROBIT KELOMPOK". The menu bar includes View, Procs, Objects, Print, Name, Freeze, Transform, Edit+/-, Smpl+/-, InsDel, Transpose, and Title. The data table has the following content:

obs	ZI 5	PENDPTN	PENGLMN
1	4.040000	4.000000	6.000000
2	4.290000	1.000000	4.000000
3	4.610000	3.000000	7.000000
4	4.920000	3.000000	7.000000
5	5.330000	4.000000	8.000000
6	5.000000	3.000000	6.000000
7	4.950000	3.000000	6.000000
8	5.130000	1.000000	4.000000
9	5.000000	1.000000	3.000000
10	4.850000	1.000000	3.000000
11	5.000000	4.000000	7.000000
12	4.820000	4.000000	7.000000
13	5.000000	3.000000	7.000000
14	6.230000	1.000000	2.000000
15	5.960000	1.000000	2.000000

Gambar 4.3 Tampilan Pengentrian Data

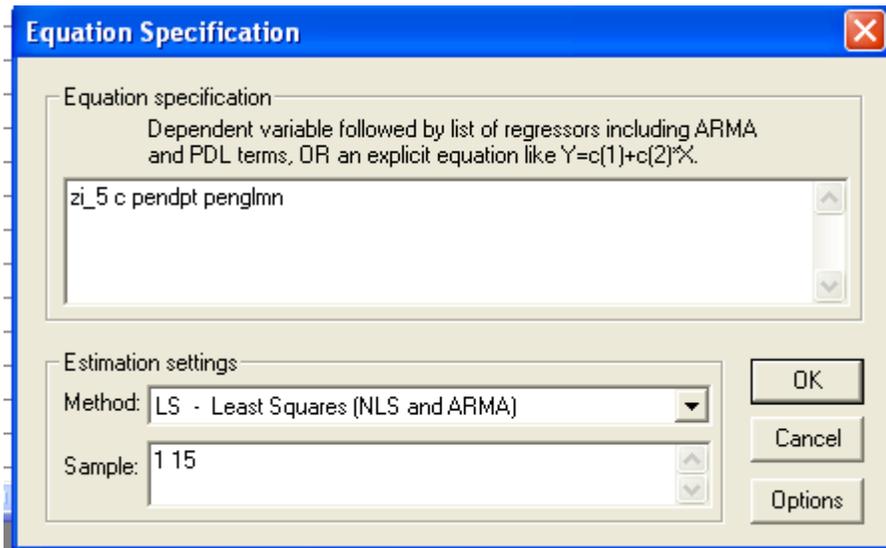
3. Memberi nama pada masing-masing variabel, seperti di atas

4. Melakukan analisis



Gambar 4.4 Tampilan Penganalisan Data

Isikan variabel sebagai berikut.



Gambar 4.5 Tampilan Pemilihan Alat Analisis

## 5. Menginterpretasi hasil

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.700000	0.414905	13.73809	0.0000
PENDPTN	0.105833	0.263966	0.400935	0.6955
PENGLMN	-0.180833	0.167512	-1.079522	0.3016

R-squared	0.211030	Mean dependent var	5.008667
Adjusted R-squared	0.079535	S.D. dependent var	0.547773
S.E. of regression	0.525538	Akaike info criterion	1.728068
Sum squared resid	3.314283	Schwarz criterion	1.869678
Log likelihood	-9.960512	F-statistic	1.604854
Durbin-Watson stat	0.617887	Prob(F-statistic)	0.241192

Gambar 4.6 Tampilan Hasil Analisis

## 2. Model Probit Data Individu

Yang dimaksud dengan data individu adalah data yang munculnya tiap individu. Dalam kasus sebelumnya, orang yang bermitra diberi kategori 1 dan yang tidak bermitra diberi kategori nol, dan sebaliknya. Selanjutnya, prosedur untuk mengestimasi model probit ini adalah melalui metode *maximum likelihood estimation*.

Secara matematis model MLE adalah model yang mengestimasi dengan kemungkinan tertinggi artinya mengestimasi variabel Y sebagaimana mungkin sama dengan variabel Y.

Misalkan distribusi normal memiliki persamaan sebagai berikut.

$$Y_i = \beta_0 + \beta_1 X_i + e$$

Variabel Y memiliki distribusi normal dengan rata-rata  $= \beta_0 + \beta_1 X_i$  dan varian  $\sigma^2$ . Misalkan distribusi probabilitas dengan rata-rata dan varian tertentu dapat ditulis sebagai berikut (Widarjono, 2007).

$$P(Y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} (Y_i - \beta_0 - \beta_1 X_i)^2 \right]$$

Fungsi likelihood adalah perkalian dari setiap probabilitas kejadian individual pada semua observasi n. Dengan demikian fungsi likelihood dapat ditulis sebagai berikut.

$$LF(Y_1, Y_2, \dots, Y_n, \beta_0, \beta_1, \sigma^2) = p(Y_1) p(Y_2) \dots p(Y_n)$$

$$= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp \left[ -\sum \left( \frac{1}{2\sigma^2} (Y_i - \beta_0 - \beta_1 X_i)^2 \right) \right]$$

Tujuan menggunakan estimasi maximum likelihood adalah mengestimasi nilai Y setinggi mungkin. Untuk memaksimumkan hal tersebut dilakukan dengan cara melakukan diferensiasi atau turunan setiap parameter dengan variabel tertentu yang disamadengankan nol. Sekarang ini software program untuk menganalisis estimasi ini sudah tersedia antara lain SPSS, e-views, stata dan lain-lain. Sekarang ini digunakan perangkat e-views untuk menganalisis hal tersebut.

Berebeda dengan metode OLS, metode estimasi MLE(*maximum likelihood estimation*) memiliki ukuran-ukuran antara lain.

1. Untuk mengestimasi pengaruh secara bersama-sama digunakan uji likelihood ratio (LR) atau sama dengan nilai F-hitung di OLS. Nilai LR ini dibandingkan dengan  $X^2$  (*Chi-square*) tabel. Jika nilai  $X^2$  hitung  $> X^2$  \_tabel maka secara bersama-sama variabel independen berpengaruh terhadap variabel dependen.
2. Untuk mengetahui besarnya pengaruh perubahan variabel x terhadap perubahan variabel Y digunakan uji Mc Fadden  $R^2$  atau disingkat (McF  $R^2$ ). Nilai McF  $R^2$  ini antara nilai 0 – 1.
3. Untuk mengetahui secara parsial pengaruh variabel X terhadap variabel Y digunakan uji z. Nilai z ini dibandingkan dengan standar normal. Jika nilai Z lebih besar dibandingkan Z normal pada tingkat signifikansi tertentu, maka secara statistic variabel independen berpengaruh terhadap variabel dependen.
4. Untuk mengetahui besarnya pengaruh masing-masing variabel digunakan koefisien regresi dari model tersebut. Nilai koefisien ini menunjukkan arah dan

besarnya pengaruh variabel independen. Jika positif, maka keputusan mengarah pada  $Y=1$ , jika negative, maka keputusan mengarah pada  $Y=0$ .

Berikut diberikan contoh.

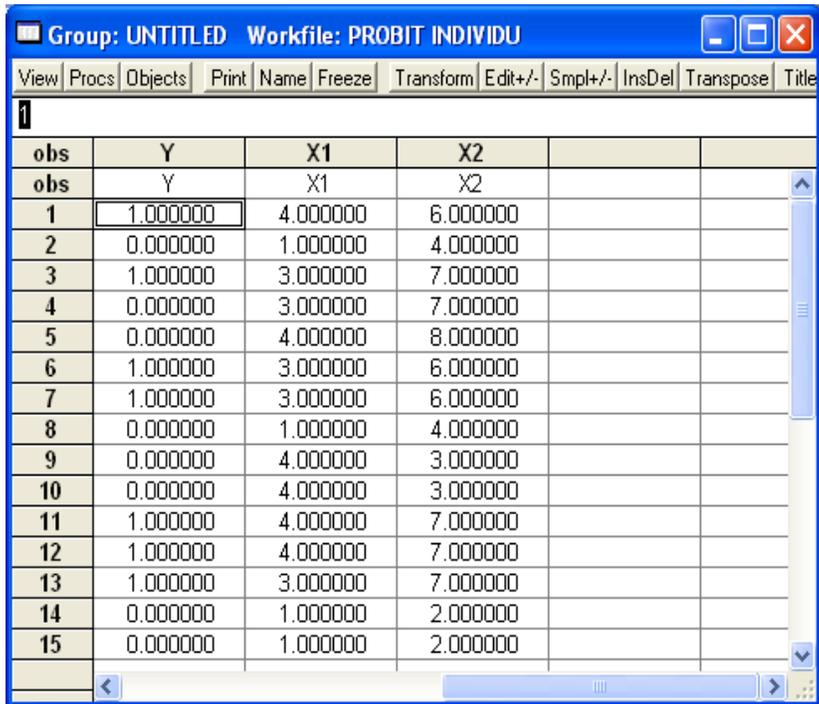
Sebuah penelitian tentang ingin mengetahui apakah faktor pendapatan dan lamanya pendidikan pengusaha kecil berpengaruh terhadap pilihan pengusaha dalam bermitra dengan perusahaan X atau tidak. Diambil 15 sampel (diasumsikan memenuhi) dari pengusaha dan didapatkan hasil sebagai berikut.

Tabel 4.3. Pengaruh Variabel Pendapatan dan Pengalaman terhadap Pilihan Petani untuk Bermitra atau Tidak

No	Pilihan	Pendapatan (Rp000000/bulan)	Pengalaman (tahun)
	Y	$X_1$	$X_2$
1	1	4	6
2	0	1	4
3	1	3	7
4	0	3	7
5	0	4	8
6	1	3	6
7	1	3	6
8	0	1	4
9	0	4	3
10	0	4	3
11	1	4	7
12	1	4	7
13	1	3	7
14	0	1	2
15	0	1	2

Berikut langkah penyelesaian dengan program e-views.

1. Memasukkan data dari ms.excell dan memberi nama pada masing-masing variabel

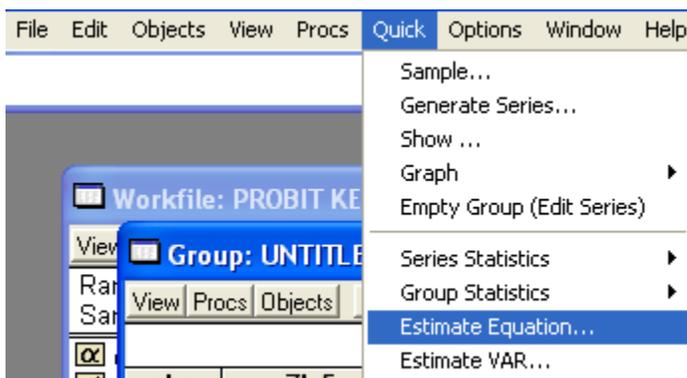


The screenshot shows the EViews software interface with a data table. The window title is "Group: UNTITLED Workfile: PROBIT INDIVIDU". The menu bar includes View, Procs, Objects, Print, Name, Freeze, Transform, Edit+/-, Smpl+/-, InsDel, Transpose, and Title. The data table has the following content:

obs	Y	X1	X2
1	1.000000	4.000000	6.000000
2	0.000000	1.000000	4.000000
3	1.000000	3.000000	7.000000
4	0.000000	3.000000	7.000000
5	0.000000	4.000000	8.000000
6	1.000000	3.000000	6.000000
7	1.000000	3.000000	6.000000
8	0.000000	1.000000	4.000000
9	0.000000	4.000000	3.000000
10	0.000000	4.000000	3.000000
11	1.000000	4.000000	7.000000
12	1.000000	4.000000	7.000000
13	1.000000	3.000000	7.000000
14	0.000000	1.000000	2.000000
15	0.000000	1.000000	2.000000

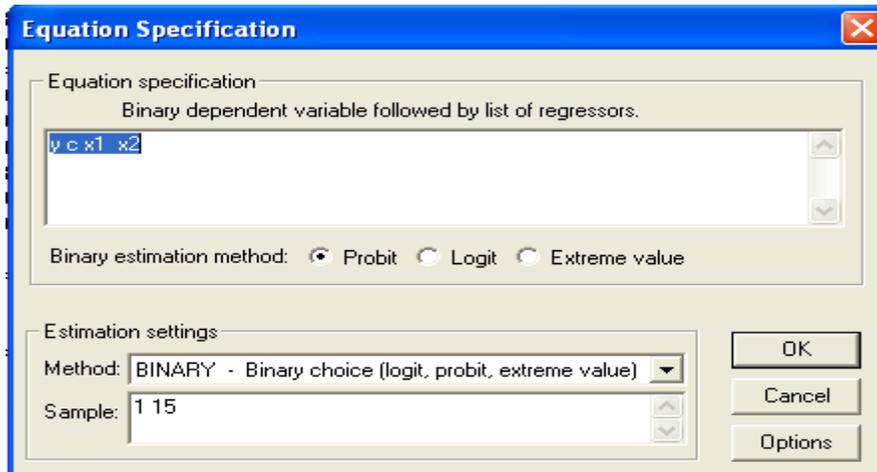
Gambar 4.7 Pengentrian Data

2. Melakukan analisis



Gambar 4.8 Langkah Penganalisisan Data

Masukkan nama variabel dalam box yang tersedia dan pilih probit.



Gambar 4.9 Pilihan Alat Analisis Data

### 3. Menginterpretasi hasil

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-3.320537	1.736707	-1.911973	0.0559
X1	0.249111	0.460201	0.541309	0.5883
X2	0.441379	0.259489	1.700956	0.0890

Mean dependent var	0.466667	S.D. dependent var	0.516398
S.E. of regression	0.428525	Akaike info criterion	1.320593
Sum squared resid	2.203604	Schwarz criterion	1.462203
Log likelihood	-6.904445	Hannan-Quinn criter.	1.319084
Restr. log likelihood	-10.36385	Avg. log likelihood	-0.460296
LR statistic (2 df)	6.918810	McFadden R-squared	0.333795
Probability(LR stat)	0.031448		

Obs with Dep=0	8	Total obs	15
Obs with Dep=1	7		

Gambar 4.10 Hasil Analisis Data

Istilah-istilah penting

*Commulative distribution function*

*Probit model*

*Logit model*

*Maximum Likelihood Estimation*

*Ordinary least square*

**V.**  
**MODEL LOGISTIK**  
**LOGISTIC MODEL**

Standar Kompetensi	:	Mahasiswa dapat mempraktekkan Aplikasi E-Views untuk menganalisis Regresi Model Logistik
Kompetensi dasar	:	1. Mahasiswa memahami konsep dasar ekonometrik regresi model logistik. 2. Mahasiswa mampu menggunakan regresi model logistik di bidang ekonomi pertanian. 3. Mahasiswa mampu mengaplikasikan e-views untuk pengujian regresi model logistik.
Metode Pembelajaran	:	- Metode SCL (Student Centered Learning) - Pendekatan: Case Study, Discovery Learning

**6.1 Pendahuluan**

Model CDF digunakan jika variabel Y berbentuk binomial yang biasanya merupakan respon kualitatif. Pada pembahasan sebelumnya, model probit mendasarkan pada distribusi normal. Sedangkan pada model logistic variabel Y memiliki distribusi logistic. Pada model ini variabel Y berbentuk probabilitas dengan pendekatan logaritma.

**6.2 Model**

Untuk menyelesaikan model logit terlebih dahulu dibuat model fungsi probabilitik logistic kumulatifnya. Sebagai gambaran dan perbandingan, berikut ini diberikan model probit. Kembali ke contoh ilustrasi awal, bahwa probabilitas untuk bermitra tergantung dari variabel-variabel yang mempengaruhinya. Dalam suatu kelompok dengan jumlah 30 orang, jumlah orang yang bermitra sebanyak 6 orang maka probabilitasnya adalah 0,2 (6/30). Asumsi yang digunakan dalam probit model adalah

mengikuti distribusi normal maka model fungsi probabilistic kumulatifnya adalah sebagai berikut.

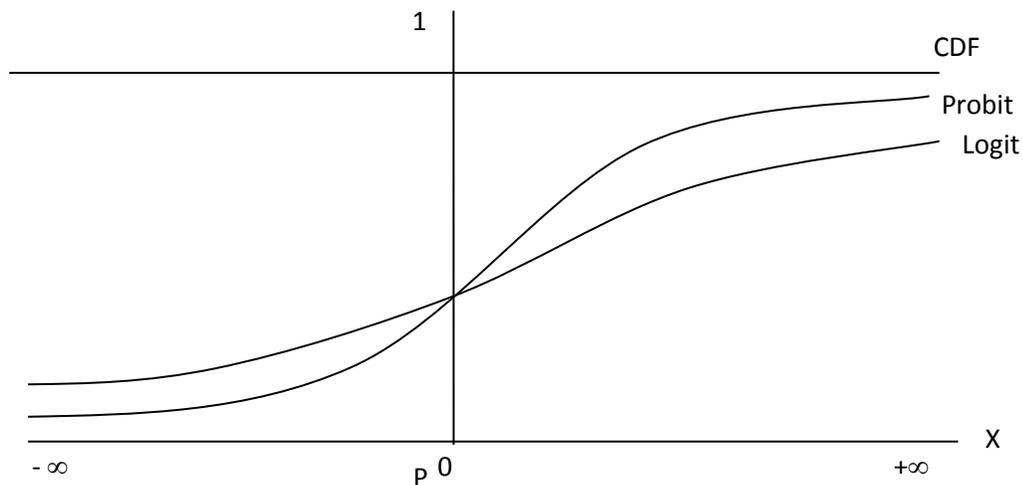
$$P_i = f(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\beta_0 + \beta_1 X_1)} e^{-z^2/2} dz$$

Model probit lebih menjelaskan bahwa Probabilitas merupakan fungsi  $Z_i$  dimana dapat diselesaikan dengan integral dari tidak terhingga sampai ke- $z$  dengan pendekatan  $e^{z^2/2}$ . Sedangkan di model logistic digunakan pendekatan sebagai berikut.

$$P_i = f(Z_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 = \frac{1}{1 + e^{-Z_i}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1)}} \text{ (Gujarati, 2006)}$$

Probabilitas merupakan fungsi dari  $Z_i$  dimana dapat diselesaikan dengan pembagian  $1/(1 + e^{-(\beta_1 X_1 + \beta_2 X_2)})$ . Jika angka-angka dimasukkan fungsi juga mengikuti commulatif distribution function.

Secara grafis, dapat dijelaskan sebagai berikut.



Terlihat bahwa grafi model probit mendekati 0 dan 1 dibandingkan dengan model logit yang relative lamban. Hal ini juga berpengaruh terhadap model estimasi dan hasil yang didapatkan.

$P_i = \frac{1}{1+e^{-Z_i}}$  supaya dapat sama dengan satu, maka sebelah kanan dan kiri dikalikan dengan  $1+e^{-Z_i}$

$$P_i(1+e^{-Z_i}) = 1$$

Angka satu berarti probabilitas adalah 100%, sedangkan untuk mencari probabilitas dari Y maka dapat dibagi dengan  $P_i-1$

$$\frac{(1+e^{-Z_i})P_i}{P_i-1} = \frac{1}{P_i-1}$$

$$e^{-Z_i} \cdot P_i = 1 - P_i$$

$$\frac{1}{e^{Z_i}} = \frac{(1-P_i)}{P_i}$$

$$\frac{e^{Z_i}}{1} = \frac{P_i}{(1-P_i)}$$

Model diatas dapat ditransformasikan ke model logaritma menjadi

$$Z_i = \ln\left(\frac{P_i}{1-P_i}\right)$$

Oleh karenanya model logit dapat diselesaikan sebagai berikut.

$$Z_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Dari model tersebut diketahui bahwa  $P_i$  tidak linear baik pada  $X$ inya maupun pada parameternya.

Estimasi model logit dapat dibedakan baik pada data kelompok maupun data individu, seperti pada model probit sebelumnya. Pendekatan data kelompok artinya data yang ada digunakan adalah proporsi dari data yang telah dikelompokkan. Alat analisis yang digunakan adalah *ordinary least square* seperti biasa. Sedangkan jika data yang digunakan adalah data individu, maka alat analisis yang digunakan adalah *maximum likelihood estimation (MLE)*. Berikut dibahas satu persatu.

### **1. Model Logit Data Kelompok**

Data untuk analisis logistik sama dengan untuk analisis probit. Misalkan penelitian yang dilakukan pada 15 desa di Kecamatan tertentu. Di setiap desa diambil sampel sebesar menurut proporsinya masing-masing. Adapun datanya sebagai berikut.

Tabel 5.1. Pendapatan, pengalaman dan jumlah yang bermitra pada masing-masing desa.

No	Jumlah Sampel masing-masing	Jumlah yang bermitra	Pendapatan (Rp000000/bulan))	Pengalaman (tahun)
	N	n	X1	X2
1	30	5	4	6
2	25	6	1	4
3	20	7	3	7
4	15	7	3	7
5	16	10	4	8
6	20	10	3	6
7	25	12	3	6
8	20	11	1	4
9	30	15	1	3
10	45	20	1	3
11	30	15	4	7
12	35	15	4	7
13	40	20	3	7
14	45	40	1	2
15	30	25	1	2

Langkah-langkah untuk menganalisis model tersebut adalah

1. Cari probabilitas dari kejadian yang dipersyaratkan misalkan probabilitas yang bermitra.
2. cari nilai Zi dengan formulasi  $Z_i = \ln (P_i/(1-P_i))$

Tabel 5.2 Penentuan nilai Z

No	Jumlah sampel	Jumlah yang bermitra	Probabilitas	1-p	p/1-p	Ziln(Pi/(1-pi))
1	30	5	0.17	0.83	0.20	-1.61
2	25	6	0.24	0.76	0.32	-1.15
3	20	7	0.35	0.65	0.54	-0.62
4	15	7	0.47	0.53	0.88	-0.13
5	16	10	0.63	0.38	1.67	0.51
6	20	10	0.50	0.50	1.00	0.00
7	25	12	0.48	0.52	0.92	-0.08
8	20	11	0.55	0.45	1.22	0.20
9	30	15	0.50	0.50	1.00	0.00
10	45	20	0.44	0.56	0.80	-0.22
11	30	15	0.50	0.50	1.00	0.00
12	35	15	0.43	0.57	0.75	-0.29
13	40	20	0.50	0.50	1.00	0.00
14	45	40	0.89	0.11	8.00	2.08
15	30	25	0.83	0.17	5.00	1.61

3. Regresikan variabel X terhadap Zi dengan pendekatan OLS (*ordinary least square*). Hasil analisis dapat dilihat sebagai berikut.

Dependent Variable: ZI

Method: Least Squares

Date: 10/31/10 Time: 22:24

Sample: 1 15

Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.192568	0.687776	1.733950	0.1085
PENGLMN	-0.303014	0.277681	-1.091231	0.2966
PENDPTN	0.171610	0.437569	0.392189	0.7018
R-squared	0.219430	Mean dependent var		0.020000
Adjusted R-squared	0.089334	S.D. dependent var		0.912900
S.E. of regression	0.871169	Akaike info criterion		2.738895
Sum squared resid	9.107228	Schwarz criterion		2.880505
Log likelihood	-17.54171	F-statistic		1.686686
Durbin-Watson stat	0.610084	Prob(F-statistic)		0.226190

Gambar 5.1 Hasil Analisis Model

4. Lakukan analisis dari hasil perhitungan tersebut.

a.  $R^2$

b. F-statistik

c. t-hitung

d. Koefesien regresi.

Coba bandingkan hasil perhitungan di atas dengan hasil perhitungan model probit.

### A. Model Logit Data Individu

Yang dimaksud dengan data induvidu adalah data yang munculnya tiap individu. Sama pada model probit dari kasus sebelumnya orang yang bermitra diberi kategori 1 dan yang tidak bermitra diberi kategori nol, dan sebaliknya. Selanjutnya, prosedur untuk mengestimasi model probit ini adalah melalui metode *maximum likelihood estimation*.

Secara matemastis model MLE adalah model yang mengestimasi dengan kemungkinan tertinggi artinya mengestimasi variabel Y sebagaimana mungkin sama dengan variabel Y.

Misalkan distribusi normal memiliki persamaan sebagai berikut.

$$Y_i = \beta_0 + \beta_1 X_i + e$$

Variabel Y memiliki distribusi normal dengan rata-rata= $\beta_0 + \beta_1 X_i$  dan varian  $\sigma^2$ . Misalkan distribusi probabilitas dengan rata-rata dan varian tertentu dapat ditulis sebagai berikut (Widarjono, 2007).

$$Z_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Tujuan menggunakan estimasi maximum likelihood adalah mengestimasi nilai Y setinggi mungkin. Untuk memaksimalkan hal tersebut dilakukan dengan cara melakukan diferensiasi atau turunan setiap parameter dengan variabel tertentu yang disamadengankan nol.

Sekarang ini software program untuk menganalisis estimasi ini sudah tersedia antara lain SPSS, e-views, stata dan lain-lain. Sekarang ini digunakan perangkat e-views untuk menganalisis hal tersebut.

Ukuran yang digunakan di model logit sama dengan yang digunakan di model logit yaitu

L-R ratio, McFadden  $R^2$ , Z-hitung dan koefisien regresi. Khusus untuk koefisien regresi perlu adanya transformasi dari koefisien regresi logaritmik menjadi koefisien regresi bilangan riil dengan cara mencari eksponensial dari koefisien regresi.

1. Uji likelihood ratio (LR) atau sama dengan nilai F-hitung di OLS digunakan untuk mengestimasi pengaruh secara bersama-sama variabel x terhadap variabel Y. Nilai LR ini dibandingkan dengan  $X^2$  (*Chi-square*) tabel. Jika nilai LR ratio-hitung  $> X^2$  tabel maka secara bersama-sama variabel independen berpengaruh terhadap variabel dependen.
2. Mc. Fadden  $R^2$  atau disingkat (McF  $R^2$ ) digunakan untuk mengetahui besarnya pengaruh perubahan variabel x terhadap perubahan variabel Y digunakan. Nilai McF  $R^2$  ini antara nilai 0 – 1.
3. Untuk mengetahui secara parsial pengaruh variabel X terhadap variabel Y digunakan uji z. Nilai z ini dibandingkan dengan standar normal. Jika nilai Z lebih besar dibandingkan Z normal pada tingkat signifikansi tertentu, maka secara statistic variabel independen berpengaruh terhadap variabel dependen.
4. Untuk mengetahui besarnya pengaruh masing-masing variabel digunakan koefisien regresi dari model tersebut. Namun koefisien regresi yang didapatkan

adalah koefesien regesi yang sudah dieksponensialkan. Nilai koefesien ini menunjukkan arah dan besarnya pengaruh variabel independen. Jika positif, maka keputusan mengarah pada  $Y=1$ , jika negative, maka keputusan mengarah pada  $Y=0$ .

### Contoh

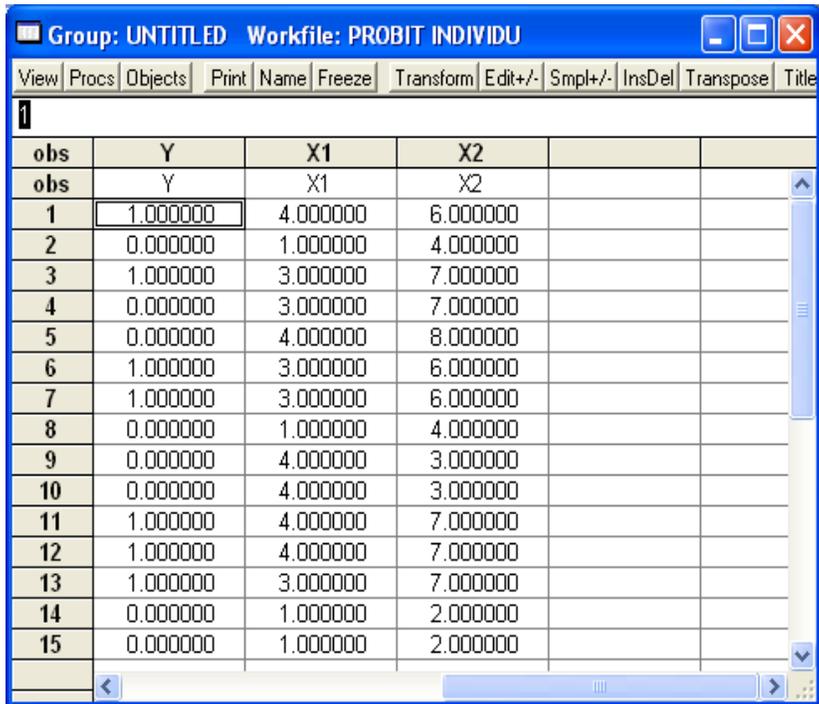
Sebuah penelitian ingin mengetahui apakah faktor pendapatan dan lamanya pendidikan pengusaha kecil berpengaruh terhadap pilihan pengusaha dalam bermitra dengan perusahaan X atau tidak. Diambil 15 sampel (diasumsikan memenuhi) dari pengusaha dan didapatkan hasil sebagai berikut.

Tabel 5.2 Pengaruh Faktor Pendapatan dan Pengalaman terhadap Pilihan Bermitra

No	Pilihan Y	Pendapatan (Rp000000/bulan) X1	Pengalaman (tahun) X2
1	1	4	6
2	0	1	4
3	1	3	7
4	0	3	7
5	0	4	8
6	1	3	6
7	1	3	6
8	0	1	4
9	0	4	3
10	0	4	3
11	1	4	7
12	1	4	7
13	1	3	7
14	0	1	2
15	0	1	2

Berikut langkah penyelesaian dengan program e-views.

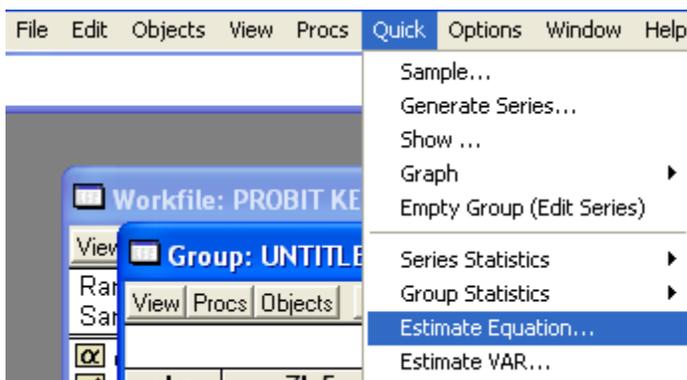
4. Memasukkan data dari ms.excell dan memberi nama pada masing-masing variabel



The screenshot shows the EViews software interface. The title bar reads "Group: UNTITLED Workfile: PROBIT INDIVIDU". The menu bar includes View, Procs, Objects, Print, Name, Freeze, Transform, Edit+/-, Smpl+/-, InsDel, Transpose, and Title. The main window displays a table with 15 observations and three variables: Y, X1, and X2. The data is as follows:

obs	Y	X1	X2
1	1.000000	4.000000	6.000000
2	0.000000	1.000000	4.000000
3	1.000000	3.000000	7.000000
4	0.000000	3.000000	7.000000
5	0.000000	4.000000	8.000000
6	1.000000	3.000000	6.000000
7	1.000000	3.000000	6.000000
8	0.000000	1.000000	4.000000
9	0.000000	4.000000	3.000000
10	0.000000	4.000000	3.000000
11	1.000000	4.000000	7.000000
12	1.000000	4.000000	7.000000
13	1.000000	3.000000	7.000000
14	0.000000	1.000000	2.000000
15	0.000000	1.000000	2.000000

5. Melakukan analisis



Masukkan nama variabel dalam box yang tersedia dan pilih logit.



Gambar 5.1 Pilihan Alat Analisis

Hasil analisis dapat dilihat sebagai berikut.

Dependent Variable: Y  
 Method: ML - Binary Logit (Quadratic hill climbing)  
 Date: 10/31/10 Time: 23:57  
 Sample: 1 15  
 Included observations: 15  
 Convergence achieved after 5 iterations  
 Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-5.692925	3.232359	-1.761229	0.0782
X1	0.448096	0.795255	0.563463	0.5731
X2	0.758052	0.456923	1.659037	0.0971
Mean dependent var	0.466667	S.D. dependent var		0.516398
S.E. of regression	0.423884	Akaike info criterion		1.316650
Sum squared resid	2.156134	Schwarz criterion		1.458260
Log likelihood	-6.874873	Hannan-Quinn criter.		1.315141
Restr. log likelihood	-10.36385	Avg. log likelihood		-0.458325
LR statistic (2 df)	6.977953	McFadden R-squared		0.336649
Probability(LR stat)	0.030532			
Obs with Dep=0	8	Total obs		15
Obs with Dep=1	7			

Gambar 5.2 Hasil Analisis

## VI.

### PENDUGAAN PARAMETER PADA MODEL SIMULTAN

Standar Kompetensi	:	Mahasiswa dapat menganalisis model simultan
Kompetensi dasar	:	1. Mahasiswa menjelaskan contoh perekonomian dengan model simultan 2. Mahasiswa mampu menganalisis persamaan simultan. 3. Mahasiswa mampu menginterpretasikan model analisis simultan
Metode Pembelajaran	:	- Metode SCL (Student Centered Learning) - Pendekatan: Case Study, Discovery Learning

#### Pendahuluan

Banyak dijumpai dalam kehidupan sehari-hari kondisi perekonomian yang sangat kompleks dan beragam. Misalkan kondisi pasar beras yang sangat kompleks dimana permasalahan beras tidak hanya pada penawaran saja, akan tetapi juga permintaan. Juga permasalahan beras dipengaruhi oleh impor dan ekspor beras. Pada pembahasan-pembahasan sebelumnya telah dibahas secara parsial (pendekatan terpisah) masing-masing permodelan. Secara parsial permintaan beras dipengaruhi oleh harga beras itu sendiri, harga komoditas lain seperti jagung, ikan, dan daging; pendapatan penduduk, selera masyarakat dan ekspektasi masyarakat pada masa yang akan datang. Di sisi lain, secara parsial diketahui bahwa penawaran beras dipengaruhi oleh harga beras, harga bahan baku, jumlah penduduk, dan ekspektasi pada masa yang akan datang. Pendekatan-pendekatan tersebut tepat adanya, dan dapat dibahas lebih detail.

Namun demikian, ada pendekatan lain yang dapat merangkum seluruh kondisi tersebut (permintaan dan penawaran beras) secara bersama (simultan). Karenanya dalam bagian ini dikenalkan model simultan. Jadi persamaan simultan adalah persamaan yang terdiri dari dua atau lebih persamaan.

### **Model Simultan (Penjelasan Singkat)**

Model simultan merupakan model yang menggambarkan kondisi perekonomian riil yang lebih kompleks dalam sebuah model pendugaan parameter. Dikatakan model yang lebih kompleks karena terdapat model-model yang secara parsial menganalisis perekonomian. Model ini sesuai digunakan untuk menganalisis kondisi perekonomian secara general. Model ini lebih banyak digunakan untuk menduga parameter yang bersifat *time series*, meskipun ada juga yang menggunakannya untuk parameter yang bersifat *cross section*. Untuk mempermudah pemahaman digunakan model simultan, yaitu model keragaan beras. Model ini terdiri dari 3 persamaan struktural (masing-masing persamaan yang ada dalam persamaan simultan) yaitu persamaan penawaran, permintaan, impor, dan stok. Masing-masing persamaan struktural digambarkan sebagai berikut.

$$\ln S_t = a_0 + a_1 \ln P_d + a_2 \ln P_{d(t-1)} + a_3 \ln A + v_1$$

$$\ln D_t = b_0 + b_1 \ln P_d + b_2 \ln P_p + b_3 \ln Y + v_2$$

$$\ln I_t = c_0 + c_1 \ln NT + c_2 \ln Q + c_3 P_w + v_3$$

di mana

$S_t$  = Penawaran beras

$P_d$  = harga beras domestik

$D_t$  = Permintaan beras

$P_{d(t-1)}$  = harga beras domestik tahun

sebelumnya

$I_t$  = Impor beras

$A$  = luas lahan

$P_p$  = jumlah penduduk

$P_w$  = harga dunia

$Y$  = pendapatan perkapita

$NT$  = nilai tukar

$Q$  = produksi beras dalam negeri

Selain persamaan structural, dalam model simultan terdapat persamaan yang dikenal dengan persamaan identitas. Persamaan identitas tidak memiliki parameter, akan tetapi merupakan fungsi identitas (penambahan, pengurangan, perkalian dan pembagian) dari persamaan atau variabel lainnya.

Contoh dari persamaan identitas adalah sebagai berikut.

$$ST_t = S_t - D_t + I_t$$

Dimana  $ST_t$  adalah stok pada tahun tertentu ( $t$ ) merupakan fungsi identitas dari penawaran dikurangi permintaan ditambah dengan impor. Contoh lain dari persamaan identitas adalah persamaan keseimbangan perekonomian makro pendekatan Keynes

$$Y = C + I + G + (X - M)$$

Di mana  $Y$  adalah keseimbangan output yang merupakan fungsi identitas dari konsumsi investasi dan pengeluaran pemerintah ditambah dengan ekspor bersih.

Langkah-langkah dalam analisis persamaan simultan

1. Mengidentifikasi model

- Identifikasi model
  - Uji Simultanitas
2. Ketepatan Penaksiran Model
- Uji F
  - Uji hipotesis
  - Uji validasi model
3. Simulasi model (tambahan)
- Ex-post
  - Ex-ante

## Identifikasi Model

Sebelum mengidentifikasi model perlu diketahui tentang jenis-jenis variabel yang ada dalam model. Terdapat dua jenis variabel penting dalam persamaan simultan, yaitu variabel endogen, variabel predetermined. Variabel endogen adalah variabel yang dipengaruhi oleh variabel lain dalam model, sedangkan variabel predetermined adalah variabel yang besarnya ditentukan terlebih dahulu. Dalam persamaan, tidak digunakan variabel eksogen dengan alasan bahwa terkadang terdapat variabel eksogen di satu persamaan yang merupakan variabel endogen pada persamaan lain.

Dari persamaan diatas, dapat diketahui bahwa variabel endogen berjumlah 3 ( $P_d$  dan  $Q(S_t, D_t)$  dan  $I$ ). Dituliskan  $Q$  karena dari 3 persamaan structural yang ada dapat ditentukan dengan mencari  $Q$  keseimbangan. Dalam persamaan  $I$ , variabel  $Q$  didapatkan dari persamaan  $S_t$  dan  $D_t$ . Jumlah variabel predetermined dalam model sebanyak 6 ( $P_{d(t-1)}$ ;  $A$ ;  $P_p$ ;  $Y$ ;  $NT$ ,  $P_w$ ). Selain itu terdapat 3 konstanta ( $a_0$ ;  $b_0$ ;  $c_0$ ), 8 parameter ( $a_1$ ;  $a_2$ ;  $a_3$ ;  $b_1$ ;  $b_2$ ;  $b_3$ ;  $c_1$ ;  $c_2$ ;  $c_3$ ), dan 3 disturbance/error ( $v_1, v_2; v_3$ ).

Suatu model dikatakan model persamaan simultan jika telah memenuhi uji identifikasi model. Identifikasi model ini dilakukan dengan dua cara (1) cara penurunan (*reduced form*) dan identifikasi urutan dan ranking (*order and rank conditions of identification*). Cara penurunan (*reduced form*) dilakukan dengan mensubstitusikan persamaan-persamaan structural tersebut untuk mendapatkan parameter dari variabel predetermined. Cara kedua adalah dengan menentukan identifikasi urutan dan ranking. Cara kedua lebih banyak digunakan karena lebih mudah dan cepat. Ukuran dari uji identifikasi adalah *under identified*, *exactly identified*, dan *overidentified*. Sebuah persamaan dikatakan simultan jika uji identifikasi menunjukkan *exactly identified* dan *overidentified*.

Untuk mengetahui apakah persamaan yang dibangun merupakan persamaan simultan perlu diketahui hal-hal berikut.

1. Jumlah variabel endogen dalam model (semua persamaan), disimbolkan dengan M.
2. Jumlah variabel endogen dalam persamaan tertentu, disimbolkan dengan m.
3. Jumlah variabel predetermined dalam model, disimbolkan dengan K.
4. Jumlah variabel predetermined dalam persamaan tertentu, disimbolkan dengan k.

#### **Kondisi urutan (*Order condition*)**

Kriteria pengambilan keputusan untuk model adalah

- Model dikatakan *underidentified* jika  $M < M-1$
- Model dikatakan *just identified* jika  $M = M-1$
- Model dikatakan *over identified* jika  $M > M-1$

Kriteria pengambilan keputusan untuk persamaan adalah

- Persamaan dikatakan *underidentified* jika  $K-k < m-1$
- Persamaan dikatakan *just identified* jika  $K-k = m-1$
- Persamaan dikatakan *over identified* jika  $K-k > m-1$

Dari model di atas, diketahui bahwa model memiliki variabel endogen (M) sebanyak 3 artinya M lebih besar dari M-1; artinya model adalah simultan. Selanjutnya, pengujian masing-persamaan diketahui bahwa  $M = 3, K = 6,$

- Persamaan penawaran,  $m = 2, k = 2$  sehingga  $K-k \geq m-1 \rightarrow 6-2 \geq 2-1 \rightarrow 4 > 1$  overidentified
- Persamaan permintaan,  $m = 2, k = 2$  sehingga  $K-k \geq m-1 \rightarrow 6-2 \geq 2-1 \rightarrow 4 > 1$  overidentified
- Persamaan impor,  $m = 2, k = 2$  sehingga  $K-k \geq m-1 \rightarrow 6-2 \geq 2-1 \rightarrow 4 > 1$  overidentified

### Kondisi Ranking (The Rank Identification)

Kondisi ranking mensyaratkan bahwa jika dan hanya jika determinan dari koefisien (endogen maupun predetermined) yang tidak dicakup pada persamaan tertentu tetapi dicakup pada persamaan lainnya adalah bernilai bukan nol. Untuk mengetahui hal tersebut persamaan structural di atas perlu dibuat dalam matrik sebagai berikut.

Persamaan	1	St	Dt	It	Pd	Pdt- 1	A	Pp	Y	NT	Pw
St	a0	1	0	0	a1	a2	a3	0	0	0	0
Dt	bo	0	1	0	b1	0	0	b2	b3	0	0
I	co	c2	c2	1	0	0	0	0	0	c1	c3

Sumber: persamaan pada model simultan di atas

Ambil contoh pada persamaan pertama, nilai 0 pada persamaan pertama tetapi bukan nol pada persamaan lainnya, dapat dilihat pada tabel di atas yang berarsir. Karena variabel endogen dalam model sama dengan tiga, maka matriknya berukuran 2x2. Ambil contoh pada kolom Y dan NT.

$$\begin{bmatrix} b_3 & 0 \\ 0 & c_1 \end{bmatrix}$$

Determinan dari matrik tersebut bukan bernilai nol. Artinya bahwa persamaan tersebut adalah *over identified*.

### Uji Simultanitas

Uji simultanitas digunakan untuk mengetahui apakah variabel error dalam persamaan tertentu berpengaruh pada variabel endogen sebelah kiri. Jika ada pengaruh, maka model dikatakan simultan, jika tidak maka model dikatakan tidak simultan. Untuk mengetahui hal tersebut digunakan uji Hausman Specification error test.

Ambil contoh persamaan structural

$$\ln I_t = c_0 + c_1 \ln NT + c_2 \ln Q + c_3 P_w + v_3$$

Q merupakan variabel endogen dari persamaan berikut (St dan Dt merupakan bentuk lain dari Q) .

$$\ln S_t = a_0 + a_1 \ln P_d + a_2 \ln P_{d(t-1)} + a_3 \ln A + v_1$$

$$\ln D_t = b_0 + b_1 \ln P_d + b_2 \ln P_p + b_3 \ln Y + v_2$$

Menurut uji Hausman  $v_1$  atau  $v_2$  harus memiliki error yang berpengaruh secara statistic pada  $S_t$  atau  $D_t$ . Jika kondisi ini terpenuhi maka model tersebut bersifat simultan dan sebaliknya.

Contoh hasil analisis

$$\ln S_t = 10.2 + 6.7 \ln P_d + 5.1 \ln P_{d(t-1)} + 1.24 \ln A + 0.2 v_1$$

$$P(\text{sig}) = (0.001) \quad (0,001) \quad (0,120) \quad (0,250) \quad (0,000)$$

$$D_t = 6,4 + 1,1 \ln P_d + 1,7 \ln P_p + 2,3 \ln Y + 5,2 v_2$$

$$P(\text{sig}) = (0.001) \quad (0,121) \quad (0,014) \quad (0,140) \quad (0,000)$$

Berdasarkan hasil analisis diketahui bahwa error ( $v_1$  dan  $v_2$ ) secara statistic signifikan, sehingga disimpulkan bahwa error berpengaruh secara signifikan terhadap variabel endogen, karenanya disimpulkan bahwa ada simultansi dalam model yang dibangun tersebut.

## Penaksiran Model Simultan

Uji identifikasi (kondisi rank dan order) merupakan salah satu uji apakah model yang dibuat bersifat simultan atau tidak. Jika model tidak simultan maka ordinary least-square (OLS) merupakan penaksir terbaik. Selanjutnya, jika model adalah simultan (*exactly* dan *over identified*) maka penaksiran penggunaan OLS bersifat bias. Hal ini karena error berpengaruh pada variabel endogen. Karenanya penggunaan penaksiran OLS tidak bisa dianjurkan. Karenanya pendugaan parameter yang digunakan adalah model *two-stage least square (2SLS)*, *three-stage least square (3SLS)*, *seemingly unrelated regression estimator (SURE)*.

*Two-stage least square-TSLS (Kuadrat terkecil dua tahap).*

Analisis TSLS merupakan analisis dari persamaan structural dengan menggunakan metode kuadrat terkecil dua tahap. Artinya pada tahap pertama, mencari bentuk *reduced form* variable endogen yang ada dalam variabel endogen lainnya. Selanjutnya pada tahap kedua meregresikan variabel endogen yang sudah didapatkan ke dalam persamaan structuralnya.

Kembali pada contoh sebelumnya.

$$\ln S_t = a_0 + a_1 \ln P_d + a_2 \ln P_{d(t-1)} + a_3 \ln A + v_1$$

$$\ln D_t = b_0 + b_1 \ln P_d + b_2 \ln P_p + b_3 \ln Y + v_2$$

$$\ln I_t = c_0 + c_1 \ln NT + c_2 \ln Q + c_3 P_w + v_3$$

Dengan menggunakan metode TSLS, maka langkah-langkah yang dilakukan adalah

1. Tahap pertama: mencari melakukan analisis regresi untuk mencari P dan Q yang merupakan bentuk reduced form (bentuk turunan model simultan) tersebut dengan metode OLS.
2. Tahap kedua: melakukan analisis dengan memasukkan hasil estimasi Pd (Pd estimasi) ke persamaan structural St, dan Persamaan structural Dt. Juga pada tahap kedua juga memasukkan hasil estimasi Q pada persamaan It.

Sekarang ini beberapa software yang digunakan untuk menganalisis model simultan adalah E-Views, dan SAS/ETS, SPSS, R statistic.

Berikut ini hasil contoh analisis

The SAS System  
The SYSLIN Procedure  
Two-Stage Least Squares Estimation

Model DEMAND\_B  
Dependent Variable KB  
Label KONSUMSI BERAS

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1,3379E8	33447253	19,74	0,0171
Error	3	5084315	1694772		
Corrected Total	7	1,3887E8			

Root MSE 1301,83396 R-Square 0,96339  
Dependent Mean 32546,6213 Adj R-Sq 0,91457  
Coeff Var 3,99991

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variable Label
Intercept	1	62132,72	229996,8	0,27	0,8046	Intercept
HB	1	-28,6271	94,89160	-0,30	0,7826	HARGA ECERAN BERAS
HJGG	1	18,54475	57,92188	0,32	0,7698	HARGA JAGUNG
PDD	1	0,112492	0,032941	3,41	0,0420	POPULASI PENDUDUK
PDRB	1	0,001324	0,001723	0,77	0,4983	PRODUK DOMESTIK REGIONAL BRUTO

Durbin-Watson 2,303207  
Number of Observations 8  
First-Order Autocorrelation -0,2193

Beberapa hal yang penting terkait hasil analisis diatas adalah F-hitung, t-tes dan arah koefisien regresi (Uji hipotesis). Jika model belum sesuai dengan spesifikasi yang diinginkan peneliti, maka perlu dilakukan perbaikan model lagi. Perbaikan model dapat dilakukan dengan beberapa hal antara lain dengan menambah variabel (jika masih ada variabel yang perlu ditambahkan), menambah variabel lag (variabel yang sama pada periode sebelumnya) dan seterusnya.

### Uji Validasi.

Uji validasi digunakan untuk mengetahui apakah model yang dibangun memiliki ketepatan. Terdapat beberapa uji validasi model persamaan simultan antara lain root mean square error (RMSE), root mean square percent error (RMSPE), dan Theil's inequality (U).

- RMSE adalah rata-rata kuadrat dari perbedaan nilai taksiran dengan nilai observasi suatu variabel. Jika nilai RMSE semakin kecil maka model akan semakin bagus.
- RMSPE adalah kuadrat dari proporsi perbedaan nilai taksiran dengan nilai observasi dengan nilai observasi suatu variabel. Jika nilai semakin kecil maka model akan semakin bagus.
- Theil's inequality (U) adalah perbandingan RMSE dengan penjumlahan nilai taksiran rata-rata kuadrat nilai observasi rata-rata dan kuadrat nilai observasi rata-rata suatu model atau variabel. Jika nilai proporsi ini kurang dari atau sama dengan 0,2 maka model yang dibangun merupakan model yang baik.

Statistics of fit								
Variable	N	Mean Error	Mean % Error	Mean Abs Error	Mean Abs % Error	RMS Error	RMS % Error	R-Square
LAP	8	-0,2638	0,1487	228,9	3,5709	270,5	4,2957	0,9482
Q	8	0,0217	0,6039	0,0448	1,2216	0,0495	1,3484	0,8786
HB	8	-0,3811	0,0173	29,2722	0,8655	41,1242	1,2322	0,9249
KB	8	11,6877	0,1061	804,8	2,4473	991,9	2,9556	0,9433
QP	8	152,8	0,7852	1101,7	4,5475	1229,4	5,2105	0,9226
QB	8	91,6668	0,7852	661,0	4,5475	737,6	5,2105	0,9226

Theil Forecast Error Statistics										
Variable	N	MSE	MSE Decomposition Proportions						Inequality Coef	
			Corr (R)	Bias (UM)	Reg (UR)	Dist (UD)	Var (US)	Covar (UC)	U1	U
LAP	8	73191,9	0,97	0,00	0,00	1,00	0,01	0,99	0,0385	0,0193
Q	8	0,00246	0,95	0,19	0,01	0,80	0,00	0,80	0,0135	0,0067
HB	8	1691,2	0,97	0,00	0,09	0,91	0,18	0,82	0,0121	0,0060
KB	8	983894	0,97	0,00	0,02	0,98	0,00	1,00	0,0302	0,0151
QP	8	1511429	0,96	0,02	0,01	0,98	0,00	0,98	0,0478	0,0238
QB	8	544114	0,96	0,02	0,01	0,98	0,00	0,98	0,0478	0,0238

Dari hasil di atas diketahui bahwa nilai U thail masing-masing variabel adalah (0,03; 0,01; 0,012; 0,03;0,04; dan 0,04) artinya berada dibawah 0,2.

## SOAL-SOAL

Berikut diberikan persamaan-persamaan dalam model simultan

$$\ln S_t = a_0 + a_1 \ln P_d + a_2 \ln P_{d(t-1)} + a_3 \ln A + a_4 \ln P_{ip} + v_1$$

$$\ln D_t = b_0 + b_1 \ln P_d + b_2 \ln P_p + b_3 \ln Y + v_2$$

$$\ln I_t = c_0 + c_1 \ln NT + c_2 \ln Q + c_3 P_w + v_3$$

di mana

$S_t$  = Penawaran beras

$P_d$  = harga beras domestik

$D_t$  = Permintaan beras  
sebelumnya

$I_t$  = Impor beras

$P_p$  = jumlah penduduk

$Y$  = pendapatan perkapita

$Q$  = produksi beras dalam negeri

$P_{d(t-1)}$  = harga beras domestik tahun  
sebelumnya

$A$  = luas lahan

$P_w$  = harga dunia

$NT$  = nilai tukar

$P_{ip}$  = harga input pupuk

Coba identifikasi apakah model tersebut *underidentified*, *exactly identified*, atau *over identified* dengan metode kondisi order (*order condition*)

## DAFTAR PUSTAKA

Eviews User Guide:2<sup>nd</sup> Edition, United States of America

Green W.H., 1993, *Econometric Analysis:Second Editon*, Maxwell Macmillan International Publishing Group: New York

Gujarati Damodar N., 2003, *Basic Econometric:Fourth Edition*, Mc Graw Hill, New York

Kuncoro M., 2001, *Metode Kuantitatif*, UMP AMP YKPN, Yogyakarta

Ramanathan R., 1989, *Introductory Econometrics with Aplication: Fourth Editions*, The Dryden Press, United States of America

Sritua Arief, 1977, *Metode Penelitian Ekonomi*, Universitas Indonesia Press, Jakarta

Walter Enders, 2004, *Applied Econometric Time Series: Second Edition*, Wiley Series in Probability and Statistic, United States of America