

## XALQARO NORDIK UNIVERSITETI

### Iqtisodiyot va pedagogika fakulteti, Iqtisodiyot va biznesni boshqarish kafedrası

Fan o'qituvchisi: Sabirov Xasan Nusratovich

---

#### Mavzu: Ko'p omilli regression modellar

##### Reja:

1. Big andy's burger barn ma'lumotlari asosida model tuzish
2. Eng kichik kvadratlar bashorati
3. Tasviriy statistika ma'lumotlarini olish aniqligi

##### 5.1. Big andy's burger barn ma'lumotlari asosida model tuzish

Oddiy chiziqli regressiyada erksiz o'zgaruvchining o'rtacha qiymati doimiy va bitta erkli o'zgaruvchining chiziqli funktsiyasi sifatida modellashtiriladi. Ko'p chiziqli regressiya modelida erkli o'zgaruvchilar sonini kengaytiradi. Shunday qilib, bu oddiy, ammo muhim kengaytma bo'lib, chiziqli regressiyani juda kuchli qiladi.

Ushbu bobda foydalanilgan misol Big Andy's Burger Barn uchun savdo modeli tuziladi. Big Andy gamburgerining sotuvi uning narx va reklama darajasiga bog'liqligini ifodalaniadi. Shunday qilib, model ikkita ta'sir etuvchi omil va o'zgarmasni (constant)ni o'z ichiga oladi:

$$SALES = \beta_1 + \beta_2 PRICE + \beta_3 ADVERT + e$$

bu yerda SALES-ma'lum bir shahardagi oylik sotuvlar hajmi (1000 AQSh dollari miqdorida), PRICE - gamburgerning narxi (dollarda) va ADVERT - reklama xarajatlari (ming dollarlarda) shuningdek,  $i = 1, 2, \dots, N$ .

Birinchiidan, Stata-ni ishga tushiring va buyruq satridan ishchi katalogni ma'lumotlar fayllaringiz joylashgan joyga o'zgartiring yoki Internet manzilida saqlangan fayllardan foydalanmoqchi bo'lsangiz, Stata jurnali faylini yozmoqchi bo'lgan joy. Barcha misollarimizda bo'lgani kabi,

```
cd c:\data\poe4stata
```

Biz ochmoqchi bo'lgan faylni toping, bu yerda *andy.dta* ni tanlaymiz va **Open** tugmasini bosib. Har qanday modelni taxmin qilishdan oldin, ma'lumotlarning dasturiy ta'minotigizga to'g'ri yuklanganligini tekshirish yaxshi amaliyotdir. Bunday holda,

umumiy statistikani tekshiring va dastlabki bir nechta kuzatuvlarni sanab o'ling. Ular qo'llanmaning 5.1-jadvalidagilarga mos kelishi kerak. Asosan, siz ular mantiqiy yoki mantiqsizligini bilish uchun umumiy statistikaga qarashimiz kerak. Sizda kerakli kuzatuvlar soni bormi? O'zgaruvchilar diapazoni o'rtacha ko'rinadimi - bu holda savdo, narxlar va reklama xarajatlari ijobiymi?

Eslatib o'tamiz, *summarize* buyrug'i umumiy statistik ma'lumotlarning asosiy to'plamini va *list* ma'lumotlar to'plamidagi o'zgaruvchilar hamda kuzatishlarni ifodalaydi. Ro'yxatga qo'shilgan *in 1/5* qismi Stataga chop etishni 1 dan 5 gacha bo'lgan kuzatishlar bilan cheklashni aytadi. Natijalar:

. summarize

Variable	Obs	Mean	Std. Dev.	Min	Max
sales	75	77.37467	6.488537	62.4	91.2
price	75	5.6872	.518432	4.83	6.49
advert	75	1.844	.8316769	.5	3.1

. list in 1/5

	sales	price	advert
1.	73.2	5.69	1.3
2.	71.8	6.49	2.9
3.	62.4	5.63	.8
4.	67.4	6.22	.7
5.	89.3	5.02	1.5

Odatda vaqtni tejaydigan buyruq satridan regressiyani baholashni tanlashingiz mumkin.

*regress sales price advert*

Agar sintaksisni eslay olmasangiz, ochiladigan menyulardan foydalaning. **Statistics > Linear models and related > Linear regression** -ni tanlang. Bu 2.4-bo'limda ko'rsatilgan **regress** dialog oynasini ochadi. **Sales** ni bog'liq o'zgaruvchi sifatida kiriting, **price** va **advert** ni mustaqil o'zgaruvchi sifatida kiriting hamda "yuborish" tugmasini bosing. Esda tutingki, Stata modelga kesishuvni joylashtiradigan doimiy(constant) o'zgaruvchini o'z ichiga oladi. Agar sizda modeldagi kesishuvni olib tashlash uchun jiddiy sabab bo'lmasa, har doim bittasini kiritishingiz kerak. Demak, bu Stata uchun standart hisoblanadi. Natija quyida keltirilgan:

Source	SS	df	MS			
Model	1396.53893	2	698.269465	Number of obs =	75	
Residual	1718.94294	72	23.8742075	F( 2, 72) =	29.25	
Total	3115.48187	74	42.1011063	Prob > F =	0.0000	
				R-squared =	0.4483	
				Adj R-squared =	0.4329	
				Root MSE =	4.8861	

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
price	-7.907854	1.095993	-7.22	0.000	-10.09268	-5.723032
advert	1.862584	.6831955	2.73	0.008	.500659	3.22451
_cons	118.9136	6.351638	18.72	0.000	106.2519	131.5754

Kesish uchun parametr  $\beta_2$  va  $\beta_3$  qiymatlar uchun mos ravishda *price* va *advert* bilan belgilanadi. Dispersiya-kovariatsiya matritsasi eng kichik kvadratlar usuli sizning modelingiz parametrlarini o'lchashga qodir bo'lgan aniqlikni ifodalaydi. Eng kichik kvadratlarning aniqligi bir qancha narsalarga, jumladan, ma'lumotlaringizning o'zgaruvchanligiga ( $\sigma^2$ ), namunangizning o'lchamiga va mustaqil o'zgaruvchilarning raqamli qiymatlari nazarda tutilgan "tajriba" dizayniga bog'liq. Ushbu ma'lumot kesma, har bir qiymat va ular orasidagi har qanday kovariatsiyaning o'lchovini o'z ichiga olgan taxminiy variatsiya-kovariatsiya matritsasi bilan umumlashtiriladi.

$$cov(b_1, b_2, b_3) = \begin{bmatrix} var(b_1) & cov(b_1, b_2) & cov(b_1, b_3) \\ cov(b_1, b_2) & var(b_2) & cov(b_2, b_3) \\ cov(b_1, b_3) & cov(b_2, b_3) & var(b_3) \end{bmatrix}$$

Eng kichik kvadratlar hisoblagichining dispersiyalari diagonalga to'g'ri keladi va har bir juftlik orasidagi kovariatsiya pastki uchburchakda bo'ladi.

Regressiyadan keyin dispersiya-kovariatsiya matritsasi taxminini topish uchun quyidagi buyruqdan foydalaning va natijani oling.

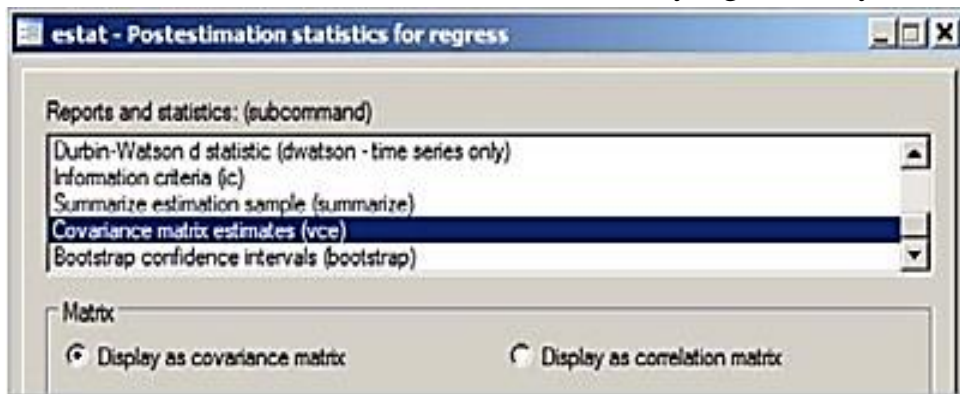
```
. estat vce

Covariance matrix of coefficients of regress model
```

e(v)	price	advert	_cons
price	1.2012007		
advert	-.01974215	.46675606	
_cons	-6.7950641	-.7484206	40.343299

Shunday qilib,  $b_2$  ning taxminiy dispersiyasi 1,20 ga,  $b_3$  bilan taxminiy kovariatsiyasi esa 0,0197 ga teng. Diagonal elementlarning kvadrat ildizlarini olish eng kichik kvadratlar standart xatolarni keltirib chiqaradi va buni osongina tekshirish mumkin, yuqoridagi regressiya jadvalidagi natijalarga mos keladi.

**Statistics >Linear Models and related >Regression diagnostics > Specification tests, etc.**, tanlab, ochiladigan menyular yordamida dialog oynasini chaqirish mumkin, bu esa regressiya uchun taxminiy statistik ma'lumotlar dialog oynasini ochiladi. Quyida ko'rsatilganidek, Kovarians matritsasini baholash (**vce**) buyrug'idan foydalanishiz mumkin:



## 5.2. Eng kichik kvadratlar bashorati

Hisoblangan qiymatlarni topishni 2-bobda ko'rib chiqqandik va bu yerda qayta ko'rib chiqamiz. Aytaylik, Big Andy 5,50 dollarlik narxda va 1200 dollarlik reklama xarajatlari uchun savdo daromadini hisoblangan qiymatlarni topib olamiz. Ushbu hisoblangan qiymatni quyidagi buyruq orqali topib olamiz:

*di \_b[\_cons] + \_b[price]\*5.50 + \_b[advert]\*1.2*

**Results** oynasida quyidagi natijalar aks ettiriladi:

```
. di _b[_cons] + _b[price]*5.50 + _b[advert]*1.2  
77.655513
```

Ushbu misolda, berilgan narx va reklama xarajatlari uchun hisoblangan qiymat sotish 77,66 ni tashkil qiladi.

Shartli prognozlarini yaratish uchun ma'lumotlar to'plamiga kuzatishlar va hisoblangan qiymat buyrug'ini qo'shadigan 4-bobda o'rganilgan buyruqdan ham foydalanish mumkin. Ma'lumotlar muharriri **Data>Data editor** ni oching va yangi ma'lumotlar qatoriga mustaqil o'zgaruvchilarning kerakli qiymatlarini qo'shing (bu yerdagi kuzatish 76). Buni **Command** oynasidan ham ifodalash mumkin. Avval **76**-kuzatishni ifodalash uchun *set obs 76* buyrug'idan foydalaning. So'ngra *replace* buyrug'idan foydalanib, yangi kuzatish uchun mustaqil o'zgaruvchilarning kerakli qiymatlarini kiriting. Bu quyidagi dastlabki uchta qatorda amalga oshiriladi

*set obs 76*

*replace price = 5.50 in 76*

*replace advert = 1.2 in 76*

*predict yhat*

*list yhat in 76*

*price* va *advert* o'zgaruvchilarga kerakli qiymatlari kiritilgandan so'ng, Stata-ning *predict* buyrug'i orqali hisoblangan qiymatni yaratish uchun foydalanish mumkin. Bashorat qilish buyrug'i Big Andy regressiyasi oxirgi hisoblangan deb taxmin qiladi va u regressiyadan bashorat qilingan qiymatlarni olish uchun standart variant, **xb** dan foydalanadi. Yana bir bor, ushbu **list yhat** ma'lumotlarning 76-kuzatuvini ro'yxatga olish uchun *in 76* shartidan foydalaning.

```
. list yhat in 76
```

	yhat
76.	77.65551

### 5.3. Tasviriy statistika ma'lumotlarini olish aniqligi

Ushbu va boshqa statistik ma'lumotlarni qo'lda hisoblash uchun zarur bo'lgan ko'plab natijalar Stata tomonidan ichki hisob kitoblar natijalarida saqlanadi. Baholashdan foydalanish natijasida olingan natijalar mazmunini ko'rish

*ereturn list*

qisman natijalar ifodalaniadi:

scalars:

```
e(N) = 75
e(df_m) = 2
e(df_r) = 72
e(F) = 29.24785947967357
e(r2) = .4482577622149436
e(rmse) = 4.886123970679952
e(mss) = 1396.538929773235
e(rss) = 1718.942936893432
```

Tenglamaning error (xato) dispersiyasini baholash uchun tenglamadan foydalaning:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^N \hat{e}_i^2}{N - K}$$

Bu yerda  $\hat{e}_i^2$  eng kichik kvadratlarning kvadrati qoldiqlaridir. Asosan, numerator faqat erkinlikning qoldiq darajalariga bo'lingan regressiyadan olingan kvadrat xatolar yig'indisidir. Erkinlikning qoldiq darajalari  $e(df\_r)$  da, kvadratik xatolar yig'indisi esa  $e(rss)$  da ifodalanadi. Shunday qilib,

*scalar sighat2* =  $e(rss)/e(df\_r)$

*scalar list sighat2*

Natijasi:

```
. scalar list sighat2
sighat2 = 23.874207
```

Ushbu ma'lumotlarning barchasini Stata tomonidan ishlab chiqarilgan dispersiya jadvali tahlilidan olish mumkin, u quyida ko'rsatilgan.

	SSE	N-K	SSE/(N-K)		
<pre>. regress sales price advert</pre>					
Source	SS	df	MS		
Model	1396.53921	2	698.269603	Number of obs = 75	
Residual	1718.94281	72	23.8742057	F( 2, 72) = 29.25	
Total	3115.48202	74	42.1011083	Prob > F = 0.0000	
				R-squared = 0.4483	
				Adj R-squared = 0.4329	
				Root MSE = 4.8861	
sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
price	-7.907856	1.095993	-7.22	0.000	-10.09268 -5.723034
advert	1.862584	.6831955	2.73	0.008	.5006587 3.224509
_cons	118.9136	6.351638	18.72	0.000	106.2519 131.5754

Ko'rib turganingizdek, taxminiy farq 23,874 ni tashkil qiladi. Ildiz MSE bu raqamning kvadrat ildizidir va sariq rang bilan belgilangan jadvalning o'ng tomonidagi chop etishda ko'rinadi.

#### 5.4. Ishonch intervallari

Regressiya jadvali sizga eng kichik kvadratlar hisobini va taxminiy standart xatolarni beradi. Berilgan standart xato eng kichik kvadratlar qiziqish parametrini qanchalik aniq o'lchashga qodirligini baholashdir. Ishonch oralig'i xuddi shunday maqsadga xizmat qiladi, garchi uni izohlash ancha sodda bo'lsa-da, chunki u sizga yuqori va pastki chegaralarni beradi, ular orasida noma'lum parametr berilgan ehtimollik bilan yotadi. Faraz qilaylik,

kimdir Big Andy's Burger Barn'da ma'lum darajadagi reklama uchun narx o'zgarishiga sotishdan tushgan tushumning javobi  $\beta_2$  uchun 95% intervalli taxminni baholaydi. Bu jarayon 3-bobdagi tahlil bilan bir xil. Bu safar kritik qiymatlarni olish uchun foydalaniladigan t-taqsimoti  $N-K=75-3=72$  erkinlik darajasiga ega bo'ladi.

95% ishonch oralig'i matndagi (5.13) tenglamaga asoslanadi:

$$P(-t_c < t_{(72)} < t_c) = 0.95$$

Bu tenglamani qanoatlantiradigan  $t_{(72)}$  - taqsimotdan, uni  $t_c$  deb ataydigan kritik qiymat kerak. Doimiy  $t_c$  - bu t-distribution dan  $\alpha/2$  kritik qiymat va  $\alpha$  "rad etish" maydoni (ishonch oralig'idan tashqaridagi maydon) bilan bog'liq bo'lgan jami kerakli ehtimollikdir. Kritik qiymat  $t_c$  ni statistik jadvalda topish mumkin yoki Stata-dagi **invttail(df,  $\alpha/2$ )** funksiyasidan foydalanib,  $t_{(72)}$  - taqsimotdan 0,025 kritik qiymat 1,993 ekanligini aniqlash mumkin. Bir oz algebra natijasida  $\beta_2$  ning 95% intervalli baholovchisi, uning eng kichik kvadratlar hisoblagichi  $b_2$  va hisoblangan standart xatosi,  $se(b_2)$ :

$$[b_2 - 1.993 * se(b_2), b_2 + 1.993 * se(b_2)]$$

Quyida siz taxminiy ishonch oralig'i uchun pastki va yuqori so'nggi nuqtalarni yaratish uchun Stata buyruqlaridan foydalanamiz:

```
scalar lb = _b[price] - invttail(e(df_r),.025) * _se[price]
scalar ub = _b[price] + invttail(e(df_r),.025) * _se[price]
scalar list lb ub
```

```
. scalar list bL bU
      bL = -10.092676
      bU = -5.7230322
```

Bunday holda, biz Big Andy regressiya tenglamasining bir nechta saqlangan natijalaridan foydalandik. Bularga narx bo'yicha koeffitsientning eng kichik kvadratlari bahosi, **\_b[price]** , uning taxminiy standart xatosi, **\_se[price]** va regressiyadan ozod bo'lishning qoldiq darajalari (N-K), **e(df\_r)**. kiradi.

Har safar chiziqli regressiya modelini taxmin qilganingizda 95% ishonch oralig'i sukut bo'yicha hisoblanadi. Interval regressiya natijalarining oxirgi ikki ustunida ko'rinadi. Regressiyadan so'ng **level()** opsiyasidan foydalanib, ehtimolliklarni 90% ga o'zgartirishingiz mumkin. Misol uchun, regressiya bilan 90% intervallarni olish uchun quyidagi buyruqdan foydalaning:

```
regress sales price advert, level(90)
```

Intervalni menyu tizimi orqali ham o'zgartirish mumkin. "**Statistics > Linear models and related > Linear regression**" -ni ochish uchun ochiladigan menyudan "Chiziqli regressiya" dialog oynasini tanlang - "Reporting" yorlig'ini tanlang va bu yerda ko'rsatilganidek, ishonch darajasini o'zingiz xohlagan darajaga o'zgartiring:



#### 5.4.1. Parametrlarning chiziqli birikmasi uchun ishonch oralig'i

Big Andy reklama xarajatlarini 800 dollarga oshirishni va narxni 40 sentga tushirishni rejalashtirmoqda. Kutilayotgan savdo hajmining o'zgarishi

$$\lambda = -0.4\beta_2 + 0.8\beta_3$$

parametrlarning chiziqli birikmasidir. Stata parametrlarning chiziqli kombinatsiyasi uchun turli statistik ma'lumotlarni hisoblaydigan o'rnatilgan **lincom** buyruqlarini o'z ichiga oladi; **lincom** har qanday baholashdan so'ng koeffitsientlarning chiziqli birikmalari uchun nuqta baholarini, standart xatolarni, **t** yoki **z** statistikasini, p-qiymatlarini va ishonch oraliqlarini hisoblaydi. Big Andy sotuvining eng kichik kvadratlari  $\lambda$  yordamida taxmin qilish uchun:

*lincom -0.4\*price+0.8\*advert, level(90)*

Bunday holda, ishonch oralig'i uchun qamrov ehtimolini 90% ga o'rnatish uchun ixtiyoriy **level(90)** buyrug'i kiritilgan. Natija:

```
. lincom -0.4*price+0.8*advert, level(90)
( 1)  - .4*price + .8*advert = 0
```

sales	Coef.	Std. Err.	t	P> t	[90% Conf. Interval]
(1)	4.653209	.7096133	6.56	0.000	3.470785 5.835633

Stata buyruqni ekranga aks ettiradi va chiziqli kombinatsiyani takrorlaydi.  $\lambda$  ning taxminiy qiymati 4,65, uning taxminiy standart xatosi 0,7096 va 90% ishonch oralig'i (3,471, 5,835).

Xuddi shu narsani bajarish uchun algebra va arifmetikadan foydalanish mumkin. Stata hisoblash ishlarini osonlashtirsa-da, foydalanuvchi algebrani taqdim etishi kerak. Bu yerda  $\hat{\lambda}$  ning standart xatosini hisoblash uchun eng kichik baholovchining taxminiy dispersiyalari va kovariatsiyalari qo'llaniladi.

*matrix cov=e(V)*

*scalar lambda = -0.4\*\_b[price]+0.8\*\_b[advert]*

*scalar var\_lambda = (-0.4)^2\*cov[1,1]+(0.8)^2\*cov[2,2]+2\*  
(-0.4)\*(0.8)\*cov[1,2]*

**matrix cov=e(V)** bayonoti **e(V)** da saqlangan dispersiya kovariatsiyasini **cov** etiketli matritsaga yozadi. Uchinchi va to'rtinchi qatorlarda Stata matritsasini indekslash imkoniyatlari kovariatsiya matritsasiidan  $\hat{\lambda}$  dispersiyasini hisoblashda foydalanish uchun kerakli elementlarni tanlash uchun ishlatiladi. Masalan, **cov[1,1]** **cov** ning birinchi qatori va

birinchi ustunidagi elementga ishora qiladi, ya'ni  $\beta_2$ ning dispersiyasi. Esda tutingki, Stata avtomatik ravishda tenglama oxiriga doimiy qo'yadi, shuning uchun  $\beta_2$ aslida koeffitsient vektorining birinchi elementida joylashgan.

Dispersiya hisoblangandan so'ng, uning kvadrat ildizini oling va standart xatoni olish uchun  $\lambda$  bahosini natijaga bo'ling. U yerdan ishonch oralig'i odatdagi usulda hisoblanadi.

*scalar se = sqrt(var\_lambda)*

*scalar t = lambda/se*

*scalar lb = lambda-invttail(e(df\_r),.05)\*se*

*scalar ub = lambda+invttail(e(df\_r),.05)\*se*

Natija olishda **invttail(e(df\_r),.05)**  $t_{(72)}$  – taqsimotidan 5% kritik qiymatni olish uchun ishlatiladi. Natijalar **lincom** natijalariga deyarli mos keladi.

```
. scalar list lambda var_lambda se t lb ub
      lambda = 4.6532091
var_lambda = .50355097
      se = .70961325
      t = 6.5573876
      lb = 3.4707851
      ub = 5.8356332
```

## 5.5 Gipoteza testlari

### 5.5.1 Ikki tomonlama testlar

Ikki tomonlama muqobilga nisbatan  $\beta_2=0$  bo'lgan testning t nisbati  $\beta_2 \neq 0$

$$t = \frac{b_2 - 0}{se(b_2)} \sim t_{(N-K)}$$

$\beta_2 = 0$  to'g'ri degan nol gipoteza mavjud bo'lsa. Hisoblangan qiymat:

$$t = \frac{-7.908}{1.096} = -7.215$$

Stata buni buyruq satridan osongina hisoblaydi. p-qiymati **tail(df,stat)** buyrug'i yordamida olinadi, bu yerda **df** - erkinlik darajalari va **stat** - test statistikasining qiymati. Bu funksiya **df** erkinlik darajasi bilan *t-distribution* uchun **stat** ning o'ng tomonidagi quyruq maydonini o'lchaydi. Ikki tomonlama test uchun t ning ijobiy qiymatining quyruq maydonini ikki barobarga oshirish orqali hisoblangan statistik (**stat**) ning chap tomonidagi maydonni o'z ichiga oladi. Demak, **2\*ttail(72,abs(t1))** dan foydalaning.

To'liq Stata buyrug'i quyidagicha:

*scalar t1 = (\_b[price]-0)/\_se[price]*

*scalar p1 = 2\*ttail(72,abs(t1))*

*scalar list t1 p1*

$\beta_3 = 0$  bo'lgan ikki tomonlama test xuddi shunday amalga oshiriladi va to'liq natijalar ushbu bobning oxirida topiladi.

### 5.5.2. Bir tomonlama testlar

Muqobilga qarshi  $\beta_2 \geq 0$  degan nol gipotezani sinab ko'rish uchun!  $\beta_2 < 0$ , xuddi shu statistikadan foydalaning:



$$t = \frac{b_2 - 0}{se(b_2)} \sim t_{(N-K)}$$

Hisoblash bir xil bo'lsa-da, kritik qiymat endi t- distribution ning chap dumida:

*scalar t1 = (\_b[price]-0)/\_se[price]*

*scalar crit = -invttail(e(df\_r),.05)*

*scalar pval = 1-ttail(e(df\_r),t1)*

**ttail(e(df\_r),t1)** ni tasdiqlashi **df\_r** erkinlik darajasiga ega t-tasodifiy o'zgaruvchining **t1** dan katta bo'lish ehtimolini hisoblaydi. t-distribution simmetrik bo'lgani uchun va biz **t1** dan kichik bo'lish ehtimolini xohlaymiz, biz 1 dan **ttail(e(df\_r),t1)** ni ayiramiz.

Agar p-qiymati testning 5% ahamiyatlilik darajasidan kam bo'lsa, u holda statistik ahamiyatga ega bo'lgan ushbu darajadagi nol gipotezani rad eting. Aks holda, ushbu ahamiyat darajasida nolli gipotezani rad etib bo'lmaydi.

Sinovni o'tkazishning boshqa usuli - statistik ma'lumotlarni 5% kritik qiymat bilan solishtirish. Stata kritik qiymatni **invttail(e(df\_r),.05)** funksiyasi yordamida ishlab chiqaradi. Ushbu hisob-kitoblarning natijalari:

```
. scalar list t1 crit pval
      t1 = -7.2152415
      crit = -1.6662937
      pval = 2.212e-10
```

Reklama samaradorligi testi  $\beta_3 \leq 1$  muqobilga qarshi  $\beta_3 > 1$  degan nol gipotezani hisobga oladi. Buyruqlar quyidagicha:

*scalar t2 = (\_b[advert]-1)/\_se[advert]*

*scalar crit = invttail(e(df\_r),.05)*

*scalar pval = ttail(e(df\_r),t2)*

Muqobil variant t-distribution ning o'ng tomonida joylashganligi sababli buyruq biroz shaffofroq. Yana **invttail** t-distribution uchun 5% o'ng dum kritik qiymatini va **ttail t2** ning o'ng tomonidagi maydonni hisoblaydi. Ushbu buyruqlar quyidagi natijalarni ifodalaydi:

```
. scalar list t2 crit pval
      t2 = 1.2625732
      crit = 1.6662937
      pval = .10540831
```

### 5.5.3. Chiziqli parametirni tekshirish

Big Andy kompaniyasining marketing bo'yicha maslahatchisining ta'kidlashicha, narxni 20 sentga tushirish reklama xarajatlarini 500 dollarga oshirishdan ko'ra savdo daromadini oshirish uchun samaraliroq bo'ladi. Model nuqtai nazaridan  $-0.2\beta_2 > 0.5\beta_3$  deb hisoblaydi. Nol gipoteza  $-0.2\beta_2 - 0.5\beta_3 \leq 0$  va muqobil  $-0.2\beta_2 - 0.5\beta_3 > 0$  ga teng.

Yana bir bor, **lincom** buyrug'i foydalidir, chunki u chiziqli kombinatsiyani va uning standart xatosini baholaydi. Bir tomonlama test sifatida, **lincom**-ning topilgan natijasiga

tayangandan ko'ra, to'g'ridan-to'g'ri tanqidiy qiymatni yoki p-qiymatini hisoblash yaxshidir, bu sizga ikki tomonlama t-testi uchun faqat p-qiymatini beradi. Birinchidan, **lincom**-dan foydalaning:

**lincom -0.2\*price-0.5\*advert**

natija esa:

( 1) - .2\*price - .5\*advert = 0

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
(1)	.6502787	.4009846	1.62	0.109	-.1490694	1.449627

Hisoblangan koeffitsient 0,650 ni standart xatolik 0,400 ga bo'lish natijasida t nisbati 1,62 ga teng bo'ladi. Bir tomonlama p-qiymati ikki tomonlama qiymatning yarmini tashkil etadi, agar taxminiy chiziqli kombinatsiya kutilgan belgiga ega bo'lsa. Chalkashmaslik uchun biz quyida ko'rsatilganidek, bir tomonlama testlar uchun mos kritik qiymat yoki p-qiymatini hisoblashni tavsiya qilamiz.

*scalar t = r(estimate)/r(se)*

*scalar crit = invttail(r(df),.05)*

*scalar pval = ttail(r(df),t)*

*scalar list crit t pval*

```
. scalar list crit t pval
      crit = 1.6662937
         t = 1.6217052
        pval = .05461891
```

Yangi element **r(estimate)**, **r(df)** va **r(se)** ni o'z ichiga oladi. Bu **lincom** tomonidan hisoblangan taxminiy chiziqli parametr va standart xato. Hisoblashdan keyin mavjud bo'lgan **e()** buyruqlari singari, boshqa ko'plab buyruqlar ham (masalan, **test** va **lincom**) keyingi hisoblash uchun natijalarni saqlaydi. Buyruqlar:

*return list*

natija:

```
scalars:
          r(df) = 72
          r(se) = .4009845529554849
          r(estimate) = .65027873012098
```

Bu shuni ko'rsatadiki, erkinlik darajalari (N-K), baholash va uning standart xatosi **lincom** buyrug'idan oldingi natijalar qismida saqlangan. Natijalarni takror ko'rish uchun, saqlangan natijalar to'plamida nima borligini tekshirish uchun **ereturn list** yoki **return list** buyruqlaridan foydalanishingiz mumkin.

## 5.6. Polinomial tenglamalar

Mustaqil va bog'liq o'zgaruvchilar o'rtasidagi chiziqli bo'lmagan munosabatlarga ruxsat berishning usullaridan biri modelga regressorlarning polinomlarini kiritishdir. Ushbu misolda qo'shimcha 1000 dollarlik reklamaning marjinal ta'siri ko'proq reklama ishlatilganligi sababli kamayishi kutilmoqda. Model quyidagicha hosil bo'ladi:

$$SALES = \beta_1 + \beta_2 PRICE + \beta_3 ADVERT + \beta_4 ADVERT^2 + e$$

Buni baholashning ikki yo'li mavjud. Birinchidan, yangi  $ADVERT^2$  o'zgaruvchisini yaratish, uni modelga qo'shish va parametrlarni baholash mumkin. Bu quyidagicha:

*generate a2 = advert\*advert*

*reg sales price advert a2*

Buyruq natijasida quyidagi natijalar ifodalaniladi:

*. reg sales price advert a2*

Source	SS	df	MS			
Model	1583.39744	3	527.799145	Number of obs =	75	
Residual	1532.08443	71	21.5786539	F( 3, 71) =	24.46	
Total	3115.48187	74	42.1011063	Prob > F =	0.0000	
				R-squared =	0.5082	
				Adj R-squared =	0.4875	
				Root MSE =	4.6453	

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
price	-7.64	1.045939	-7.30	0.000	-9.725543	-5.554457
advert	12.15124	3.556164	3.42	0.001	5.060447	19.24203
a2	-2.767964	.9406241	-2.94	0.004	-4.643515	-.8924123
_cons	109.719	6.799045	16.14	0.000	96.16212	123.2759

**advert\*advert** ko'paytirish orqali yaratilgan **a2** o'zgaruvchisi ba'zan o'zaro ta'sir o'zgaruvchisi deb ataladigan oddiy misoldir. O'zaro ta'sir o'zgaruvchisi haqida fikr yuritishning eng oddiy usuli shundaki, uning bog'liq o'zgaruvchiga ta'siri boshqa o'zgaruvchiga bog'liq - ikkita o'zgaruvchi bog'liq o'zgaruvchining o'rtacha qiymatini aniqlash uchun o'zaro ta'sir qiladi. Bunday holda, reklamaning o'rtacha savdoga ta'siri reklamaning o'zi darajasiga bog'liq.

O'zgaruvchilar o'zaro ta'sir qilganda, boshqa reklama birligining marjinal ta'siri (\$1000) hisob-kitob asosida qo'lda hisoblanishi kerak. Reklama daromadiga nisbatan o'rtacha savdoning qisman hosilasini olish mumkin:

$$\frac{\partial E(SALES)}{\partial ADVERT} = \beta_3 + 2\beta_4 ADVERT$$

Marjinal ta'sirning kattaligi parametrlarga, shuningdek, reklama darajasiga bog'liq. Misolda u ikki nuqtada baholanadi,  $ADVERT=0.5$  va  $ADVERT=2$ . Buyruq quyidagicha:

*scalar me1 = \_b[advert]+2\*(.5)\*\_b[a2]*

*scalar me2 = \_b[advert]+2\*(2)\*\_b[a2]*

va natijalar:

```
. scalar list me1 me2
      me1 =  9.3832736
      me2 =  1.079383
```

### 5.6.1. Optimal reklama: parametrlarning chiziqli bo'lmagan kombinatsiyasi

Ushbu misolda reklamaning optimal darajasi aniq sof sotishni maksimal darajada oshiradi. Endi yana bir dollar sarf-xarajatlar bir dollar savdoni qo'shadigan nuqtaga qadar reklama qiladi - va bundan ortiq emas. Bu nuqtada marjinal effekt birga teng,

$$\beta_3 + 2\beta_4 ADVERT = 1$$

ADVERTni parametrlar bo'yicha yechish

$$ADVERT_0 = \frac{1 - \beta_3}{2\beta_4}$$

model parametrlarida chiziqli bo'lmagan. Reklamaning optimal darajasining izchil bahosini eng kichik kvadratlar baholarini parametrlar uchun almashtirish orqali olish mumkin. Delta usuli orqali standart xatoni baholash ba'zi hisob-kitoblarni talab qiladi, ammo Stata **nlcom** buyrug'ini o'z ichiga oladi, u baholashdan keyin parametrlarning chiziqli bo'lmagan birikmalarini, shuningdek standart xato, t-nisbat va ishonch oralig'ini hisoblaydi.

**nlcom (1-\_b[advert])/(2\*\_b[a2])**

```
. nlcom (1-_b[advert])/(2*_b[a2])
```

```
_nl_1: (1-_b[advert])/(2*_b[a2])
```

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
_nl_1	2.01434	.128723	15.65	0.000	1.757673	2.271006

Ushbu hisob-kitoblarga ko'ra, reklamaning optimal darajasi 2014 dollarni tashkil qiladi.

### 5.6.2. O'zaro ta'sirlar uchun omil o'zgaruvchilardan foydalanish

O'zaro ta'sirga ega modellarni baholash va marjinal ta'sirlarni baholashning eng yaxshi usuli bu Stata omil o'zgaruvchilari va o'rnatilgan **margins** buyrug'idan foydalanishdir. Stata-da o'zgaruvchilar bilan o'zaro ta'sir qilish juda oson, ammo siz nimani nazarda tutayotganingizni taxmin qilish uchun biroz ehtiyot bo'lishingiz kerak. Stata omil o'zgaruvchilari toifali o'zgaruvchilardan indikator o'zgaruvchilarni, toifali o'zgaruvchilar ko'rsatkichlarining o'zaro ta'sirini, toifali va uzluksiz o'zgaruvchilarning o'zaro ta'sirini va uzluksiz o'zgaruvchilarning (polinomiylar) o'zaro ta'sirini yaratadi. Ularga bir nechta boshqa buyruqlar bilan bir qatorda ko'p baholash va baholash buyruqlari bilan ruxsat beriladi

Stata-da faktor-o'zgaruvchilar o'z operatorlariga ega:

**Operator Description**

<b>i.</b>	<b>unary operator to specify indicators</b>
<b>c.</b>	<b>unary operator to treat as continuous</b>
<b>#</b>	<b>binary operator to specify interactions</b>
<b>##</b>	<b>binary operator to specify factorial interactions</b>

*Unary operator* (birlik operator)- bu faqat bitta raqam yoki o'zgaruvchiga ta'sir qiladigan matematik operatsiya uchun math-speak - operand deb ataladi. Misol sifatida raqamdan oldin paydo bo'lgan manfiy belgi bo'lishi mumkin; shunday foydalanilganda u raqamni oladi va uni manfiy qiladi. Demak, -2 ifodasida manfiy ishora 2 raqamida ishlaydigan unar operator bo'lib, uni manfiy qiladi. Ikkilik operatorlar ikkita o'zgaruvchida ishlaydi. Masalan, 1-2 ifodadagi - 1 dan 2 ni ayirib, 1 va 2 da ishlaydi.

Stata bir nechta unar operatorlarni o'z ichiga oladi, ular orasida **i.** va **c.** O'zgaruvchi nomidan oldin foydalanilganda, bu operatorlar Stata-ga o'zgaruvchiga uzluksiz (masalan, **c.price**) yoki diskret (**i.price**) sifatida qarash kerakligini ko'rsatadi. Ikkinchi holda, **i.price** PRICE ning turli qiymati uchun indikator o'zgaruvchisini yaratadi. O'zgaruvchi uzluksiz va ma'lumotlar to'plamidagi har bir qiymat noyob bo'lganligi sababli, Stata modelda foydalanish uchun N ko'rsatkich o'zgaruvchilari to'plamini yaratadi - ehtimol siz odatdagidek qilmoqchi bo'lgan narsa emas. Boshqa tomondan, ma'lumotlaringizdagi diskret o'zgaruvchilar uchun indikator o'zgaruvchilari to'plamlarini yaratishda foydali bo'ladi. Bunga ba'zi misollar ushbu qo'llanmaning boshqa joylarida keltirilgan.

Hozirgi holatda ADVERT uzluksiz (**c.advert** ga ekvivalent) va biz uni o'zi bilan o'zaro aloqada bo'lishni xohlaymiz (**c.advert**). Ikkilik operator # Stata-da ikkita omilli o'zgaruvchilarni ko'paytirish uchun ishlatiladi. Shu tarzda u ikki o'zgaruvchining o'zaro ta'sirini yaratadi. Shunday qilib,  $ADVERT^2$  uzluksiz o'zgaruvchisi **c.advert#c.advert** yordamida yaratiladi

O'zgaruvchilarni shu tarzda belgilashning kamida ikkita sababi bor. Birinchidan, ma'lumotlar to'plamiga kiritish uchun o'zaro ta'sir o'zgaruvchisini maxsus yaratishga hojat yo'q; **c.advert#c.advert** to'g'ridan-to'g'ri modelga uni **regress** buyrug'ining o'zgaruvchilar ro'yxatiga kiritish orqali qo'shilishi mumkin. Ikkinchidan, shu tarzda yaratilgan o'zgaruvchilar marjinal effektlarni hisoblash uchun Stata-ning o'rnatilgan **margins** buyrug'idan foydalanishi mumkin. Oddiy ko'paytirish \* o'rniga omil o'zgaruvchilarni ko'paytirish uchun # dan foydalanish Stataga uni ko'paytma qilgan ikkita o'zgaruvchining o'zaro ta'sir qilganligini bildiradi. Bu sizga hech qanday hisob-kitob qilmasdan marjinal effektlarni kuzatib borish imkonini beradi. Asosiysi, ushbu omil o'zgaruvchilari yordamida o'zaro ta'sirlarni yaratish va chekkalarda to'g'ri variantlardan foydalanish mumkin.

Regressiyani baholash uchun buyruq

***regress sales price advert c.advert#c.advert***

va natijalari quyidagilardan iborat:

`. regress sales price advert c.advert#c.advert`

Source	SS	df	MS			
Model	1583.39741	3	527.799136	Number of obs =	75	
Residual	1532.08446	71	21.5786543	F( 3, 71) =	24.46	
Total	3115.48187	74	42.1011063	Prob > F	= 0.0000	
				R-squared	= 0.5082	
				Adj R-squared	= 0.4875	
				Root MSE	= 4.6453	

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
price	-7.64	1.045939	-7.30	0.000	-9.725543	-5.554457
advert	12.15124	3.556164	3.42	0.001	5.060446	19.24203
c.advert#c.advert	-2.767963	.940624	-2.94	0.004	-4.643514	-.892412
_cons	109.719	6.799045	16.14	0.000	96.16212	123.2759

E'tibor bering, bu regressiya natijalari **a2** qo'lda yaratilgan va modelga kiritilgan natijalar bilan bir xil. O'zaro ta'sir muddati bo'yicha koeffitsient endi **c.advert#c.advert** sifatida aniqlanadi. To'lov marjinal effektini hisoblash uchun marjlardan foydalanish bilan birga keladi. 0,5 va 2 reklama darajalarida marjinal ta'sirni baholash olinadi:

```
. margins, dydx(advert) at(advert=(.5 2))
Average marginal effects          Number of obs   =       75
Model VCE      : OLS
Expression     : Linear prediction, predict()
dy/dx w.r.t.   : advert
1._at         : advert           =       .5
2._at         : advert           =       2
```

		dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
advert	_at						
	1	9.383273	2.636965	3.56	0.000	4.214916	14.55163
	2	1.079383	.7019353	1.54	0.124	-.2963846	2.455151

Chegaralarning sintaksisi ba'zi tushuntirishlarni talab qiladi. Birinchi variant - **dydx (advert)**. Bu qavs ichidagi o'zgaruvchining (**advert**) bir birlik o'sishining qaram o'zgaruvchining o'rtacha qiymatiga marjinal ta'sirini so'raydi. Ikkinchi variant Stata-ga **advert** ning qaysi qiymatlarida marjinal effektini baholashini aytadi. Har holda, marjlar birinchi usul bilan bir xil natijani berdi va buni deyarli hech qanday dasturlashsiz va matematikasiz qilinadi. Bonus sifatida **margins** standart xatolik, t-nisbati, ikki tomonlama p-qiymati va ishonch oralig'i taxminini ham ifodalaniadi.

### 5.7. O'zgaruvchilar orasidagi o'zaro ta'sirlar

O'zgaruvchilar orasidagi o'zaro ta'sir ko'p nomli atamalar yaratish uchun oldingi bo'limda tanishtirilgandi. Tushuncha juda umumiy bo'lib, bir o'zgaruvchining o'zgarishining bog'liq o'zgaruvchining o'rtacha qiymatiga ta'siri boshqa o'zgaruvchiga bog'liq bo'lgan har qanday vaziyatga nisbatan qo'llanilishi mumkin. Ko'rib chiqilgan asosiy model quyidagicha:

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + e$$

Inson yoshi ulg'aygan sari uning pitsaga sarflashga bo'lgan cheklov moyilligi pasayadi - bu koeffitsientning inson yoshiga bog'liqligini anglatadi.

$$\beta_3 = \beta_4 + \beta_5 AGE$$

Buni modelga almashtirish orqali ifodalaymiz:

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_4 INCOME + \beta_5 (INCOME * AGE) + e$$

Bu o'zaro ta'sir o'zgaruvchisi bo'lgan  $INCOME * AGE$  yangi o'zgaruvchini taqdim etadi. Ushbu modeldagi  $AGE$  ning birlik o'sishining marjinal ta'siri  $INCOME$ Ega bog'liq va  $INCOME$  ning o'sishining marjinal ta'siri  $AGE$ Ega bog'liq.

O'zaro ta'sir Stata-da **generate** buyrug'i yordamida yaratilishi mumkin, ammo omil o'zgaruvchilardan foydalanish yaxshiroq tanlovdir. Bu o'rnatilgan **margins** buyrug'i yordamida marjinal effektlarni baholashga imkon beradi.

use pizza4, clear

regress pizza age income c.age#c.income  
 margins, dydx(age) at(income=(25 90))

Regressiyadan olingan natija:

```
. regress pizza age income c.age#c.income
```

Source	SS	df	MS			
Model	367043.25	3	122347.75	Number of obs =	40	
Residual	580608.65	36	16128.0181	F( 3, 36) =	7.59	
Total	947651.9	39	24298.7667	Prob > F =	0.0005	
				R-squared =	0.3873	
				Adj R-squared =	0.3363	
				Root MSE =	127	

pizza	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	-2.977423	3.352101	-0.89	0.380	-9.775799	3.820952
income	6.979905	2.822768	2.47	0.018	1.255067	12.70474
c.age# c.income	-.1232394	.0667187	-1.85	0.073	-.2585512	.0120725
_cons	161.4654	120.6634	1.34	0.189	-83.25131	406.1822

va margins buyrug'idan esa:

```
. margins, dydx(age) at(income=(25 90))
```

Average marginal effects  
 Model VCE : OLS  
 Expression : Linear prediction, predict()  
 dy/dx w.r.t. : age

Number of obs = 40

1.\_at : income = 25  
 2.\_at : income = 90

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
age						
_at						
1	-6.058407	2.390502	-2.53	0.011	-10.74371	-1.373109
2	-14.06896	4.171058	-3.37	0.001	-22.24409	-5.893842

Yana bir misol, *cps\_small.dta* ma'lumotlar to'plamidan foydalanadigan 4-bobda ko'rib chiqilgan ish haqi modeliga ta'lim va tajriba o'rtasidagi o'zaro ta'sirni ko'rib chiqamiz.

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4(EDUC * EXPER) + e$$

use cps4\_small, clear

gen lwage = ln(wage)

regress lwage educ exper c.educ#c.exper

```
. regress lwage educ exper c.educ#c.exper
```

Source	SS	df	MS			
Model	65.5449479	3	21.848316	Number of obs =	1000	
Residual	271.235783	996	.272325084	F( 3, 996) =	80.23	
Total	336.780731	999	.337117849	Prob > F =	0.0000	
				R-squared =	0.1946	
				Adj R-squared =	0.1922	
				Root MSE =	.52185	

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.0949385	.0146246	6.49	0.000	.06624	.123637
exper	.0063295	.0066985	0.94	0.345	-.0068153	.0194743
c.educ# c.exper	-.0000364	.0004838	-0.08	0.940	-.0009858	.0009129
_cons	1.392318	.2066447	6.74	0.000	.986809	1.797827

va - *regress lwage educ exper c.educ#c.exper c.exper#c.exper*

```
. regress lwage educ exper c.educ#c.exper c.exper#c.exper
```

Source	SS	df	MS			
Model	82.3591698	4	20.5897924	Number of obs =	1000	
Residual	254.421562	995	.255700062	F( 4, 995) =	80.52	
Total	336.780731	999	.337117849	Prob > F =	0.0000	
				R-squared =	0.2445	
				Adj R-squared =	0.2415	
				Root MSE =	.50567	

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.1271953	.0147188	8.64	0.000	.0983118	.1560789
exper	.0629807	.0095361	6.60	0.000	.0442676	.0816938
c.educ# c.exper	-.0013224	.0004949	-2.67	0.008	-.0022935	-.0003513
c.exper# c.exper	-.0007139	.000088	-8.11	0.000	-.0008867	-.0005412
_cons	.5296774	.2267415	2.34	0.020	.084731	.9746237

Yana bir bor ta'kidlaymizki, faktor o'zgaruvchilari uchun Stata unary operator laridan foydalanish bu regressiyani baholashni oson qiladi.

### 5.8. Modellar orasidagi moslikni baholash

Biz 4-bobda ko'rganimizdek, regressiyaning mosligi dispersiya jadvali tahlilidan olingan statistik ma'lumotlarga asoslanadi.  $R^2$  regressiya hisobiga bog'liq bo'lgan bog'liq o'zgaruvchidagi namunaviy o'zgarishlar ulushini o'lchaydi. Xuddi shu narsa ko'p regressiya modeli uchun ham amal qiladi. Yagona farq shundaki, endi modelda bir nechta tushuntiruvchi o'zgaruvchilar mavjud. Shunday qilib, 4-bobdagi munosabatlardan foydalanish mumkin. Erksiz o'zgaruvchi (SST) kvadratlarining umumiy yig'indisini (SSR) ichida ifodalaniladi va tushuntirilmagan (SSE) o'zgarishlarga ajrating.

$$SST = SSR + SSE$$

Keyin moslik yaxshiligi o'lchovi  $R^2$  bo'ladi

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

Principle of Econometrics adabiyotida ishlab chiqilgan va bu yerda qo'llaniladigan yozuvlar bilan Stata-ning dispersiya jadvalini tahlilidan foydalanganda ehtiyot bo'lishingiz kerak. Stata kvadratlarining namunaviy yig'indisi sifatida tushuntirilgan o'zgarishlarga ishora qiladi. Matn buni kvadratlar yig'indisi regressiya yoki SSR deb ataydi. Bog'liq o'zgaruvchining tushunarsiz o'zgarishi eng kichik kvadrat qoldiqlarida olinadi. Sizing darslingizda bu kvadrat xatolar yig'indisi, SSE deb ataladi. Stata bu tushunarsiz o'zgarishlarni kvadratlarining qoldiq yig'indisi deb ataydi.

Concept	POE	Stata
Unexplained variation	SSE (sum of squared errors)	Residual SS
Explained variation	SSR (sum of squared regression)	Model SS
Total variation	SST	Total SS

Bu yo'ldan chiqib, siz  $R^2$  ni bir qancha usullar bilan olishingiz mumkin. Birinchisi, dispersiya jadvalining tahlilidan foydalanish va uni qo'lda hisoblash. Ikkinchisi, uni



regressiya natijasidan to'g'ridan-to'g'ri olish mumkin, chunki u chiziqli regressiya uchun hisoblangan natijalarda o'z ifodasini topadi. Uchinchi - oldingi regressiyadan saqlangan natijalar yordamida hisoblash uchun buyruq satridan foydalanish. Albatta, bu variant birinchisidan osonroq, lekin ikkinchisidan qiyinroq! Shunday qilib, Stata **e(mss)** da saqlaydigan kvadratlarning model yig'indisi va **e(rss)** da saqlangan kvadratlarning qoldiq yig'indisidan foydalanib, biz  $R^2$  ni hisoblaymiz va ko'rsatamiz. E'tibor bering, buyruq bo'yicha hisoblangan  $R^2$  ham Stata xotirasida saqlanadi va uni **e(r2)** yordamida qayta ifodalashimiz mumkin.

*use andy, clear*

*reg sales price advert*

```
. reg sales price advert
```

Source	SS	df	MS			
Model	1396.53893	2	698.269465	Number of obs =	75	
Residual	1718.94294	72	23.8742075	F( 2, 72) =	29.25	
Total	3115.48187	74	42.1011063	Prob > F =	0.0000	
				R-squared =	0.4483	
				Adj R-squared =	0.4329	
				Root MSE =	4.8861	

sales	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
price	-7.907854	1.095993	-7.22	0.000	-10.09268	-5.723032
advert	1.862584	.6831955	2.73	0.008	.500659	3.22451
_cons	118.9136	6.351638	18.72	0.000	106.2519	131.5754

*di "R-square " e(mss)/(e(mss)+e(rss))*

*di "R-square " 1-e(rss)/(e(mss)+e(rss))*

```
. di "R-square " e(mss)/(e(mss)+e(rss))
R-square .44825776

. di "R-square " 1-e(rss)/(e(mss)+e(rss))
R-square .44825776
```

$R^2$  ning moslik mezoni sifatidagi kamchiliklari shundaki, modelga regressorlarni qo'shish har doim moslikni yaxshilaydi degan tushunchadan yiroqdir. Tegishli moslik o'lchovi regressorlarni qo'shish uchun kichik o'zgarishni qo'llaydi, modelga aloqador bo'lmagan mustaqil o'zgaruvchilar qo'shilganligi sababli qayta hisoblangan  $R^2$  kichikroq bo'lishi mumkin. Hisoblangan  $R^2$

$$\bar{R}^2 = 1 - \frac{SSE/(N - K)}{SST/(N - 1)}$$

Ushbu statistika Stata **regress** buyrug'i tomonidan buyruq bo'yicha ham xabar qilinadi. E'tibor bering, sozlangan R-squared **R**-kvadratidan kichikroq. Aslida, R-kvadrat hech qachon ko'p chiziqli regressiyada qayta hisoblangan R-squared dan kichik bo'lmaydi, chunki taxminiy moslik tushuntirish o'zgaruvchilarni qo'shish uchun ifodalaniladi.

## V bob mavzularini mustahkamlash uchun savollar

1. Ko'p chiziqli regressiya va Oddiy chiziqli regressiyaning farqli jihatlarini aniqlang.
2. Tenglamaning error (xato) dispersiyasini baholash uchun qanday formuladan foydalaniladi?
3. taxminiy ishonch oralig'i uchun pastki va yuqori so'nggi nuqtalarni yaratish uchun qanday buyruq amalga oshiriladi?

4. Ikki tomonlama muqobilga nisbatan  $\beta_2=0$  bo'lgan testning t nisbati qaday bo'ladi?
5. Muqobilga qarshi  $\beta_2 \geq 0$  degan nol gipotezani sinab ko'rish uchun qanday formuladan foydalanish kerak?
6. Agar p-qiymati testning 5% ahamiyatlilik darajasidan kam bo'lsa qanday amal bajarish talab etiladi?
7. Chiziqli parametirni tekshirishda lincom buyrug'i qanday amal bajaradi?
8. parametrlarning chiziqli bo'lmagan kombinatsiyasi haqida nimalarni bilib oldingiz?
9. O'zaro ta'sirlar uchun omil o'zgaruvchilardan foydalanish qanday buyruqlar orqali bajariladi?
10. O'zgaruvchilarni c.advert#c.advert tarzida belgilashning sabablarini tushuntiring.
11. O'zaro ta'sir Stata-da yaratish qaysi buyruq orqali amalga oshiriladi?
12. Modellar orasidagi moslikni baholashni tusuntirib bering.