4 Gaia

Murrizketa linealen kontrasteak eta aurresanak

Aurkibidea

4.1 Sarrera

Gai honetan, koefizienteen edozein murrizketa linealen kontrastea egiteko estatistiko orokor bat aztertuko dugu. Aurreko gaietan azaldutako esanguratasun kontrasteak izaten direnez ekonometrian gehien erabiltzen direnak, erregresioak egiten dituen edozein programa informatikok automatikoki kontraste horiek egiteko informazioa ematen du. Hala ere, badira beste kontraste interesgarri batzuk ere, ekonomilariek portaera ekonomikoari buruzko teoriak garatu eta ebaluatzen baitituzte eta hipotesien kontrasteak, teoria hauek ebaluatzeko prozedurak dira.

Atal honetan ikusiko dugun bezala, estatistiko bakar batekin koefizienteen murrizketa linealen multzo bat kontrasta daiteke. Jakin, aurretik ikusitako banakako esanguratasun kontrasterako estatistikoa eta baterako esanguratasun kontrastekoa, jarraian lortuko dugun estatististikotik eratorriak izan daitezkeela. Aldi berean, Gretl erabiliz kontraste hauek nola egin aztertuko dugu.

4.2 Murrizketa linealen kontraste orokorra

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_K X_{Ki} + u_i \quad i = 1, 2, \ldots, N$$

Aurreko erregresio linealeko eredu orokorrean, koefizienteen edozein q murrizketa lineal ($R\beta = r$ adierazpen matriziala izanik) kontrastatu nahi izanez gero: ezagunak diren konstanteen matrizea R ($q \times K$ ordenakoa) eta konstante ezagunen bektorea r ($q \times 1$ ordenakoa) definitu behar ditugu kontrastatu nahi diren murrizketen arabera. Hipotesi hutsan murrizketak matrizialki ($R\beta = r$) jarri behar direnez, jarraian adibide batzuen bitartez adierazten da nola osatu R eta r:

1. X_2 aldagai azaltzailearen banakako esanguratasun kontraste batean:

$$H_0 : \beta_2 = 0 \quad \text{vs} \quad H_a : \beta_2 \neq 0$$

$$R\beta = r \quad \Rightarrow \quad \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}}_{\beta} = \underbrace{(0)}_{r}$$

Hipotesi hutsa $H_0 : \beta_2 = 8$ bada, aldatu beharreko gauza bakarra $r = (8)$ izango litzateke.

2. Baterako esanguratasun kontraste batean: $H_0 : \beta_2 = \beta_3 = \ldots = \beta_K = 0$

$$R\beta = r \quad \Rightarrow \quad \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & & 1 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}}_{\beta} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{r}$$

3. Koefizienteen konbinazio linealen murrizketak:

   (a) Murrizketa bat: $H_0 : \beta_2 + \beta_3 = 7 \quad \text{vs} \quad H_a : \beta_2 + \beta_3 \neq 7$

$$R\beta = r \quad \Rightarrow \quad \underbrace{\begin{bmatrix} 0 & 1 & 1 & 0 & \cdots & 0 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{bmatrix}}_{\beta} \underbrace{(7)}_{r}$$

   (b) Murrizketa bat baino gehiago:

$$H_0 : \begin{cases} 5\beta_2 + 4\beta_3 & = & 5 \\ \beta_1 - 4\beta_4 & = & 6 \end{cases} \quad \text{vs} \quad H_a : \begin{cases} 5\beta_2 + 4\beta_3 & \neq & 5 \\ \beta_1 - 4\beta_4 & \neq & 6 \end{cases} \quad \text{edota}$$

$$R\beta = r \quad \Rightarrow \quad \underbrace{\begin{bmatrix} 0 & 5 & 4 & 0 & \cdots & 0 \\ 1 & 0 & 0 & -4 & \cdots & 0 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{bmatrix}}_{\beta} \underbrace{\begin{bmatrix} 5 \\ 6 \end{bmatrix}}_{r}$$

Teoriara bueltatuz, ikusi dugunez, oinarrizko hipotesien menpean KTA estimatzailearen banaketa $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2(X'X)^{-1})$ da. Adierazitako murrizketa linealen kontrasteak egiteko, ordea, $R\hat{\beta}$-ren banaketan oinarrituko gara:

$$\underset{(q \times 1)}{R\hat{\beta}} \sim \mathcal{N}(\underset{(q \times 1)}{R\beta}, \underbrace{\sigma^2 R(X'X)^{-1}R'}_{(q \times q)}) \tag{4.1}$$

Estatistiko honekin eta $\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{N-K}$ estimatzailearekin, Wald-en F estatistiko bezala ezagutzen den hurrengo estatistikoa osatuko genuke:

$$\boxed{F = \frac{\left(R\hat{\beta} - r\right)' \left(R\,(X'X)^{-1}\,R'\right)^{-1} \left(R\hat{\beta} - r\right)/q}{\hat{\sigma}^2} \overset{H_0}{\sim} \mathcal{F}_{(q, N-K)}} \tag{4.2}$$

H_0 ez baldin bada egiazkoa, $(R\hat{\beta} - r)$ diferentzia handia izango denez, F estatistikoak balio altuak hartuko ditu eta H_0 baztertuko genuke α esangura-mailarekin. Hau da, hipotesi hutsa baztertuko dugu baldin eta $F > \mathcal{F}_{(q,N-K)\alpha}$ bada eta ez dugu baztertuko bestelako kasuan.

Bada ordea estatistiko honen baliokidea den beste estatistiko bat: murriztu gabeko ereduko hondar karratuen batura eta murriztutako ereduko hondar karratuen baturan oinarritzen den estatistikoa. Ikus dezagun beste adierazpen hau lortzeko burutu behar diren pausuak.

4.3 Murrizketei baldintzaturiko Karratu Txikienen Arruntetako estimatzailea

Erregresio linealeko eredu orokorra izanik (matrizialki $Y = X\beta + u$), ereduko koefizienteei buruzko informazioa eskuragarri baldin bada, oso garrantzitsua da informazio berri hori egiazkoa den edo ez kontrastatzea zeren β ezezaguna denez, ez dakigu ziurtasunez $R\beta = r$ ematen den. Kontrastearen ondorioz, hau da, hipotsi hutsa ez bada baztertzen, informazioa egiazkoa dela ondorioztatuko dugu α esanguratasun mailarekin eta murrizketa kontuan hartuta lortuko genukeen estimatzaile berria (Murriztutako Karratu Txikienen Arruntetako estimatzailea ($\hat{\beta}_{MKTA}$ edo $\hat{\beta}_M$)), KTA estimatzailea baino efizienteagoa izango da, bariantza txikiagoa izango baitu.

Ereduko koefizienteak murrizketa hauek kontuan hartuta estimatu nahi izanez gero ($\hat{\beta}_M$), murrizketa hauei baldintzatutako Hondar Karratuen Batura minimo egin beharko genuke. Minimizazio prozeduraren ondorioz eta zenbait eragiketa eginik, honako emaitza lortuko genuke:

$$\hat{\beta}_M = \hat{\beta}_{KTA} + (X'X)^{-1}R'\big[R(X'X)^{-1}R'\big]^{-1}(r - R\hat{\beta}_{KTA}) \tag{4.3}$$

non aurreko atalean ikusi dugun bezala, koefizienteen edozein murrizketa multzo $R\beta = r$ bezala adierazten dugun eta R eta r matrizeak ezagunak diren, murrizketen arabera osatuak izanik.

KTA estimatzailearekin bezala, posible da KTM estimatzailearekin Hondar Karratuen Batura edota mugatze koefizientea lortzea:

$$HKB_M = \hat{u}'_M\hat{u}_M = (Y - X\hat{\beta}_M)'(Y - X\hat{\beta}_M) \quad \text{eta} \quad R^2_M = 1 - (\hat{u}'_M\hat{u}_M / \textstyle\sum(Y_t - \bar{Y})^2)$$

Dena den, bide hau jarraituz estimazioak lortzeko egin beharreko eragiketa kopurua handia da eta eskuz egin beharko genituzke. Bada ordea, Gretl erabiliz $\hat{\beta}_M$ lortzeko eta bide batez, HKB_M eta R^2_M lortzeko beste metodo alternatibo bat: $R\beta = r$ murrizketak ereduan ordezkatuz, **murriztutako eredua** lortuz ($Y_* = X_*\beta_* + u_*$) eta bertako koefizienteak KTA bitartez estimatuz ($\hat{\beta}_* = (X'_*X_*)^{-1}X'_*Y_*$). Horrela, hasierako ereduko (**murriztu gabeko eredua**) koefiziente batzuk bakarrik estimatuko genituzkeenez, gainontzekoak lortzeko murrizketetara bueltatuko ginateke.

Aukera hau adibide batekin ikusiko dugu. Izan bedi ondorengo erregresio eredua:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad i = 1, 2, \ldots, N$$

non $\beta_2 = 5$ dela dakigun. Informazio hau ereduan barneratzen badugu, murriztutako eredua lortuko dugu. Aipatu bezala, eredu honetan Hondar Karratuen Batura minimo eginez, bertako koefizienteak estimatu ahal izango ditugu, hauek murrizketak betetzen dituztelarik.

Eredu honetan ordea, ez daude murriztu gabeko ereduko koefiziente guztiak eta murrizketetara zuzendu beharko gara guztiak lortzeko. Ikus dezagun:

Hasierako eredua: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

Murriztutako eredua:
$$\begin{aligned} Y_i &= \beta_1 + \beta_2 X_{2i} + 5X_{3i} + u_i^* \\ Y_i - 5X_{3i} &= \beta_1 + \beta_2 X_{2i} + u_i^* \\ Y_i^* &= \beta_1 + \beta_2 X_{2i} + u_i^* \end{aligned}$$

non $Y_i^* = Y_i - 5X_{3i}$ den. Hondar Karratuen Batura minimo eginik, β_1 eta β_2 estimatu ahal izango ditugu $\hat{\beta}_1^M$ eta $\hat{\beta}_2^M$ lortuz. Gainontzekoak murrizketetara bueltatuz lortuko genituzke. Horrela, murriztutako estimatzailea $\hat{\beta}_M = \begin{bmatrix} \hat{\beta}_1^M & \hat{\beta}_2^M & 5 \end{bmatrix}'$ izango da.

Eredu berdinarekin jarraituz, beste adibide bat ondorengoa litzateke: $\beta_1 + 3\beta_2 = 4$ murrizketari loturikoa edo baliokidea dena, $\beta_1 = 4 - 3\beta_2$.

Hasierako eredua: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

Murriztutako eredua:
$$\begin{aligned} Y_i &= \beta_1 + (4 - 3\beta_3)X_{2i} + \beta_3 X_{3i} + u_i^* \\ \underbrace{Y_i - 4X_{2i}}_{Y_i^*} &= \beta_1 + \beta_3(\underbrace{X_{3i} - 3X_{2i}}_{X_i^*}) + u_i^* \end{aligned}$$

Azken eredu honetan $\hat{\beta}_1^M$ eta $\hat{\beta}_3^M$ estimatzen dira eta ondoren $\hat{\beta}_2^M = 4 - 3\hat{\beta}_3^M$ lortzen da.

Murriztutako eredua estimatzerakoan, erregresio horretan lortuko genukeen Hondar Karratuen Batura ($HKB_* = \hat{u}_*'\hat{u}_* = Y_*'Y_* - \hat{\beta}_*'X_*'Y_*$) eta lehen lortutako HKB_M berdinak dira; hau da, hasierako ereduan murriztutako estimatzailearekin lortutako HKB eta murriztutako eredukoak berdinak dira. Bestalde, murriztutako ereduko aldagai azaldua eta murriztu gabekoa berdinak badira, murriztutako ereduko mugatze koefizientea (R_*^2) eta lehenago definitu duguna KTM estimatzailearekin (R_M^2), berdinak dira. Guzti hau jakinik eta Gretl erabiliz, eragiketak asko errazten dira, baina lehenik, murriztutako ereduko aldagai berrien behaketak kalkulatu beharko genituzke eta aurreko gaietan ikusitakoa aplikatu: aldagai berriak definitu eta murriztutako ereduan KTA aplikatu.

Esan bezala, aipatutako HKB hauek erabiliz, posible da ere koefizienteen murrizketa linealen kontrastea burutzea, hau da, posible da lehen adierazitako estatistikoaren beste adierazpide baliokide bat lortzea: halaber, hipotesi hutsagatik murriztutako ereduko HKBn ($\hat{u}_M'\hat{u}_M$) eta murriztu gabeko ereduko HKBn ($\hat{u}'\hat{u}$), non $\hat{u}_M = Y - X\hat{\beta}_M$ eta $\hat{u} = Y - X\hat{\beta}_{KTA}$ diren, oinarrituriko estatistikoa ondorengoa da:

$$\boxed{F = \frac{(\hat{u}_M'\hat{u}_M - \hat{u}'\hat{u})/q}{\hat{u}'\hat{u}/(N-K)} \overset{H_0}{\sim} \mathcal{F}_{(q,N-K)}} \tag{4.4}$$

non $\hat{u}_M'\hat{u}_M = (Y - X\hat{\beta}_M)'(Y - X\hat{\beta}_M)$ edota murriztutako ereduan ($Y_* = X_*\beta_* + u_*$), $\hat{u}_*'\hat{u}_* = Y_*'Y_* - \hat{\beta}_*'X_*'Y_*$. Lehen bezala, hipotesi hutsa baztertuko dugu baldin eta $F > \mathcal{F}_{(q,N-K)\alpha}$ bada eta ez dugu baztertuko bestelako kasuan.

Honen baliokidea den beste estatistiko bat mugatze koefizienteetan oinarritzen dena da. Alde batetik, murriztu gabeko eredua KTA bitartez estimatuz lortuko genukeena ($R^2 = R^2_{MG}$) eta bestalde, KTM bitartez lortuko genukeena (R^2_M) (azken hau murriztutako ereduko mugatze koefizientearen (R^2_*) berdina da, murrizketa kontuan hartuta aldagai azaldua ez bada aldatzen, hau da $Y = Y_*$ denean, $R^2_M = R^2_*$ izango da). Honakoa da aipaturiko estatistikoa:

$$F = \frac{(R^2_{MG} - R^2_M)/q}{(1 - R^2_{MG})/(N-K)} \overset{H_0}{\sim} \mathcal{F}_{(q,N-K)} \tag{4.5}$$

Kontrastea aurreko estatistikoarekin bezala burutzen delarik.

Azken estatistiko hau aprobetxatuz, aldagai azaltzaileen baterako esanguratasun kontrastea aztertu genuenean proposatzen genuen estatistikoa gogoratuko dugu. Erregresio lineal eredu orokor batean izango genukeen hipotesi hutsa

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_K = 0$$

izango litzateke eta honi atxikitutako murrizturiko eredua $Y_t = \beta_1 + u_t$ izango da non $R^2_* = 0 = R^2_M$ izango den, horrelako kontraste baterako proposatzen genuen estatistikoa justifikatuz:

$$F = \frac{R^2/(K-1)}{(1-R^2)/(N-K)} \overset{H_0}{\sim} \mathcal{F}_{(K-1,N-K)} \tag{4.6}$$

4.4 Murrizketa linealen kontrasteak Gretl erabiliz

Murrizketa linealen kontrasteak formalki nola egingo genituzkeen azaldu dugu baina zenbait programa informatikok, besteak beste Gretl, estatistiko hauen balioa lortzeko aukera eskeintzen du (Wald-en testa), murrizketen kontrasteak egiteko aukera asko erreztuz.

Erregresio lineal eredu orokor batean eta Gretl programa erabiliz, koefizienteak nola adierazten ditugun eta programak nola adierazten dituen kontuan hartu beharra dago. Gretl programan, guk β_1 bezala definitu dugun koefizientea (hau da termino konstantea) $b1$ da, gure β_2, $b2$ izango da, β_3 berriz $b3$ eta antzera ereduko koefiziente guztiekin. Hau da:

Eredu orokorra: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_K X_{Ki} + u_i$

Gretl programan: $Y = b1 + b2 * X_2 + b3 * X_3 + \ldots + bK * X_K + u$

Horrela, koefizienteen edozein murrizketa ($R\beta = r$) adierazterakoan, kontuan izan beharko dugu deitura desberdintasunak eta biderkaketak nola adierazten diren. Adibidez, $\beta_1 + 5\beta_3 = 4$ eta $\beta_2 + 4,5\beta_4 = 2$ murrizketak, Gretlekin $b1 + 5 * b3 = 4$ eta $b2 - 4,5 * b4 = 2$ idatzi behar ditugu. Murrizketa bakoitza ekuazio bakar batean idatzi behar dugu, berdintzaren ezker aldean parametroen konbinazio lineala adieraziz eta eskubian dagokion zenbakia.

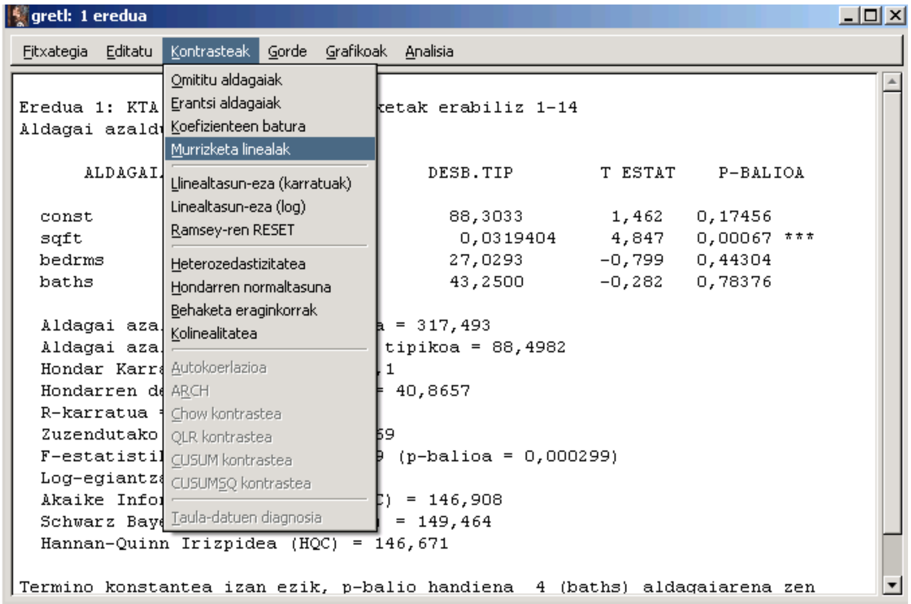
Etxebizitzen prezioen adibideko ereduan[1]

$$PRICE_i = \beta_1 + \beta_2 SQFT_i + \beta_3 BEDRMS_i + \beta_4 BATHS_i + u_i$$

ondorengo irudian azaltzen den bezala, eredua KTA bitartez estimatu ondoren lortzen den lehiatilan kontrasteak egin daitezke: *Kontrasteak* → *Murrizketa linealak*.

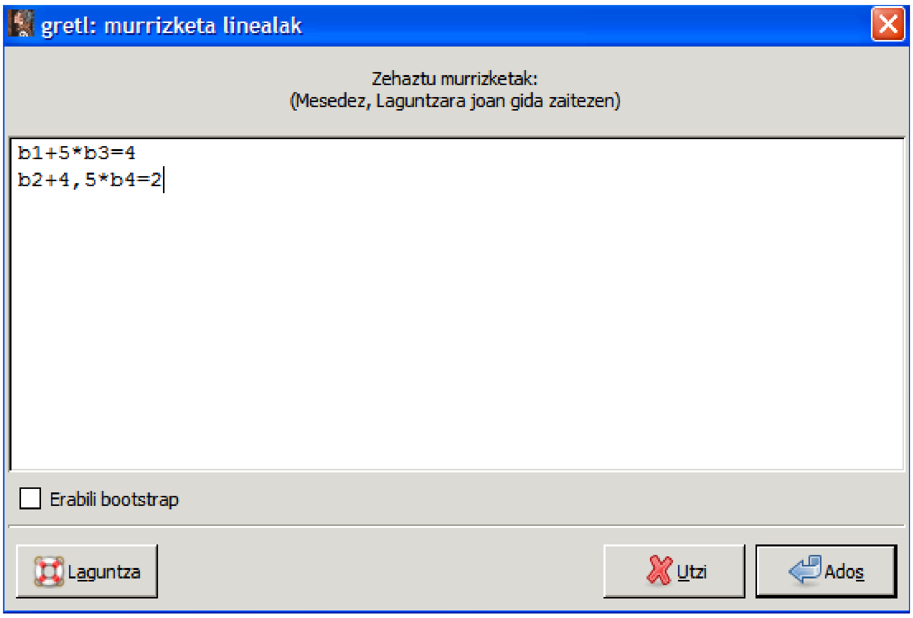
[1] Ramanathaneko (2002) data4-1.gdt datu-fitxategia.

4.1 Irudia: Murrizketa linealen kontrasteak



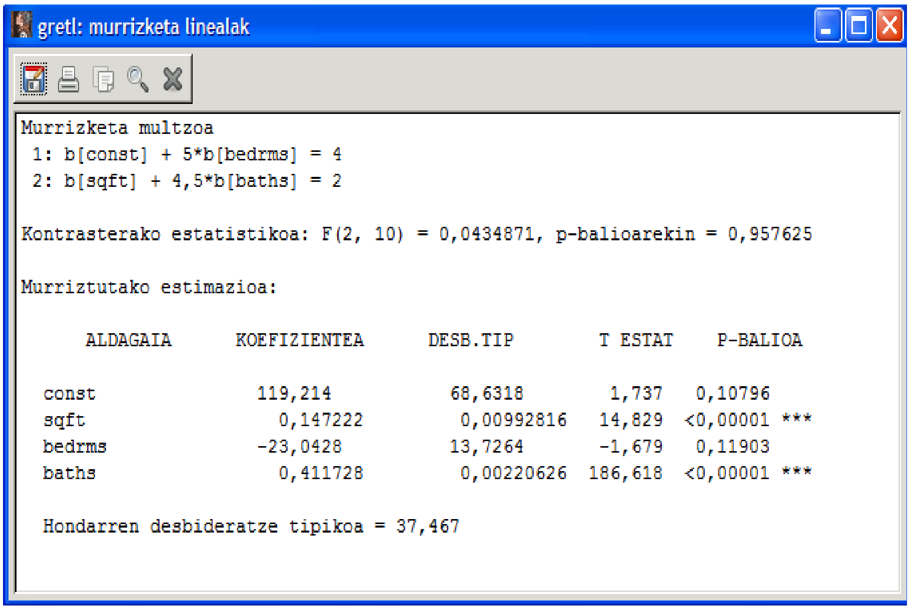
Ondoren interesatzen zaizkigun murrizketak adieraziko ditugu:

4.2 Irudia: Murrizketa linealen zehazpena



Ados klikatzean, murrizketak kontuan hartuz lortzen diren murriztutako estimazio emaitzak agertzen dira. Hau da Murriztutako Karratu Txikienen estimatzailea. Bertan, Wald-en F

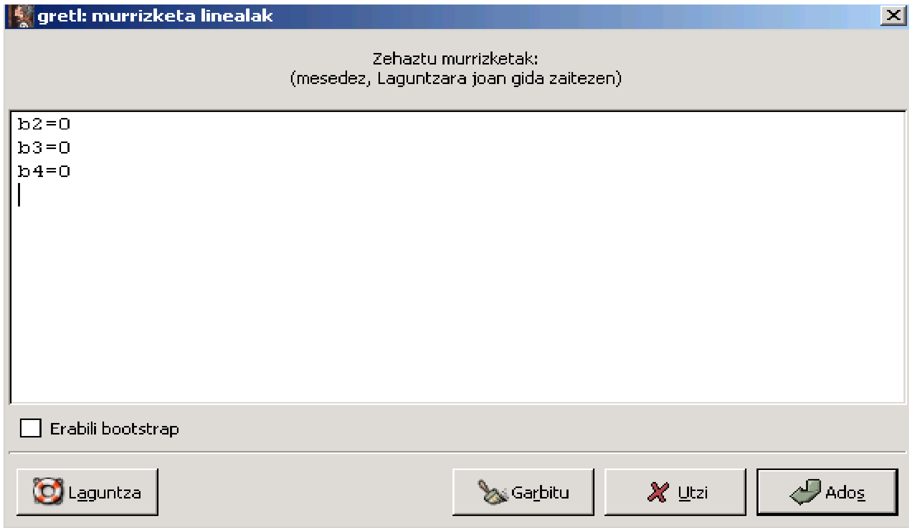
4.3 Irudia: Murrizketa linealen kontrastearen emaitza



estatitsikoaren balioa ere agertzen da (F=0,04389) eta dagokion p-balioa, kontrastea burutu ahal izateko.

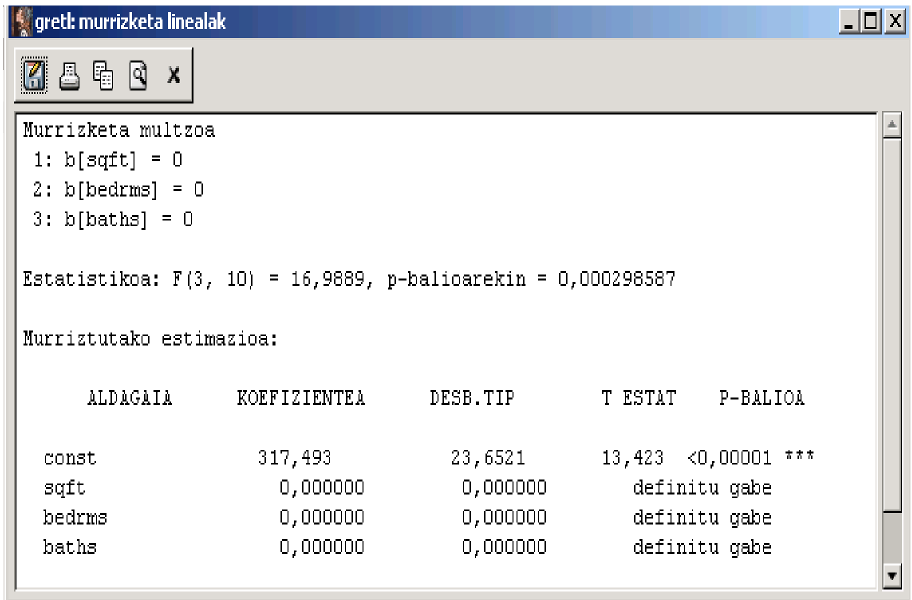
Horrela koefizienteen edozein murrizketa linealen kontrasteak burutu ditzakegunez, jada egin ditugun banakako eta baterako esanguratasun kontrasteak ere egin ditzakegu. Baterako esanguratasun kontrastean adibidez, horrela adieraziko genuke:

4.4 Irudia: Baterako esanguratasunaren murrizketen zehazpena



Ondorengo emaitza lortuz:

4.5 Irudia: Baterako esanguratasun kontrastearen emaitza



Bertan ikusten dugun bezala, oraingoan proposatzen dugun Wald-en F estatistikoaren balioa 16,9889 da, hirugarren gaian ikusi genuen bezala. Horregatik, kontrastea era berean egin dezakegu, estatistikoaren balioarekin batera, alboan dagokion p-balioa agertzen baita. Dakusagunez, $F = 16,9889 > \mathcal{F}_{(3,10)0,05} = 3,71$ da eta hipotesi hutsa baztertu egiten da $\alpha = 0.05$eko esangura-mailarekin, ondorioz aldagai azaltzaileak batera nabariak direlarik. Bide batez, p-balioa erabiliz, $\alpha = 0,01$ balioa baino txikiagoa denez, hipotesi hutsa baztertu egingo dugu $\alpha = 0,01$eko esangura-mailarekin eta baita $\alpha = 0,1$ eta $\alpha = 0,05$ekin.

4.5 Puntuzko eta tartezko aurresanak

Ekonometriaren helburu nagusia ereduaren estimazio on baten lorpena dela pentsatzen bada ere, sarritan aurresan zehatzak lortzea ere oso garrantzitsua da. Eredua zuzenki zehaztu ondoren, parametroak estimatu ditugu eta kontrasteak eginez, ereduari oniritzia eman diezaiokegu edo ez. Emaitza ezezkoa bada, eredua berriro zehaztu beharko dugu eta pausu guztiak berregin. Eredua onartuz gero, aurresanak egiteko erabili dezakegu edota "zein izango litzatekeen aldagai azalduaren balioa aldagai azaltzaileek balio konkretu bat hartzen badute"motako galderei erantzun ahal izango diegu.

Suposa dezagun ondorengo eredua T behaketekin estimatu dela:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki} + u_i \quad i = 1, \ldots, N$$

hurrengo lagineko erregresio funtzioa lortuz:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \ldots + \hat{\beta}_K X_{Ki}$$

Orduan, aldagai azaltzaileen behaketa berri bat emanik,

$$X'_p = \begin{bmatrix} 1 & X_{2p} & \cdots & X_{Kp} \end{bmatrix} \qquad p \notin \{1, 2, \ldots, N\}$$

KTAn bitartez estimatutako eredua erabili daiteke aldagai azalduak izango duen balioa aurresateko (puntuzko aurresana), hau da, balio hauek lagineko erregresio funtzioan ordezkatuko genituzke, $\hat{Y}_p$ lortuz.

$$\hat{Y}_p = X'_p \hat{\beta}_{KTA}$$

Edo beste era batera:

$$\hat{Y}_p = \hat{\beta}_1 + \hat{\beta}_2 X_{2p} + \ldots + \hat{\beta}_K X_{Kp}.$$

Hala ere, aldagai azalduaren p momentuko benetako balioa jakiterakoan, errore bat egin dugula antzemango dugu (aurresanaren errorea: $e_p = Y_p - \hat{Y}_p$), aldagai azaltzaileen balioetan errorea dagoelako, β koefizienteen estimatzaileak erabili ditugulako, Y_p u_p-ren menpean dagoelako (hau da behaketa horri dagokion perturbazioaren menpe), etab. Horregatik, komenigarria izaten da tartezko aurresan bat egitea, nolabait aurresanaren zehaztasun neurri bat kontuan hartzen baitu.

Tartezko aurresana lortzeko, aurresanaren errorearen banaketan oinarrituko gara, u_p eta $\hat{\beta}$ aldagai aleatorio normalak badira, aurresanaren errorea ere normala baita:

$$e_p \sim \mathcal{N}(0, \sigma^2(\, 1 + X_p' \left(X'X \right)^{-1} X_p \,)).$$

Orokorrean σ^2 ezezaguna izaten denez, bere estimatzaile alboragabea erabiliz, ondorengo banaketa lortuko genuke:

$$\frac{e_p}{\hat{\sigma}\sqrt{1 + X_p' \,(X'X \,)^{-1} X_p}} \sim t_{(N-K)}.$$

Eta hemendik, puntuzko aurresanaren inguruan, aldagai azalduak p momentuan hartuko duen balioaren aurresanaren tarte bat lortuko genuke $1-\alpha$ konfidantza mailarekin, ondorengo adierazpena lortuz:

$$KT(Y_p)_{1-\alpha} = \left(\hat{Y}_p - \, t_{(N-K)\frac{\alpha}{2}} \, \hat{\sigma}_{e_p} \;,\; \hat{Y}_p + \, t_{(N-K)\frac{\alpha}{2}} \, \hat{\sigma}_{e_p} \right)$$

non $\hat{\sigma}^2_{e_p} = \hat{\sigma}^2(\, 1 + X_p' \,(X'X \,)^{-1} X_p \,)$ eta $X_p' = \begin{bmatrix} 1 & X_{2p} & \cdots & X_{Kp} \end{bmatrix}$ diren.

4.6 Puntuzko eta tartezko aurresanak Gretl erabiliz

Aurreko atalean aurresanak nola lortzen diren aztertu dugu baina eskuz egin nahi izanez gero, egin beharreko eragiketak asko dira. Gretlek aurresanak lortzeko aukera eskeintzen duenez, etxebizitzen prezioen ariketa erabiliko dugu Gretlen jarraitu beharreko pausuak zein diren jakiteko. Demagun honako eredua

$$PRICE_i = \beta_1 + \beta_2 SQFT_i + \beta_3 BEDRMS_i + \beta_4 BATHS_i + u_i \tag{4.7}$$

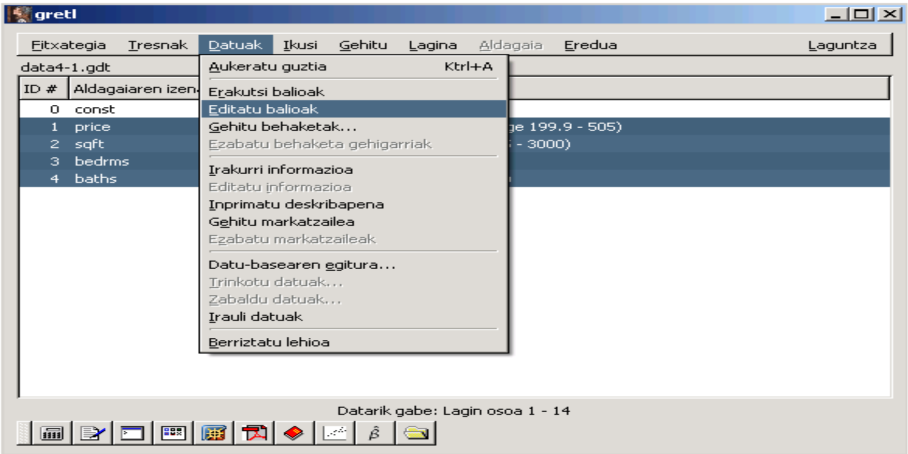
dugula eta beste etxebizitza baten ondorengo informazioa: $SQFT = 3200$, $BEDRMS = 5$ eta $BATHS = 3$ izanik, bere prezioa aurresan nahi dugula. Horretarako (4.7) eredua erabiliz ezaugarri hauetako etxebizitzaren prezioaren aurresan bat egin dezakegu eta bide batez, aztertu ea eskatzen diguten prezioa ($PRICE = 500$) arrazoizkoa den edo ez.

Gretlen bitartez aurresan hau egiteko lehen pausua datu berriak datu-basean barneratzea da: *Datuak* $\rightarrow$ *Aukeratu guztia* egin ondoren, *Datuak* $\rightarrow$ *Editatu balioak* aukeran

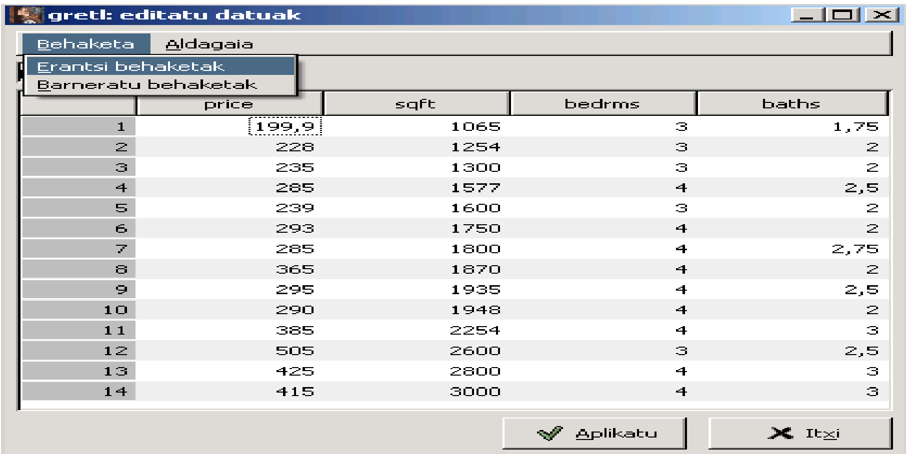
Behaketa $\rightarrow$ *Erantsi behaketak* klikatu

eta behaketa berria barneratuko dugu. Hau egiteko, barneratuko ditugun behaketa kopurua (adibidean behaketa bat) zein den adierazi behar dugu:

4.6 Irudia: Balioak editatzea



4.7 Irudia: Behaketak erantsi

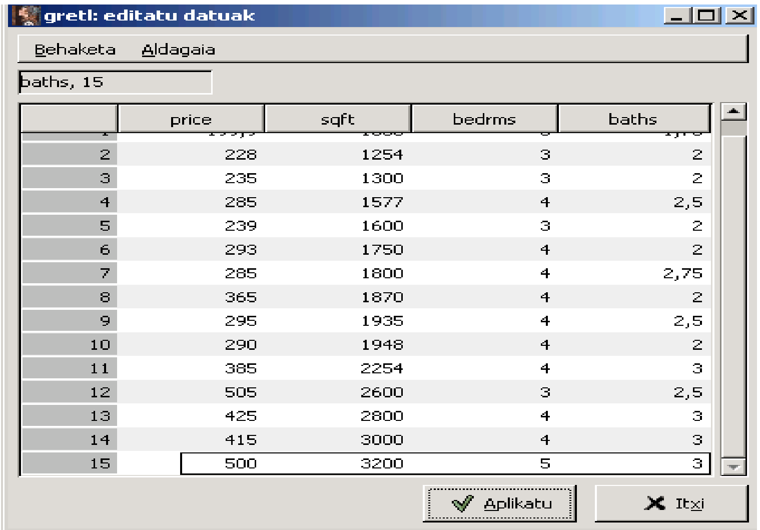


4.8 Irudia: Eransteko behaketen kopurua

eta bakoitzari dagokion lekuan ezagutzen ditugun balioak barneratzen ditudu azken errenkada osatuz. Jarraian, *Aplikatu* klikatuz.

Ondoren, kontu handia izan behar dugu, gure eredua estimatzerakoan, behaketa berri hau ez baitugu kontuan hartu behar (lagina 14 behaketez osaturik zegoen). Horregatik, *Lagina* →

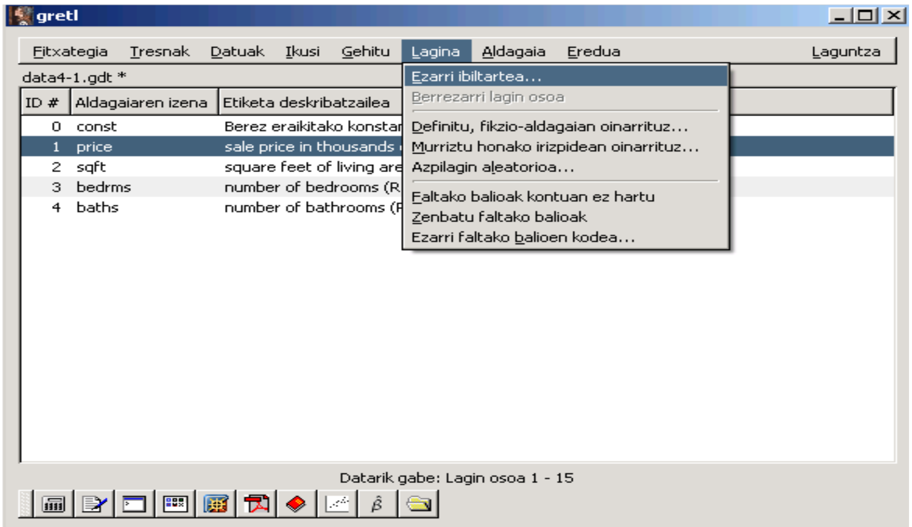
4.9 Irudia: Datu berrien ikuspena



	price	sqft	bedrms	baths
2	228	1254	3	2
3	235	1300	3	2
4	285	1577	4	2,5
5	239	1600	3	2
6	293	1750	4	2
7	285	1800	4	2,75
8	365	1870	4	2
9	295	1935	4	2,5
10	290	1948	4	2
11	385	2254	4	3
12	505	2600	3	2,5
13	425	2800	4	3
14	415	3000	4	3
15	500	3200	5	3

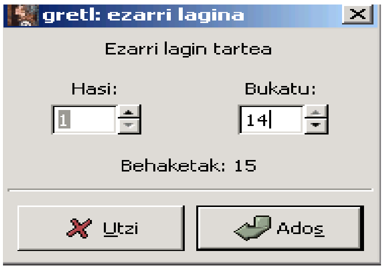
Ezarri ibiltartea aukeran

4.10 Irudia: Ibiltartearen ezarpena



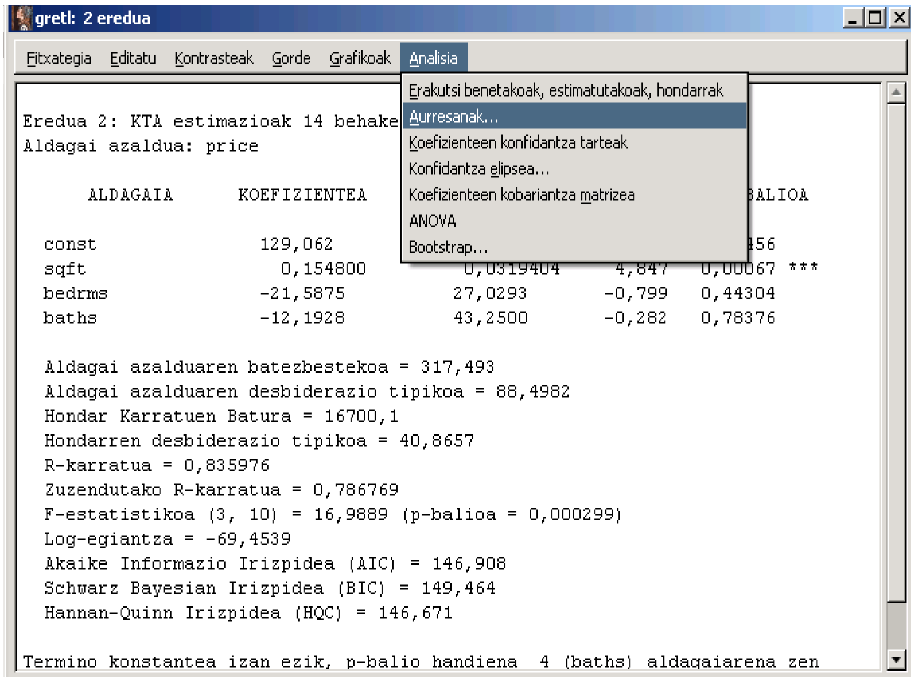
lagina nola osatzen den adieraziko dugu, gure kasuan 14 behaketez:

4.11 Irudia: Lagin tartearen ezarpena



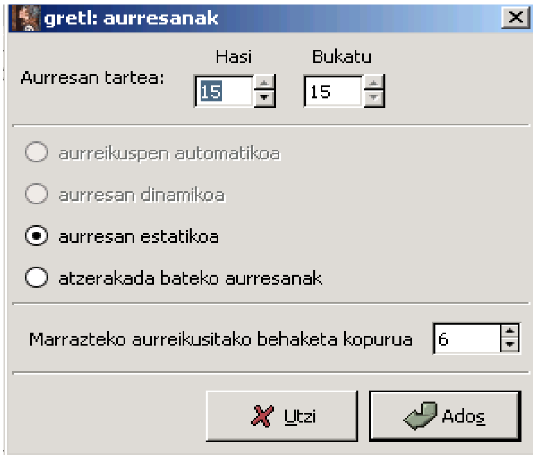
Ados klikatu ondoren, aurreko gaietan ikusi bezala, eredua KTA bitartez estimatuko dugu eta emaitzen lehiatilan *Analisia* → *Aurresanak* klikatuko.

4.12 Irudia: Aurresanak



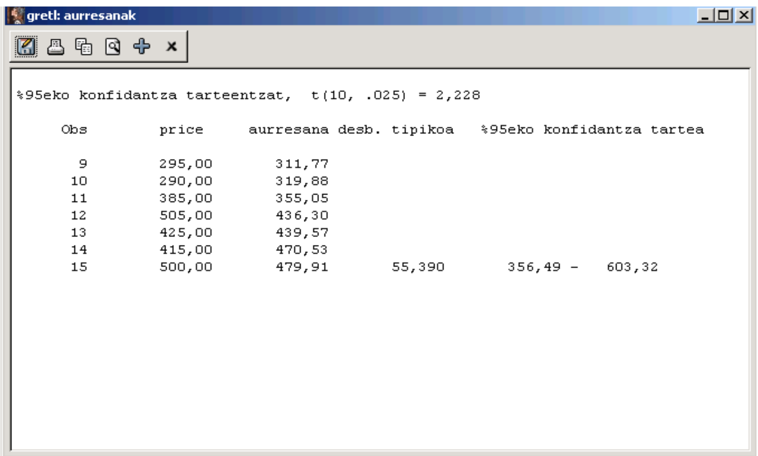
Lehiatila berrian 15. behaketa sartuko dugu, ondoren *Ados* klikatuz.

4.13 Irudia: Aurresan tartearen ezarpena



Lehiatila berrian, interesatzen zaizkigun balioak azken errenkadan daude. Hau da, ezaugarri horietako etxebizitza baten prezioaren puntuzko aurresana $479,91$ da edota %95eko konfidantza mailarekin $356,49$ eta $603,32$ bitartean egongo denez, 500 prezioa arrazoizkoa dela deritzogu.

4.14 Irudia: Puntuzko eta tartezko aurresanen balioak

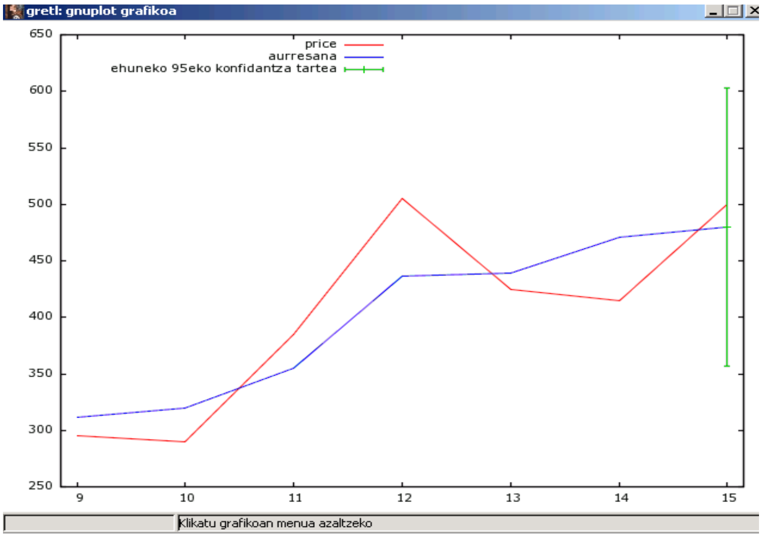

gretl: aurresanak

%95eko konfidantza tarteentzat, t(10, .025) = 2,228

Obs	price	aurresana	desb. tipikoa	%95eko konfidantza tartea
9	295,00	311,77		
10	290,00	319,88		
11	385,00	355,05		
12	505,00	436,30		
13	425,00	439,57		
14	415,00	470,53		
15	500,00	479,91	55,390	356,49 - 603,32

Aurreko emaitzarekin batera beste irudi bat agertzen da. Bertan orain arte lortutakoa laburbiltzen da, halaber, interesatzen zaigun aldagaiaren balioaren garapena (gorriz), aurresanak (urdinez) eta %95eko konfidantza mailako konfidantza tartea (berdez). Askotan lagungarria izaten da emaitzak horrela ikustea, zenbakiz osatutako taula batean baino hobeto nabari direlako aztergai dugun aldagaiaren ezaugarri nagusienak.

4.15 Irudia: Aurresanen grafikoa



Bibliografia

Ramanathan, R. (2002), *Introductory Econometrics with Applications*, 5. ed., South-Western, Ohio.