

**EREMU TXIKIEN ESTIMAZIOA EUSKAL AUTONOMIA ERKIDEGOKO
INDUSTRI INKESTAN**



**EUSKAL ESTATISTIKA ERAKUNDEA
INSTITUTO VASCO DE ESTADISTICA**

Donostia-San Sebastián, 1
01010 VITORIA-GASTEIZ
Tel.: 945 01 75 00
Faxa: 945 01 75 01
E-mail: eustat@eustat.es
www.eustat.es

AURKEZPENA

EUSTATEk jakin badaki gero eta gehiago eskatzen direla bereizituago egiten diren kalitatezko estatistikak. Horrexegatik, duela bi urte EUSTATEko sail ezberdinetako eta Unibertsitateko kideez osatutako ikerketa talde bat eratu zuen honako helburu hauekin: estatistika-eragiketetan estimazioak egiteko teknikak hobetzea eta eremu txikietan ereduetan oinarrituriko estimaziorako teknikak estatistika ofizialean sartzea.

Lan honetan, Eustat-ek eremu txikien gaineko estimazioa egiteko sistema bat definitzen duen lehenengo inkesta jasotzen da. Urteko Industri Inkestan ageri da eragiketa hori, batez ere Eustatek egindako inkesta ekonomikoen artean garrantzi handiena duena delako eta euskal ekonomian funtsezko sektorea delako.

Argitalpen honetan, Euskal Autonomia Erkidegoko 2002 eta 2003 urteetako Industri Inkestarako eskualde-estimazioak aurkezten dira. Lan-mota hauek berriak eta metodologia aldetik konplexuak direla kontuan hartuta, agiri honen helburu nagusia erabili diren estimatzaileak eta ereduak modu xehekatuan deskribatzea da, bai balio osoetarako bai batez besteko akats koadratikoetarako.

Argitalpen honen edukiaren bidez, orduan, teknika horiek Eustaten beste inkesta batzuetan aplikatzeko proiektua abiarazi nahi da eta, halaber, eremu txikietan metodoak ezagutu eta erabiltzeko interesa duten erabiltzaile guztientzat probetxugarria izango den materiala eskaini.

Vitoria-Gasteiz, 2006ko urtarrilla

JOSU IRADI ARRIETA

Zuzendari Orokorra

Aurkibidea

AURKEZPENA	1
AURKIBIDEA	3
1.- EUSKAL AUTONOMIA ERKIDEGOKO INDUSTRI INKESTA	4
1.1.- EUSKAL AUTONOMIA ERKIDEGOKO INDUSTRI INKESTAREN DESKRIPZIOA	4
1.2.- EUSKAL AUTONOMIA ERKIDEGOKO INDUSTRI INKESTAN ERABILITAKO ZENBATESLEAK	6
2.- INDUSTRI INKESTARAKO EREMU TXIKIETAN ESTIMATZEKO SISTEMA	9
2.1.- SARRERA	9
2.2.- EREDU LINEAL MISTOA	11
2.3.- ERAGIN FINKODUN EREDU LINEALA	21
2.4.- SEKTOREAREN ARABERAKO GUZTIRAKOAK	23
2.5.- KALIBRATZE-PROZESUA	24
2.6.- INDUSTRI INKESTAKO ESTIMAZIO PLANA	24
3.- EREMU TXIKIEN ESTIMAZIO TEKNIKEN APLIKAZIOA EUSKAL AEKO INDUSTRI INKESTAN. 2002 ETA 2003.	27
3.1.-ESKUALDEKAKO ESTIMAZIOAK EUSKAL AUTONOMI ERKIDEGOKO INDUSTRI INKESTAN. 2002 ETA 2003.	27
3.2.-BALIO ERANTSIA FAKTOREEN KOSTUETAN ETA LANGILE OKUPATUAK ESKUALDEKA INDUSTRI INKESTAN. 2002 ETA 2003. ONDORIOAK	32
4.- ONDORIOAK	34
BIBLIOGRAFIA	35

Euskal Autonomia Erkidegoko industri inkesta

1.1.- Euskal Autonomia Erkidegoko industri inkestaren deskripzioa

Aurrekariak

Estatistika hau 1981. urtean egin zen lehenengo aldiz. Funtsezko helburua beti izan da EAEko industriaren egitura xehetasunez ezagutzea, sektore honek garrantzi handia duelako balio erantsiari eta enpleguari dagokienez. Horretarako beharrezkoa den oinarritzko informazioa galera eta irabazien kontuko partida garrantzitsuenetatik lortzen da, partida horietako makromagnitude nagusienak estimatzean.

Estatistika-eragiketa hau Nekazaritza eta Arrantza Saileko Estatistika eta Sektore-azterketa Zerbitzuarekin lankidetzan egiten da, bera delako sail horretako estatistika-organo berezia.

Ezaugarri teknikoak

Unibertsoa: Biztanleria-eremua jarduera nagusia industriakoa duten establezimenduetara mugatzen da, betiere sortutako balio erantsiaren arabera neurtuta.

1993ko Ekonomi Jardueren Sailkapen Nazionalaren arabera (aurrerantzean 93-EJSN), hurrengo atalak ditu:

C atala: ateratze-industria;

D atala: manufakturatako industria;

E atala: energia elektrikoaren, gasaren eta uraren ekoizpena eta banaketa.

Geografia-eremua: Euskal Autonomia Erkidegoan dauden estatistika-unitateak, egoitza edo kudeatzailatza EAetik kanpo egon arren.

Denbora-eremua: erreferentziazat hartzen den aldia urte naturaleko ekitaldi ekonomikoa da. Salbuespenez, establezimenduren batek urte naturalekoak ez diren denbora-aldien gaineko kontabilitatea aurkezten badu urte horietan amaitzen diren ekitaldiei buruzko informazioa jasoko da.

Inkestaren markoa

Inkestaren testuingurua EUSTATen Jarduera Ekonomikoen Gidazerrenda da. Horri esker laginketa-akatsak mugatuko dituen probabilitate-laginketa bat egin daiteke.

Estatistika-unitatea

Estatistika-unitatea izango da unitate bezala definitzen den eta kokaleku geografiko berean egonda, soilik edo batez ere, jarduera bat edo gehiago lantzen dituen establezimendua.

Lagin-diseinua eta estrapolazioa

Probabilitate-laginketa bi fasetan egin daiteke: lehenengoan, 19 langile baino gehiago dituzten unitate guztiak hautatuko dira "bat" probabilitatearekin; bigarren fasean, ausazko laginketa estratifikatua egingo da hurrengo estratifikazio-aldagaiekin:

a) Lurralde historikoa: Araba, Bizkaia eta Gipuzkoa.

b) Jarduera: Ekonomi Jardueren Sailkapen Nazionala (EJSN-93) azpimailan, hau da, 5 digiturekin. Gero, hedapenerako EUSTATEko A84 sailkapen normalizatua erabiliko da. A84 sailkapena A60aren desagregazio bat da (EJSN-93 2 digiturekin), Euskal Autonomia Erkidegoko ekonomi egitura bereziaren arabera.

Hautatu den laginak 3.000 estatistika-unitate ditu gutxi gorabehera.

Estrapolazioa egin aurretik lagineko-establezimenduak post-estratifikatu egiten dira, hiru lurralde historikoen (Araba, Bizkaia, Gipuzkoa), 5 digituko EJSN-93 azpimailaren eta bost establezimendu-tamainen arabera. Hona hemen establezimenduen tamainak:

1 - 19 langile;
20 - 49 langile;
50 - 99 langile;
100 - 499 langile;
500 langile edo gehiago.

Lagineko datuetatik estimazioetara igarotzeko, estratu bakoitzeko estimazio-matrize bat erabiliko da. Estimazioak lortzeko erabiltzen den aldagaia industri establezimenduetan lanean dagoen pertsona kopurua da. Aldagai horren erabilera justifikatuta dago, inkestak aztertu nahi dituen aldagai ekonomiko nagusien gehien lotzen zaiena delako.

1.2.- Euskal Autonomia Erkidegoko industri inkestan erabilitako zenbatesleak

19 langile baino gehiagoko estatistika-unitate guztiak autoponderatuak dira, eta horrenbestez 1-19 langileko enplegu estratua estimatzea izango da interesgarriena. Ondoren, jardueraren azpimaila jakin bateko (EJSN-93, 5 digiturekin) edozein y aldagairen guztirakoaren eta horren aldakuntza-koefizientearen estimazioa deskribatzen dira.

Egun, y aldagaiaren guztirakoa estimatzeko, arrazoiaren zeharkako estimatzailea edo estimatzaile sintetikoa erabiltzen da; halaber, establezimenduetako langile kopurua informazio osagarri bezala erabiliko da. Azterketa deskribatzaile-exhaustibo bat egin ostean hartu zen erabaki hori, azterketan egiaztatu baitzen establezimenduetako langile kopuruaren eta Industri Inkestako aldagai nagusien magnitudearen artean korrelazio positibo handia zegoela.

Industri Inkestaren kasuan interesa duen edozein y aldagairen arrazoiaren zeharkako estimatzailea, x aldagai lagungarria dagoenean, mota jakin bateko erregresio lineal heterozedastikoaren ereduaz lagunduko da. Hona hemen mota hori:

$$y_{hj} = x_{hj}\beta + \varepsilon_{hj} \quad \text{hurrengoarekin:} \quad \text{var}(\varepsilon_{hj}) = \sigma^2 x_{hj} . \quad (1)$$

Kasu horretan, h estratua da eta j estatistika-unitatea. Estratuak lurralde historikoak dira (Araba, Bizkaia eta Gipuzkoa), hemendik aurrera azpimaila bakar batek bakarrik

izango baitu interesa. Eredu lineal bakunak (1) errore heterozedastikoak ditu, eta bariantza, industri establezimenduetako langile kopuruaren funtzio lineal positiboa da.

h lurralde historikoan eman den azpimaila bateko y aldagaiaren guztirakoaren estimatzailea honelaxe ateratzen da:

$$\hat{t}_{yh.SYN} = X_h \hat{\beta} = X_h \frac{\sum_{h=1}^H \sum_{j=1}^{n_h} w_{hj} y_{hj}}{\sum_{h=1}^H \sum_{j=1}^{n_h} w_{hj} x_{hj}},$$

$X_h = \sum_{j=1}^{n_h} x_{hj}$, w_{hj} delakoa j unitatearen laginketa-pisua da h lurralde historikoan; x_{hj} delakoak h lurralde historikoko j establezimenduko enplegua jasotzen du eta n_h laginaren tamaina da baina h lurralde historikoan.

Arrazoiaren zeharkako estimatzaileak duen bariantzaren estimatzailea honelaxe hurbil daiteke:

$$\hat{\text{var}}(\hat{t}_{yh.SYN}) \approx N^2 \left(1 - \frac{n}{N}\right) \frac{1}{n} \left(\frac{X_h}{\sum_{h=1}^H \sum_{j=1}^{n_h} w_{hj} x_{hj}} \right)^2 \hat{\text{var}}(\varepsilon),$$

$\hat{\text{var}}(\varepsilon)$ datu guztiak dituen (1) eredu heterozedastikoko hondarren lagin-bariantza da (hau da, hondarrak Euskal Autonomia Erkidego osoan kalkulatzen dira, ez bakarrik h lurralde historikoan). Gainerako idazkera ohikoa da.

Särndal and Hidioglou (1989) lanean estimatzaile sintetikoaren alborapenaren hurbilketa bat ematen da; horren arabera $E\left(\hat{t}_{yh.SYN}\right) - t_{yh.SYN} \approx -\sum_{j=1}^N \hat{\varepsilon}_j$ eta

$\hat{\varepsilon}_j = y_j - x_j \hat{\beta}$. Ondorioz, estimatzailea gutxi gorabehera alboragabea izango da $\sum_{j=1}^N \hat{\varepsilon}_k = 0$ dela egiaztatzen bada. Normalean ez da baldintza hori betetzen. Eredua ez bada behar bezala egokitzen intereseko eremuan, hondarren batura zerotik urrun egon daiteke eta horrela alborapen handia erakutsi. Bestela, alborapen mugatua espero dezakegu. Horregatik, batez besteko errore koadratikoa estimatzea komeni da, estimatzailearen doitasunaren adierazle bezala. Honelaxe emango da:

$$MSE(\hat{t}_{yh.SYN}) = \text{var}(\hat{t}_{yh.SYN}) + (\text{alborapena}_{h.SYN})^2,$$

eta hurrengo adierazpenaren bidez estimatzen da:

$$\hat{MSE}(\hat{t}_{yh.SYN}) = \hat{\text{var}}(\hat{t}_{yh.SYN}) + \left(\sum_{j=1}^{n_h} \hat{\varepsilon}_j \right)^2,$$

$\hat{\varepsilon}_j = y_j - x_j^T \hat{\beta}$, $j = 1, \dots, n$ delakorako, laginketako datu guztien bidez estimatu den

(1) ereduari esker lortutako hondarrak dira, baina lurralde historiko bakoitzean soilik batzen dira lurralde horri dagozkionak. Aldakuntza-koefizientea honelaxe definituko da:

$$\hat{cv}(\hat{t}_{yh.SYN}) = \frac{\hat{rmse}(\hat{t}_{yh.SYN})}{\hat{t}_{yh.SYN}},$$

Kasu horretan, $\hat{rmse}(\hat{t}_{yh.SYN}) = \sqrt{\hat{MSE}(\hat{t}_{yh.SYN})}$.

Industri Inkestarako eremu txikietan estimatzeko sistema

2.1.- Sarrera

Euskal Autonomia Erkidegoak 20 administrazio-eskualde ditu, hain zuzen:

- Araba: Arabako Ibarrak/Valles Alaveses, Arabako Lautada/Llanada Alavesa, Arabako Mendialdea/Montaña Alavesa, Arabako Errioxa/Rioja Alavesa, Gorbeia Inguruak/Estribaciones del Gorbea eta Kantauri Arabarra/Cantábrica Alavesa.
- Bizkaia: Arratia-Nerbioi/Arratia-Nervión, Bilbo Handia/Gran Bilbao, Durangaldea/Duranguésado, Enkarterriak/Encartaciones, Gernika-Bermeo, Markina-Ondarroa eta Plentzia-Mungia.
- Gipuzkoa: Bidasoa Beherea/Bajo Bidasoa, Deba Beherea/Bajo Deba, Deba Garaia/Alto Deba, Donostialdea/Donostia-San Sebastián, Goierri, Tolosaldea/Tolosa eta Urola-Kostaldea/Urola Costa.





Euskal Autonomia Erkidegoko industri jarduera ez dago berdin banatuta 20 administrazio-eskualdeetan, hain zuzen, eskualde batetik bestera industri sektorearen garrantzia eta tamaina oso ezberdina da. Izan ere, eskualde batzuetan industri jarduera oso urria da. Horrelakoetan, eskualdeko estimazioa burutzeko eremu txikien estimazio teknikak aplikatu behar dira nahitaez, betiere ereduetan oinarrituta. Kalitatezko eskualde-estimazioak egiteko beharrezko laginen tamaina handitzea oso garestia izango litzateke.

Eremu txikietako ereduaren azpian biztanleei buruzko datu guztiak biltzen dituen beste eredu bat dago, baina hau laginaren datuekin estimatzen da (Rao, 2003). EUSTATEk, Industri Inkestan eskualde-estimazioak lortzeko bi eredu mota erabiltzen ditu: bata eragin finkoak dituen erregresio linealeko eredu, eta bestea, eragin finkoak eta ausazkoak dituen erregresio linealeko eredu edo eredu mistoa.

Eredu mistoan, aurrealeak eragin finkoetarako termino bat du eta beste bat d ($d = 1, \dots, t$) eskualde bakoitzeko elementuetarako. Azken termino hori ausazko eraginek (v_d) osatzen dute, eta eskualde bereko datu guztiek ausazko eragin bera dutelarik. Eragin finkoak dituen ereduaren ez dira ezberdintzen eskualde bakoitzerako terminoak, alderdi sistematikoa bera delako eskualde guztientzat. Alabaina, berezitasuna lortzeko, denentzat estimatu den koefizientea eskualde bakoitzeko informazio osagarri eta zehatzean proiektatu behar da.

2.2.- Eredu lineal mistoa

1. Proiekziozko bertsioa

Jardueraren azpimaila jakin bateko (EJSN-93, 5 digiturekin) N establezimenduek osatzen duten biztanleria hartuko da abiapuntutzat. d ($d = 1, \dots, t$) eskualde bakoitzean biztanlerian N_d establezimendu daude, beraz $N = \sum_d N_d$. Azpimaila horretan n establezimenduk osatzen dute lagina; horietatik n_d d eskualdekoak direlarik. Hurrengo eredu lineal, misto eta heterozedastikoa proposatzen da:

$$y_{dj} = \beta_0 + \beta_1 x_{dj} + v_d + e_{dj}, \quad d = 1, \dots, t \quad j = 1, \dots, n_d, \quad (2)$$

d eskualdeko j establezimendurako, y_{dj} intereseko aldagaiak hartzen duen balioa da eta x_{dj} establezimenduan dagoen langile kopurua. d eskualdean, guztira, n_d establezimendu daude laginean. Ereduaren eragin finkoak β_0 eta β_1 dira. d eskualdeko establezimendu guztientzako ausazko eragina bera da, v_d alegia; eta e_{dj} , establezimendu bakoitzak dituen ausazko erroreak dira. Gainera, $v_d \subset N(0, \sigma_v^2)$ eta $e_{dj} \subset N(0, \sigma_e^2 c_{dj}^{-1})$ independenteak direla uste da. Datuetako heterozedastizitatea zuzentzeko $c_{dj} = 1/x_{dj}$ pisuak erabiltzen dira. $c_{dj} = 1 \quad \forall d, j$ ematen denean, eredu hau eta Battese *et al* (1988) lanean proposatutakoa antzekoak dira.

(2) ereduari dagokion superpopulazio eredua matrize eran dago idatzita eta honelaxe adierazten da:

$$Y = X\beta + Zv + \varepsilon, \quad v \subset N(0, \sigma_v^2 I_t), \quad \varepsilon \subset N(0, \sigma_e^2 C^{-1}), \quad (3)$$

$C = \text{diag}(c_{dj})$ ($d = 1, \dots, t$) ereduaren pisu-matrizea da eta j ($j = 1, \dots, N_d$) establezimendua. $Y = (Y_1^T, \dots, Y_t^T)^T$ bektorea ($N \times 1$) bektorea da. Horren Y_d^T osagaiak eskualde bakoitzerako interesa duen aldagaiaren balioak dira; $\beta = (\beta_0, \beta_1)^T$, ereduaren koefizienteen bektorea da; X , ($N \times 2$) diseinu-matrizea da eta ondoko aldagaiek osatzen dute: aldagai independenteari dagokion bateko zutabe bat eta aldagai lagungarriari dagokion beste bat, hau da, kasu honetan establezimendu bakoitzeko langile kopurua. $Z = \text{diag}(1_{N_d})$, $d = 1, \dots, t$, $N \times t$ diseinu-matrizea da, diagonalak, blokeetan banatuta eta ausazko eraginei lotuta. Hau da, d eskualde bakoitzerako, Z

matrizeak N_d dimentsioa duen $1_{N_d} = (1, \dots, 1)^l$ bektoreak zehazten duen bateko zutabe bat dauka. $v = (v_1, \dots, v_t)^l$ eskualde bereko N_d elementuentzako komunak diren ausazko eraginak dira eta $\varepsilon = (\varepsilon_1^l, \dots, \varepsilon_t^l)^l$ ausazko errorearen bektorea da. Gainera, $\varepsilon_d = (\varepsilon_{d_1}, \dots, \varepsilon_{d_{N_d}})^l$.

Teoria bateratzeko eta (3) ereduaren proiektiozko eta auresateko bertsioak aurkezteko, lagina osatzen duen zatian eta laginetik kanpoko zatian bereiztuko da:

$$\begin{pmatrix} Y_s \\ Y_r \end{pmatrix} = \begin{pmatrix} X_s \\ X_r \end{pmatrix} \beta + \begin{pmatrix} Z_s \\ Z_r \end{pmatrix} v + \begin{pmatrix} \varepsilon_s \\ \varepsilon_r \end{pmatrix},$$

s eta r azpiindizeek laginean parte hartzen duten eta parte hartzen ez duten establezimenduak adierazten dituzte, hurrenez hurren. Orduan, laginketa-eredua honela idatz daiteke:

$$Y_s = X_s \beta + Z_s v + \varepsilon_s, \quad v \subset N(0, \sigma_v^2 I_t), \quad \varepsilon_s \subset N(0_s, \sigma_e^2 C_s^{-1}),$$

$C_s = \text{diag}(c_{dj} = 1/x_{dj})$, $d = 1, \dots, t_s$, $j = 1, \dots, n_d$ eta t_s laginean atera diren eskualde kopuru osoa da. Y_s -ren bariantzen eta kobariantzen matrizea $\text{var}(Y_s) = V_s = Z_s \sigma_v^2 Z_s^l + \sigma_e^2 C_s^{-1} = \text{diag}(V_1, \dots, V_{t_s})$ bezala adieraz daiteke; horrela, $V_d = \sigma_e^2 C_d^{-1} + \sigma_v^2 1_{n_d} 1_{n_d}^l$ eta $C_d = \text{diag}(c_{d_1}, \dots, c_{d_{n_d}})_{n_d \times n_d} = \text{diag}(c_{n_d})$. $\sigma^2 = (\sigma_e^2, \sigma_v^2)$ bariantzaren osagaiak ezagutzen ditugula uste badugu, eragin finkoen estimatzailea eta horren bariantzen eta kobariantzen matrizea karratu txiki orokorren bidez lor daitezke:

$$\tilde{\beta} = (X_s^l V_s^{-1} X_s)^{-1} X_s^l V_s^{-1} Y_s, \quad \text{var}(\tilde{\beta}) = \Phi_s = (X_s^l V_s^{-1} X_s)^{-1},$$

$$V_s^{-1} = \text{diag}(V_1^{-1}, \dots, V_d^{-1}, \dots, V_{t_s}^{-1}), \quad V_d^{-1} = \frac{1}{\sigma_e^2} \left(C_d - \frac{\gamma_{dc}}{c_d} c_{n_d} c_{n_d}^l \right),$$

$$\gamma_{dc} = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2 / c_d} \text{ eta } c_d = \sum_{j=1}^{n_d} c_{dj}.$$

$1_{n_d}^l = (1, \dots, 1)$ delakoak n_d dimentsioa duenean, ausazko eraginaren auresatea $\hat{v}_d = \sigma_v^2 1_{n_d}^l V_d^{-1} (Y_d - X_d^l \hat{\beta}) = \hat{\gamma}_{dc} (\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta})$ -ren bidez lortzen da. Kasu horretan, $\bar{y}_{dc} = \frac{1}{c_d} \sum_{j=1}^{n_d} c_{dj} y_{dj}$, eta $\bar{x}_{dc}^l = \frac{1}{c_d} \sum_{j=1}^{n_d} c_{dj} x_{dj}^l = (1, \bar{x}_{dc})$, $x_{dj}^l = (1, x_{dj})$ -rekin. d

eskualdean interesa duen aldagaiaren batez bestekoaren proiektziozko auresalea honelaxe aterako da:

$$\hat{y}_d^* = \bar{X}_{d(p)}^l \hat{\beta} + \hat{\gamma}_{dc} (\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta}), \quad d = 1, \dots, t, \quad (4)$$

$\hat{\beta} = \tilde{\beta}(\hat{\sigma}_e^2, \hat{\sigma}_v^2)$, $\bar{X}_{d(p)}^l = (1, \bar{x}_{d(p)})$ bariantzako osagaien estimatzaileen bidez ebaluatu da; eta, $\bar{x}_{d(p)} = \frac{\sum_{j \in N_d} x_{dj}}{N_d}$, d eskualdean dagoen langile kopuruaren

populazio-batez bestekoa da populazioaren establezimendu guztietarako: d eskualdeko lagina osatzen dutenak eta osatzen ez dutenak. Guztirakoaren bertsioa honela aterako da:

$$\hat{t}_d^* = X_{d(p)}^l \hat{\beta} + N_d \hat{\gamma}_{dc} (\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta}), \quad d = 1, \dots, t, \quad (5)$$

$X_{d(p)}^l = (N_d, X_{d(p)})$, N_d establezimendu kopuru osoa da eta, $X_{d(p)}$, d eskualdeko langile kopuru osoa (populazioa). (4) auresalea mota jakin bateko erregresio-estimatzaile orokor baten batura ponderatu bezala ere adieraz daiteke. Hona hemen aipatu mota:

$$\bar{y}_{dc} + (\bar{X}_{d(p)}^l - \bar{x}_{dc}^l)^l \hat{\beta},$$

$\bar{X}_{d(p)}^l \hat{\beta}$ erregresioko estimatzaile sintetikoaren bidez ere adieraz daiteke, horrela

$$\hat{y}_d = \gamma_{dc} (\bar{y}_{dc} + (\bar{X}_{d(p)}^l - \bar{x}_{dc}^l)^l \hat{\beta}) + (1 - \gamma_{dc}) \bar{X}_{d(p)}^l \hat{\beta}, \quad d = 1, \dots, t, \quad (6)$$

$0 \leq \gamma_{dc} \leq 1$ pisuak σ_v^2 bariantzaren proportzioa neurtzen du, $\sigma_v^2 + \sigma_e^2 / c_d$ bariantza osoari dagokionez. Ereduaren bariantza txikia denean, γ_{dc} txikiak dira eta osagai sintetikoari pisu gehiago ematen zaio. Era berean, erregresioko zenbatesle orokorrari pisu gehiago ematen zaio c_d handiagoa den heinean. $c_{dj} = 1$ denean, erregresioko estimatzaile orokorra gutxi gorabehera diseinuarekiko alboragabea izango da, baldin eta n_d nahiko handia bada. Oro har, alboragabea da izan diren eraginaren pean dagoen ereduaren azpian v_d , $\tilde{\beta}$ β -rekiko alboragabea izateko baldintzatuta badago. Hau da, BLUP (4) (Best Linear Unbiased Predictor) auresalearen baldintza alboragabea izatea da, $\bar{X}_{d(p)}^l \hat{\beta}$ osagai sintetikoa dagoelako. Laginketa ausazko bakuna denean eta $c_{dj} = 1$, BLUP estimatzailea trinkoa edo konsistentea izango da diseinuarekiko $\bar{Y}_d$ eskualdearen batez bestearantzako n_d hazten denean. Hauxe da horren arrazoia: $\gamma_d \rightarrow 1$.

Eredua homozedastikoa denean eta $c_{dj} = 1$ orduan $c_d = n_d$, batez bestekoaren proiektziozko auresalea honelaxe aterako da:

$$\hat{y}_d^* = \bar{X}_{d(p)}^l \hat{\beta} + \hat{\gamma}_d (\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta}) = \hat{\gamma}_d \bar{y}_{dc} + (\bar{X}_{d(p)}^l - \hat{\gamma}_d \bar{x}_{dc}^l) \hat{\beta},$$

$\hat{\gamma}_d$ -k auresalea modelizatzearekin lortutako ziurgabetasuna neurtzen du eta itxura hau hartzen du:

$$\hat{\gamma}_d = \frac{\text{cov}(v_d, \bar{u}_{d.})}{\text{var}(\bar{u}_{d.})} = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + \hat{\sigma}_e^2 / n_d}.$$

2. Auresateko bertsioa

Eskualdearen araberrako $f_d = n_d / N_d$ laginketa-proporzioa esangarria denean, literaturak gomendatzen du proiektziozko bertsioaren ordez auresateko bertsioa erabiltzea guztirakoaren auresatea edo d eskualdearen batez bestekoa lortzeko. Lagina osatzen duten eta osatzen ez duten establezimenduak bereizten ditu bertsio honek. Horrela, lagina osatzen dutenen auresatea lagina bera da eta lagina osatzen ez dutenen zatia proiektziozko auresalearen bidez auresaten da.

Auresateko bertsioa lortu ahal izateko $\sum_{j \in N_d} y_{dj} = \sum_{j \in d_r} y_{dj} + \sum_{j \in d_s} y_{dj}$ guztirakoa deskonposatu behar da; hain zuzen, d_s -k d eskualdeko lagina adierazten du eta d_r -k d eskualdekoak ez diren gainerako establezimenduak. Jarraian, populazioaren batez bestekoa deskonposatzen da:

$$\bar{y}_d = \frac{\sum_{j \in d_r} y_{dj} + \sum_{j \in d_s} y_{dj}}{N_d} = \frac{(N_d - n_d) \bar{y}_{dr} + n_d \bar{y}_{ds}}{N_d} = (1 - f_d) \bar{y}_{dr} + f_d \bar{y}_{ds}. \quad (7)$$

d eremuko batez bestekoaren auresateko estimatzailea, $d = 1, \dots, t$ guztirako, honelaxe adierazten da:

$$\hat{y}_d = (1 - f_d) \hat{\bar{y}}_{dr} + f_d \bar{y}_{ds} = (1 - f_d) \hat{y}_{dr}^* + f_d \bar{y}_{ds}.$$

$\hat{\bar{y}}_{dr}^*$ -ren ordeaz $\bar{X}_{dr}^l \hat{\beta} + \hat{v}_d$ adierazpena jarriko da, $\bar{X}_{dr}^l = (1, \bar{x}_{dr})$ eta $\bar{x}_{dr} = \frac{\sum_{j \in d_r} x_{dj}}{N_d - n_d}$ izanik. Orduan, hurrengoa dator:

$$\hat{\bar{y}}_d = \hat{\bar{Y}}_d = (1 - f_d) [\bar{X}_{d(p_r)}^l \hat{\beta} + \hat{\gamma}_{dc} (\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta})] + f_d \bar{y}_{ds}, \quad (8)$$

Guztirakoaren auresateko bertsiora garamatza:

$$\hat{t}_d = X_{d(p_r)}^l \hat{\beta} + (N_d - n_d) \hat{\gamma}_{dc} \left(\bar{y}_{dc} - \bar{x}_{dc}^l \hat{\beta} \right) + \sum_{j=1}^{n_d} y_{dj}, \quad d = 1, \dots, t \quad (9)$$

$\hat{\beta} = \hat{\beta}_c(\hat{\sigma}_e^2, \hat{\sigma}_v^2)$, bariantzako osagaien estimazioekin ebaluatu da eta, $X_{d(p_r)}$, d eskualdeko laginetik kanpoko establezimendu guztietako langile-kopurua da.

2.1. Batez bestekoaren eta guztirakoaren estimazioak, lurralde historikoaren arabera eta Euskal Autonomia Erkidegoan

Bai auresateko bertsioan bai proiektiozkoan, lurralde historikoen arabera eta Euskal Autonomia Erkidegoan dauden batez bestekoak eta guztirakoak kalkulatzeko hurrengoak bete behar dira.

Lurralde historikoen arabera batez bestekoaren estimatzailea honelaxe ateratzen da:

$$\hat{\bar{y}}_h = \frac{1}{N_h} \sum_{d \in h} N_d \hat{\bar{y}}_d = \frac{1}{N_h} \sum_{d \in h} \hat{t}_d, \quad (10)$$

$d \in h$ -k adierazten du batura h estratuko eskualde guztietan egiten dela (kasu honetan $h = 1, 2, 3$ lurralde historikoak dira) eta, $N_h = \sum_{d \in h} N_d$, h lurralde historikoko populazioaren guztirakoa da.

Lurralde historikoen arabera guztirakoaren estimatzailea honelaxe ateratzen da:

$$\hat{t}_h = \sum_{d \in h} \hat{t}_d. \quad (11)$$

Euskal Autonomia Erkidegoko batez bestekoaren estimatzailea honelaxe ateratzen da:

$$\hat{\bar{y}} = \frac{1}{N} \sum_{h=1}^3 N_h \hat{\bar{y}}_h, \quad (12)$$

$$N = \sum_{h=1}^3 N_h \text{ delarik.}$$

Euskal Autonomia Erkidegoko guztirakoaren estimatzailea honelaxe ateratzen da:

$$\hat{t} = \sum_{h=1}^3 N_h \hat{\bar{y}}_h = \sum_{h=1}^3 \hat{t}_h, \quad (13)$$

3. Bariantzaren osagaien estimazio ponderatua

Bariantzaren osagaiak estimatzeko metodo bat baino gehiago egon arren agiri honetan momentuen metodoa erabiltzen da estimaziorako, metodo hori ez dagoelako, egiantz handiena (MV) edo mugatutako egiantz handiena (REML) bezala, normaltasun-hipotesien menpean.

3.1. Momentuen metodoa

Momentuen metodoa (Searle et al. 1992) σ_v^2 eta σ_e^2 estimatzeko erabil daiteke, karratuen baturen esperantzak beraien estimatzaileekin berdinduz. Honelaxe adieraziko da beraz:

$$\hat{\sigma}_e^2 = \frac{\mathcal{E}_s' C_e \mathcal{E}_s}{n - \text{rank}(X_s, Z_s)} = \frac{\mathcal{E}_s' C_e \mathcal{E}_s}{n - t_s - 2}, \quad (14)$$

$n = \sum_{d=1}^t n_d$ eta $\text{rank}(X_s, Z_s)$, (X_s, Z_s) matrizearen hein zabaldua da. Horrenbestez, emandako estimatzailea lortzeko beste modu bat lagineko datuak (mendeko aldagaia) datu lagungarriak (enplegua) eta eskualdeko ausazko eraginak auresale gisa (aldagai independenteak) osaturiko erregresio ponderatuaren hondar-bariantza kalkulatzeko litzateke. Ausazko bariantza ondorengo adierazpenaren bidez lor daiteke:

$$\hat{\sigma}_v^2 = \max(\tilde{\sigma}_v^2, 0) = \max\left(\frac{1}{n_{*c}} \left\{ \sum_{d=1}^{t_s} \sum_{j=1}^{n_d} c_{dj} s_{dj}^2 - (n-k-1) \hat{\sigma}_e^2 \right\}, 0\right), \quad (15)$$

Kasu horretan:

$$n_{*c} = \text{tr}(M_c Z_s Z_s^{\top}),$$

$$M_c = C_s - C_s X_s (X_s^{\top} C_s X_s)^{-1} X_s^{\top} C_s,$$

$s_{dj} = y_{dj} - x_{dj} \hat{\beta}_0 = y_{dj} - x_{dj} (X_s^{\top} C_s X_s)^{-1} X_s^{\top} C_s Y_s$, X_s -rekiko Y_s -ren erregresio ponderatuaren hondarrak dira, $C_s = \text{diag}(c_{dj} = 1/x_{dj})$, $d = 1, \dots, t_s$ eta $j = 1, \dots, n_d$ pisuak dituzte. $\hat{\sigma}_v^2 = \max(\tilde{\sigma}_v^2, 0)$ estimatzaile moztua alboratua baina tinkoa da t_s gehitzen denean.

4. Batez besteko errore koadratikoa, eskualdeen arabera

Kackar and Harville (1984) lanean egiaztatu zen normaltasun-hipotesien menpean BLUP $t(\hat{\theta}, Y)$ auresale generikoaren batez besteko errore koadratikoa honela ateratzen dela:

$$MSE[t(\hat{\theta}, Y)] = MSE[t(\theta, Y)] + E[t(\hat{\theta}, Y) - t(\theta, Y)]^2, \quad (16)$$

$\theta = \sigma^2 = (\sigma_v^2, \sigma_e^2)$ da, baldin eta $\hat{\theta} = \hat{\sigma}^2 = (\hat{\sigma}_v^2, \hat{\sigma}_e^2)$ traslazioen aurrean aldatzen ez bada. Batez bestekoaren auresalerako, $t(\hat{\theta}, Y) = \hat{\bar{y}}_d$, $\hat{\bar{y}}_d$ delakoa $\bar{y}$ -ren auresalea da, bariantzaren osagaiak ezagutzen direnean. Horrenbestez, $t(\theta, Y) = \tilde{\bar{y}}_d$. Batez bestekoaren auresalearen MSE honela ateratzen da:

$$MSE[\hat{\bar{y}}_d] = E[\bar{y}_d - \tilde{\bar{y}}_d]^2 + E[\hat{\bar{y}}_d - \tilde{\bar{y}}_d]^2. \quad (17)$$

Henderson-ek (1975) $MSE[\tilde{\bar{y}}_d] = g_{1d}(\sigma^2) + g_{2d}(\sigma^2)$ -ri adierazpen bat eman zion baina (17)-ren bigarren terminoa, $g_{3d}(\sigma^2)$ izenekoa, ez da erraz kalkulatu. Soilik kasu berezi batzuetan kalkula daiteke erraz. Kackar-ek eta Harville-k (1984), Taylor-en seriezko garapenean oinarritutako adierazpen bat lortu zuten:

$$E[\hat{\bar{y}}_d - \tilde{\bar{y}}_d]^2 \approx E[h_d(\theta)(\hat{\theta} - \theta)]^2,$$

hurrengoarekin: $h_d(\theta) = \frac{\partial t_d(\theta)}{\partial \theta}$. Prasad-ek eta Rao-k (1990) geroago hurbilketa bat proposatu zuten:

$$\text{tr}\{A_d(\hat{\theta})E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)']\} \approx \text{tr}\{(\sqrt{b_d})V_d(\sqrt{b_d})'E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)']\},$$

$\sqrt{b_d} = \text{col}_{1 \leq d \leq p} \frac{\partial b_d}{\partial \theta_j}$ eta p bariantzaren osagai kopurua da. $g_{2d}(\sigma^2)$ eta $g_{3d}(\sigma^2)$ -ren estimatzaileak $g_{2d}(\hat{\sigma}^2)$ eta $g_{3d}(\hat{\sigma}^2)$ dira. Estimatzailer hauek zuzenak dira $O_p(t^{-1})$ ordenaraino (hemen t eskualde kopurua da eta ez aurrealea); izan ere, $\hat{\sigma}^2$, σ^2 -ren estimatzailea tinkoa da. Alabaina, $g_{1d}(\hat{\sigma}^2)$ ez da $g_{1d}(\sigma^2)$ -ren estimatzaile zuzena, bere alborapena $O(t^{-1})$ ordenakoa baita eta hori lortzeko $g_{1d}(\sigma^2)$ Taylor-en seriezko garapena erabili behar da σ^2 -ren inguruan eta horren esperantza hartuta. Eragiketa batzuk egin ondoren ondokoa lortzen da:

$$E[g_{1d}(\hat{\sigma}^2)] - g_{1d}(\sigma^2) = -g_{3d}(\hat{\sigma}^2) + O(t^{-1}).$$

4.1. Proiekziozko bertsioa

Prasad-ek eta Rao-k (1990), (17)ren estimatzaile bat ematen dute proiekziozko bertsioan. Bertsio hori erabil daiteke, hain zuzen, bariantzako osagaien estimatzaileak REML bidez edo momentuen metodoaren bidez lortzen direnean. Hona hemen eragiketa hori:

$$M\hat{S}E[\hat{y}_{d(p)}] = g_{1d}(\hat{\sigma}^2) + g_{2d}(\hat{\sigma}^2) + 2g_{3d}(\hat{\sigma}^2),$$

Kasu horretan:

$$g_{1d}(\hat{\sigma}^2) = (1 - \hat{\gamma}_{dc})\hat{\sigma}_v^2,$$

$$g_{2d}(\hat{\sigma}^2) = (\bar{X}_{d(p)} - \hat{\gamma}_{dc}\bar{X}_{dc})' \hat{\Phi}_c (\bar{X}_{d(p)} - \hat{\gamma}_{dc}\bar{X}_{dc}),$$

$$g_{3d}(\hat{\sigma}^2) = \hat{\gamma}_{dc}(1 - \hat{\gamma}_{dc})^2 \hat{\sigma}_e^{-4} \hat{\sigma}_v^{-2} h(\hat{\sigma}^2),$$

$$h(\hat{\sigma}^2) = \hat{\sigma}_e^4 \text{var}(\hat{\sigma}_v^2) + \hat{\sigma}_v^4 \text{var}(\hat{\sigma}_e^2) - 2\hat{\sigma}_v^2 \hat{\sigma}_e^2 \text{cov}(\hat{\sigma}_e^2, \hat{\sigma}_v^2),$$

$$\text{var}(\hat{\sigma}_e^2) = 2(n - t_s - k)^{-1} \hat{\sigma}_e^4 = 2d_e^{-1} \hat{\sigma}_e^4,$$

$$\text{var}(\hat{\sigma}_v^2) = 2n_{*c}^{-2} [(n - t_s - k)^{-1} (t_s - 1)(n - k - 1) \hat{\sigma}_e^4 + 2n_* \hat{\sigma}_e^2 \hat{\sigma}_v^2 + n_{**c} \hat{\sigma}_v^4],$$

$$\text{cov}(\hat{\sigma}_e^2, \hat{\sigma}_v^2) = -(t_s - 1)n_{*c}^{-1} \text{var}(\hat{\sigma}_e^2),$$

$$n_{*c} = \text{tr}(M_c Z_s Z_s^T),$$

$$n_{**c} = \text{tr}(M_c Z_s Z_s^T)^2,$$

$$M_c = C_s \left(I - X_s (X_s^T C_s X_s)^{-1} X_s^T C_s \right),$$

eta k aldagai lagungarrien kopurua da, kasu honetan $k = 1$.

Sarri, $\hat{y}_d$ estimatzailea doitzeko neurri bezala batez besteko errore koadratikoaren (RMSE) erro karratua erabiltzen da. Honela adierazten da:

$$RMSE[\hat{y}_d] = \sqrt{MSE[\hat{y}_d]}.$$

Eskualde bakoitzerako guztirakoaren auresalearen MSE estimatzeko, batez bestekoaren MSEaren estimatzailea eskualdeko populazio-tamainaren karratuaz biderkatu behar da, N_d^2 . Hain zuzen ere,

$$MSE[\hat{t}_d] = N_d^2 [g_{1d}(\hat{\sigma}^2) + g_{2d}(\hat{\sigma}^2) + 2g_{3d}(\hat{\sigma}^2)],$$

eta guztirakoaren batez besteko errore koadratikoaren erro karratua estimatzeko hurrengo adierazpena erabili behar da:

$$RMSE[\hat{t}_d] = \sqrt{MSE[\hat{t}_d]}.$$

Aldakuntza-koefizientea honelaxe definituko da:

$$CV[\hat{t}_d] = \frac{RMSE[\hat{t}_d]}{\hat{t}_d}.$$

4.2. Auresateko bertsioa

Aurresateko bertsioan, (8) aurrealearen batez besteko errore koadratikoaren estimatzailea (baliozkoa dena bariantzako osagaien estimatzaileak REML metodoaren bidez edo momentuen metodoaren bidez lortzen direnean) honelaxe adierazten da:

$$M\hat{S}E[\hat{y}_d] = g_{1d}(\hat{\sigma}^2) + g_{2d}(\hat{\sigma}^2) + 2g_{3d}(\hat{\sigma}^2) + g_{4d}(\hat{\sigma}^2),$$

Kasu horretan:

$$g_{1d}(\hat{\sigma}^2) = (1 - f_d)^2 (1 - \hat{\gamma}_{dc}) \hat{\sigma}_v^2,$$

$$g_{2d}(\hat{\sigma}^2) = (1 - f_d)^2 \left[(\bar{X}_{d(p_r)} - \hat{\gamma}_{dc} \bar{X}_{dc}) \hat{\Phi}_s (\bar{X}_{d(p_r)} - \hat{\gamma}_{dc} \bar{X}_{dc}) \right],$$

$$g_{3d}(\hat{\sigma}^2) = (1 - f_d)^2 c_d^{-1} (\hat{\sigma}_v^2 + \hat{\sigma}_e^2 / c_d)^{-3} [\hat{\sigma}_e^4 \text{var}(\hat{\sigma}_v^2) + \hat{\sigma}_v^4 \text{var}(\hat{\sigma}_e^2) - 2\hat{\sigma}_e^2 \hat{\sigma}_v^2 \text{cov}(\hat{\sigma}_e^2, \hat{\sigma}_v^2)]$$

$$g_{4d}(\hat{\sigma}^2) = \sigma_e^2 N_d^{-2} \sum_{j \in p_r} c_{dj}^{-1},$$

ausazko eraginen, eragin finkoen, bariantzako osagaien eta ereduaren pisuen estimazioak batez besteko errore koadratikoari egindako ekarpenak dira. Gainera, p_r laginekoak ez diren zentsu-establezimenduen d eremua da eta, $c_{dj} = 1/x_{dj}$ delakoan, x_{dj} lagina osatzen duten establezimenduetako zentsu-enplegua da. Estimazioekin hasi baina lehen agregazioak egiten badira gutxieneko tamaina bat lortzeko, lagina osatzen ez duen populazio bezala proiektzioa egin den azpimaila hartuko da, eta ez handiagoa den agregazioarena.

Eskualde bakoitzerako guztirakoaren aurrealearen MSEa estimatzeko, batez bestekoaren MSEaren estimatzailea eskualdeko populazio-tamainaren karratuaz biderkatu behar da, N_d^2 . Hain zuzen ere,

$$M\hat{S}E[\hat{t}_d] = N_d^2 [g_{1d}(\hat{\sigma}^2) + g_{2d}(\hat{\sigma}^2) + 2g_{3d}(\hat{\sigma}^2)], \quad (18)$$

eta guztirakoaren batez besteko errore koadratikoaren erro karratua estimatzeko hurrengo adierazpena erabili behar da:

$$R\hat{M}S\hat{E}[\hat{t}_d] = \sqrt{M\hat{S}E[\hat{t}_d]}. \quad (19)$$

Aldakuntza-koefizientea honelaxe definituko da:

$$CV[\hat{t}_d] = \frac{R\hat{M}S\hat{E}[\hat{t}_d]}{\hat{t}_d}.$$

2.3.- Eragin finkodun eredu lineala

Eragin finkodun superpopulazio eredua ondorengoa da:

$$Y = X\beta + \varepsilon, \quad \varepsilon \subset N(0, \sigma_e^2 C^{-1}), \quad (20)$$

$C = \text{diag}(c_{dj})$ ereduaren pisu-matrizea da; d , ($d = 1, \dots, t$) eskualdea da eta, j ($j = 1, \dots, N_d$) establezimendua da. $Y = (Y_1^T, \dots, Y_t^T)$ bektorea ($N \times 1$) bektorea da, eta horren Y_d^T osagaiak d eskualde bakoitzerako intereseko aldagaietan ikusten diren balioak dira; β ereduko koefiziente finko bakarra da; X , aldagai lagungarriaren ($N \times 1$) zutabe bektorea da, hau da, enpleguarena; eta $\varepsilon_d^T = (\varepsilon_{d_1}, \dots, \varepsilon_{d_{N_d}})$ -ean $\varepsilon_d^T = (\varepsilon_{d_1}, \dots, \varepsilon_{d_{N_d}})$ dagoela kontuan hartuta, ausazko errorearen bektorea da.

Eredu mistoan egindako deskonposizioaren antzera, lagina osatzen duen zatia eta osatzen ez duen zatia jarraian adierazi bezala bereiztu daitezke:

$$\begin{pmatrix} Y_s \\ Y_r \end{pmatrix} = \begin{pmatrix} X_s \\ X_r \end{pmatrix} \beta + \begin{pmatrix} \varepsilon_s \\ \varepsilon_r \end{pmatrix}$$

s eta r azpiindizeek lagineko eta laginetik kanpoko establezimenduak adierazten dituzte, hurrenez hurren. Orduan, eragin finkodun langiketa-eredua honela idatz daiteke:

$$Y_s = X_s \beta + \varepsilon_s \quad \varepsilon_s \subset N(0_s, \sigma_e^2 C_s^{-1}), \quad (21)$$

$C_s = \text{diag}(c_{dj})$, $d = 1, \dots, t_s$, $j = 1, \dots, n_d$ eta t_s laginean parte hartzen duten eskualdeen kopurua da. Hedatuz gero, (21) eredu honela adieraziko da:

$$y_{dj} = \beta x_{dj} + e_{dj} \quad d = 1, \dots, t, \quad j = 1, \dots, n_d \quad (22)$$

j eskualdeko d establezimendurako, y_{dj} intereseko aldagaiak hartzen duen balioa da eta x_{dj} establezimenduan dagoen langile kopurua. d eskualdean laginean dauden establezimendu kopuru osoa n_d da, β ereduaren eragin finko bakarra da eta $e_{dj} \subset N(0, \sigma_e^2 c_{dj}^{-1})$ ausazko erroreak dira. Gainera, $e_{dj} \subset N(0, \sigma_e^2 c_{dj}^{-1})$. Datuetako heterozedastizitatea zuzentzeko $c_{dj} = 1/x_{dj}$ pisuak erabiltzen dira.

N_d , d eremuko unitateen kopuru osoa bada, d eremuko populazioaren batez bestekoa ondoren adierazi bezala aterako da:

$$\bar{Y}_d = \frac{1}{N_d} \sum_{d=1}^{N_d} y_{dj} = f_d \bar{y}_{ds} + (1 - f_d) \bar{y}_{dr}, \quad (23)$$

$f_d = n_d / N_d$, $\bar{y}_{ds}$ lagineko unitateen batez bestekoa da eta $\bar{y}_{dr}$ laginekoak ez diren unitateena. (23)aren bigarren terminoa ezezaguna denez, estimaturiko bere balioarekin ordezkatu da. (7)aren antzeko moduan atera den (23)aren estimatzailea honelaxe adieraziko da:

$$\hat{\bar{y}}_d^F = f_d \bar{y}_{ds} + (1 - f_d) \bar{X}_{d(p_r)} \hat{\beta}, \quad (24)$$

$\bar{X}_{d(p_r)} = \sum_{j \in d_r} x_{dj} / (N_d - n_d)$, d eremuan laginekoak ez diren unitateen langile kopuruaren populazio batez bestekoa da.

β -ren estimatzailea $\hat{\beta} = (X_s^T C_s X_s)^{-1} (X_s^T C_s Y_s) = \sum_{d=1}^{t_s} \sum_{j=1}^{n_d} y_{dj} / \sum_{d=1}^{t_s} \sum_{j=1}^{n_d} x_{dj}$ bezala adierazten da; azken hau, β -aren gutxieneko karratu orokorren arabera estimatzailea da. $Var(\hat{\beta}) = \sigma^2 (X_s^T C_s X_s)^{-1} = \sigma^2 / \sum_{d=1}^{t_s} \sum_{j=1}^{n_d} x_{dj}$ bariantzen eta kobariantzen matrizea da.

d eskualderako guztirakoaren estimatzailea honela aterako da:

$$\hat{t}_d^F = \sum_{j=1}^{n_d} y_{dj} + X_{d(p_r)} \hat{\beta}, \quad (25)$$

eta lurralde historikoen arabera batez bestekoaren estimatzailea honelaxe ateratzen da:

$$\hat{\bar{y}}_h^F = \frac{1}{N_h} \sum_{d \in h} N_d \hat{\bar{y}}_d^F, \quad (26)$$

$d \in h$ -k adierazten du batura h estratuko eremu guztietan egiten dela (kasu honetan $h = 1, 2, 3$ lurralde historikoak dira) eta, $N_h = \sum_{d \in h} N_d$, h lurralde historikoko populazioaren guztirakoa da.

Lurralde historikoen arabera guztirakoaren estimatzailea honelaxe ateratzen da:

$$\hat{t}_h^F = \sum_{d \in h} N_d \hat{y}_d^F = \sum_{d \in h} \hat{t}_d^F, \quad (27)$$

$d \in h$ -k adierazten du batura h estratuko eremu guztietan egiten dela (kasu honetan $h=1,2,3$ lurralde historikoak dira) eta, $N_h = \sum_{d \in h} N_d$, h lurralde historikoko populazioaren guztirakoa da.

Euskal Autonomia Erkidegoko batez bestekoaren eta guztirakoaren estimatzaileak, hurrenez hurren, ondokoak dira:

$$\hat{y}^F = \frac{1}{N} \sum_{h=1}^3 N_h \hat{y}_h^F, \quad \hat{t}^F = \sum_{h=1}^3 \hat{t}_h^F. \quad (28) (29)$$

(24) eskualdeen arabera batez bestekoaren estimatzaileen batez besteko errore koadratikoak, aurrerako bertsioan, honela ematen dira:

$$M\hat{S}E[\hat{y}_d^F] = E\left[\left(\hat{y}_d^F - \bar{Y}_d\right)^2\right] = (1 - f_d)^2 \left[\bar{X}_{d(p_r)}^2 \text{vâr}(\hat{\beta}) \bar{X}_{d(p_r)}\right] + \frac{\sigma^2}{N_d^2} \sum_{j \in p_r} x_{dj} \quad (30)$$

p_r , laginetik kanpoko zentsu-enplegua da. Estimazioekin hasi baino lehen agregazioak egiten badira gutxieneko tamaina bat lortzeko, laginetik kanpoko populazio bezala hartuko da proiektzioa egin den azpimaila, eta ez handiagoa den agregazioarena.

Eskualde osorako MSEa honen bidez estimatzen da:

$$M\hat{S}E[\hat{t}^F] = N_d^2 (1 - f_d)^2 \left[\bar{X}_{d(p_r)}^2 \text{vâr}(\hat{\beta}) \bar{X}_{d(p_r)}\right] + \sigma^2 \sum_{j \in p_r} x_{dj} \quad (31)$$

2.4.- Sektorearen arabera guztirakoak

Orain arte aurkeztu diren eredu bidez guztirakoaren aurrerak lortu daitezke, eskualdeen, lurralde historikoen eta Euskal Autonomia Erkidegoaren arabera, azpimaila bakoitzerako; baita errore estandarrek (RMSE) eta aldakuntza-koefizienteak ere. Sektore bakoitzeko guztirakoak kalkulatzeko, A84 sailkapena erabiliz adibidez, sektore

bera osatzen duten azpimailetak bakoitzeko eskualdeen arabera lortu diren guztirakoak agregatu baino ez da egin behar; agregazio hori eskualdeen, lurralde historikoen eta Euskal Autonomia Erkidegoaren arabera egin daiteke. Sektoreko errore estandarrak, azpimaila bakoitzeko MSEen baturaren erro karratuaren bidez lortzen dira. Sektore berean gerta daiteke azpimaila batzuk eredu mistoen arabera estimatu izana eta beste batzuk eredu finkoen arabera. Bigarren agregazio mailarako emaitzak lortzeko, A31 sailkapenerako adibidez, A84 sektoreen guztirakoak agregatu behar dira. Errore estandarrak kalkulatzeko ere sektoreak independenteak direla suposatu behar da eta, beraz, RMSE kalkulatzeko sektoreen MSEen baturaren erro karratua kalkulatu behar da.

2.5.- Kalibratze-prozesua

Kalibratzearen bidez, Industri Inkestak A84 edo beste agregazio maila batean lurralde historikoen eta Euskal Autonomia Erkidegoaren mailan ematen dituen guztirako berberak lortzen dira. t_d baderitzogu sektore jakin batek d eskualdean eredueta oinarritutako estimazioari, t_h lurralde historiko bakoitzeko eredueta oinarrituz lortutako guztirakoaren estimazioari eta C_h Industri Inkestak lurralde historiko bakoitzeko lortutako guztirakoaren estimazioari, orduan, sektore bakoitzeko kalibratu den estimazio berria honela aterako da:

$$\tilde{t}_d = t_d \frac{C_h}{t_h}.$$

Horrela, eredueta oinarrituriko lurralde historiko bakoitzeko kalibratutako $\tilde{t}_h$ guztirakoa bat dator sektore berbererako Industri Inkestaren estimatzailearen bidez lurralde historiko bakoitzeko estimatu den C_h guztirakoarekin. Izan ere:

$$\tilde{t}_h = \sum_d \tilde{t}_d = \sum_d t_d \frac{C_h}{t_h} = \frac{C_h}{t_h} \sum_d t_d = C_h.$$

2.6.- Industri Inkestako estimazio plana

EUSTATEk aplikazio informatiko bat programatu du *ad hoc* SASen, eremu txikiak estimazioen kalkulua estatistika-ekoizpenean sartzeko asmoz. Industri Inkestarako programa berezi bat da hori, baina erraz molda daitekeena bestelako inkesta ekonomikoetara. Orain arte hainbat erabaki hartu eta programatu dira:

- Bai eredu lineal mistoa bai eragin finkodun eredu lineala aurreratzeko bertsioan kalkulatu dira. Azpimaila batzuetako populazioak oso txikiak dira eta, ondorioz, laginketa-proportzioa ezin da gutxietsi.
- Aurreratzeko bertsioak bi zatitan banatzen du aurrerana: lagina osatzen duten establezimenduetarako ereduak lortutakoa eta lagina osatzen ez duten establezimenduetarako lortutakoa. Hori oso erabilgarria da Industri Inkestan, azpimaila batzuetan ohiz kanpoko establezimenduak daudelako, hots, besteekin alderatuta portaera oso ezberdina duten establezimenduak daudelako eta horiek asko alda ditzaketelako estimazioak. Establezimendu horiek, estimaziorako onargarriak ez diren establezimendu bezala sailkatzen dira eta, beraz, aurreratzeko bertsioa erabiltzean ez dira estimazioak egiteko prozeduretan kontuan hartzen. Azken aurreranean, ordea, kontuan hartzen dira. Onargarriak ez diren establezimendu gisa sailkatutakoak, gainerako lagina bezala, lagina osatzen ez dutenen aurreran batzen dira gero guztirakoaren aurrerana lortzeko.
- Eredu mistoak edo finakoak kalkulatzeko gutxieneko establezimendu kopuru bat egon behar dela proposatu da. Gutxieneko establezimendu kopuru horretara heldu ezean EJSNen agregazioak egingo dira baina digitu bat gutxiagorekin. Lehenengo eredu mistoa estimatuko da agregazio maila horretan, eta $\sigma_v^2 = 0$ edo $\sigma_e^2 = 0$ denean eragin finkodun eredu estimatuko da. Egun, gutxieneko kopuru hori 5 establezimendukoa izatea ezarri da (aldatu daitekeen arren).
- Eredu mistoa onargarria ez denean eragin finkodun eredu erabiltzea erabaki da, azpimailak ez taldekatzeko erabakiak lehentasuna du eta. Horrenbestez, nahiago da adibidez 5 digitutan eragin finkodun eredu erabiltzea eta ez 3 digitutan eredu mistoa.
- Agregazio bat egiten denean, honen bidez ereduaren koefizienteak estimatu ahal izango dira; baina aurreran aintzat hartu den azpimaila bakoitzean bereizita egingo dira.
- “Langile kopurua” izeneko aldagai lagungarria erabiltzean ereduaren heterozedastizitatea sartzen da, normalean y aldagai-erantzunak aldakortasun handiagoa izango duelako langile kopurua gehitzen doan heinean. Orduan, eragin finkodun eredu eta eredu misto guztietan akatsaren bariantza langile kopuruarekiko proportzionala izango da.
- Azpimaila bakoitzean, lurralde historikoen eta Euskal Autonomia Erkidegoaren araberrako guztirakoak modu agregatuan lortzen dira, eskualdeen araberrak egindako estimazioak oinarritzat hartuta. A84 sektorearen araberrako guztirakoak lortzeko, azpimailan lortutako aurreran agregatu behar dira. Gauza bera egin behar da lurralde historikoen eta EAEaren araberrako guztirakoekin. Baita beste agregazio mota batzuekin ere.
- Azpimaila bakoitzean, lurralde historikoen mailako aurreran batez besteko errore koadratikoen baturaren erro karratuak kalkulatzeko formula bereziak

aplikatzen dira; izan ere, ezin da hori lortu eskualdeen arabera aurrean batez besteko akats koadratikoen baturaren erro karratuarekin. Hori horrela da azpimaila bakoitzean eskualdeen arabera estimazioak ez direlako eredu batean ere independenteak.

- Alabaina, lurralde historiko bakoitzeko RMSEen estimazioak azpimailen arabera egin ostean, EAEaren arabera estimazioak zuzenean kalkulatu ahal izango dira orain independentzia-hipotesia betetzen delako. Azpimailen arabera kalkuluak aintzat hartuta, A84 aldagaiaren edo beste edozein taldekatzeren RMSEak zuzenean ateratzen dira; hots, sektore bakoitza osatzen duten azpimailen MSEen baturaren erro karratua kalkulatu besterik ez da egin behar.

Eremu txikien estimazio tekniken aplikazioa Euskal AEko Industri Inkestan. 2002 eta 2003.

3.1.-Eskualdekako estimazioak Euskal Autonomi Erkidegoko Industri Inkestan. 2002 eta 2003.

Ondoren, arestian Euskal Autonomia Erkidegoko Industri Inkestan aipatu den estimazio sistema erabiliz lortutako 2002 eta 2003ko estimazioak aurkezten dira. Honako makromagnitudeak argitaratzea erabaki da

- Balio erantsia faktoreen kostuan
- Salmenta garbiak
- Ustiapeneko soberakin gordina
- Langileriaren kostuak
- Inbertsioa
- Zerga aurreko emaitza

Halaber, eskualdeetako enpleguaren gaineko estimazioak ere ematen dira, Euskal Autonomia Erkidegoko Industri Inkestan erabilitako eremu txikietako eredueta kanpoko aldagai bezala erabili baita.

Industriako langile okupatuak, lurralde historikoaren eta eskualdearen arabera. 2002-2003

	2002	2003	Δ 03/02
Euskal AE	246.063	248.922	1,2
Araba	47.804	49.019	2,5
Arabako Ibarrek	1.889	1.953	3,4
Arabako Lautada	30.978	31.912	3,0
Arabako Mendialdea	354	294	-16,9
Errioxa Arabarra	3.778	4.042	7,0
Gorbeia Inguruak	3.942	3.949	0,2
Kantauri Arabarra	6.863	6.869	0,1
Bizkaia	103.749	104.619	0,8
Arratia-Nerbioi	5.173	5.021	-2,9
Bilbo Handia	60.622	61.312	1,1
Donostialdea	21.614	21.736	0,6
Enkartzioak	2.128	2.139	0,5
Gernika-Bermeo	4.418	4.612	4,4
Markina-Ondarroa	3.892	4.056	4,2
Plentzia-Mungia	5.902	5.743	-2,7
Gipuzkoa	94.510	95.284	0,8
Bidasoa Beherea	5.490	5.602	2,0
Deba Beherea	10.452	10.643	1,8
Deba Garaia	20.275	20.091	-0,9
Donostia-San Sebastián	25.464	25.508	0,2
Goierrri	12.831	13.233	3,1
Tolosa	8.001	8.185	2,3
Urola-Kostaldea	11.997	12.022	0,2

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

**Industriako balio erantsia faktore-kostutan eta aldakuntza-koefizientea (cv)
lurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003**

	2002	cv	2003	cv	Δ 03/02
Euskal AE	13.008.214	0,01	13.371.649	0,01	2,8
Araba	2.676.614	0,00	2.747.477	0,00	2,6
Arabako Ibarrek	89.743	0,03	106.061	0,02	18,2
Arabako Lautada	1.708.733	0,01	1.763.392	0,01	3,2
Arabako Mendialdea	17.455	0,19	14.211	0,16	-18,6
Errioxa Arabarra	243.155	0,02	270.594	0,04	11,3
Gorbeia Inguruak	216.196	0,02	210.776	0,01	-2,5
Kantauri Arabarra	401.332	0,01	382.443	0,01	-4,7
Bizkaia	5.371.448	0,01	5.561.854	0,01	3,5
Arratia-Nerbioi	252.658	0,02	259.622	0,01	2,8
Bilbo Handia	3.259.555	0,01	3.395.279	0,01	4,2
Donostialdea	1.109.460	0,01	1.113.048	0,01	0,3
Enkartzioak	154.460	0,05	155.544	0,04	0,7
Gernika-Bermeo	180.951	0,03	199.627	0,02	10,3
Markina-Ondarroa	149.595	0,03	167.664	0,02	12,1
Plentzia-Mungia	264.769	0,01	271.070	0,01	2,4
Gipuzkoa	4.960.152	0,01	5.062.318	0,01	2,1
Bidasoa Beherea	218.659	0,02	231.525	0,02	5,9
Deba Beherea	482.592	0,02	482.471	0,02	0,0
Deba Garaia	996.600	0,01	1.013.516	0,00	1,7
Donostia-San Sebastián	1.585.031	0,01	1.591.077	0,01	0,4
Goierrri	655.356	0,01	682.139	0,01	4,1
Tolosa	422.004	0,02	442.768	0,01	4,9
Urola-Kostaldea	599.910	0,01	618.822	0,01	3,2

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

Industriako salmenta garbiak eta aldakuntza-koefizientea (cv)

Iurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003

	2002	cv	2003	cv	Δ 03/02
Euskal AE	42.393.031	0,01	43.768.410	0,01	3,2
Araba	8.336.340	0,01	8.614.962	0,01	3,3
Arabako Ibarrek	321.016	0,02	390.287	0,02	21,6
Arabako Lautada	5.521.407	0,01	5.564.479	0,01	0,8
Arabako Mendialdea	46.122	0,17	40.494	0,17	-12,2
Errioxa Arabarra	603.353	0,02	702.555	0,03	16,4
Gorbeia Inguruak	705.855	0,02	763.485	0,01	8,2
Kantauri Arabarra	1.138.586	0,01	1.153.661	0,01	1,3
Bizkaia	19.341.307	0,01	20.078.450	0,01	3,8
Arratia-Nerbioi	743.317	0,03	822.265	0,02	10,6
Bilbo Handia	12.733.700	0,01	13.268.830	0,01	4,2
Donostialdea	3.518.902	0,02	3.562.602	0,01	1,2
Enkartzioak	422.018	0,05	417.302	0,03	-1,1
Gernika-Bermeo	673.681	0,03	727.858	0,02	8,0
Markina-Ondarroa	495.976	0,03	524.603	0,03	5,8
Plentzia-Mungia	753.713	0,03	754.989	0,02	0,2
Gipuzkoa	14.715.384	0,01	15.074.998	0,01	2,4
Bidasoa Beherea	581.630	0,03	630.369	0,02	8,4
Deba Beherea	1.273.758	0,03	1.238.176	0,04	-2,8
Deba Garaia	3.134.142	0,01	3.202.536	0,01	2,2
Donostia-San Sebastián	4.600.028	0,01	4.557.629	0,01	-0,9
Goierrri	2.039.115	0,01	2.224.916	0,01	9,1
Tolosa	1.239.249	0,04	1.315.686	0,02	6,2
Urola-Kostaldea	1.847.462	0,02	1.905.687	0,03	3,2

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

Industriako langileria-kostuak eta aldakuntza-koefizientea (cv)

Iurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003

	2002	cv	2003	cv	Δ 03/02
Euskal AE	7.398.012	0,00	7.808.500	0,00	5,5
Araba	1.477.956	0,00	1.576.926	0,00	6,7
Arabako Ibarrek	57.176	0,01	61.286	0,01	7,2
Arabako Lautada	969.544	0,00	1.024.924	0,00	5,7
Arabako Mendialdea	7.968	0,06	7.258	0,04	-8,9
Errioxa Arabarra	88.179	0,03	106.669	0,03	21,0
Gorbeia Inguruak	119.351	0,01	126.500	0,01	6,0
Kantauri Arabarra	235.738	0,00	250.291	0,00	6,2
Bizkaia	3.108.386	0,00	3.273.633	0,00	5,3
Arratia-Nerbioi	154.341	0,01	155.158	0,01	0,5
Bilbo Handia	1.879.256	0,01	1.976.019	0,01	5,1
Donostialdea	638.435	0,01	675.569	0,01	5,8
Enkartzioak	57.201	0,02	62.725	0,02	9,7
Gernika-Bermeo	114.381	0,01	125.955	0,01	10,1
Markina-Ondarroa	93.588	0,01	102.151	0,01	9,1
Plentzia-Mungia	171.183	0,02	176.056	0,01	2,8
Gipuzkoa	2.811.670	0,00	2.957.941	0,00	5,2
Bidasoa Beherea	138.692	0,01	150.936	0,01	8,8
Deba Beherea	303.162	0,01	319.017	0,01	5,2
Deba Garaia	625.672	0,00	652.605	0,00	4,3
Donostia-San Sebastián	749.887	0,01	775.827	0,01	3,5
Goierrri	403.843	0,00	436.067	0,01	8,0
Tolosa	240.291	0,01	259.569	0,01	8,0
Urola-Kostaldea	350.123	0,01	363.920	0,01	3,9

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

Industriako ustiapeneko soberakin gordina eta aldakuntza-koefizientea (cv)
Iurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003

	2002	cv	2003	cv	Δ 03/02
Euskal AE	5.610.202	0,01	5.563.149	0,01	-0,8
Araba	1.198.658	0,01	1.170.551	0,01	-2,3
Arabako Ibarrek	32.931	0,07	44.762	0,05	35,9
Arabako Lautada	739.335	0,01	737.566	0,02	-0,2
Arabako Mendialdea	9.530	0,34	7.186	0,29	-24,6
Errioxa Arabarra	156.177	0,03	164.114	0,05	5,1
Gorbeia Inguruak	95.502	0,04	83.599	0,03	-12,5
Kantauri Arabarra	165.183	0,01	133.325	0,03	-19,3
Bizkaia	2.263.062	0,01	2.288.221	0,01	1,1
Arratia-Nerbioi	98.254	0,04	104.809	0,03	6,7
Bilbo Handia	1.376.066	0,02	1.418.524	0,01	3,1
Donostialdea	474.025	0,02	438.507	0,02	-7,5
Enkartzioak	97.238	0,07	92.094	0,05	-5,3
Gernika-Bermeo	66.908	0,07	73.657	0,05	10,1
Markina-Ondarroa	56.424	0,06	64.772	0,05	14,8
Plentzia-Mungia	94.146	0,03	95.858	0,04	1,8
Gipuzkoa	2.148.482	0,01	2.104.377	0,01	-2,1
Bidasoa Beherea	79.729	0,05	79.886	0,04	0,2
Deba Beherea	178.555	0,04	164.393	0,05	-7,9
Deba Garaia	373.600	0,02	358.072	0,02	-4,2
Donostia-San Sebastián	831.687	0,02	814.317	0,02	-2,1
Goierri	254.308	0,02	247.468	0,02	-2,7
Tolosa	181.731	0,04	185.976	0,04	2,3
Urola-Kostaldea	248.871	0,03	254.264	0,03	2,2

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

Industriako inbertsio egina eta aldakuntza-koefizientea (cv)
Iurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003

	2002	cv	2003	cv	Δ 03/02
Euskal AE	3.211.667	0,02	2.717.879	0,01	-15,4
Araba	660.740	0,02	510.428	0,01	-22,7
Arabako Ibarrek	36.549	0,55	26.231	0,03	-28,2
Arabako Lautada	457.031	0,02	332.071	0,02	-27,3
Arabako Mendialdea	2.396	0,17	2.198	0,20	-8,3
Errioxa Arabarra	58.364	0,15	60.380	0,05	3,5
Gorbeia Inguruak	55.831	0,45	34.254	0,04	-38,6
Kantauri Arabarra	50.569	0,04	55.295	0,02	9,3
Bizkaia	1.459.249	0,02	1.364.705	0,01	-6,5
Arratia-Nerbioi	41.686	0,35	41.403	0,05	-0,7
Bilbo Handia	1.030.310	0,02	1.079.300	0,01	4,8
Donostialdea	248.724	0,27	165.607	0,03	-33,4
Enkartzioak	42.907	1,80	18.728	0,08	-56,4
Gernika-Bermeo	27.135	0,54	8.636	0,18	-68,2
Markina-Ondarroa	25.626	0,07	24.139	0,06	-5,8
Plentzia-Mungia	42.860	0,35	26.891	0,06	-37,3
Gipuzkoa	1.091.678	0,02	842.746	0,02	-22,8
Bidasoa Beherea	53.852	0,54	37.669	0,04	-30,1
Deba Beherea	93.377	0,23	79.199	0,05	-15,2
Deba Garaia	193.433	0,02	213.272	0,01	10,3
Donostia-San Sebastián	370.970	0,04	216.867	0,03	-41,5
Goierri	142.977	0,30	115.465	0,03	-19,2
Tolosa	137.092	1,17	79.849	0,15	-41,8
Urola-Kostaldea	99.976	0,05	100.425	0,04	0,4

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

Industriako zerga aurreko emaitza eta aldakuntza koefizientea (cv)
Iurralde historikoaren eta eskualdearen arabera (€ milakotan). 2002-2003

	2002	cv	2003	cv	Δ 03/02
Euskal AE	3.213.405	0,02	2.904.928	0,02	-9,6
			0		
Araba	772.874	0,02	579.541	0,02	-25,0
Arabako Ibarrek	13.060	0,15	18.734	0,10	43,5
Arabako Lautada	468.551	0,02	316.075	0,04	-32,5
Arabako Mendialdea	6.934	0,46	6.568	0,31	-5,3
Errioxa Arabarra	145.513	0,02	103.511	0,07	-28,9
Gorbeia Inguruak	40.156	0,07	61.471	0,03	53,1
Kantauri Arabarra	98.660	0,02	73.182	0,05	-25,8
Bizkaia	1.344.669	0,02	1.274.144	0,02	-5,2
Arratia-Nerbioi	64.725	0,06	67.487	0,04	4,3
Bilbo Handia	838.007	0,02	774.138	0,02	-7,6
Donostialdea	263.632	0,03	243.464	0,03	-7,7
Enkartzioak	63.818	0,10	62.697	0,08	-1,8
Gernika-Bermeo	37.564	0,11	31.403	0,11	-16,4
Markina-Ondarroa	23.769	0,13	34.147	0,08	43,7
Plentzia-Mungia	53.155	0,04	60.808	0,06	14,4
Gipuzkoa	1.095.862	0,03	1.051.243	0,03	-4,1
Bidasoa Beherea	37.231	0,10	35.167	0,09	-5,5
Deba Beherea	101.389	0,07	91.380	0,09	-9,9
Deba Garaia	139.336	0,04	127.648	0,04	-8,4
Donostia-San Sebastián	429.737	0,03	418.766	0,03	-2,6
Goierri	127.438	0,04	108.258	0,05	-15,1
Tolosa	108.981	0,06	123.785	0,05	13,6
Urola-Kostaldea	151.750	0,05	146.238	0,05	-3,6

Iturria: EUSTAT. Industriaren eta Eraikuntzaren Kontuak

3.2.-Balio erantsia faktoreen kostuetan eta Langile okupatuak eskualdeka Industri Inkestan. 2002 eta 2003. Ondorioak.

Industri Inkestako bi aldagai garrantzizkoenen (Balio erantsia faktoreen kostuetan eta Langile okupatuak) emaitzen analisitik ateratako ondorio nagusienak azaltzen dira ondoren.

Lurralde historikoa baino txikiagoak diren agregazio-mailentzako estatistika-datuak lortzeko aukera ematen du metodologia berri eta konplexu horrek. Izan ere, eskuarki maila horietarako definitzen dira estatistika-eragiketen laginak. Horrela, azken urteotan hainbat gizarte-eragilek planteatutako datu desagregatuen estatistika-eskari gero eta handiagoari erantzun nahi zaio.

Euskadiko industriaren eskualdekako datuen 2003rako estimazioan azpimarratzekoa da Bilbo Handiak (%25,4) hartu zuela industriako balio erantsiaren laurden bat. Arabako Lautada (%13,2) eta Donostia-San Sebastián (%11,9) eskualdeak kokatu ziren jarraian. Hiru eskualde horiek Euskal AEko industriako balio erantsiaren erdia baino gehixeago (%50,5) ekarri zuten 2003an, eta %50,4 2002an. Industriako enplegatuei dagokienez, hiru eskualde horiek Euskal AEko industriako langileen %47,7 enplegatu zuten.

Durungaldeak, Deba Garaiak eta Kantauri Arabarrak Euskal AEko industriako balio erantsiaren %18,8 ekarri zuten eta sektore horretako langileen %19,6 okupatu zituzten. Beraz, industriaren %69,3 Euskadiko hoge eskualdeetatik seitan kontzentratu zen. Hala ere, kontzentrazio-dispersio fenomeno horrek alde handiak ageri ditu lurralde batetik bestera.

Esaterako Arabako lurralde historikoan, industriako balio erantsiaren ia bi herenak (%64,2) Arabako Lautadan kontzentratu ziren, baina Arabako Mendialdean %0,5 besterik ez. Lurralde horretako gainerako lau eskualdeek Kantauri Arabarraren %13,9 eta Arabako Ibarren %3,9 arteko industriako balio erantsia izan zuten.

Industriaren eskualdekako kontzentrazioaren eragina handia da Bizkaian ere. Lurraldeko industria-jardueraren %81,0 bi eskualdetan metatu zen: Bilbo Handian (%61,0) eta Durungaldean (%20,0). Bestalde, Enkartzioak apenas heldu ziren %2,8ra. Industriako pisu handieneko hirugarren eskualdea Arratia-Nerbioi izan zen, lurralde horretako industriako balio erantsiaren %4,7rekin.

Gipuzkoan, oster, kontzentrazioa txikiagoa da eta eskualdeen arteko oreka neurritsuagoa, gainerako lurraldeetan baino. Industriaren erdia bi eskualdetan kontzentratu zen: Donostia-Sansebastianen (%31,4) eta Deba Beherean (%20,0), baina era berean badira esaterako Goerri eta Urola Kosta eskualdeak, industriako lurralde-pisu erlatiboaren %13,5ekin eta %12,2kin, hurrenez hurren. Bestalde, Gipuzkoako industrian pisu txikieneko eskualdea Bidasoa Behera dugu, baina hala ere lurraldeko industriaren %4,6 hartu zuen, Arabako Mendialdearen %0,5 eta Enkartzioetako %2,8ren aldean.

Euskal AEko industriak %2,8ko hazkunde nominala izan zuen 2003an eta batez besteko hazkunde hori gaingitu egin zuten hainbat eskualdek: Araban, Arabako Ibarrek (%18,2), Errioxa Arabarrak (%11,3) eta Arabako Lautadak (%3,2); Bizkaian, Markina-Ondarroak

(%12,1), Gernika-Ondarroak (%10,3) eta Bilbo Handiak (%4,2); Gipuzkoan, Bidasoa Behereak (%5,9), Tolosak (%4,9), Goierrik (%4,1) eta Urola Kostak (%3,2).

2003an enplegatutako langileen bilakaerari dagokionez, Arabako eskualde guztiek profil positiboa izan zuten, Arabako Mendialdeak izan ezik. Bizkaian bi eskualdek bakarrik (Arratia-Nerbioi eta Plentzia-Mungia) egin zuten atzera industriako enpleguan, eta beste guztiek handitu egin zuten industria-okupazioa. Gipuzkoan, berriz, Deba Goiena izan zen industriako langileak gutxitu zituen eskualde bakarra, gainerako seiak gehitu egin baitzituzten.

Ondorioak

Agiri honetan deskribatzen diren eremu txikietako ereduak, Industri Inkestako makromagnitude ezberdinen eskualde mailan guztirakoen estimazioak lortzeko aukera ematen dute. Oro har, lortutako emaitzak oso onak dira, estimatu diren aldakuntza-koefizienteak benetan neurrizkoak direlako laginaren irudikapena eta populazioaren tamaina altu samarra diren eskualdeetan. Horrelakorik gertatzen ez den eskualdeetan, estimatutako aldakuntza-koefizienteak altuagoak dira.

Urte honetatik aurrera Eustatek Industri Inkestan eskualde mailako estimazioak egingo ditu, urtero agiri honetan aurkeztu den desagregazio mailarekin edo hori baino handiagoarekin. Horrela, gerora, eskualdeko estimazioen denborazko serieak izango ditugu eskura eta horien bidez egungoak baino konplexuagoak diren analisi koiunturalak egin ahal izango dira, Euskal Autonomia Erkidegoko makromagnitude ekonomiko nagusien bilakaerari buruzko informazio gehiago lortuz. Informazio hori guztia oso erabilgarria izango da lurralde mailako politika ezberdinak eratu ahal izateko.

Ezinbestean aipatu behar da gainera, eremu txikiak estimatzeko proiektu hau egitean zehatz-mehatz berrikusi behar izan direla Industri Inkestako jasokari-prozesuak; ondorioz, estimatzaileak kalkulatzeko eta laginketa-akatsak kalkulatzeko metodologiak hobetu egin dira.

Eustat oraindik aztertzen dabil eremu txikietan estimatzeko ereduaren metodologia, laster aurrerapauso bat eman eta desagregazio maila are handiagoarekin kalitatezko estimazioak eskaini ahal izateko. Horretaz gain, metodologia hori Eustaten beste inkesta batzuetan erabiltzen ari da; ez bakarrik industriaren esparruari lotutako inkestetan, estimatu beharreko aldagai guztiak edo horietako batzuk diskretuak diren inkestetan ere erabiltzen ari delako.

Bibliografia

[1] Battese, G.E., Harter, R.M and Fuller, W.A. (1988). An Error-Components Model for Prediction of Country Crop Areas Using Survey and Satellite Data. *Journal of the American Statistical Association*, **83**, 28-36.

[1] Henderson, C.R. (1975). Best Linear Unbiased Estimation and Prediction Under a Selection Model. *Biometrics*, **31**, 423-447.

[2] Kackar, R.N. and Harville, D.A.. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association*.

[2] Pfefferman, D. (2002). Small Area Estimation – New Developments and Directions. *International Statistical Review*, **70**, 125-143.

[3] Prasad, N.G.N and Rao, J.N.K (1990). The Estimation of Mean Squared Error of Small Area Estimators. *Journal of the American Statistical Association*, **85**, 163-171.

[4] Rao, J.N.K. (2003). Small Area Estimation. Wiley Series in Survey Methodology.

[5] Särndal, C.E., Hidirogloy, M.A. (1989). Small Domain Estimation: A Conditional Analysis. *Journal of the American Statistical Association*, **84**, 266-275.

[6] Searle, S.R., Casella, G. and McCulloch, C.E. (1992). Variance Components. Wiley Series in Probability and Statistics.