

**“PENDUGAAN PARAMETER REGRESI
MENGGUNAKAN SOFTWARE R “**



DISUSUN OLEH :

A ROFIQI MAULANA

125090500111025

DOSEN : SAMINGUN HANDOYO, S.SI, M.CS

**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA**

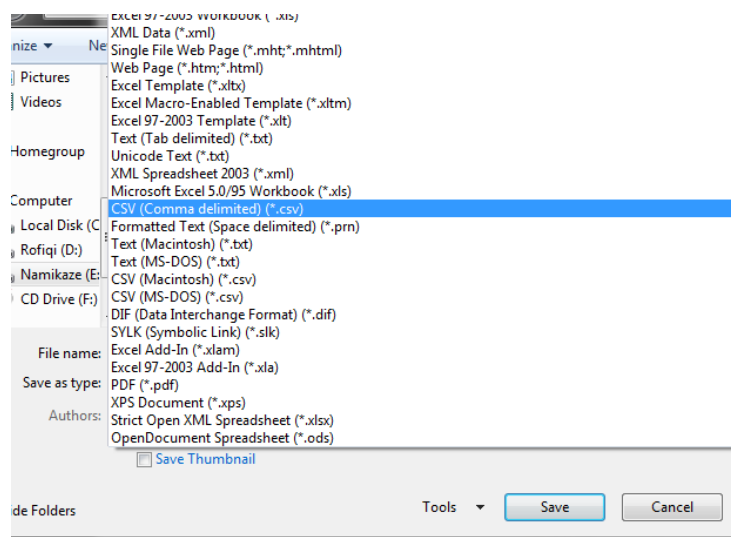
2014

A. Proses Input Data ke R

Melakukan input data ke **R** bisa dari berbagai ekstensi software lain seperti SAS, Stata, SPSS, Systat, Excel, dan CSV (Comma Separated Value). Untuk Stata dan Systat menggunakan packages “foreign”. Sedangkan untuk SPSS dan SAS dapat menggunakan packages “Hmisc”. Untuk data yang disimpan di Excel akan lebih mudah jika juga menyimpannya dalam format CSV. Untuk format Excel, membutuhkan instalasi package `xlsx`, `xlsxjar`, dan `rJava`. Input data ke R juga bisa dilakukan secara langsung menggunakan **R Commander**.

1) Input Data dari file .csv

Mengubah ekstensi file .xlsx ke .csv



Input data dari file.csv lebih mudah daripada file excel. Jika sudah mempunyai file excel, maka dapat di *save as* ke file .csv.

Melakukan input data ke R dengan cara berikut :

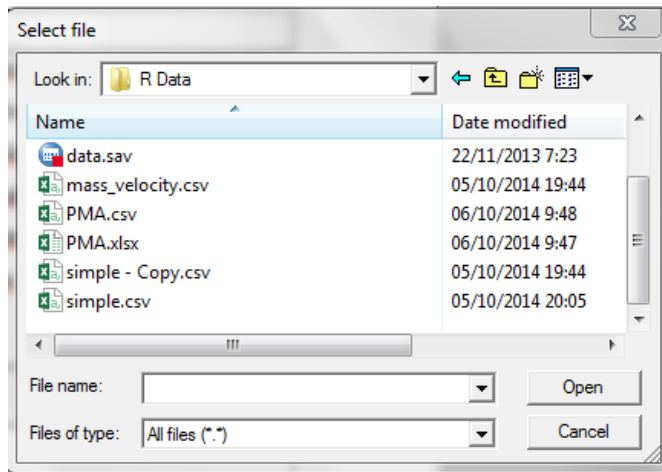
```
> #Input data dari file .csv  
> data1=read.csv(file.choose(),header=T,sep=";")
```

Penjelasan :

data1 merupakan nama yang diberikan terhadap file yang akan dibaca di **R**. Perintah *read.csv* melakukan perintah membaca data ber ekstensi .csv yang telah disimpan di suatu tempat tertentu. *file.choose()* berfungsi untuk memilih data yang akan dibuka. Perintah *header=T* jika ada nama/judul di setiap kolom. Perintah *sep=";"* merupakan

pembacaan data tiap kolom dan tanda ; (titik koma) sebagai pemisah antar data maupun antar kolom

Setelah di enter, akan muncul kotak dialog *Select file* dan pilih *file* yang akan digunakan dan klik *Open*.



Untuk menampilkan data di **R**, maka klik data1 (“ sesuai nama file yang diberikan pada coding di **R**”)

```
> #Untuk menampilkan data, ketik data1
> data1
```

	PMA	G	CPI	EX	ER
1	2561.2	26378.80	7.82	148537.0	8389
2	1974.3	28901.00	7.25	154776.2	8465
3	1836.1	30758.40	6.37	147345.7	8908
4	4835.6	35365.90	4.84	148857.4	8285
5	1563.0	29670.84	5.72	151536.7	9415
6	2246.1	30871.24	6.71	162529.9	9387
7	4814.7	29090.80	6.27	182514.0	9170
8	2053.5	35503.88	6.41	184040.3	9480
9	4282.3	26823.27	7.76	184892.6	9290
10	2645.2	28813.63	8.41	191189.9	10310
11	4735.2	34640.96	7.65	204876.2	9075
12	2370.6	44347.70	17.79	212717.2	9830
13	2916.6	30202.70	16.90	206730.1	9713
14	3608.2	37102.47	15.51	212898.2	9300
15	4586.1	35237.25	14.87	221837.2	9235
16	6059.7	45314.16	6.36	224553.5	9118
17	11135.5	31021.72	6.02	226790.9	9020
18	6382.1	37537.80	6.25	234934.5	9054
19	9046.0	38522.13	6.51	238286.8	9137
20	7105.5	39145.81	6.73	263606.1	9419

Mengaktifkan tiap-tiap variabel pada data1

```
> #Mengaktifkan variabel pada data1
> attach(data1)
The following objects are masked from data1 (position 3):

    CPI, ER, EX, G, PMA
> |
```

Penjelasan :

perintah attach() berfungsi untuk mengaktifkan variabel-variabel pada data tersebut sehingga dapat digunakan untuk proses perhitungan. Jika tidak dilakukan attach data maka akan terjadi sebagai berikut (variabel tidak ditemukan)

```
> mean(PMA)
Error in mean(PMA) : object 'PMA' not found
> summary(G)
Error in summary(G) : object 'G' not found
> |
```

2) Input data dari file excel

Melakukan input data dari file excel membutuhkan instalasi packages `xlsxjars` dan `rJava`. Untuk melakukannya, dapat langsung memilih direktori file tersebut. Sehingga diharuskan tahu letak file yang disimpan. Untuk menampilkan data dan mengaktifkan variabel-variabel di dalamnya sama seperti cara sebelumnya.

```
> #Input data dari file excel
> data1 = read.xlsx("D:/Adw.xlsx",1)
> data1
```

	No	Jumlah. Produksi
1	1	637.0
2	2	2984.0
3	3	5573.1
4	4	13635.2
5	5	16871.6
6	6	17377.2
7	7	16603.9
8	8	16207.3
9	9	12513.9
10	10	11795.2
11	11	1758.0
12	12	5789.5
13	13	15312.2
14	14	15959.2
15	15	17895.4
16	16	14314.3
17	17	14425.1
18	18	11280.2
19	19	6435.6
20	20	1194.9
21	21	465.7

B. Pendugaan Parameter Regresi

Pendugaan parameter regresi, secara langsung dapat menggunakan perintah `lm` (*linier model*) dan menggunakan **R** Commander. Hal tersebut sangat berguna untuk mengecek pembuatan suatu coding pendugaan parameter regresi telah benar atau salah. Pendugaan parameter regresi linier sederhana dapat dilakukan dengan pendekatan matriks maupun non matriks. Untuk regresi linier berganda, cukup sulit untuk menerapkan pendekatan non matriks.

1) Pendugaan Parameter Regresi Linier Sederhana

Data

```
> data2=read.table(file.choose(),header=T,sep=";")
> data2
      PMA      G
1  2561.2 26378.8
2  1974.3 28901.00
3  1836.1 30758.4
4  4835.6 35365.9
5  1563.00 29670.84
6  2246.1 30871.24
7  4814.7 29090.8
8  2053.5 35503.88
9  4282.3 26823.27
10 2645.2 28813.63
11 4735.2 34640.96
12 2370.6 44347.7
13 2916.6 30202.7
14 3608.2 37102.47
15 4586.1 35237.25
16 6059.7 45314.16
17 11135.5 31021.72
18 6382.1 37537.8
19 9046.00 38522.13
20 7105.5 39145.81
> attach(data2)
```

Hasil cek regresi

```
> fit = lm(PMA ~ G)
> coef(fit)
(Intercept)      G
-604.1372205    0.1463757
> |
```

Pendugaan Parameter Pendekatan non matriks

```
> #Membuat function untuk Pendugaan parameter koef. regresi
> b = function(x,y){
+   (sum(x*y)-(sum(x)*sum(y)/length(x)))/(sum(x*x)-length(x)*(mean(x)^2))
+ }
> b(G,PMA)
[1] 0.1463757
>
> #Membuat function pendugaan parameter untuk konstanta regresi
> a = function(x,y){
+   (mean(y)-b(x,y)*mean(x))
+ }
> a(G,PMA)
[1] -604.1372
> |
```

Penjelasan :

Untuk menghitung koef. Regresi (tentunya harus mengetahui rumus dan perhitungannya), dapat menggunakan function yang nantinya akan fungsi tersebut akan dipanggil dengan cara “ function(x,y){proses perhitungan}” dengan x = variabel independent dan y = variabel dependent.

Pendugaan Parameter Pendekatan matriks

```
> #Membuat function dari matriks x
> x = function(x){
+   cbind(seq(1),x)
+ }
>
> #Function Pendugaan parameter regresi pendekatan matriks
> beta = function(x,y){
+   solve(t(x)%*%x)%*%t(x)%*%y
+ }
> x(G)
      x
[1,] 1 26378.80
[2,] 1 28901.00
[3,] 1 30758.40
[4,] 1 35365.90
[5,] 1 29670.84
[6,] 1 30871.24
[7,] 1 29090.80
[8,] 1 35503.88
[9,] 1 26823.27
[10,] 1 28813.63
[11,] 1 34640.96
[12,] 1 44347.70
[13,] 1 30202.70
[14,] 1 37102.47
[15,] 1 35237.25
[16,] 1 45314.16
[17,] 1 31021.72
[18,] 1 37537.80
[19,] 1 38522.13
[20,] 1 39145.81
> beta(x(G),PMA)
      [,1]
-604.1372205
x      0.1463757
> |
```

Penjelasan :

Untuk melakukan pendugaan matriks menggunakan pendekatan matriks, maka dibuat terlebih dahulu matriks X yang berisi $[1 | X_1 | X_2 \dots | X_n]$. Seq(1) berfungsi memberi nilai 1 pada kolom pertama dan cbind(seq(1),x) menyatukannya dalam 1 matriks X. Pendugaan parameter pendekatan matriks menggunakan rumus $B = (X'X)^{-1}X'Y$. Mencari invers matriks menggunakan solve() dan transpose = t(). Untuk memanggil fungsi tersebut dengan x(variabel independent) dan beta(x(variabel independent),variabel dependent).

2) Pendugaan Parameter Regresi Linier Berganda

Data

```
> data3=read.table(file.choose(),header=T,sep=";")
> data3
```

	PMA	G	CPI	EX	ER
1	2561.2	26378.80	7.82	148537.0	8389
2	1974.3	28901.00	7.25	154776.2	8465
3	1836.1	30758.40	6.37	147345.7	8908
4	4835.6	35365.90	4.84	148857.4	8285
5	1563.0	29670.84	5.72	151536.7	9415
6	2246.1	30871.24	6.71	162529.9	9387
7	4814.7	29090.80	6.27	182514.0	9170
8	2053.5	35503.88	6.41	184040.3	9480
9	4282.3	26823.27	7.76	184892.6	9290
10	2645.2	28813.63	8.41	191189.9	10310
11	4735.2	34640.96	7.65	204876.2	9075
12	2370.6	44347.70	17.79	212717.2	9830
13	2916.6	30202.70	16.90	206730.1	9713
14	3608.2	37102.47	15.51	212898.2	9300
15	4586.1	35237.25	14.87	221837.2	9235
16	6059.7	45314.16	6.36	224553.5	9118
17	11135.5	31021.72	6.02	226790.9	9020
18	6382.1	37537.80	6.25	234934.5	9054
19	9046.0	38522.13	6.51	238286.8	9137
20	7105.5	39145.81	6.73	263606.1	9419

Hasil cek koefisien regresi

```
> lm(PMA ~ G+CPI+EX+ER)
```

Call:

```
lm(formula = PMA ~ G + CPI + EX + ER)
```

Coefficients:

(Intercept)	G	CPI	EX	ER
1.285e+04	-1.029e-01	-2.039e+02	7.701e-02	-1.991e+00

```
> |
```


Pendugaan Parameter Regresi Linier Berganda

```
> #Function untuk membuat matriks x
> x = function(x1,x2,x3,x4){
+   cbind(seq(1),x1,x2,x3,x4)
+ }
>
> #Function Untuk pendugaan parameter regresi
> beta = function(x,y){
+   solve(t(x)%*%x)%*%t(x)%*%y
+ }
>
```

Penjelasan :

Terdapat sedikit perbedaan source code pendugaan parameter regresi linear sederhana terhadap regresi linier berganda yaitu pada function untuk membuat matriks X. Terdapat penambahan kolom matriks sesuai banyaknya variabel independent. Sedangkan function pada pendugaan parameternya sama dengan cara di regresi linier sederhana.

Matriks X dan nilai penduga beta

```
> x(G,CPI,EX,ER)
      x1      x2      x3      x4
[1,] 1 26378.80  7.82 148537.0  8389
[2,] 1 28901.00  7.25 154776.2  8465
[3,] 1 30758.40  6.37 147345.7  8908
[4,] 1 35365.90  4.84 148857.4  8285
[5,] 1 29670.84  5.72 151536.7  9415
[6,] 1 30871.24  6.71 162529.9  9387
[7,] 1 29090.80  6.27 182514.0  9170
[8,] 1 35503.88  6.41 184040.3  9480
[9,] 1 26823.27  7.76 184892.6  9290
[10,] 1 28813.63  8.41 191189.9 10310
[11,] 1 34640.96  7.65 204876.2  9075
[12,] 1 44347.70 17.79 212717.2  9830
[13,] 1 30202.70 16.90 206730.1  9713
[14,] 1 37102.47 15.51 212898.2  9300
[15,] 1 35237.25 14.87 221837.2  9235
[16,] 1 45314.16  6.36 224553.5  9118
[17,] 1 31021.72  6.02 226790.9  9020
[18,] 1 37537.80  6.25 234934.5  9054
[19,] 1 38522.13  6.51 238286.8  9137
[20,] 1 39145.81  6.73 263606.1  9419

> beta(x(G,CPI,EX,ER),PMA)
      [,1]
x1 1.285331e+04
x2 -1.029397e-01
x3 -2.038508e+02
x4  7.701077e-02
x5 -1.990833e+00
```

C. Kesimpulan

1. Hasil dari perhitungan manual menggunakan coding sama dengan perhitungan secara langsung. Pada regresi linier sederhana didapatkan model berikut

$$\text{PMA} = -604,1372205 + 0,1463757G$$

2. Hasil dari perhitungan manual menggunakan coding sama dengan perhitungan secara langsung. Pada regresi linier berganda didapatkan model berikut

$$\text{PMA} = 12853,31 - 0,1029G - 203,85\text{CPI} + 0,07701\text{EX} - 1,9908\text{ER}$$