

XALQARO NORDIK UNIVERSITETI

Iqtisodiyot va pedagogika fakulteti, Iqtisodiyot va biznesni boshqarish kafedrası

Fan o'qituvchisi: Sabirov Xasan Nusratovich

Mavzu: Chiziqsiz regressiyalar va modellashtirish masalalari

Reja:

1. Eng kichik kvadratlarni hisoblash
2. Moslikni o'lchash
3. Ma'lumotlarni kattalashtirish va o'zgartirish ta'sirlari
4. Qoldiqlarni tahlil olish
5. Polinomial modellar
6. Log-log modeli

4.1. Eng kichik kvadratlarni hisoblash

Biz II-bobda \hat{Y} hisoblangan masalasiga to'xtalgandik. Endi biz prognozning aniq xatoligini yoki bashorat qilishning aniqligini o'lchash sifatida prognozning standart xatosini topamiz. Ishchi katalogingizga o'ting, bo'lim uchun ishchi faylni oching va oziq-ovqat xarajatlari ma'lumotlarini joylashtiring.

version 11.1

capture log close

set more off

log using chap04_food, replace text

use food, clear

Keling, haftalik daromadi 2000 dollar bo'lgan uy xo'jaliklari uchun uy -ro'zg'or xarajatlarining taxminiy qiymatini topaylik.

4.1.1. Ma'lumotlarni tahrirlash

41-kuzatuvda $income=20$ ni kiritib, ma'lumotlarni tahrir qilamiz. Bosqichlar ushbu qo'llanmaning 2.5-bo'limida to'liq tushuntirilgan.

edit

set obs 41

replace income = 20 in 41

4.1.2. Regressiyani baholash va keyingi baholash natijalarini olish

Endi biz quyidagi buyruq orqali oziq-ovqat xarajatlarining regressiyasini hisoblaymiz.

quietly regress food_exp income

Hisoblashdan keyingi *predict* buyrug'idan foydalanib, hisoblangan qiymatlarni (**yhat**), eng kichik kvadratlar qoldiqlarini (**ehat**) topamiz. Bularni hisoblashni biz 2.4.1 - bo'limda to'liq to'xtalib o'tgandik.

predict yhat

predict ehat, residuals

Biz qo'shadigan yangi buyruq - *stdf*, bu prognozning standart xatosini hisoblab chiqadi. Biz prognoz qilmoqchi bo'lgan yangi kuzatuv topamiz:

$$y_0 = \beta_1 + \beta_2 x_0 + e_0$$

bu yerda e_0 - tasodifiy xato. $E(y_0) = \beta_1 + \beta_2 x_0$ va $E(e_0) = 0$ deb faraz qilamiz. Shuningdek, biz e_0 ning regressiya xatoliklari bilan bir xil dispersiyaga ega deb faraz qilamiz, $var(e_0) = \sigma^2$ va e_0 namuna ma'lumotlarining bir qismi bo'lgan tasodifiy xatolar bilan bog'liq emas, shuning uchun $cov(e_0, e_i) = 0$, $i = 1, 2, \dots, N$. y_0 ning eng kichik kvadratlar orqali hisoblanib mos keladigan regressiya chizig'i topiladi:

$$\hat{y}_0 = b_1 + b_2 x_0$$

Ya'ni, hisoblangan qiymat \hat{y}_0 eng kichik kvadratlar orqali hisoblangan chiziqning $x = x_0$ nuqtasida berilgan. Ushbu bashoratchi qanchalik yaxshi ishlashini baholash uchun biz eng kichik kvadratlar qoldig'iga o'xshash prognoz xatosini aniqlaymiz,

$$f = y_0 - \hat{y}_0 = (\beta_1 + \beta_2 x_0 + e_0) - (b_1 + b_2 x_0)$$

Ushbu prognoz xatosining taxminiy farqi

$$var(\widehat{f}) = \hat{\sigma}^2 \left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right]$$

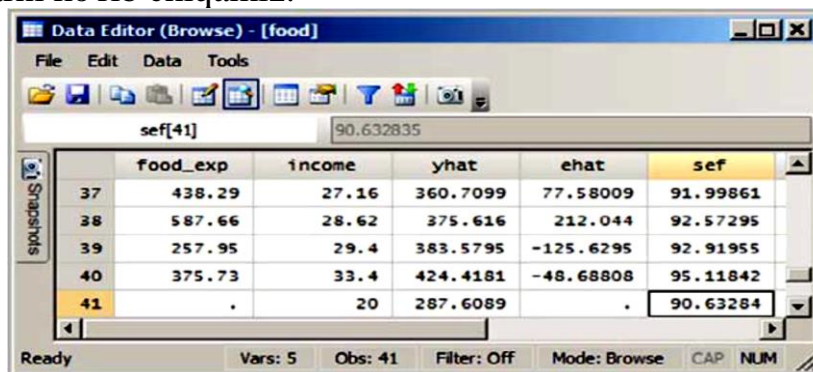
Bu taxminiy dispersiyaning kvadrat ildizi prognozning standart xatosidir

$$se(f) = \sqrt{var(\widehat{f})}$$

Quyidagi buyruq yordamida hisoblangan miqdor topiladi.

predict sef, stdf

Ma'lumotlarni ko'rib chiqamiz:



	Food_exp	income	yhat	ehat	sef
37	438.29	27.16	360.7099	77.58009	91.99861
38	587.66	28.62	375.616	212.044	92.57295
39	257.95	29.4	383.5795	-125.6295	92.91955
40	375.73	33.4	424.4181	-48.68808	95.11842
41	.	20	287.6089	.	90.63284

yhat tuzilgan modelning hisoblangan qiymati va **sef** buyrug'i orqali prognozning standart xatosi qiymatlari hisoblanadi. **Food_exp** bog'liq o'zgaruvchi va eng kichik kvadrat qoldiqlari 41 kuzatuv uchun qiymatlar topilmagan.

4.1.3. \hat{Y} hisoblangan qiymat oralig'ini yaratish

t_c kritik qiymatini t taqsimotidan $100(1 - \frac{\alpha}{2})$ -protsentil deb belgilab, $100(1 - \alpha)\%$ prognozlash oralig'ini quyidagicha olishimiz mumkin.

$$\hat{y}_0 \pm t_c se(f)$$

Biz t -kritik qiymatni **invttail(n,p)** yordamida yaratamiz.

scalar tc = invttail(38,.025)

di "t critical value 97.5 percentile = " tc

```
. di "t critical value 97.5 percentile = " tc
t critical value 97.5 percentile = 2.0243942
```

Hisoblangan qiymat - 2.0243942ga teng. So`ngra biz bu qiymatni **yhat** va **sef** bilan birgalikda bashorat qilish oralig'ining pastki va yuqori chegarasi bo'lgan yangi o'zgaruvchilarni yaratish uchun ishlatamiz. Yaratish buyrug'i **gen** ga qisqartirilib topish mumkin.

*gen lb = yhat - tc*sef*

*gen ub = yhat + tc*sef*

Biz bu o'zgaruvchilar qiymatlarini 41-kuzatuvda ro'yxat yordamida ko'rishimiz ham mumkin. Biz **list** buyrug'i orqali 41-kuzatuvda **income**, **lb**, **yhat** va **ub** o'zgaruvchilarini miqdorini ifodalaymiz.

list income lb yhat ub in 41

```
. list income lb yhat ub in 41
```

	income	lb	yhat	ub
41.	20	104.1323	287.6089	471.0854

Biz 41-kuzatuvni endi ishlatmaymiz, shuning uchun uni ma'lumotlar faylidan olib tashlaymiz.

drop in 41

drop buyrug'i bir nechta funktsiyalarga ega. Ma'lumotlar faylidan, biz yuqorida aytganimizdek, ba'zi kuzatuvlarni o'chirish uchun ishlatilishimiz mumkin va undan o'zgaruvchilarni tushirish uchun foydalanish mumkin. Sintaksisni ko'rish uchun **help drop** buyrug'iga to`xtalib qo`shimcha ma'lumotlarni olishingiz mumkin.

4.2. Moslikni o'lchash

Muvofiqlik o'lchovi R^2 kvadratlarning parchalanish yig'indisi bilan izohlanadi.

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum \hat{e}_i^2$$

Bu "kvadratlar yig'indisi":

- $\sum (y_i - \bar{y})^2 =$ kvadratlarning umumiy yig'indisi = SST: o'rtacha qiymat bo'yicha y ning umumiy o'zgarish o'lchovi.

- $\sum (\hat{y}_i - \bar{y})^2 =$ hisoblangan regressiya orqali kvadratlar yig'indisi = SSR: hisoblangan y qiymatining umumiy y ning o'rtacha qiymatidan ayirmasining kvadratining summasini izohlanadi. Shuningdek, "hisoblangan kvadratlar yig'indisi" deb ham ataladi.

- $\sum \hat{\epsilon}_i^2 =$ xatolikning kvadratlar yig'indisi = SSE: regressiyada izohlanmagan, y ning umumiy o'rtacha o'zgaruvchanligi. Kvadratlarning tushuntirilmagan yig'indisi, kvadrat qoldiqlar yig'indisi yoki kvadrat xatolar yig'indisi sifatida ham tanilgan.

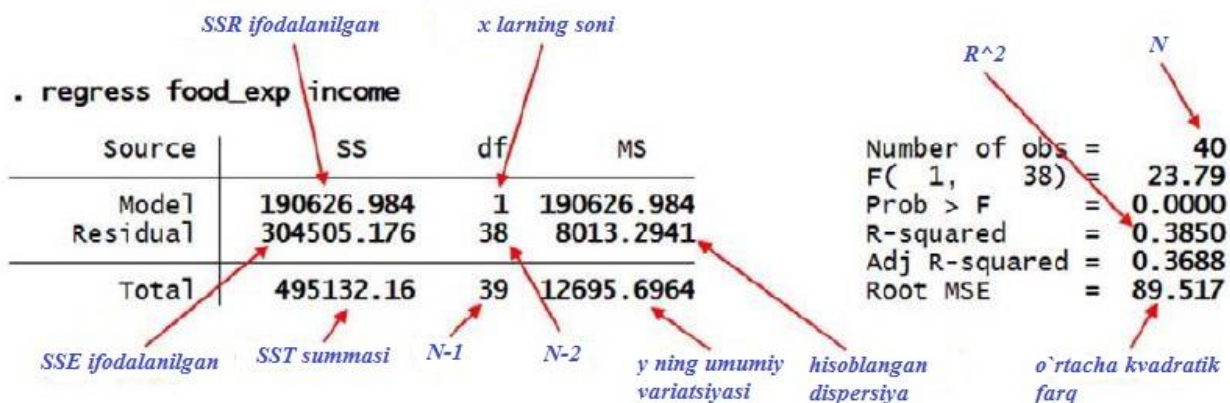
Ushbu qisqartmalar yordamida kvadratlarning parchalanishi yig'indisi hosil bo'ladi.

$$SST = SSR + SSE$$

U holda moslik o'lchovi R^2 bo'ladi.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

Regressiya baholanganda, bu miqdorlarning barchasi regressiya koeffitsientlarining yuqori qismida **Analysis of Variance Table** paydo bo'ladi. Quyidagi rasmdagi ma'lumotlar (**Model, Residual, Total**) kvadratlarning parchalanish yig'indisi bilan bir xil bo'linishni anglatadi. **SS** deb belgilangan ustun kvadratlar yig'indisiga ishora qiladi. **df** ustuni erkinlik darajalari uchun. "Erkinlik darajalari" modeli 1 ga teng, chunki modelda bitta tushuntirish o'zgaruvchisi mavjud. Erkinlikning qoldiq darajalari $N-2$ ni tashkil qiladi, bu kuzatishlar sonidan model parametrlari soni ayirmasidir. **MS** ustuni **Mean Square** hisoblanadi. Ushbu ustunda **SS** ustunining **df** ustuniga bo'lingan nisbati mavjud. Shunday qilib, **Residual MS** - bu taxminiy xato dispersiyamizdir.



4.2.1. Korrelyatsiya va determinatsiya R^2

Shuningdek, Principle of Econometrics adabiyotida qayd etilishicha, oddiy regressiya modelida R^2 - y va x o'zgaruvchilar orasidagi korrelyatsion bog'lanishning kvadrati. Bundan tashqari, umuman olganda, R^2 - y va uning hisoblangan qiymati \hat{y} o'rtasidagi kvadratik korrelyatsiyasi ham deb ataladi. Stata ning menyusidan foydalanib quyidagi ketma-ketlik asosida topiladi:

Statistics > Summaries, tables, and tests > Summary and descriptive statistics > Pairwise correlations

Pairwise correlations muloqot oynasida ochiladigan ro'yxat yordamida o'zgaruvchilar nomlarini kiriting.

Ok tugmasini bosing. So'ngra Stata buyrug'i quyidagicha namoyon bo'ladi:

pwcorr food_exp income yhat

Shuningdek, natijalar oynasida Pirson korrelatsiya koeffitsienti namoyon bo'ladi.

```
. pwcorr food_exp income yhat
```

	food_exp	income	yhat
food_exp	1.0000		
income	0.6205	1.0000	
yhat	0.6205	1.0000	1.0000

So`ngra *food_exp* va *income* o'rtasidagi kvadratik bog'liqlikni aniqlang.

$$r_{xy}^2 = 0,62^2 = 0,385 = R^2$$

4.3. Ma'lumotlarni kattalashtirish va o'zgartirish ta`sirlari

Regressiya modelidagi ma'lumotlar hech qanday haqiqiy ma'noni yo'qotmasdan kattalashtirish mumkin. O'zgaruvchilarning miqyosi o'ta katta yoki juda kichik sonlarsiz oqilona o'lchov birliklariga olib kelishi kerak. O'lchovli va o'zgartirilgan o'zgaruvchilar *generate* buyrug'i yordamida osongina yaratiladi.

Oziq - ovqat xarajatlari to'g'risidagi ma'lumotlar uchun o'zgaruvchan daromad 100 AQSh dollari birliklari bilan o'lchanadi, shuning uchun daromad = 20 xonadonning oylik daromadi 2000 AQSh dollarini tashkil etadi. Daromad o'zgaruvchisini miqdorini dollarga o'zgartirish uchun buyruqlar oynasiga quyidagini kiriting.

$$gen\ inc_dollar = income * 100$$

Ushbu yangi daromad o'zgaruvchisi bilan oziq - ovqat xarajatlari regressiyasini hisoblang. E'tibor bering, biz regress buyrug'ini reg buyrug`iga qisqartirdik va quyidagi natijalarga erishdik:

```
reg food_exp income
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
income	10.20964	2.093264	4.88	0.000	5.972052 14.44723
_cons	83.416	43.41016	1.92	0.062	-4.463279 171.2953

```
reg food_exp inc_dollar
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
inc_dollar	.1020964	.0209326	4.88	0.000	.0597205 .1444723
_cons	83.416	43.41016	1.92	0.062	-4.463279 171.2953

Regressiya natijalarida daromad koeffitsientining o'zgarishi va uning standart xatosi, shu bilan ishonch oralig'i qayd etiladi.

log close

4.3.1. Liner-log funktsiyasining shakli

Chiziqli regressiya modeli o'zgaruvchilarning oddiy o'zgarishi yordamida chiziqli bo'lmagan o'zgaruvchilar o'rtasidagi munosabatni ifodalashi mumkin. Oziq-ovqat xarajatlari misoli uchun chiziqli-log modelini quyidagicha ifodalash mumkin:

$$FOOD_EXP = \beta_1 + \beta_2 \ln(INCOME) + e$$

Ushbu modelni baholash uchun biz $\ln(\text{income})$ bo'lgan yangi o'zgaruvchini yaratamiz va keyin eng kichik kvadrat regressiyasini qo'llaymiz.

*log using chap04_linlog, replace text
use food, clear*

```

gen lnincome = ln(income)
reg food_exp lnincome
. reg food_exp lnincome

```

Source	SS	df	MS			
Model	176519.828	1	176519.828	Number of obs =	40	
Residual	318612.333	38	8384.53507	F(1, 38) =	21.05	
				Prob > F =	0.0000	
				R-squared =	0.3565	
				Adj R-squared =	0.3396	
				Root MSE =	91.567	
Total	495132.16	39	12695.6964			

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnincome	132.1659	28.80461	4.59	0.000	73.85397	190.4777
_cons	-97.18645	84.23744	-1.15	0.256	-267.7162	73.34333

4.3.2-rejada biz foydalanadigan **loglinear** model regressidan mos qiymatlar va qoldiqlarni olish uchun taxmin qilishdan keyingi buyruq **predict** dan foydalanamiz.

predict lyhat

predict lehat, resid

Chiziqli logli model uchun qiyalik $\beta_2/INCOME$. Buni hisoblash uchun Stata dasturidan qanday foydalanishimiz mumkin? Avval daromad uchun qiymatni tanlashimiz kerak. Namunaviy o'rtacha qiymat, boshqa aniq qiymat qiziqirmasa, yaxshi tanlov bo'ladi. Daromadlarning umumiy statistikasini hisoblaymiz.

summarize income

Stata hisoblangan bir necha natijalarni saqlaydi. Buni quyidagi buyruqni kiritish orqali ko'rishingiz mumkin:

return list

```
. summarize income
```

Variable	Obs	Mean	Std. Dev.	Min	Max
income	40	19.60475	6.847773	3.69	33.4

```
. return list
```

scalars:

```

r(N) = 40
r(sum_w) = 40
r(mean) = 19.60475
r(Var) = 46.89198967948718
r(sd) = 6.847772607168492
r(min) = 3.69
r(max) = 33.4
r(sum) = 784.1899999999999

```

Yuqoridagi natijalarni ko'ramiz, unda bir qancha skalarlar ifodalanilgan, shu jumladan o'rtacha qiymat ham. Uni skalyar sifatida saqlashda quyidagi buyruqni kiriting.

scalar xbar = r(mean)

Endi slop formulasini hisoblash uchun **lincom** buyruqdan foydalaning.

lincom lnincome/xbar

```
. lincom lincome/xbar
( 1) .051008*lincome = 0
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	6.741522	1.469267	4.59	0.000	3.767147	9.715897

O'rtacha daromad darajasida har haftalik uy xo'jaliklari daromadi 100 dollarga oshishi, oziq -ovqat xarajatlarini 6,74 dollarga ko'payishiga olib keladi. Hisoblangan chiziqli-log modelning parametri boshqa qiymatlarda ham hisoblanishi mumkin.

lincom lincome/10

lincom lincome/20

lincom lincome/30

```
. lincom lincome/10
```

```
( 1) .1*lincome = 0
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	13.21659	2.880461	4.59	0.000	7.385397	19.04777

```
. lincom lincome/20
```

```
( 1) .05*lincome = 0
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	6.608293	1.440231	4.59	0.000	3.692698	9.523887

```
. lincom lincome/30
```

```
( 1) .0333333*lincome = 0
```

food_exp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	4.405528	.9601537	4.59	0.000	2.461799	6.349258

4.3.2. Hisoblangan liner-log modelini grafigini chizish

Oldingi bo'limda hisoblangan qiymatlar va qoldiqlardan foydalanib, biz **twoway** grafiklardan foydalangan holda ikkita chizmani topamiz, birinchisi nuqtali chizma, ikkinchisi esa chiziqli chizma. Raqamlar chiroyli tarzda ifodalandi, chunki ma'lumotlar INCOME bo'yicha tartiblangan.

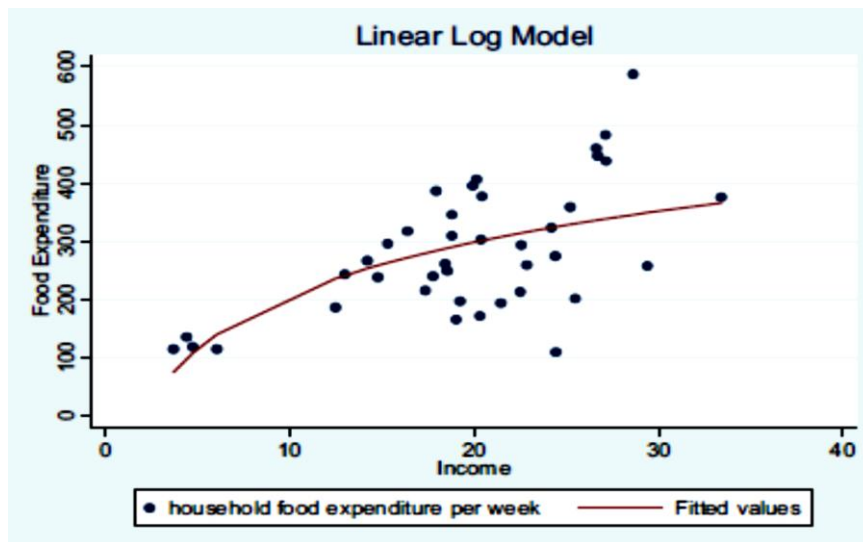
```
twoway (scatter food_exp income, sort) ///
```

```
(line lyhat income, sort lwidth(medthick)), ///
```

```
xtitle(Income) ytitle(Food Expenditure) ylabel(0(100)600) ///
```

```
title(Linear Log Model)
```

```
graph save linlog, replace
```



Bu munosabatlar to'g'ri chiziqli emas. Ma'lumotlar markaziga yaqin joyda o'rnatilgan chiziq juda tik emas. Koeffisient 6,74, oddiy chiziqli regressiyadan kichikroq. Buni aniqroq ko'rish uchun keling, uchta chizmanini ustiga qo'yamiz. Natijani topish va bashorat qilingan qiymatlar va qoldiqlarni olish uchun **quietly** variantni ishlatib, chiziqli munosabatlarni baholaymiz.

quietly reg food_exp income

predict yhat

predict ehat, resid

Mos keladigan chiziqni boshqa yangi buyruq bilan qo'shing, **lpattern(dash)**.

twoway (scatter food_exp income, sort) ///

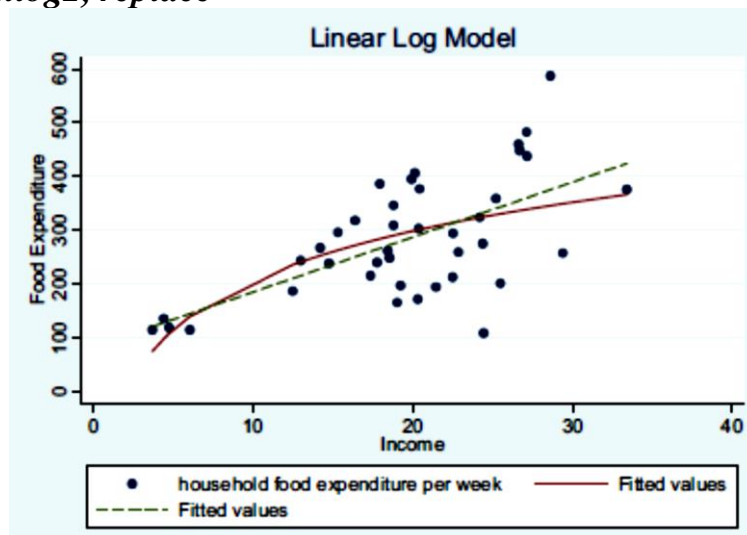
(line lyhat income, sort lwidth(medthick)) ///

(line yhat income, sort lpattern(dash) lwidth(medthick)), ///

xtitle(Income) ytitle(Food Expenditure) ylabel(0(100)600) ///

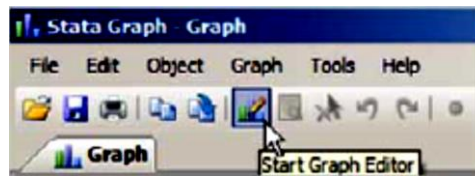
title(Linear Log Model)

graph save linlog2, replace

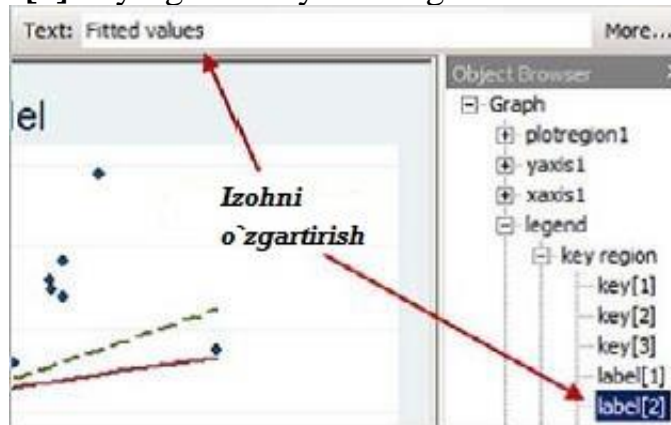


4.3.3. Grafiklarni tahrirlash

Graph Editor yordamida grafikning aspektlarini interaktiv tarzda o'zgartirish mumkin.

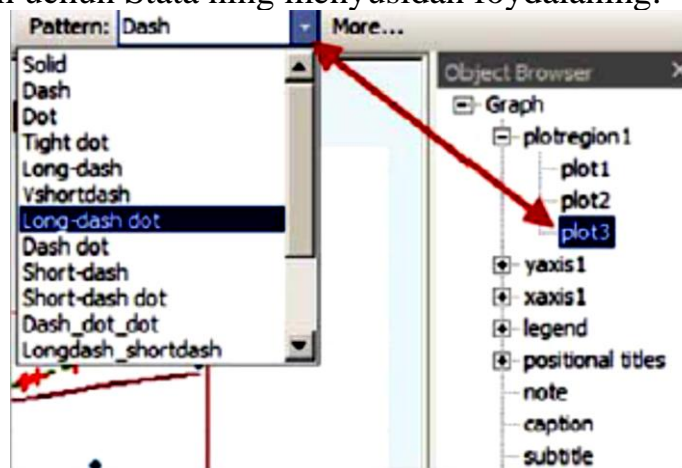


O'ng tomonda batafsilroq ro'yxatga ochiladigan elementlarni bildiruvchi "+" bilan grafik elementlarning ro'yxati mavjud. Elementlarni bosish ular bilan bog'langan ajratilgan maydonlarni ko'rsatadi. Chiziqli-log modeli uchun "Hisoblangan qiymatlar" izohini o'zgartirish uchun **label[2]** buyrugidan foydalaning.



label [2] ni tanlash chiziqni rangi, o'lchami va matni uchun yuqori tahrirlash imkoniyatlarini ko'rsatadi. Faqat kerakli matni kiriting va Enter tugmasini bosing holos.

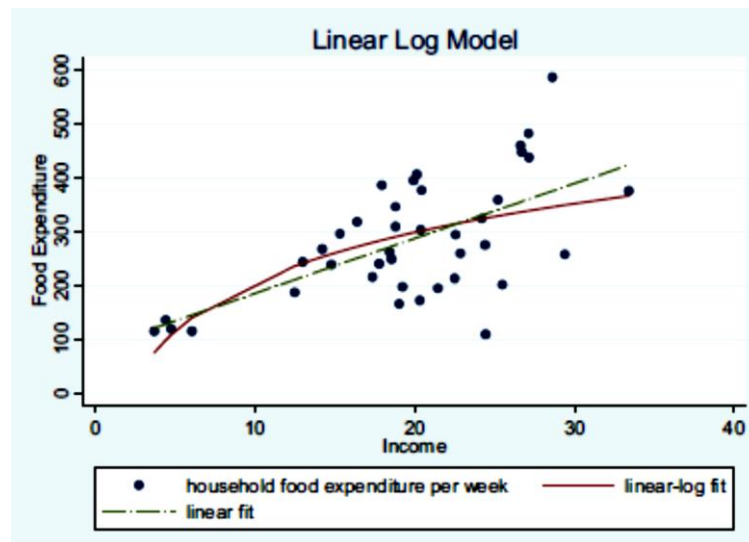
Plotregion va **plot3**ni tanlang. **Plot3** tanlanganda, chiziqli o'rnatilgan chiziq tanlanganligini ko'rsatadigan kesishishlarni ko'rsatadi. Chiziq shaklini to'g'ri chiziqdan boshqasiga o'zgartirish uchun Stata ning menyusidan foydalaning.



O'zgartirishlar kiritish uchun **Apply**-ni tanlang. Keyin o'zgartirishlaringizni **Save** orqali saqlang va **Stop** orqali plot muharririni to'xtating.



Bizning grafik natijada quyidagicha ifodalaniladi:



Graph Editor da grafikningizning deyarli barcha jihatlarini boshqarish mumkin. Siz kiritgan tahrirlarni **graph recorder** yordamida yozib olish va boshqa grafiklarda qayta qo'llash mumkin. Ushbu xususiyat haqida ko'proq ma'lumot olish uchun **help graph recorder** ga kirib o'rganishingiz mumkin.

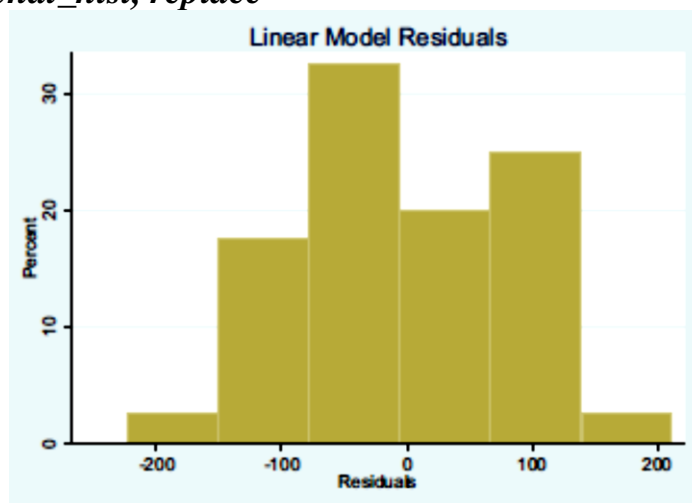
4.4. Qoldiqlarni tahlil olish

Tanlangan funksional shaklning adekvatligini aniqlashning asosiy elementlaridan biri bu model qoldiqlarini tahlil qilishdir. Yuqoridagi 4.1.2-rejada chiziqli model uchun eng kichik kvadratlar qoldiqlarini hisoblab chiqdik. Ular xotiradagi ma'lumotlar to'plamida saqlab qoyilgandi. Agar yo'q bo'lsa, oziq-ovqat xarajatlarining chiziqli modelini qayta hisoblang va biz **ehat** deb ataydigan qoldiqlarni hisoblang. Agar sizda kengaytirilgan natijalar bo'lsa, keyinroq foydalanish uchun kerak bo'ladigan o'zgaruvchilarni saqlashni unutmang.

Birinchi qadam sifatida eng kichik kvadratlar qoldiqlarining gistogrammasini tuzing. Biz har doim natija normal taqsimotni eslatib, qo'ng'iroq shaklida bo'lishiga umid qilamiz.

histogram ehat, percent title(Linear Model Residuals)

graph save olsehat_hist, replace



Ikkinchidan, biz qoldiqlarni batafsil umumlashtirishimiz mumkin

summarize ehat, detail

```
. summarize ehat, detail
```

Residuals				
	Percentiles	Smallest		
1%	-223.0255	-223.0255		
5%	-133.9407	-142.2519		
10%	-115.6628	-125.6295	Obs	40
25%	-52.94326	-119.058	Sum of wgt.	40
50%	-6.324473		Mean	4.77e-08
		Largest	Std. Dev.	88.3619
75%	68.72928	117.4039		
90%	112.8848	120.0951	Variance	7807.825
95%	121.0697	122.0443	Skewness	-.0973187
99%	212.044	212.044	Kurtosis	2.989034

4.4.1. Jarque-Bera testi

Qoldiqlarni normal taqsimlanganini batafsil tahlil qilish uchun Jarque-Bera testidan foydalaniladi, bunda uning tarkibiy qismlari bo'lgan Skewness va Kurtosis koeffitsientlarini o'z ichiga oladi. Y o'zgaruvchisi uchun namuna jarayonlarini quyidagicha aniqlang:

$$\tilde{\mu}_2 = \sum (Y_i - \bar{Y})^2 / N = \tilde{\sigma}^2$$

$$\tilde{\mu}_3 = \sum (Y_i - \bar{Y})^3 / N$$

$$\tilde{\mu}_4 = \sum (Y_i - \bar{Y})^4 / N$$

E'tibor bering, ushbu hisob-kitoblarni biz N ga bo'lamiz. Markaziy momentlarning ushbu namunaviy baholaridan foydalanib, biz skewness koeffitsienti (S) va kurtoz koeffitsienti (K) ning taxminiy qiymatini olishimiz mumkin.

$$\widehat{skewness} = S = \frac{\tilde{\mu}_3}{\tilde{\sigma}^3}$$

$$\widehat{kurtosis} = K = \frac{\tilde{\mu}_4}{\tilde{\sigma}^4}$$

Yuqorida ta'kidlab o'tilganidek, umumlashtirishdan foydalanilganda, bir qator elementlar saqlanadi. Quyidagi buyruqni kiriting:

```
return list
```

Ba`zi elementlar takrorlanadi:

```
. return list
```

```
scalars:
```

```
      r(N) = 40  
      r(sum_w) = 40  
      r(mean) = 4.76837158203e-08  
      r(var) = 7807.824984910715  
      r(sd) = 88.36189781184373  
      r(skewness) = -.0973187365457582  
      r(kurtosis) = 2.989033831200444
```

Qaytarilgan N, skewness va kurtoz qiymatlaridan foydalanib, Jarque-Bera statistikasini hisoblash uchun Stata (yoki kalkulyator) dan foydalanishimiz mumkin.

$$JB = \frac{N}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right)$$

Statistikani yaratish va uni ko'rsatish uchun buyruqlar:

```
scalar jb = (r(N)/6)* (r(skewness)^2) + ((r(kurtosis)-3)^2)/4 )
```

```
di "Jarque-Bera Statistic = " jb
```

Natijalar oynasida quyidagi ko'rinishda natija chiqadi:

```
. di "Jarque-Bera Statistic = " jb  
Jarque-Bera Statistic = .06334
```

4.4.2. χ^2 -kvadrat taqsimotining kritik qiymatlari

Ushbu taqsimot uchun kritik qiymat 2 daraja erkinlik bilan χ^2 -kvadrat taqsimotidan kelib chiqadi. Kritik qiymatlar **inv** tipidagi funksiyalar yordamida hisoblanadi. To'g'ri funktsiya turini topish uchun **help scalar**, so'ngra biz oldin bir necha marta qilganimizdek, belgilash tugmasini bosamiz. Yana bir tezkor alternativ – kirish usulidir:

```
db scalar
```

"db" "muloqot oynasi" degan ma'noni anglatadi. Stata buyrug'i:

```
scalar chic = invchi2tail(2,.05)
```

```
di "Chi-square(2) 95th percentile = " chic
```

Natijalar oynasida quyidagi hisoblangan qiymat:

```
. scalar chic = invchi2tail(2,.05)  
. di "Chi-square(2) 95th percentile = " chic  
Chi-square(2) 95th percentile = 5.9914645
```

Hisoblangan Jarque-Bera statistikasining qiymati = 0,06334 testning kritik qiymati = 5,99 dan ancha past, shuning uchun biz regressiya xatolari normal taqsimlangan degan gipotezani rad eta olmaymiz.

4.4.3. χ^2 -kvadrat taqsimotining p-qiymatlari

χ^2 -kvadrat kritik qiymatini olish o'rniga, biz Jarque-Bera test statistik qiymati uchun qiymatni olish orqali testni amalga oshirishimiz mumkin. To'g'ri funktsiya turini topish uchun **db skaler** dan foydalanamiz. Stata buyrug'I esa:

```
scalar pvalue = chi2tail(2,jb)
```

```
di "Jarque-Bera p-value = " pvalue
```

Olingan p-qiymati

```
. scalar pvalue = chi2tail(2,jb)  
. di "Jarque-Bera p-value = " pvalue  
Jarque-Bera p-value = .96882624
```

p-qiymati 0,9688 ga teng bu esa I-turdagi xatolik $\alpha = 0,05$ ehtimolidan katta bo'lganligi sababli, biz avval aytib o'tganimizdek, normallikni rad eta olmaymiz.

```
log close
```

4.5. Polinomial modellar

Polinom modeliga misol sifatida biz G'arbiy Avstraliyaning ba'zi fermalarida vaqt o'tishi bilan bug'doy hosili orasidagi bog'liqlikni ko'rib chiqamiz. Ma'lumotlar **wa_wheat.dta** deb nomlangan fayldan foydalanamiz Stata xotirasini tozalang va ushbu faylni ishga tushiring. O'zgaruvchilar tavsifi va umumiy statistikasi natijasini oling.

```
log using chap04_wheat, replace text
use wa_wheat, clear
describe
summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
northampton	48	1.168654	.4250324	.3024	2.3161
chapman	48	1.072385	.3328069	.4167	2.0244
mullewa	48	.9840625	.3352854	.3965	1.7992
greenough	48	1.15306	.3653873	.4369	2.2353
time	48	24.5	14	1	48

Xulosa statistik ma'lumotlar shuni ko'rsatadiki, 4 xil fermada t=48 kuzatuv mavjud. Shunday qilib, bizning terminologiyamiz *Principle of Econometrics, 4e* nashirdagi adabiyot bilan parallel bo'lishi uchun, Greenough fermasidagi ishlab chiqarishni o'rganish uchun *YIELD* deb nomlangan yangi o'zgaruvchini yaratamiz.

```
gen yield = greenough
```

```
label variable yield "wheat yield greenough shire"
```

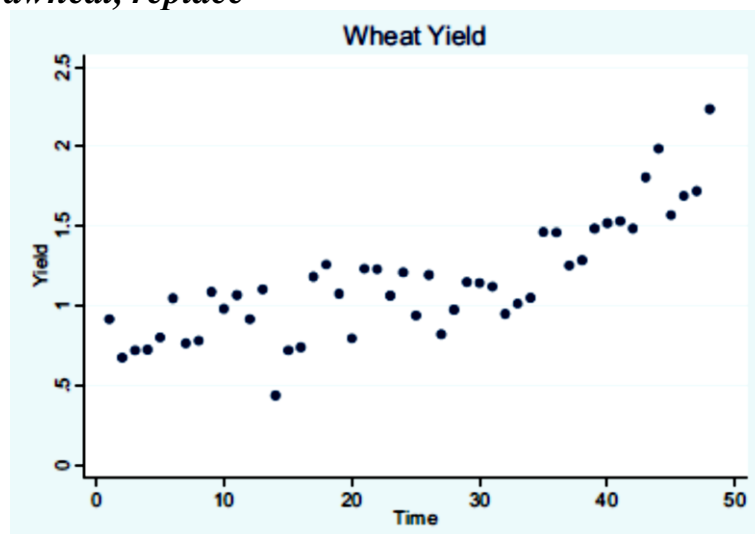
YIELD va *TIME* o'rtasidagi munosabatni ko'rsatadigan nuqtali grafikni tuzamiz.

```
twoway (scatter yield time, sort), ///
```

```
xtitle(Time) ylabel(0(.5)2.5) ytitle(Yield) ///
```

```
title(Wheat Yield)
```

```
graph save wawheat, replace
```



4.5.1. Chiziqli munosabatni baholash va tekshirish

Bug'doy hosildorligining o'sish sur'ati, ayniqsa, davr oxiriga kelib ortgan. *YIELD* va *TIME* o'rtasidagi chiziqli munosabat mos kelmaydi, chunki u o'zgaruvchan nuqtalarga mos kelmaydi. Ko'rishimiz uchun *YIELD* va *TIME* o'rtasidagi chiziqli regressiyani hisoblaymiz va hisoblangan qiymat va eng kichik kvadratlar qoldiqlarini hisoblab olamiz.

```
reg yield time
```

```
predict yhat
```

```
predict ehat, residuals
```

Baholash natijalari quyidagicha bo'ladi:

reg yield time

Source	SS	df	MS			
Model	4.07486005	1	4.07486005	Number of obs =	48	
Residual	2.20000947	46	.047826293	F(1, 46) =	85.20	
Total	6.27486952	47	.133507862	Prob > F =	0.0000	
				R-squared =	0.6494	
				Adj R-squared =	0.6418	
				Root MSE =	.21869	

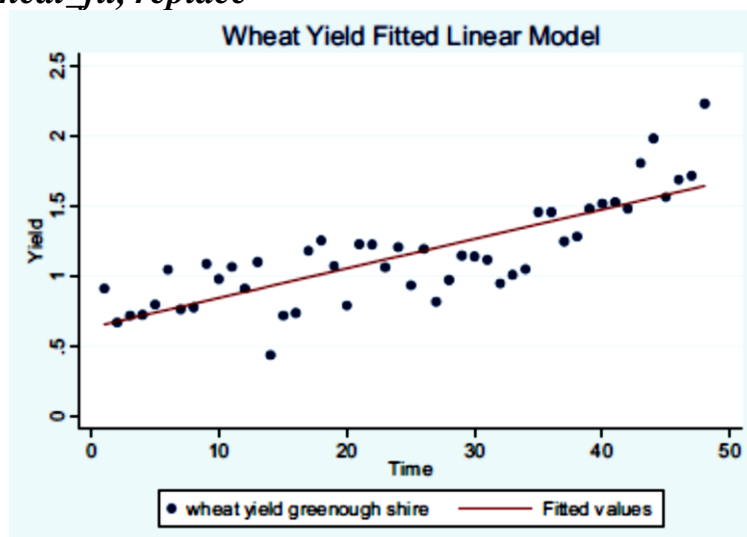
yield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
time	.0210319	.0022785	9.23	0.000	.0164455	.0256184
_cons	.6377778	.0641305	9.94	0.000	.5086898	.7668658

Hisoblangan eng kichik kvadratlar chizig'i va ma'lumotlarga mos chiziq bilan bir grafikda ifodalaymiz.

```

twoway (scatter yield time, sort) ///
(line yhat time, sort lwidth(medthick)) , ///
xtitle(Time) ytitle(Yield) ylabel(0(.5)2.5) ///
title(Wheat Yield Fitted Linear Model)
graph save wheat_fit, replace

```



Izoh: Uzoq qoldiqli o'zgaruvchilar grafik ostidagi matnni o'qib bo'lmisligiga olib kelishi mumkin. Grafik muharriri yordamida grafikni tahrirlang va yangi belgini kiriting.

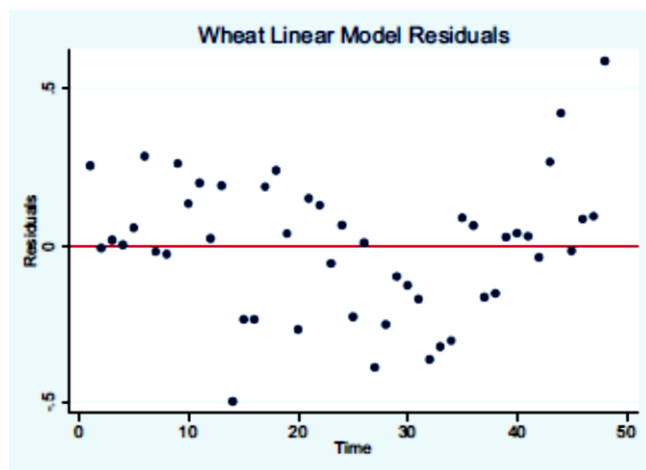
Amaliyot orqali siz vaqt davrining boshida va oxirida ijobiy qoldiqlar klasterini va o'rtadagi salbiy qoldiqlar klasterini aniqlay olasiz. Buni osonroq ko'rish uchun biz qoldiqlarni vaqtga nisbatan chizishimiz mumkin. Oddiy grafikdan foydalanib olish mumkin.

```

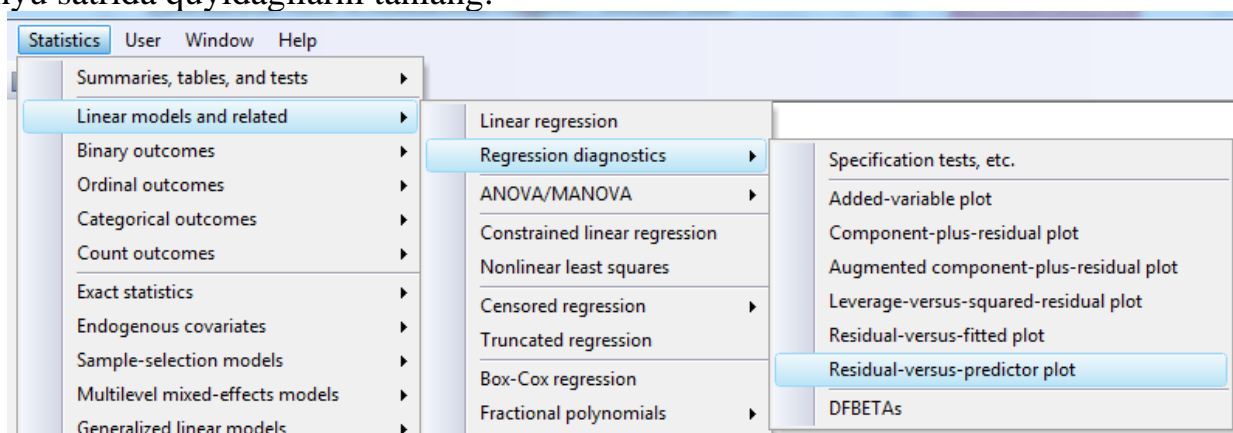
twoway (scatter ehat time, sort) , ///
xtitle(Time) ytitle(Residuals) yline(0) ///
title(Wheat Linear Model Residuals)
graph save wheat_ehat, replace

```

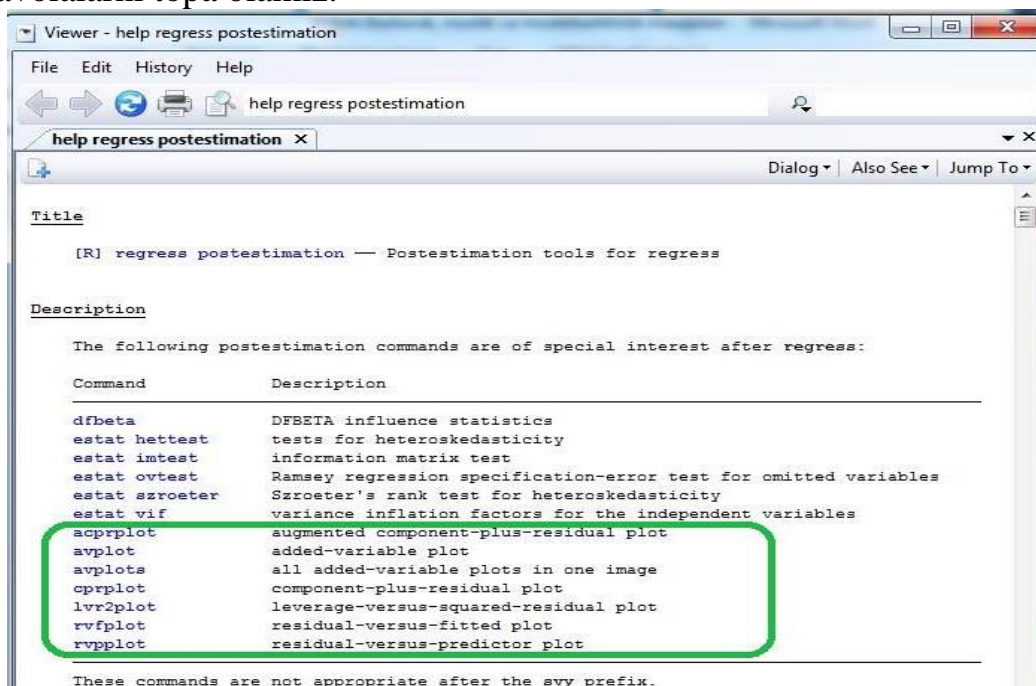
yline(0) varianti gorizontaal mos yozuvlar chizig'ini nolga tenglashtiradi.



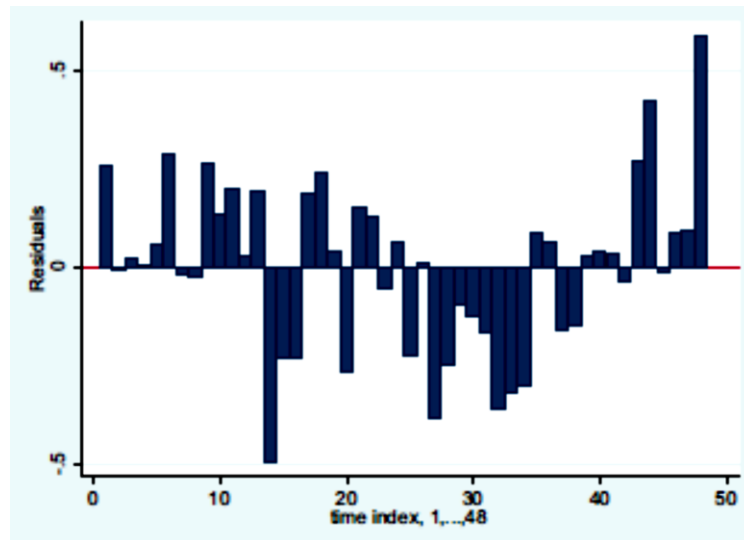
Shu bilan bir qatorda, Stata ba'zi o'rnatilgan qoldiq diagnostika grafiklariga ega. Stata menyu satrida quyidagilarni tanlang:



help regress postestimation yordamida biz dialoglar va mavjud grafiklarning tavsiflariga havolalarni topa olamiz.



Shuningdek, nuqtali grafikni o'rniga shtrixli grafikni olish uchun *recast (bar)* buyrug'idan foydalanish mumkin,
rvpplot time, recast(bar) yline(0)



Ushbu grafik vaqt davrining ikki oxiridagi ijobiy qoldiqlar to'plamini va markazdagi salbiy qoldiqlarning katta klasterini aniq ko'rsatadi. Qoldiqlardagi nuqtalar istalmagan va yaxshiroq funktsional shaklni topish zarurligini ko'rsatishi mumkin.

4.5.2. Kubik tenglamani baholash va tekshirish

$time0=time/100$ o'zgaruvchisini yarating. Bu o'lchov kub o'zgaruvchining juda katta bo'lishiga yo'l qo'ymaydi va taxminiy koeffitsientni nisbatan kattaroq qiladi. Kub tenglamani faktor o'zgaruvchilari yozuvidan foydalanib baholang va o'rnatilgan qiymatlar va qoldiqlarni oling, ularni *yhat3* va *ehat3* deb nomlang, chunki *yhat* va *ehat* oldingi regressiyada mavjud edi.

```
generate time0=time/100
list yield time0 in 1/5
```

	yield	time0
1.	.9141	.01
2.	.6721	.02
3.	.7191	.03
4.	.7258	.04
5.	.7998	.05

```
summarize time0
```

variable	Obs	Mean	Std. Dev.	Min	Max
time0	48	.245	.14	.01	.48

```
reg yield c.time0#c.time0#c.time0
```



```
. reg yield c.time0#c.time0#c.time0
```

Source	SS	df	MS	Number of obs = 48		
Model	4.71126527	1	4.71126527	F(1, 46) =	138.60	
Residual	1.56360425	46	.033991397	Prob > F =	0.0000	
Total	6.27486952	47	.133507862	R-squared =	0.7508	
				Adj R-squared =	0.7454	
				Root MSE =	.18437	

yield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
c.time0#c.time0#c.time0	9.681516	.8223546	11.77	0.000	8.026202	11.33683
_cons	.8741166	.0356307	24.53	0.000	.8023958	.9458374

Eng kichik kvadratlar qoldiqlari va taxmin qilingan qiymatlarni hisoblaymiz.

```
predict yhat3
```

```
predict ehat3, residuals
```

Faktor belgilarining afzalligi shundaki, chekkalar biz uchun qiyalikni to'g'ri hisoblab chiqadi.

```
margins, dydx(*) at(time=(0.15 0.30 0.45))
```

		Delta-method		z	P> z	[95% Conf. Interval]	
		dy/dx	Std. Err.				
time0	_at						
	1	.6535024	.0555089	11.77	0.000	.5447069	.7622979
	2	2.61401	.2220357	11.77	0.000	2.178827	3.049192
	3	5.881521	.4995804	11.77	0.000	4.902361	6.86068

Ma'lumotlarning grafigi va o'rnatilgan chiziq kubik funktsiyaning egriligi davr boshida va oxirida ma'lumotlarga yaxshiroq mos kelishini ko'rsatadi.

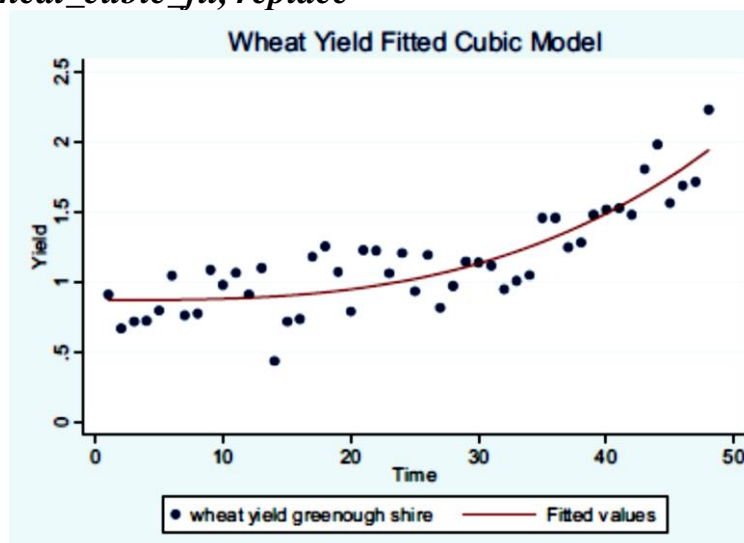
```
twoway (scatter yield time, sort) ///
```

```
(line yhat3 time, sort lwidth(medthick)), ///
```

```
xtitle(Time) ytitle(Yield) ylabel(0(.5)2.5) ///
```

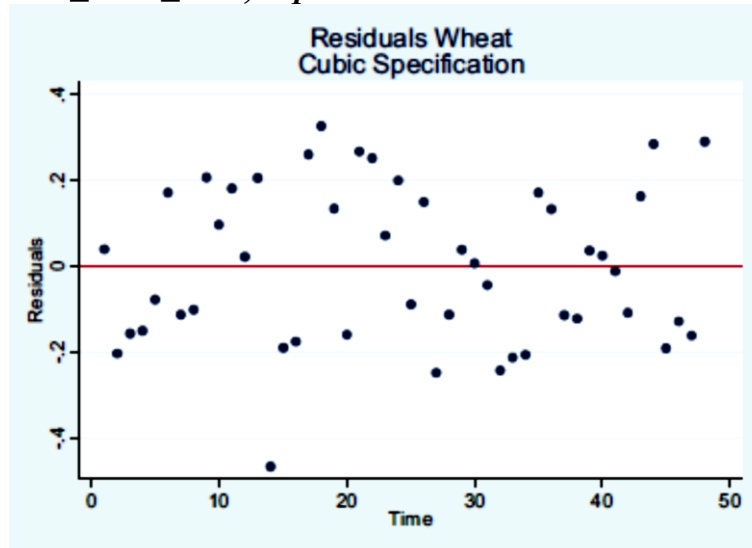
```
title(Wheat Yield Fitted Cubic Model)
```

```
graph save wheat_cubic_fit, replace
```



Qoldiq diagnostika grafikdan foydalanib, biz ijobiy va salbiy qoldiqlarning qisqarishini aniqlaymiz.

```
twoway (scatter ehat3 time, sort) , ///
xtitle(Time) ytitle(Residuals) yline(0) ///
title("Residuals Wheat" "Cubic Specification")
graph save wheat_cube_ehat, replace
```



4.5.3. Hosildorlikning log-liner o'sish modelini baholash

Oldingi bo'limning kubik modeliga muqobil ravishda log-chiziqli o'sish modelidan foydalanish ham mumkin.

```
gen lyield = ln(yield)
reg lyield time
. reg lyield time
```

Source	SS	df	MS			
Model	2.93313558	1	2.93313558	Number of obs =	48	
Residual	1.82466561	46	.039666644	F(1, 46) =	73.94	
Total	4.75780119	47	.101229813	Prob > F =	0.0000	
				R-squared =	0.6165	
				Adj R-squared =	0.6082	
				Root MSE =	.19916	

lyield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
time	.0178439	.0020751	8.60	0.000	.0136669	.0220208
_cons	-.3433665	.0584042	-5.88	0.000	-.460928	-.2258049

Biz yillik o'sish sur'atini 1,78% deb hisoblaymiz. log-chiziqli funktsional shakl haqida keyingi bo'limda yanada chuqurroq to`xtalib o`tamiz.

```
log close
```

4.6. Log-liner oylik maosh tenglamasini hisoblash

Yangi faylda ishlashni boshlang, bunda **cps4_small.dta** ma'lumotlar faylidan foydalaning. Shuningdek, o'zgaruvchilar tavsifini ifodalang va tasviriy statistikasi toping. **EDUC** qiymatlarini jadvalda ifodalang.

```
log using chap04_lwage, replace text
use cps4_small, clear
```

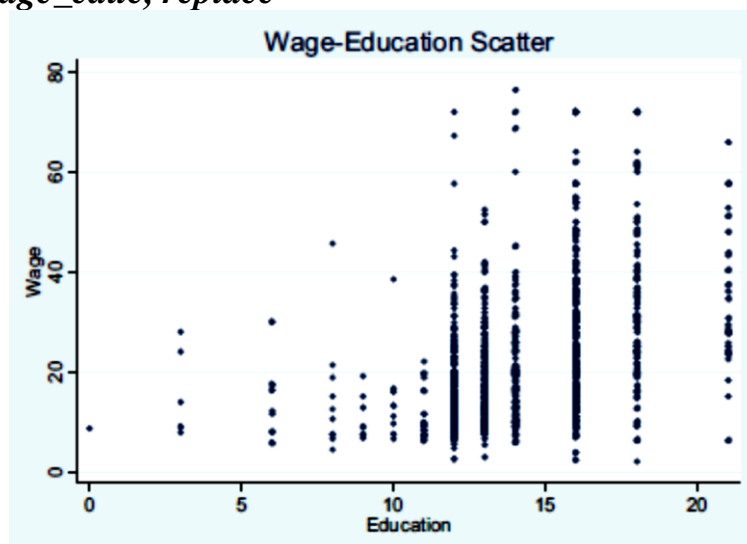
describe
summarize
tabulate educ

Biz foydalanadigan o'zgaruvchilar quyidagilar:

variable name	storage type	display format	value label	variable label
wage	double	%10.0g		earnings per hour
educ	byte	%8.0g		years of education

msize(small) varianti bilan kichik belgilardan foydalangan holda ish haqining ta'lim yillariga nisbatan nuqtali grafigini tuzamiz.

```
twoway (scatter wage educ, msize(small)), ///
xtitle(Education) ytitle(Wage) ///
title(Wage-Education Scatter)
graph save wage_educ, replace
```



Ushbu ma'lumotlardan foydalanib, biz ish haqi va ta'lim yillari o'rtasidagi bog'liqlikni o'rganamiz. Ish haqining ta'limga qarshi grafigi ta'limning yuqori darajalarida ba'zi ish haqining tarqoqlik tendentsiyasini ko'rsatadi. Bu turdagi maoshlar, ish haqi va daromad ma'lumotlariga xosdir.

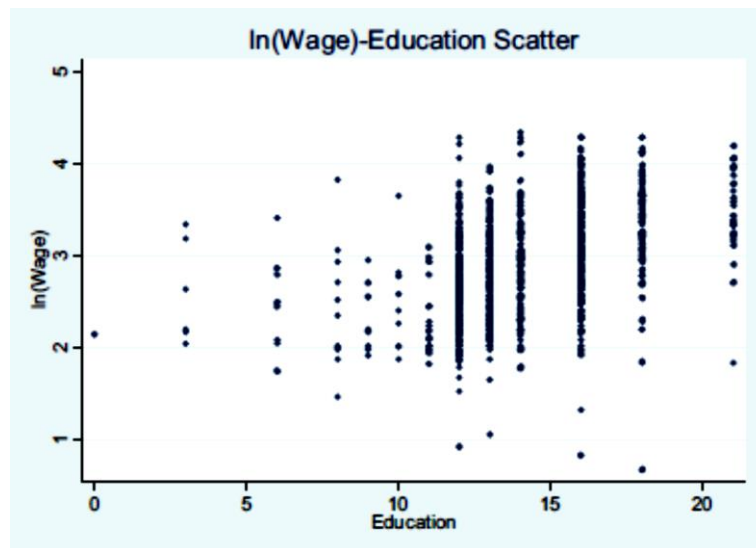
4.6.1. Log-liner modeli

Bu log-chiziqli model, ya'ni ta'sir etuvchi o'zgaruvchini o'zgartirish orqali amalga oshiriladi.

gen lwage = ln(wage)

Endi ta'limga qarshi *ln(wage)* nuqtali grafigni tayyorlaymiz. O'zgartirilgan ma'lumotlar sezilarli darajada kamroq tarqoqlikga ega.

```
twoway (scatter lwage educ, msize(small)), ///
xtitle(Education) ytitle(ln(Wage)) ///
title(ln(Wage)-Education Scatter)
graph save lwage_educ, replace
```



Ushbu misolda biz 12 yillik ta'limdagi ma'lumotga ega bo'lgan shaxslarning ish haqini taxminiy hisoblaymiz. Shunday qilib, 1001 kuzatuviga $EDUC = 12$ qiymatini qo'shish orqali ma'lumotlar faylini tahrirlaymiz.

edit

set obs 1001

replace educ=12 in 1001

Log-chiziqli regressiyani hisoblaymiz va hisoblangan qiymatlarni, eng kichik kvadratlar qoldiqlarini va prognozning standart xatosini hisoblaymiz.

reg lwage educ

. reg lwage educ

Source	SS	df	MS			
Model	60.015841	1	60.015841	Number of obs =	1000	
Residual	276.76489	998	.27731953	F(1, 998) =	216.41	
Total	336.780731	999	.337117849	Prob > F =	0.0000	
				R-squared =	0.1782	
				Adj R-squared =	0.1774	
				Root MSE =	.52661	

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.0904082	.0061456	14.71	0.000	.0783484	.1024681
_cons	1.609444	.0864229	18.62	0.000	1.439853	1.779036

Hisoblangan qiymatlarni, qoldiqlarni va kelajakda foydalanish uchun prognozning standart xatosini hisoblaymiz.

predict lwagehat

predict ehat, residuals

predict sef, stdf

Regressiya natijalariga ko'ra, biz har bir qo'shimcha ta'lim yili ish haqining taxminan 9% oshishiga olib kelishini taxmin qilishimiz mumkin.

Quyida biz hisoblashda xato atamasining taxminiy dispersiyasidan foydalanamiz, shuning uchun regressiya natijalari Stata xotirasida bo'lsa, biz undan foydalanamiz. Regressiyadan keyin Stata tomonidan saqlangan elementlar ro'yxatini ko'rish uchun *ereturn list* buyrug'idan foydalanamiz.

ereturn list

```
. ereturn list
```

```
scalars:
```

```
    e(N) = 1000  
    e(df_m) = 1  
    e(df_r) = 998  
    e(F) = 216.4140443756974  
    e(r2) = .1782044973688944  
    e(rmse) = .5266113647668089  
    e(mss) = 60.01584096379838  
    e(rss) = 276.764890442558
```

Qaytarilgan skalarlar orasida erkinlik darajalari (df_r) va kvadrat qoldiqlar yig'indisi (ssr) mavjud. Xato atamasining taxminiy dispersiyasi bo'lgan skalarini hisoblaymiz.

```
scalar sig2 = e(rss)/e(df_r)
```

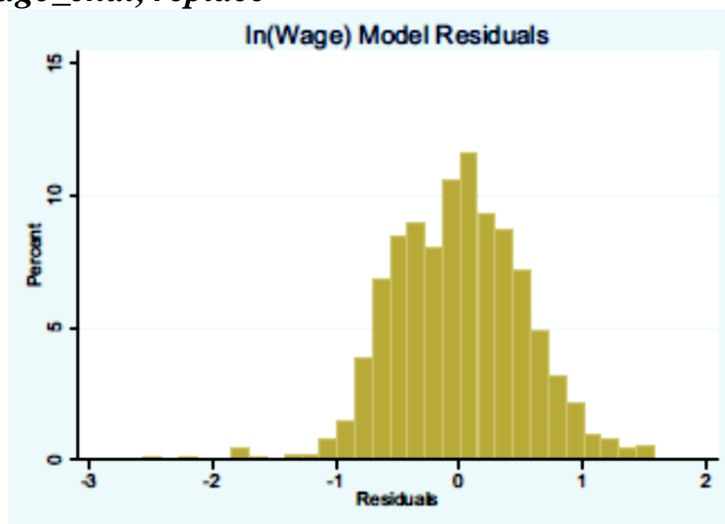
```
di "sigma-hat squared = " sig2
```

```
. di "sigma-hat squared = " sig2  
sigma-hat squared = .27731953
```

Endi oldingi kabi buyruqlardan foydalanib, gistogramma tuzamiz va Jarque-Bera testini normallik uchun tekshiramiz.

```
histogram ehat, percent title(log(Wage) Model Residuals)
```

```
graph save lwage_ehat, replace
```



```
summarize ehat, detail
```

```
scalar jb = (r(N)/6)* (r(skewness)^2) + ((r(kurtosis)-3)^2)/4
```

```
di "Jarque-Bera Statistic = " jb
```

```
scalar chic = invchi2tail(2,.05)
```

```
di "Chi-square(2) 95th percentile = " chic
```

```
scalar pvalue = chi2tail(2,jb)
```

```
di "Jarque-Bera p-value = " pvalue
```

```

. di "Jarque-Bera Statistic = " jb
Jarque-Bera Statistic = 27.528329

. scalar chic = invchi2tail(2,.05)

. di "Chi-square(2) 95th percentile = " chic
Chi-square(2) 95th percentile = 5.9914645

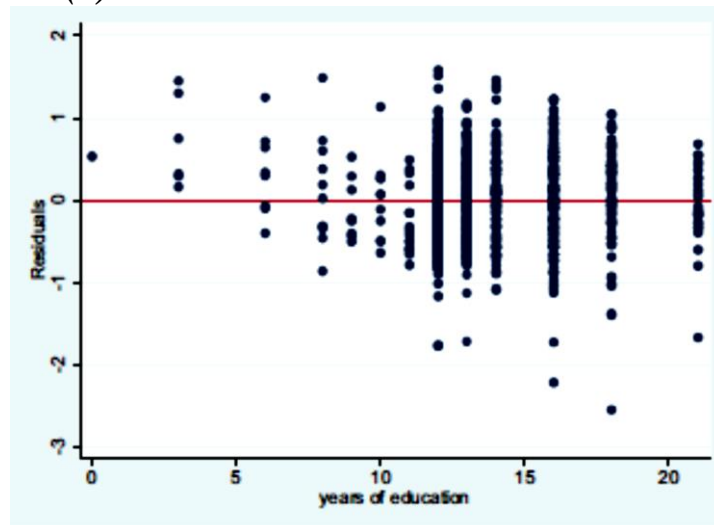
. scalar pvalue = chi2tail(2,jb)

. di "Jarque-Bera p-value = " pvalue
Jarque-Bera p-value = 1.053e-06

```

Natijada qoldiqlarning normalligi rad etiladi. So`ngra qoldiqlarning taqsimlanish yordamida grafikni olamiz.

rvpplot(educ, yline(0))



Qoldiqlar ham pasayish tendentsiyasini ifodalayapti, bu esa ta'lim oshgani sayin ish haqi oshishini ko'rsatadi.

4.6.2. Hisoblangan ish haqlarini hisoblash

Log-chiziqli modeldan hisoblangan qiymat $\ln(y) = \widehat{b_1} + b_2x$ dir. y bog'liq o'zgaruvchining bashoratini olish uchun biz "natural logorifimlangan" qiymatini olgan holda logorifimlangan qiymatini olamiz.

$$\hat{y}_n = \exp(\ln(\widehat{y})) = \exp(b_1 + b_2x)$$

Stata ning *exp()* funksiyasi eksponensial funktsiyani ifodalaydi. Hisoblangan *ln(wage)* dan foydalanish uchun logorifimlangan o'zgaruvchisini yaratamiz.

gen yhatn = exp(lwagehat)

Ko'pchilik namunalarda yaxshiroq hisoblangan qiymatga erishish uchun "to'g'rilangan" deb nomlanuvchi hisoblangan qiymatlar hisoblanadi.

$$\hat{y}_c = \widehat{E(y)} = \exp\left(b_1 + b_2x + \frac{\hat{\sigma}^2}{2}\right) = \hat{y}_n e^{\hat{\sigma}^2/2}$$

Eslatib o'tamiz, biz hisoblangan xato dispersiyasi $\hat{\sigma}^2$ ni *sig2* deb hisoblab chiqdik. Ushbu koeffitsientni "tuzatish koeffitsienti" hisoblangan qiymatlar yordamida hosil qilish mumkin.

di "correction factor = " exp(sig2/2)

```
. di "correction factor = " exp(sig2/2)
correction factor = 1.1487332
```

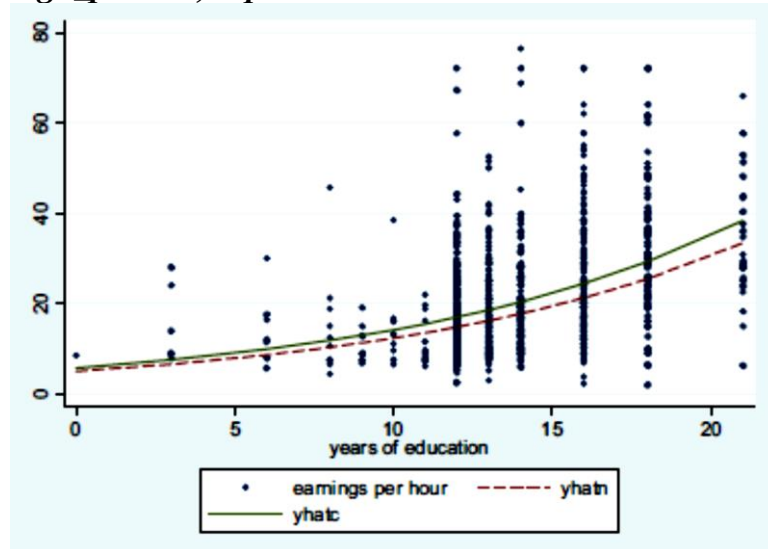
```
gen yhatc = yhatn*exp(sig2/2)
```

Tuzatish koeffitsienti 1,1487 ni tashkil qildi va shuning uchun "tuzatilgan" hisoblangan qiymatlar "tabiiy" qiymatlardan kattaroqdir. Bu har doim shunday bo'ladi, chunki har qanday $a > 0$ qiymati uchun tuzatish koeffitsienti $e^a > 1$ va taxminiy dispersiya $\hat{\sigma}^2$ har doim noldan katta.

4.6.3. Ish haqi grafiklarini tuzish

Ikki hisoblangan qiymatlarni taqqoslaydigan grafikda ko'rish yanada yaxshi xulosalar chiqarish mumkin.

```
twoway (scatter wage educ, sort msize(small)) ///
(line yhatn educ, sort ///
lwidth(medthick) lpattern(dash)) ///
(line yhatc educ, sort lwidth(medthick) lpattern(solid))
graph save lwage_predict, replace
```



Grafik *yhatc* har doim *yhatn* dan katta ekanligini ko'rsatadi. Tabiiy va hisoblangan qiymatlarni ko'rish uchun biz 1001 kuzatuvini qiymatini ko'rib o'tamiz.

```
list educ yhatn yhatc in 1001
```

	educ	yhatn	yhatc
1001.	12	14.7958	16.99643

Bu qiymatlar mantiqiyimi? Bizning namunamizda 12 yillik ta'lim ma'lumotga ega bo'lganlar uchun ish haqi bo'yicha umumiy statistikani hisoblash orqali biz taxminiy tekshirishimiz mumkin.

```
summarize wage if educ==12 in 1/1000
```

```
. summarize wage if educ==12 in 1/1000
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	328	15.99329	8.843706	2.5	72.13

Biz prognozlar 12 yillik ta'limga ega bo'lgan shaxslarning haqiqiy o'rtacha ish haqiga mos kelishini ko'rishimiz mumkin.

4.6.4. Determinatsiya koeffitsienti R^2

Determinatsiya koeffitsienti R^2 o'lchovi "eng yaxshi" hisoblangan qiymatni baholovchi **yhatc** va o'zgaruvchan **income** o'rtasidagi korrelyatsiya kvadratidir. Bu holda **yhatc** va **yhatn** faqat multiplikativ konstanta bilan farqlanadi, shuning uchun ularning **income** bilan bog'liqligi bir xil bo'ladi. Korrelyatsiyalarni hisoblash uchun quyidagi buyruqdan foydalaning:

correlate wage yhatn yhatc

di "r2g = " r(rho)^2

```
. correlate wage yhatn yhatc
(obs=1000)
```

	wage	yhatn	yhatc
wage	1.0000		
yhatn	0.4312	1.0000	
yhatc	0.4312	1.0000	1.0000

```
. di "r2g = " r(rho)^2
r2g = .18593072
```

4.6.5. Log-liner modeldagi bashorat oraliqlari

Log-chiziqli modeldagi bashorat oralig'i sifatida quyidagi formula tuzilgan.

$$\left[\exp(\widehat{\ln(y)} - t_c se(f)), \exp(\widehat{\ln(y)} + t_c se(f)) \right]$$

U "tabiiy" hisoblangan qiymatga va yuqoridagi 4.6.1-bo'limda tuzilgan prognozning standart xatosiga asoslanadi. Birinchidan, 998 daraja erkinlik bilan t-taqsimotidan 97,5 foizli ehtimollikni hisoblaymiz.

scalar tc = invttail(998,.025)

ln(income) ning pastki va yuqori prognoz oraliqlarini topamiz.

*gen lb_lwage = lwagehat - tc*sef*

*gen ub_lwage = lwagehat + tc*sef*

Ekspontensial funktsiyadan foydalanib logorifim qiymatlarini topamiz.

gen lb_wage = exp(lb_lwage)

gen ub_wage = exp(ub_lwage)

1001-kuzatish uchun bashorat oralig'ining qiymatlarini ko'rib o'tamiz.

list lb_wage ub_wage in 1001

```
. list lb_wage ub_wage in 1001
```

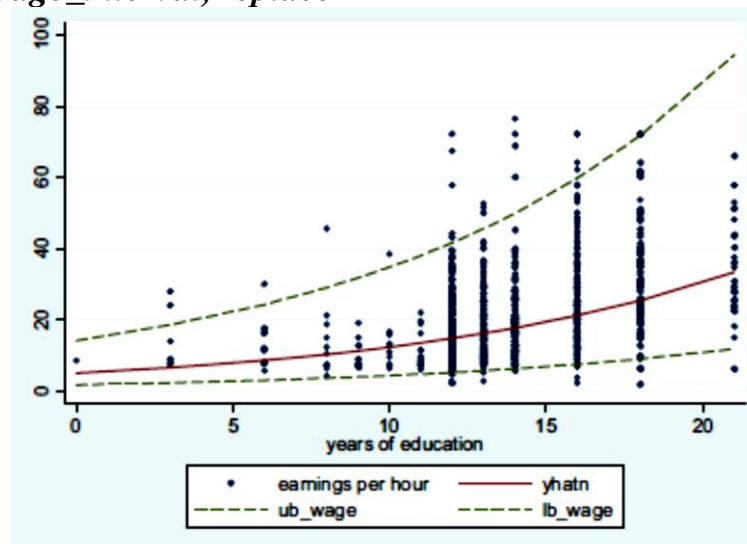
	lb_wage	ub_wage
1001.	5.260397	41.61581

Ish haqi ma'lumotlari, tabiiy hisoblangan qiymat va ta'limga qarshi intervalli hisoblangan qiymat bilan grafik yaratamiz.


```

twoway (scatter wage educ, sort msize(small)) ///
(line yhatn educ, sort lwidth(medthick) lpattern(solid)) ///
(line ub_wage educ, sort lcolor(forest_green) lwidth(medthick)) ///
lpattern(dash)) ///
(line lb_wage educ, sort lcolor(forest_green) lwidth(medthick)) ///
lpattern(dash))
graph save lwage_interval, replace

```



Natijalarga erishilgandan so`ng ishchi faylini yopish tavxiya qilinadi va buni quyidagi buyruq orqali amalga oshiring.

log close

4.7. Log-log modeli

Log-log funksiyasi $\ln(y) = \beta_1 + \beta_2 \ln(x)$ talab tenglamalari va ishlab chiqarish funksiyalarini hisoblash uchun keng qo'llaniladi. "Log-log" nomi logarifm tenglamaning har ikki tomonida paydo bo'lishidan kelib chiqadi. Ushbu modeldan foydalanish uchun y va x ning barcha qiymatlari ijobiy bo'lishi kerak. Ushbu egri chiziqlarning qiyaliklari har bir nuqtada o'zgaradi, lekin elastiklik doimiy va β_2 ga teng. Log-log funksiyasi haqida fikr yuritishning foydali usuli uning elastikligi $\frac{dy}{dx} = \beta_2 \left(\frac{y}{x}\right)$ ni yaqinroq tekshirishdan kelib chiqadi.

Buni $\beta_2 = \left(\frac{dy}{y}\right) / \left(\frac{dx}{x}\right)$ ga o'zgartiring. Shunday qilib, log-log funksiyasining elastikligi doimiy nisbiy o'zgarishlarni ko'rsatadi, chiziqli funktsiya esa doimiy mutlaq o'zgarishlarni ko'rsatadi. Yangi ishchi faylini oching va *newbroiler.dta* ma'lumotlaridan foydalaning.

log using chap04_loglog, replace text

use newbroiler, clear

describe

summarize

Bunda sifat o'zgaruvchilarni ham o'z ichiga olgan.

variable name	storage type	display format	value label	variable label
year	float	%9.0g		year
q	float	%9.0g		per capita consumption of boneless chicken, pounds
y	float	%9.0g		per capita real disposable income, 1996 = 100
p	float	%9.0g		real price (index) of fresh chicken

Miqdor va narxning logarifmini tuzing va log-log modelini baholang. Bu yerda logga ekvivalent bo'lgan ln log funksiyasidan foydalanamiz.

gen lq = ln(q)

gen lp = ln(p)

reg lq lp

. reg lq lp

Source	SS	df	MS			
Model	7.36410139	1	7.36410139	Number of obs =	52	
Residual	.69608941	50	.013921788	F(1, 50) =	528.96	
Total	8.0601908	51	.158042957	Prob > F =	0.0000	
				R-squared =	0.9136	
				Adj R-squared =	0.9119	
				Root MSE =	.11799	

lq	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lp	-1.121358	.0487564	-23.00	0.000	-1.219288	-1.023428
_cons	3.716944	.0223594	166.24	0.000	3.672034	3.761854

Hisoblangan elastiklik -1,12 ga teng hisoblanadi.

Hisoblangan qiymatlar va grafikni tuzamiz. Plot buyruqlari biroz uzun, shuning uchun biz ularni bu yerda ikkinchi qatorga ifodaladik, lekin uni do-faylda ifodalamaymiz.

predict lqhat

scalar sig2 = e(rss)/e(df_r)

*gen qhatc = exp(lqhat)*exp(sig2/2)*

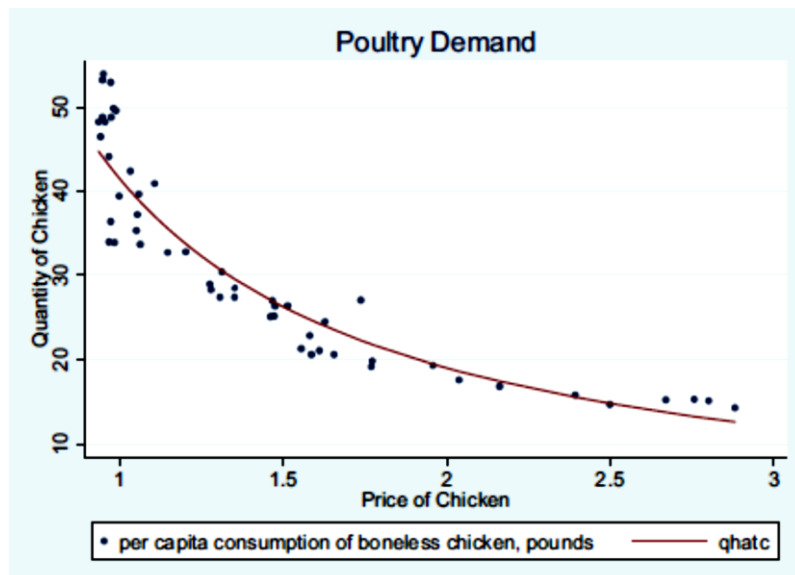
twoway (scatter q p, sort msize(small) lwidth(medthick) ///

lpattern(solid)) ///

(line qhatc p, sort lwidth(medthick)), ///

xtitle(Price of Chicken) ytitle(Quantity of Chicken) ///

title(Poultry Demand)



Determinatsiya koeffitsienti R^2 ni hisoblaymiz.

correlate q qhatc

di "r2g = " r(rho)^2

log close

```
. correlate q qhatc
(obs=52)
```

	q	qhatc
q	1.0000	
qhatc	0.9390	1.0000

```
. di "r2g = " r(rho)^2
r2g = .88177576
```

IV bob mavzularini mustahkamlash uchun savollar

1. Regressiyani baholash uchun qanday amallar ketma ketligi bajarilishi kerak?
2. \hat{Y} hisoblangan qiymat oralig'ini yaratish buyrug'i qanday?
3. drop buyrug'i funksiyalarini sanang.
4. Muvofiqlik o'lchovini ifodalang.
5. oddiy regressiya modelida R^2 nimani ifodalaydi?
6. O'lchovli va o'zgartirilgan o'zgaruvchilar qanday buyruq orqali yaratiladi?
7. Hisoblangan liner-log modelini grafigini chizish ketma ketligini yozing.
8. Grafik yaratilganidan so'ng uni qayta tahrirlash qanday amalga oshiriladi?
9. Jarque-Bera testi vazifasi nima?
10. Kub tenglamani faktor o'zgaruvchilari yozuvidan foydalanib baholash qanday ketma ketlik asosida amalga oshiriladi?
11. Hisoblangan Ish haqi grafklarini tuzishni amalda bajaring.
12. Log-log funksiyasi qay hollarda qo'llaniladi?