O’ZBEKISTON RESPUBLIKASI

OLIY VA O’RTA MAXSUS TA’LIM VAZIRLIGI

TOSHKENT DAVLAT IQTISODIYOT

UNIVERSITETI

R.H. ALIMOV, D.M. RASULEV, X.N. SOBIROV

EKONOMETRIK MODELLASHTIRISHDA STATA DASTURIDAN FOYDALANISH BO’YICHA AMALIY QO’LLANMA

O’ZBEKISTON RESPUBLIKASI
OLIY VA O’RTA MAXSUS TA’LIM VAZIRLIGI

TOSHKENT DAVLAT IQTISODIYOT UNIVERSITETI

R.H. ALIMOV, D.M. RASULEV, X.N. SOBIROV

EKONOMETRIK MODELLASHTIRISHDA STATA DASTURIDAN FOYDALANISH BO’YICHA AMALIY QO’LLANMA

Toshkent - 2019

Alimov R.H., Rasulev D.M., Sobirov X.N. Ekonometrik modellashtirishda Stata dasturidan foydalanish bo'yicha amaliy qo'llanma. - T.: TDIU, 2019. - 92 bet.

Mazkur amaliy qo'llanmada ekonometrik modellashtirish jarayonida Stata dasturining imkoniyatlari yoritib berilgan. Qo'llanmada Stata dasturiga ma'lumotlarni kiritish, ekonometrik modellarni tuzish, olingan natijalarni turli testlar yordamida tekshirish hamda prognozlash masalalari keltirilgan.

Qo'llanma bakalavr talabalari uchun mo'ljallangan bo'lib, u ekonometrika sohasidagi fanlarni o'rganishda tavsiya etiladi.

Bundan tashqari mazkur o'quv qo'llanmadan professor - o'qituvchilar hamda tadqiqotchilar ham foydalanishlari mumkin.

Taqrizchilar:

S.O.Xomidov − PhD, TDIU, "Ekonometrika" kafedrasi katta o`qituvchisi;

D.K.Abdullayeva − i.f.n., ,G.V.Plexanov nomidagi Rossiya Iqtisodiyot Universiteti "Raqamli iqtisodiyot va matematik fanlar" kafedrasi muduri.

Toshkent davlat iqtisodiyot universiteti kengashining 2019 yil 30 maydagi №11 sonli qaroriga asosan chop etishga tavsiya etildi.

Foydalanuvchiga!

Qadirdon kitobxonlar,

Avvalambor, Ekonometrik modellashtirishda Stata dasturidan foydalanish bo'yicha amaliy qo'llanmadan foydalanayotganingiz uchun sizga tashakur bildiramiz. Ushbu amaliy qo`llanma ta`lim ehtiyojidan kelib chiqqan holda yaratildi. Qo`llanma ekonometrika va stata dasturi bilan bog'liq amaliy muammolarni hal qilish va ma'lumotlar to'plamlari bilan ishonchli ishlashga imkon beruvchi amaliy va qulay ko`nikmani hosil qilishni ta'minlaydi.

Umid qilamizki, ushbu qo`llanmadan foydalangan holda ekonometrika haqida tushunchaga ega bo`lasiz va stata dasturi imkoniyatlaridan foydalanishni o'rganasiz. Bu darslikning bir turi bo'lib, Stat Data Release 12 dasturidan foydalangan holda qo`llanmadagi misollarni qanday bajarish kerakligini buyruqlar va menyular asosiya ko'rsatib beradi. Ushbu kitob ekonometrika fanini o`rganayotgan talabalar, shuningdek, o'qituvchilar va Stata dasturidan ekonometrik va statistik tahlil uchun foydalanmoqchi bo'lganlar uchun foydalidir. Ushbu qo`llanma 3ta bobdan iborat. Birinchi bob Stata dasturi haqida qisqacha ma`lumot deb nomlanadi va bu bobda STATA dasturinini ishga tushirish, o`chirish, ma`lumotlarni yuklash, tasvirlash va natijalarni saqlash haqida tushunchaga ega bo`lasiz. Ikkinchi bob Stata dasturida regression model tuzish deb nomlangan va bu qismda Klasik chiziqli modelni yaratish, tahliliy statistik ma`lumotlarini olish, ma`lumotlar asosida diagrammalar va grafiklar tuzish, hisoblangan qiymat va qoldiqlarni aniqlash, elastiklik qiymatini hisoblashni, variatsiya va kovarvatsiyani hisoblashni, prognoz qiymatlarni topishni, kvadratik va logarifmik modellarni tuzishni va sifat o`zgaruvchilardan foydalanish imkoniyatlari tushuntirib o`tilgan. Uchinchi bob esa Intervalli baholash va gipoteza testlari deb nomlandi va t-taqsimotning kritik qiymatlarini hisoblash, gipoteza testlarini va p-qiymatlarni hisoblash mavzulari ketma-ket yoritilgan. Stata dasturida alohida o`ringa ega bo`lgan ko`p omilli modellar, Heterokidastiklik, vaqtli qatorlar, autoregressiya va ARCH modellari haqida amaliy qo`llanmaning II-qismida yoritiladi.

Stata bugungi kunda juda ko'p turli xil ilmiy fanlarda qo'llaniladigan juda

kuchli dastur hisoblanadi. Ushbu Veb-sayt: http://www.stata.com orqali juda ko'p ma`lumotlarni topishingiz mumkin. UCLA-saytda vizual manbalardan biri: http://www.ats.ucla.edu/stat/stata/ hisoblanadi. Ushbu veb-saytdan foydalanishingizni tavsiya qilamiz. Ushbu Stata kompyuter dasturidan tashqari, shunga o'xshash EViews, Excel, Gretl va Shazam dasturiy ta'minot paketlari mavjud.

Biz ushbu amaliy qo`llanmaga izohlarni va takomillashtirish bo'yicha takliflarni mamnunyat bilan qabul qilamiz. Savollarimizga javob berish va o`quv qo`llanmani takomillashtirishga beradigan takliflaringizni quyudagi elektron pochta manziliga yozishingizni so`raymiz. Navbatdagi izlanishlarimizda Stata dasturiy paketining ekonometrik modellashtirishdagi qo`shimcha imkoniyatlarini yoritish, shunga o`xshash dasturiy ta`minot paketlarining qiyosiy tahlili rejalashtirilgan va sizlar bergan takliflaringiz orqali takomillashtiriladi.

khasansabirov19@gmail.com

MUNDARIJA

KIRISH

Bugungi kunda ekonometrik usullar iqtisodiy tahlil va prognozlashning eng muhim instrumentiga aylanib bormoqda. Ekonometrik tahlilda statistik axborotlarni qayta ishlash va tahlil qilishda bir qator universal dasturiy mahsulotlar mavjud. Ko'rib chiqilayotgan masalalar ko'lamiga qarab, ular ekonometrik usullarni o'rganishda nafaqat talabalarga, balki statistik ma'lumotlardan foydalanib, iqtisodiy tahlil va prognozlash masalalarini hal qilishda ilmiy izlanuvchilar hamda iqtisodchilarga ham foydali bo'lishi mumkin.

Mazkur qo'llanmada barcha asosiy ekonometrik hioblar Stata amaliy paketida amalga oshirish ko`zda tutilga. Stata dasturini tanlash undan foydalanishning qulayligi va ekonometrik modellashtirish amaliyotida keng qo'llanishi bilan asoslanadi.

Stata dasturi ekonometrik modellashtirishning barcha bosqichlarini o`zida mujassamlashtirilgan. Unda ma`lumotlarni kiritish, ular asosida tavsifiy statistikalarni o`tkazish, korrelatsion tahlil, juft va ko`plikdagi regression tahlillarni o`tkazish, fazoviy va vaqtli qatorlar bo`yicha turli xil grafiklarni olish, olingan natijalarni turli testlar yordamida tekshirish imkoniyati mavjud.

Foydalanuvchi dasturda keltirilgan qulay menyu yordamida turli buyruqlar bilan ishlash, natijalarni taqqoslash uchun oynalarni yonma-yon qo`yish, yangi vaqtli qatorlarni tuzish, olingan natijalarni ishchi faylda obyekt sifatida saqlash va zarur vaqtda ulardan foydalanishi mumkin.

I-BOB. STATA DASTURI HAQIDA QISQACHA MA`LUMOT

1.1 STATA dasturini ishga tushirish va o`chirish

Stata dasturini bir necha usul bilan ishga tushurish mumkin. Birinchisi, kompyuteringiz ish stolida (1.1-rasm) ikki marta bosishingiz mumkin bo'lgan yorliq bo'lishi mumkin va Stata / SE Release 12 uchun quyidagicha ko'rinishda bo`ladi:

1.1-rasm. Stata dasturining yorlig'i

Stata dasturining oldingi versiyalariga o'xshash ko'rinishdagi belgilari mavjud, lekin, bu albatta, boshqa raqamga ya`ni "StataSE-12" qo`shimcha funksiyalariga ega. Shu bilan bir qatorda, Windows menusidan foydalanib, **Start > All Programs > Stata 12** ni bosishingiz mumkin.

Ikkinchi usul esa *** .dta** kengaytmasi bilan Stata ma'lumotli faylini topish va uni ikki marta bosib Stata dasturini ishga tushurish mumkin.

Dasturning ko`rinishi.

Stata dasturi ishga tushirilgandan so'ng, oynalarni o'z ichiga olgan Stata dasturinig asosiy oynasi paydo bo'ladi (1.2-rasm). Bu asosiy oynada quyidagicha oynalar mavjud:

Command (buyruq**) -** bunda Stata buyruqlari yoziladi.

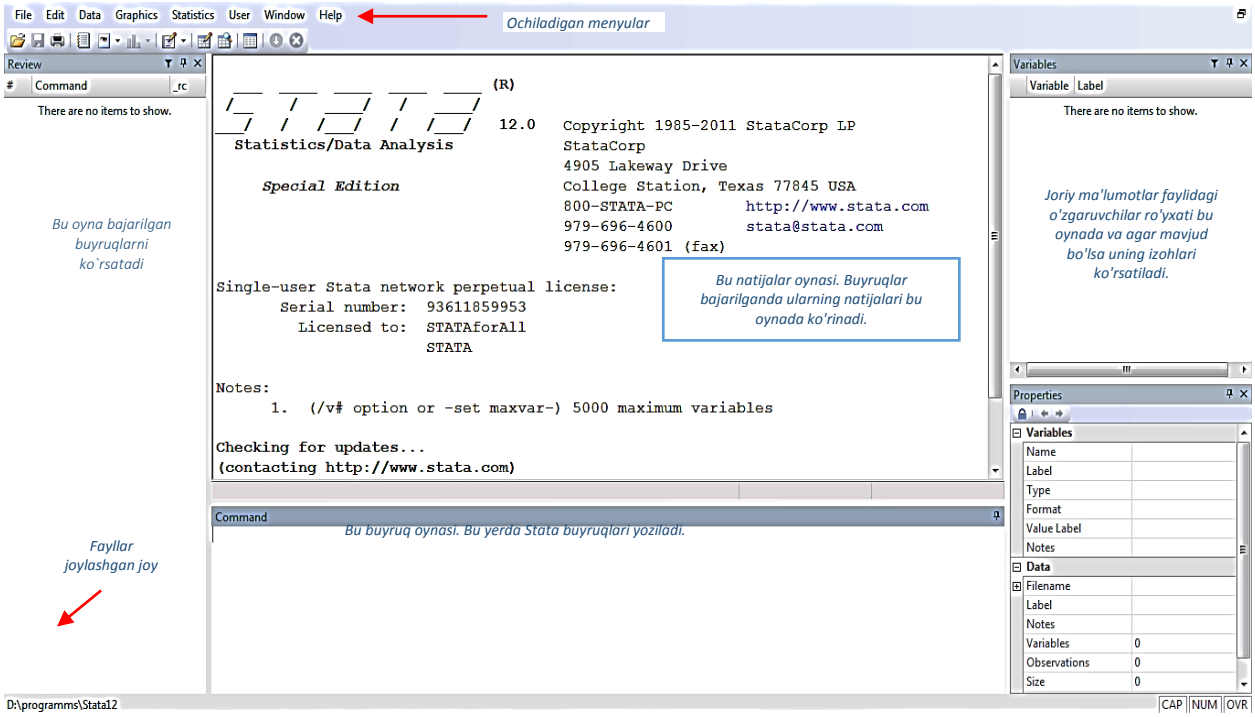
Results (natijalar**) -** buyruqlar natijalari va xato buyruqlar bu oynada ko'rinadi.

Review (ko'rib chiqish**) -** yaqinda bajarilgan buyruqlar ro'yxati oynasi.

Variables (o'zgaruvchilar**)** - ma'lumotlar va teglardagi o'zgaruvchining nomlari bu oynada ko`rinadi (yaratilgan bo'lsa).

Yuqorida "Stata"da ochiladigan menyulardir. Biz ulardan ko'pchiligini

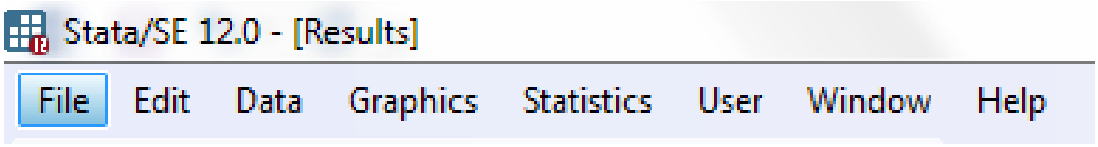
o'rganib chiqamiz. Chap pastki burchakda “Stata”da grafikalar, ma'lumotlar fayllari va hokazolarni saqlaydigan ishchi katalogning mavjud joyidir.



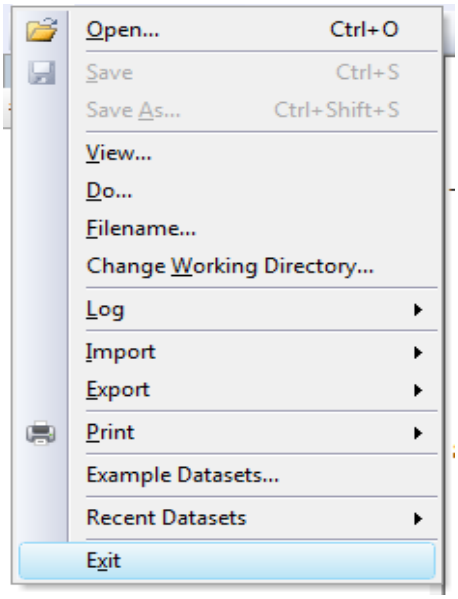
1.2-rasm. Stata dasturi oynasi

STATA - dasturini o'chirish.

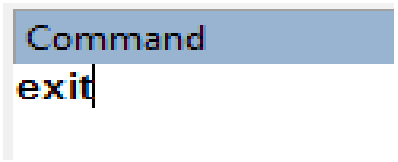
Stata dasturini to`xtatish uchun faylni bosib ya`ni:



va File menu sidagi Exit ni tanlab bosish kifoya.



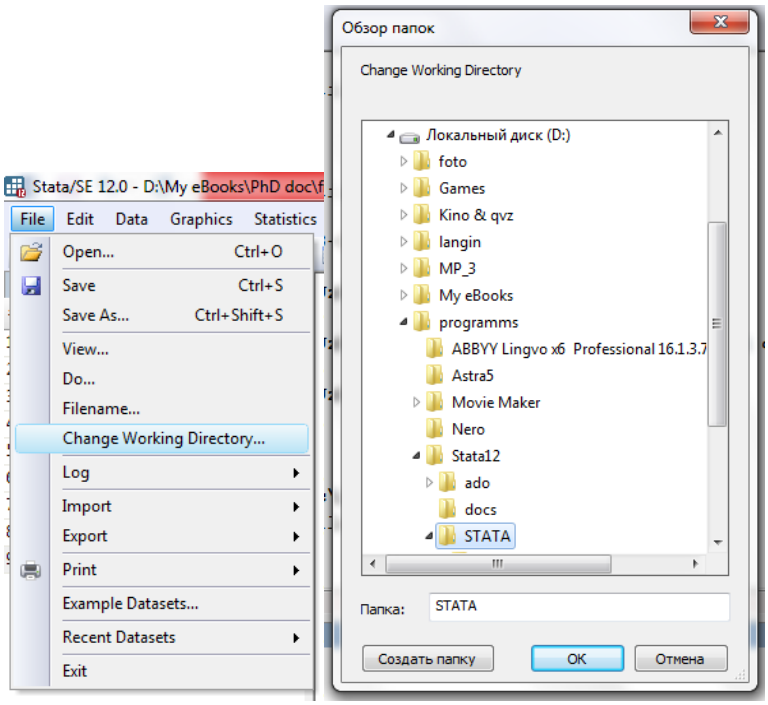
Biz File > Exit kabi ketma - ket bosish buyruqlarini bildiramiz. Shu bilan bir qatorda, shunchaki buyruqlar oynasiga “exit” so’zini yozib, Enter klavishasini bosish ham kifoya qiladi:



Ishchi fayl.

Ma'lumotni qulay faylga ko'chirishingiz kerak. Buni qanday amalga oshirilishi kompyuter tizimiga bog'liq bo'ladi. Ushbu Windows - bazasidagi barcha ma'lumot va natijalar fayllari uchun biz **d:\programms\Stata12\STATA** pastki katalogidan foydalanamiz. Biz buni qulay bo`lishi uchun qilyapmiz va agar siz laboratoriya sharoitida bo'lsangiz, bu juda yaxshi tanlov emas.

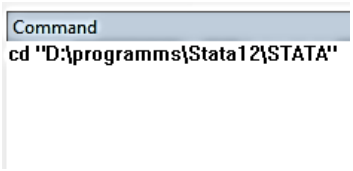
Agar siz laboratoriya xonalaridagi kompyuterda ishlayotgan bo'lsangiz, siz **"fleshka"** yoki **"Hard"** kabi saqlash qurilmasiga ega bo'lishingiz kerak. Ular Stata dasturidagi ma'lumotlar faylli va dastur fayllarini saqlash uchun yetarlicha katta bo`lishi kerak. Ishchi faylda ishlash uchun **File (Fayl) > Working Directory of Change**ni oching. Natijada fayllar oynasida siz tanlagan joyga o'ting va **OK** tugmasini bosing. Ushbu ketma-ketlik 3-rasmda izohlangan.



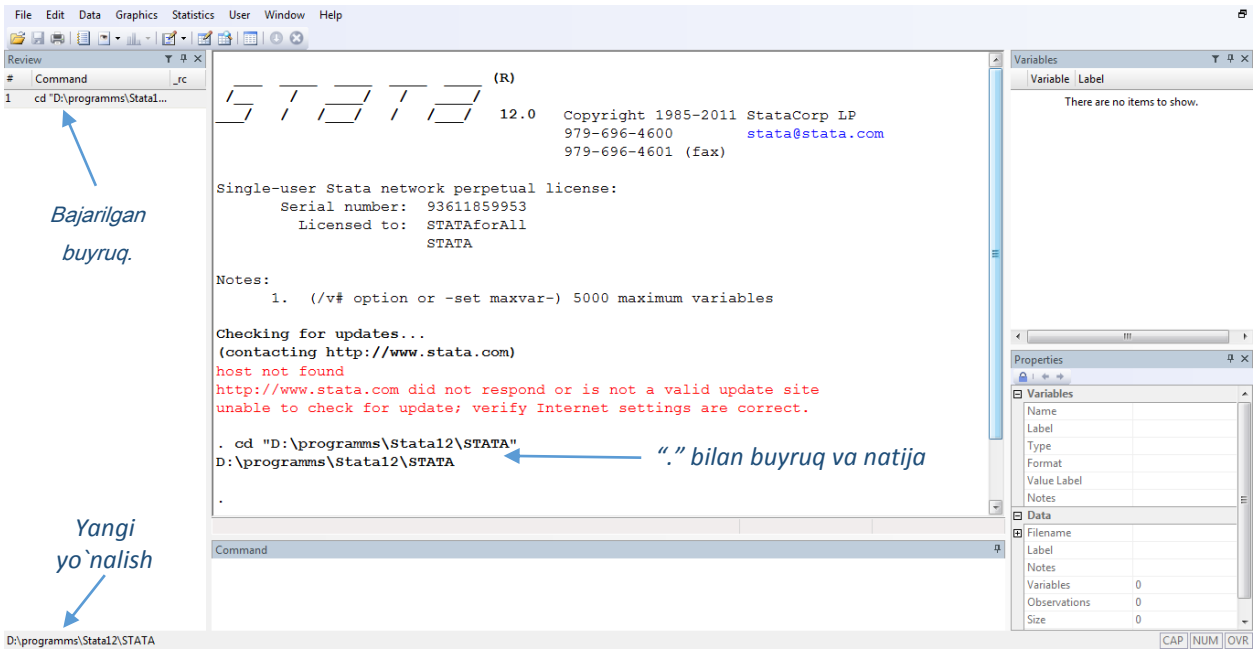
3-rasm. Stata ischi faylida ishlash

Mazkur ishni Stata dasturida **Command** oynasi orqali ham bajarish mumkin, yani: cd “D:\programms\Stata12\STATA”.

Yuqoridagi buyruqni **Command** oynasiga kiriting va **Enter** tugmasini bosing.



Mazkur buyruq natijasi esa 4-rasmda ko`rsatilgan:



4-rasm. Stata dasturi menyusi

Results (natijalar) oynasida buyruq qayta aks ettiriladi va u **Review** (tadqiq) oynasida ham ko'rinadi. Yangi yo'l ekranning chap pastki qismida ko'rsatiladi.

1.2. STATA dasturida ma`lumotlarni kiritish va yuklash.

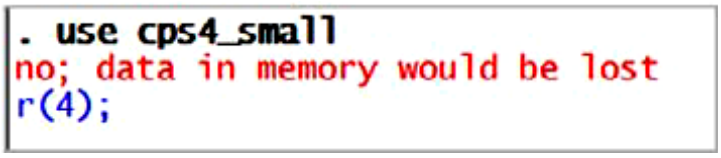
Stata ma'lumotlar fayllarini ochish yoki yuklashning bir necha yo'llari mavjud. Ularning ikki usulini tushuntiramiz.

Fayldan foydalanish. Stata dasturi ishga tushgandan so`ng, Stata ma'lumotlar

fayllarini saqlagan joyingizga o`zingizning ishchi katalogngizni joylashtiring. So`ngra **Command** oynasida **use cps4_small** dan foydalaning va **Enter** tugmasini bosing.

5-rasm. Stata dasturida use cps4_small buyrug'idan foydalanish

Agar ma'lumotlar faylingiz allaqachon ochiq bo'lsa va uni qandaydir tarzda o'zgartirgan bo'lsangiz, bunda Stata buyquqni xato ma`lumot deb quyidagicha javob beradi.



6-rasm. Stata dasturida xato ma`lumot haqidagi xabar

r (4) tugmasini bosganingizda; xato ma`lumot haqidagi xabarni Viewer oynasida o'qishingiz mumkin. Ba'zan bu foydali bo'ladi. Viewer oynasini yopish uchun X ni bosing.

Bu xususiyat siz saqlashingiz mumkin bo'lgan ma'lumotlar faylidagi o'zgarishlarni yo'qotishingizga to'sqinlik qiladi. Agar shunday bo'lsa, avvalgi ma'lumot faylini saqlashingiz yoki quyidagi buyruqni kiritishingiz mumkin:

«clear»

Yuqoridagi buyruq Stata xotirasida bo'lgan narsalarni yo'q qiladi. Ma'lumotlar faylini ochish va xotirani tozalash zarur bo'lsa, quyidagini kiriting:

use cps4_small, clear

Keyingi usul uskunalar panelini ishlatish. Stata data faylini ochishda uskunalar satridagi Open (ochish) belgisini bosing.

8-rasm. Stata dasturida uskunalar paneli

Siz ochmoqchi bo'lgan faylni toping, tanlang va Open buyrug’ini bosing. Review oynasida ko'rsatilgan Stata buyrug'i ko'rsatiladi.

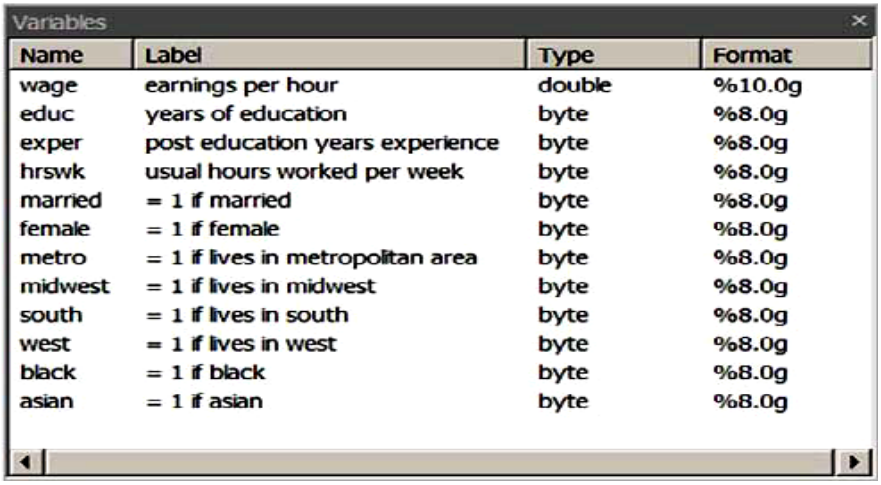
use "C:\data\poe4stata\cps4_small.dta", clear.

Ma'lumotlar faylini Stata dasturida ochishda use buyrug'i bilan erishiladi. Ma'lumotlar faylining yo'li ekranning chap pastki qismida ko'rsatiladi. Yangi ma`lumotni ochish mavjud ma'lumotlarning xotiradan o'chirilishini ko'rsatadi.

9-rasm. Stata dasturida uskunalar paneli

O`zgaruvchilar oynasi (variables window).

O`zgaruvchilar oynasida stata ma'lumotlari faylining parametrlari beriladi (10-rasm). Shuningdek, o'zgaruvchilarning izohlari ko`rsatiladi, agar ular mavjud bo'lsa, ularning turi va formati ham ko'rsatilgan bo`ladi.

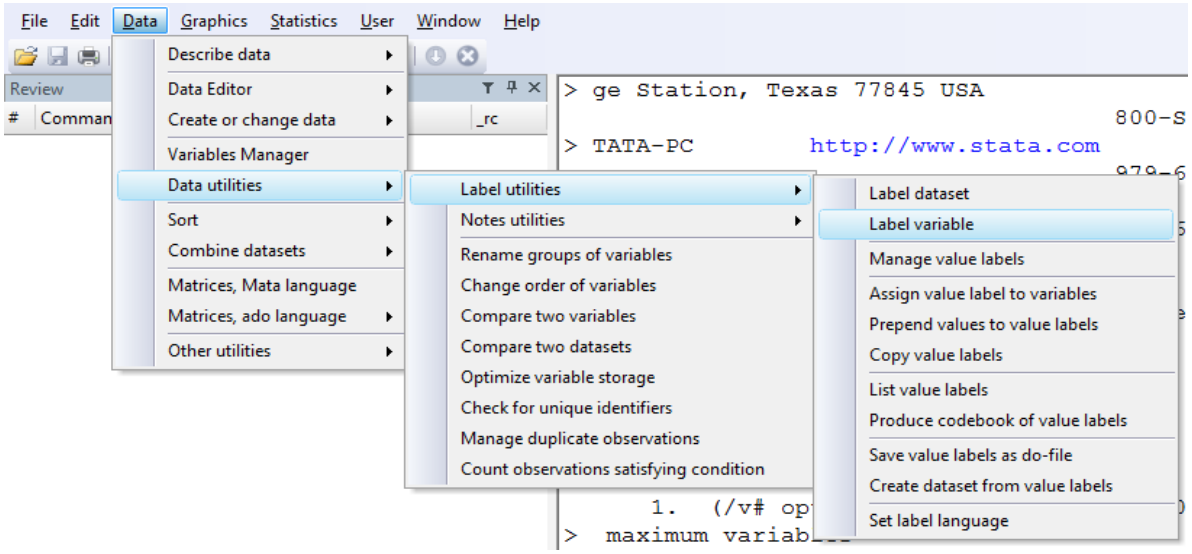


10-rasm. O`zgaruvchilar oynasida stata ma'lumotlari faylining parametrlari

Izohlar foydalidir va ularni osongina qo'shilishi, o'zgartirilishi yoki o'chirilishi mumkin.

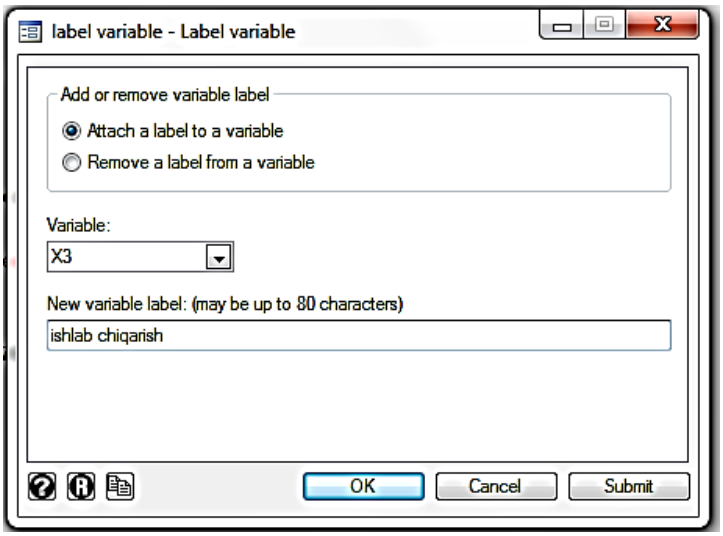
Bitta o`zgaruvchiga izoh joylashtirish. Stata ning yuqoridagi menusidan **Data > Data utilities > Label utilities > Label variable** ni tanlang. Bu ketma-ketlik 6-rasmda izohlangan.

6-rasm



Natijada paydo bo`lgan oynada, yangi o'zgaruvchiga izoh yozish uchun **Attach a label to a variable** ni tanlab so`ngra o'zgaruvchi tanlanadi va **New variable label** joyiga, o`zingiz xohlagan izoh qoldirishingiz mumkin (7-rasm). Undan keyin **OK** tugmasini bosish kifoya.

7-**rasm**



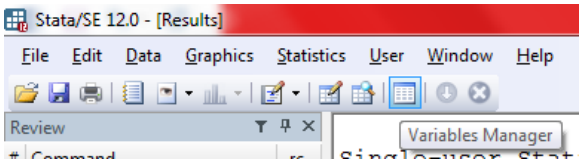
Yuqoridagi ketma-ketlik o`rniga siz **Command** oynasiga quyidagini

kiritishingiz mumkin. Bu quyidagicha bo`ladi:

label variable X3 "ishlab chiqarish"

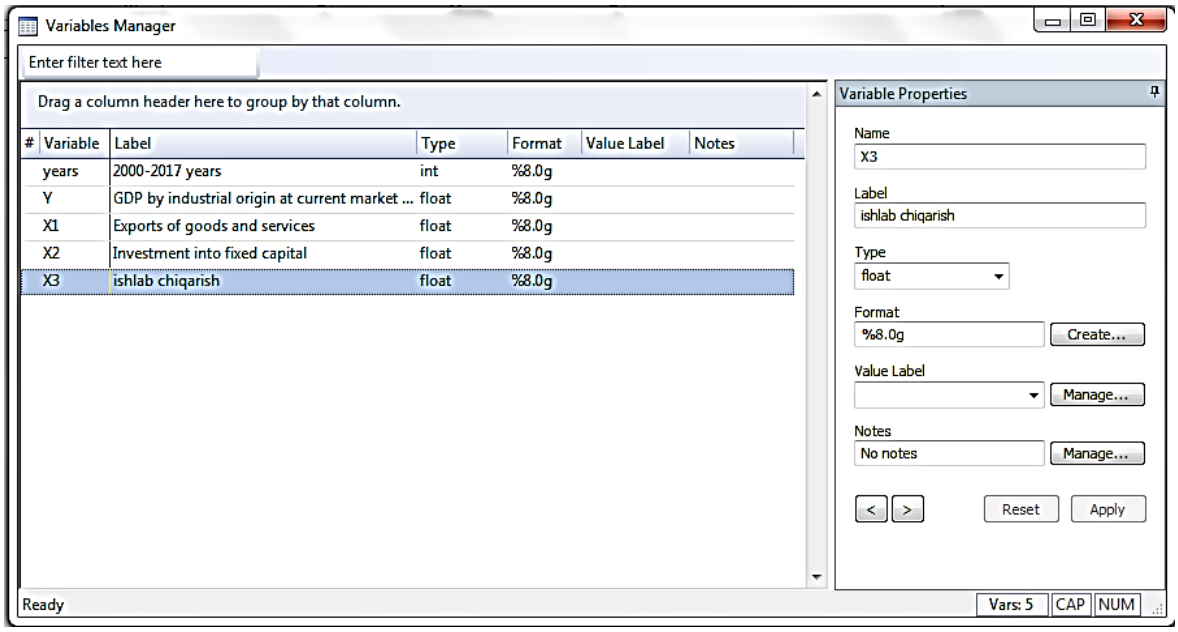
Command oynasiga buyruqni yozganingizdan so`ng **Enter** ni bosing. Bu esa izohni yaratadi va X3 uchun yangi izoh yoziladi. Muloqot oynasidan siz yozgan izohni olib tashlashni ham tanlashingiz mumkin.

Shuningdek, izohlar menejeridan ham foydalanish mumkin. Izohlar boshqaruvchisiga boshqarish uchun yagona joy - **Variables Manager**. Stata ning yuqoridagi menu sidan quyidagi belgini bosing.



Variables Manager o'zgaruvchilari xususiyatlarini ochish uchun **Variable Properties** ni bosing. Bu yerda o'zgaruvchining nomini o'zgartirish, izohlar qo'shish va o'zgaruvchining formatini va guruhlarini boshqarishingiz mumkin (8-rasm).

8-rasm



1.3. MA'LUMOTLARNI TASVIRLASH VA TASVIRIY STATISTIKASINI OLISH

Ma'lumotlar faylini har doim ko`rish va tasvirlash uchun bir necha usullar mavjud. Avval quyidagi buyruqni kiriting

describe

Bu esa xotiradagi ma'lumotlar majmuasining xulosasini, shu jumladan o'zgaruvchilar ro'yxati, ular haqidagi ma'lumotlar va ularning izohlarini tasvirlaydi. Natijalarning bir qismini 9-rasmda ko`rishinggiz mumkin.

9-rasm

```
. describe

Contains data from 2017.dta
  obs:            18
 vars:             5                          1 Mar 2019 15:38
 size:           324
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
years           int    %8.0g                  2000-2017 years
Y               float  %8.0g                  GDP by industrial origin at current market
                                                prices
X1              float  %8.0g                  Exports of goods and services
X2              float  %8.0g                  Investment into fixed capital
X3              float  %8.0g                  Manufacturing
-------------------------------------------------------------------------------
Sorted by:
```

Keyin quyidagi buyruqni kiriting:

Summarize

Natijalar oynasida biz qisqacha xulosa statistikasini ko`ramiz. Bu esa 10-rasmda tasvirlangan.

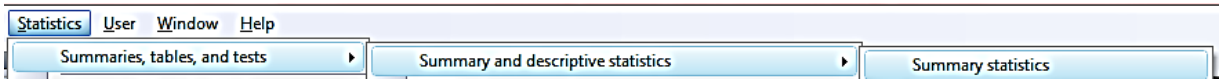
10-rasm

```
. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       years |         18      2008.5     5.338539      2000       2017
           Y |         18    73.49739     75.99263     3.256    254.043
          X1 |         18    19.23783     18.00062      .864     73.392
          X2 |         18    17.95961     19.63861      .744     68.424
          X3 |         18    13.18078     12.69302      .462     42.611
```

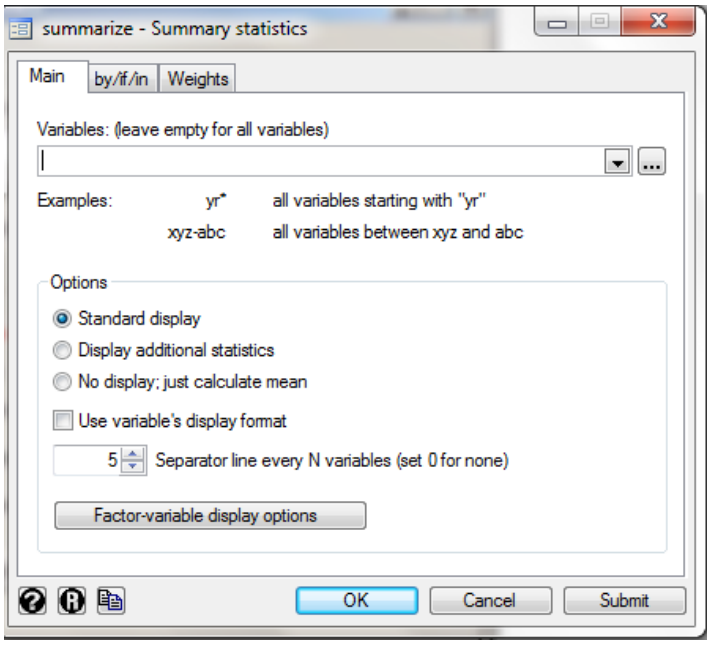
Agarda Stata buyrug'ini unutib qo'ysangiz, yuqoridagi menyulardan foydalanib qisqacha xulosa statistikasini olishingiz mumkin. Siz Stata menyusi

ro'yxatidagi Statistics ni bosing.

Siz undan foydalanishingiz mumkin bo'lgan statistik tahlillarning uzoq ro'yxatini ko`rasiz. Keyin undan **Summaries, tables, and tests** ni tanlang so`ngra **Summary and descriptive statistics** va **Summarry statistics** ni tanlang.

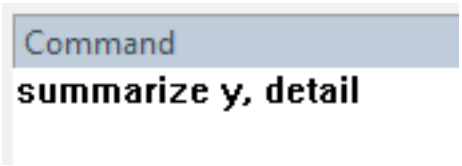


Ko'p variantli ko'rsatadigan dialog oynasi ochiladi. Asosiy abstrakt statistikasi uchun hech qanday tanlov kerak emas. OK-ni tanlang. Stata avtomatik ravishda ma'lumotlarni to'plamidagi barcha o'zgaruvchilar uchun qisqacha xulosa statistikasini taqdim etadi. **Variables** joyiga o'zingizga kerakli o`zgaruvchilar nomlarini yozish orqali yoki individual o'zgaruvchilarni tanlashingiz ham mumkin. **Standard display** kuzatuvlar soni, arifmetik o'rtacha, standart og'ish, o`zgaruvchilarning minimal va maksimal qiymatini beradi.



Tasviriy statistikaning sintaksisi

Quyidagi misollarni sintaksis xususiyatlaridan foydalanib ko'rib chiqamiz. Shuning uchun **Command** oynasiga quyidagi buyrug`ni kiriting va Enter tugmasini bosing.



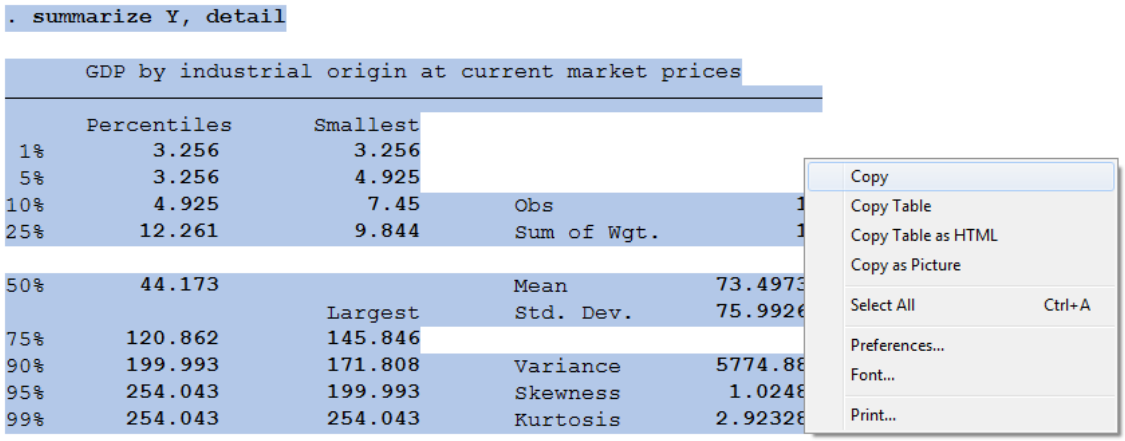
Summarize Y, detail buyrug`i o'zgaruvchan **Y(GDP)** uchun batafsil abstrakt statistikasini hisoblaydi. Eng kichik va eng kattalikdagi **Y** o`zgaruvchining foizlari haqida ma'lumot olishingiz mumkin bo'lgan qo'shimcha ma`lumotlar statistikasi (masalan, skewness va kurtosis) bilan birga ko'rsatiladi. Eslatib o'tamiz, Stata oldingi nuqta (.) bilan berilgan buyruqqa qaytadi.

```
. summarize Y, detail

      GDP by industrial origin at current market prices
-------------------------------------------------------------
      Percentiles      Smallest
 1%         3.256          3.256
 5%         3.256          4.925
10%         4.925           7.45       Obs                  18
25%        12.261          9.844       Sum of Wgt.          18

50%        44.173                      Mean           73.49739
                          Largest       Std. Dev.      75.99263
75%       120.862         145.846
90%       199.993         171.808       Variance       5774.881
95%       254.043         199.993       Skewness       1.02485
99%       254.043         254.043       Kurtosis       2.923281
```

Natijalar oynasining quyi chap qismida – **more -** ko`rsangiz, natijada: Results oynasi to'lgan bo'lsa to'xtatib turadi va qo'shimcha natijalar paydo bo'lishi uchun – **more -** so`zini bosishingiz kerak.

1.4. NATIJALARNI SAQLASH USULI

Nusxa olish va joylashtirish. Umumiy tartib sifatida tavsiya etilmasa ham, natijalarni **Result** oynasinidan belgilab olib sichqonchani o'ng tugmasini bosing.



Bu sizning matni matn sifatida nusxalash (**Ctrl + C**) va qisqa (**Ctrl + V**) yoki pastki belgini bosish orqali hujjatga joylashtirishni beradi.

Agar siz matnni qayta ishlash hujjatiga joylashtirsangiz, to`g`ri joylashtirilgan Stata natijalari noto'g'ri va o'qish uchun qiyinlashishi mumkin. Natijalarning bir qismi o'xshash bo'lishi mumkin. Bu quyidagi ko`rinishda bo`ladi:

Variable	Obs	Mean	Std. Dev.	Min	Max
years	18	2008.5	5.338539	2000	2017
Y	18	73.49739	75.99263	3.256	254.043
X1	18	19.23783	18.00062	.864	73.392
X2	18	17.95961	19.63861	.744	68.424
X3	18	13.18078	12.69302	.462	42.611

Bu shriftni o'zgartiradigan matn protsessorga bog'liq. Siz standart matn uchun Times New Roman shriftini ishlatishingiz mumkin bo'lsa, Stata da natijalar chiqishi uchun Courier New shriftidan foydalanilgan. Shrift hajmini moslash uchun 8 yoki 9 hajmgacha kamaytirishingiz kerak bo'ladi. Shunda natija qisman to`g`ri chiqishi mumkin:

Variable	Obs	Mean	Std. Dev.	Min	Max
years	18	2008.5	5.338539	2000	2017
Y	18	73.49739	75.99263	3.256	254.043
X1	18	19.23783	18.00062	.864	73.392
X2	18	17.95961	19.63861	.744	68.424
X3	18	13.18078	12.69302	.462	42.611

Belgilangan materialni rasm sifatida nusxalash sizning ma`lumotlaringizni hujjatga joylashtirilganda to`g`ri va aniq ko'rinishdagi rasm bo`lib joylashadi. Rasm sifatida joylashishi, bu sizning ma`lumotlaringizni tahrir qilish imkoniga ega bo`lmaysiz.

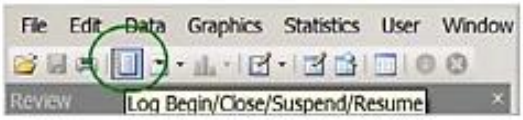
II-BOB. OZIQ-OVQAT SANOATINING DINAMIK MA`LUMOTLARI ASOSIDA STATA DASTURIDA REGRESSION MODELLAR TUZISH

2.1 KLASIK CHIZIQLI MODELNI YARATISH UCHUN MA`LUMOTLARINI YUKLASH

“Principles of Econometrics”, nomli adabiyotning 4-nashrining 1-bobining bir necha rejalarida, uy xo`jaliklarining haftalik xarajatlari bilan daromadlari o`rtasidagi munosabatlarning iqtisodiy modeli ko`rsatib o`tilgan va shu asnoda stata dasturida oziq-ovqat sanoatining dinamik ma`lumotlari asosida model tuzishga harakat qilamiz. Birinchidan Stata dasturini ishga tushuramiz va ishchi katalogini kiritamiz. Buning qanday qilinishi kompyuterning operatsion tizimiga va stata dasturining do-file lari joylashgan joyiga bog`liq.

Ishchi katalog turi quyidagicha: **cd “C:\Users\User\Documents\stata”** yuqoridagi ishchi katalogni buyruqlar oynasiga kiriting va **Enter** ni bosing. Yoki Stata ochiladigan menyusidagi **File > Change Working Directory** katalogini tanlang.

Yangi muammoga duch kelsangiz, ochiq bo`lgan har qanday jurnal faylini o`chirib tashlashingiz kerak. Yangi bir stata log faylini boshlash uchun uskunalar paneli belgisiga bosing.



Bir kunlik faylingiz ochiq bo`lsa, sizga ba`zi imkoniyatlarni beradigan dialog paydo bo`ladi. Yangi jurnal faylini ishga tushirishdan avval siz eski faylni yopishingiz kerak. Yoki buyruq oynasida quyidagi buyruqni kiriting:

log close

Stataning xotirasini tozalash uchun quyidagini kiriting:

clear

2.1.1. KUNDALIK LOG FAYLNI ISHGA TUSHIRISH VA OCHISH

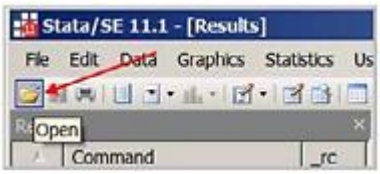
Stata log faylini ishga tushirish yoki yopish uchun uskunalar paneli belgisiga bosing yoki quyudagi buyruqni kiriting:

log using chap02, replace text

Ushbu **log** fayli mavjud katalogda matn formatida ochiladi. Variantni almashtirish **chap02.log** ning avvalgi versiyasiga yozilib, o`chiriladi.

Izoh: Foydalanuvchilar har bir bo`lim yoki bo`limning bir qismi uchun log faylini ochishlari kerak. Oldingi boblar uchun log fayllarini ochish va yopishni sizga eslatib qo`yamiz, lekin keyingi boblarda bajarib bo`lmaydi. Kundalik log fayllardan foydalanish odat qiling.

Oziq-ovqat xarajatlari misoli haqidagi ma`lumotlar stata nomli faylida va food_exp faylida joylashgan. Stata ma`lumotlar faylini ochish uchun uskunalar panelidagi **Open** (foydalanish) tugmasini bosing



Food_exp.dta toping, uni tanlang va **Open** tugmasini bosing. Yoki buyruqlar oynasida, joriy papkadan ma`lumotlar faylini ochish uchun quyidagi buyruqni kiriting:

use food_exp

Agar Stata xotirasini bir vaqtning o`zida yangi ma`lumotlar faylini ochish hamda tozalash zarur bo`lsa, quyidagicha buyruqni kiriting:

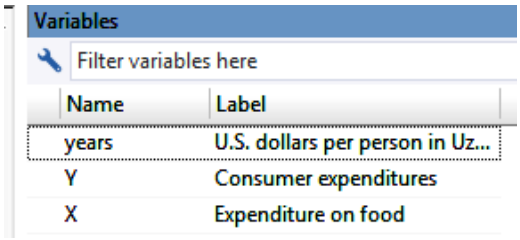
use food_exp, clear

Oldindan ochilgan ma`lumotlar to`plamini xotiradan o`chiradi. Biroq, yangi ma`lumotlar faylini ochishdan oldin "tozalovchi" ma`lumot faylini amalga oshirish xavfsizroqdir.

Stata dasturida internet saytidan ma`lumotlarni yuklash ham mumkin. Buning uchun quyidagi buyruqni kiring:

use http://www.stata.com/texts/s4poe4/food

O`zgaruvchilar oynasida ikkita parametr ro`yxatga kiritiladi: **Y** va X izohi bilan birga. O`zgaruvchilar Turi va Formati haqida boshqa ma`lumotlar ham paydo bo`lishi mumkin. Biz ustunlarni to`ldirish uchun tanladik.



2.1.2 MA`LUMOTLARNI TASVIRLASH VA RO`YXATLASH

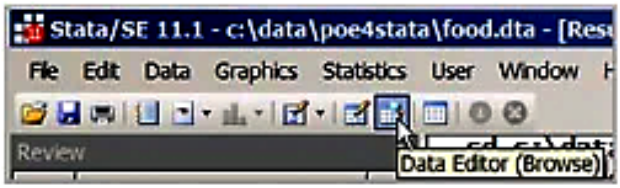
Har bir yangi masalani boshlashdan oldin ma`lumotni ko`rib tekshirib chiqish kerak. Buyruqlar oynasiga quyidagi buyruqni kiriting:

describe

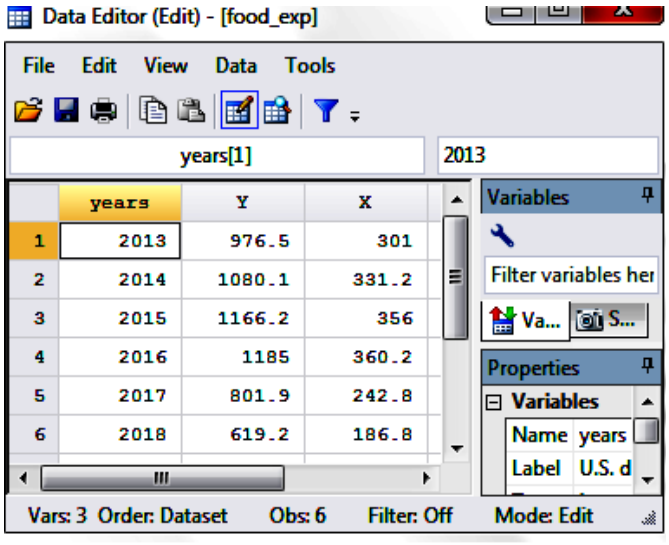
Ushbu o`zgaruvchilar haqida ko`proq ma`lumot olish uchun Buyruqlar oynasida **help describe** buyrug`ini kiriting. Oddiy sarlavha uchun hech narsa talab qilinmaydi, shuning uchun **OK** tugmasini bosing.

```
Contains data from food_exp.dta
  obs:             6
 vars:             3                          5 Nov 2019 16:24
 size:            60
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
years           int     %8.0g                 U.S. dollars per person in Uzbekistan
Y               float   %8.0g                 Consumer expenditures
X               float   %8.0g                 Expenditure on food
-------------------------------------------------------------------------------
```

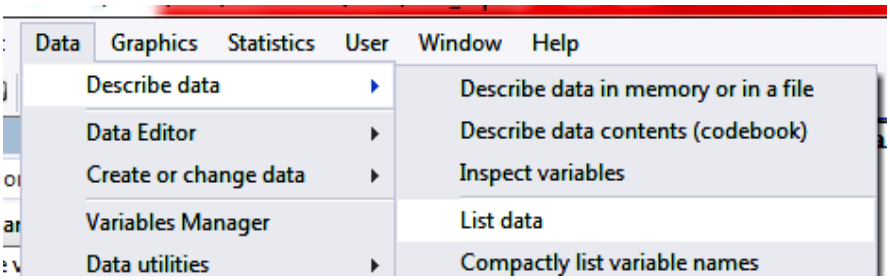
Yuqoridagi ma`lumotlar fayli **food_exp.dta** haqida umumiy ma`lumot. Biz ma`lumotlarimizni tekshiramiz. Ma`lumotlar brauzeri (**Data Editor**) dan foydalaning.



Ma`lumotlar qiymatlarini ko`rish imkonini beruvchi elektron ko`rinishi ochiladi.



Data Editor (Browser)ni yopish uchun yuqoridagi “x” tugmasini bosing. Agarda siz u ma`lumotlarni chop qilishni yoki ba`zilarini ro`yxatlamoqchi bo`lsangiz yuqoridagi menu oynasidan **Data > Describe data > List data** ni bosing.



Ochilgan dialog oynasida, o`zgaruvchini tanlang va Natijalar oynasida barcha ma`lumotlarni ro`yxatlash uchun OK tugmasini bosing. Stata buyrug`i ro`yxatlash. Ro`yxatlash buyrug`i sintaksisi quyidagicha:

list [varlist] [if] [in] [, options]

Muayyan o`zgaruvchilar qiymatlarini ro`yxatlash uchun o`zgarmaydigan nomlarini kiriting. Ro`yxatlangan qiymatlar oralig`i muayyan chiziqlarni ko`rsatish uchun mantiqiy "if" yoki "in" yordamida o`zgartirilishi mumkin. Misol uchun:

list in 1/5

list Y in 1/5

list Y if X <= 360

Natijalar oynasida quyidagilarni ko`rsatadi:

```
list in 1/5
```

	years	Y	X
1.	2013	976.5	301
2.	2014	1080.1	331.2
3.	2015	1166.2	356
4.	2016	1185	360.2
5.	2017	801.9	242.8

```
list Y in 1/5
```

	Y
1.	976.5
2.	1080.1
3.	1166.2
4.	1185
5.	801.9

```
list Y if X <= 360
```

	Y
1.	976.5
2.	1080.1
3.	1166.2
5.	801.9
6.	619.2

Natijalar oynasi to`lsa, siz –**more-** so`zini ko`rasiz va ko`proq ma`lumotni ko`rish uchun uni bosishingiz kerak shunda barcha ma`lumotlarni topa olasiz. Bu –**more-** so`zi pauzani bildiradi. Yana bosing yoki bo`sh joyni bosing. Stata log faylini ishga tushurgandan keyin **set more off** buyrug`ini kiritgan bo`lsa, pauza xususiyatini o`chirib qo`yadi.

2.2. MA`LUMOTLARNING TAHLILIY STATISTIKASI

Keling, ma`lumotlarning ta`rif faylida bildirilgan bir xil xulosa statistik qiymatlarga ega ekanligini tekshiring. Yuqoridagi menyulardan foydalanib ishlating.

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

Natijada muloqot oynasidagi ma`lumotlar to`plamidagi barcha o`zgaruvchilar bo`yicha abstrakt statistika uchun o`zgaruvchilarni tanlang va OK tugmasini bosing. Bundan tashqari, dialog oynasini ochish uchun **db summarize** yoki **db su** buyruqlarini kiritishingiz ham mumkin shunda dialog oynasi paydo bo`ladi. Yoki buyruqlar oynasida quyidagi buyruqni kiriting:

Summarize

```
summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
years	6	2015.5	1.870829	2013	2018
Y	6	971.4833	222.7953	619.2	1185
X	6	296.3333	68.90558	186.8	360.2

Xulosa statistika buyrug`i sintaksisi quyidagicha:

summarize [varlist] [if] [in] [weight] [, options]

Asosiy variant bizga batafsil abstrakt statistikani olish imkonini beradi. Stata buyrug`i oynasida quyidagini kiriting:

summarize Y, detail

```
. summarize Y, detail

                    Consumer expenditures
-------------------------------------------------------------
      Percentiles      Smallest
 1%         619.2          619.2
 5%         619.2          801.9
10%         619.2          976.5       Obs                   6
25%         801.9         1080.1       Sum of Wgt.           6

50%        1028.3                      Mean           971.4833
                          Largest       Std. Dev.      222.7953
75%        1166.2          976.5
90%          1185         1080.1       Variance       49637.74
95%          1185         1166.2       Skewness      -.5982541
99%          1185           1185       Kurtosis       1.936811
```

Natijalar oynasida ma`lumotlarni eng ko`p miqdori, eng kichik va eng katta kuzatishlar, kuzatuvlar soni (Obs) va e`tiborga olinmaydigan WGT summasi ko`rsatilgan. Stata siz tushunmagan ko`p narsalarni e`lon qiladi. Haqiqat siz bilgan narsalarni aniqlashga qodir. Misol uchun, natijalar o`z ichiga quyidagilarni oladi:

```
Mean           971.4833
Std. Dev.      222.7953

Variance       49637.74
```

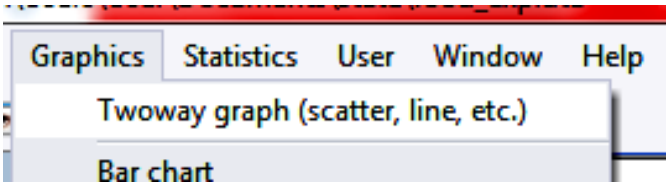
Food_exp o`zgaruvchisi uchun xulosa statistikasi.

- O`rtacha qiymat, $\bar{y} = \sum y_i / N$
- Std. Dev. Variantsning kvadrat ildizlari bo`lgan namunali standart og`ish
- Oddiy variatsianing turi, $var(y) = \sum (y_i - \bar{y})^2 / N - 1$

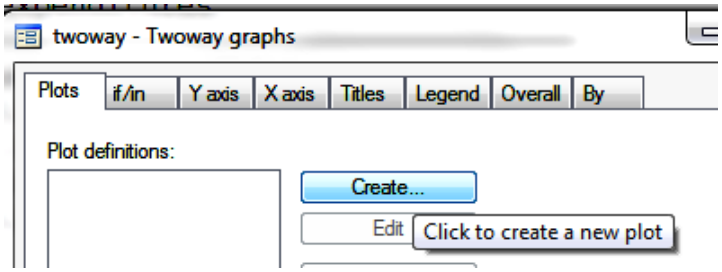
Skewness va Kurtosis qiymatlari keyinroq muhokama qilinadi.

2.3 MA`LUMOTLAR ASOSIDA SCATTER DIAGRAMSINI YARATISH USULLARI

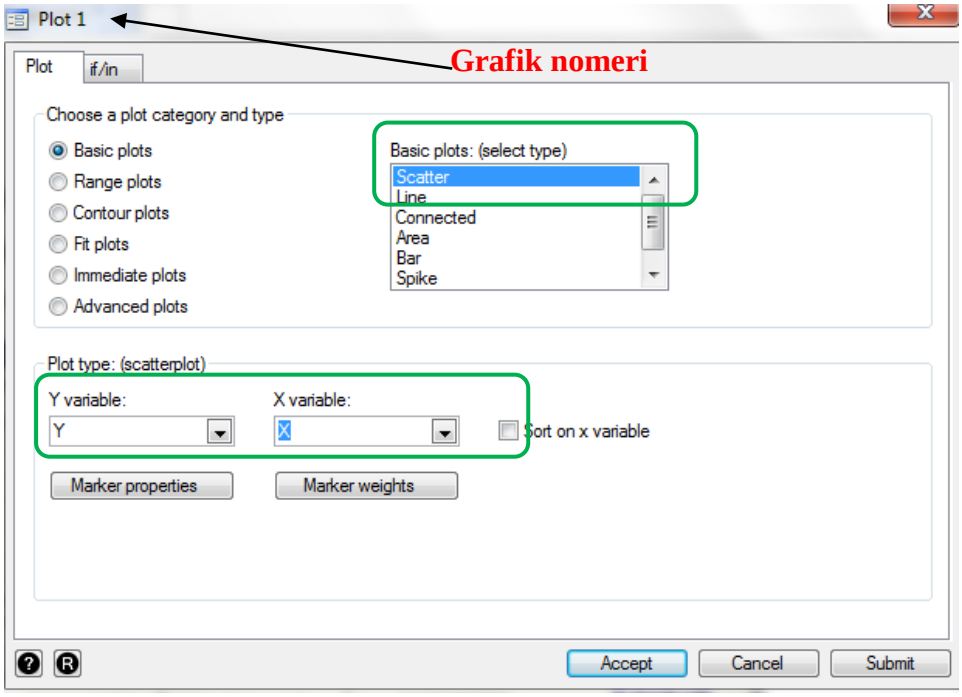
Oddiy regressiya modelida ma`lumotlar qiymatlarini Scatter diagramasida chizish muhimdir. Stata ochiladigan menyusida **Graphics> Twoway graph (scatter, line, etc.)** ni tanlang. Qo`shimcha tafsilotlarni bilish uchun buyruqlar oynasiga **help twoway** buyrug`ini kiriting.



Muloqot oynasida **Create** tugmasini bosing.



Natijada paydo bo`lgan dialog oynasida **Basic plots, Scatter** ni tanlab, keyin ochiladigan o`qlarni foydalanib, Y o`zgaruvchisini (vertikal o`q) va X o`zgaruvchisini (gorizontal o`q) tanlang.



Submit tugmasini bosganingizda, Scatter diagrammasi yaratiladi. Stata buyruqlar oynasiga quyidagi buyruqni kiritib yaratsa ham bo`ladi.

twoway (scatter Y X)

Agar **Accept** tugmasini bossangiz Plot-1-buyruq belgilari oynasida paydo bo`ladi va **OK**-tugmasini bosganingizda grafika yaratiladi.

Grafikni stata grafik kengaytmasi *** .gph**-dan foydalanib ya`ni diskda standart katalogga saqlash uchun

graph save food1, replace

Grafikni va saqlashni **save** opsiyasidan foydalanib, bir qadamda bajarish mumkin

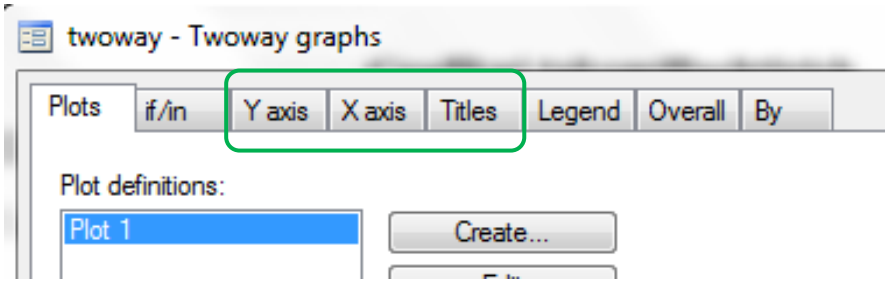
twoway (scatter Y X), saving(food1, replace)

Diskni saqlashning o`rniga xotiraga nom variantini ishlatishingiz ham mumkin. Agar siz laboratoriya sharoitida bo`lsangiz, bu qulay bo`lishi mumkin.

twoway (scatter Y X), name(food1, replace)

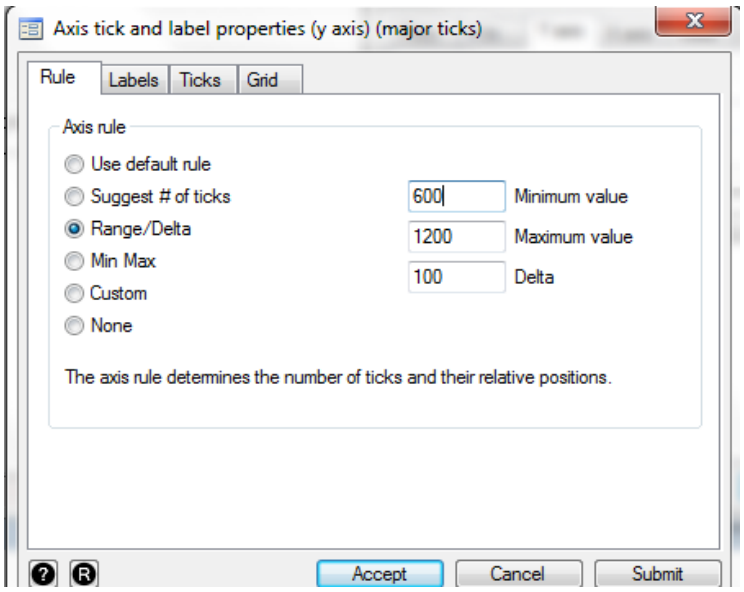
Grafikni takomillashtirish

Grafikni kuchaytirish uchun 1-satrda dialog oynasida Accept ni bosing. Grafika dialog oynasida bu Plot 1 deb ataluvchi plot ta`rifi yoki profilini yaratadi.



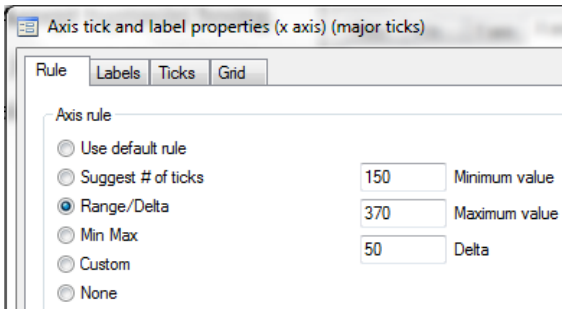
Y o`qi yorlig`ini bosing. Natijada dialog oynasidan quyidagini tanlang.

Bir nechta variant mavjud, ammo vertikal o`qning oralig`ini belgilaylik.

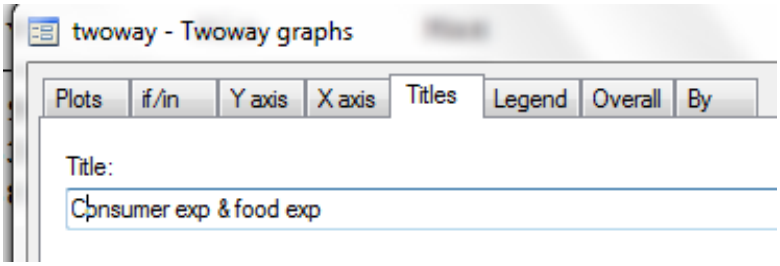


Range / Delta ni bosing. Xulosa statistikasidan **Y** ning minimal va maksimal qiymatlarini ko`rgandik. Ma`lumotlar butunligini ko`rishimiz uchun Minimal qiymat 600 va Eng yuqori 1200 qiymatini kiriting. **Delta** – o`qlar uchun o`lchov birliklari - belgilanadigan belgilar orasidagi bo`shliq. Ushbu qiymatini 100 ga to`g`rilang. **Accept** tugmasini bosing.

Ushbu jarayonni X o`qi uchun maksimal qiymat 370 va **Delta** 50 dan foydalanib takrorlang. **Accept** tugmasini bosing.

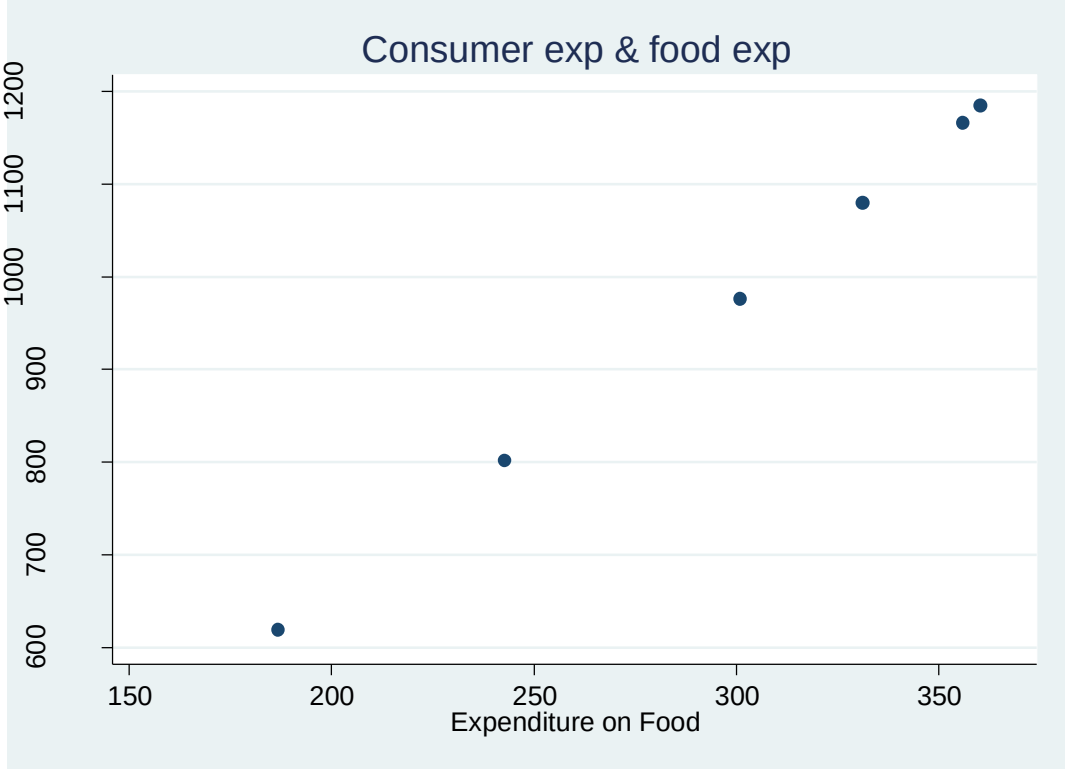


Grafikga nom qo`shish uchun **Twoway Graph** dialog oynasidagi **Titles** yorlig`ini bosing.



Grafik yaratish va oynani yopish uchun **OK** tugmasini bosing. Bu Y-o`qi va X-o`qi bilan chiroyli nuqtalarni ishlab chiqaradi va ma`lumotlar to`plamidagi

o`zgaruvchilar izohlari bilan belgilanadi.



Ushbu nuqtali grafikni yaratish uchun foydalaniladigan Stata buyrug`i quyidagicha:

twoway (scatter Y X), ytitle(Consumer expenditure) ///
ylabel(600(100)1200) xtitle(Expenditure on Food) ///
xlabel(150(50)370) title(Consumer exp & food exp)

Ushbu buyruqdagi "///" belgilar keyingi qatordan buyruqni davom ettiradi. Buyruqlar uzunligi yoki buyruqning muayyan qismlaridan keyin ularga izohlar kiritilishi foydalidir. Fikrlarni ushbu belgilar orqali "/ * .... * / " ifodalaniladi va bu foydalidir, chunki u har qanday joyga joylashtirilishi mumkin va tarkib Stata dasturi tomonidan e`tiborga olinmaydi. Muqobil sharhning buyruqlarini muhokama qilish uchun **help comments** buyrug`ini kiriting.

twoway (scatter Y X), oddiy nuqtali grafik uchun ishlatiladigan buyruqlar.

Bu vergul muhim ahamiyatga ega va qo`llaniladigan variantlarni bildiradi.

ylabel(600(100)1200) Y o`qi oralig`ini, 600 dan 1200 gacha va 100 asosiy nuqtalar oralig`ini bildiradi.

xlabel(150(50)370) 150 - 370 oralig`ini belgilaydi va X o`qi uchun 50 sonini

oralig` hisoblanadi.

title(Consumer exp & food exp) asosiy grafik nomini bildiradi.

Yana bir marta siz ikkita buyruqqa saqlash parametrini qo`shishingiz yoki xotiraga saqlash uchun variant nomini ishlatishingiz yoki grafikni saqlash buyrug`idan foydalanishingiz mumkin. Grafikni saqlash buyruqlaridan ushbu bobda faylda foydalanamiz.

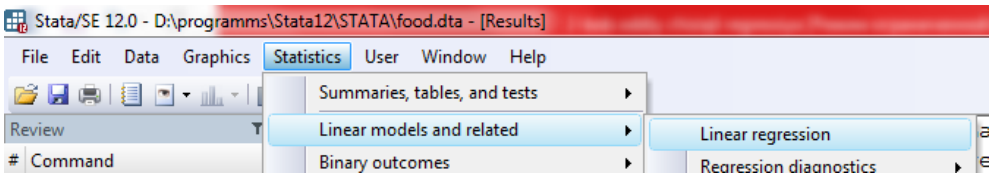
2.4 STATA DASTURIDA REGRESSION MODEL YARATISH USULI

Oddiy chiziqli regressiya modeli:

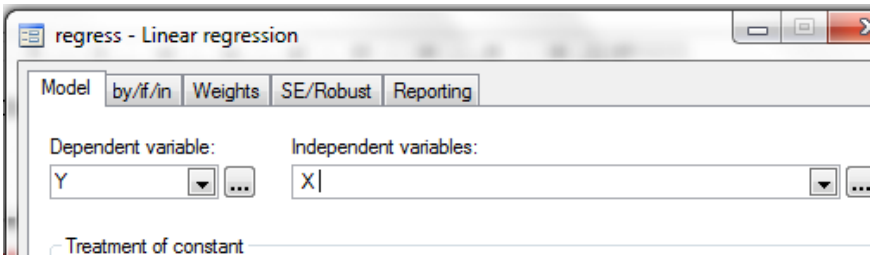
$$y = \beta_1 + \beta_2 x + e$$

Asosiy omil, y (**consumer expenditure**) va ta`sir etuvchi omil x (**expenditure on food**) bo`yicha ma`lumot berilgan holda, noma`lum parametrlarni topish va baholash uchun Stata dasturidan foydalanamiz. Regression tahlilni amalga oshirish uchun yuqoridagi menu dan foydalanilgan holda quyidagicha amalga oshiramiz:

Statistics > Linear models and related > Linear regression



Regress - Linear regression muloqot oynasida **Y** -ga bog`liq bo`lgan o`zgaruvchini tanlang. Bu esa chap tomonda joylashgan regressiya modelidagi asosiy omildir. Ta`sir etuvchi (o`ng tomon) o`zgaruvchisi sifatida **X** ni tanlang (yoki kiriting). So`ngra **OK** tugmasini bosing. Stata dasturi avtomatik tarzda hisob-kitoblarni amalga oshiradi.



Shu bilan bir qatorda, quyidagi buyruq orqali amalga oshirish ham mumkin.

regress depvar [indepvars] [if] [in] [weight] [, options]

Help regress buyrug`i orqali qo`shimcha ma`lumotga ega bo`lasiz. Bu shuni ko`rsatadiki, asosiy omil **depvar** birinchi bo`lib joylashtiriladi, keyin ta`sir etuvchi o`zgaruvchi yoki o`zgaruvchilar (ko`p ommilli regressiya uchun), **indepvars** lar kiritiladi. Agar xohlasangiz, siz **if** yoki **in** bilan cheklash, kuzatishlardagi hajmini belgilash yoki ulardan foydalanish imkoniyatlarini qo`llashingiz mumkin. Ushbu variantlarni (**if** va **in** buyruqlarini) kelgusi boblarda ko`proq muhokama qilamiz.

Iste`mol va oziq-ovqat xarajatlari uchun oddiy regressiya buyrug`i mavjud

regress Y X

buni qisqartirilishi ham mumkin

reg Y X

Natijalar oynasida **regress** uchun Stata buyrug`i berilgan va regression natijalari ko`rsatiladi. Bu narsalarni hammasini o`rganamiz, ammo hozircha **Y** va **Coef** deb nomlangan birinchi ikkita ustunni ko`rib chiqamiz.

Birinchi ustunda o`zgaruvchilarning nomlari berilgan. Kompyuter dasturlari biz "$\beta_1$" va "$\beta_2$" deb chiziqli regressiya tenglamasida o`zgaruvchilarning nomlarini belgilab olganimizni bilmaydi va shuning uchun "β_1" va "β_2" lar o`rniga o`zgaruvchilarning nomlari bilan ifodalaydi. Stata birinchi navbatda tenglama parametrlarni **Coef.** so`zi ostida beradi. Y-endogen **_cons**, ya`ni "doimiy" yoki "asosiy omil" deb nomlanadi, bu **y-endogen** uchun yana bir umumiy nom.

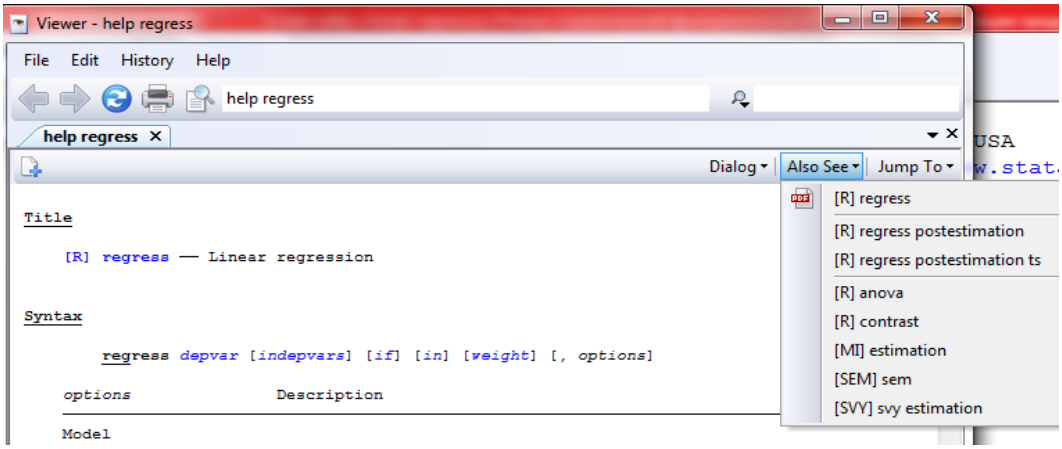
```
. regress Y X
```

Source	SS	df	MS	Number of obs	=	6
				F(1, 4)	=	5315.64
Model	248002.086	1	248002.086	Prob > F	=	0.0000
Residual	186.620714	4	46.6551785	R-squared	=	0.9992
				Adj R-squared	=	0.9991
Total	248188.707	5	49637.7414	Root MSE	=	6.8305

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X	3.232126	.0443313	72.91	0.000	3.109042	3.355209
_cons	13.69666	13.42954	1.02	0.365	-23.58972	50.98304

2.4.1. HISOBLANGAN QIYMAT ($\hat{Y}$) VA QOLDIQLAR

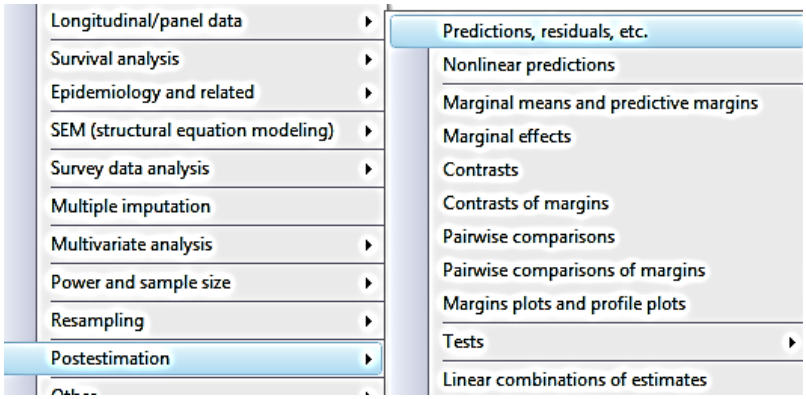
Hisoblangan y qiymatlari va qoldiqlarni **"postestimation"** buyruqlaridan foydalanib topiladi. Ularni **"postestimation"** parametri deb atashadi, chunki ular regressiya modelini baholashadi. Bunga qo`shimcha ma`lumotni, **help regress** buyrug`ini kiritgandan so`ng postestimation parametrlari havolasini oynaning o`ng tomoni yuqorisida va burchagida ko`rishingiz mumkin.



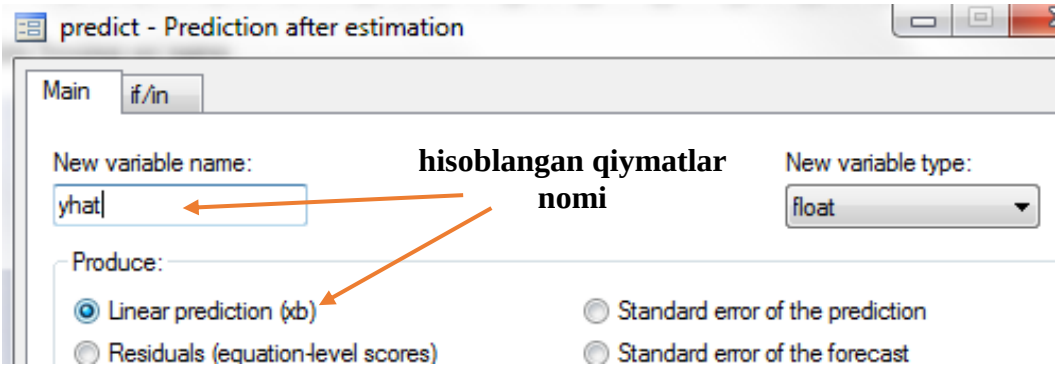
Stata dasturidagi yuqoridagi **menu** lar orqali bajarish yo`li:

Statistics > Postestimation > Predictions, residuals, etc.

Statistics menu sini bosing so`ngra:



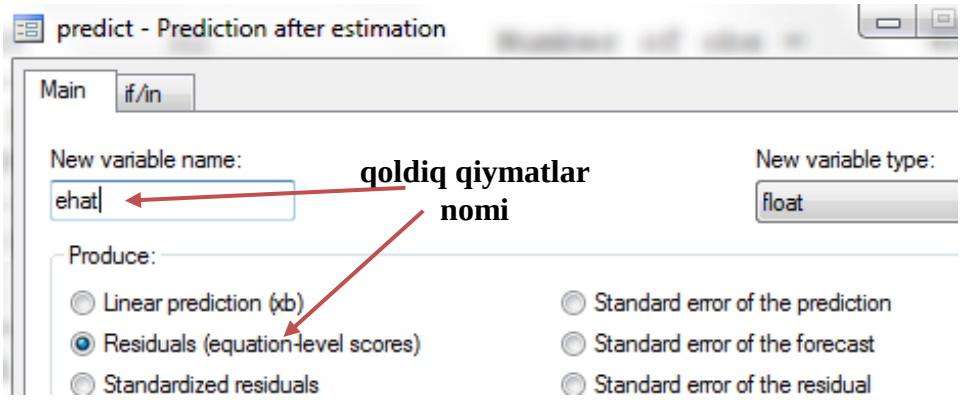
Paydo bo`lgan muloqot oynasida bir nechta muqobillar mavjud. Hisoblangan qiymatni **($\hat{Y}$)** hisoblash uchun **Linear prediction** ni belgilang va hisoblangan qiymatlar nomini kiriting, so`ngra **Ok** tugmasini bosing.



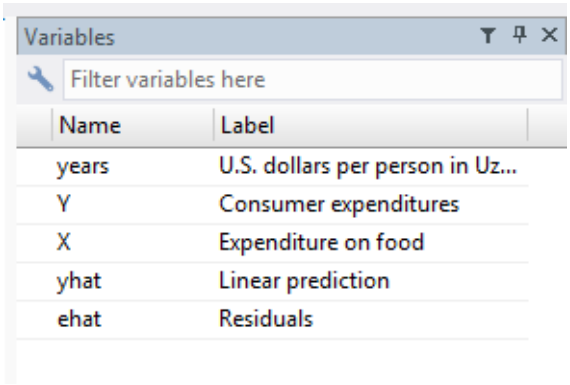
Qoldiqlarni olish uchun,

$$\widehat{e_i} = y_i - \widehat{y_i} = y_i - b_1 - b_2 x_i$$

dialog oynasini qayta oching, **Residuals (equation-level scores)** ni belgilang va qoldiqlarning nomini kiriting ya`ni **ehat** deb so`ngra **Ok** tugmasini bosing.

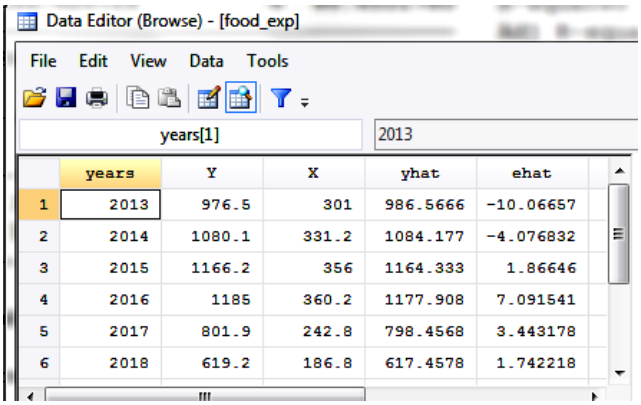


Ushbu buyruqlar o`zgaruvchilari oynasida ko`rinadi ya`ni ikkita yangi o`zgaruvchi, **yhat** va **ehat** paydo bo`ladi.



Ma`lumotlar Brauzerini (**Data editor (Browse)**) tanlab, yoki quyidagi buyruqni ishlatib ushbu qiymatlarni tekshiring

browse



	years	Y	X	yhat	ehat
1	2013	976.5	301	986.5666	-10.06657
2	2014	1080.1	331.2	1084.177	-4.076832
3	2015	1166.2	356	1164.333	1.86646
4	2016	1185	360.2	1177.908	7.091541
5	2017	801.9	242.8	798.4568	3.443178
6	2018	619.2	186.8	617.4578	1.742218

Hisoblangan $\hat{y}$ qiymatlari va qoldiqlarni hisoblaydigan Stata komandalari asosiy **predict** buyrug`i yordamida taxmin qilinadi. Ma`lumotlar brauzerni yoping va buyruqlar oynasiga quyidagini kiriting:

help predict

Predict buyrug`ining asosiy sintaksi:

predict [type] newvar [if] [in] [, single_options]

Nimalar kerak o`zgaruvchining nomi, **newvar** va turi. Hisoblangan y qiymatlarini olish uchun "**x** marta **b**" uchun qisqa bo`lgan **xb-** qisqartmani ishlating.

predict yhat, xb

Bashoratlarni olish, aslida, bu buyruq uchun amrimahol, shuning uchun biz quyidagidan foydalangan afzal:

predict yhat

Qolgan qoldiqlarni olish uchun **options** ga **residuals** dan foydalaning. Oziq-ovqat sarf-xarajatlari modelli uchun bu:

predict ehat, residuals

Qoldiqlarni bir necha turlarda yozish mumkin minimal darajasiga qisqartmasi **r** yoki **res** yoki **resid** ga o`xshab biroz qisqartirilishi mumkin.

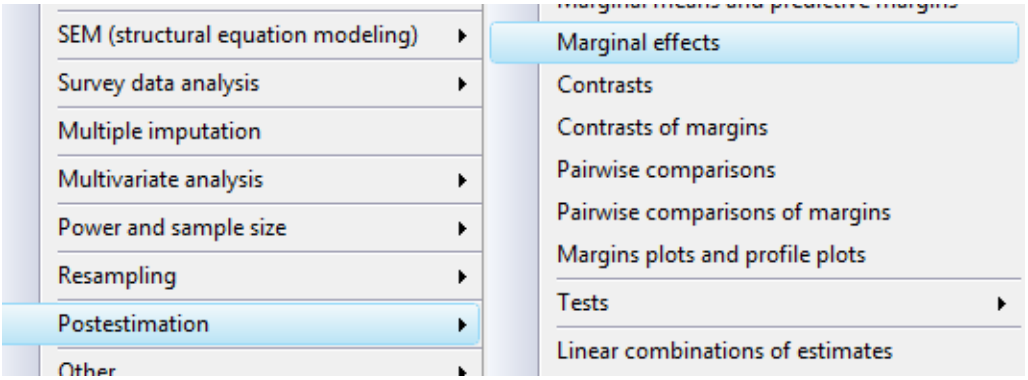
2.4.2 ELASTIKLIKNI HISOBLASH

Parametr tahminlarini va o`zgaruvchilar uchun umumiy statistikani hisobga olgan holda, iste`mol sarf-xarajatlarining oziq-ovqat xarajatlariga nisbatan moslashuvchanligi kabi boshqa miqdorlarni osonlikcha hisoblashimiz mumkin.

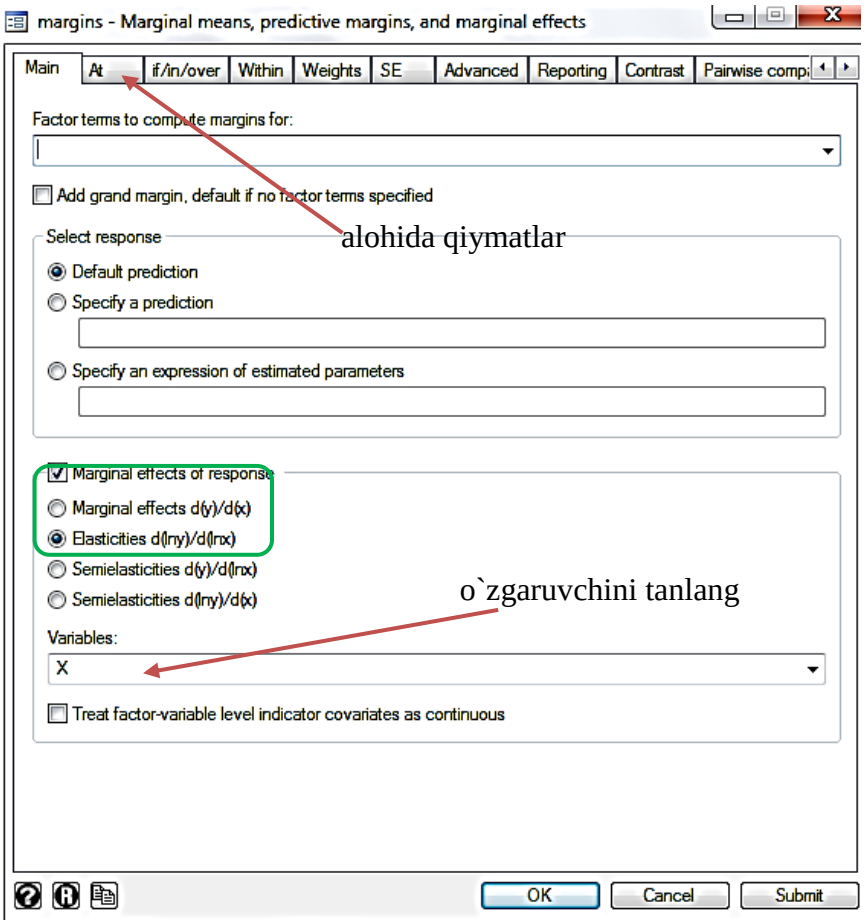
$$\hat{\varepsilon} = b_2 \cdot \frac{\bar{x}}{\bar{y}} = 3.2321 \times \frac{296.33}{971.48} = 0.98$$

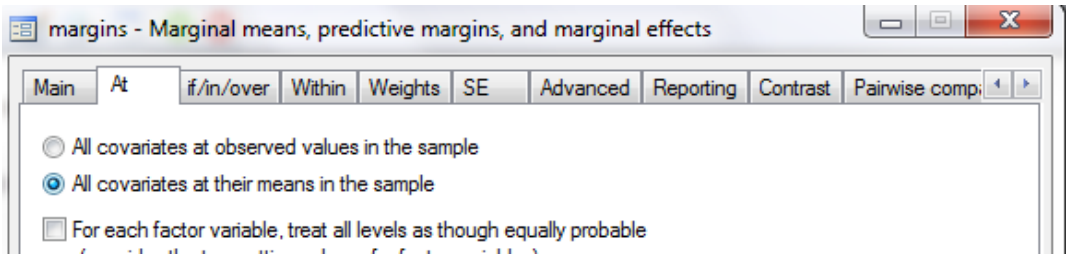
Stata dasturining **postestimation** komandalaridan biri bu elastiklikni avtomatik ravishda hisoblash imkonini beradi. Bu quyidagicha:

Statistics > Postestimation > Marginal effects.



Ochilgan muloqot oynasida **Elasticities** va **Variable** ni tanlang. Bizning oddiy regressiya modelimizda bitta o`zgaruvchi ya`ni **X** ni kiriting. Elastiklikni baholash uchun **At** menu sini tanlang va **All covariates at their means in the sample** ni belgilang.





Natijalar oynasida biz shartli marginal ta`sirlarni ko`ramiz:

```
. margins, eyex( X) atmeans

Conditional marginal effects                    Number of obs    =          6
Model VCE    : OLS

Expression   : Linear prediction, predict()
ey/ex w.r.t. : X
at           : X               =     296.3333 (mean)
```

	Delta-method ey/ex	Std. Err.	z	P>|z|	[95% Conf.	Interval]
X	.9859013	.0138154	71.36	0.000	.9588236	1.012979

Margins dialog oynasi juda kuchli va siz tushunmaydigan hisob-kitoblarni amalga ham oshiradi. Biz uni bir nechta nuqtalar qilish uchun taqdim etamiz. Iqtisodiy tahlilda bir o`zgaruvchining natijasi boshqa bir o`zgaruvchiga ta`sirini ko`rasiz, shuningdek barchasi o`zgarmas bo`lsa bu **marginal effect** deb ataladi. Regression modeldan, bu qiyalik parametridir va oddiy regresyon modeli hosilasi *dy/dx* hisoblanadi. Stata bu qiymatni murakkab modellarda ham hisoblab chiqadi va biz ushbu xususiyatga keyinroq ko`rib chiqamiz.

Linear (to`g`ri chiziqli) regressiya modelida elastiklik quyidagicha:

$$\varepsilon = \frac{\Delta y/y}{\Delta x/x} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y}$$

Deltani lotin shaklida “d” harfi bilan o`zgartiring.

$$\varepsilon = \frac{dy/y}{dx/x} = \frac{dy}{dx} \cdot \frac{x}{y} = \beta_2 \frac{x}{y}$$

To`g`ri chiziqli regressiya modelida **dln (y) = dy/y** va **dln (x) = dx/x** qilib hisoblanadi.

$$\varepsilon = \frac{dln(y)}{dln(x)} = \frac{dy/y}{dx/x}$$

Regressiyadan keyin ushbu elastiklik uchun Stata post-estimation buyrug`i quyidagicha:

margins, eyex(income) atmeans

Stata faqat elastiklik haqida xabar bermasdan, shuningdek bu miqdor uchun standart xatolikni va ishonch oralig`ini hisoblaydi. Bu miqdor hisoblab chiqilgan oziq-ovqat mahsulotlariga xarajat qiymatini ko`rsatadi, bu esa namunadagi o`rtacha qiymatdir. Odatda moslashuvchanlik, "vositalar nuqtasida" $(\bar{x},\bar{y})$=(296.33;971.48) da hisoblab chiqiladi, chunki u regression yo`nalishdagi vakolatli nuqtadir.

Bir nuqtada elastiklikni hisoblash o`rniga, y va x ning har bir qiymatida moslashuvchanlikni topishimiz va keyinchalik bu miqdorni barcha kuzatishlar bo`yicha o`rtacha hisoblashimiz mumkin. Bunga O`rtacha marginal ta`sir (**Average marginal effect**) deyiladi, bu holda o`rtacha elastiklik. Ya`ni quyidagicha

$$AME = \bar{\varepsilon} = \frac{1}{N}\sum_{i=1}^{N} b_2 \frac{x_i}{y_i}$$

Stata dasturi quyidagicha hisoblaydi

$$\widehat{AME} = \frac{1}{N}\sum_{i=1}^{N} b_2 \frac{x_i}{\hat{y}_i} = \frac{1}{N}\sum_{i=1}^{N} b_2 \left(\frac{x_i}{b_1 + b_2 x_i}\right)$$

Bu miqdorni topish uchun Stata **margins** buyrug`ida ilgari ishlatilgan "**atmeans**" so`zi ishlatmaydi.

margins, eyex(income)

Natijalar, endi yuqoridagi kabi, shartli marginal ta`sirlardan ko`ra o`rtacha marginal ta`sirlar farqlidir.

```
. margins, eyex( X)

Average marginal effects                        Number of obs    =          6
Model VCE    : OLS

Expression   : Linear prediction, predict()
ey/ex w.r.t. : X
```

	Delta-method ey/ex	Std. Err.	z	P>|z|	[95% Conf.	Interval]
X	.985126	.0145496	67.71	0.000	.9566092	1.013643

Ushbu hisobni tekshirishimiz mumkin. Regressiyadan so`ng, Stata xotirasida hisoblangan koeffitsientlarni saqlab qoladi. Regression koeffitsientlari **_b[varname]** deb nomlanadi. Iste`mol sarfini kamaytirgandan so`ng, taxminan **_b [X]** sifatida saqlanadi. Keyinchalik namunadagi har bir kuzatuv uchun baholangan moslashuvchanlikni hisoblash mumkin. Bu esa quyidagicha buyruq orqali amalga oshadi:

generate elas = _b[X]*X/yhat

O`rtacha elastiklik

summarize elas

Natijalar oynasida quyidagi paydo bo`ladi:

```
generate elas = _b[X]*X/yhat

summarize elas
```

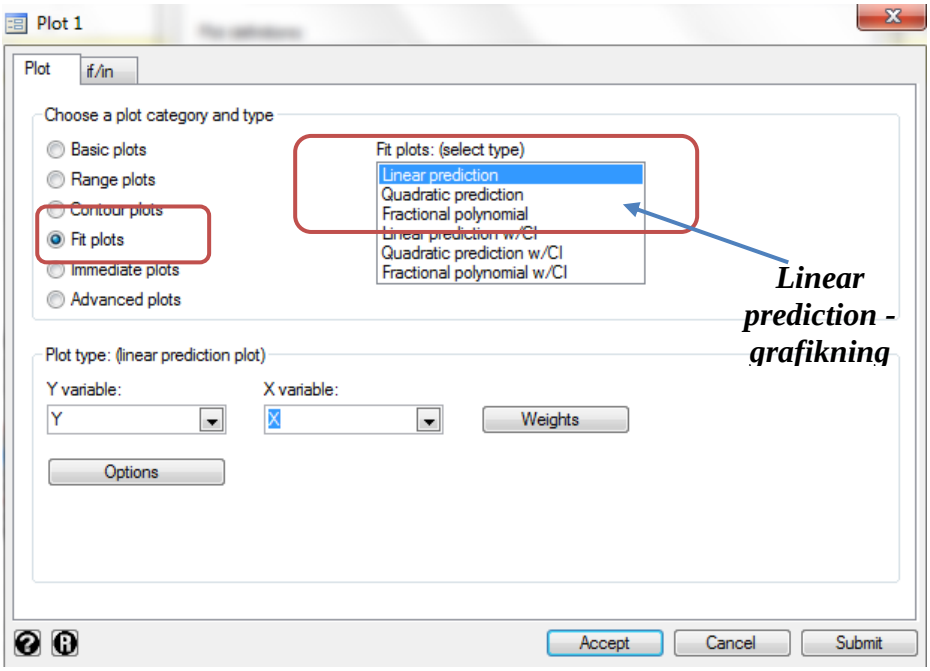
Variable	Obs	Mean	Std. Dev.	Min	Max
elas	6	.985126	.0041189	.9778176	.9883721

Ushbu natijalarda **Std. Dev.** standart og`ishni anglatadi. Margins buyrug`i **Delta-metod Std. Err.** deb atalgan narsalarni ishlab chiqaradi. Ular bir xil emas. Delta usuli POE4 kitobining 5-bobida keltirilgan.

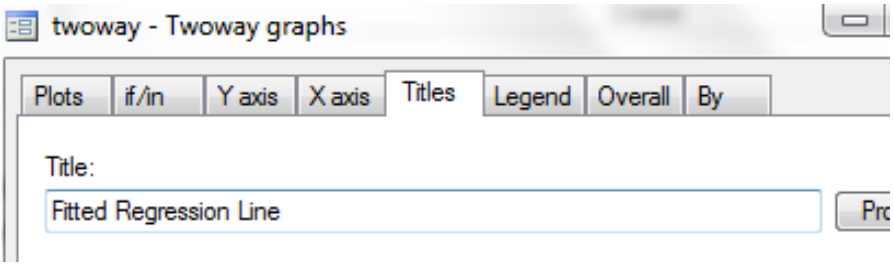
2.4.3 HISOBLANGAN REGRESSION MODEL CHIZIG`INI GRAFIKDA JOYLASHTIRISH

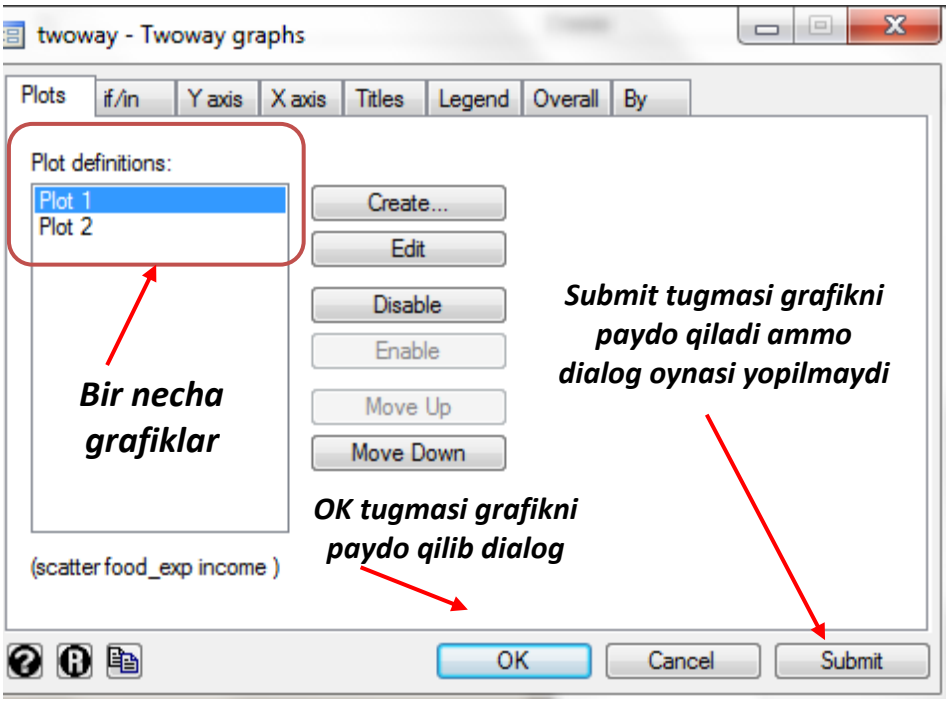
Hisoblangan regressiya chizig`ini tuzish uchun Stata dasturining yuqoridagi ochiladigan menu sidan **Graphics > Twoway graph (scatter, line, etc.)** tanlang.

Agar siz bu mashg`ulotni davom ettirmoqchi bo`lsangiz va nuqtali grafikni 2.3-bo`limga kiritgan bo`lsangiz, u holda 1-grafikda nuqtali diagrammasi allaqachon mavjud. Biz chiziqli prognoz uchun yangi grafik yaratamiz. **Create** tugmasini bosing. **Plot 2** dialog oynasida **Fit plots** tanlang va **Linear prediction** bilan birgalikda Y o`zgaruvchi **(consumer exp)** va X o`zgaruvchi **(food exp)**sini kiriting.

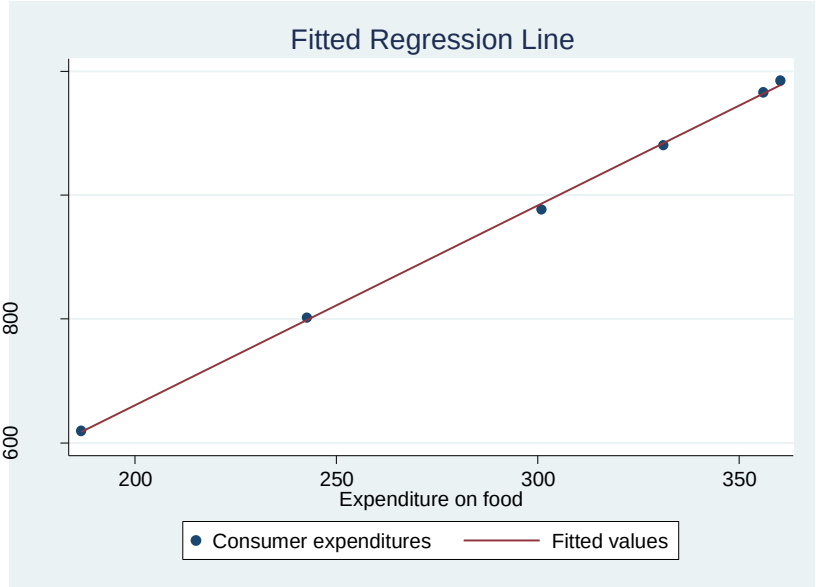


Accept tugmasini bosing. **Titles** so`zini ustiga bosib Plot-2 ni nomni o`zgartiring ya`ni yangi nom kiriting. **Plots** menu siga qayting. **Submit** tugmasini bosish orqali grafikani yaratadi va grafikni ko`rsatadi.





Qachonki bir nechta grafiklar mavjud va xotirada bo`lsa, Stata bir grafini boshqasiga joylashtiradi. Tayyorlangan grafik:



Buyruqlar oynasida biz quyidagi buyruqni kiritsa ham bo`ladi:

```
twoway (scatter Y X)              ///     /* basic plot control */
       (lfit Y X),                ///     /* add linear fit */
       ylabel(600(100)1200)       ///     /* label Y axis */
       xlabel(150(50)370)         ///     /* label X axis */
       title(Fitted Regression Line)      /* graph title */
```

2.4.4. VARIATSIYA VA KOVARVATSIYANI HISOBLASH

Regressiya koeffitsientlaridan tashqari yana bir muhim parametr - bu variatsion xatolik atamasi,

$$var(e_i) = \sigma^2 = E[e_i - E(e_i)]^2 = E(e_i^2)$$

Ushbu parametrni kiritish quyidagicha:

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2}$$

Yuqoridagi $\hat{e}_i = y_i - \hat{y}_i = y_i - b_1 - b_2 x_i$ bu so`nngi kvadrat qoldiqlari. "2" soni - bu $\beta_1$ va $\beta_2$ regression tenglamaning parametrlari soni. Stataning regressiya buyrug`idagi natijaning yuqori qismi yani quyidagi bu jadval **Analysis of Variance** berilgan.

Source	SS	df	MS
Model	248002.086	1	248002.086
Residual	186.620714	4	46.6551785
Total	248188.707	5	49637.7414

Kvadrat qoldiqlar summasi ***N-2*** ***Xatolikning variatsion hisobi***

Residual qatoriga e`tibor bering. **SS** bilan belgilangan ustunda kvadrat yig`indisi mavjud. 186.620714 qiymati kvadratlar qoldiqlarining yig`indisidir.

$$SSE = \sum_{i=1}^{N} \hat{e}_i^2 = SS\ Residual$$

df bilan belgilangan ustun bu (residual) erkinlik darajasi bo`lib, u bu holda N-2 = 4 bo`ladi. Modelning erkinlik darajalari 1, qaysiki biri o`zgaruvchi soni bo`lsa, boshqa biri bu parametrlar soni ya`ni $\beta_2$. Ustun boshi ya`ni **MS** o`rtacha kvadratni anglatadi. Kvadrat qoldiqlarining o`rtacha qiymati taxminiy xato variatsiasidir.

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2} = MS\ Residual = 46.6551785$$

Jadvalning boshqa qismlarini keying boblarda tushuntiriladi.

Regressiya tenglamasini tuzgandan so`ng, bu regressiyaga bog`liq bir necha statistik hisob kitoblar bor, lekin Stata dasturi ularni hisoblab ko`rsatmaydi. Oziq-ovqat mahsulotlarini sarf qilish modeli uchun eng kichkina kvadratlar variatsiasi va kovarvatsia:

$$\widehat{var(b_1)} = \hat{\sigma}^2 \left[\frac{\sum x_i^2}{N\sum(x_i - \bar{x})^2}\right] = 180.35254$$

$$\widehat{var(b_2)} = \frac{\hat{\sigma}^2}{\sum(x_i - \bar{x})^2} = 0.00196526$$

$$\widehat{cov(b_1, b_2)} = \hat{\sigma}^2 \left[\frac{-\bar{x}}{\sum(x_i - \bar{x})^2}\right] = -0.58237347$$

Bularni **estat** buyrug`i yordamida ko`rish ham mumkin. Buyruqlar oynasiga quyidagini kiriting:

help estat

Variatsiani va kovarvatsiani hisoblash uchun quyidagini kiritish kerak:

estat vce

Natijada, regressiya koeffitsientlarining hisoblangan variatsiasi diagonaliga joylashgan va regressiya koeffitsientlarining hisoblangan kovarvatsiasi "diagonal bo`lmagan" mavjud jadval ko`rsatadi. Stata hisoblagan matritsasining yuqori chap burchakdagi b_2 parametrining variatsiasidir.

$$\text{Stata covariance matrix} = \begin{bmatrix} \widehat{\text{var}(b_2)} & \widehat{\text{cov}(b_1, b_2)} \\ \widehat{\text{cov}(b_1, b_2)} & \widehat{\text{var}(b_1)} \end{bmatrix}$$

```
. estat vce

Covariance matrix of coefficients of regress model

      e(V) |          X       _cons
-----------+------------------------
         X |  .00196526
     _cons | -.58237347   180.35254
```

Variatsianing kvadrat ildizlar koeffitsientlari standart xatolikning hisoblangan

koeffitsientidir.

$$se(b_1) = \sqrt{\widehat{var}(b_1)} = 13.42954$$

$$se(b_2) = \sqrt{\widehat{var}(b_2)} = 0.0443313$$

Ular regressiya natijasida ko`rsatiladigan ya`ni Stata dasturi tomonidan avtomatik ravishda chiqariladigan **Std.Err** deb izohlanadi.

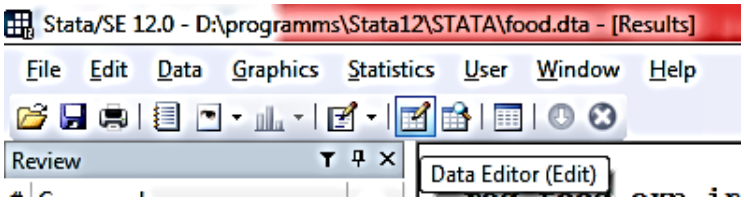
Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X	3.232126	.0443313	72.91	0.000	3.109042	3.355209
_cons	13.69666	13.42954	1.02	0.365	-23.58972	50.98304

2.5 REGRESSION MODELNING PROGNOZ QIYMATLARNI OLISHDA STATA DASTURIDAN FOYDALANISH

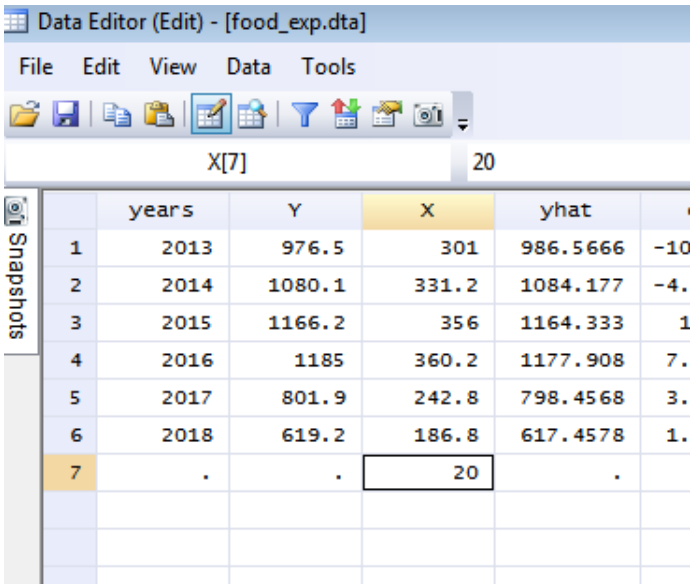
Iste`mol va oziq-ovqat sarf-xarajatlari modeliga asoslanib, oziq-ovqat xarajatlari uchun yilda 20 dollar xarajat keltiradigan taxminiy qiymatini olaylik.

$$\hat{y}_i = 13.696 + 3.21x_i = 13.696 + 3.21(20) = 77.896$$

Ma`lumotlar fayliga to`liq bo`lmagan kuzatuv qo`shish orqali biz Stata dasturini ishlashda davom ettiramiz. Stata uskunalar panelidan **Data Editor** tugmasini bosing.



Bu ma`lumotlar bazasi kabi elektron jadval ko`rinishini ochib beradi, xuddi **Data Browser** ga o`xshash, lekin bu **Data Editor** farq qiladi unda ma`lumotlarni o`zgartira olamiz. 6- ya`ni oxirgi kuzatish natijasi uchun pastga tushing, 7-qatorda daromad uchun yachekani tanlang. **"Formula bar"** oynasida 20 qiymatini kiriting va **Enter** tugmasini bosing.



	years	Y	X	yhat	
1	2013	976.5	301	986.5666	-10
2	2014	1080.1	331.2	1084.177	-4.
3	2015	1166.2	356	1164.333	1
4	2016	1185	360.2	1177.908	7.
5	2017	801.9	242.8	798.4568	3.
6	2018	619.2	186.8	617.4578	1.
7	.	.	20	.	

Data Editor endi daromadning 7-kuzatish natijasi uchun 20 qiymatga ega ekanligini ko`rsatadi, ammo boshqa qiymatlar **"."** ekanini ko`rsatadi. Stata da **"."** belgisi bir davr raqamli ma`lumotlar qiymati yo`qligini bildiradi. **Data Editor** oynasini yopish uchun x-tugmasini bosing. Shu bilan Stata buyruqlar oynasiga quyidagi buyruqlar orqali amalga oshirsa ham bo`ladi.

edit

set obs 7

replace X = 20 in 7

Ushbu **set** buyrug`i ko`plab tizim parametrlarini boshqaradi, kuzatishlar sonini qo`shadigan yordamchi buyruq ham hisoblanadi. Buyruqning joyi **Data > Create or change data > Change contents of variable** orqali ham amalga oshirish mumkin. **Replace** buyrug`ining sintaksisi:

replace oldvar = exp [if] [in] [, nopromote]

X ning 7-qatoriga qiymati 20-sonli qiymatni joylashtirdik. Bu buyruqlar haqida ko`proq ma`lumot olish uchun **help set** va **help replace** orqali ega bo`lishingiz mumkin.

Bashoratli yoki hisoblangan qiymatlarni allaqachon **yhat** deb topib olgandik,shuning uchun biz **"yhat0"** deb nomlab qiymatni topamiz. **Predict** buyug`idagi **xb** parametrini tushurib qoldirib topishimiz mumkin.

predict yhat0

Kuzatishdagi 7-X va yhat0 ma`lumotlarini ro`yxatlash.

list X yhat0 in 7

Natijalar oynasida ko`rishingiz mumkin

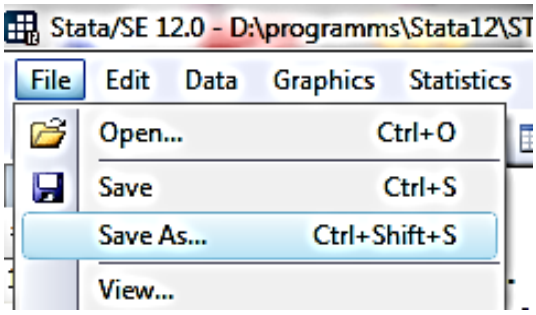
```
. list X yhat0 in 7
```

	X	yhat0
7.	20	78.33918

Har yilda $ 20 oziq-ovqat mahsulotlar xarajati uchun kutilgan yillik iste`mol xarajatlari **yhat0** qiymatida hisoblab ko`rsatilgan.

Stata dastur ma`lumoti faylini saqlash

Stata ma`lumot faylini saqlash uchun **File >Save as** ketma-ketligini amalga oshiring.



Ma`lumotlar faylini **chap02.dta** kabi yangi nom bilan saqlang. Stata dasturida quyidagi buyruq orqali amalga oshiriladi.

save chap02.dta

Oxirgi qadam sifatida **log** faylini yopasiz.

log close

2.6. CHIZIQSIZ REGRESSION MODELLARNI YARATISH USULLARI

Chiziqli regression modeli chiziqsiz, egri chiziqli, munosabatlarni kiritish uchun foydalanish mumkin. Bu muammo emas, chunki chiziqli regressiya parametrlarni chiziqsiz bo`lmagan tarzda kiritilmaydigan modelga ishora qiladi.

Chiziqli regressiya $y = \beta_1 + \beta_2 x + e$. Y va x o`zgaruvchilar boshqa o`zgaruvchilardan iborat bo`lishi mumkin. Chiziqsiz regressiya $y = \exp(\beta_1 + \beta_2 x) + e$ chunki parametrlar chiziqsiz parametr sifatida ishtirok etadi, xuddiki exponensialni shakilda.

Ikkita o`zgaruvchi polinomiya va logarifmik atamalar sifatida o`zgarib ishlatiladi. Ushbu bo`limda kvadratik va log-linear modellarni ko`rib chiqamiz.

2.6.1 KVADRATIK MODEL

Biz haqiqiy **estate data** ma`lumotlarini qo`llash orqali biz kvadratik modelni hisoblaymiz ya`ni $y = \beta_1 + \beta_2 x^2 + e$. Yangi log faylini ishga tushiring va ***br.dta*** ma`lumotni oching.

log using chap02_quad, replace text

use "C:\Users\User\Documents\stata\br.dta", clear

describe

summarize

Tasvir va xulosa statistikasi qisman,

```
              storage  display     value
variable name   type   format      label      variable label
------------------------------------------------------------------------------
price           float  %9.0g                  sale price, dollars
sqft            float  %9.0g                  total square feet

. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |       1080    154863.2    122912.8      22000    1580000
        sqft |       1080    2325.938    1008.098        662       7897
```

Yangi o`zgaruvchi yaratishning eng sodda usuli - o`zgaruvch **sqft** kvadratiga teng bo`lgan yangi **sqft2** o`zgaruvchini yaratishdir.

generate sqft2=sqft^2

Uy narxi bilan uy hajmining kvadrati bilan tuzilgan regressiya va hisoblangan qiymatlarni olish, **priceq**.

regress price sqft2

predict priceq, xb

```
. regress price sqft2
```

Source	SS	df	MS	Number of obs =	1080
				F(1, 1078) =	2425.98
Model	1.1286e+13	1	1.1286e+13	Prob > F =	0.0000
Residual	5.0150e+12	1078	4.6522e+09	R-squared =	0.6923
				Adj R-squared =	0.6921
Total	1.6301e+13	1079	1.5108e+10	Root MSE =	68207

price	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sqft2	.0154213	.0003131	49.25	0.000	.014807	.0160356
_cons	55776.56	2890.441	19.30	0.000	50105.04	61448.09

Hisoblangan chiziqni tuzish uchun biz **twoway graph** buyrug`ining bir necha yangi imkoniyatlarini ko`rsatamiz. Birinchidan, biz mavjud ma`lumotlar bilan grafik tuzamiz. Ikkinchidan, ikkita ham nuqtali grafik ham hisoblangan to`g`ri chiziqli grafikni o`rnatamiz . Boshqa grafik turlari uchun **help twoway** ga qarang. **Sort** imkoniyati qo`shilishi kerak, shuning uchun o`rnatilgan chiziq kuzatishlar muntazam ravishda, har xil o`lchamdagi uylar bir-birini ta`qib qilgan holda uyg`unlashtirilgan natija o`rniga silliq va uzluksiz bo`ladi. O`rnatilgan chiziqni asl holatidan ancha qalinroq qilish uchun, **lwidth(medthick)** so`zini qo`shing.

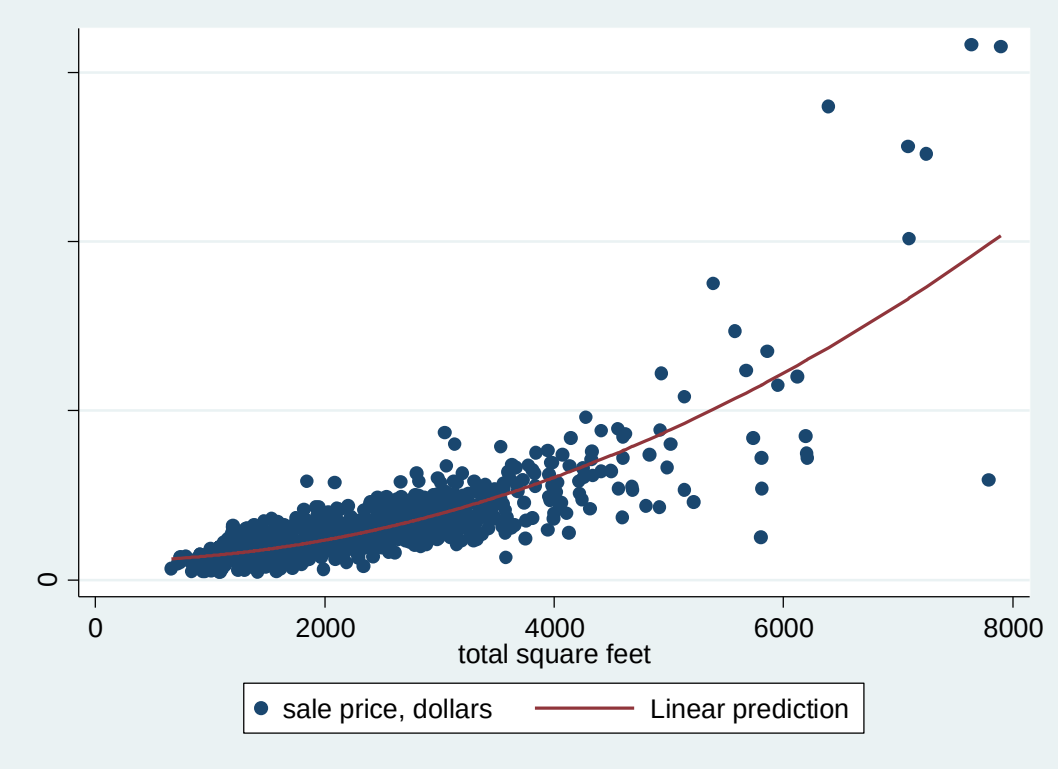
- twoway (scatter price sqft) /// /* basic plot */

(line priceq sqft, /// /* 2nd plot: line is continuous */

sort lwidth(medthick)) /* sort & change line thickness */

- graph save br_quad, replace

Natijada grafik, hisoblangan kvadrat funktsiyani egri chiziqli shaklini ko`rsatadi. Nuqtalar barqaror emas va ular katta uylar qo`shimcha kattalikdagi narxning ko`payishi ko`rsatib turibdi. Bu funksiya tobora yuqoriga o`sib bormoqda. Nuqtalar va moslashuvchanlik hisob-kitoblari biz ilgari ko`rib chiqqan "to`g`ri chiziq" dan yoki chiziqli munosabatlarning funktsional shaklidan o`zgacha.



Hisoblangan kvadrat regressiya funktsiyasi $\hat{y} = b_1 + b_2 x^2$ burchagi $d\hat{y}/dx = 2b_2 x$ ni tashkil qiladi. **x = sqft** ning turli xil qiymatlarida nuqtalarni hisoblash uchun biz oddiy **display (ma`lumot uchun help display)** buyrug`idan foydalanishimiz mumkin, bu **di** deb qisqartirilishi mumkin. Ushbu bobning avvalida mavzusida bo`lgani kabi, biz regressiya koeffitsientiga **_b [sqft2]** ko`rinishida foydalanilgandik. Nuqtalarni **sqft = 2000, 4000** va **6000** qiymatida hisoblashimiz mumkin:

di "slope at 2000 = " 2*_b[sqft2]*2000

di "slope at 4000 = " 2*_b[sqft2]*4000

di "slope at 6000 = " 2*_b[sqft2]*6000

nuqta va elastikliq qiymati:

```
. di "slope at 2000 = " 2*_b[sqft2]*2000
slope at 2000 = 61.685207

.
. di "slope at 4000 = " 2*_b[sqft2]*4000
slope at 4000 = 123.37041

.
. di "slope at 6000 = " 2*_b[sqft2]*6000
slope at 6000 = 185.05562
```

Xuddi shunday buyrug`dan foydalanib, biz regressiyadan hisoblangan qiymatlarni ham topa olamiz.

di "predicted price at 2000 = " _b[_cons]+_b[sqft2]*2000^2

di "predicted price at 4000 = " _b[_cons]+_b[sqft2]*4000^2

di "predicted price at 6000 = " _b[_cons]+_b[sqft2]*6000^2

```
. di "predicted price at 2000 = " _b[_cons]+_b[sqft2]*2000^2
predicted price at 2000 = 117461.77

.
. di "predicted price at 4000 = " _b[_cons]+_b[sqft2]*4000^2
predicted price at 4000 = 302517.39

.
. di "predicted price at 6000 = " _b[_cons]+_b[sqft2]*6000^2
predicted price at 6000 = 610943.42
```

Hisoblangan funksiya uchun elastiklik $\hat{\varepsilon} = (d\hat{y}/dx) \times (x/\hat{y}) = 2b_2x^2/\hat{y}$ dir. Ushbu qiymatlardan foydalanib koeffitsientlarni topamiz

di "elasticity at 2000 = " 2*_b[sqft2]*2000^2/(_b[_cons]+_b[sqft2]*2000^2)

di "elasticity at 4000 = " 2*_b[sqft2]*4000^2/(_b[_cons]+_b[sqft2]*4000^2)

di "elasticity at 6000 = " 2*_b[sqft2]*6000^2/(_b[_cons]+_b[sqft2]*6000^2)

```
. di "elasticity at 2000 = " 2*_b[sqft2]*2000^2/(_b[_cons]+_b[sqft2]*2000^2)
elasticity at 2000 = 1.0503027

.
. di "elasticity at 4000 = " 2*_b[sqft2]*4000^2/(_b[_cons]+_b[sqft2]*4000^2)
elasticity at 4000 = 1.6312505

.
. di "elasticity at 6000 = " 2*_b[sqft2]*6000^2/(_b[_cons]+_b[sqft2]*6000^2)
elasticity at 6000 = 1.8174084
```

Keyinchalik zamonaviy va samarali yondashuv faktor parametrlarini ishlatishdan iborat. Agar biz $x^2 = x \times x$ dan foydalansak, o`zgaruvchan **x = sqft** o`zgaruvchining funktsiyali o`zgaruvchan belgisi **c.sqft** bilan ifodalanadigan bo`lsa, biz yangi o`zgaruvchini yaratmasdan turib, kvadrat funktsiyasini to`g`ridan to`g`ri baholashimiz mumkin.

regress price c.sqft#c.sqft

```
. regress price c.sqft#c.sqft
```

Source	SS	df	MS	Number of obs =	1080
Model	1.1286e+13	1	1.1286e+13	F(1, 1078) =	2425.98
Residual	5.0150e+12	1078	4.6522e+09	Prob > F =	0.0000
				R-squared =	0.6923
Total	1.6301e+13	1079	1.5108e+10	Adj R-squared =	0.6921
				Root MSE =	68207

price	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
c.sqft# c.sqft	.0154213	.0003131	49.25	0.000	.014807	.0160356
_cons	55776.57	2890.441	19.30	0.000	50105.04	61448.09

Ushbu natija bo`yicha oldingi hisoblangan regressiya bilan bir xil, chunki siz ma`lumotni **browsing** orqali tasdiqlashingiz mumkin.

predict price2

Faktor belgilaridan foydalanishning katta afzalligi mavjud, Stata **margins** buyrug`i orqali nuqtalarni aniq hisoblashi va elastikni to`g`ri hisoblashga yordam beradi. Birinchidan, nuqtalar **dydx (*)** bilan chekka joylarni ishlatadi, * bu holatda faqatgina **sqft** bo`lgan barcha model o`zgaruvchilari uchun nuqtalar uchun so`rov ko`rsatiladi. Qulaylik bilan biz bir nechta qiymatlarni hisoblash uchun **at** so`zi bilan nuqtalarni kiritishimiz mumkin.

margins, dydx(*) at(sqft=(2000 4000 6000))

```
. margins, dydx(*) at(sqft=(2000 4000 6000))

Conditional marginal effects                    Number of obs   =       1080
Model VCE    : OLS

Expression   : Linear prediction, predict()
dy/dx w.r.t. : sqft

1._at        : sqft            =        2000

2._at        : sqft            =        4000

3._at        : sqft            =        6000
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
sqft _at						
1	61.68521	1.252385	49.25	0.000	59.23058	64.13983
2	123.3704	2.504769	49.25	0.000	118.4612	128.2797
3	185.0556	3.757154	49.25	0.000	177.6917	192.4195

Nuqtalar faqat to`g`ri hisoblanmaydi, shuningdek bizga standart xatolik va intervalli kiritish ham taqdim etiladi. Elastikliklarni toppish uchun **eyex (*)** buyrug`i ishlatiladi.

margins, eyex(*) at(sqft=(2000 4000 6000))

```
. margins, eyex(*) at(sqft=(2000 4000 6000))

Conditional marginal effects                    Number of obs   =       1080
Model VCE    : OLS

Expression   : Linear prediction, predict()
ey/ex w.r.t. : sqft

1._at        : sqft            =        2000

2._at        : sqft            =        4000

3._at        : sqft            =        6000
```

	Delta-method ey/ex	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sqft						
_at						
1	1.050303	.0336868	31.18	0.000	.9842778	1.116328
2	1.631251	.0203148	80.30	0.000	1.591434	1.671067
3	1.817408	.0112071	162.17	0.000	1.795443	1.839374

Yuqorida qayd etilgan nuqtalar va elastikliklar **shartli** hisoblanadi, chunki ular ma`lum qiymatlarda hisoblangan. O`rtacha marginal effektlar yoki o`rtacha elastikliklarni hisoblash uchun **at** buyrug`i parametridan foydalanmang.

margins, eyex(*)

```
. margins, eyex(*)

Average marginal effects                        Number of obs   =       1080
Model VCE    : OLS

Expression   : Linear prediction, predict()
ey/ex w.r.t. : sqft
```

	Delta-method ey/ex	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sqft	1.102401	.0292176	37.73	0.000	1.045135	1.159666

Ushbu O`rtacha marginal effektlar kutilgan natijalar ekanligiga ishonch hosil qilish uchun biz ularni to`g`ridan-to`g`ri hisoblash uchun oldin hisoblagan regression koeffitsientlarni hisoblashimiz mumkin. Birinchidan, alohida parametr

koeffitsientlarga qanday murojaat qilishni bilish qiyin bo`lishi mumkin. Bu regressiya parametrini **coeflegend** buyrug`i orqali hal qilinishi mumkin.

regress, coeflegend

Bu ANOVA jadvaliga va koeffitsientlarga qaytaradi.

price	Coef.	Legend
c.sqft# c.sqft	.0154213	_b[c.sqft#c.sqft]
_cons	55776.57	_b[_cons]

Faktor modeli spetsifikatsiyasi uchun **_b [varname] - _b [c.sqft # c.sqft]**.

generate elas2 = 2*_b[c.sqft#c.sqft]*(sqft^2)/price2

summarize elas2

```
. generate elas2 = 2*_b[c.sqft#c.sqft]*(sqft^2)/price2

.
. summarize elas2
```

Variable	Obs	Mean	Std. Dev.	Min	Max
elas2	1080	1.102401	.3528353	.2161448	1.890364

E`tibor bering, o`rtacha elastiklik aslida qanday **margins** hisoblangan bo`lsada, yana bir marta **summarize** standart egiluvchanliklarni hisoblab chiqsangiz va **margins** Delta-uslubidagi standart xatoni hisoblab chiqadi. Endi ishchi faylini yopishingiz mumkin.

log close

2.6.2 CHIZIQLI LOGORIFMIK MODEL

Xuddi shu ma`lumotlardan foydalanib biz log-linear modelni taxmin qilamiz. Hisoblangan chiziqli model quyidagicha bo`ladi.

$$\widehat{\ln(y)} = b_1 + b_2 x$$

Y ning hisoblangan qiymatini olish uchun eng to`g`ri yo`l - bu **antilog**ni

hisoblashdir.

$$\hat{y} = \exp\left(\widehat{\ln(y)}\right) = \exp(b_1 + b_2 x)$$

Hisoblangan log-chiziqlini egri chizig`i $d\hat{y}/dx = b_2\hat{y}$ dir va elastiklik $\hat{\varepsilon} = (d\hat{y}/dx) \times (x/\hat{y}) = b_2 x$ dir. Yangi faylni oching va **br.dta**-dan foydalaning. O`zgaruvchi **price** bo`yicha batafsil xulosa statistikasi va uning histogrammini o`ng tomondan uzun egri chiziq bilan ko`rsatilgan taqsimot borligini ko`rasiz.

log using chap02_llin, replace text

use br, clear

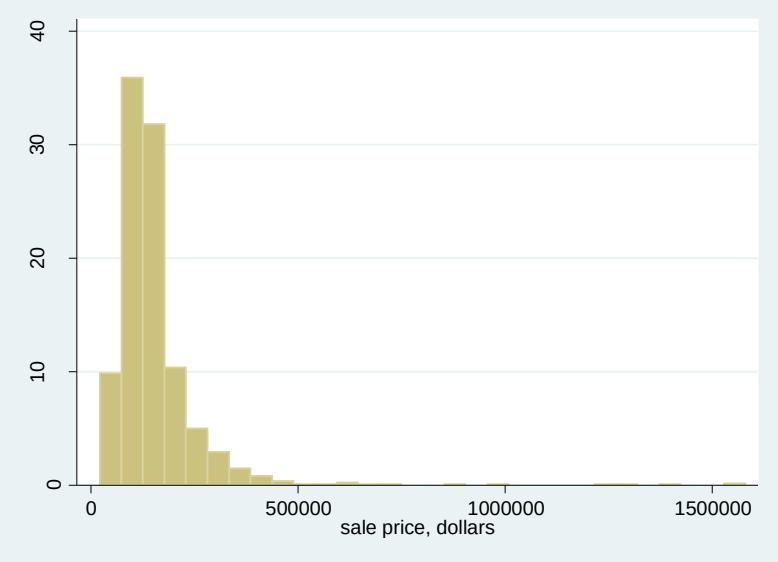
summarize price, detail

```
. summarize price, detail

                    sale price, dollars
-------------------------------------------------------------
      Percentiles      Smallest
 1%         31000         22000
 5%       59897.5         22000
10%         74450         22654       Obs                1080
25%         99000         23000       Sum of Wgt.        1080

50%        130000                     Mean           154863.2
                        Largest       Std. Dev.      122912.8
75%        170325       1280000
90%        244200       1400000       Variance       1.51e+10
95%        315000       1575000       Skewness       6.291909
99%        610000       1580000       Kurtosis       60.94976
```

Shuni ta`kidlash kerakki, natijaga asosan o`rtacha qiymatdan (50 foiz) juda katta qiymatga ega bo`lib, 1,5 million dollarni tashkil etadi. Skewness koeffitsienti noldan ko`ra ijobiydir biz kutgan edik nosimmetrik tarqalishi xuddi normal taqsimotga o`xshashini.

histogram price, percent

graph save price, replace



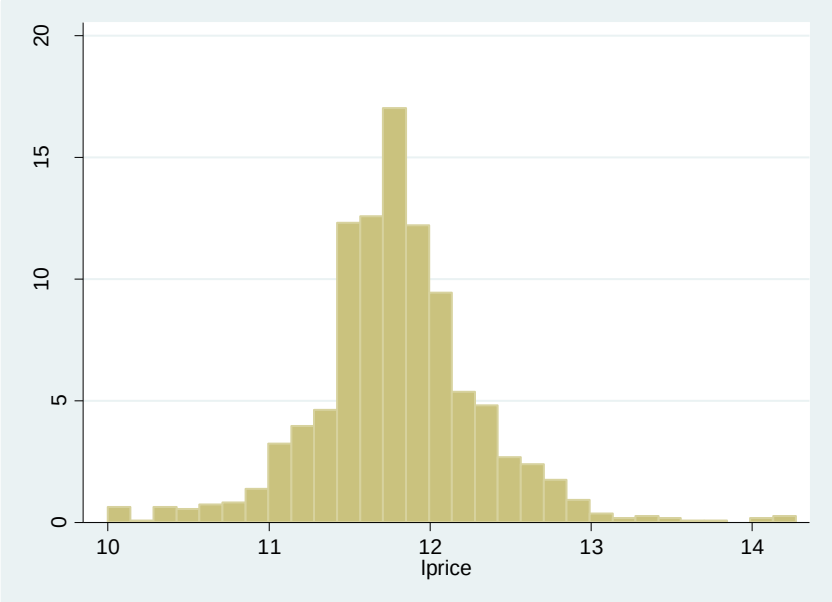
Endi narxni logorifimlangan qiymatini yaratamiz va uning histogrammasini tuzamiz.

generate lprice = ln(price)

histogram lprice, percent

graph save lprice, replace

Quyidagi histogramma simmetrikdir, lekin qo`ng`iroq shakilday emas ya`ni normal taqsimlanmagan.



Log-linear regressiya modeli

reg lprice sqft

```
. reg lprice sqft
```

Source	SS	df	MS
Model	185.472091	1	185.472091
Residual	111.400275	1078	.103339773
Total	296.872366	1079	.275136577

Number of obs	= 1080
F(1, 1078)	= 1794.78
Prob > F	= 0.0000
R-squared	= 0.6248
Adj R-squared	= 0.6244
Root MSE	= .32147

lprice	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sqft	.0004113	9.71e-06	42.36	0.000	.0003922	.0004303
_cons	10.8386	.0246075	440.46	0.000	10.79031	10.88688

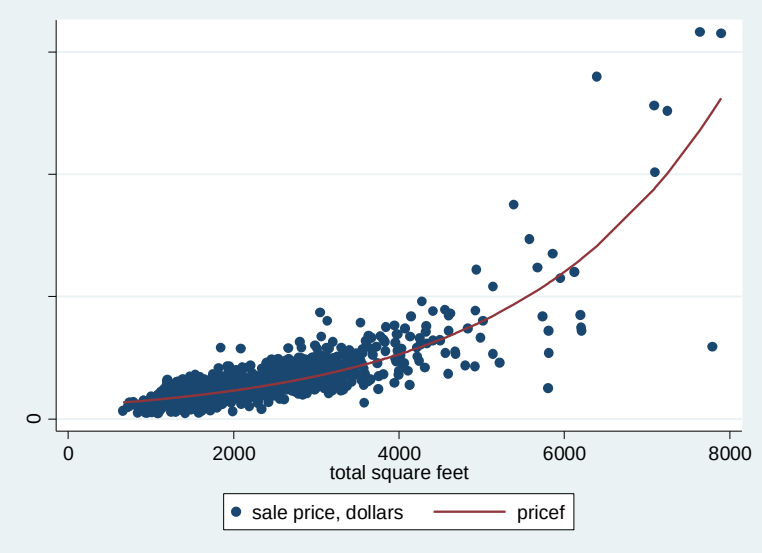
Hisoblangan qiymatlarni topamiz.

predict lpricef, xb

generate pricef = exp(lpricef)

O`zgaruvchi **pricef**- hisoblangan (yoki prognoz qilingan) narxdir. Hisoblangan egri chiziq grafigi.

```
reg lprice sqft
predict lpricef, xb
generate pricef = exp(lpricef)
twoway (scatter price sqft) ///
    (line pricef sqft, sort lwidth(medthick))
graph save br_loglin, replace
```



Hisoblangan egri chiziq chiziqli munosabat emas chunki nurlar va elastiklik aniq nuqtalarda hisoblanishi kerak. Shuni ta`kidlash kerakki, bu yerda qiymat

mavjud emas **margins** hisoblari amalga oshirilmaydi, chunki Stata logaritmik transformatsiyalangan ta`sir etuvchi o`zgaruvchini o`qiy olmaydi. Shu sababdan biz nuqtalarni va elastiklikni aniq tarzda hisoblashimiz kerak. E`tibor beraylikki, nuqtalar bir uyning kattaligi bo`yicha prognoz baholar bilan emas, balki ma`lum bir uy narxini hisobga olgan holda hisoblab chiqilgan.

di "slope at 100000 = " _b[sqft]*100000

di "slope at 500000 = " _b[sqft]*500000

di "elasticity at 2000 = " _b[sqft]*2000

di "elasticity at 4000 = " _b[sqft]*4000

```
. di "slope at 100000 = " _b[sqft]*100000
slope at 100000 = 41.126885

.
. di "slope at 500000 = " _b[sqft]*500000
slope at 500000 = 205.63442

.
. di "elasticity at 2000 = " _b[sqft]*2000
elasticity at 2000 = .82253769

.
. di "elasticity at 4000 = " _b[sqft]*4000
elasticity at 4000 = 1.6450754
```

Shuningdek, hisoblangan har bir uyning narxi bo`yicha o`rtacha marginal ta`sirlarni hisoblashimiz mumkin.

generate me = _b[sqft]*pricef

summarize me

```
. generate me = _b[sqft]*pricef

.
. summarize me
```

Variable	Obs	Mean	Std. Dev.	Min	Max
me	1080	61.00072	42.91725	27.51118	539.2198

Xuddi shunday o`rtacha elastiklik ham topa olamiz.

generate elas = _b[sqft]*sqft

summarize elas

```
. generate elas = _b[sqft]*sqft

.
. summarize elas
```

Variable	Obs	Mean	Std. Dev.	Min	Max
elas	1080	.9565858	.4145993	.27226	3.24779

Ushbu ishchi faylini yopishingiz mumkin.

log close

2.7 REGRESSION JARAYONDA KO`RSATKICH O`ZGARUVCHILARNI ISHLATISH

Ko`rsatkichli o`zgaruvchilar odatda ikki tomonlama 0-1 o`zgaruvchilardir. Ular ko`chmas mulk modelidagi joy kabi sifat omillarini ko`rsatish uchun regressiyada foydalanish mumkin. Yangi ishchi faylni va ma`lumotlar **utown.dta** faylini oching. Ma`lumotlarni **describe** va **summarize** qiling.

log using chap02_indicator, replace text

use utown, clear

describe

summarize

Ma`lumotlar shunu tasvirlaydiki, **utown** o`zgaruvchisi 1 soni bilan agarda uy universitetga yaqin bo`lsa, va boshqa hollarda 0 sono bilan ifodalangan.

variable name	storage type	display format	value label	variable label
price	double	%10.0g		house price, in $1000
sqft	double	%10.0g		square feet of living area, in 100s
age	byte	%8.0g		house age, in years
utown	byte	%8.0g		=1 if close to university
pool	byte	%8.0g		=1 if house has pool
fplace	byte	%8.0g		=1 if house has fireplace

Xulosa statistikasi shuni ko`rsatadiki, 1000ta kuzatuv, taxminan 52% i universitetga yaqin mahallada joylashgan.

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	1000	247.6557	42.19273	134.316	345.197
sqft	1000	25.20965	2.91848	20.03	30
age	1000	9.392	9.426728	0	60
utown	1000	.519	.4998889	0	1
pool	1000	.204	.4031706	0	1
fplace	1000	.518	.4999259	0	1

Umumiy grafikda umumiy bahoga ega bo`lgan uy narxlarining histogrammasini yaratish ularni osonlik bilan taqqoslash imkonini beradi. Avval $ 12000 [**width(12)**] kengligdagi va $ 130000 dan boshlanadigan [**start (130)**], misol minimalidan past bo`lgan qutilari yordamida ikki alohida histogramma yarating va saqlang. $ 12000 miqyosdagi yirik konteynerlardan foydalanish sinov va xatolarga asoslangan edi. **xlabel** tasdiqlash indikatori 130 dan 350 gacha 24 kenglikda bo`lgan ustunlar orqali ifodalanishi kerakligini ko`rsatib turibdi. **Histogram** buyrug`i mantiqiy operatoridan foydalananing, masalan, **if utown == 0** bo`lsa, **utown** o`zgaruvchi tomonidan ikkita belgilangan ma`lumotlarni tanlang.

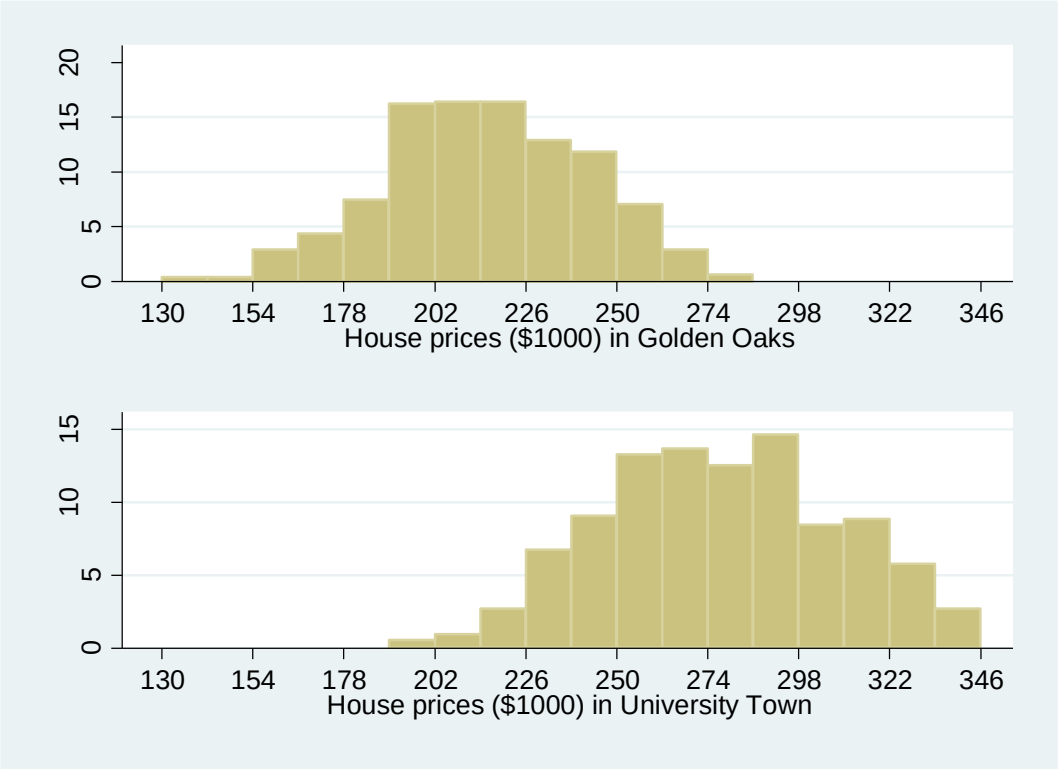
```
histogram price if utown==0, width(12) start(130) percent ///
   xtitle(House prices ($1000) in Golden Oaks) ///
            xlabel(130(24)350) legend(off)
graph save utown_0, replace

histogram price if utown==1, width(12) start(130) percent ///
   xtitle(House prices ($1000) in University Town) ///
            xlabel(130(24)350) legend(off)
graph save utown_1, replace
```

Saqlab olingan grafikalarni **graph combine** buyrug`i yordamida birlashtiriladi. Ularni bir-biriga biriktirish uchun ularni bitta ustunga **[col(1)]** ga kiritishni tanlaymiz. **help graph combine** buyrug`idan foydalanib, **"... iscale (1) da matn va belgilar aslida bir xil o`lchamlarda bo`lishi kerakligini anglatadi" deb izoh qoldirilgan**. Grafik nomlari ularni *** .gph** fayllari sifatida identifikatsiya

qilish uchun kodirovkalarga qo`yiladi.

graph combine "utown_0" "utown_1", col(1) iscale(1)

graph save combined, replace



Grafika shuni ko`rsatadiki, universitet mahallasidagi narxlar boshqa mahallalardagi uylarga qaraganda yuqori qiymatga ega.

Ikki grafika yaratib ularni birlashtirmasdan, Stata ning **by** buyrug`ini ma`lumotlar to`plamlari orqali takrorlaydigan juda muhim xususiyatdan foydalanishimiz mumkin. Undan foydalanishda izohlardan foydalanish muhim ahamiyatga ega. Keyingi ikkita bayonnomada avval biz "Golden Oaks" uchun 0, "university town" uchun "1" deb nomlangan izohli belgilarni yaratamiz. Ikkinchidan, biz izohni o`zgaruvchi **UTOWN** ga qo`llaymiz.

label define utownlabel 0 "Golden Oaks" 1 "University Town"

label value utown utownlabel

Histogramma ma`lumotlarning pastki qismlari **UTOWN** tomonidan aniqlanganligini bildiruvchi parametrdan foydalaning.

histogram price, by(utown, cols(1)) ///

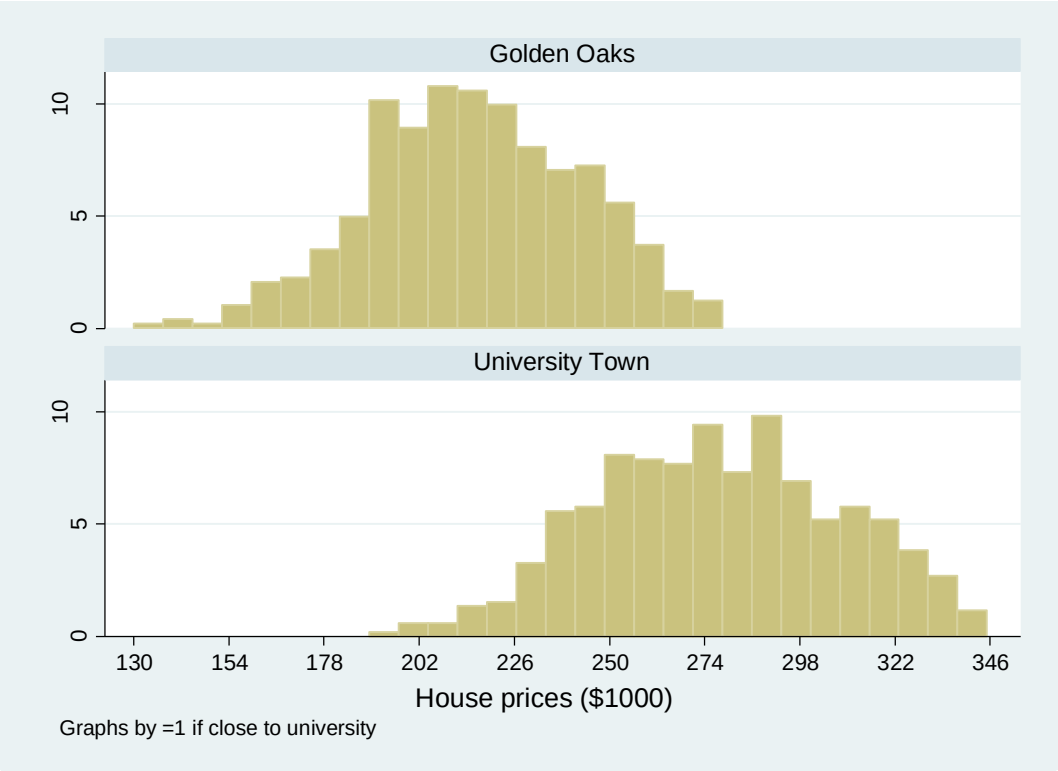
start(130) percent ///

xtitle(House prices ($1000)) ///

xlabel(130(24)350) legend(off)

graph save combined2, replace

Olingan grafik narx uchun bir o`lchov va **by** opsiyaga ko`ra izoh ko`rsatilgan.



Bundan tashqari, alohida mahallalarda narxlar bahosi statistikasini ham o`rganish mumkin.

summarize price if utown==0

summarize price if utown==1

```
. summarize price if utown==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	481	215.7325	26.73736	134.316	276.977

```
.
. summarize price if utown==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	519	277.2416	30.78208	191.57	345.197

Shunga qaramay, **by** optsiyasidan foydalanish oson va samaraliroq. Buyruq

bo`yicha ishlatilayotganda ma`lumotlarni kuzatishning pastki qismlarini belgilaydigan o`zgarmaydigan qiymatiga qarab tartiblash kerak. Tartiblash quyidagi buyruqlar yordamida amalga oshirilishi mumkin.

by utown, sort: summarize price

```
. by utown, sort: summarize price

-------------------------------------------------------------------------------
-> utown = Golden Oaks

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |        481    215.7325    26.73736     134.316    276.977

-------------------------------------------------------------------------------
-> utown = University Town

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |        519    277.2416    30.78208      191.57    345.197
```

Yuqoridagidan farqli o`laroq, buyruqni qisqaroq qilish ham mumkin.

bysort utown: summarize price

```
. bysort utown: summarize price

-------------------------------------------------------------------------------
-> utown = Golden Oaks

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |        481    215.7325    26.73736     134.316    276.977

-------------------------------------------------------------------------------
-> utown = University Town

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |        519    277.2416    30.78208      191.57    345.197
```

Taqqoslama o`zgaruvchisi sifatida indikator o`zgaruvchisi bilan regressiya odatdagidek bir xil sintaksisga ega, lekin biz **i.utown** buyrug`idan foydalangan bo`lardik.

regress price utown

```
. regress price utown
```

Source	SS	df	MS
Model	944476.744	1	944476.744
Residual	833969.397	998	835.640678
Total	1778446.14	999	1780.22637

Number of obs	= 1000
F(1, 998)	= 1130.24
Prob > F	= 0.0000
R-squared	= 0.5311
Adj R-squared	= 0.5306
Root MSE	= 28.907

price	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
utown	61.50911	1.829589	33.62	0.000	57.91882	65.09939
_cons	215.7325	1.318066	163.67	0.000	213.146	218.319

E`tibor bering, taxmin qilingan o`zgarmas atama o`rtacha narx bo`lsa, **utown = 0** ga teng bo`ladi va **utown** ning koeffitsienti namuna vositalarining orasidagi farqdir.

Ikkala o`zgaruvchini statistika kursida teng yoki tengmasligini tekshirishingiz mumkin. Stata bu tekshirishni amalga oshirish uchun **ttest** buyrug`ini o`z ichiga oladi.

ttest price, by(utown)

Statistik natijalarni olingandan so`ng ularni buyruqlarini do-file ga saqlab ishchi faylni yopishingiz mumkin.

log close

III-BOB. INTERVALLI BAHOLASH VA GIPOTEZA TESTI

Intervalli baholashlar ishonch oralig'i sifatida ham tanilgan. Stata regressiya tahlilini amalga oshirganda, uning standart chiqishi ya`ni bir qismi har bir koeffitsient uchun intervalning 95% ni baholab ko`rsatishidir. Yangi Stata fayilini yarating va ish katalogiga o'ting. Oldingi bobning 2.4 bo'limida ko'rsatilgandek, jurnal faylini oching va oziq-ovqat xarajatlari modelini baholang.

log using chap03, replace text

use "D:\programms\Stata12\STATA\food.dta", clear

reg food_exp income

Regressiya tahlilida **[95% Conf. Interval]** ni o'z ichiga oladi, tegishli koeffitsientlar uchun intervalli hisob-kitoblarning pastki va yuqori chegaralari berilgan.

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

Intervalli baholashlar **Coef. ± t-critical*Std. Err.** sifatida hisoblanadi. Koeffitsientlar qiymatlari va standart xatolar berilgan. Qolgan tarkibiy qismi ya`ni **t**-muhim ahamiyatga ega. Buni *Principles of Econometrics* kitobining 2-jadvalida yoki biz hozirda ko'rsatganimizdek, Stata yordamida topish ham mumkin.

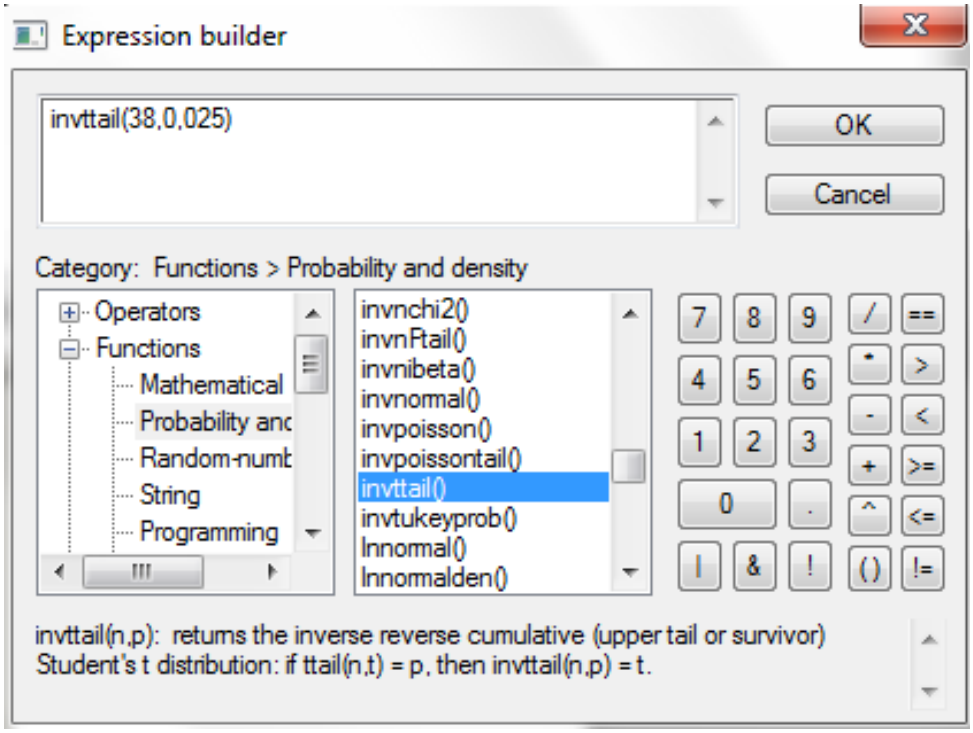
3.1.1 T-TAQSIMOTNING KRITIK QIYMATLARI

Ko'pgina ehtimolli taqsimotlarning kritik qiymatlarini hisoblash uchun biz Stata-dan foydalanishimiz mumkin, bu ko'plab kontekstlarda juda qulaydir. Kritik qiymatlar Stata da skalyar shaklida yaratilgan va ular "inverse" funksiyalarni ko'rsatuvchi umumiy prefiksi **inv** buyrug`ini o'z ichiga oladi. Muayyan skalyar qiymatni chaqirish uchun quyidagi buyruqni kiriting:

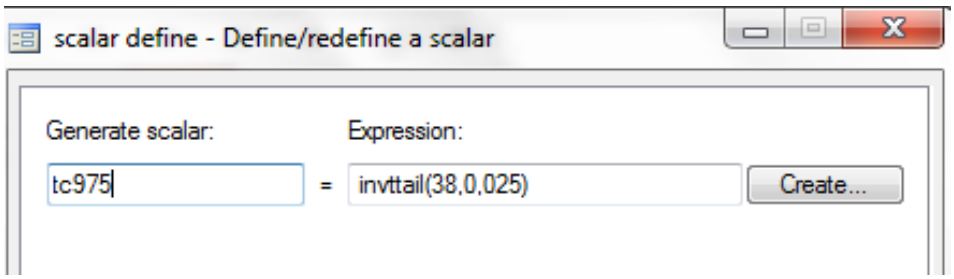
help scalar

Agar dialog oynasidan foydalanmoqchi bo'lsangiz, **Viewer** oynasidan **Dialog**

menu sidagi **define** tugmasini bosing. Shunda **scalar define** oynasi paydo bo`ladi, so`ngra **Create** tugmasini bosing. **Expression builder** oynasidan **Category** qismidan **functions** ni oching va uning ichidan **probabilitiy and density** dan foydalanib **invttail()** ni, ikki marta bosing va N-2 = 38 erkinlik darajalari va taqsimotning yuqori qismidagi ehtimollik miqdorini to'ldiring. 95% oraliq baholash: yuqoridagi buyruqlardan so`ng keyingi oynada ehtimolning 2,5% t-taqsimotining 97,5 foizini belgilaydi. So`ngra OK tugmasini bosing.



So`ngra bizda mavjud bo'lgan skalyar belgilash oynasi paydo bo`ladi:



Ushbu oynaning pastki qismidagi **"Submit"** ni bosing. Natijalar oynasida (va ko'rib chiqish oynasida) Stata buyrug'i quyidagicha ko'rsatilgan bo`ladi:

scalar define tc975 = invttail(38,0.025)

Ixtiyoriy aniqlash shart emas, shuning uchun buyruqni soddalashtirish mumkin

scalar tc975 = invttail(38,0.025)

Ushbu skalarning qiymatini ko'rish uchun biz uni natijasini ko'rsatishimiz uchun display buyrug`idan foydalanamiz.

di "t critical value 97.5 percentile = " tc975

quyidagicha natijaga erishamiz:

```
. di "t critical value 97.5 percentile = "  tc975
t critical value 97.5 percentile = 2.0243942
```

t-kritik qiymatlarning boshqa misollari:

di "t(30) 95th percentile = " invttail(30,0.05)

```
. di "t(30) 95th percentile = " invttail(30,0.05)
t(30) 95th percentile = 1.6972609
```

di "t(20) 5th percentile = " invttail(20,0.95)

```
. di "t(20) 5th percentile  = " invttail(20,0.95)
t(20) 5th percentile  = -1.7247182
```

di "t(30) 2.5th percentile = " invttail(30,0.975)

```
. di "t(30) 2.5th percentile  = " invttail(30,0.975)
t(30) 2.5th percentile  = -2.0422725
```

Oraliq interval tuzish

Regressiya natijalari va biz hozirgina 95% intervalgacha hisoblagan t-kritik qiymatdan kalkulyator yordamida ham hisoblash mumkin. Interval hisobini olish uchun Stata xotirasida saqlangan natijalardan ham foydalanishingiz mumkin. Yuqorida ta'kidlab o'tilgandek, agar regressiya aniqlansa, ma'lum natijalar saqlanib qoladi va keyinchalik ulardan foydalanish mumkin bo'ladi. Hisoblangan koeffitsientlar va standart xatolar navbati bilan **_b [varname]** va **_se [varname]** sifatida saqlanadi. Oziq-ovqat xarajatlari tenglamasini hisoblab chiqqandan so'ng, nuqta koeffitsienti **_b [income]** deb nomlanadi va taxmin qilinadigan nuqta **_b [_cons]**. Ularning standart xatolari **_se [income]** va **_se [_cons]** dir. Qo'shimcha ma'lumot olish uchun buyruqlar oynasida **help _variables** buyrug`ini kiriting. 95% oraliq hisob-kitoblar $b_k \pm t_c se(b_k)$ formula orqali hisoblanadi. Nuqta uchun

intervalni hisoblashning yuqori chegarasi va pastki chegarasi quyidagicha topiladi:

scalar ub2 = _b[income] + tc975*_se[income]

scalar lb2 = _b[income] - tc975*_se[income]

Bularning natijalarini quyidagi buyruq orqali namoyish etilishi mumkin.

di "beta 2 95% interval estimate is " lb2 " , " ub2

va buning natijasi:

```
. di "beta 2 95% interval estimate is " lb2 " , " ub2
beta 2 95% interval estimate is 5.9720525 , 14.447233
```

3.2 GIPOTEZA TESTLARI

Parametrlar to'g'risidagi gipoteza testlari uchun ishlatiladigan t-statistikani regressiya chiqishidagi kalkulyator va statistik jadvalning t-kritik qiymati yordamida hisoblash mumkin. Ammo ushbu bo'limda biz Stata yordamida testning statistik qiymatlari, tanqidiy qiymatlari va p-qiymatlarini hisoblab chiqamiz. Misol sifatida biz oziq-ovqat xarajatlari regressiyasining modelini davom ettiramiz.

$H_0: \beta_2 = 0$ nol gipotezani muqobil $H_0: \beta_2 > 0$ gipotezaga qarshi tekshiramiz. t-statik qiymatni va kritik qiymatni ishlatib qurishimiz va namoyish etishimiz mumkin

scalar tstat0 = _b[income]/_se[income]

di "t statistic for Ho: beta2=0 = " tstat0

di "t(38) 95th percentile = " invttail(38,0.05)

```
. scalar tstat0 = _b[income]/_se[income]

.
. di "t statistic for Ho: beta2=0 = " tstat0
t statistic for Ho: beta2=0 = 4.8773806

.
. di "t(38) 95th percentile = " invttail(38,0.05)
t(38) 95th percentile = 1.6859545
```

E'tibor bering, kritik qiymat t-taqsimotning o'ng tomonida keladi va biz

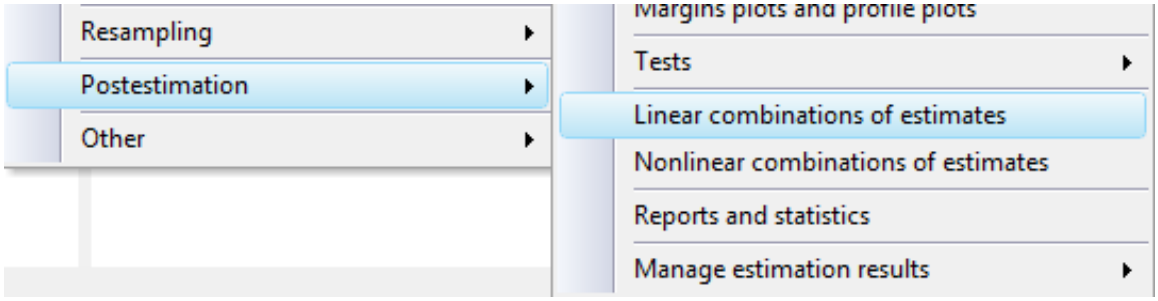
kritik qiymatni topish uchun **invttail** buyrug'idan foydalanamiz. Koeffitsientlar nolga teng bo`lganda gipotezaning t-statistik qiymatlari Stata tomonidan avtomatik ravishda ishlab chiqariladi, ya`ni regressiya modelida **"t"** deb ko'rsatilgan ustunda hisoblanganda.

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

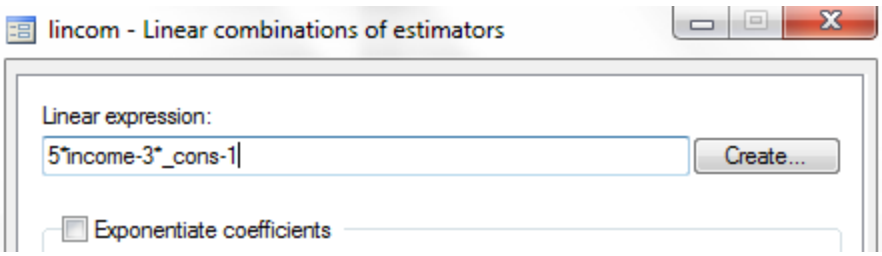
Koeffitsientlar nolga teng bo`lganda gipotezaning t-statistik qiymatlari

Ba'zida koeffitsientlar haqida yanada murakkab farazlarni sinab ko'rishni istaymiz va ularni baholashda keyingi buyruq bu **lincom** yordamida amalga oshirish mumkin. Stata ochiladigan menyusidan ko`rsatilgan ketma-ketlikni tanlang:

Statistics > Postestimation > Linear combinations of estimates.



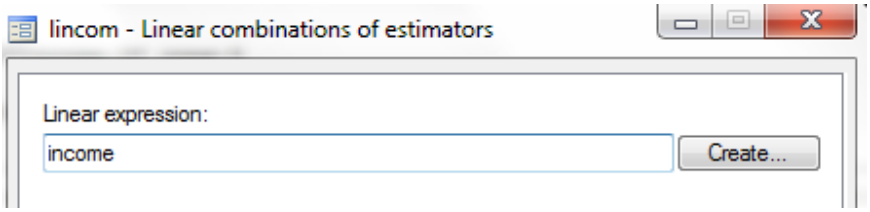
Muloqot oynasida siz taxmin qilmoqchi bo'lgan gipotezani kiriting. Bu esa $5b_2 - 3b_1 = 1$ kabi b_1 va b_2 koeffitsientlarini o'z ichiga olgan chiziqli kombinatsiya deb ataladigan har qanday chiziqli ifoda bo'lishi ham mumkin. Albatta, Stata b_1 va b_2 ga to'g'ridan-to'g'ri murojaat qilmaydi, aksincha **_b [income]** va **_b [_cons]** orqali. Agar buyruq **lincom** bo'lsa, **5 * income - 3 * _cons-1** kabi o'zgaruvchan nomlarga murojaat qilishimiz kerak. Buyruq ifoda qiymatini va uning standart xatosini hisoblab chiqadi va t-statistik va interval bahosini hisoblab ko`rsatadi.




```
. lincom 5*income-3*_cons-1

 ( 1)  5*income - 3*_cons = 1
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-200.1998	140.1664	-1.43	0.161	-483.9518	83.55225

Boshqa misol sifatida shunchaki kiriting



OK tugmasini bosing. So`ngra Natijalar oynasida biz kiritgan Stata buyrug'i va regressiya qiymatlariga o`xshash quyidagi natija paydo bo`ladi.

```
. lincom income

 ( 1)  income = 0
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	10.20964	2.093264	4.88	0.000	5.972052	14.44723

3.2.1 IQTISODIY GIPOTEZANING RIGHT-TAIL TESTI

$H_0: \beta_2 \leq 5.5$ muqobil gipotezaga qarshi $H_1: \beta_2 > 5.5$ sinovini o'tkazish uchun biz yana test statistikasini va 0.01 o'ng tomonning kritik qiymatini hisoblab chiqamiz.

scalar tstat1 = (_b[income]-5.5)/_se[income]

E'tibor bering, biz ish tartibini boshqarish uchun qavslardan foydalanganmiz.

di "t-statistic for Ho: beta2 = 5.5 is " tstat1

di "t(38) 99th percentile = " invttail(38,0.01)

bu buyruqlar quyidagi natijalarni hisoblab ko`rsatadi

```
. di "t-statistic for Ho: beta2 = 5.5 is " tstat1
t-statistic for Ho: beta2 = 5.5 is 2.2499045

. di "t(38) 99th percentile = " invttail(38,0.01)
t(38) 99th percentile = 2.4285676
```

So`ngra **lincom** buyrug`idan foydalaning

lincom income-5.5

Natija shuni ko'rsatadiki, b_2-5.5 qiymati hisoblangan **[Coef.]** va t-statistik **[t]** hisoblab chiqilgan, shuningdek 95% interval bilan hisoblangan.

```
. lincom income-5.5
 ( 1)  income = 5.5
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	4.709643	2.093264	2.25	0.030	.4720525	8.947233

-5.5 dagi qiymati

t-statistik qiymati

3.2.2 IQTISODIY GIPOTEZANI LEFT-TAIL TESTI

Left-tail testini tekshirish uchun $H_1: \beta_2 < 15$ muqobil faraz bilan $H_0: \beta_2 \geq 15$ nol gipoteza bilan hisoblab chiqamiz. Buyruqlar ketma-ketligi quyidagicha:

scalar tstat2 = (_b[income]-15)/_se[income]

di "t-statistic for Ho: beta2 = 15 is " tstat2

0.05 kritik qiymatini hisoblash uchun yana **invttail** buyrug`idan foydalanamiz. Left-tail testini tekshirib ko`rishimizga qaramay, bizga bir vaqtning o`zida right-tail qismidagi 95% li ehtimollik kerak bo`ladi. Shu asnoda quyidagi buyruqga yuzlanamiz:

di "t(38) 5th percentile = " invttail(38,0.95)

quyidagi natijaga erishamiz:

```
.
. di "t-statistic for Ho: beta2 = 15 is " tstat2
t-statistic for Ho: beta2 = 15 is -2.2884634

. di "t(38) 5th percentile = " invttail(38,0.95)
t(38) 5th percentile = -1.6859545
```

So`ngra **lincom** buyrug`idan foydalanamiz:

lincom income-15

```
. lincom income-15

 ( 1)  income = 15
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-4.790357	2.093264	-2.29	0.028	-9.027948	-.5527666

3.2.3 IQTISODIY GIPOTEZANI IKKI TOMONLAMA TEKSHIRISH

Ikki tomonlama sinov kritik qiymatlarni hisoblashdan tashqari bitta tomon sinov bilan bir xil natijaga ega. 0,05 ahamiyatlilik darajasidagi sinov uchun t-taqsimotning 2,5 va 97,5 foizi bo'lishi kerak. $H_0: \beta_2 = 7.5$ muqobil gipotezaga qarshi $H_1: \beta_2 \neq 7.5$ sinovini o'tkazish uchun quyidagi buyruqlardan foydalaning

scalar tstat3 = (_b[income]-7.5)/_se[income]

di "t-statistic for Ho: beta2 = 7.5 is " tstat3

di "t(38) 97.5th percentile = " invttail(38,0.025)

di "t(38) 2.5th percentile = " invttail(38,0.975)

natija quyidagicha

```
. di "t-statistic for Ho: beta2 = 7.5 is " tstat3
t-statistic for Ho: beta2 = 7.5 is 1.2944586

.
. di "t(38) 97.5th percentile = " invttail(38,0.025)
t(38) 97.5th percentile = 2.0243942

.
. di "t(38) 2.5th percentile = " invttail(38,0.975)
t(38) 2.5th percentile = -2.0243942
```

Stata **invttail** buyrug'i t-taqsimotning yuqori qismi bilan ishlaganligi sababli, pastki qismi uchun kritik qiymatlarni hisoblash chalkash bo'lishi mumkin. Lekin har doim funksiyani ishlashni davom ettiring. T-taqsimot nosimmetrik va 90, 95, 97.5 va 99-foizlar ijobiy qiymatlar, 1, 2,5, 5 va 10-foizlar esa salbiydir. **Lincom** buyrug`idan foydalaning:

lincom income-7.5

```
. lincom income-7.5
( 1)  income = 7.5
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	2.709643	2.093264	1.29	0.203	-1.527948	6.947233

3.3 p-QIYMATLAR (p-values)

p-qiymatlarni osonlikcha hisoblash qobiliyati Stataning kuchli xususiyatidir. Eslatib o'tamiz

- agar $H_1: \beta_k > c$, p = t ning o`ng tomon ehtimolligi
- agar $H_1: \beta_k < c$, p = t ning chap tomon ehtimolligi
- agar $H_1: \beta_k \neq c$, p = o`ng tomon |t| va chap tomon -|t| ehtimolligining summasi

Gipotezalarni sinash uchun p-qiymat qoidasi: Nol gipotenuzani rad etiladi qachonki p-qiymat past yoki unga teng bo'lganda ya`ni α darajadagi qiymatga. Ya'ni, agar p≤α bo'lsa, H_0 ni rad etadi. Agar p>α bo'lsa, H_0 ni rad qilinmaydi.

T-taqsimotining kritik qiymatlarini jadvalda ko'rish yoki **invttail** funktsiyasi yordamida hisoblash mumkin. Ammo kompyuter yordamida p-qiymatlarni hisoblash kerak. Stata **ttail** funktsiyasidan foydalanadi. Buyruqning sintaksisi va uning ta'rifi buyruqlar oynasida **help ttail** buyrug`ini kiritish orqali ko`rish mumkin.

ttail(n,t) “returns the reverse cumulative (upper-tail) Student's t

distribution; it returns the probability T > t.”

bu erda **n** - erkinlik darajalarining soni va **t** - t-statistikaning qiymati. **ttail** funktsiyasi yana bir marta yuqori darajadagi ehtimollik qiymatini hisoblab ko`rsatadi.

3.3.1 p-QIYMATNING O'NG TOMON QIYMATNI TEKSHIRISH

Yuqoridagi 3.2.2 qismida biz $H_0: \beta_2 \leq 5.5$ nol gipotezaga qarshi muqobil gipoteza $H_1: \beta_2 > 5.5$ sinov bajargandik. T-statistik qiymatdan foydalanib hisoblagandik:

scalar tstat1 = (_b[income]-5.5)/_se[income]

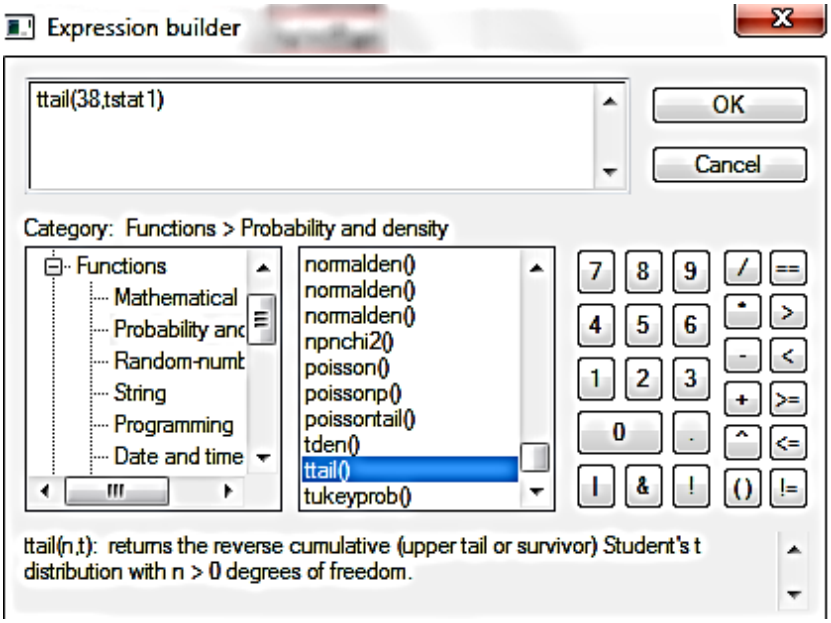
p-qiymatidan foydalanib hisoblaymiz va ko`rsatib o`tamiz:

di "p value right tail test ho:beta2 = 5.5 is " ttail(38,tstat1)

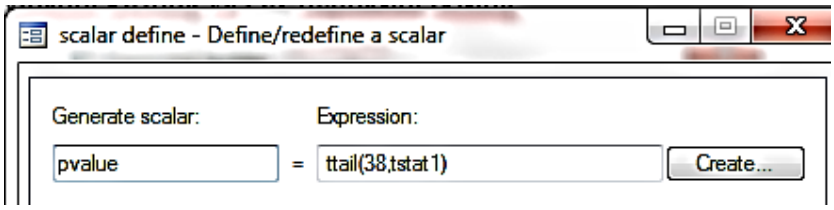
Natija esa quyidagicha:

```
. di "p value right tail test ho:beta2 = 5.5 is " ttail(38,tstat1)
p value right tail test ho:beta2 = 5.5 is .01516329
```

Eslatib o'tamiz, agar siz sintaksisni unutib qo'ysangiz, **help scalar** buyrug'ini kiritib, undan **Dialog** menyusinidan **scalar define** dialog oynasini ochishinggiz mumkin. **Create** tugmasini va keyin **Expression builder** muloqot oynasida **Probability functions** ni tanlang va o`ng tomonga o'ting. Vazifalarning ta'riflari qutining pastki qismida ko'rinadi, bu juda yaxshi yordam beradi. Keyin funktsiya nomini ikki marta bosing va **Expression builder** oynasida **n** erkinlik darajalari va **t** t-statistik qiymatini kiriting va OK tugmasini bosing.



Olingan natija uchun nom kiriting va OK tugmasini bosing



3.3.2 p-QIYMATNING CHAP TOMON QIYMATNI TEKSHIRISH

Chap tomonning sinov qiymatini ko'rsatish uchun yuqoridagi 3.2.3 bo'limidagi misoldan foydalaning. $H_0: \beta_2 \geq 15$ nol gipoteza $H_1: \beta_2 < 15$ muqobil faraz bilan bo'lsin! . Hisoblaymiz:

scalar tstat2 = (_b[income]-15)/_se[income]

Chap tomon test p-qiymati t-taqsimotning chap yoki pastki qismida joylashgan. Buyruqni kiriting:

di "p value left tail test ho:beta2 = 15 is " 1-ttail(38,tstat2)

Biz **1-ttail(38,tstat2)** dan foydalanishimiz kerak, chunki tstat2 ning chap tomoning hududini topishni xohlaymiz qaysiki **ttail (38, tstat2)** tstat2 ning o'ng tomonini hisoblab chiqadi. Natijada

```
. di "p-value for left-tail test ho:beta2 = 15 is " 1-ttail(38,tstat2)
p-value for left-tail test ho:beta2 = 15 is .01388071
```

3.3.3 p-QIYMATNING IKKI TOMONLI TESTINI TEKSHIRISH

Yuqoridagi 3.2.4-bo'limda biz $H_1: \beta_2 \neq 7.5$ muqobil gipotezaga qarshi nol gipoteza $H_0: \beta_2 = 7.5$ ga tekshirgan edik. t-statistik dan foydalanib hisoblab chiqilgan

scalar tstat3 = (_b[income]-7.5)/_se[income]

Ikki tomonli test uchun p-qiymati t ning o'ng va chap tomonidagi maydonlar yig'indisidir. Quyidagidan foydalanamiz:

scalar phalf = ttail(38,abs(tstat3))

Ushbu buyruq t-statistikaning mutlaq qiymati (abs funktsiyasi) o'ng tomonidagi t-taqsimotning yuqori qismidagi qismini ya`ni p-qiymatini yarmini hisoblaydi. Shu sababdan ushbu qiymatni 2 ga ko'paytiring va natijani ko`ring.

scalar p3 = 2*phalf

di "p value for two tail test ho:beta2 = 7.5 is " p3

Natija esa:

```
.
. di "p value for two tail test ho:beta2 = 7.5 is " p3
p value for two tail test ho:beta2 = 7.5 is .20331828
```

Albatta, alohida hisob-kitoblar talab qilinmaydi. Shuning uchun bitta buyruq sifatida kiritib topsa ham bo`ladi.

di "p value for ho:beta2 = 7.5 is " 2*ttail(38,abs(tstat3))

```
. di "p value for ho:beta2 = 7.5 is " 2*ttail(38,abs(tstat3))
p value for ho:beta2 = 7.5 is .20331828
```

3.3.4 STATA TOMONIDAN p-QIYMATLARNI TOPISH

Qachonki regressiya aniqlansa va baholashdan keyingi buyruq **lincom** dan foydalansa, p-qiymatlari ko`rsatib o`tiladi. Masalan, regressiya natijasi:

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

koeffitsient nol bo`lganda ikki tomonlama testing p-qiymati

P>|t| deb belgilangan ustun bu koeffitsient nolga teng ekanligi haqidagi nol gipoteza uchun ikki tomonli test p-qiymatidir. Belgining o'zi **P** ehtimolligi t-statistik qiymatining |t| qiymatidan kattaroq o'rtacha qiymatiga olinadi. Bu, haqiqatan ham t ning musbat qiymatidan katta va t ning manfiy qiymatidan kichikroq bo'lgan ikkita iboraga aylanadi. Agar ahamiyatlilikning bir tomonli testi talab qilinsa, p-qiymat -ikki tomonli testning p-qiymati, modomiki taxmin alternativ gipoteza qondiradi, chunki u taqsimotning bitta tomonida sodir bo'ladi.

Lincom ishlatilganda ham bir xil natijalar paydo bo`ladi.

```
. lincom income-15
 ( 1)  income = 15
```

gipoteza β_2 -15=0 bo`lgandagi ikki tomonlama p-qiymat

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-4.790357	2.093264	-2.29	0.028	-9.027948	-.5527666

b_2 -15 bo`lgandagi hisoblangan qiymat

b_2 -15 bo`lgandagi standart xatolik

gipoteza β_2 -15=0 bo`lgandagi t-statistik qiymat

β_2 -15=0 ning 95% li interval qiymati

3.3.5 CHIZIQLI FUNKSIYALARNING PARAMETRLARINI TEKSHIRISH VA BAHOLASH

Ko`pincha umumiy chiziqli gipoteza ikkala parametrni ham o'z ichiga oladi va shunday aytilishi mumkin:

$$H_0: c_1\beta_1 + c_2\beta_2 = c_0$$

bu erda c_0, c_1 va c_2 konstantalar ko'rsatilgan. Ushbu gipotezani tekshirishda t-statistika dan foydalaniladi

$$t = \frac{(c_1 b_1 + c_2 b_2) - c_0}{se(c_1 b_1 + c_2 b_2)}$$

Bir va ikki tomonli alternativalarni (i) - (iii) rad etish hududlari 3.3-bo'limda tavsiflanganlar bilan bir xil, xulosalar ham xuddi shunday izohlanadi. t-statistikaning aniqlangan xatosi bu ildizning kvadratidir

$$\widehat{var[c_1 b_1 + c_2 b_2]} = c_1^2 \widehat{var(b_1)} + c_2^2 var(b_2) + 2\widehat{c_1 c_2 cov(b_1, b_2)}$$

Ushbu qiymatni hisoblash uchun **estat vce** yordamida hisob-kitobdan keyingi olingan eng kam kvadratlarning hisoblangan kovoriyalik matritsasidan foydalaning.

```
. estat vce

Covariance matrix of coefficients of regress model
```

e(V)	income	_cons
income	4.3817522	
_cons	-85.903157	1884.4423

Lincom yordamida biz $c_1\beta_1 + c_2\beta_2$ kabi chiziqli funksiyani taxmin qilishimiz va chiziqli farazning umumiy shaklini sinab ko'rishimiz mumkin. Masalan, Agar c_1 = 1 va c_2= 20 bo'lsa, u holda

lincom _cons + income*20

```
. lincom _cons + income*20

 ( 1)  20*income + _cons = 0
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	287.6089	14.17804	20.29	0.000	258.9069	316.3108

Ushbu chiziqli funksiya 250 ga teng bo`ladi nol gipotezani tekshirganda va undan foydalaning

lincom _cons + income*20 - 250

```
. lincom _cons + income*20 - 250

 ( 1)  20*income + _cons = 250
```

food_exp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	37.60886	14.17804	2.65	0.012	8.906915	66.31081

Mustaqil yechish uchun masalalar.

1-Masala.

	Yalpi ichki mahsulot (y)	Tovar va xizmatlar eksporti (x1)
2000	3.256	0.864
2001	4.925	1.517
2002	7.450	2.353
2003	9.844	3.667
2004	12.261	4.930
2005	15.923	6.028
2006	21.125	7.820
2007	28.190	11.372
2008	38.970	15.144
2009	49.376	17.276
2010	62.388	20.668
2011	78.764	25.717
2012	97.929	25.717
2013	120.862	30.100
2014	145.846	31.414
2015	171.808	32.225
2016	199.993	36.077
2017	254.043	73.392

Yuqoridagi jadval ma`lumotidan foydalanib quyidagilarni toping.

- Qisqacha xulosa statistikasi ma`lumotlarini;
- SCATTER diagrammasini yaratish;
- Chiziqli regression model tuzing;
- Hisoblangan yhat ni toping;
- Bashoratli qiymatlarni hisoblang.

2-Masala. Yuqoridagi jadvaldan y va x1 o`zgaruvchilari uchun quyidagilarni toping.

- Kvadrati regression model tuzing;
- Elastiklik qiymatini toping;

- T-taqsimotning kritik qiymatlarini hisoblang;
- Gipoteza testlarini amalga oshiring;
- p-qiymatni toping.

3-Masala.

	Yalpi ichki mahsulot (y)	Asosiy kapitalga investitsiya (x2)
2000	3.256	0.744
2001	4.925	1.321
2002	7.450	1.526
2003	9.844	1.978
2004	12.261	2.629
2005	15.923	3.165
2006	21.125	4.041
2007	28.190	5.903
2008	38.970	9.555
2009	49.376	12.531
2010	62.388	15.338
2011	78.764	17.953
2012	97.929	22.797
2013	120.862	28.694
2014	145.846	35.233
2015	171.808	41.670
2016	199.993	49.771
2017	254.043	68.424

Yuqoridagi jadval ma`lumotidan foydalanib quyidagilarni toping.

- Qisqacha xulosa statistikasi ma`lumotlarini;
- SCATTER diagrammasini yaratish;
- Chiziqli regression model tuzing;
- Hisoblangan yhat ni toping;
- Bashoratli qiymatlarni hisoblang.

4-Masala. Yuqoridagi jadvaldan y va x2 o`zgaruvchilari uchun quyidagilarni

toping.

- Kvadrati regression model tuzing;
- Elastiklik qiymatini toping;
- T-taqsimotning kritik qiymatlarini hisoblang;
- Gipoteza testlarini amalga oshiring;
- p-qiymatni toping.

5-Masala.

	Yalpi ichki mahsulot (y)	Ishlab chiqarish (x3)
2000	3.256	0.462
2001	4.925	0.696
2002	7.450	1.079
2003	9.844	1.553
2004	12.261	2.147
2005	15.923	3.371
2006	21.125	4.597
2007	28.190	5.907
2008	38.970	9.148
2009	49.376	11.651
2010	62.388	15.115
2011	78.764	13.443
2012	97.929	16.148
2013	120.862	20.637
2014	145.846	25.363
2015	171.808	29.370
2016	199.993	33.956
2017	254.043	42.611

Yuqoridagi jadval ma`lumotidan foydalanib quyidagilarni toping.

- Qisqacha xulosa statistikasi ma`lumotlarini;
- SCATTER diagrammasini yaratish;
- Chiziqli regression model tuzing;

- Hisoblangan yhat ni toping;
- Bashoratli qiymatlarni hisoblang.

6-Masala. Yuqoridagi jadvaldan y va x3 o`zgaruvchilari uchun quyidagilarni toping.

- Kvadrati regression model tuzing;
- Elastiklik qiymatini toping;
- T-taqsimotning kritik qiymatlarini hisoblang;
- Gipoteza testlarini amalga oshiring;
- p-qiymatni toping.

7-Masala.

	Yalpi ichki mahsulot (y)	Oziq-ovqat, ichimliklar va tamaki mahsulotlarini ishlab chiqarish(x4)
2000	3.256	0.250
2001	4.925	0.355
2002	7.450	0.644
2003	9.844	0.751
2004	12.261	0.774
2005	15.923	0.909
2006	21.125	1.308
2007	28.190	1.715
2008	38.970	2.554
2009	49.376	3.235
2010	62.388	6.843
2011	78.764	8.981
2012	97.929	10.544
2013	120.862	13.700
2014	145.846	17.177
2015	171.808	21.889
2016	199.993	26.782
2017	254.043	28.332

Yuqoridagi jadval ma`lumotidan foydalanib quyidagilarni toping.

- Qisqacha xulosa statistikasi ma`lumotlarini;
- SCATTER diagrammasini yaratish;
- Chiziqli regression model tuzing;
- Hisoblangan yhat ni toping;
- Bashoratli qiymatlarni hisoblang.

8-Masala. Yuqoridagi jadvaldan y va x4 o`zgaruvchilari uchun quyidagilarni toping.

- Kvadrati regression model tuzing;
- Elastiklik qiymatini toping;
- T-taqsimotning kritik qiymatlarini hisoblang;
- Gipoteza testlarini amalga oshiring;
- p-qiymatni toping.

Foydalanilgan adabiyotlar ro`yxati

1. O`zbekiston Davlat statistika qo`mitasining 2010-2018 yillardagi "O`zbekiston Respublikasi yillik statistik to`plami" . http://www.stat.uz

2. Lee C. Adkins, R. Carter Hill. Using Stata for Principles of Econometrics, Fourth Edition. Wiley, 2011

3. A.Ishnazarov, Sh.Nurullayeva, M.Mo`minova. "Ekonometrika asoslari" fani bo`yicha o`quv-uslubiy majmua. Toshkent: Iqtisodiyot, 2017 yil.

4. *Introduction to econometrics* / James H. Stock, Mark W. Watson. 2nd ed., Boston: Pearson Addison Wesley, 2007.

5. *Data analysis using regression and multilevel/hierarchical models* / Andrew Gelman, Jennifer Hill. Cambridge ; New York : Cambridge University Press, 2007.

6. Gujarati D.N. "Basic Econometrics". McGraw-Hill, 4th edition, 2003 (Gu), Inc.p. 90.

7. William H. Greene. Econometric analysis, seventh edition. USA, 2012.

8. Stata basics (Stata 13 or 14), Ursina Kuhn and Oliver Lipps, FORS, Lausanne. Version August 2016

Foydali internet saytlari

www.Stata.com

http://www.stata.com/links/video-tutorials/

http://www.ats.ucla.edu/stat/stata

http://dss.princeton.edu/training/

http://www.ats.ucla.edu/stat/stata/

http://fmwww.bc.edu/GStat/docs/StataIntro.pdf

http://stata.com/support/faqs/

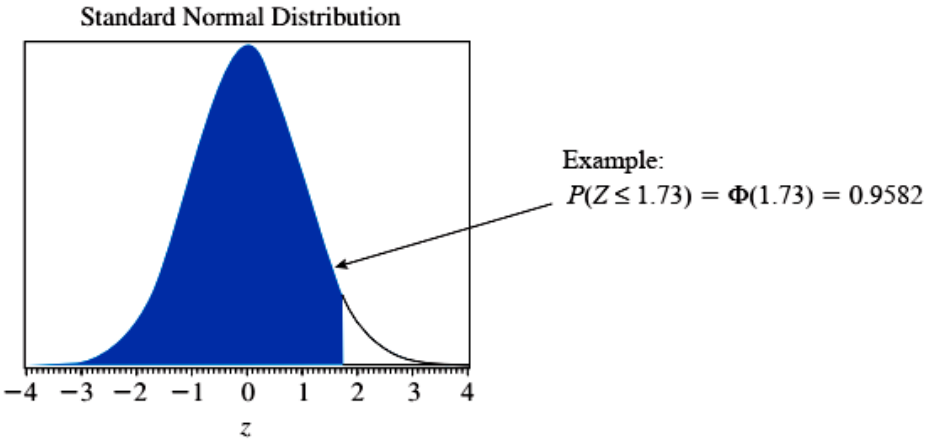


Table 1 **Cumulative Probabilities for the Standard Normal Distribution**
$\Phi(z) = P(Z \le z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Source: This table was generated using the SAS® function PROBNORM.

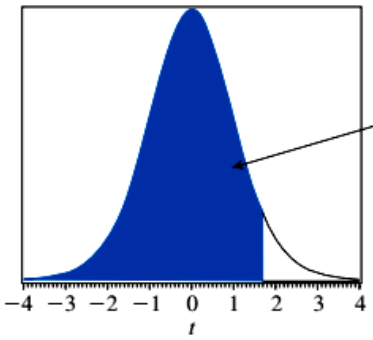


Table 2 **Percentiles of the *t*-distribution**

df	$t_{(0.90,df)}$	$t_{(0.95,df)}$	$t_{(0.975,df)}$	$t_{(0.99,df)}$	$t_{(0.995,df)}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
31	1.309	1.696	2.040	2.453	2.744
32	1.309	1.694	2.037	2.449	2.738
33	1.308	1.692	2.035	2.445	2.733
34	1.307	1.691	2.032	2.441	2.728
35	1.306	1.690	2.030	2.438	2.724
36	1.306	1.688	2.028	2.434	2.719
37	1.305	1.687	2.026	2.431	2.715
38	1.304	1.686	2.024	2.429	2.712
39	1.304	1.685	2.023	2.426	2.708
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
∞	1.282	1.645	1.960	2.326	2.576

Source: This table was generated using the SAS® function TINV.

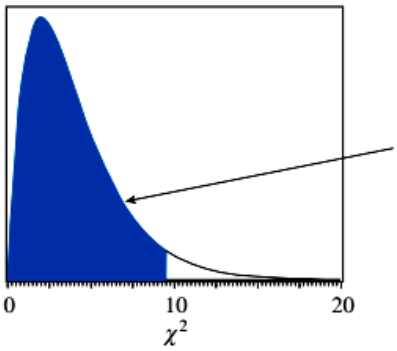


Table 3 Percentiles of the Chi-square Distribution

df	$\chi^2_{(0.90,df)}$	$\chi^2_{(0.95,df)}$	$\chi^2_{(0.975,df)}$	$\chi^2_{(0.99,df)}$	$\chi^2_{(0.995,df)}$
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28.300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.144	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.289	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.652	40.646	44.314	46.928
26	35.563	38.885	41.923	45.642	48.290
27	36.741	40.113	43.195	46.963	49.645
28	37.916	41.337	44.461	48.278	50.993
29	39.087	42.557	45.722	49.588	52.336
30	40.256	43.773	46.979	50.892	53.672
35	46.059	49.802	53.203	57.342	60.275
40	51.805	55.758	59.342	63.691	66.766
50	63.167	67.505	71.420	76.154	79.490
60	74.397	79.082	83.298	88.379	91.952
70	85.527	90.531	95.023	100.425	104.215
80	96.578	101.879	106.629	112.329	116.321
90	107.565	113.145	118.136	124.116	128.299
100	118.498	124.342	129.561	135.807	140.169
110	129.385	135.480	140.917	147.414	151.948
120	140.233	146.567	152.211	158.950	163.648

Source: This table was generated using the SAS® function CINV.

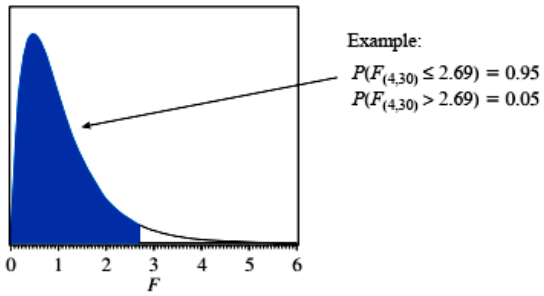


Table 4 95th Percentile for the *F*-distribution

v_2/v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	250.10	252.20	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.48	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	3.01	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.70	2.62	2.54
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.07
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.84
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.92	1.82	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.62
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.04	1.96	1.88	1.79	1.68	1.56
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.74	1.64	1.51
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	1.97	1.89	1.81	1.71	1.60	1.47
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.87	1.78	1.69	1.58	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.65	1.53	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.55	1.43	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.00

Source: This table was generated using the SAS® function FINV.

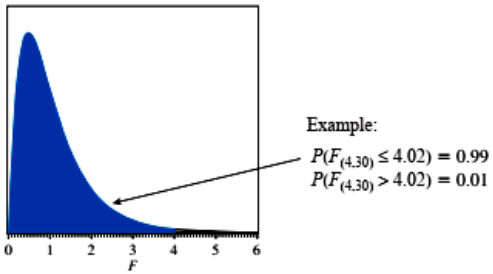


Table 5 99th Percentile for the *F*-distribution

v_2/v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6157.28	6208.73	6260.65	6313.03	6365.87
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.47	99.48	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	z26.87	26.69	26.50	26.32	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.84	13.65	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.38	9.20	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	5.99	5.82	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.20	5.03	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	3.91
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.87
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.42
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.54	2.36	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.01
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88	2.74	2.60	2.44	2.28	2.10	1.89
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.20	2.02	1.80
45	7.23	5.11	4.25	3.77	3.45	3.23	3.07	2.94	2.83	2.74	2.61	2.46	2.31	2.14	1.96	1.74
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.56	2.42	2.27	2.10	1.91	1.68
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.86	1.66	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.00

Source: This table was generated using the SAS® function FINV.

Ilovalar

The Rules of Summation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

$$\sum_{i=1}^{n} a = na$$

$$\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} (ax_i + by_i) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} (a + bx_i) = na + b\sum_{i=1}^{n} x_i$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

$$\sum_{i=1}^{2}\sum_{j=1}^{3} f(x_i, y_j) = \sum_{i=1}^{2} [f(x_i, y_1) + f(x_i, y_2) + f(x_i, y_3)]$$
$$= f(x_1, y_1) + f(x_1, y_2) + f(x_1, y_3)$$
$$+ f(x_2, y_1) + f(x_2, y_2) + f(x_2, y_3)$$

Expected Values & Variances

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n)$$
$$= \sum_{i=1}^{n} x_i f(x_i) = \sum_x x f(x)$$

$$E[g(X)] = \sum_x g(x) f(x)$$

$$E[g_1(X) + g_2(X)] = \sum_x [g_1(x) + g_2(x)] f(x)$$
$$= \sum_x g_1(x) f(x) + \sum_x g_2(x) f(x)$$
$$= E[g_1(X)] + E[g_2(X)]$$

$E(c) = c$
$E(cX) = cE(X)$
$E(a + cX) = a + cE(X)$
$\text{var}(X) = \sigma^2 = E[X - E(X)]^2 = E(X^2) - [E(X)]^2$
$\text{var}(a + cX) = E[(a + cX) - E(a + cX)]^2 = c^2\text{var}(X)$

Marginal and Conditional Distributions

$f(x) = \sum_y f(x, y)$ for each value X can take

$f(y) = \sum_x f(x, y)$ for each value Y can take

$$f(x|y) = P[X = x|Y = y] = \frac{f(x, y)}{f(y)}$$

If X and Y are independent random variables, then $f(x,y) = f(x)f(y)$ for each and every pair of values x and y. The converse is also true.

If X and Y are independent random variables, then the conditional probability density function of X given that

$$Y = y \text{ is } f(x|y) = \frac{f(x, y)}{f(y)} = \frac{f(x)f(y)}{f(y)} = f(x)$$

for each and every pair of values x and y. The converse is also true.

Expectations, Variances & Covariances

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$
$$= \sum_x \sum_y [x - E(X)][y - E(Y)] f(x, y)$$

$$\rho = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

$E(c_1X + c_2Y) = c_1E(X) + c_2E(Y)$
$E(X + Y) = E(X) + E(Y)$

$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z)$
$+ 2ab\text{cov}(X,Y) + 2ac\text{cov}(X,Z) + 2bc\text{cov}(Y,Z)$

If X, Y, and Z are independent, or uncorrelated, random variables, then the covariance terms are zero and:

$$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z)$$

Normal Probabilities

If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

If $X \sim N(\mu, \sigma^2)$ and a is a constant, then

$$P(X \geq a) = P\left(Z \geq \frac{a - \mu}{\sigma}\right)$$

If $X \sim N(\mu, \sigma^2)$ and a and b are constants, then

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right)$$

Assumptions of the Simple Linear Regression Model

SR1	The value of y, for each value of x, is $y = \beta_1 + \beta_2 x + e$
SR2	The average value of the random error e is $E(e) = 0$ since we assume that $E(y) = \beta_1 + \beta_2 x$
SR3	The variance of the random error e is $\text{var}(e) = \sigma^2 = \text{var}(y)$
SR4	The covariance between any pair of random errors, e_i and e_j is $\text{cov}(e_i, e_j) = \text{cov}(y_i, y_j) = 0$
SR5	The variable x is not random and must take at least two different values.
SR6	(*optional*) The values of e are *normally distributed* about their mean $e \sim N(0, \sigma^2)$

Least Squares Estimation

If b_1 and b_2 are the least squares estimates, then

$$\hat{y}_i = b_1 + b_2 x_i$$
$$\hat{e}_i = y_i - \hat{y}_i = y_i - b_1 - b_2 x_i$$

The Normal Equations

$$Nb_1 + \Sigma x_i b_2 = \Sigma y_i$$
$$\Sigma x_i b_1 + \Sigma x_i^2 b_2 = \Sigma x_i y_i$$

Least Squares Estimators

$$b_2 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma (x_i - \bar{x})^2}$$
$$b_1 = \bar{y} - b_2\bar{x}$$

Elasticity

$$\eta = \frac{\text{percentage change in } y}{\text{percentage change in } x} = \frac{\Delta y/y}{\Delta x/x} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y}$$

$$\eta = \frac{\Delta E(y)/E(y)}{\Delta x/x} = \frac{\Delta E(y)}{\Delta x} \cdot \frac{x}{E(y)} = \beta_2 \cdot \frac{x}{E(y)}$$

Least Squares Expressions Useful for Theory

$$b_2 = \beta_2 + \Sigma w_i e_i$$

$$w_i = \frac{x_i - \bar{x}}{\Sigma(x_i - \bar{x})^2}$$

$$\Sigma w_i = 0, \quad \Sigma w_i x_i = 1, \quad \Sigma w_i^2 = 1/\Sigma(x_i - \bar{x})^2$$

Properties of the Least Squares Estimators

$$\text{var}(b_1) = \sigma^2 \left[\frac{\Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2} \right] \quad \text{var}(b_2) = \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}$$

$$\text{cov}(b_1, b_2) = \sigma^2 \left[\frac{-\bar{x}}{\Sigma(x_i - \bar{x})^2} \right]$$

Gauss-Markov Theorem: Under the assumptions SR1–SR5 of the linear regression model the estimators b_1 and b_2 have the *smallest variance of all linear and unbiased estimators* of β_1 and β_2. They are the Best Linear Unbiased Estimators (BLUE) of β_1 and β_2.

If we make the normality assumption, assumption SR6, about the error term, then the least squares estimators are normally distributed.

$$b_1 \sim N\left(\beta_1, \frac{\sigma^2 \Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2}\right), b_2 \sim N\left(\beta_2, \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}\right)$$

Estimated Error Variance

$$\hat{\sigma}^2 = \frac{\Sigma \hat{e}_i^2}{N - 2}$$

Estimator Standard Errors

$$\text{se}(b_1) = \sqrt{\widehat{\text{var}(b_1)}}, \quad \text{se}(b_2) = \sqrt{\widehat{\text{var}(b_2)}}$$

t-distribution

If assumptions SR1–SR6 of the simple linear regression model hold, then

$$t = \frac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-2)}, \quad k = 1, 2$$

Interval Estimates

$$P[b_2 - t_c \text{se}(b_2) \leq \beta_2 \leq b_2 + t_c \text{se}(b_2)] = 1 - \alpha$$

Hypothesis Testing

Components of Hypothesis Tests

1. A *null* hypothesis, H_0
2. An *alternative* hypothesis, H_1
3. A test *statistic*
4. A *rejection* region
5. A conclusion

If the null hypothesis $H_0 : \beta_2 = c$ is ***true***, then

$$t = \frac{b_2 - c}{\text{se}(b_2)} \sim t_{(N-2)}$$

Rejection rule for a two-tail test: If the value of the test statistic falls in the rejection region, either tail of the *t*-distribution, then we reject the null hypothesis and accept the alternative.

Type I error: The null hypothesis is *true* and we decide to *reject* it.

Type II error: The null hypothesis is *false* and we decide *not* to reject it.

***p*-value rejection rule:** When the *p*-value of a hypothesis test is *smaller* than the chosen value of α, then the test procedure leads to *rejection* of the null hypothesis.

Prediction

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

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

A $(1 - \alpha) \times 100\%$ confidence interval, or prediction interval, for y_0

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

Goodness of Fit

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

$$SST = SSR + SSE$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = (\text{corr}(y, \hat{y}))^2$$

Log-Linear Model

$$\ln(y) = \beta_1 + \beta_2 x + e, \quad \widehat{\ln(y)} = b_1 + b_2 x$$

$100 \times \beta_2 \approx \%$ change in y given a one-unit change in x.

$$\hat{y}_n = \exp(b_1 + b_2 x)$$

$$\hat{y}_c = \exp(b_1 + b_2 x)\exp(\hat{\sigma}^2/2)$$

Prediction interval:

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

Generalized goodness-of-fit measure $R_g^2 = (\text{corr}(y, \hat{y}_n))^2$

Assumptions of the Multiple Regression Model

MR1 $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i$

MR2 $E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \Leftrightarrow E(e_i) = 0.$

MR3 $\text{var}(y_i) = \text{var}(e_i) = \sigma^2$

MR4 $\text{cov}(y_i, y_j) = \text{cov}(e_i, e_j) = 0$

MR5 The values of x_{ik} are not random and are not exact linear functions of the other explanatory variables.

MR6 $y_i \sim N[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2]$
$\Leftrightarrow e_i \sim N(0, \sigma^2)$

Least Squares Estimates in MR Model

Least squares estimates $b_1, b_2, \ldots, b_K$ minimize $S(\beta_1, \beta_2, \ldots, \beta_K) = \Sigma(y_i - \beta_1 - \beta_2 x_{i2} - \cdots - \beta_K x_{iK})^2$

Estimated Error Variance and Estimator Standard Errors

$$\hat{\sigma}^2 = \frac{\Sigma \hat{e}_i^2}{N - K} \quad \text{se}(b_k) = \sqrt{\widehat{\text{var}(b_k)}}$$

Hypothesis Tests and Interval Estimates for Single Parameters

Use t-distribution $\quad t = \dfrac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-K)}$

t-test for More than One Parameter

$$H_0 : \beta_2 + c\beta_3 = a$$

When H_0 is true $\quad t = \dfrac{b_2 + cb_3 - a}{\text{se}(b_2 + cb_3)} \sim t_{(N-K)}$

$$\text{se}(b_2 + cb_3) = \sqrt{\widehat{\text{var}(b_2)} + c^2\widehat{\text{var}(b_3)} + 2c \times \widehat{\text{cov}(b_2, b_3)}}$$

Joint F-tests

To test J joint hypotheses,

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

To test the overall significance of the model the null and alternative hypotheses and F statistic are

$$H_0 : \beta_2 = 0,\ \beta_3 = 0, \ldots, \beta_K = 0$$
$$H_1 : \textit{at least one} \text{ of the } \beta_k \text{ is nonzero}$$
$$F = \frac{(SST - SSE)/(K-1)}{SSE/(N-K)}$$

RESET: A Specification Test

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i \qquad \hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3}$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + e_i, \quad H_0 : \gamma_1 = 0$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + e_i, \quad H_0 : \gamma_1 = \gamma_2 = 0$

Model Selection

$\text{AIC} = \ln(SSE/N) + 2K/N$

$\text{SC} = \ln(SSE/N) + K\ln(N)/N$

Collinearity and Omitted Variables

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$

$$\text{var}(b_2) = \frac{\sigma^2}{(1 - r_{23}^2)\sum (x_{i2} - \bar{x}_2)^2}$$

When x_3 is omitted, $\text{bias}(b_2^*) = E(b_2^*) - \beta_2 = \beta_3 \dfrac{\widehat{\text{cov}(x_2, x_3)}}{\widehat{\text{var}(x_2)}}$

Heteroskedasticity

$$\text{var}(y_i) = \text{var}(e_i) = \sigma_i^2$$

General variance function

$$\sigma_i^2 = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS})$$

Breusch-Pagan and White Tests for H_0: $\alpha_2 = \alpha_3 = \cdots = \alpha_S = 0$

When H_0 is true $\quad \chi^2 = N \times R^2 \sim \chi^2_{(S-1)}$

Goldfeld-Quandt test for $H_0 : \sigma_M^2 = \sigma_R^2$ versus $H_1 : \sigma_M^2 \neq \sigma_R^2$

When H_0 is true $\quad F = \hat{\sigma}_M^2/\hat{\sigma}_R^2 \sim F_{(N_M - K_M, N_R - K_R)}$

Transformed model for $\text{var}(e_i) = \sigma_i^2 = \sigma^2 x_i$

$$y_i/\sqrt{x_i} = \beta_1(1/\sqrt{x_i}) + \beta_2(x_i/\sqrt{x_i}) + e_i/\sqrt{x_i}$$

Estimating the variance function

$$\ln(\hat{e}_i^2) = \ln(\sigma_i^2) + v_i = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i$$

Grouped data

$$\text{var}(e_i) = \sigma_i^2 = \begin{cases} \sigma_M^2 & i = 1, 2, \ldots, N_M \\ \sigma_R^2 & i = 1, 2, \ldots, N_R \end{cases}$$

Transformed model for feasible generalized least squares

$$y_i \Big/ \sqrt{\hat{\sigma}_i} = \beta_1 \left(1 \Big/ \sqrt{\hat{\sigma}_i}\right) + \beta_2 \left(x_i \Big/ \sqrt{\hat{\sigma}_i}\right) + e_i \Big/ \sqrt{\hat{\sigma}_i}$$

Regression with Stationary Time Series Variables

Finite distributed lag model

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + v_t$$

Correlogram

$$r_k = \sum (y_t - \bar{y})(y_{t-k} - \bar{y}) / \sum (y_t - \bar{y})^2$$

For $H_0 : \rho_k = 0, \quad z = \sqrt{T} r_k \sim N(0, 1)$

LM test

$y_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t \quad$ Test $H_0 : \rho = 0$ with t-test

$\hat{e}_t = \gamma_1 + \gamma_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t \quad$ Test using $LM = T \times R^2$

AR(1) error $\quad y_t = \beta_1 + \beta_2 x_t + e_t \quad e_t = \rho e_{t-1} + v_t$

Nonlinear least squares estimation

$$y_t = \beta_1(1 - \rho) + \beta_2 x_t + \rho y_{t-1} - \beta_2 \rho x_{t-1} + v_t$$

ARDL(p, q) model

$$y_t = \delta + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + v_t$$

AR(p) forecasting model

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + v_t$$

Exponential smoothing $\quad \hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}$

Multiplier analysis

$$\delta_0 + \delta_1 L + \delta_2 L^2 + \cdots + \delta_q L^q = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_p L^p) \times (\beta_0 + \beta_1 L + \beta_2 L^2 + \cdots)$$

Unit Roots and Cointegration

Unit Root Test for Stationarity: Null hypothesis:

$$H_0 : \gamma = 0$$

Dickey-Fuller Test 1 (no constant and no trend):

$$\Delta y_t = \gamma y_{t-1} + v_t$$

Dickey-Fuller Test 2 (with constant but no trend):

$$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$$

Dickey-Fuller Test 3 (with constant and with trend):

$$\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + v_t$$

Augmented Dickey-Fuller Tests:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \Delta y_{t-s} + v_t$$

Test for cointegration

$$\Delta \hat{e}_t = \gamma \hat{e}_{t-1} + v_t$$

Random walk: $\quad y_t = y_{t-1} + v_t$

Random walk with drift: $\quad y_t = \alpha + y_{t-1} + v_t$

Random walk model with drift and time trend:

$$y_t = \alpha + \delta t + y_{t-1} + v_t$$

Panel Data

Pooled least squares regression

$$y_{it} = \beta_1 + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it}$$

Cluster robust standard errors $\quad \text{cov}(e_{it}, e_{is}) = \psi_{ts}$

Fixed effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{1i}$ not random

$y_{it} - \bar{y}_i = \beta_2(x_{2it} - \bar{x}_{2i}) + \beta_3(x_{3it} - \bar{x}_{3i}) + (e_{it} - \bar{e}_i)$

Random effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{it} = \bar{\beta}_1 + u_i$ random

$y_{it} - \alpha\bar{y}_i = \bar{\beta}_1(1 - \alpha) + \beta_2(x_{2it} - \alpha\bar{x}_{2i}) + \beta_3(x_{3it} - \alpha\bar{x}_{3i}) + v_{it}^*$

$$\alpha = 1 - \sigma_e \Big/ \sqrt{T\sigma_u^2 + \sigma_e^2}$$

Hausman test

$$t = (b_{FE,k} - b_{RE,k}) \Big/ \left[\widehat{\text{var}(b_{FE,k})} - \widehat{\text{var}(b_{RE,k})}\right]^{1/2}$$