

ISTIANAH SURURY, SKM, M.EPID



BUKU PRAKTIKUM STATA



**PROGRAM STUDI KESEHATAN MASYARAKAT
FAKULTAS KESEHATAN MASYARAKAT
UNIVERSITAS MUHAMMADIYAH JAKARTA
2020**

DAFTAR ISI

TIM PENYUSUN.....	iii
DAFTAR ISI.....	iii
DAFTAR GAMBAR.....	iv
DAFTAR TABEL.....	iv
DAFTAR GRAFIK.....	iv
CHI-SQUARE	7
REGRESI LOGISTIK	8
REGRESI LOGISTIK BERGANDA	9
REGRESI LINIER BERGANDA.....	12
1. Uji Normalitas.....	12
2. Seleksi Bivariat	13
3. Analisis Regresi Linier Berganda	16
4. Uji Asumsi Regresi Linier	18
a) Asumsi Eksistensi	18
b) Asumsi Normalitas	19
c) Asumsi Homoscedascity.....	20
d) Asumsi Linearitas	20
e) Asumsi Independensi.....	21
f) Asumsi Kolinearitas	21
5. Pemodelan Regresi Linier.....	21
DAFTAR PUSTAKA	23

DAFTAR GAMBAR

Gambar 1. Kolom Variabel.....	5
Gambar 2. Kolom Command.....	6
Gambar 3. Tabel Deskriptif.....	6
Gambar 4. Tabel Hasil Uji Chi Square.....	7
Gambar.5 Tabel Regresi Logistik.....	8
Gambar.6 Tabel Regresi Logistik Berganda.....	9
Gambar. 7 Tabel Regresi Logistik Berganda (Pemodelan Akhir).....	10
Gambar 10. Asumsi Eksistensi.....	19
Gambar 11. Uji Asumsi dengan Shapiro Wilk.....	19
Gambar 12. Asumsi Homoscedasity.....	20
Gambar 13. Asumsi Kolinearitas.....	21

DAFTAR TABEL

Tabel 1. Seleksi bivariat antara <i>birth weight</i> dengan <i>low birth weight</i>	13
Tabel 2. Seleksi bivariat antara <i>birth weight</i> dengan <i>race</i>	14
Tabel 3. Seleksi bivariat antara <i>birth weight</i> dengan <i>smoke</i>	14
Tabel 4. Seleksi bivariat antara <i>birth weight</i> dengan <i>history premature labor</i>	15
Tabel 5. Seleksi bivariat antara <i>birth weight</i> dengan <i>history of hypertention</i>	15
Tabel 6. Seleksi bivariat antara <i>birth weight</i> dengan <i>history of uterine irritability</i>	16
Tabel 7. Seleksi bivariat antara <i>birth weight</i> dengan <i>visits physician 1st trimeste</i>	16
Tabel 8. Regresi Linier Berganda.....	17
Tabel 9. Regresi Linier Berganda.....	17
Tabel 10. Asumsi Linearitas.....	20
Tabel 11. Pemodelan Akhir Uji Regresi Linier.....	22

DAFTAR GRAFIK

Grafik 1. Grafik Uji Normalitas.....	12
Grafik 2. Uji Asumsi dengan Histogram pada Variabel Residual.....	19

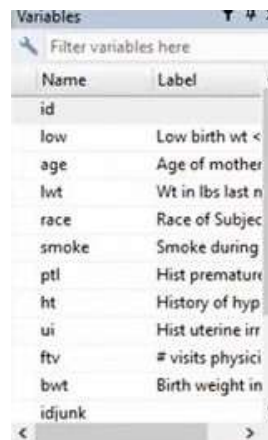
Stata adalah suatu perangkat lunak statistika terintegrasi dan lengkap yang menyediakan berbagai fitur untuk analisis data, manajemen data, dan grafik. Penyimpanan Output dalam stata tidak dapat membuat output secara otomatis, tetapi kita harus menyimpan diawal sebelum menganalisis atau disebut juga dengan *log file*. Berikut merupakan cara penyimpanan output (*log file*):

1. Klik "*file*" (bagian kiri atas)
2. Klik "*log*" kemudian pilih "*Begin*"
3. Kemudian buatlah file dengan format nama bebas dalam bentuk file *log*
4. File sudah tersimpan dan sudah siap menganalisis

Sebelum melakukan analisis, bukalah file data yang akan di analisis dengan cara sebagai berikut:

1. Klik gambar berbentuk file (bagian kiri atas), kemudian buka file yang akan di analisis
2. Dapat dilihat dibagian bawah yang bertulisan "*command*" merupakan tempat anda menuliskan perintah untuk menganalisis data atau dapat digunakan pada fitur yang berada dibagian atas.
3. Pada bagian kanan atas dengan tulisan "*variables*" itu merupakan jenis variabel apa saja yang dapat kita analisis.

Gambar 1. Kolom Variabel



Setelah dibuka data set yang akan di analisis, jika kita ingin mengetahui gambaran dan kelengkapan data mengenai data set yang akan di analisis yaitu anda dapat mengetik *sum* pada kolom "*command*" dan akan muncul gambaran data mulai dari rata-rata, standar deviasi, nilai minimal maksimum, dan juga jumlah sample pada setiap variabel.

Gambar 2. Kolom Command

Apabila kita ingin melihat gambaran secara deskriptif dari data set yang akan dianalisis dari satu variabel, maka ketik pada kolom "command" → ex variabel; tab low (sesuaikan dengan variabel pada data set) → enter. Maka kita dapat melihat gambaran deskriptif atau dalam bentuk persentase dari variabel yang ingin kita analisis.

Gambar 3. Tabel Deskriptif

The screenshot shows the Stata Command window with the following commands:

```

1. list
2. use "C:\Users\Aria\Downloads\low-ha20"
3. tab low
4. tab low (Described)
5. label define low 0 "normal"
6. tab race
7. tab pt
8. tab smoke
9. tab pt
10. tab pt
11. tab w
12. tab ft
13. tab low smoke
14. tab smoke low
15. tab def low 0 "yes"
16. tab def smoke 0 "n."
17. tab smoke
18. tab smoke low
19. label define low 0 "110"
20. label define smoke 0 "110"
21. tab race low

```

The first table, titled "Smoke during pregnancy?", shows the following data:

	low	Total
0	88	29
1	88	28
Total	138	88

The second table, titled "Low birth wt <1000 grams?", shows the following data:

	low	Total
0	73	22
1	18	11
2	42	22
Total	138	88

Cara menuliskan label dari setiap variabel, kita dapat melakukan dengan cara yaitu mengetik → label define low 0 "normal" → enter. Untuk menentukan label, maka disesuaikan dengan koding dari masing-masing variabel.

CHI-SQUARE

Untuk menganalisis bivariat menggunakan uji chi-square dapat dilakukan dengan cara sebagai berikut:

1. Sebelum melakukan analisis, tentukan variabel dependen dan independen anda yang akan anda lakukan analisis. Ex variabel; low smoke (sesuaikan dengan variabel dependen dan independen anda)
2. Lalu, ketik pada kolom "*command*" → "tab low smoke" atau "tab smoke low" (sesuaikan dengan nama variabel anda) → enter. (akan muncul tabel 2x2)

Cara diatas hanya dapat melihat jumlah atau deskriptif dari hasil uji chi square saja, apabila ingin melihat nilai *p value* dan nilai OR (ex: *study case control*) yaitu ketik pada kolom "*command*" → "cc low smoke" → enter. Lalu akan muncul tabel 2x2 sesuai dengan variabel yang akan kita analisis. Pada tabel ini akan didapatkan nilai *p value* dan nilai OR.

Gambar 4. Tabel Hasil Uji Chi Square

The screenshot shows the STATA command window with the following commands and results:

```

1. tab low
2. tab smoke
3. cc smoke low
  
```

The results for `tab low` are:

	Exposed	Unexposed	Total	Exposed
Cases	82	29	111	0.5045
Controls	46	89	135	0.3363
Total	128	118	246	0.5203

The results for `cc smoke low` are:

	Exposed	Unexposed	Total	Exposure
Cases	35	46	81	0.4321
Controls	28	89	117	0.2393
Total	63	135	198	0.3182

The Odds Ratio (OR) for the case control analysis is 2.02194, with a 95% confidence interval of 1.02902 to 3.98994. The p-value is 0.0402.

Catatan : cc → *Case Control*

REGRESI LOGISTIK BERGANDA

Dalam melakukan uji multivariat, dapat menggunakan uji regresi logistik berganda sesuai dengan nilai *p value* pada masing-masing variabel yang masuk untuk di uji multivariat (*p value* < 0,25). Uji regresi logistik pada stata dapat dilakukan dengan cara sebagai berikut:

1. Sebelum melakukan analisis, tentukan variabel dependen dan independen anda yang akan anda lakukan analisis multivariat (sesuaikan dengan variabel dependen dan independen anda). Ex variabel yang masuk seleksi multivariat yaitu low, smoke, race, lwt, ptl, ui
2. Ketik pada kolom "command" → "logistic low smoke race lwt ptl ui" → enter.

Gambar.6 Tabel Regresi Logistik Berganda

low	smoke	race	lwt	ptl	ui
0.000000	0.000000	-0.28	0.000	0.000000	0.000000
0.000000	0.000000	0.000	0.000	0.000000	0.000000
0.000000	0.000000	0.000	0.000	0.000000	0.000000
0.000000	0.000000	0.000	0.000	0.000000	0.000000
0.000000	0.000000	0.000	0.000	0.000000	0.000000
0.000000	0.000000	0.000	0.000	0.000000	0.000000

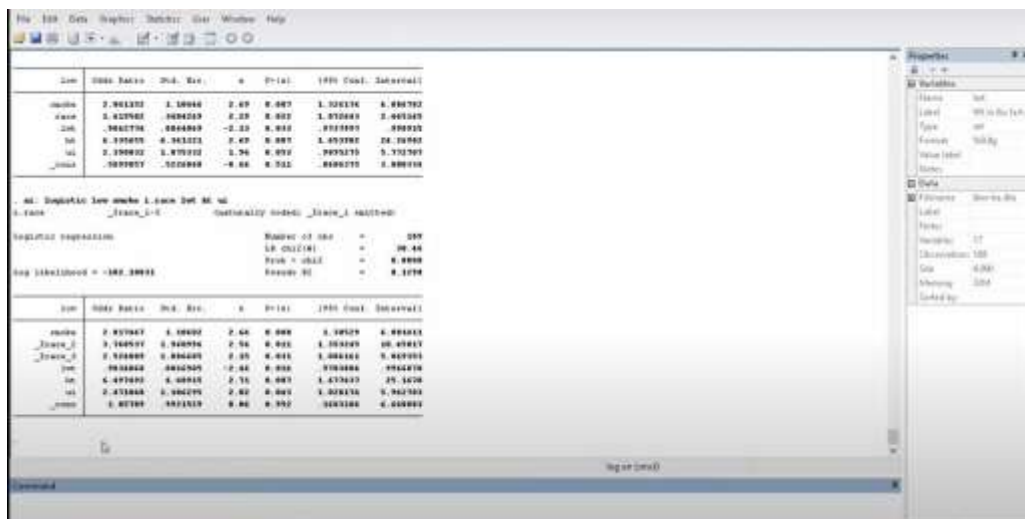
Pada uji multivariat dapat dilihat dari gambar diatas nilai *p value* dari masing-masing variabel yaitu dapat dilihat dari nilai $P > |z|$. Pada gambar diatas merupakan pemodelan pertama. Jika ingin diketahui model akhir dari uji multivariat ini yaitu dengan cara mengeluarkan variabel yang memiliki nilai *p value* > 0,05 dengan membandingkan juga nilai OR. Apabila

ditemukan perubahan nilai $OR > 10\%$ setelah mengeluarkan variabel, maka variabel tersebut harus dimasukkan kembali pada pemodelan dan menjadi variabel konfounding.

Cara untuk mengeluarkan variabel yang memiliki nilai $p\ value > 0,25$ dapat menggunakan cara yang sama seperti diatas, yaitu ketik pada kolom "command" \rightarrow "logistik (jangan cantumkan variabel yang memiliki $p\ value > 0,05$) low, smoke, race, lwt, ptl, ui" \rightarrow enter. Apabila ditemukannya variabel yang memiliki lebih dari 2 kategori, kemudian ingin diketahui nilai $p\ value$ dan OR dari masing-masing kategori tersebut, dapat dilakukan cara sebagai berikut:

1. Sebelum melakukan analisis, tentukan variabel dependen dan independen anda yang akan anda lakukan analisis multivariat (sesuaikan dengan variabel dependen dan independen anda). Ex variabel yang masuk seleksi multivariat yaitu low, smoke, race, lwt, ptl, ui.
2. Lalu, pada variabel yang memiliki lebih dari 2 kategori, dapat ditambahkan "i.variabel" ex: i.race pada kolom "command" \rightarrow "xi : logistik low smoke i.race lwt ptl ui" \rightarrow enter. Maka akan muncul tabel.

Gambar. 7 Tabel Regresi Logistik Berganda (Pemodelan Akhir)



Interpretasi :

Berdasarkan gambar diatas, didapatkan bahwa faktor risiko ht (hipertensi) memiliki nilai OR sebesar 6,497 yang artinya ibu yang memiliki riwayat hipertensi memiliki odds 6,497 kali lebih tinggi memiliki bayi BBLR dibandingkan pada ibu yang tidak memiliki riwayat hipertensi, dan

variabel lainnya sama seperti diatas interpretasinya. (diurutkan dari mulai nilai OR tertinggi sampai terendah).

REGRESI LINIER BERGANDA

Regresi linier berganda merupakan perluasan dari regresi linier sederhana. Analisisnya berfokus untuk mengetahui hubungan antara beberapa variabel bebas dengan satu variabel terikat. Uji ini digunakan untuk melihat besar pengaruh setiap variabel bebas terhadap variabel terikat. Untuk melakukan regresi linier, variabel terikat bersifat numerik, dan variabel bebas juga bersifat numerik (minimal pada satu variabel).

1. Uji Normalitas

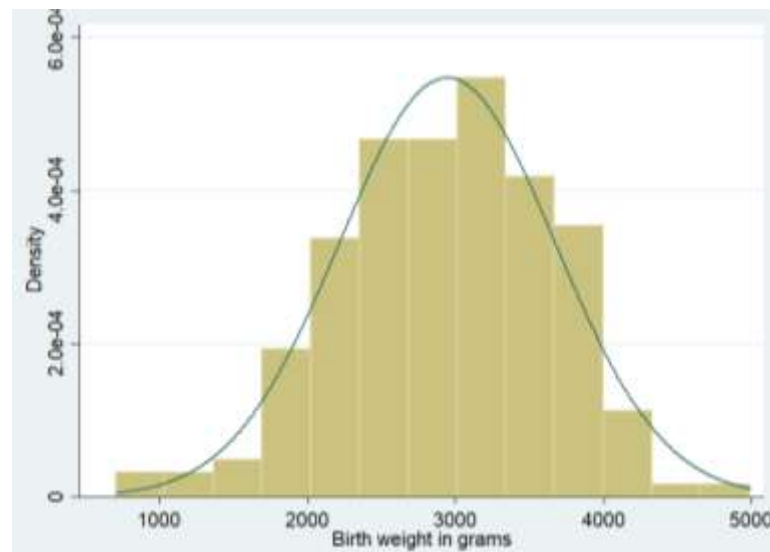
Uji regresi linier dapat dilakukan pada data yang berdistribusi normal. Maka dari itu langkah pertama yang dilakukan ialah melakukan Uji Normalitas pada variabel terikat. Dengan cara:

➤ Ketik pada kolom *Command*

“histogram (variabel terikat), normal” → “histogram bwt, normal”

Output:

Grafik 1. Grafik Uji Normalitas



Berdasarkan grafik, data dinyatakan berdistribusi normal karena kurva mempunyai bentuk menyerupai lonceng.

2. Seleksi Bivariat

Setelah data telah dipastikan berdistribusi normal, selanjutnya dilakukan seleksi bivariat pada semua variabel bebas. Hal yang ditetapkan pada seleksi bivariat ini ialah nilai p pada variabel bebas harus $< 0,25$ untuk lanjut pada tahap analisis multivariat.

Langkah-langkah:

- Ketik pada kolom *Command*
“regress + var terikat + var bebas” → “regress bwt lwt” (untuk variabel bebas dengan 2 kategori)
- Klik *enter*, lalu akan muncul *output*:

Tabel 1. Seleksi bivariat antara *birth weight* dengan *low birth weight*

```
. regress bwt lwt
```

Source	SS	df	MS	Number of obs	=	189
Model	3448881.3	1	3448881.3	F(1, 187)	=	6.69
Residual	96468171.3	187	515872.574	Prob > F	=	0.0105
				R-squared	=	0.0345
				Adj R-squared	=	0.0294
Total	99917052.6	188	531473.684	Root MSE	=	718.24

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lwt	4.429264	1.713025	2.59	0.010	1.049927 7.8086
_cons	2369.672	228.4306	10.37	0.000	1919.04 2820.304

Dilihat dari nilai $P < 0,25$ maka variabel lwt dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command*
“regress + var terikat + i.var bebas” → “regress bwt i.race” (untuk variabel bebas dengan > 2 kategori)
- Klik *enter*, lalu akan muncul *output*:

Tabel 2. Seleksi bivariat antara *birth weight* dengan *race*

. regress bwt i.race

Source	SS	df	MS	Number of obs	=	189
Model	5070607.63	2	2535303.82	F(2, 186)	=	4.97
Residual	94846445	186	509927.124	Prob > F	=	0.0079
				R-squared	=	0.0507
				Adj R-squared	=	0.0405
Total	99917052.6	188	531473.684	Root MSE	=	714.09

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
race					
2	-384.0473	157.8744	-2.43	0.016	-695.5019 -72.59266
3	-299.7247	113.6776	-2.64	0.009	-523.9878 -75.4615
_cons	3103.74	72.88169	42.59	0.000	2959.959 3247.521

Dilihat dari nilai $P < 0,25$ maka variabel *race* dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command*
“regress bwt smoke”
- Klik *enter*, lalu akan muncul *output*:

Tabel 3. Seleksi bivariat antara *birth weight* dengan *smoke*

. regress bwt smoke

Source	SS	df	MS	Number of obs	=	189
Model	3573406.24	1	3573406.24	F(1, 187)	=	6.94
Residual	96343646.4	187	515206.665	Prob > F	=	0.0092
				R-squared	=	0.0358
				Adj R-squared	=	0.0306
Total	99917052.6	188	531473.684	Root MSE	=	717.78

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
smoke	-281.7133	106.9687	-2.63	0.009	-492.7338 -70.69274
_cons	3054.957	66.93324	45.64	0.000	2922.915 3186.998

Dilihat dari nilai $P < 0,25$ maka variabel *smoke* dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command*
“regress bwt ptl”
- Klik *enter*, lalu akan muncul *output*:

```
. regress bwt ptl
```

Source	SS	df	MS	Number of obs	=	189
Model	2392205.02	1	2392205.02	F(1, 187)	=	4.59
Residual	97524847.6	187	521523.249	Prob > F	=	0.0335
				R-squared	=	0.0239
				Adj R-squared	=	0.0187
Total	99917052.6	188	531473.684	Root MSE	=	722.17

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ptl	-228.6506	106.7603	-2.14	0.034	-439.26 -18.04114
_cons	2989.418	56.53494	52.88	0.000	2877.89 3100.947

Tabel 4. Seleksi bivariat antara *birth weight* dengan *history premature labor*

Dilihat dari nilai $P < 0,25$ maka variabel pwt dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command*
“regress bwt ht”
- Klik *enter*, lalu akan muncul *output*:

Tabel 5. Seleksi bivariat antara *birth weight* dengan *history of hypertension*

```
. regress bwt ht
```

Source	SS	df	MS	Number of obs	=	189
Model	2132014.49	1	2132014.49	F(1, 187)	=	4.08
Residual	97785038.2	187	522914.643	Prob > F	=	0.0449
				R-squared	=	0.0213
				Adj R-squared	=	0.0161
Total	99917052.6	188	531473.684	Root MSE	=	723.13

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ht	-435.5607	215.7094	-2.02	0.045	-861.0973 -10.02413
_cons	2972.311	54.35366	54.68	0.000	2865.086 3079.536

Dilihat dari nilai $P < 0,25$ maka variabel ht dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command*
“regress bwt ui”
- Klik *enter*, lalu akan muncul *output*:

```
. regress bwt ui
```

Source	SS	df	MS	Number of obs	=	189
Model	8028747.44	1	8028747.44	F(1, 187)	=	16.34
Residual	91888305.2	187	491381.311	Prob > F	=	0.0001
				R-squared	=	0.0804
				Adj R-squared	=	0.0754
Total	99917052.6	188	531473.684	Root MSE	=	700.99

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ui	-580.1801	143.5318	-4.04	0.000	-863.3298 -297.0304
_cons	3030.609	55.24543	54.86	0.000	2921.624 3139.593

Tabel 6. Seleksi bivariat antara *birth weight* dengan *history of uterine irritability*

Dilihat dari nilai $P < 0,25$ maka variabel *ui* dapat masuk dalam pemodealan multivariat.

- Ketik pada kolom *Command* “regress bwt ftv”
- Klik *enter*, lalu akan muncul *output*:

Tabel 7. Seleksi bivariat antara *birth weight* dengan *visits physician 1st trimeste*

```
. regress bwt ftv
```

Source	SS	df	MS	Number of obs	=	189
Model	339165.165	1	339165.165	F(1, 187)	=	0.64
Residual	99577887.5	187	532502.072	Prob > F	=	0.4258
				R-squared	=	0.0034
				Adj R-squared	=	-0.0019
Total	99917052.6	188	531473.684	Root MSE	=	729.73

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ftv	40.09714	50.24218	0.80	0.426	-59.01716 139.2114
_cons	2912.833	66.38875	43.88	0.000	2781.866 3043.8

Dilihat dari nilai $P > 0,25$ maka variabel *ftv* tidak dapat masuk dalam pemodealan multivariat.

3. Analisis Regresi Linier Berganda

- a) Seleksi multivariat variabel bebas

Seleksi ini dilakukan untuk melihat variabel bebas yang paling berpengaruh pada variabel terikat. Metode yang akan digunakan ialah *method backward* yang cara kerjanya melakukan seleksi bertahap pada variabel bebas. Variabel bebas yang akan dikeluarkan ialah variabel dengan korelasi parsial terkecil dengan variabel terikat. Eliminasi akan dilakukan pada variabel dengan nilai $p > 0,10$.

Langkah-langkah:

- Ketik pada kolom *Command*

“regress + var terikat + semua var bebas” → “regress bwt lwt i.race smoke ptl ht ui”
 Variabel “ftv” tidak dimasukkan, karena pada seleksi bivariat nilai p variabel ftv > 0,25.

➤ Klik *enter*, lalu akan muncul *output*:

Tabel 8. Regresi Linier Berganda

```
regress bwt lwt i.race smoke ptl ht ui
```

Source	SS	df	MS	Number of obs	=	189
Model	24125126.7	7	3446446.67	F(7, 181)	=	8.23
Residual	75791926	181	418739.922	Prob > F	=	0.0000
				R-squared	=	0.2415
				Adj R-squared	=	0.2121
Total	99917052.6	188	531473.684	Root MSE	=	647.1

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lwt	4.154202	1.685962	2.46	0.015	.8275338 7.480869
race					
2	-475.2293	145.8684	-3.26	0.001	-763.0505 -187.4081
3	-347.4696	112.6619	-3.08	0.002	-569.7691 -125.17
smoke	-344.8647	105.314	-3.27	0.001	-552.6657 -137.0636
ptl	-53.77393	100.5334	-0.53	0.593	-252.1421 144.5942
ht	-582.1721	200.0786	-2.91	0.004	-976.9585 -187.3856
ui	-509.4444	137.7991	-3.70	0.000	-781.3437 -237.545
_cons	2851.922	245.5702	11.61	0.000	2367.373 3336.47

Berdasarkan tabel, variabel ptl tidak memenuhi persyaratan (p > 0,10) sehingga tidak terseleksi karena memiliki korelasi parsial terkecil terhadap variabel bwt.

➤ Ketik pada kolom *Command*

“regress bwt lwt i.race smoke ht ui”

➤ Klik *enter*, lalu akan muncul *output*:

Tabel 9. Regresi Linier Berganda

```
regress bwt lwt i.race smoke ht ui
```

Source	SS	df	MS	Number of obs	=	189
Model	24005323.8	6	4000887.3	F(6, 182)	=	9.59
Residual	75911728.8	182	417097.411	Prob > F	=	0.0000
				R-squared	=	0.2403
				Adj R-squared	=	0.2152
Total	99917052.6	188	531473.684	Root MSE	=	645.83

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lwt	4.239475	1.675113	2.53	0.012	.9343367 7.544613
race					
2	-475.808	145.578	-3.27	0.001	-763.0456 -188.5703
3	-349.9976	112.3417	-3.12	0.002	-571.6572 -128.3379
smoke	-354.8997	103.426	-3.43	0.001	-558.9679 -150.8315
ht	-585.1125	199.6104	-2.93	0.004	-978.9606 -191.2644
ui	-524.4395	134.6525	-3.89	0.000	-790.1201 -258.7588
_cons	2837.638	243.6347	11.65	0.000	2356.926 3318.35

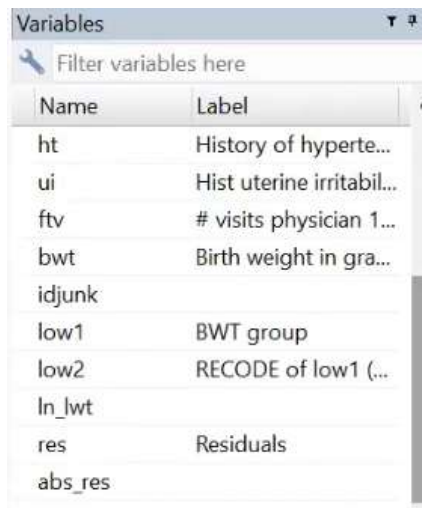
Berdasarkan tabel, semua variabel bebas sudah memenuhi nilai p untuk multivariat. Jika dilihat dari nilai p, maka **variabel ui (*hist of uterine irritability*) paling berpengaruh** karena memiliki nilai p paling kecil. Selain itu dilihat dari nilai koefisien, **variabel ht (*hipertensi*) memiliki pengaruh paling besar** yang artinya ibu dengan riwayat hipertensi memiliki pengaruh terhadap berat badan anaknya sebesar 585 gram.

4. Uji Asumsi Regresi Linier

Uji asumsi regresi linier yang diuji adalah pada variabel residu dari model regresi, dengan tahapan berikut:

1. Tahap pertama yang dilakukan yaitu membuat variable residual baru dengan cara command → **“predict res, r”**
2. Membuat variable residual absoult dengan cara, command → **“gen abs_res=abs(res)”** digunakan untuk uji linieritas.

Gambar 9. Membuat Varabel Baru Residual



Name	Label
ht	History of hyperte...
ui	Hist uterine irritabil...
ftv	# visits physician 1...
bwt	Birth weight in gra...
idjunk	
low1	BWT group
low2	RECODE of low1 (...)
ln_lwt	
res	Residuals
abs_res	

Catatan: Untuk uji asumsi yang dilakukan bukan hanya menilai variabel dependen dan independen saja tetapi residu pemodelan yang dibuat. Maka dari itu, uji asumsi dilakukan setelah pemodelan.

a) Asumsi Eksistensi

Tahapan yang dilakukan yaitu membuat summarize dari variabel residual dengan klik pada bagian kolom command → **“sum res”** → untuk melihat eksistensi dari mean dan standar deviasi.

Gambar 10. Asumsi Eksistensi

```
. sum res
```

Variable	Obs	Mean	Std. Dev.	Min	Max
res	189	-4.72e-07	635.4414	-1843.33	1630.907

log on (smf)

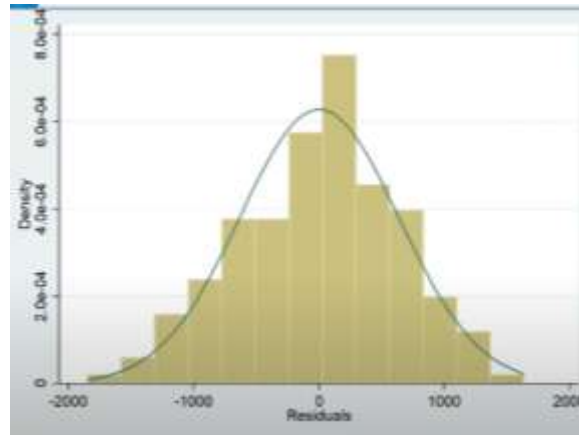
Command

Catatan: **Syarat uji eksistensi**, di lihat dari nilai mean dan standar deviasi mendekati 0 maka asumsi eksistensinya terpenuhi.

Penjelasan : pada gambar 10 diketahui bahwa memenuhi asumsi eksistensi dengan mean - 4.72 dan standar deviasi 635.44

b) Asumsi Normalitas

Pada uji ini dilakukan uji normalitas pada variabel residual dengan cara, command→ **“histogram res, normal”**.

Grafik 2. Uji Asumsi dengan Histogram pada Variabel Residual

Asumsi normalitas dapat dilakukan dengan diagram lonceng atau shapiro wilk dan uji kolmogorov smirnov. Bila, menggunakan nilai shapiro wilk, dapat dilakukan dengan comand → **“swilk res”**.

Gambar 11. Uji Asumsi dengan Shapiro Wilk

```
swilk res
```

Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	z	Prob>z
res	189	0.99400	0.852	-0.366	0.64300

Catatan: >0.05 Normal, dapat dilihat dari probability z score.

Penjelasan: dapat dilihat probabilitas variabel residu > 0.05 berarti berdistribusi normal pada uji shapiro wilks.

c) Asumsi Homoscedascity

Pada tahapan ini dilakukan dengan cara klik pada bagian kolom command → “**estat hettest**”.

Gambar 12. Asumsi Homoscedascity

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of bwt

      chi2(1)      =      0.06
      Prob > chi2   =      0.8009
```

Catatan: asumsi ini dilihat dengan scarlet plot atau dapat dilihat dari nilai chi square. Bila nilai >0.05 berarti **homogen** dan nilai < 0.05 berarti **heterogen**.

Penjelasan: pada uji asumsi homoscedascity didapatkan homogen dengan probability chi square 0.8009.

d) Asumsi Linearitas

Asumsi linieritas dilakukan untuk melihat hubungan antara variabel prediktor terhadap variabel outcome dengan analisis anova. Dapat dikatakan hubungan ini linearitas dengan nilai $p < 0.05$. selain itu uji asumsi ini dapat dilakukan dengan prosedur “**compare mean**”. Asumsi linearitas dapat dilahat dari nilai probabilitinya, dengan cara klik pada bagian comand → “**regress bwt lwt i.race smoke ht ui**”

Tabel 10. Asumsi Linearitas

```
regress bwt lwt i.race smoke ht ui
```

Source	SS	df	MS	Number of obs =	189
Model	24005323.8	6	4000887.3	F(6, 182) =	9.59
Residual	75911728.8	182	417097.411	Prob > F =	0.0000
Total	99917052.6	188	531473.684	R-squared =	0.2403
				Adj R-squared =	0.2152
				Root MSE =	645.83

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lwt	4.239475	1.675113	2.53	0.012	.9343367 7.544613
race					
2	-475.808	145.578	-3.27	0.001	-763.0456 -188.5703
3	-349.9976	112.3417	-3.12	0.002	-571.6572 -128.3379
smoke					
ht	-354.8997	103.426	-3.43	0.001	-558.9679 -150.8315
ui	-585.1125	199.6104	-2.93	0.004	-978.9606 -191.2644
_cons	-524.4395	134.6525	-3.89	0.000	-790.1201 -258.7588
	2837.638	243.6347	11.65	0.000	2356.926 3318.35

Penjelasan: Pada gambar di atas dapat dilihat dari nilai probabilitas yang menunjukkan nilai $p < 0.05$ yang berarti terdapat hubungan yang linier antara variabel residu dengan variabel outcome.

e) Asumsi Independensi

Asumsi independensi di gunakan pada studi longitudinal untuk melihat pajanan pada outcome, asumsi ini dilakukan untuk melihat waktu. Tidak perlu dilakukan independensi bila tidak ada variabel waktu.

f) Asumsi Kolinearitas

Asumsi ini dilakukan dengan cara klik pada bagian kolom comand → “estat vif”

Gambar 13. Asumsi Kolinearitas

```
. estat vif
```

Variable	VIF	1/VIF
lwt	1.18	0.845540
race		
2	1.14	0.877701
3	1.31	0.764157
smoke	1.15	0.865985
ht	1.07	0.931491
ui	1.04	0.964465
Mean VIF	1.15	

Catatan: bila nilai VIF (Varian Factors) atau tolerance > 10 maka terjadi multikolinearitas, untuk memenuhi asumsi kolinearitas nilai $VIF \leq 10$.

Penjelasan : Pada tabel didapatkan bahwa nilai VIF sebesar 1.15 yang berarti tidak ada multikolinearitas di dalam pemodelan regresi linier tersebut.

5. Pemodelan Regresi Linier

Pada model regresi linear memformulasikan hubungan antara satu variabel terikat kontinyu Y dengan beberapa variabel bebas/prediktor X_i yang berskala interval. Model regresi Poisson memformulasikan hubungan antara beberapa variabel bebas kategorikal atau konyinyu dengan Log (rate) terjadinya kejadian Y (binary).

```
regress bwt lwt i.race smoke ht ui
```

Source	SS	df	MS	Number of obs =	189
Model	24005323.8	6	4000887.3	F(6, 182)	= 9.59
Residual	75911728.8	182	417097.411	Prob > F	= 0.0000
				R-squared	= 0.2403
				Adj R-squared	= 0.2152
Total	99917052.6	188	531473.684	Root MSE	= 645.83

bwt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lwt	4.239475	1.675113	2.53	0.012	.9343367 7.544613
race					
2	-475.808	145.578	-3.27	0.001	-763.0456 -188.5703
3	-349.9976	112.3417	-3.12	0.002	-571.6572 -128.3379
smoke	-354.8997	103.426	-3.43	0.001	-558.9679 -150.8315
ht	-585.1125	199.6104	-2.93	0.004	-978.9606 -191.2644
ui	-524.4395	134.6525	-3.89	0.000	-790.1201 -258.7888
_cons	2837.638	243.6347	11.65	0.000	2356.926 3318.35

Tabel 11. Pemodelan Akhir Uji Regresi Linier

Berat bayi= 2837,64+ (4,24 lwt)+ (-475,81 black)+(-349,9 other)+(-354,9 smoke)+(-585,11 ht)+(-524,44 ui)

Interpretasi: Model ini dapat memprediksi 24% berat bayi, setelah mengontrol variabel seperti berat ibu, ras, status merokok, hipertensi, dan riwayat infeksi saluran kemih. Berat bayi dapat di prediksi dengan model berikut, 2837,24 ditambah 4,24 berat ibu dikurang 475,81 gram jika ibu memiliki ras hitam, dikurang 354,9 gram jika ibu memiliki ras lainnya, dikurang 354,9 gram jika ibu merokok, dikurang 585,11 gram jika ibu mempunyai riwayat hipertensi dan dikurang 524,44 gram jika ibu memiliki gangguan infeksi saluran kemih.

DAFTAR PUSTAKA

- Hastono, S. P., 2006. *Analisis Data*. Jakarta: Fakultas Kesehatan Masyarakat Universitas Indonesia.
- Widarsa, K. T., Kurniasari, N. M. D. & Mulyawan, K. H., 2017. *Modul Praktikum Stata*. Denpasar: Program Studi Kesehatan Masyarakat Universitas Udayana.