

ESTATISTIKA ETA DATUEN ANALISIA

VIII: Bi aldagai kualitatiboen baterako azterketa: **asoziazioa**

Egilea: Josemari Sarasola



Gizapedia

gizapedia.hirusta.io

- 8.1 Aldagai bakunaren azterketatik binakako aldagaien azterketara
- 8.2 Datuetatik kontingentzia-taulara
- 8.3 Asoziazioaren azterketa esploratzailea: portzentajeak
- 8.4 Asoziazioa eta independentzia
 - 8.4.1 Maiztasun teorikoak
- 8.5 Asoziazio-neurriak
 - 8.5.1 Khi-karratu
 - 8.5.2 Khi-karratuan oinarritutako neurriak
 - 8.5.2.1 Phi
 - 8.5.2.2 Kontingentzia-koefizientea
 - 8.5.2.3 Cramér-en V
 - 8.5.3 PRE neurriak: lambda
- 8.6 Asoziazio ordinala: Goodman eta Kruskal-en gamma
- 8.7 Interakzioa
- 8.8 Ariketak

8. gaia: Bi aldagai kualitatiboen baterako azterketa: asoziazioa

8.1 Aldagai bakunaren azterketatik binakako aldagaien azterketara

Aurreko ikasgaietan aldagai bakunak aztertu dira, haien zentroa, sakabanatzea eta beste ezaugarriak ikertuz. Ikasgai honetan aldagaiak binaka aztertzen hasi behar dugu ikasten, eta horretarako beraien arteko erlazio estatistikoa ikertuz (nola lotzen den hezkuntza-maila sexuarekin, adibidez). Ikasgai honetan **atributu edo aldagai kualitatiboen arteko erlazio estatistikoa, asoziazioa** alegia, ikasi behar dugu, zehatzago **asoziazioaren sendotasuna** (zenbaterainokoa den erlazioa sendoa) eta **asoziazioaren norabidea** (erlazioa nolakoa den alegia; adibidez, emakumeak edo gizonak diren hezkuntza-maila altuagoa dutenak).

8.2 Datuetatik kontingentzia-taulara

Bi aldagai kualitatiboei buruzko datuak [kontingentzia-taula](#) (egin klik) direlakoetan biltzen dira, azterketa abiarazi aurretik. Kontingentzia-taulak [sarrera biko taulak](#) (egin klik) besterik ez dira, non bi atributuetako kategoriak gurutatu eta kategoria-konbinazio bakoitzeko maiztasunak zehazten diren.

Adibidea: Ikasgela birtual batean ikastaroa hasi zutenen sexua eta ikastaroa bukatu zuten jaso da (g: gizon, e: emakume / b: bai, e: ez):

Sexua: g-g-e-e-g-g-e-e-g-g-e-e-g-g-e
Bukatu al zuen?: b-e-e-e-b-e-b-e-b-e-b-e-b-e-b

Datuei dagokien kontingentzia-taula osatu behar da.

Kategoria konbinazio bakoitzean datuak zenbatuta, taula hau suertatzen da:

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	5	4	9
Emakumea	2	6	8
Guztira	7	10	17

Aldagai bakoitzeko maiztasun totalak, beste aldagaia kontuan hartu gabe, [baztermaiztasunak](#) dira (egin klik).

8.3 Asoziazioaren azterketa esploratzailea

8.3.1 Portzentajeak

Kontingentzia-taula batean portzentajeak edo maiztasun erlatiboak aldagai bati zein besteari buruz kalkula daitezke, maiztasun gurutzatu bakoitza baztermaiztasunekin, alde batekoekin zein bestekoekin, zatituz:

- sexuari buruz,

Sexua	Bukatu al zuen?	
	Bai	Ez
Gizona	%71.4	%40
Emakumea	%28.6	%60
Guztira	%100	%100

Honela interpretatzen da:

- Ikastaroa bukatu zutenen artean, gizonak %71.4 dira, eta emakumeak %28.6.
- Ikastaroa bukatu ez zutenen artean, gizonak %40 dira, eta emakumeak %60.

- ikastaroa bukatzeari buruz,

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	%55.5	%44.4	%100
Emakumea	%25	%75	%100

Honela interpretatzen da:

- Gizonen artean, ikastaroa bukatu zutenak %55.5 dira, eta bukatu ez zutenak %44.4.
- Emakumeen artean, ikastaroa bukatu zutenak %25 dira, eta bukatu ez zutenak %75.

Portzentaje horiek *asoziazioaren norabidea esploratzeko* erabiltzen dira bereziki, asoziazioaren nondik norako nagusiak ezagutzeko. Adibidean, asoziazioaren norabideari buruz, emakumeek ikastaroa ez bukatzeko joera dutela ikus daiteke.

Asoziazioaren sendotasunari buruz jakiteko ere erabil daitezke, horri buruz zerbait zehaztea zaila den arren: gizon eta emakumeen artean bukatze-portzentajeak oso desberdinak dira eta beraz aurreratu daiteke asoziazioa sendoa edo argia dela, lagin errorearen erreserbapean betiere.

8.3.2 Odds ratio edo probabilitate-zatidura

2x2 tauletan, aldagaiak kausa-efektu erlazio batez loturik daudenean, aldagai independentea eta aldagai dependentea bereizi daitezkeenean, ohikoa da asoziazioa esploratzeko **odds ratio edo probabilitate-zatidura** kalkulatzeko.

Probabilitate-zatidura kalkulatzeko taula aldagai independente edo kausaren modalitateen arabera banatu behar da, eta ondoren modalitate bakoitzeko aldagai dependentean interesekoa den modalitatearen **probabilitate empirikoa** kalkulatu da.

Arestiko adibidea harturik, aldagai independentea sexua. Beraz, horren arabera banatuko dugu taula: gizonen artean, ikastaroa bukatzeko probabilitatea $5/9=0.55$ da; emakumeen artean, ikastaroa bukatzeko probabilitatea $2/8=0.25$. Probabilitate-zatidura $0.55/0.25=2.2$ da. Honela interpretatzen da: gizonek 2.2 aldiz probabilitate handiagoa dute emakumeek baino ikastaroa gainditzeko.

8.4 Asoziazioa eta independentzia

Bi atributuen artean **independentzia** dagoela esaten da, batere asoziaziorik ez dagoenean, hau da, atributu bateko zein besteko portzentajeak berdinak direnean beste atributuaren arabera.

Adibidez, har dezagun kontingentzia-taula hau:

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	30	15	45
Emakumea	60	30	90
Guztira	90	45	135

Taula 8.1: Portzentajeak: sexuari buruz

Taula 8.2: Portzentajeak: ikastaroari buruz

Sexua	Bukatu al zuen?		Sexua	Bukatu al zuen?		Guztira
	Bai	Ez		Bai	Ez	
Gizona	%33.3	%33.3	Gizona	%66.6	%33.3	%100
Emakumea	%66.6	%66.6	Emakumea	%66.6	%33.3	%100
Guztira	%100	%100				

Hurrengo orrialdeko portzentaje-tauletan ikus daitekeenez, ez dago inongo ezberdintasunik gizon eta emakumeen artean, ezta ikastaroa bukatu zutenen eta ez zutenen artean. Beraz, bi atributuen artean erabateko independentzia dago.

8.4.1 Maiztasun teorikoak

Kontingentzia-taula bat edukita, bi atributuen arteko independentziaren egoeran izango liratekeen maiztasunak jakitea bilatzen da batzuetan. Independentzia-egoerako maiztasun balizko horiei **maiztasun teoriko (ingelesez, *expected*)** deritze.

Maiztasun teorikoak kalkulatzeko kontingentzia-taula bateko bazter-maiztasunak soilik behar dira. Gelaska jakin baterako maiztasun teorikoa zutabeko eta errenkadako bazter-maiztasunen biderketa lagin-tamaina osoarekin zatituz kalkulatzeko da.

Adibidea: Aurreko adibideetan landutako kontingentzia-aula honetako maiztasun teorikoak kalkulatu behar dira:

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	5	4	9
Emakumea	2	6	8
Guztira	7	10	17

Arestian, azaldu bezala, gelaska bakoitzean zutabeko eta errenkadako baztermaiztasunak bidertuz eta emaitza lagin-tamainarekin zatituz:

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	$\frac{7 \times 9}{17} = 3.7$	$\frac{10 \times 9}{17} = 5.3$	9
Emakumea	$\frac{7 \times 8}{17} = 3.3$	$\frac{10 \times 8}{17} = 4.7$	8
Guztira	7	10	17

Taula teoriko honetan, portzentaje-aula biak kalkulaturik, independentzia dagoela ikusiko genuke.

8.5 Asoziazio-neurriak

Ikasiko ditugun asoziazio-neurriek asoziazioaren sendotasuna neurtzen dute, baina orokorrean ez dute zuzenean informaziorik ematen asoziazioaren norabideari buruz. Horretarako, portzentajeak azter daitezke.

8.5.1 Khi-karratu

Maiztasun enpirikoak edo behatuak (ingelesez *observed frequencies* edo **O**) datuetatik eratortzen diren diren maiztasun errealak dira.

Maiztasun enpirikoak eta maiztasun teorikoak (independentzia-egoeran itxaroten direnak, ingelesez *expected frequencies* edo **E) zenbat eta desberdinagoak diren, hainbat eta sendoagoa da atributuen arteko asoziazioa. Khi-karratu** (X^2 , grezierazko alfabetoko khi letratik), ingelesez *chi-square*, O maiztasun enpirikoen eta E teorikoen arteko diskrepantzia aldea neurtzen duen estatistiko bat da:

$$X^2 = \sum_{\text{gelaska}} \frac{(O - E)^2}{E}$$

Adibidea: Aurreko adibideko kontingentzia-taulatik, kalkulatu khi-karratu estatistikoa.

Aurretik maiztasun enpirikoak eskura eta maiztasun teorikoak kalkulatu jada, honela kalkulatzen da khi-karratu estatistikoa:

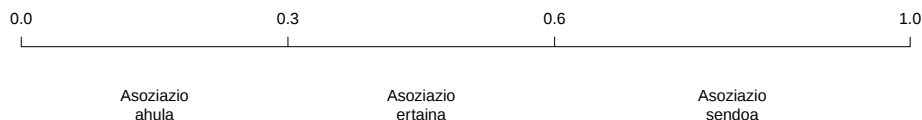
$$X^2 = \sum \frac{(O - E)^2}{E} = \frac{(5 - 3.7)^2}{3.7} + \frac{(4 - 5.3)^2}{5.3} + \frac{(2 - 3.3)^2}{3.3} + \frac{(6 - 4.7)^2}{4.7} = 1.64$$

Khi-karratu estatistikoak balio positiboak hartzen ditu beti (0 independentzia-egoeran). Zenbat eta handiagoa izan, orduan eta asoziazio sendoagoa izango da, baina ez du goi-mugarik (eta beraz ezin eman asoziazio sendoa adierazten duen baliorik).

Asoziazioaren norabideari buruz, $\frac{(O-E)}{\sqrt{E}}$ balioei erreparatuko diegu: taulako balio handienek gelaska horietako bi kategoriak loturik daudela adierazten dute, eta txikieneek, zeinu negatiboarekin, kategoria horiek ez daudela loturik.

8.5.2 Khi-karratuan oinarritutako neurriak

Goi-mugarik ez duenez, ez dago asoziazioa handitzat jo daitekeen khi-karratu estatistikoaren baliorik. Hori dela eta, asoziazioa [0-1] tartera (0: independentzia, 1: erabateko asoziazioa) mugatzen duten asoziazio-neurri batzuk asmatu dira, jarraian ikasiko ditugunak, asoziazioaren sendotasuna neurtzeko. Orokorrean, ikasiko ditugun neurrietan asoziazio-maila 0.3 baino txikiagoa denean, asoziazio ahultzat joko du, 0.3-0.6 bitartean ertaintzat eta 0.6tik gora sendotzat.



Irudia 8.1: Asoziazioaren sendotasunerako neurriak interpretatzeko erregela orokorra.

Edonola ere, neurri horietako balioak interpretatzean antzeko ikerketak ere hartu behar dira kontuan. Adibidez, 0.8ko asoziazio-maila txikia izan daiteke, antzeko ikerketetan 0.9tik gorako balioak eskuratu badira. Horrela, asoziazio-maila bati buruzko interpretazioa *antzeko ikerketen erreserbapean* izango da bati, *lagin-erorrea* ere ahaztu gabe.

8.5.2.1 Phi

2×2 tauletarako bakarrik erabiltzen da. Honela kalkulatzen da, n lagin-tamaina izanik:

$$\phi = \sqrt{\frac{X^2}{n}}$$

8.5.2.2 Kontingentzia-koefizientea

Edozein tamainako tauletarako erabiltzen da. Honela kalkulatzen da:

$$C = \sqrt{\frac{X^2}{X^2 + n}}$$

Erabateko asoziazioaren kasuan ere, koefizientea ez da 1 baliora heltzen, $\sqrt{\frac{m-1}{m}}$ baliora taula karratuen kasuan (m izanik errenkada edo zutabe kopurua) eta $\left(\frac{r}{r+1} \times \frac{c}{c+1}\right)^{1/4}$ baliora taula ez-karratuen kasuan (r eta c errenkada- eta zutabe-kopurua izanik, hurrenez hurren). Koefizientea $[0-1]$ tartera mugatzeko, koefizientea *normalizatze*ko alegia, koefizientearen balioa koefizientearen balio maximoarekin zatitu behar da, besterik gabe.

8.5.2.3 Cramér-en V

Edozein tamainako tauletarako erabiltzen da. Honela kalkulatzen da:

$$V = \sqrt{\frac{X^2}{n(m-1)}}$$

m izanik, errenkada- eta zutabe-kopuruetan txikiena.

8.5.3 Asoziazioaren norabidea khi-karratuaren emaitzetatik

Khi-karratu eta horrekin loturiko neurriak asoziazio estatistikoaren intentsitatea neurtzeko erabiltzen direla ikusi dugu aurreko atalean. Baina asoziazioaren beste ikuspegia, *nolako erlazioa* dagoen bi aldagai kualitatiboen artean, asoziazioaren norabidea alegia, aztertzeko ere balia dezakegu.

Horreta khi-karratu kontribuzioen karratuak kalkulatzeko ditugu gelaska bakoitzeko, hau da, balio hauek:

$$\frac{O - E}{\sqrt{E}}$$

Eta balio horiei zeinu hau emanez erregela honen arabera:

$$O > E \longrightarrow +\frac{O - E}{\sqrt{E}} ; O < E \longrightarrow -\frac{O - E}{\sqrt{E}}$$

Balio haueetatik guztietatik hartuz balio handiena (gelaska hori urdinez inguratuz) eta txikiena (gelaska hori gorritz inguratuz), lotura estueneko kategoriak, gehien erakartzen direnak alegia, eta elkarrengandik gehien aldentzen diren kategoriak antzemango ditugu horrela, hurrenez hurren.

8.5.4 PRE neurriak: lambda

Lambda PRE (*Proportional Reduction in Error*) motako neurri bat da, aldagai kualitatibo baten balioa ezagutzeak bestea aurrasatean dakarren errore-murrizketa kalkulatzen duena. Horretarako, bi aldagai kualitatiboetan *aldagai independentea edo kausa* eta *aldagai dependentea edo ondorioa* ezarri behar dira aldez aurretik.

Lehenbizi, aldagai dependenteak zein balio hartu behar duen aurrasan behar da, aldagai independenteak batere kontuan hartu gabe, hau da, bazter-maiztasunetan maiztasun handieneko kategoria hartuz. Beste bazter-maiztasunen batura aurrasan horretan guztira sortutako errore-kopurua da.

Bigarrenik, aldagai dependentearen balioa aurrasan behar da, baina aldagai independenteak hartzen duen balio bakoitzeko, hau da, aldagai independentearen kategoria ezberdinak ezagunak direla kontuan harturik, hurrenez hurren. Horretarako maiztasun handiena hartuko da kategoria bakoitzeko errenkada edo zutabea, eta erroreak beste maiztasunen batura izango da.

Aldagai independentea kontuan hartu gabe ken aldagai independentea kontuan hartuta sortutako erroreak kalkulatzen dira, aldagai independentearen balioa ezagutzeak dakarren errore-murrizketa alegia. Errore-murrizketa hori erlatiboki kalkulatuko da, aldagai independenteak kontuan hartu gabe sortzen den errore-kopuruarekiko alegia. Errore-murrizketa erlatibo hori izango da lambda.

Lambdak $[0-1]$ tarteko balioak hartzen ditu eta horren balioa beste asoziazio-neurrietan bezala interpretatzen da, 0-0.3-0.6-1 erregelaren arabera.

Adibidea: Kontingentzia-taula honetako lambda kalkulatu behar da:

Sexua	Bukatu al zuen?		Guztira
	Bai	Ez	
Gizona	5	4	9
Emakumea	2	6	8
Guztira	7	10	17

Aldagai independentea sexua da, eta ikastaroa bukatzea edo ez dependentea. Izan ere, sexua izan daiteke ikastaroa bukatzearen eragilea edo kausa eta ez alderantziz.

Aldagai independentea, sexua alegia, kontuan hartu gabe, 7 lagunek ikastaroa bukatu zutela dakigu, eta 10ek ez. Egin beharreko aurrasana sexua ez dakigun pertsona bati buruz ikastaroa ez duela bukatuko izan behar da, baina horretan 7 errore egingo ditugu 17 lagunen artean, 7 lagunek ikastaroa bukatu egin zutelako. Beraz, it sexua kontuan hartu gabe 7 errore egiten dira.

Pertsonaren sexua jakinik, berriz,

- pertsona gizona dela badakigu, ikastaroa bukatuko duela aurrean behar da, 5 direlako bukatu zutenak, eta 4 ez, baina horrela 4 pertsona horien kasuan kale egiten dugu, 4 errore alegia;
- pertsona emakumea dela badakigu, ikastaroa bukatuko ez duela aurrean behar da, 6 direlako bukatu ez zutenak, eta 2 bai, baina horrela 2 pertsona horien kasuan kale egiten dugu, 2 errore alegia.

Beraz, pertsonaren sexua jakinik, guztira $4+2=6$ errore egiten dira. Sexua jakin gabe 7 errore egiten genituenek, $7-6=1$ errore gutxiago egiten dugu sexua jakinda, hau da $1/7=0.14$. Beraz,

$$\lambda = 0.14$$

Hala, bi aldagaien arteko asoziazioa ahula dela ondorioztatzen dugu (0.3 baino txikiagoa denez), antzeko ikerketen eta lagin-errorearen erreserbapean.

Taula 8.3: **Lambdarako kalkulu-taula azkarra.** Bete beharrekoak azpimarratuta dauden hutsuneak dira, hemen letra etzanarekin bete direnak.

<u>sexua</u> (ald. indep.) kontuan hartu gabe	<u>sexua</u> kontuan hartuta	errore-aldea (amaiera)
→ auresana: <u>Ez</u> → erroreak: <u>7</u>	• gizona dela jakinda: → auresana: <u>Bai</u> → erroreak: <u>4</u> • emakumea dela jakinda: → auresana: <u>Ez</u> → erroreak: <u>2</u> • erroreak guztira: <u>4+2=6</u>	Zenbat errore gutxiago? <u>7-6=1</u> Batekotan? <u>$\lambda=1/7=0.14$</u>

Aldagai independentea zein den garbi ez dagoenean, lambda bi norabideetan kalkulatzeko da egokiena.

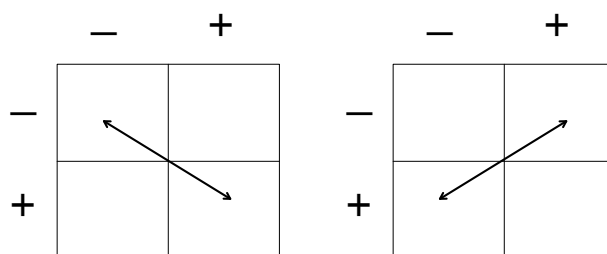
8.6 Asoziazio ordinala: Goodman eta Kruskal-en gamma

Atributuen kategoriak mailakatu daitezkeenean, aldagaia ordinala dela esaten da (adibidez, hezkuntza-maila, kalitate-maila, ...). Kontingentzia-taula batean bi aldagaiak ordinalak direnean, ohiko asoziazio-neurriez gainera **Goodman eta Kruskal-en gamma** ere kalkula daiteke:

$$\gamma = \frac{k - d}{k + d}$$

non k konkordantzien kopurua den, eta d diskordantzien kopurua. $[-1,1]$ tarteko balioak hartzen ditu eta negatiboa denean, asoziazioa negatiboa dago (aldagai bat igotzean, bestea jaitsi egiten da) eta positiboa denean, asoziazio positiboa (aldagai bat igotzean, bestea igo egiten da). Asoziazioaren sendotasuna koefizientearen balio absolutuari erreparatuz kalkulatzen da, eta beste asoziazio-neurrietan bezala interpretatzen da.

Konkordantzia bat taulako bi elementuen artean dagoen asoziazio positiboa dela esan daiteke: elementu batetik bestera aldagai bat *igotzean*, beste aldagaia ere *igo* egiten denean. Konkordantzien kopurua kalkulatzeko aski da, taulako elementuen artean osa daitezkeen *igo-igo* bikoteak zenbatzea. *Diskordantzia* suertatzen da, berriz, aldagai bat igotzean elementu batetik bestera, beste aldagaia *jaitsi* egiten denean. Diskordantzien kopurua kalkulatzeko aski da *igo-jaitsi* bikoteak zenbatzea.



Irudia 8.2: Konkordantzia duten elementuak ezkerrean, eta diskordantzia dutenak eskuinean.

Adibidea: Kontingentzia-aula honetan ikasle batzuen ingelesa-maila eta ingelesez emandako ikastaro bat gaititu duten jaso dira. Dagokion Goodman eta Kruskalen gamma kalkulatu eta interpretatu.

Ingeles-maila	Ikastaroa gaititu al du?	
	Ez	Bai
Baxua	8	4
Altua	5	12

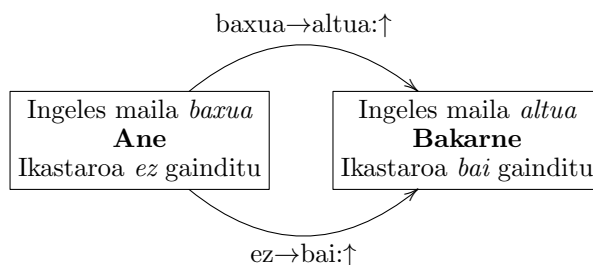
Konkordantziak ingeles-maila baxua duten eta ikastaroa gaititu ez duten ikasleen eta ingeles-maila altua eta ikastaroa gaititu duten ikasleen artean gertatzen dira. Ikasle horien guztien artean bikoteak eginez, $8 \times 12 = 96$ konkordantzia daude.

Diskordantziak ingeles-maila altua duten eta ikastaroa gaititu ez duten ikasleen eta ingeles-maila baxua eta ikastaroa gaititu duten ikasleen artean gertatzen dira. Ikasle horien guztien artean bikoteak eginez, $4 \times 5 = 20$ diskordantzia daude.

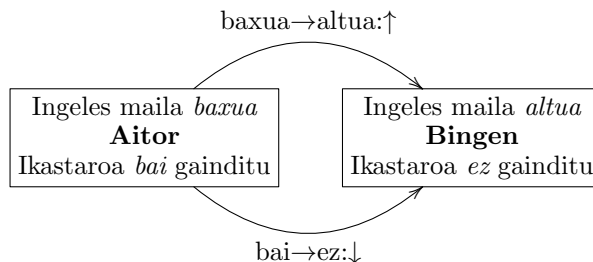
Goodman eta Kruskal-en gamma honela kalkulatu da:

$$\gamma = \frac{96 - 20}{96 + 20} = 0.65$$

Positiboa denez, ingeles-maila handia izateak ikastaroa gaititzea ekartzen du orokorrean, eta lotura hori sendoa da gainera, antzeko ikerketen eta lagin-erreorearen erreserbapean.



Irudia 8.3: Ane eta Bakarne alderaturik, **konkordantzia** dago batengandik bestearengana igarotzerakoan, aldagai batek *gora* egitean, besteak ere *gora* egiten duelako.



Irudia 8.4: Aitor eta Bingen alderaturik, **diskordantzia** dago batengandik bestearengana igarotzerakoan, aldagai batek *gora* egitean, besteak *behera* egiten duelako.

8.7 Interakzioa

Interakzioa hirugarren aldagai kualitatibo bat kontuan ez hartzeagatik bi aldagai kualitatiboaren artean agertzen den asoziazio ahula da. Hain zuzen, datuak hirugarren aldagai horren arabera banaturik, asoziazio sendoa da, baina efektu banatu horiek konpentsatzen dira eta hala itxurazko asoziazio ahula sorrarazten da. Adibidez, taula honetan agerikoa da sexuaren eta tabakismoaren arteko asoziazio ahula:

Taula 8.4: Lurralde osoan jasotako datuak

Sexua	Erretzen al du?	
	Bai	Ez
Gizona	33	31
Emakumea	32	35

Baina pertsonak bizitokiaren arabera bereizten badira:

Taula 8.5: Hiri handietan jasotako datuak

Sexua	Erretzen al du?	
	Bai	Ez
Gizona	28	4
Emakumea	8	26

Taula 8.6: Herri txikietan jasotako datuak

Sexua	Erretzen al du?	
	Bai	Ez
Gizona	5	27
Emakumea	24	9

Nabari da orain sexu-tabakismo erlazioa, aurkako norabidea duena hirietan eta herri txikietan: gizonak dira hirietan erretzen dutenak, eta emakumeak herri txikietan. Bizitokia interakzioa eragiten duen aldagaia da beraz, eta horren arabera datuak bereizten ez badira, konklusio okerrera goaz, sexuak ez duela tabakismoarekin loturarik alegia, konpentsazio efektu batengatik.

8.8 Ariketak

1. Kafetegi batera joaten diren bezeroen sexua eta tea edo kafea eskatu duten jaso dira. Emaitzak kontingentzia taula honetan jaso dira:

Sexua	Zer eskatu du?	
	Tea	Kafea
Gizona	186	332
Emakumea	566	387

- (a) Taulako portzentajeak kalkulatu behar dira norabide bietan. Horiek kalkulatzeko zeintzuk balio falta zaizkio taulari?
 - (b) Portzentajeen tauletatik portzentaje bana interpretatu.
 - (c) Asoziazioaren norabidea azaldu.
 - (d) Kafea eskatzen duten bezeroei fideltasuna saritzeko oparia eskaini behar bazaie, zein sexutarako egin beharko litzateke bereziki?
 - (e) Hiriko emakumeen artean sustapen-kanpaina bat zein produkturekin egin beharko litzateke?
 - (f) Tea eskatzen dutenen artean *odds ratio* delakoa kalkulatu eta interpretatu.
2. Sexuari eta aukeratutako batxilergo-adarrari buruzko datuak jaso eta kontingentzia-
taula batean bildu dira:

Sexua	Zer ikasi du?			
	Zientziak	Gizartea	Teknologia	Guztira
Gizona	274	86	225	585
Emakumea	312	165	186	663
Guztira	586	251	411	1248

- (a) Khi-karratuan oinarritutako asoziazio neurriak kalkulatu eta interpretatu.
- (b) Asoziazioaren norabidea ere aztertu behar da, gelaskak behar den moduan koloreztatuz.

3. 25 urteko gazteen artean sexua eta ikasketa-maila galdetu zaizkie:

Sexua	Zer ikasi du?			
	DBH	Lanb. Hez.	Unibertsitatea	Guztira
Gizona	265	238	386	889
Emakumea	186	436	198	820
Guztira	451	674	584	1709

- (a) Lambda eta gamma (ahal bada) kalkulatu eta interpretatu.
- (b) Lambda eta gamma (ahal bada) kalkulatu eta interpretatu, taulan akatsa dagoela jakinik: sexuaren ordezkari lan-egoera azaldu behar da eta "gizona"ordezkari "lanik gabe" eta "emakume"ordezkari "lanean".
4. Herrialde batean, telefoniako operadoreek mugikorraren eskaintzak egiten dizkiete bezeroei. 18-30 urte bitarteko gazteei darabilten mugikorraren marka eta kontratatuta duten operadorea galdetu zaizkie:

Marka	Operadorea			Guztira
	A	B	C	
Alfa	86	202	145	433
Beta	121	111	136	368
Gamma	44	108	168	320
Guztira	251	421	449	1121

Lambda eta gamma (ahal bada) kalkulatu eta interpretatu.

5. Lurrin-denda batean lurrin bat erosi duten 30-50 urte bitarteko emakume guztiei aurreko aldia zein lurrin marka erosi duten galdetu zaie, une horretan erosi duten lurrinaren marka jasotzearekin batera. Emaitzak hauek dira (zutabeetan oraingo marka, errenkadetan aurreko marka):

Aurreko marka	Oraingo marka				Guztira
	Dior	Chanel	Shiseido	Loewe	
Dior	46	58	21	16	141
Chanel	67	87	36	25	215
Shiseido	54	42	67	8	171
Loewe	13	25	45	98	181
Guztira	180	212	169	147	708

- (a) Zein da bezeroen leialtasun handiena duen marka?
- (b) Khi-karratuan oinarritutako asoziazio-neurriak kalkulatu eta interpretatu, asoziazioa dakarten gelaskak eta euren asoziazio norabidea nabarmenduz.
- (c) Lambda kalkulatzeko, zein aldagai hartuko genuke independentetzat?
6. Enpresa txiki, ertain eta handiei azken urtean euren mozkinak igo ala jaitsi diren galdetu zaie, ISO 9000 kalitate agiria dutenak eta ez dutenak bereiziz taula bietan:

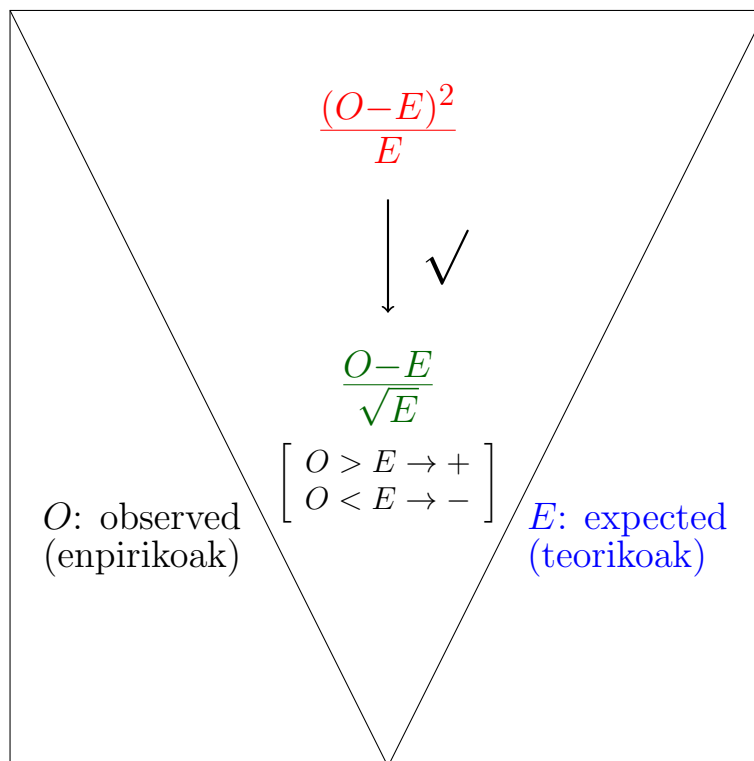
• ISO EZ

Tamaina	Mozkinak		Guztira
	Igo	Jaitsi	
Txikia	126	15	141
Ertaina	43	68	111
Handia	3	28	31
Guztira	172	111	283

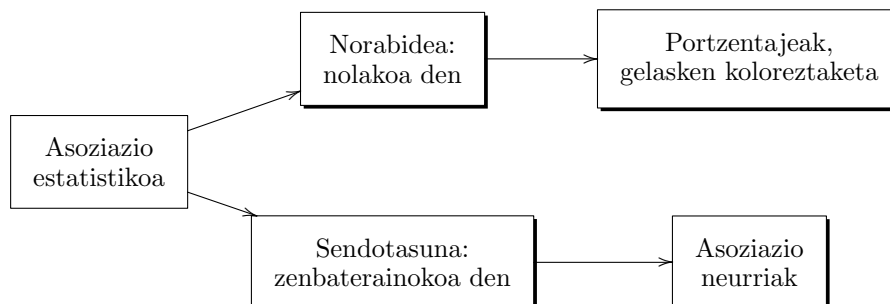
• ISO BAI

Tamaina	Mozkinak		Guztira
	Igo	Jaitsi	
Txikia	12	152	164
Ertaina	64	50	114
Handia	23	2	25
Guztira	99	204	303

Bi tauletarako gamma, lambda eta khi-karratuan oinarritutako neurriak kalkulatu eta interpretatu, norabidea ere aztertuz (khi-karratuaren kasuan, gelaskak koloreztatuz). Halaber, neurri horiek kalkulatu bi taulak bateratuta eta emaitzak interpretatu aurreko emaitzen aldean, interakzioa izateari buruz.



Irudia 8.5: Khi-karraturako kalkuluak gelaskatik gelaskara egiten dira: gelaska hiru zatitan egin eta irudiko balioak kalkulatu behar dira: O maiztasun enpirikoak hartu, eta E maiztasun teorikoak kalkulaturik (urdinez), $\frac{(O-E)^2}{E}$ kalkulatzten dira (gorriz). Balio horien batura eginez, gelaska guztietarako, khi-karratu kalkulatzten da. Balio horiexen erroa eginez, $\frac{O-E}{\sqrt{E}}$ balioak eskuratzen dira, dagokien zeinuarekin, asoziazioaren norabidea jakiteko baliagarri izango zaizkigunak.



Irudia 8.6: Asoziazio estatistikoaren bi aspektuak