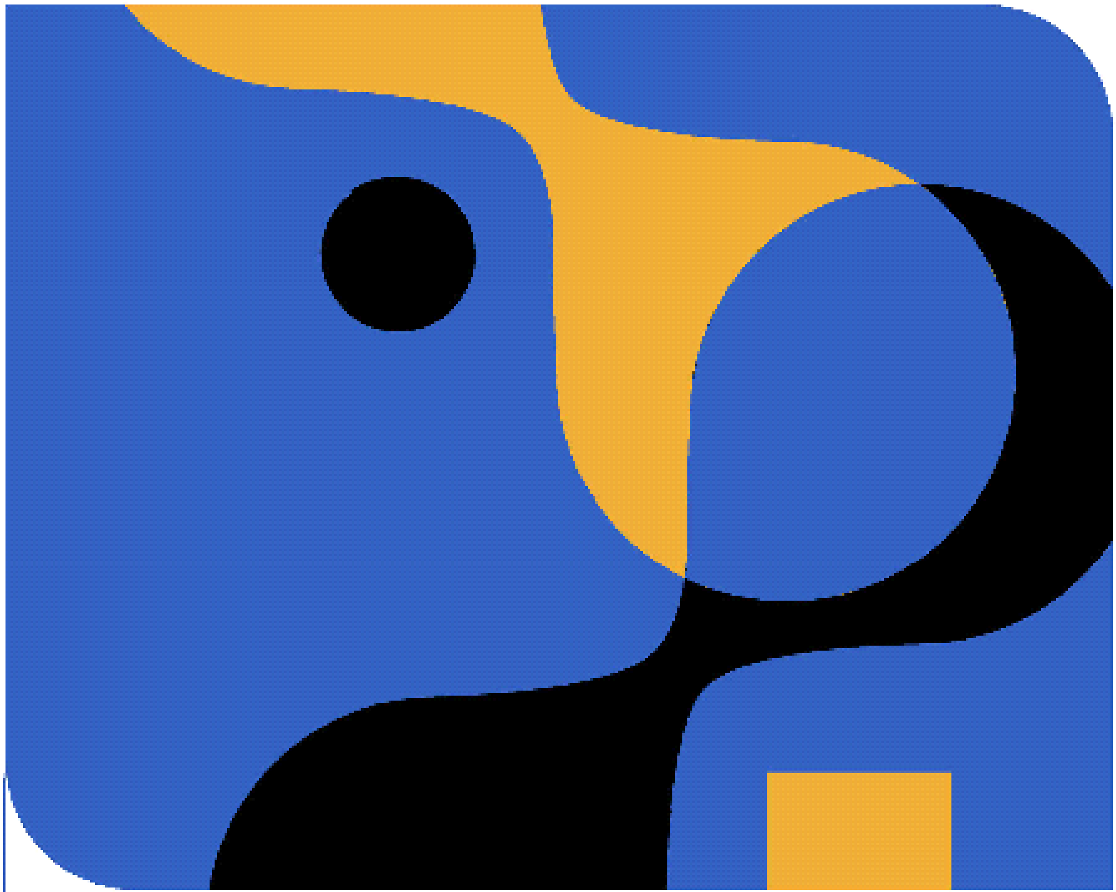
EUSKO JAURLARITZA
GOBIERNO VASCO

KULTURA SAILA

DEPARTAMENTO DE CULTURA

SOZIOLINGUISTIKA KLUSTERRA



ZUZENEKO BEHAKETAZ HIZKUNTZA-ERABILERA NEURTZEKO METODOLOGIAREN EREDU MATEMATIKOA

LAGINKETA ETA ESTIMAZIOA

Yosu Yurramendi Mendizabal *(Konputazio Zientziak eta Adimen Artifiziala Saila, Euskal Herriko Unibertsitatea)*
Olatz Altuna Zumeta *(Soziolinguistika Klusterra)*

2009ko abendua

Gaien Aurkibidea

1 Sarrera

Euskara zenbat erabiltzen den neurtzeko saiakera ugari egin dira azken urteotan. Batzuk eremu txikietan egin dira: ikastetxeetan edo lantegietan, esate baterako; beste batzuk, ordea, eremu handiagoetan: udalerrietan, eskualdeetan edota Euskal Herri osoan. Saio horien artean, zuzeneko behaketan oinarritutakoak dira gure interesa piztu dutenak. *Zuzeneko behaketa* esaten zaio, izan ere, hiztuna jardun naturalean ari dela, behatzaile batek jardun horri buruzko informazioa jasotzen duelako, inolako interferentziarik gabe.

Jasotako informazioarekin datu-multzoa osatzen da eta, datuen azterketatik, erabilpenaren ideia bat lortzen da. Emaitzak ez dira erabat zehatzak, behatutakoa ez delako beha daitekeen guztia. Hala ere, neurketa horiek ezinbesteko baliabideak dira diagnostiko egokiak egiteko eta euskararen aurrera begirako ildo eraginkorrak zehazteko.

Soziolinguistika Klusterrak, Hizkuntza Politikako Sailburuordetzaren laguntzarekin, ahozko erabilera neurtzeko metodo hauetan sakondu nahi izan du, eragile ezberdinek baliatutako metodologiak bilduz eta aztertuz. Orain arte, metodologia hauetan lagina, gehienetan, datu-bilketa burutzen den ordu kopurura mugatu da, eta neurketak irauten duen denboran jasotzen diren elkarrizketek osatu dute. Horrela sortutako laginetatik ez da, besterik gabe, zehaztasun-mailarik edo errore-tarterik eratortzen. Ezin jakin, beraz, lortutako informazioak zenbaterainoko fidagarritasuna duen.

Oinarrizko eredu formal edo matematiko bat beharko genuke gabezia horri aurre egiteko. Eredu formal horrek modua eman behar liguke bi galdera nagusiri erantzuteko:

- Zenbat elkarrizketa neurtu edo behatu behar dira, jasotako informazioa adierazgarria izan dadin?
- Lortzen den informazioak zein zehaztasun-maila edo errore-tartea du erabilera errealarekiko?

Bi gaiak elkarri lotuta daude, bistan denez.

Egindako lana txosten honetan aurkezteko, honako bide hau jarraitu dugu. Sarreraren ondoren, 2. puntuan, eredu matematikoaren oinarriak azaltzen dira. 3. puntuan, binakako elkarrizketen analisi matematikoa jorratzen da. 4. puntuan, binakako elkarrizketak ez ezik, unibertsoko elkarrizketa guztiak hartzen dira kontuan. Bukatzeko, taulak eta ondorio nagusiak aurkezten dira.

2 Eredu matematikoaren oinarriak

Planteatutako bi galderei zorroztasunez heltzeko, beharrezkoa da testuingurua formalizatzea, hau da, elkarrizketetako hizkuntza-erabilera *eredu matematiko* batez adieraztea.

Lehenik eta behin, aztergai nagusia ondo zedarritzen saiatuko gara. Elkarrizketak dira azterketa-unitatea. Gure ikuspegitik, elkarrizketa bat biren edo gehiagoren artean gertatzen den ahozko harreman linguistikoa da. Datuak biltzean honako arau hauek finkatu ditugu:

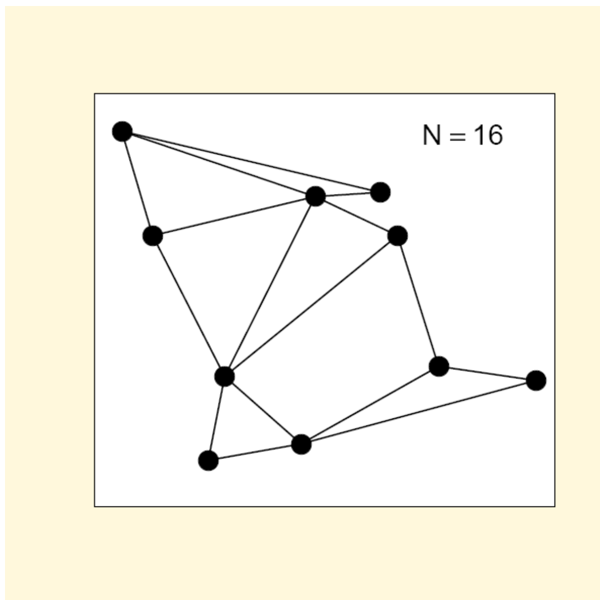
- Elkarrizketa guztien iraupena berdintzat hartzen da. Motzak ala luzeak diren, ez da aintzat hartzen. Behatzaileek ez diote erreparatzen elkarrizketen iraupenari.
- Elkarrizketa *bera* behin baizik ez da kontatzen. Horrek esan nahi du hizlaria eta entzuleak pertsona berak baldin badira eta erabilitako hizkuntza aldatzen ez bada, behin bakarrik kontatzen dela elkarrizketa hori. Murriztapen honek garrantzia du formalizazioari begira, izan ere, horrela, gerta litezkeen elkarrizketa guztien kopurua taldeko kideen kopuruaren

araberakoa da, eta denboraren aldagaia alde batera utz daiteke. Parte-hartzaileen artean beste batek hitza hartzeak (parte-hartzaile *aktiboa* aldatzeak) ez du esan nahi beste elkarrizketa bat hasi denik, non-eta ez den hizkuntza aldatzen.

- Elkarrizketak euskaraz ala erdaraz dira. Behatzaileek bi aukera baino ez dituzte elkarrizketa-hizkuntza jasotzeko. Elkarrizketa bereko hizkuntz alternantzia nahasia ez da jasotzen.
- Elkarrizketetako parte-hartzaileen kopurua ez da finkoa: bik edo gehiagok har dezakete parte elkarrizketetan.

Neurtze-arau hauek kontuan hartuta egin da formalizazio-saiakera. *Grafo* bidezko adierazpidea hautatu da giza-sare batean gertatzen diren elkarrizketak errepresentatzeko (ikus 1 irudia). Grafo horretan *erpinak* (nodoak) dira gizataldeko partaideak, eta erpinen arteko *ertzak* (arkuak) partaideen artean gerta litezkeen elkarrizketak.

Gizatalde baten baitako elkarrizketak horrelako grafoen bidez adierazteko, saiatu gara datu-



Irudia 1: Grafo bidezko adierazpidea.

bilketarako ezarritako arauak eredu formalera eramaten. Formalizazio-ahalegin orok berarekin dakar errealitatetik eredu baten baitara egokitu beharra, eta, hortaz, murriztapen edo sinplifikazio batzuk ezinbestean definitu behar dira. Ezarritako murriztapen hauek ez diote, gure ustez, ereduaren orokortasunari eta egokitasunari kalterik egiten.

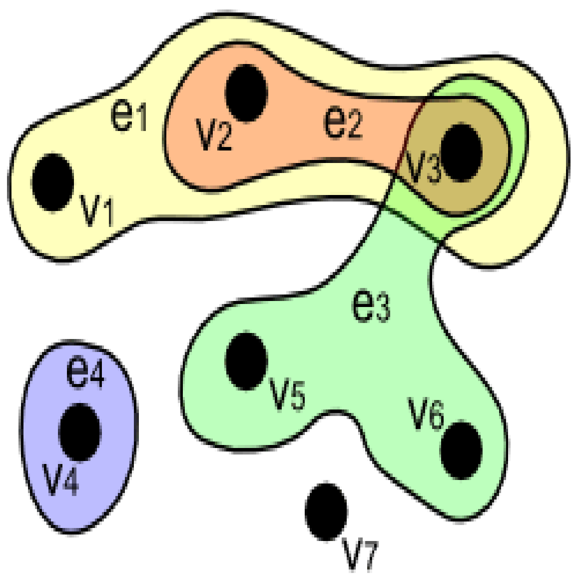
Murriztapenok deskribatzean saiatu gara, bide batez, datu-bilketan finkatutako arauekin loturak egiten (zerikusi estua baitute) eta grafoen interpretaziorako lagungarriak izan daitezkeen azalpenak ematen.

- Grafoko arkuek ez dute pisurik. Arku *ez-ponderatuak* dira, eta bi erpinen arteko *lotura* baizik ez dute adierazten, hau da, bi norbanakoen arteko elkarrizketa. Pisurik ez izateak elkarrizketen iraupena (edo beste edozein ezaugarri kuantifikagarri) aintzat ez hartzearen baliokidea da. Indar berdinekoak dira guztiak, nahiz eta badakigun, hizkuntza-erabilerari dagokionez, sinplifikazio handia dela.
- Elkarrizketa bera behin bakarrik kontatzen denez, aski da arku bakarra bi erpinen arteko elkarrizketa formalizatzeko (arku gehiago jartzeak elkarrizketa bera behin baino gehiagotan

agertzen dela esan nahiko luke). Grafoak une batean gertatzen diren elkarrizketak adierazten ditu. Bestalde, esan dugu datu-bilketako baldintzetan hizlaria aldatzeak ez duela beste elkarrizketa bat eragiten. Hori horrela, arku *ez-zuzenduak* aski dira elkarrizketak grafoan adierazteko. Bi hasierako sinplifikazio hauen arabera, esan dezakegu, esate baterako, eredu formal honetan minutu bakarreko bost elkarrizketek bost minutuko elkarrizketa batek halako bost balio dutela.

- Arkuei hizkuntzaren etiketa esleitzen zaie, horrela adieraziz zein hizkuntza erabili den elkarrizketa bakoitzean. Arkuen hizkuntza bitarra da: euskara ala erdara. Ez dago beste aukerarik. Biltze-arauetan esan dugunarekin bat etorriz, alternantzia nahasi etengabea ez da errepresentatu grafoan. Adierazpide grafikoan, euskaraz etiketatutako arkuak gorriz adierazten dira, eta erdaraz etiketatutakoak, beltzez.

- Elkarrizketa guztien parte-hartzaileen kopurua (bi, hiru, lau eta abar) ez da berdina izaten. Beraz, grafo klasikoak baino *hipergrafoak* beharko lirateke elkarrizketa mota horiek guztiak eredutu edo formalizatzeko (2 irudia).



Irudia 2: Hipergrafoa. *http://en.wikipedia.org/wiki/File:Hypergraph.gif* helbidetik hartua.

- Hasierako modelizazioan eta analisi matematikoan, elkarrizketa guztiak bi norbanakoen artekoak direla joko dugu. Sinplifikazio handia da, baina hasierako hurbilpen gisa beharrezkoa iruditzen zaigu. Abiapuntu horretan oinarrituta, saiatuko gara aurrerago sinplifikazio hau baztertzen eta eredu konplexuago bat eraikitzen.

- Ezinbestean *lagin-grafo* batekin lan egin beharra dago. Izan ere, binakako elkarrizketak baizik hartzen ez badira ere, ezinezkoa baita errealitatearen argazki osatu bat egitea, batez ere gizataldea txikia ez denean. Asko jota, egin daitekeen gauza bakarra da une batean ondo erreparatu une horretako elkarrizketei, hau da, partzialki ezagutu grafoa, eta ezagutza horretatik abiatuta grafo osoa irudikatu. Ikuspegi estatistiko batetik, denbora batean eta toki jakin batean gertatzen diren elkarrizketei erreparatzea elkarrizketen unibertsotik *lagin* bat hartu eta behatzea da. Laginarekin osatzen den lagin-grafoa grafo osoaren *adierazgarria* izatea nahi da, bere estimazio bat alegia, eta, beraz, ***zorizkoa*** beharko du izan.

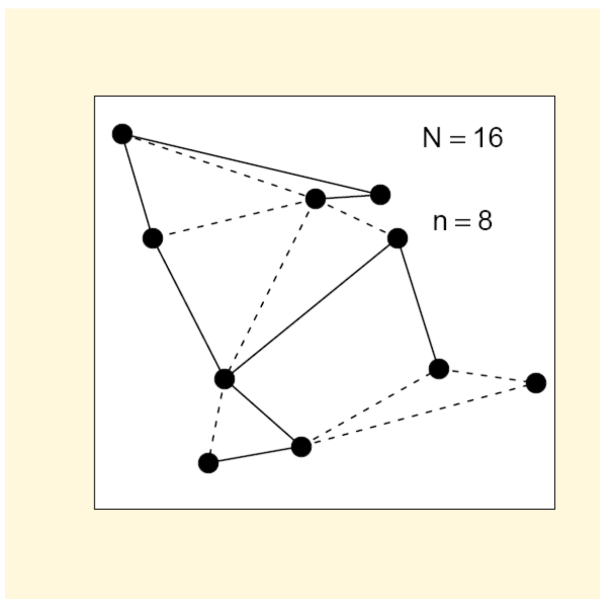
Lagina zorizkoa dela esaten denean, esan nahi da unibertsoko binakako elkarrizketa guztiek gertatzeko aukera (probabilitate, zori) berdina dutela behatzaileek erreparatu eta jasotzen dituzten unean eta tokian. Ohartzekoa da, batetik, lagin batean norbanako batek elkarrizketa bat baino gehiagotan har dezakeela parte eta, bestetik, behin elkarrizketa bat kontatuz gero ezin denez berriro kontatu, laginketa *itzulerarik gabekoa* dela.

- Giza-sare batean edozein norbanako ez da taldeko norbanako guztiekin erlazionatzen, azpitalde batekin baizik, batez ere taldea handia denean (*Dunbar*en zenbakia, *http://en.wikipedia.org/wiki/Dunbars number*). Dunbar-en zenbakiak adierazten du zenbat pertsonarekin eduki ditzakeen harreman egonkorrak norbanako batek. Ez da zenbaki zehatza, baina 150eko balioa hartzen da hurbilpen egokitzat.
  Batzuk jende gehiagorekin erlazionatzen dira beste batzuk baino, eta erlazio batzuk maizago gertatzen dira beste batzuk baino, norbanakoaren nortasuna eta bere giza-inguruaren arabera.

2.1 Grafoko ezaugarri esanguratsuenak

Formalizazio bidean, N batez izendatuko dugu unibertsoko elkarrizketa kopurua, eta n batez zorizko laginaren tamaina edo jasotako elkarrizketen kopurua (3 irudia). Elkarrizketa guztiek $\frac{n}{N}$eko probabilitatea (zoria) dute aukeratuak edo erreparatuak izateko.

Laginketaren unitate estatistikoak, beraz, elkarrizketak izango dira, grafoaren N arkuak edo



Irudia 3: N, elkarrizketen unibertsoaren tamaina, eta n, laginarena.

ertzak, alegia.

Praktikan, elkarrizketen unibertsoa zein den ez da ezaguna, ez eta N ere. Baina estatistikoki ezagutza lortzeko nahikoa izaten da handia dela jakitea. Edo bestela, unibertsoaren eredu zehatz batzuen arabera unibertsoa ezaguntzat jo daiteke, eta horren arabera ondorioak atera. Ondorioen ontasuna ereduaren egokitasunari lotuta egongo da.

Elkarrizketen grafo honetan, euskararen erabilera zenbaterainokoa den ezagutzeko, euskarazko eta erdarazko elkarrizketak zenbatzen dira. N_B batez izendatuko dugu euskarazko elkarrizketa-kopurua, eta n_B batez n tamainako laginean euskaraz erreparatutakoak.

4 irudian arku gorriz agertzen dira euskaraz jasotako elkarrizketak eta beltzez erdaraz jasotakoak.

Gizatalde baten euskararen erabilera elkarrizketen unibertsoan euskaraz gertatzen direnen ehunekoa da: $p = \frac{N_B}{N}$. Era berean defini daiteke laginean antzemandakoen ehunekoa: $\hat{p} = \frac{n_B}{n}$.

p ezezaguna da, eta laginaren bitartez kalkulatutako $\hat{p}$-k p horren *estimazio* bat ematen du.

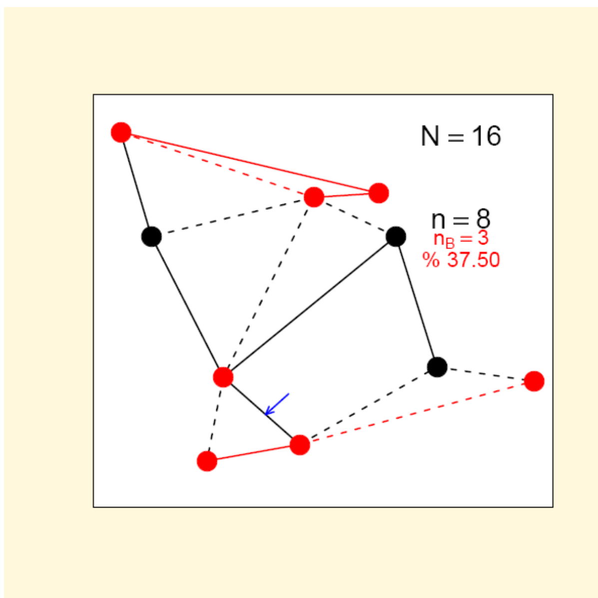
Orain artean definitutako eredu formalean eta aldagaien definizioan aurrera egiteko, badu garrantzirik gizataldeko euskaldunen edo elebidunen jokabideak. Deskribapenaren une honetan beste sinplifikazio bat egin beharrean gaude:

- Gizataldea bitan dago banatuta: A hizkuntza bakarrik ezagutzen dutenak, eleb$A$karrak, eta A eta B hizkuntzak ezagutzen dituztenak, ele$B$idunak. [1]

Ereduan ez dago, beraz, B hizkuntza bakarrik ezagutzen duenik, ez eta hizkuntza baten ezagutza partziala duenik ere. Gure kasuan A erdara da eta B euskara.

Gizataldeko bi azpitaldeen jokabideak garrantzia baduela esaten dugunean esan nahi dugu eleb$A$karrek eleb$A$karrekin (irudian beltzez) eta ele$B$idunekin (irudian gorriz) izan ditzaketen elkarrizketak nolakoak izango diren argi dagoela (erdaraz), baina ele$B$idunen arteko jokabide linguistikoa ez dagoela hain garbi (euskaraz ala erdaraz izan daiteke). Gezi urdin batez adierazten da, 4 irudian, pertsona ele*b*idunen arteko erdarazko elkarrizketa.

Jakina, halako ezagutza lortzeko ezinbestekoa da gizataldearen hizkuntzarekiko ezagutza no-

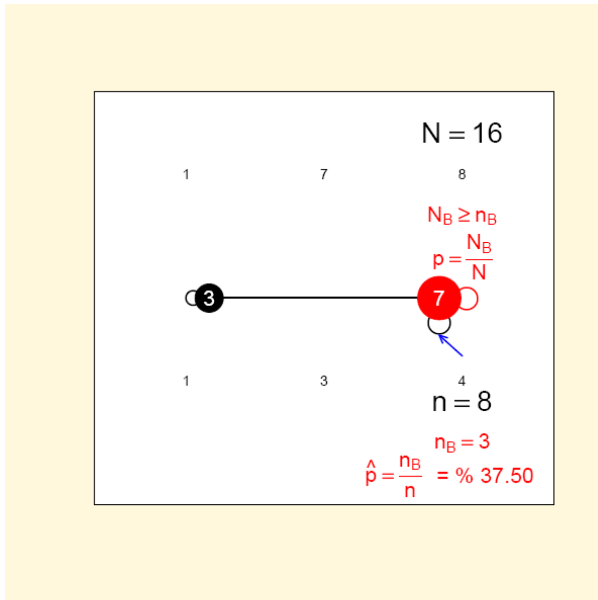


Irudia 4: Gorriz, pertsona ele$B$idunak eta euskarazko elkarrizketak; beltzez, pertsona eleb$A$karrak eta erdarazko elkarrizketak.

lakoa den jakitea. Zenbait kasutan talde jakin batean ele$B$idunak nortzuk diren jakitea badago; adibidez, talde txikietan, nahiz erabilera ezaguna ez izan.

[1] Anbiguotasun bat sortzen da A eta B letra larriak erabiltzean, azpitalde bat ote den ala azpitalde batek erabiltzen duen hizkuntza, baina korrespondentzia argi eta garbia denez, erdaldunak (elebakarrak) erdaraz (A) eta euskadunak (elebidunak) euskaraz (B) eta erdaraz, ulergarritasunaren mesedetan bere horretan utziko ditugu bi letrak, eta testuinguruaren arabera ulertuko ditugu. Are gehiago A erdaraz bakarrik mintzatzen den azpitaldearen tamaina izango da, eta B bi hizkuntzetan mintza daitekeen azpitaldearena. Gizatalde osoaren partaide kopurua, beraz, $(A + B)$ da.

Aipatutako ikuspegi soziolinguistikotik, norbanakoen jokabidea ezagutzea baino, talde osoarena da ezagutu nahi dena, baita aipatutako azpitalde ele$B$idunarena ere. Ondorioz, grafo osoa laburtu egiten da eta interesa duten azpitaldeen bitartez adierazten den grafo laburra antolatzen da (5 irudia). Grafo labur honetan erpinek azpitaldeak adierazten dituzte, eta bakoitzak bildutako



Irudia 5: Azpitaldeen grafo laburra.

partaideen pisua edo kopurua izango du. Arkuek (elkarrizketak) ere pisu ezberdinekoak izango dira, partaideek emandako ertz-kopuruen batura edo maiztasunen araberakoak. Ondorioz, grafo laburrean ertz batzuk 'bakarrizketa' modukoak dira: azpitalde batek bere buruarekin izan ditzaketenak.

Hasierako grafoaren laburpen hau eginez gero, talde txikietan ez ezik, talde handietan ere, zenbait kasutan, badago jakitea erpinen pisua zein den. Adibidez, estatistika ofizialen bitartez ezaguna izan daiteke gizatalde jakin batean zenbat eleb$A$kar eta zenbat ele$B$idun dagoen. Informazio hau garrantzizkoa izan daiteke laginaren tamaina erabakitzerakoan, aurrerago ikusiko denez.

2.2 Hasierako helburuak, eredu honetan

Gizatalde osoan euskararen erabilera zenbatekoa den ezagutzea da helburu nagusia, eta emaitza horretara indukzioaren bidez hurbiltzen gara, azaldu den legez. Azken emaitzak ezin du izan erabat zehatza. Horrenbestez:

- $\hat{p}$ $p$ren estimazio bat da. Irudiko adibidean: %37.5.
- Estimazio sakonagoa izaten da *zehaztasun-maila* bat gehitzen zaionean $\hat{p}$ren balioari: $\hat{p} \pm \delta$. Honekin esan nahi da $p$ren balioa $(\hat{p} - \delta, \hat{p} + \delta)$ tartean dagoela. Adibidean: $\%37.5 \pm \%12$.
- Baina hori ere ez da ziurra, eta orduan zehaztasun-maila horri ziurtasun- edo *konfiantza-maila* bat gehitu behar zaio: p $(\hat{p}-\delta, \hat{p}+\delta)$ tartean dago α' konfiantzaz. Adibidean: $\%37.5 \pm \%12$ %90eko konfiantzaz. Honek esan nahi du n tamainako laginketa behin eta berriz,

askotan, egingo bagenu %90etan harrapatuko genukeela $p$ren benetako balioa eraikitako balio-tarteetan.

p erabilera-mailarentzat lor daitezkeen zehaztasun- eta konfiantza-mailekin badute zerikusia, jakina, gerta litezkeen elkarrizketa-kopuruak (N) eta behatutakoen kopuruak (n).

Ereduaren argibide nagusiak azaldu ondoren, txosten honen helburu gisa planteatu ditugun bi galderak honela birformula daitezke:

- N (elkarrizketen unibertsoaren tamaina) jakin baterako, δ (zehaztasun-maila) eta α' (konfiantza-maila) finkatuz gero, erabaki euskararen erabilera-mailaren ($p$ren) estimaziorako zenbatekoa izan behar den gutxieneko n balioa (laginaren tamaina, erreparatu beharreko elkarrizketa-kopurua).
- N tamainako elkarrizketen unibertsotik zoriz lortutako n tamainako lagin batean euskararen erabilera-maila ($\hat{p}$) kalkulatu eta α' konfiantza-maila finkatuz gero, erabaki zein den δ zehaztasun-maila (errore-tasa ere esaten zaio: $p \pm \delta$).

Bi problemak lotuta daude, jakina.

3 Binakako elkarrizketak

Bi problema horien ebazpidea binakako elkarrizketen unibertsoan aztertuko dugu lehenbizi, atal honetan. Esan dugu euskararen erabilerari buruz bi azpitalde bereizten direla: erdaldunak (A norbanako eleb$A$karrak) eta euskaldunak (B norbanako ele$B$idunak). Bi hizkuntza dira, beraz, kontuan hartuko direnak: erdara (A) eta euskara (B), eta elkarrizketak bi azpitalde horietako kideen arteko binakakoak izango dira. Eredurik sinpleena da.

Jatorrizko bi galderen erantzuna 3.5 eta 3.6 ataletan aurkezten da. Horren aurretik, erreferentziatzat hartuko ditugun muturreko kasu batzuk landuko ditugu, jatorrizko galderen azalpenean lagungarri izango direlakoan. Muturreko kasu hauek formulatzeko beste hiru aurre-baldintza definitu ditugu. Hauexek:

- Norbanako bakoitzak izan dezakeen elkarrizketa-kopurua finkatu egiten da.
- Euskaldunak euskaldunekin euskaraz mintzatzen dira <u>beti</u>, eta gainontzeko elkarrizketak erdaraz egiten dira.
- A eta B kopuruak ezagunak dira.

Ikusiko dugunez, baldintza hauen guztien menpean ez da laginik behar euskararen erabilera-maila ezagutzeko. Aski izango da elebidunen proportzioa jakitearekin erabilera-proportzioa kalkulatzeko. Bi kasu aztertuko ditugu, muturreko portaerak definituz. Horrela jokatuta, erabilera-proportzioa zedarritzeko aukera izango dugu, eta muga horiek baliagarriak izango zaizkigu laginketaren tamaina kalkulatzeko.

3.1 Azpitaldeak erabat integratuta

A eta B azpitaldeak erabat integratuta egoteak esan nahi du denek denekin komunikatzeko joera dutela, azpitalde berekoekin komunikatzeko (edo, alderantziz, beste azpitaldekoekin komunikatzeko) joera markaturik ez dagoela. Integrazio erabatekoa dago bi azpitaldeen artean, jokabide endogamikoen edo exogamikoen gainetik.

Hori horrela, analisi matematikoari heltzeko, bi lan-hipotesi hartuko ditugu. Lehena: norbanako bakoitza beste gainontzeko guztiekin mintza daiteke (komunikatzeko joera erabatekoa). Bigarrena: norbanako bakoitza norbanako-kopuru jakin batekin mintza daiteke (komunikatzeko joera partziala). Azter ditzagun bi eredu hauek banan-banan.

3.1.1 Komunikatzeko joera erabatekoa

Taldeko norbanako bakoitza beste gainontzeko guztiekin mintza daiteke (6 irudia).

Problema hain sinpleki azalduta, elkarrizketen unibertsoa zenbaki konbinatorio hau da:

$$N = C(A+B,2) = \tfrac{1}{2} \cdot (A+B) \cdot (A+B-1)$$

Beste formula baliokide hau interesgarriagoa da:

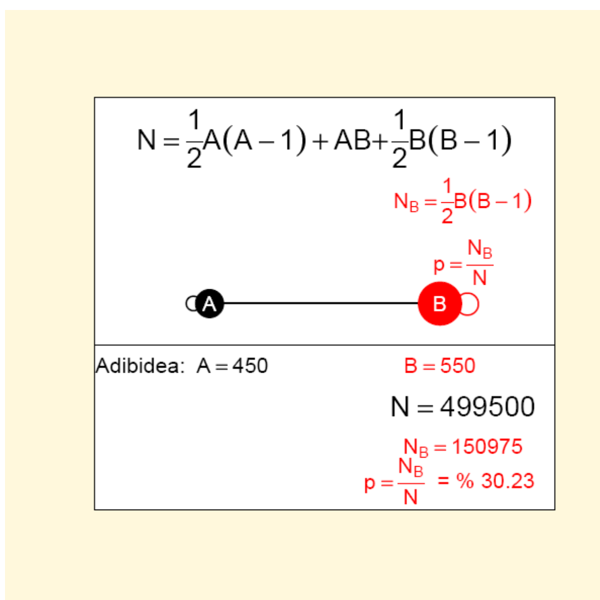
$$\begin{aligned} N = C(A+B,2) &= C(A,2) + C(A,1) \cdot C(B,1) + C(B,2) = \\ &= \tfrac{1}{2} \cdot A \cdot (A-1) + A \cdot B + \tfrac{1}{2} \cdot B \cdot (B-1) = \\ &= \tfrac{1}{2} \cdot (A \cdot (A-1) + 2 \cdot A \cdot B + B \cdot (B-1)). \end{aligned}$$

Determinismo honen arabera, populazioaren erabilera-maila (p), binakako elkarrizketei dagokionez, hauxe da:

$$p = \frac{B \cdot (B-1)}{A \cdot (A-1) + 2 \cdot A \cdot B + B \cdot (B-1)}$$

Adibidez: $A = 450, B = 550$ badira, $N = C(1000,2) = 499500$ da, eta $p = 0.3022523 \cong \%30.23$.

Kasu hau beste A eta B balio batzuekin ere kalkula daiteke. Esate baterako, proportzioak

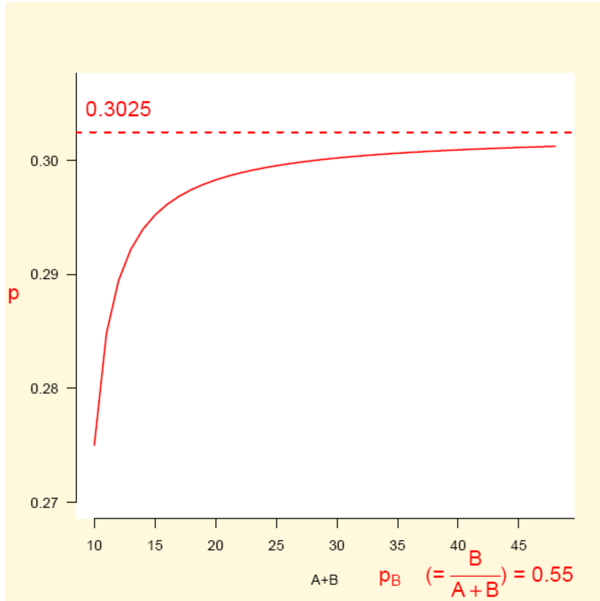


Irudia 6: Grafo laburra: azpitaldeak erabat integratuta.

mantenduz, $A = 45$ eta $B = 55$ badira, erabilera-maila %30 izango da, eta $A = 4500$ eta $B = 5500$ badira, %30.25.

Muturreko erabilera-maila, beraz, A eta B balioen arteko proportzioaren arabera dagoela ematen du.

Gainera, zenbat eta handiagoak izan A eta B, beraien proportzioari eutsiz, p limite baterantz



Irudia 7: $p_B$ri eutsita, p limite baterantz doa $(A+B)$ handitu ahala.

joaten da; hau da, limite horretatik gero eta gertuago dago $(A+B)$ren handitzearekin batera (esango genuke $(A+B)$ 'infiniturantz' joaten denean). Adibidean, $\frac{B}{A+B} = 0.55(\%55)$ dela, jomuga %30.25 da (ikus 7 irudia). Beste proportzio batzuk erabiliz, gauza bera antzeman daiteke. Adibidez, $\frac{B}{A+B} = 0.30(\%30)$ balitz, erabilera-mailaren limitea %9.00 izango litzateke, eta $\frac{B}{A+B} = 0.80(\%80)$ izango balitz, %64.00.

Orokorrean, $p_B = \frac{B}{A+B}$ izanda, limite hori p_B^2 da, ele$B$idunen proportzioaren balio karratua, hain zuzen ere.

$$B = p_B \cdot (A+B) \quad A = (1-p_B)\cdot(A+B)$$
$$p = \frac{B\cdot(B-1)}{A\cdot(A-1)+2\cdot A\cdot B+B\cdot(B-1)}$$
$$p = \frac{(A+B)\cdot p_B^2 - p_B}{(A+B)-1}$$
$$\lim_{(A+B)\to\infty} p = p_B^2$$

8 irudian ikus daiteke nola aldatzen den $p = p_B^2$ limite p_B ezberdinentzako.

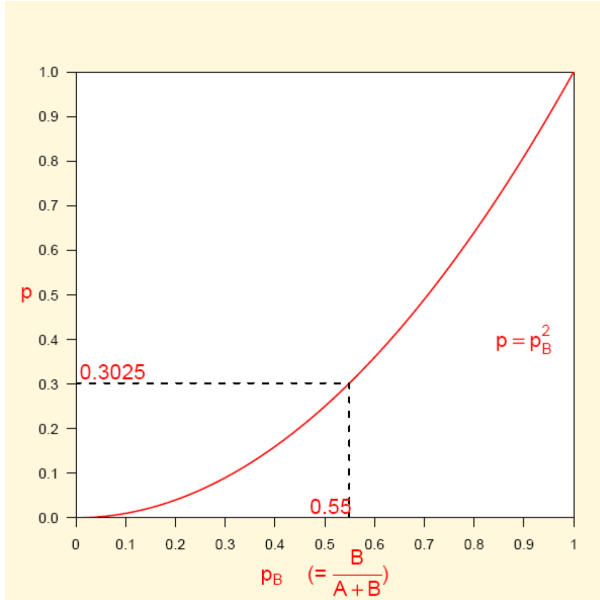
Limite horiek ontzat har daitezke gizataldearen ia edozein $(A+B)$ kopururentzat. Izan ere, froga daiteke, adibidez, 26 baino handiagoa izanez gero dagokion probabilitatea limite horretatik 0.01 baino gertuago dagoela.

Limitearen hurbiltasuna:

$$B = p_B \cdot (A+B) \quad A = (1-p_B)\cdot(A+B)$$
$$p_B^2 - \frac{B\cdot(B-1)}{A\cdot(A-1)+2\cdot A\cdot B+B\cdot(B-1)} < \epsilon$$
$$p_B^2 - \frac{(A+B)\cdot p_B^2 - p_B}{(A+B)-1} < \epsilon$$
$$(A+B) > 1 + \frac{p_B\cdot(1-p_B)}{\epsilon}$$

$p_B\cdot(1-p_B)$ren balio handiena $p_B = 0.5$ denean gertatzen da: $(A+B) > 1 + \frac{0.5\cdot0.5}{\epsilon} = 1 + \frac{1}{4\cdot\epsilon}$
Adibidez, $\epsilon = 0.01 \quad (A+B) > 26 \quad \epsilon = 0.001 \quad (A+B) > 251$

Halako behe-borneak guztiz orokorrak dira, baina zehaztu daitezke egoerak finkatuz gero. Adibidez:



Irudia 8: $p$ren bilakaera, $p_B$ren arabera.

$$p_B = 0.20 \;\; \epsilon = 0.01 \text{ baldin badira, } (A+B) > 17. \text{ Ia beti, beraz.}$$
$$p = 0.10 \;\; \epsilon = 0.01 \text{ baldin badira, } (A+B) > 10 \text{ (beti),}$$
$$\text{eta } p_B = 0.05 \;\; \epsilon = 0.01 \text{ baldin badira, } (A+B) > 5 \text{ (beti).}$$

3.1.2 Komunikatzeko joera partziala

Egoera honetan norbanakoen balizko mintzakideen kopurua finkatu egingo dugu, eta M deituko diogu balio horri. Hortaz, taldeko norbanako bakoitza beste gainontzeko Mrekin mintza daiteke ($M \leq (A+B-1)$).

Norbanakoen komunikatzeko gaitasuna murriztu egin da, baina guztiek elkarrizketa-kopuru berdina dute. Grafoko ereduaren arabera, nodo bakoitzak arku-kopuru berdina izango luke, eta grafoak erregular izaten jarraituko luke. Baldintza honek kalkuluak erraztu egiten ditu:

$$N = \tfrac{1}{2} \cdot (A+B) \cdot M \leq \tfrac{1}{2} \cdot (A+B) \cdot (A+B-1) = C(A+B, 2)$$

M eta $(A+B-1)$ balioen arteko proportzioa K ehuneko batez adieraz daiteke: $\frac{M}{A+B-1} = \frac{K}{100}$. ‘Komunikatzeko ahalmena’ deituko diogu Kri. Adibidez, $K = 25$ hartuz gero, norbanako bakoitza gizataldearen %25ekin elkarrizketatu zitekeela ulertuko genuke. Ondorioz, elkarrizketen kopurua proportzio berdinean murriztuko da:

$$N = \tfrac{1}{2}(A+B) \cdot (A+B-1) \cdot \tfrac{K}{100} = C(A+B, 2) \cdot \tfrac{K}{100}$$

Euskarazko elkarrizketak kalkulatzerakoan eragina du, bistan denez, mintzakideak zein azpitaldetakoak (Akoak edo Bkoak) diren. Hori dela eta, Mren banaketa finkatu beharrean gaude. Har dezagun honako hipotesi hau:

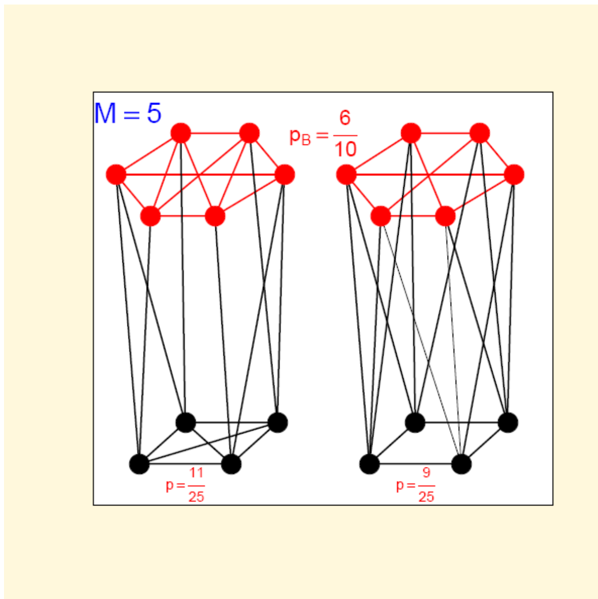
- Ele$B$idun bakoitzaren M mintzakideak proportzionalki banatzen dira bi azpitaldeetan.

Honekin esan nahi da edozein norbanakoren mintzakideak bi azpitaldeen tamainaren proportzioan banatzen direla, azpitalde batekoekin elkarrizketatzeko joera markaturik gabe (*orekatua*, nolabait). 3.1.1 ataleko egoera, komunikatzeko joera erabatekoa, honako honen kasu partikular bat besterik ez da ($M = A + B - 1$).

Horren arabera, A azpitaldeko norbanakoen M elkarrizketak honela banatuko dira: A azpitaldekoekin $\frac{A-1}{A+B-1} \cdot M = (A-1) \cdot \frac{K}{100}$ eta Bkoekin $\frac{B}{A+B-1} \cdot M = B \cdot \frac{K}{100}$, eta B azpitaldeko norbanakoena: A azpitaldekoekin $\frac{A}{A+B-1} \cdot M = A \cdot \frac{K}{100}$ eta Bkoekin $\frac{B-1}{A+B-1} \cdot M = B \cdot \frac{K}{100}$.

9 irudiko bi kasuek ez dute zehatz-mehatz betetzen hipotesian planteatutakoa, baina eskuinekoa da bietatik gehien hurbiltzen dena: $A = 4, B = 6$ eta $M = 5$; eta $p = \frac{9}{25} = 0.36$. Kasu idealean, ordea, $N = C(4+6,2) = 45$ da, eta $p = \frac{C(6,2)}{C(10,2)} = \frac{15}{45} \cong \%33.33$. Azpitaldeen tamainak handiagoak direnean, probabilitateen aldetik kasu idealera gehiago hurbiltzen den kasu bat bilatzea ez da zaila.

Populazioaren euskararen erabilera-maila kalkulatzeko oraingoan ere ez da laginik behar, eta M



Irudia 9: Mintzakideen banaketa. Eskuinekoa *orekatuagoa* da.

edo Krekin ez du zerikusirik:

$$p = \frac{\frac{1}{2} \cdot B \cdot (B-1) \cdot \frac{K}{100}}{\frac{1}{2} \cdot A \cdot (A-1) \cdot \frac{K}{100} + A \cdot B \cdot \frac{K}{100} + \frac{1}{2} \cdot B \cdot (B-1) \cdot \frac{K}{100}} = \frac{B \cdot (B-1)}{A \cdot (A-1) + 2 \cdot A \cdot B + B \cdot (B-1)} = p_B^2$$

Honenbestez, komunikatzeko joera partziala izanik ere, elkarrizketak modu orekatuan banatzen badira azpitaldeen artean, $p = p_B^2$ betetzen da.

3.2 Azpitaldeak erabat desintegratuta

Esango dugu A eta B azpitaldeak erabat desintegratuta daudela elkarrizketa guztiak azpitalde barrukoak badira eta ez badago elkarrizketarik azpitaldeen artean: ele$B$idunak elebidunekin eta eleb$A$karrak elebakarrekin mintzo dira. Azpitaldeak 'itxiak' dira.

Aurreko ereduan, azpitaldeak integratuta zeudela, egindako sinplifikazioei eutsiko diegu:

- Norbanako guztiak norbanako-kopuru berdinarekin erlazionatzen dira.
- Euskaldunak euskaldunekin euskaraz mintzatzen dira <u>beti</u>, eta gainontzeko elkarrizketak erdaraz egiten dira.

- A eta B kopuruak ezagunak dira.

Bi lan-eredu definituko ditugu honakoan ere. Lehena: norbanako bakoitza (bere azpitaldeko) beste gainontzeko guztiekin mintza daiteke (komunikatzeko joera erabatekoa). Bigarrena: norbanako bakoitza (bere azpitaldeko) norbanako-kopuru jakin batekin mintza daiteke (komunikatzeko joera partziala).

3.2.1 Komunikatzeko joera erabatekoa

Eredu honetan, taldeko norbanako bakoitza bere azpitaldeko gainontzekoekin mintza daiteke. Kontuan hartu behar da, ordea, mintzakide-kopurua ez dela berdina izango bi azpitaldeetan, bien tamainak, A eta B, berdinak ez badira behintzat:

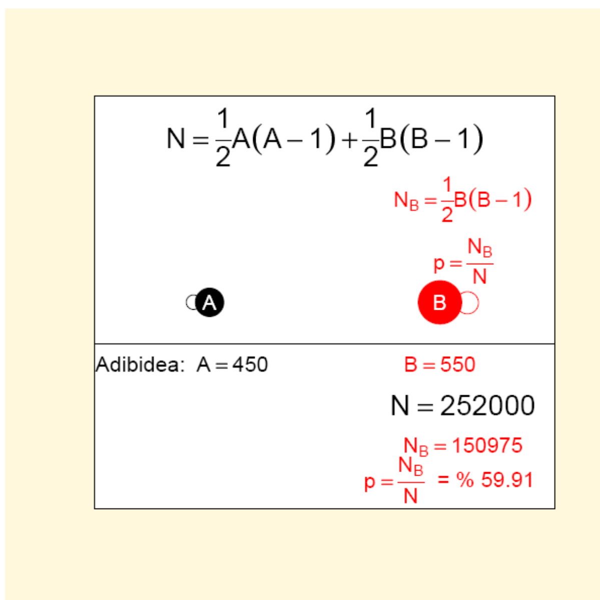
$$N = C(A,2) + C(B,2) = \frac{1}{2} \cdot A \cdot (A-1) + \frac{1}{2} \cdot B \cdot (B-1)) = C(A+B,2) - A \cdot B$$

Kasu honetan erabilera-maila honelaxe kalkulatzen da:

$$p = \frac{B \cdot (B-1)}{A \cdot (A-1) + B \cdot (B-1)}$$

Har ditzagun honakoan ere $A = 450, B = 550$ balioak adibide gisa. Kasu honetan $N = C(450,2) + C(550,2) = 252000$ da, eta $p = 0.5991071 \cong \%59.91$.

Ohartzekoa da azpitaldeen arteko integrazioa erabatekoa zen kasuarekin alderatuz gero, erabilera-

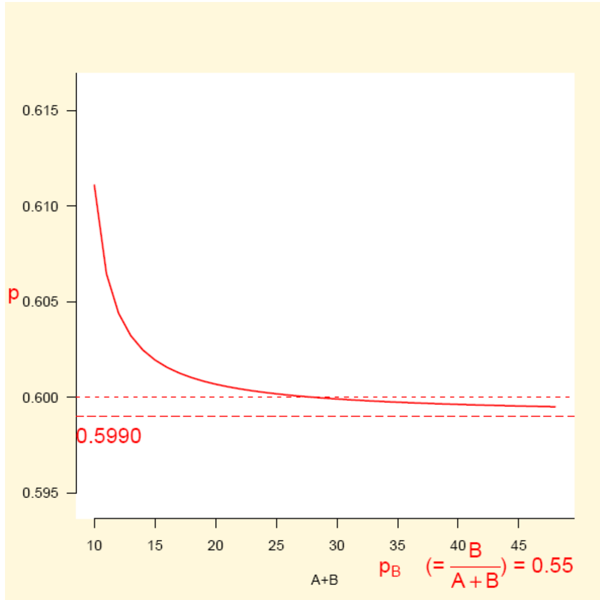


Irudia 10: Grafo laburra: azpitaldeak erabat desintegratuta.

maila beti izango dela handiagoa, kalkulua egitean proportzioaren izendatzailea txikiagoa delako; lehen, $p = \frac{B \cdot (B-1)}{A \cdot (A-1) + 2 \cdot A \cdot B + B \cdot (B-1)}$; orain, $p = \frac{B \cdot (B-1)}{A \cdot (A-1) + B \cdot (B-1)}$.

Muturreko kasu honetan ere beste A eta B kopuru batzuekin kalkula daiteke erabilera-maila, beren arteko proportzioari eutsiz. Esate baterako, $A = 45$ eta $B = 55$ badira, erabilera-maila %60 izango da , eta $A = 4500$ eta $B = 5500$ badira, %59.90. Muturreko kasu honen erabilera-maila, beraz, A eta Bren arteko proportzioaren arabera dago.

Beste proportzio batzuekin ere gauza bera gertatzen da, hots, erabilera-mailak limite jakin bat du. Adibidez, $p_B = 0.30$ (%30) balitz, erabilera-mailaren limitea %15.52 izango litzateke, eta



Irudia 11: $p_B$ri eutsita, p limite baterantz doa $(A+B)$ handitu ahala.

$p_B = 0.80$ (%80) izango balitz, %94.12. Edozein p_B proportziotarako $p$ren konbergentzia bat dago, 3.1 ataleko muturreko kasuan bezalaxe. Zenbat eta handiagoak izan A eta B, beraien arteko proportzioari eutsiz gero, orduan eta gertuago dago p limite jakin batetik: $p_B = \frac{B}{A+B}$ izanda, muga hori $\frac{p_B^2}{1-2 \cdot p_B \cdot (1-p_B)}$ da.

$$B = p_B \cdot (A+B) \quad A = (1-p_B) \cdot (A+B)$$
$$p = \frac{B \cdot (B-1)}{A \cdot (A-1) + B \cdot (B-1)}$$
$$p = \frac{(A+B) \cdot p_B^2 - p_B}{2 \cdot (A+B) \cdot p_B^2 - 2 \cdot (A+B)) \cdot p_B + (A+B) - 1}$$
$$\lim_{(A+B) \to \infty} p = \frac{p_B^2}{1-2 \cdot p_B \cdot (1-p_B)}$$

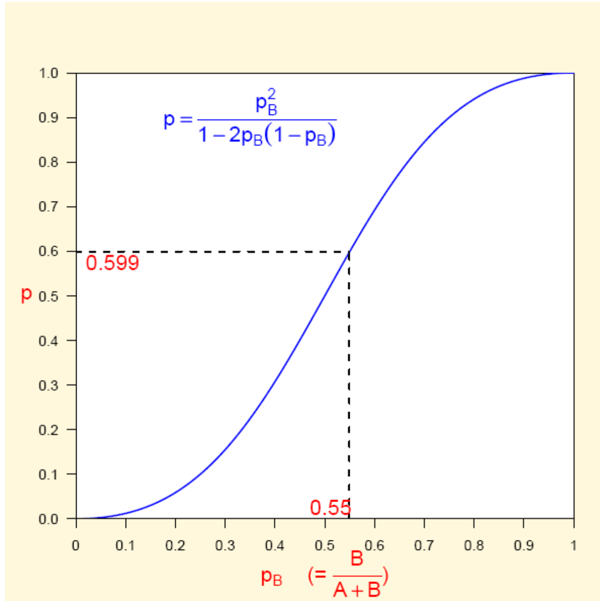
Adibidean, $p_B = 0.55$ (%55)) dela, limitea %59.90 da.

12 irudian p_B ezberdinentzako limitea azaltzen da.

Gizataldeek ez dute zertan handiak izan limite hau praktikoki erabili ahal izateko, azken batean, $(A+B) > 27$) nahikoa baita balio horretatik gutxienez %1 (0.01)era egoteko:

$$B = p_B \cdot (A+B) \quad A = (1-p_B) \cdot (A+B)$$
$$\frac{p_B^2}{2 \cdot p_B^2 - 2 \cdot p_B + 1} - \frac{B \cdot (B-1)}{A \cdot (A-1) + B \cdot (B-1)} < \epsilon$$
$$\frac{p_B^2}{2 \cdot p_B^2 - 2 \cdot p_B + 1} - \frac{(A+B) \cdot p_B^2 - p_B}{2 \cdot (A+B) \cdot p_B^2 - 2 \cdot (A+B)) \cdot p_B + (A+B) - 1} < \epsilon$$
$$(A+B) > \frac{1}{2 \cdot p_B^2 - 2 \cdot p_B + 1} + \frac{2 \cdot p_B^3 - 3 \cdot p_B^2 + p_B}{(2 \cdot p_B^2 - 2 \cdot p_B + 1)^2} \cdot \frac{1}{\epsilon}$$

Adibidez, $\epsilon = 0.01$ hartuz gero, $(A+B) > 26$ ateratzen da, eta $\epsilon = 0.001$ekin $(A+B) > 251$, azpitaldeak integratuta zeuden muturreko kasuan bezalaxe.



Irudia 12: $p$ren bilakaera, $p_B$ren arabera.

3.2.2 Komunikatzeko joera partziala

Taldeko partaide bakoitza bere azpitaldeko Mrekin bakarrik mintza daiteke: $M \leq min(A, B)-1$. Kasu honetan elkarrizketa kopurua hau da:

$$N = \frac{1}{2} \cdot A \cdot M + \frac{1}{2} \cdot B \cdot M$$

Alabaina, lehen bezala, Mk ez du eraginik erabilera-tasaren balioan:

$$p = \frac{\frac{1}{2} \cdot B \cdot M}{\frac{1}{2} \cdot A \cdot M + \frac{1}{2} \cdot B \cdot M} = \frac{B}{A+B} = p_B$$

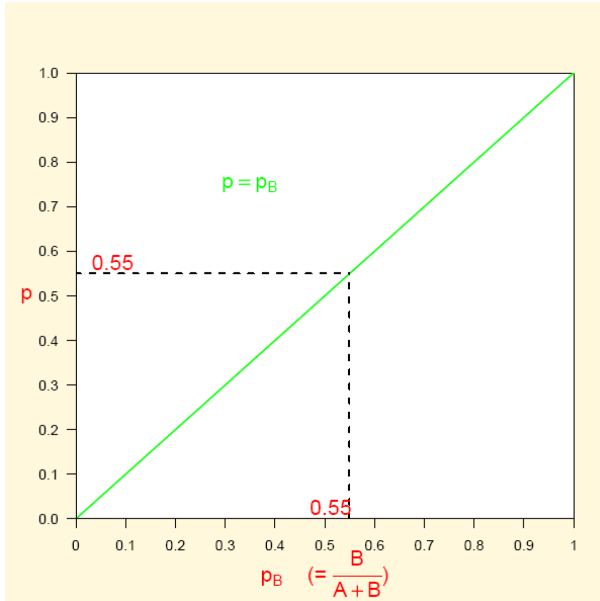
13. irudian p_B ezberdinentzako limitea azaltzen da. Hemen nabarmendu behar dena da, A eta B ezberdinak izanez gero, nahiz M bientzat berdina izan, K, komunikatzeko ahalmena, oso ezberdina izan daitekeela. Izan ere, eleb$A$karren azpitaldean $K_A = \frac{M}{A-1} \cdot 100$ izango da, eta ele$B$idunenean $K_B = \frac{M}{B-1} \cdot 100$.

3.3 Zenbait ohar

3.3.1 Probabilitateen araberako ikuspegia

Euskararen erabilera-mailari mugak jartzeko, ikusi berri dugunez, nahikoa da p_A ($\frac{A}{A+B}$) eta p_B ($\frac{B}{A+B}$) balioen arteko proportzioa ezagutzea. Are gehiago, $(A+B)$ tamainak ez du handia izan behar proportzio horiek jakinda erabilera-mailaren limitea nahiko zehatza izan dadin. Hori horrela da bai A eta B azpitaldeak integratuta daudenean, bai desintegratuta daudenean.

Integrazio-maila erabatekoa denean, eta norbanako bakoitzak komunikatzeko duen jokabidea orekatua bada (bere azpitaldekoenganako joera markaturik gabe, ezta beste azpitaldekoenganakoa ere), limitea p_B^2 da. Honek badu azalpen erraz bat probabilitateen bitartez: gizataldetik bi norbanako zori hutsez aukeratuko bagenitu, eta $p_B = \frac{B}{A+B}$ baldin bada ele$B$idun bat aukeratzeko probabilitatea, biak elebidunak izateko probabilitatea (edo, ezarri ditugun baldintzen



Irudia 13: $p$ren bilakaera, $p_B$ren arabera, azpitaldeak desintegratuta, eta komunikatzeko joera partziala, baina denentzat berdina, denean.

arabera, elkarrizketa euskaraz izatekoa) p_B^2 izango litzateke. Formularen irakurketan bi zorizko hautaketak *independente*tzat hartzen dira (lehenengo hautaketak ez du eraginik bigarren hautaketan).

Emaitza hau hobeto ulertzearren, berriro azpimarratuko dugu hasieratik ibilitako bidea.

Gizataldearen banaketa bitarra ezaguna bada (A eta B ezagunak) Nren handitasuna antzeman daiteke, praktikoki karratuaren erlazioa betetzen baitu ($N = C(A+B,2) = \frac{1}{2} \cdot (A+B) \cdot (A+B-1) \cong \frac{1}{2} \cdot (A+B)^2$, edo bestela $\frac{1}{2} \cdot (A+B)^2 \cdot \frac{K}{100}$). Gizataldea handia ez bada ere, elkarrizketa-kopurua oso handia gerta daiteke, eta kalkuluetarako infinitutzat hartu ($A = 450$, $B = 550$ azpitaldeek $N = 499500$ binakako elkarrizketa eman dezakete, eta $K = 10$ bada ere, $N = 49950$ balioak oso handia izaten jarraitzen du). Dena dela, Nk ez du handia izan behar lortutako formula ia akatsik egin gabe erabili ahal izatzeko.

Probabilitate-kalkuluaren ikuspegitik esan daiteke $A = 450$ bola beltz eta $B = 550$ bola gorri dituen kutxa batetik *itzuli gabeko* bi ateralditan bi gorri (bi euskaldun behar direnez elkarrizketa bat euskaraz lortzeko) suertatzeko probabilitatea $\frac{550}{1000} \cdot \frac{549}{999} = 0.3022523 \cong 0.3023$ dela. Populazioa handia izanda, ez du garrantzi praktikorik bolak itzultzeak ala ez. Hortaz, ateratako lehenengo bola kutxara *itzulita*, biak gorriak suertatzeko probabilitatea: $\frac{550}{1000} \cdot \frac{550}{1000} = 0.3025 \cong 0.3023$ da. Ateratako lehenengo bola kutxara ez itzulita ere, horrek ez du eragin praktikorik bigarren ateraldiko emaitzan.

N handia izanez gero, tamaina zehatzak axolik ez duenez, ez da beharrezkoa A eta Bren tamainak ezagutzea, beren arteko proportzioak baizik. Azpitaldeen kopuru absolutuen ordez erlatiboak ezagutzen direnean, adibidez, $A = \%45$ eta $B = \%55$, populazioa infinitutzat hartzen da. Halaber, gehienezko euskarazko elkarrizketen probabilitatea kalkulatzerakoan ere antzeko arrazoibidea erabiltzen da: $0.55 \cdot 0.55 = 0.3025$, eta antzeko emaitza lortzen, beraz.

Halako probabilitate-kalkuluak zuzenean egiten direnean, azpitaldeen tamainari erreparatu gabe alegia, atzean dagoen sinplifikazioa da norbanako guztiak komunikazio orekatu bat dutela beste norbanakoekin.

Azpitaldeak erabat desintegratuta daudenean, norbanakoen komunikatzeko joera erabatekoa dela joz gero, emaitzen ulerkera zailagoa da probabilitateen ikuspegitik. Kontuan hartu behar da, orokorrean, azpitalde batekoek komunikatzeko gaitasun handiagoa dutela beste azpitaldekoek baino, beren azpitaldea handiagoa delako. Beraz, azpitalde handiko norbanakoek elkarrizketa gehiagotan parte hartzeko aukera dute, eta elkarrizketak unitate estatistikoak izanda, hautatuak izateko probabilitate handiagoa dute (guztienak ez dira berdinak, uniformetasunik ez dago norbanakoengan). Ondorioz, ez da hain zuzena p eta p_B probabilitateen arteko erlazioa. Salbuespena, jakina, $A = B$ kasuan gertatzen da. Kasu horretan, bi azpitaldekoek norbanako-kopuru berdinarekin komunikatzeko gaitasuna dutela, $p_B = 0.50$ eta $p = p_B = 0.50$ da.

Aldiz, eredu desintegratuan, komunikatzeko joera partziala eta guztientzat berdina izateak esan nahi du, probabilitatearen ikuspegitik, norbanako guztiek aukeratuak izateko probabilitate berdina dutenez (elkarrizketetan parte hartzeko gaitasun bera dutelako), elkarrizketa bat euskaraz gertatzeko probabilitatea eta euskaldun bat (ele$B$idun) hautatua izateko probabilitatea berdinak direla. Behin euskaldun bat hautatuz gero, 'beti' euskaraz mintzatzen denez (euskaldunak euskaldunekin baizik ez dira mintzatzen), beste edozeinekin izango duen elkarrizketa euskaraz izango da.

3.3.2 Euskaldunak euskaldunekin ez dira 'beti' mintzatzen euskaraz

Euskararen erabilera-maila muturreko egoerak definituz zedarritu dugu. Muturreko egoera horietarako honako sinplifikazio hau erabili dugu: euskaldunak euskaraz mintzatzen dira 'beti' euskaldunekin. Beste modu batera esanda: euskaldunek euskararekiko duten leialtasuna erabatekoa da. Errealitatera hurbiltzeko, beharrezkoa da baldintza gogor eta irreal hori leuntzea.

Leialtasuna erabatekoa ez, partziala izaki, beharrezkoa da *leialtasun-maila*ren kontzeptua analisi honetara ekartzea. Xabier Isasik eta Arantxa Iriartek (BAT aldizkariaren 28. alean, 54. orria, 1998) Txillardegiren eredu matematikoa aztertzean, honela definitzen dute *leialtasuna*: '*zehaztutako giza talde jakineko elebidunen artean B hizkuntza erabiltzeko batenaz besteko joera*'. Batez besteko hori norbanako guztiei aplikatuko zaie eredu honetan.

Horrela formulatuko dugu *leialtasun-maila*, L, balioa:

- Euskaldunek euskaldunekin euskaraz izaten dituzten elkarrizketak ehuneko L dira.

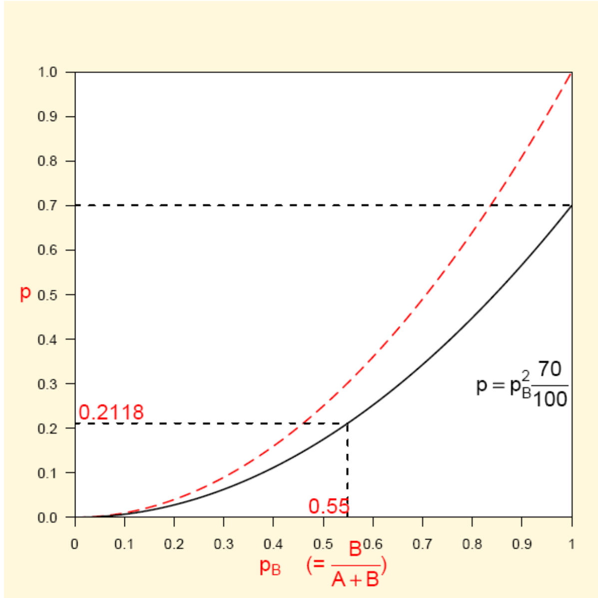
Leialtasun-maila ez bada erabatekoa ($L < 100$), oro har, euskararen erabilera-maila txikiagoa izango da.

Sinplifikazio bat leundu egin dugu, baina hala ere zenbakitan adierazten dugu, eta moldaketa honek sinplifikazio bat izaten jarraitzen du. Lehengo 'beti'ko kasua, orain kasu partikular bat besterik ez da: $L = 100$.

Orain $N_B = B \cdot (B-1) \cdot \frac{L}{100}$ da eta, beraz, $p = p_0 \cdot \frac{L}{100}$, non p_0 aurreko analisietan (leialtasun-maila erabatekoa zenekoetan) aipatutako p balioa den (ikus 14 irudia). Ohartzekoa da: $p \leq p_0$.

Leialtasun-maila (L) tipikotzat edo estandartzat har daiteke, eta, orokorra dela kontsideratuz, norbanako guztiei aplikatu. Matematikoki: $p_B = p_{B0} \cdot \sqrt{\frac{L}{100}}$, non p_{B0} aurreko analisietan aipatutako p_B balioa den, eta horrela $p = p_B^2$ izaten jarraitzen du.

Noski, errealitatean lagin baten bitartez neurtzen den $\hat{p}$ balioa p erabilera-mailaren estimazioa da, eta tartean egon daitekeen L ehunekoaren balioa ez da ezaguna izaten.



Irudia 14: $p$ren bilakaera, $p_B$ren arabera, L leialtasun-maila 70 dela.

3.3.3 Gizarteratze-maila

Soziolinguistikaren arloan, *gizarteratze-maila* edo *gizarteratze-tasa* landu izan da lehen ere. Adibidez: '*Txillardegiren hitzetan gizarteratze-tasa euskaldunen hizkuntza harremanen sareetara bildutako erdaldunen proportzioa da. Neurri batean gure Txepetxen trinkoketa datorkigu gogora, anisotropiaren adierazlea izan daiteke. Euskaldunok zein neurritan bizi diren beraien artean eta zenbat erdaldun erakartzen duten haien hizkuntza harremanen sareetara*' (Isasi eta Iriarte, BAT aldizkariaren 28. alean, 58. orria, 1998). 3.1 eta 3.2 ataletako analisietan lau muturreko egoera landu dira: azpitaldeen arteko integrazio-mailak (erabat integratuta ala erabat desintegratuta) eta norbanakoen komunikazio-joerak (erabatekoa ala partziala eta orekatua) konbinatuz sortzen diren egoerak, hain zuzen ere. Lau kasu horietan, probabilitate-azterketak argia erakutsi digu, eta euskararen erabilera-mailari nolabaiteko bi borne (goi- eta behe-borneak) jarri dizkiogu. Esan dezakegu errealitatea bi mutur horien artekoa dela.

Muturreko bi kasu mota izan dira aztertuak eta biak dira adierazgarriak hizkuntza gizarteratzeko ereduei gagozkiela. Alde batetik, bi azpitaldeen arteko komunikazioa orekatua denean, bi hizkuntzen gehieneko gizarteratzea adierazten da. Kasu honetan, euskararen erabilera-maila p_B^2 da (kontuan hartu gabe leialtasun-maila). Beste aldetik, bi azpitaldeak erabat bereizita daudenean, bi hizkuntzen gutxieneko gizarteratzea adierazten da. Honako honetan euskararen erabilera-maila p_B edo $\frac{p_B^2}{1-2\cdot p_B\cdot(1-p_B)}$ da.

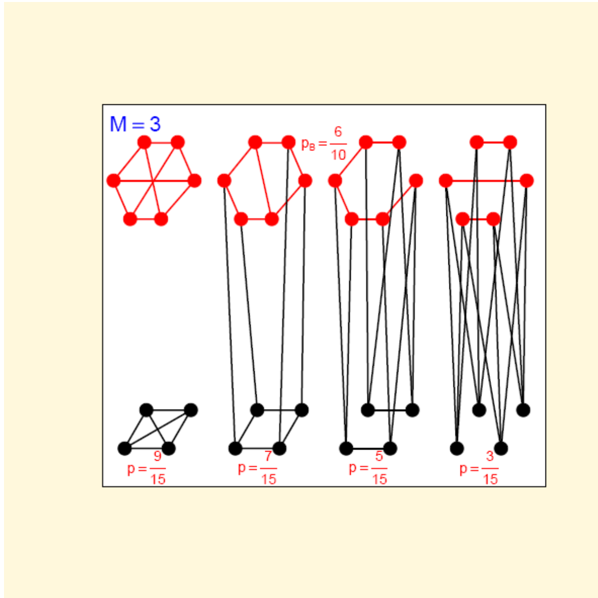
$$p_B^2 \leq p_B \text{ eta } p_B^2 \leq \frac{p_B^2}{1-2\cdot p_B\cdot(1-p_B)}$$

Ondorioa argi eta garbi azaltzen zaigu: baldintza horietan, *bi azpitaldeen komunikazioa erabat orekatuta dagoenean euskararen erabilera txikiagoa da, bi azpitaldeak beren barrurantz 'itxita' daudenean baino.*

15 irudiko adibidean gizarteratze-maila ezberdinak irudikatzen dira. Komunikatze-maila bat ezarriz gero ($M = 3$; hau da, $K = \frac{M}{A+B-1} = \frac{3}{9} = \%33.33$) irudikatuak diren kasuek erakusten dute, ezkerretik eskuinera, bi azpitalde desintegratu (1.a), endogamiarako joera ez oso markatua (2.a), gizarteratze orekatua (3.a) eta joera exogamikoa (4.a). Erabilera-mailak beheranzko joera

du 1.tik 4.era.

Muturreko kasu hauek ereduen bitartez azter daitezke:



Irudia 15: Gizarteratze-maila desberdinak, $M = 3$ baliorako.

$N = (A + B) \cdot M \cdot \frac{1}{2} \quad p_B = \frac{B}{A+B} = \frac{B \cdot M}{2 \cdot N} \quad p = \frac{N_B}{N}$

- $A \leq B$ denean ($p_B \geq 0.50$):
  - $M \leq B$ bada: $N_B = N - A \cdot M \quad p = \frac{B-A}{B+A} = 2 \cdot (p_B - 0.5)$.
  - $M > B$ bada: $N = A \cdot B + A \cdot \frac{M-B}{2} + B \cdot \frac{M-A}{2} \equiv (A + B) \cdot M \cdot \frac{1}{2} \quad p = \frac{B \cdot (M-A)}{(A+B) \cdot M} = p_B \cdot (1 - \frac{A}{M}) \quad \frac{1}{2} \cdot p_B \leq p \leq p_B$
- $A > B$ denean ($p_B < 0.50$):
  - $M \leq A$ bada: $p = 0$.
  - $M > A$ bada: $p = \frac{A \cdot (M-B)}{(A+B) \cdot M} = (1 - p_B) \cdot (1 - \frac{B}{M}) \quad \frac{1}{2} \cdot (1 - p_B) \leq p \leq (1 - p_B)$

Praktikan gizarteratze-maila zein den aldez aurretik erabakitzea ez da zeregin erraza, gizataldearen araberakoa izango da-eta (hiria, herria, lantegia, eta abar).

3.4 Laginketaren problematika

Hasieran aipatu dira zein diren elkar lotuta dauden laginketaren bi problema nagusiak: euskararen erabilera-mailaren ($p$ren) estimaziorako zenbatekoa izan behar den gutxieneko laginaren tamaina (n), eta laginaren n tamaina eta konfiantza-maila bat finkatuz gero zein den lortutako euskararen erabilera-mailaren ($\hat{p}$ren) errore-tasa.

Bada bi horien aurretik argitu beharreko beste problema bat: laginketa itzuleraduna da ala itzulera gabekoa. Esan nahi da, ea zoriz hartzen den laginean erreparatzen diren elkarrizketak errepika daitezkeen ala ez.

Lehendabizi azken problema honi egingo diogu aurre, eta, gero, beste aurreneko biak ebazteko bidea zabalduko dugu.

Ez du ematen elkarrizketa berari behin baino gehiagotan erreparatzea egokia denik euskararen erabilera-maila neurtzeko. Elkarrizketen unibertsoa handia denean arazo bat izan daiteke errepikatze ahal horri ekiditea, baina azalduko dugu kasu horretan errepikatzeak ez duela garrantzi praktiko handirik, besteak beste errepikatzea zaila gertatzen delako.

Elkarrizketen itzulera gabeko laginketa zoriz egiten bada N elkarrizketen unibertsoan (N_B dira euskaraz, proportzioz $p = \frac{N_B}{N}$, beraz), n elkarrizketari erreparatuz (n_B dira euskaraz, proportzioz $\hat{p} = \frac{n_B}{n}$), berez, euskarazko elkarrizketen kopurua probabilitate-lege *hipergeometrikoa*ren arabera lortzen da. Zenbaki konbinatorioak erabiliz hauxe da bere formula:

$$Pr(n_B = k) = \frac{C(N_B,k) \cdot C(N-N_B,n-k)}{C(N,n)}$$

Adibidea: N, N_B eta n ezagututa ($N = 16$, $N_B = 5$ $p = \frac{N_B}{N} = 0.3125$ (%31.25) $n = 8$) kalkula daitezke lege hipergeometrikoaren probabilitateak:

n_B	0	1	2	3	4	5
$\hat{p}$	0.0000	**0.1250**	**0.2500**	**0.3750**	**0.5000**	0.6250
$\hat{p}$	%0.00	**%12.50**	**%25.00**	**%37.50**	**%50.00**	%62.50
Probb.	0.0128	**0.0128**	**0.3590**	**0.3590**	**0.0128**	0.0128

Honen bitartez esan dezakegu $n = 8$ tamainako lagin bat hartuz gero, laginean aurkituko den euskara-erabilera 0.125 (%12.50) eta 0.500 (%50.00) bitartekoa (0.3125 ± 0.1875 edo $\%31.25 \pm \%18.75$) izango dela %97.44 ($0.9744 = 0.0128 + 0.3590 + 0.3590 + 0.0128$) konfiantzaz.

N handia bada eta N_B ez oso txikia, lege hipergeometrikoaren probabilitateak lege *binomiala*renen bitartez kalkula daitezke; hau da, itzulera gabeko laginketa itzulera duen laginketatzat har daiteke (unitateak *independente*ki hartuko balira bezala) kalkuluetan oso akats txikia eginaz. Lege binomialaren kalkuluak formula honen bitartez egiten dira:

$$Pr(n_B = k) = C(n,k) \cdot p^k \cdot (1-p)^{n-k}$$

Adibidea: Probabilitate-banaketa hipergeometrikoa eta binomiala kalkulatuz gero, $N = 1200$ eta $N_B = 420$ ($p = \frac{N_B}{N} = \frac{420}{1200} = 0.35$), eta $n = 100$ izanda, argi geratzen da beren hurbiltasuna:

n_B	26	27	28	29	30	31	32	33	34	35
$\hat{p}$	0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34	**0.35**
Prob.hiperg.	0.012	0.019	0.027	0.038	0.049	0.061	0.071	0.080	0.086	0.087
Prob.binom.	0.014	0.021	0.029	0.039	0.049	0.060	0.070	0.077	0.082	0.083

n_B	36	37	38	39	40	41	42	43	44	
$\hat{p}$	0.36	0.37	0.38	0.39	0.40	0.41	0.42	0.43	0.44	Σ
Prob.hiperg.	0.085	0.078	0.069	0.059	0.047	0.036	0.027	0.019	0.013	0.963
Prob.binom.	0.081	0.076	0.067	0.058	0.047	0.037	0.028	0.021	0.014	0.953

Esan daiteke laginaren euskararen-erabilera 0.26 eta 0.44 bitartean (0.35 ± 0.09) suertatuko dela, gutxi gorabehera %95eko konfiantzaz.

Beraz, N handia denean laginketa itzulera gabekoa izan ala ez izan, kalkuluen ikuspegitik, ez da garrantzitsua. Beste modu baten esanda, proportzioa kalkulatzeko berdin da elkarrizketa jakin bati behin baino gehiagotan zoriz erreparatu ahal izatea.

N handia izanda, n txikia ez denean (erreferentziatzat hartzen da $n > 25$) eta p ez oso handia, ezta oso txikia ere (erreferentzitzat hartzen da $n \cdot p > 5$ eta $n \cdot (1-p) > 5$) probabilitate-lege binomiala probabilitate-lege normal edo gausstarraren bitartez kalkula daiteke.

Aurreko adibide bera: Lege binomialaren probabilitateak (hipergeometrikoaren oso antzekoak gertatu direnak) $p = 0.35$ eta $n = 100$ izanda, lege normalaren bitartez hurbiltzean bere bi parametrotzat bi zenbaki hauek $\mu = n \cdot p = 35$ eta $\sigma^2 = n \cdot p \cdot (1-p) = 12.25$ hartzen badira:

n_B $\hat{p}$	26 0.26	27 0.27	28 0.28	29 0.29	30 0.30	31 0.31	32 0.32	33 0.33	34 0.34	35 0.35
Prob.binom.	0.014	0.021	0.029	0.039	0.049	0.060	0.070	0.077	0.082	0.083
Prob.norm.	0.014	0.021	0.029	0.038	0.048	0.059	0.069	0.076	0.082	0.083

n_B $\hat{p}$	36 0.36	37 0.37	38 0.38	39 0.39	40 0.40	41 0.41	42 0.42	43 0.43	44 0.44	Σ
Prob.binom.	0.081	0.076	0.067	0.058	0.047	0.037	0.028	0.021	0.014	0.953
Prob.norm.	0.082	0.076	0.069	0.059	0.048	0.038	0.029	0.021	0.014	0.955

Hurbilketa hau egiteko $n_B$ren balio diskretu bakoitza unitate bateko zabalera duen tartetzat hartzen da. Esate baterako, lege binomialak $n_B = 30$ balioari 0.049 probabilitatea esleitzen dio, eta lege normalak 0.048 probabilitatea $[29.5, 30.5]$ balio-tarteari. Lege normala zenbaki erreal guztien multzoaren gainean definitzen da (lege binomiala eta hipergeometrikoa, aldiz, diskretuak dira eta zenbaki arrunten gainean definitzen dira). Horrek badu bere abantaila: lege normalaren bitartez erraz kalkula daitezke $\hat{p}$rentzako balio-tarteak, aldez aurretik konfiantza-maila ezarriz gero. Adibidez, kalkulatua izan da 0.26 eta 0.44 balio-tarteari ($[0.26, 0.44]$, 0.35 ± 0.09) gutxi gorabehera %95-eko konfiantza-maila dagokiola, hau da, $0.35 \pm 1.96 \cdot \sqrt{\frac{0.35 \cdot 0.65}{100}}$, 1.96 balioa %95-eko konfiantza-mailari dagokion balio *normala*, baina lege bera aplikatuz esan daiteke %99-eko konfiantza-mailari dagokion balio *normala* 2.575 balioa dela, eta, beraz, laginean gertatuko den euskara-erabilera $0.35 \pm 2.575 \cdot \sqrt{\frac{0.35 \cdot 0.65}{100}}$ izango dela, hau da, (0.35 ± 0.12), beste inolako hausnarketarik eta kalkulurik egin behar gabe.

N handia ez denean, ordea, hurbilketa binomiala ez da bidezkoa. Hala ere, N txikia ez izanda ere, probabilitate-lege hipergeometrikoa lege normalaren bitartez hurbil daiteke. Lege hipergeometrikoaren bariantza lege binomialarena denez faktore biderkagai bat ezik, lege normalaren hurbilketan faktore hau ere hartzen da kontuan. Bariantzan parte hartzen duen delako faktore hori $\frac{N-n}{N-1}$ da.

Adibidea: $N = 400$ eta $N_B = 140$ ($p = \frac{N_B}{N} = \frac{140}{400} = 0.35$), eta $n = 100$

n_B $\hat{p}$	26 0.26	27 0.27	28 0.28	29 0.29	30 0.30	31 0.31	32 0.32	33 0.33	34 0.34	35 0.35
Prob.hiper.	0.009	0.015	0.023	0.034	0.047	0.061	0.075	0.086	0.094	0.096
Prob.norm.	0.009	0.015	0.023	0.034	0.047	0.060	0.074	0.086	0.093	0.096

n_B $\hat{p}$	36 0.36	37 0.37	38 0.38	39 0.39	40 0.40	41 0.41	42 0.42	43 0.43	44 0.44	Σ
Prob.hiper.	0.093	0.085	0.073	0.060	0.046	0.034	0.023	0.015	0.009	0.978
Prob.norm.	0.093	0.086	0.074	0.060	0.047	0.034	0.023	0.015	0.009	0.978

Aurreko bidean, ezagunak izanda N, N_B ($p = \frac{N_B}{N}$, euskararen erabilera, beraz) eta n, jakin daiteke probabilitate-hizkera erabiliz (konfiantza-maila) zein izango den $\hat{p}$ balioa gutxi gorabehera, zehaztasun-maila batez. Laginketaren bi problema nagusiak alderantzizko norabidean daude: N_B eta n ezezagunak izanda (p, beraz), konfiantza- eta zehaztasun-mailak finkatuz gero, jakin daiteke zein diren gutxieneko n laginaren tamaina, eta erreparatutako $n_B$ren ezagutzaren bitartez ($\hat{p}$, beraz) $p$ren gutxi gorabeherako balioa (estimazioa, $\hat{p}$ eta errore-tasa edo zehaztasun-maila).

3.5 Laginaren tamaina

Euskarazko p erabilera-maila ezezaguna denean, n tamainako zorizko lagin batez estimatzen da. Estimazioa zehatzagoa izan dadin n elkarrizketari erreparatzen zaie, ea erdaraz ala euskaraz egin diren, eta hor gertatzen den proportzioak balio du $p$ren delako estimazio hori egiteko ($\hat{p}$). Estimazioaren kalitatea, $n$rekin lotuta egoteaz gain (zenbat eta handiagoa den n, orduan eta hobea da laginaren kalitatea), bi alorretan islatzen da: estimaziorako nahi den α' *konfiantza-maila*, eta estimazioak izango duen δ *zehaztasun-maila*. Gakoa da, zehaztasun-maila jakin baterako eta konfiantza-maila baterako, zenbateko lagina jaso behar den erabakitzea: gutxiegi hartuz gero, zehaztasun- edota konfiantza-maila txikiak lortuko dira, eta gehiegi hartuz gero, garestiegi gertatuko da.

Aurreko atalean ikusi denez, N handia denean eta $n > 25$, probabilitate-lege normala erabil daiteke kalkuluetan oso akats txikiak eginez, eta orduan laginketa itzulera gabekoa ala itzuleraduna den jakiteak ez du garrantzi handirik kalkuluetan. N handia ez denean, aldiz, $n > 25$ bada, probabilitate-lege normala erabil daiteke ere, baina itzulera gabeko laginketari zor zaion faktore bat hartu behar da kontuan.

Beste aldetik, oso txikia ez den gizatalde batean elkarrizketen unibertsoa handia izango da. Esate baterako, 25 norbanako duen gizatalde batean guztira binakako $C(25, 2) = 300$ elkarrizketa izan daitezke, komunikatzeko ahalmena $K = 100$ baldin bada, eta 180 elkarrizketa, $K = 60$ bada; baina, hirunakako, launakako eta abarnakako elkarrizketak ere kontutan hartu behar dira, eta hauek guztiak binakakoak beste dira gutxi gorabehera, enpirikoki behatu denez (aurrerago aztertuko dugu problema hau). Beraz, aztergai den gizatalde txikienetako batean 400 bat elkarrizketek osatzen dute unibertsoa. N ez da txikia.

Baldintza horiek guztiak betez gero, eta zehaztasun- eta konfiantza-maila horiek finkatuz gero, $n$ren kalkulua honela egiten da:

$$n \geq \frac{(z_{1-\frac{\alpha}{2}})^2 \cdot p \cdot (1-p) \cdot N}{(z_{1-\frac{\alpha}{2}})^2 \cdot p \cdot (1-p) + \delta^2 \cdot (N-1)},$$

edo, besterik gabe (Cochran, *Técnicas de muestreo*, CECSA, 4^a impresión, 1974, 110. orr.):

$$n_0 \geq (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2 \cdot p \cdot (1-p), \text{ eta } n = \frac{n_0 \cdot N}{n_0 + (N-1)} \cong n_0 \cdot \frac{1}{1+\frac{n_0}{N}}$$

- δ, zehaztasun-maila da; laginaren emaitzak jaso ondoren estimazioan onar daitekeen errore-tasa handiena. Hau aldez aurretik finkatuz gero, bermatuta dago estimazioak, α' konfiantza-maila jakin baterako, δ errore-tasa edo gutxiago izango duela.

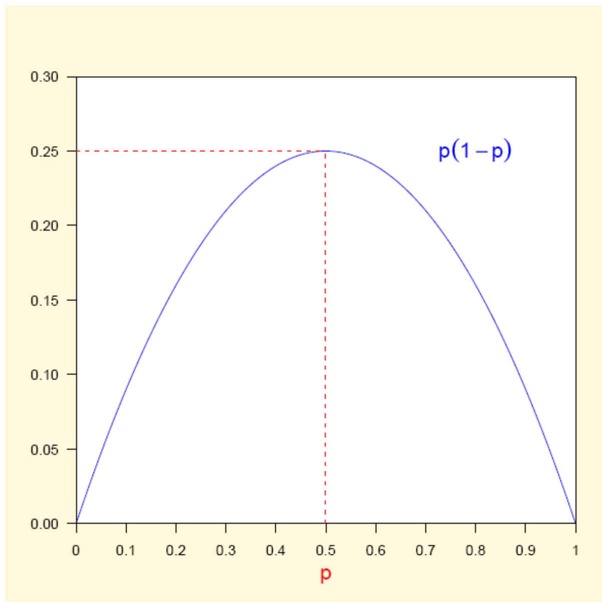
- Ohar daitekeenez, ($z_{1-\frac{\alpha}{2}}$) banaketa normalean oinarritutako hurbilketatik datorren α' konfiantza-mailari dagokion balio normala da ($\alpha = 1-\alpha'$); azpiindizea korapilatsua bada ere, tradizioz, horrela adierazten da. $z_{1-\frac{\alpha}{2}}$ balioa erabiltzea bidezkoa izan dadin nk handia behar du izan, beraz ($n > 25$ edo).

- N oso handia denean $n$ren aldean, laginaren n tamaina n_0 bihurtzen da praktikoki, n_0 $n$tik gertu baitago ($\frac{n_0}{N}$ zatikia unibertsoaren finitutasunari zor zaion faktorea da). Kasu horretan Nk ez du formulan parte hartzen eta elkarrizketen multzoa infinitua balitz bezala erabil daiteke. Hipergeometrikoaren ordez, binomiala; laginketa itzulera gabekoa izan beharrean, itzuleraduna. Handia ez denean, ordea, Nk badu bere garrantzia.

Formula horretan, ordea, $p$ren balioa ez da ezaguna. Sorgin-gurpil bat gertatzen da. Alde batetik, lagina hartu nahi da $p$ren balioa ezagutzeko, eta bestetik laginaren n tamainaren handitasuna $p$ren menpe dago. Gurpil horretatik irteteko egiten dena zera da: litezkeen $p$ren balio guztietatik $n$ren balio handiena ematen duena aukeratu. Bestela esanda, erabakirik garestiena hartzen da (zenbat eta handiagoa izan n, orduan eta garestiago da lagina lortzea), baina edozein kasutan $p$ren estimazioaren bi kalitate-adierazleak (zehaztasuna eta konfiantza) errespetatu egiten dira.

n handiena eskatzen duen erabilera-maila, orokorrean, $p = \%50$ denekoa da ($max_{p\in(0,1)}\ p \cdot (1 - p) = 0.5 \cdot (1 - 0.5) = 0.25$) (ikus 16. irudia).

Bestelako informaziorik ez badago, $p$ren balioa %50 izan daitekeenez, azaldutako formulan balio



Irudia 16: $p \cdot (1 - p)$ formularen bilakaera, $p$ren arabera.

hori erabiltzen da kalkuluak egiteko (Cochran, *Técnicas de muestreo*, CECSA, 4^a impresión, 1974, 111. orr.):

$$n_0 \geq (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2 \cdot (\frac{50}{100} \cdot (1 - \frac{50}{100})) \equiv \frac{1}{4} \cdot (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2$$

Adibidez, nahi den konfiantza-maila $\alpha' = \%95(0.95)$ ($\alpha = 0.05, 1 - \frac{\alpha}{2} = 0.975$) bada, $z_{0.975} = 1.96$, eta zehaztasun-maila (errore-tasa) $\delta = \%2(0.02)$koa bada, orduan $n_0 \geq \frac{1}{4} \cdot (\frac{1.96}{0.02})^2 = 2401$, eta $\alpha' = \%99(0.99)$ bada eta $\delta = \%3(0.03)$, orduan $n_0 \geq \frac{1}{4} \cdot (\frac{2.575}{0.03})^2 = 1841.840$.

Laginaren n tamaina elkarrizketen unibertsoaren N tamainaren arabera erabakiko da, hau da, $\frac{n_0}{N}$ faktorearen arabera. Adibidez, $N = 400$ baldin bada $n \cong n_0 \cdot \frac{1}{1+\frac{n_0}{N}} = 1841.84 \cdot \frac{1}{1+\frac{1841.84}{400}} = 328.63$. Zenbait ezaditurentzat harrigarria bada ere, zehaztasun- eta eta konfiantza-maila horiek lortzeko elkarrizketen unibertsoaren %82a baino gehiago ($\frac{328.63}{400} = 0.8216$) hartu behar da laginean.

Formula horren ondorioz zenbait taula eraikitzen dira kalkuluak aldiro ez egitearren. Gehienetan balio hauentzat izaten dira taulak: $\alpha' = \%95, \%99$, $\delta = \%2, \%3, \%4, \%5$.

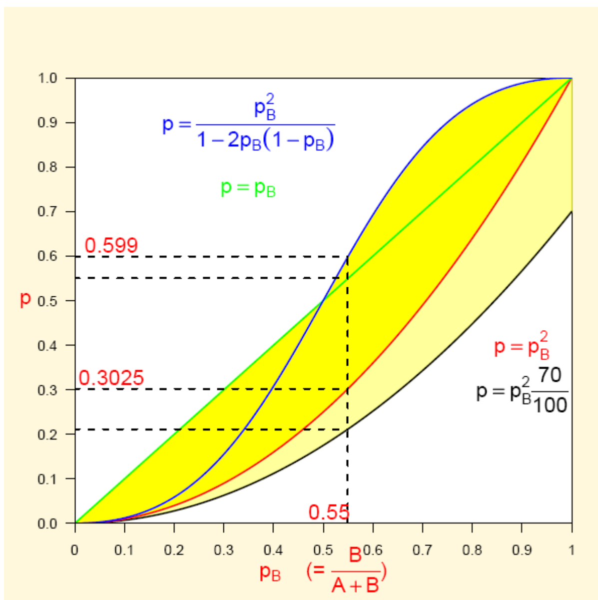
Taula horiek $p = 0.50$ edo $p = \%50$ kasurako (kasurik okerrenerako) kalkulatu dira, n handiena eskatzen duelakoan. Baina A eta B azpitaldeetako kide-kopuruak ezagutzen badira, kopuru horiek erabil daitezke $p$ren goi- eta behe-mugak ezartzeko, behin zehaztasun- eta konfiantzamailak finkatuta. Muga horiek baliatuta, laginaren tamaina beste modu batera kalkulatzen ahal da (Cochran, *Técnicas de muestreo*, CECSA, 4^a impresión, 1974, 88. orr.). Ondorioz laginketaren kostua merkatu egin daiteke.

Adibidez, $\delta = \%2(0.02)$ eta $\alpha' = \%95(0.95)$ izanda, $p \leq \%30$ izango dela ziurtasun osoa baldin badago, laginaren tamaina kalkulatzeko formulan, kasurik okerrena ez da 0.50, eremutik kanpo baitago, 0.30 baizik: $n_0 \geq (\frac{1.96}{0.02})^2 \cdot (\frac{30}{100} \cdot (1 - \frac{30}{100})) = 2016.84$, $p = 0.50$ kasuko $n_0 \geq 2401$ baino txikiagoa, alegia.

Hau da, 0.50 balioa ez baldin badago goi- eta behe-mugen artean, kasurik okerrena bi muga horietatik 0.50etik gertuen dagoena da (ikus 16 irudia).

Aztertutako eredu guztiak deterministak dira, nahiz hasierako datuak eta bukaerako erabilera-mailak proportzioen bitartez adieraziak izan. Behin baldintzak finkatuz gero, euskararen erabilera-mailaren balioa deduzitu egiten da. Errealitatea, baina, ez zaio halako eredu determinista bati zehatz-mehatz egokitzen. Eredu horiek erreferentzia-puntutzat har daitezke eta errealitatekoa horien artean kokatu.

Eredu horien arabera ele$B$idunen p_B proportzioa ezagutuz gero, p zein balioren artekoa izan



Irudia 17: $p$ren balio posibleen eremua, p_B proportzioen arabera.

daitekeen aurresatea badago: 17 irudian horiz margotu da eremu hori eta p_B abzisa duen edozein punturen ordenatua dago bere baitan. Adibideko $p_B = 0.55$ balioarentzat, $p$ren balioa 0.3025 eta 0.5990en artean dagoen edozein balio izan daiteke, betiere leialtasun-maila kontuan hartu gabe. Leialtasun-maila kontuan hartuz gero, behe-bornea txikiagoa izango da maila horren arabera (hori apalez margotua).

Laginaren tamaina kalkulatzeko kasurik okerrena, beraz, %0.50 da, aurreko bi mugen artean baitago. Baina, $p_B = 0.30$ balitz, bi mugak %9 eta %30 izango lirateke, eta, hortaz, laginaren tamaina kalkulatzeko kasurik okerrena 0.30 izango litzateke (0.09 eta 0.30 bi balio horietatik 0.50etik gertuen dagoena):

$$n_0 \geq (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2 \cdot (0.30 \cdot (1 - 0.30))$$

Beste adibide hau kontrakoa da, nolabait esateko: $p_B = 0.90$. Bi mugak muturreko bi kasuetan, %81.00 eta %98.78 dira, eta 0.50etik bietatik gertukoena 0.81 da. Laginaren tamaina kalkulatzeko formula, hortaz:

$$n_0 \geq (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2 \cdot (0.81 \cdot (1 - 0.81))$$

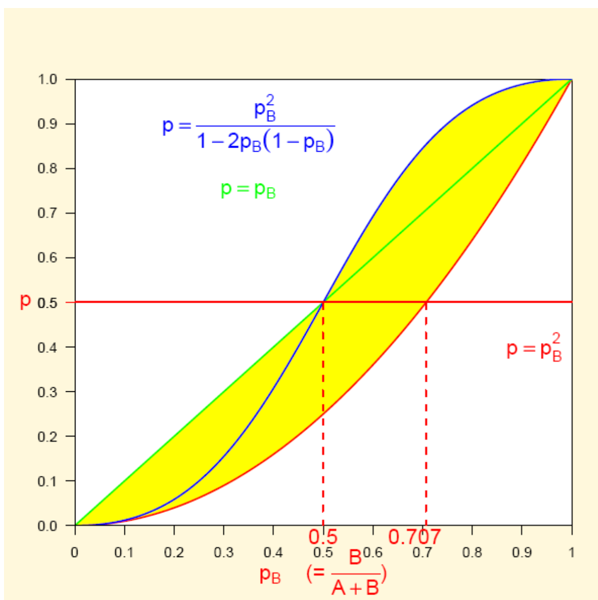
Gogoratu behar da bi balio hauek bi muturreko eredu jakinei dagozkiela: bata, gizatalde integratuko norbanako guztien artean komunikazioa orekatua denean (behe-muga, gorriz marraztutako kurba), eta bestea, gizataldea erdibituta dagoenean erabat, azpitalde bakoitzean erabateko komunikazioa dagoelarik (goi-muga, berdez). Behe-muga txikiagoa izango da leialtasun-maila ereduan sartuz gero, baita goi-muga ere leialtasun-maila ez bada $L = 100$ekoa nahiz hizkuntzataldeak 'itxiak' izan.

Eredu gehiago eraikiz mugak alda daitezke.

Laburbilduz: 0.50 proportzioa ez badago bi mugen artean, bi azpitaldeen arteko proportzioaren ezagutzak n txikiagotzen lagunduko du.

Adibide hauek direla-eta sortzen den galdera da Bren zein baliok laguntzen duten emandako formula orokorretik lortutako $n$ren tamaina txikitzen, eta zeinek ez. 18 irudiak laguntzen du erantzuna irudikatzen.

0.707 zenbakia $0.5 = p_B^2$ ekuazioaren soluzioa da ($\sqrt{0.5} \equiv 0.707$).

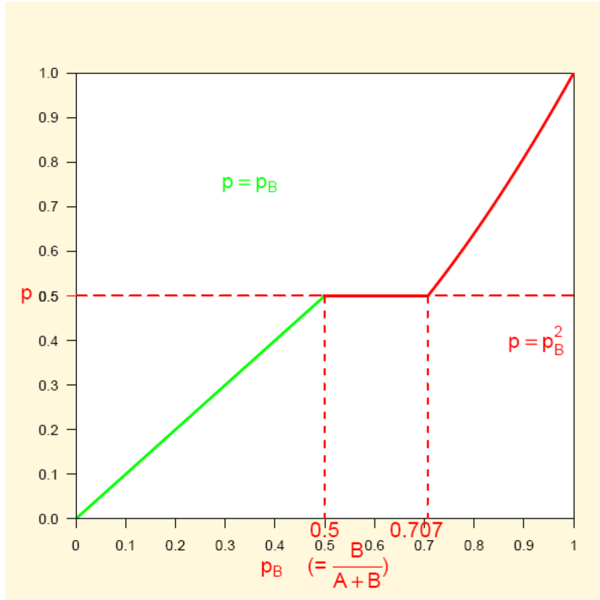


Irudia 18: $p$ren balioak kontuan hartuz, laginaren tamaina doitzeko hurbilpena.

19 irudiak erakusten du, azaldutako eredu matematikoak kontuan hartuta, $p$ren zein balio jarri behar den $n_0 = (\frac{z_{1-\frac{\alpha}{2}}}{\delta})^2 \cdot (p \cdot (1 - p))$ formulan laginaren tamaina kalkulatzeko:

1. $p_B < 0.50$ denean, $p = p_B$.
2. $0.5 \leq p_B < 0.707$ denean $p = 0.50$.
3. $p_B \geq 0.707$ denean $p = p_B^2$.

Laginaren tamaina kalkulatzeko arrazoibideetan ez dugu kontuan hartu, orain arte, *leialtasun-maila*. Faktore hau ere kontuan hartu behar dugu, ordea. Adibidez, $L = 70$ dela jakingo bagenu,



Irudia 19: $p$ren balioak, laginaren tamaina kalkulatzeko.

benetako $p_B$ren ordez $p_B \cdot \sqrt{\frac{70}{100}}$ kalkuluaren balioa jarri beharko genuke; esate baterako, benetako $p_B = 0.80$ izango balitz eta $L = 70$, kalkuluak $p_B = 0.67$ balioarekin egin beharko genituzke, eta $p_B = 0.40$ eta $L = 50$ balira, erabili beharrekoa: $p_B = 0.40 \cdot \sqrt{\frac{50}{100}} = 0.2828$.

n_0 kalkulatu ondoren, laginaren tamaina beste formula honen araberakoa da: $n = \frac{n_0 \cdot N}{n_0 + (N-1)}$. Formula horretako N balioa, dakigunez, elkarrizketen unibertsoko tamaina da. Nren balioa ezartzeko, binakako elkarrizketak baizik ez ditugu kontuan hartu orain arte. Baina, analisia osatzeko, beste guztiak zenbat diren ere jakin beharko genuke. 4 atalean landuko dugu gai hori.

Bestalde, egindako analisiak kalkulu-taulatan jaso ditugu. Taula horiek 5 atalean azaltzen dira.

3.6 Euskararen erabilera-mailaren estimazioa eta zehaztasun-maila

Erabilera-mailaren $\hat{p}$ estimazioa laginean neurtu den euskarazko elkarrizketen proportzioa ($\hat{p} = \frac{n_B}{n}$) da, eta laginaren n tamaina txikia ez den kasuetan, honela kalkulatzen da bere δ zehaztasun-maila edo 'errore-tasa' (Cochran, *Técnicas de muestreo*, CECSA, 4^a impresión, 1974, 88. orr.):

$$\hat{p} \pm \delta$$

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n-1}} \cdot \sqrt{\frac{N-n}{N-1}} + \frac{1}{2 \cdot n}$$

Adibidez, $N = 200, n = 50, \alpha = \%95, \hat{p} = \%23$ bada:
$0.23 \pm 1.96 \cdot \sqrt{\frac{0.23 \cdot (1-0.23)}{49}} \cdot \sqrt{\frac{200-50}{200-1}} + \frac{1}{2 \cdot 50} = 0.23 \pm 0.11$

N oso handia denean $n$ren aldean, $\sqrt{\frac{N-n}{N-1}} \cong \sqrt{1 - \frac{n}{N}}$, populazioa finitua izateari zor zaion faktorea 1etik gertu dago. Nk beraz ez du formulan parte hartzen eta elkarrizketen unibertsoa <u>infinitua</u> bailitzan erabil daiteke.

Formula hauek ere banaketa normalean oinarritutako hurbilketatik datoz: nk ez du txikia izan behar ($n > 25$ erreferentzia-puntutzat hartzen da). n txikia denean banaketa hipergeometrikoan oinarritutako formulak erabili behar dira (Cochran, *Técnicas de muestreo*, CECSA, 4^a impresión, 1974, 88. orr.) (Chung, DeLury, 1950 %90, 95, 99, $N = 500, 250, 10000$, Lieberman, Owen, 1961, $N \leq 100$), edo bestela, hurbilketa gisa, banaketa binomialarenak (Fisher, Yates, 1957).

4 Unibertsoko elkarrizketa-kopurua

Elkarrizketa guztiak ez dira binakakoak. Batzuk hirunaka, beste batzuk launaka eta gehiagoka ere gertatzen dira. N bada elkarrizketen unibertsoaren tamaina eta N_2, N_3, N_4 binakako, hirunakako eta abarnakako elkarrizketak, hurrenez hurren, hauxe da bilatzen dena:

$$N = N_2 + N_3 + N_4 + \ldots$$

Orain arte N_2ri buruz aritu gara, baina aurreko formulak behar bezala erabili ahal izateko, N_3, N_4 eta abarrak ere kontuan hartu behar dira.

Errealitatera hurbiltzeko kasuistika hau uztartu egin behar da binakako elkarrizketei buruz esandakoarekin. N_2, N_3, N_4 eta abarren arteko erlazioak analitikoki ezar daitezke eredu matematiko bat eraikiz, edo enpirikoki.

Bide analitikoa bukatu gabeko ikertze-bidean dago oraindik. Dena dela, bide horren hasierako urratsak aurreraxeago azalduko ditugu, beharrezkoak izango baitira Nren balioa bornatzeko.

Bide enpirikoak gizatalde desberdinen sailkaketan eta behaketen erregulartasunean oinarritutako zenbakiak eskain ditzake. Bide enpirikoa zabala da eta gizatalde mota bakoitzean ezar litezke proportzio horiek, ez baitugu uste gizatalde guztietan azaldutako binakako, hirunakako eta abarnakako proportzioak berdinak izango direnik. Esan nahi dugu gizatalde txiki batean (adibidez, 30 bat pertsonak osatua) eta gizatalde handi batean (adibidez, 30000 bat pertsonakoa) gertatzen diren proportzioak ezin direla, besterik gabe, berdintzat jo. Bide honek bestelako azterketa bat eskatzen du, gizataldeak sailkatzean oinarrituko litzatekeena.

Txillardegiren eredu matematikoan azaltzen da gizatalde handietan binakako elkarrizketak %54.88 izaten direla, hirunakakoak %28.62, eta launakakoak %16.50[2]. Hortik aurrera, egoera arruntean, taldea banatu eta elkarrizketa bat baino gehiago sortzen omen da. Proportzio horiek $p_2 = \frac{N_2}{N}$, $p_3 = \frac{N_3}{N}$, eta $p_4 = \frac{N_4}{N}$ zatikien estimazioak dira. Txillardegiren proportzioak bide enpiriko baten emaitza dira, baina ez dugu uste proportzio horiek edozein gizatalderenak direnik.

Bide analitikoa bukatu gabe dugu oraindik, eta bide enpirikoak ugari dira. Nola erabaki N zenbatekoa den? Ataka honetatik irteteko muturreko kasuetara jo dezakegu eta Nren benetako balioa bornatu edo zedarritu. Ondorioz, laginaren n tamaina kalkulatu ahal izango da kasurik okerrenean, eta tamaina horrekin euskararen erabilera-maila estimatu. Nren ezagutza sakonagoak beharrezko laginaren tamaina fintzen lagunduko du, baita estimazioaren kalitatea ere.

Orain arte N_2z aritu gara, eta $N = \frac{N_2}{p_2}$ da. N_2ren inguruko zenbait hausnarketa eginda dago aurreko ataletan, baina p_2 binakako elkarrizketen proportzioaren inguruan ez, ordea. Soziolinguistikoki, agian, onar daiteke gizatalde guztietan $p_2 \geq p_3 \geq p_4 \geq \ldots$ dela. Honek esan nahi du, launakakoetatik areagoko elkarrizketak baztertzen baldin badira, $p_2 \geq \frac{1}{3} = 0.33 = \%33.33$ dela.

Kasurik okerrena, hau da, laginaren tamaina handiena eskatzen duena, $p_2 = \frac{1}{3}$ proportzioak ematen duena da. Hortaz, elkarrizketen unibertsoko tamainaren goi-borne bat $N \leq 3 \cdot N_2$ da. Komunikatzeko joera erabatekoa denean $N_2 = C(A+B, 2)$ denez, goi-bornea $3 \cdot C(A+B, 2)$ da:

$$N^* = 1.5 \cdot (A+B) \cdot (A+B-1)$$

[2]Proportzio hauek 1993ko eta 1997ko Euskal Herriko kale-neurketetan jasotako datuetatik finkatu ziren. Neurketa haietan behatutako udalerri gehienek 1000 biztanletik gora zituzten.

Adibidez, $A+B=25$ denean $N^* \sim 900$, eta $A+B=50$ denean $N^* \sim 3675$. $A+B$ handia denean, $N^* \sim 1.5 \cdot (A+B)^2$ oso zenbaki handia da ($A+B=30000, \quad N^* \sim 1349 \cdot 10^6$).

Nren balioaren ezagutza fintzeko asmoz, saiatu gara eredu matematiko baten bitartez adierazten elkarrizketen $p_2, p_3, p_4, \cdots$ proportzioen banaketa, komunikatzeko joera erabatekoa denerako. Proportzio horiek erabakitzeko bi joera hartu behar dira kontuan. Alde batetik, antzematen da elkarrizketa gehiago izan daitezkeela hirunaka binaka baino, konbinatoriaren aldetik begiratuta, eta askoz ere gehiago launaka hirunaka baino. Beste aldetik, esan daiteke zailagoa dela hirunakako elkarrizketa jakin bat suertatzea binakako jakin bat baino, probabilitatearen aldetik ikusita.

Esate baterako, $A+B=25$ denean, binakako 300 ($C(A+B,2)=\frac{1}{2} \cdot 25 \cdot 24$) elkarrizketa desberdin dago, 2300 ($C(A+B,3)=\frac{1}{6} \cdot 25 \cdot 24 \cdot 23$) hirunakako, 12650 launakako ($C(A+B,4)=\frac{1}{24} \cdot 25 \cdot 24 \cdot 23 \cdot 22$), eta abar. Guztira, 33554406 elkarrizketa izan daitezke ($N=2^{A+B}-(A+B)-1$). Elkarrizketa guztiei gertatzeko probabilitate berdina esleituz gero, esan nahiko genuke 10 lagunen arteko elkarrizketak askoz ere sarriagotan gertatzen direla binakakoak baino, eta eskarmentuak alderantzizkoa gertatzen dela esaten digu: maizena 2koa da, gero 3koa, gero 4koa eta abar. Honek esan nahi du elkarrizketa guztien probabilitateak ez direla berdinak.

Probabilitateak esleitzeko eredu matematiko bat lantzeari ekin diogu. Komunikatzeko joera erabatekoa denean, elkarrizketa batean parte hartzeko edozein norbanakok $\frac{1}{25}$eko probabilitatea baldin badu, binakako elkarrizketa jakin bat suertatzeko probabilitateak $(\frac{1}{25})^2 \cdot (\frac{24}{25})^{23}$rekin du proportziozko lotura bat, eta hirunakako jakin bat suertatzekoak $(\frac{1}{25})^3 \cdot (\frac{24}{25})^{22}$rekin; hau da, binakako elkarrizketa jakin baten probabilitatea 24 aldiz handiagoa da hirunakako jakin batena baino, eta 576 (24^2) aldiz handiagoa launakako jakin batena baino.

Bi ikuspegi kontrajarri horiek elkartu egiten dira probabilitate-banaketa binomialean oinarritzen den eredu matematiko batean, eta horren bitartez binakako, hirunakako, launakako eta abarnakako elkarrizketak zein probabilitaterekin suerta daitezkeen kalkulatzen dira:

$$p_k = C_{A+B} \cdot C(A+B,k) \cdot (\tfrac{1}{A+B})^k \cdot (1-\tfrac{1}{A+B})^{A+B-k} \quad k=2,3,\ldots,(A+B),$$
$$C_{A+B} = 1-(1-\tfrac{1}{A+B})^{A+B-1} \cdot (2-\tfrac{1}{A+B})$$

Probabilitateek, azken batean, elkarrizketen unibertsoko proportzioak adierazten dituzte.

Adibidez, $A+B=25$ denean ($C_{A+B}=0.2642$), taula honetan agertzen diren proportzioak lortzen direla egiazta daiteke:

p_2	p_3	p_4	p_5	p_6	p_7
0.7105	0.2270	0.0520	0.0091	0.0013	0.0001

$A+B$ handia denean froga daiteke (froga *Poisson*en prozesu batean oinarritzen da) limitetzat proportzio hauek har daitezkeela:

p_2	p_3	p_4	p_5	p_6	p_7
0.6961	0.2320	0.0579	0.0117	0.0019	0.0004

Beraz, ereduaren arabera edozein gizataldetan $N=\frac{N_2}{p_2} \sim \frac{C(A+B,2)}{0.70} \sim 1.43 \cdot C(A+B,2)$ izango da, eta kontuan hartuta $C(A+B,2) \sim \frac{1}{2} \cdot (A+B) \cdot (A+B-1)$ dela, $N \sim 0.71 \cdot (A+B) \cdot (A+B-1)$, hau da, lehen kalkulatutakoaren erdia baino pixka bat gutxiago. Adibidez, $A+B=25$ denean $N \sim 426$, eta $A+B=50$ denean $N \sim 1740$. $A+B$ handia denean $N \sim 0.71 \cdot (A+B)^2$ oso zenbaki handia da ($A+B=30000, \quad N \sim 639 \cdot 10^6$).

Nren halako tamaina ematen dugunean ez dugu elkarrizketa posibleen kopurua kalkulatu binakakoak, hirunakakoak, launakakoak eta abarnakakoak soilik batuz ($C(A+B,2)+C(A+B,3)+$

$C(A+B,4)+\ldots$), beraien probabilitateak aintzat hartuta konbinatuz baizik. Esan nahi dugu, adibidez $A+B=25$ denean, elkarrizketa-kopurua $300+2300+12650+\ldots$ batura dela erabaki ordez, 426 dela ondorioztatzen dugula. Elkarrizketa posible guztien kopurua batu baizik egingo ez bagenu, elkarrizketa guztiak probabilitate berdinekoak balira bezala tratatuko genituzke; aldiz, 426 emaitzara iristeko, binakakoak %71 direla (gutxi gorabehera), eta gainontzekoak %29 (hirunakakoak %23, launakakoak %5 eta abarnakakoak %1) direla ari gara kontuan hartzen.

Orain artekoa laburbilduz, esan daiteke komunikatzeko joera erabatekoa denean Nrentzako behe- eta goi-borneak hauek direla:

$$0.71 \cdot (A+B) \cdot (A+B-1) \leq N \leq 1.5 \cdot (A+B) \cdot (A+B-1)$$

Beste aldetik, ohartzekoa da proportzio hauek oraindik urruti samar daudela Txillardegik eta Isasik enpirikoki aurkitutakoetatik. Enpirikoki aurkitutako proportzioa ($p_2 = 0.5488$) erabiliz gero, $N \sim C(A+B,2) \cdot \frac{1}{0.5488} \sim 0.91 \cdot (A+B) \cdot (A+B-1)$, honako hauek izango lirateke Nren tamainak jarritako adibideetan: $A+B=25, \quad N \sim 547$; $A+B=50, \quad N \sim 2232$, eta $A+B=30000, \quad N \sim 819 \cdot 10^6$.

Muturreko $p_2 = \frac{1}{3}$ kasuak eta eredu matematikoak ematen ez dutenez enpirikoki jasotako datuen berri, argi dago, ereduan beste aldagairen bat sartu behar dela. Argi dago, halaber, ereduan hartu dugun komunikatzeko erabateko joerarena ez dela hipotesi erabat erreala.

Soziolinguistikoki, agian, onar daiteke gizatalde handietan p_2 handiagoa dela gizatalde txikietan baino, errazagoa baita gizatalde txikietan hirunakako, launakako eta abarnakako elkarrizketak burutzea. Honekin esan nahi dugu $p_2 = 0.5488$ proportzioa gizatalde handi batean lortu baldin bada, gizatalde txikiago batean p_2 txikiagoa izango dela, eta, ondorioz, N handiagoak izango direla proportzio horrekin kalkulatu direnak baino. Elkarrizketa burutze hori errazagoa izan daiteke gizatalde txikietan edozein norbanakok, batez bestean, gizataldeko gainontzeko norbanakoekin komunikatzeko gaitasun handiagoa duelako gizatalde handietan baino. Beste alde batetik, ordea, komunikatzeko gaitasuna erabatekoa ez denez N txikiagoa izango da. Nola uztartu kontrajarrita dauden bi joera hauek, N handiagoa eta N txikiagoa?

Komunikatzeko ahalmenak sartu beharra daude eredu matematikoan zenbaki enpirikoetara hurbiltzeko. Gai hau da azterbidean duguna oraindik, eta zabaldutako bide horretan ere jarraitzeko asmoa badugu, batez ere lortutako emaitzak nolakoak izan diren ikusita.

Komunikatzeko joera erabatekoa denean $N_2 = C(A+B,2)$ da, eta partziala denean, aldiz, aurreko atal batean azaldu denez, $N_2 = C(A+B,2) \cdot \frac{K}{100}$, K komunikatzeko ahalmena izanik. Beraz:

$$N = C(A+B,2) \cdot \frac{K/100}{p_2}$$

Adibidez, gizataldeen tamainak $(A+B)$, komunikatzeko ahalmenak $(\frac{K}{100})$ eta p_2 proportzioak emanda, honako taula hauek N elkarrizketa-kopuruak erakusten dituzte:

$p_2 = \frac{1}{3}$

	$\frac{K}{100}$			
$A+B$	%100	%80	%60	%0.05
25	900	720	540	
50	3675	2940	2205	
30000	$1350 \cdot 10^6$			674978

$p_2 = 0.5488$

	$\frac{K}{100}$			
$A+B$	%100	%80	%60	%0.05
25	547	438	328	
50	2232	1786	1340	
30000	$819 \cdot 10^6$			409500

Atal honen laburpen gisa zera esan daiteke: elkarrizketa guztiak kontuan hartuta, eta ez bakarrik binakakoak, kasurik okerrenean $N = 3 \cdot C(A+B, 2)$ hartu behar da, baldin eta beste zantzurik ez badugu kopuru hori errealitatera gehiago hurbiltzeko. Errealitatearen informazio zehatzagoa bagenu, 3 biderkagaiaren ordez $\frac{K/100}{p_2}$ faktore zuzentzailea aplikatu beharko genuke, $\frac{K}{100}$ eta p_2 proportzioen balioak errealitateko zantzuetara egokitu eta gero.

5 Taulak

Orain arte egindako azterketa matematikoari erabilera praktikoa eman nahian, taulatan laburbilduko ditugu eredu matematikoko adierazle garrantzitsuenak eta beren balioak eta korrespondentziak.

Gutxieneko δ zehaztasun-maila eta α' konfiantza-maila finkatuz gero, laginaren n tamaina erabakitzeko taulak elkarrizketen unibertsoaren Nren baitan daude. Baina populazioaren ezaugarriek lagundu dezakete tamaina hori txikitzen. Hain zuzen ere, zeregin horretan eragina duten bost ezagutza mota aipatu dira eredu honetan:

1. **Binakako elkarrizketen p_2 proportzioa**. Zenbat eta p_2 txikiagoa izan (hirunakako, launakako eta abarnakoekiko), orduan eta handiagoa izango da Nren tamaina. $p_2 \geq \frac{1}{3}$ dela jotzen da. Kasu enpiriko batean behatu da $p_2 = 0.55$.

2. **Komunikatzeko ahalmena ($\frac{K}{100}$)**. N, elkarrizketen unibertsoaren kopurua, txikiagotu dezake faktore honek. Nren tamaina kalkulatzeko, N_2 (binakako elkarrizketen kopurua) abiapuntutzat hartu eta $\frac{K}{100}$ aldiz txikiagoa egiten du. Ezer ez bada ezaguna, $n$ren tamaina erabakitzeko kasurik okerrena edo garestiena $K = 100$ kasua da. Komunikatzeko joera erabatekoa denean ($K = 100$), hau da, denek denekin hitz egiten dutenean, eta $p_2 = 0.3333$ kasuan, kasurik okerrenean ($\frac{K/100}{p_2} = 3$), alegia, Nren goi-muga $1.50 \cdot (A+B) \cdot (A+B-1)$ da. Nren behe-muga $0.71 \cdot (A+B) \cdot (A+B-1)$ da ($K = 100$ eta $p_2 = 0.70$).

3. **Gizarteratze-maila**. $p$rentzako goi- eta behe-mugak ezartzen laguntzen du. Ez du oraindik zenbakizko adierazlerik, baina ezagunak eta aztertuak dira muturreko kasuak, bai gizatalde erabat integratuetan bai erabat desintegratuetan.

4. **Gizasareko ele$B$idunen p_B proportzioa**. Euskararen erabilera-maila, hau da, elkarrizketen euskarazko p proportzioa mugatuta dago $p_B$ren arabera: goi-muga eta behe-muga.

5. **Leialtasun-maila ($\frac{L}{100}$)**. Elkarrizketen euskarazko p proportzioa txikiagotu egiten da (ehuneko L jaisten da) leialtasun-maila txikiagotu ahala. Leialtasun-mailak goi- eta behe-mugak txikiagotzen ditu. $L = 100$ balioak ez du p proportzioa txikiagotzen.

Hona adibide batzuk laginak behar duen tamaina ezagutzeko:

1. $\alpha' = \%95$ eta $\delta = 0.05$. $A+B = 100$, p_B ezezaguna ($p_B = 0.5$ aplikatuko da), $K = 100$, $p_2 = \%33.33$, $L = 100$.
$N = 1.5 \cdot (A+B) \cdot (A+B-1) = 14850$,

$n_0 = (\frac{1.96}{0.10})^2 \cdot 0.50 \cdot (1-0.50) = 384.16$,
$n = 384.16 \cdot \frac{1}{1+\frac{384.16}{14850}} = 374.47$: $n = 375$.

2. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $\mathbf{p_B = 0.55}$, $K = 100$, $p_2 = \%33.33$, $L = 100$. Ele$B$idunen p_B proportzioa zehaztuta.
   $N = 1.5 \cdot (A+B) \cdot (A+B-1) = 14850$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.50 \cdot (1-0.50) = 384.16$,
   $n = 384.16 \cdot \frac{1}{1+\frac{384.16}{14850}} = 374.47$: $n = 375$.

3. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $p_B = 0.55$, $K = 100$, $\mathbf{p_2 = \%50}$, $L = 100$. Binakako proportzioen p_2 proportzioa aldatuta: $p_2 = \%50$.
   $N = 1.0 \cdot (A+B) \cdot (A+B-1) = 9900$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.50 \cdot (1-0.50) = 384.16$,
   $n = 384.16 \cdot \frac{1}{1+\frac{384.16}{9900}} = 369.81$: $n = 370$.

4. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $p_B = 0.55$, $K = 70$, $p_2 = \%50$, $L = 100$. Komunikatzeko-maila aldatuta: $K = 70$
   $N = 1.0 \cdot (A+B) \cdot (A+B-1) \cdot \frac{70}{100} = 6930$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.50 \cdot (1-0.50) = 384.16$,
   $n = 384.16 \cdot \frac{1}{1+\frac{384.16}{6930}} = 362.37$: $n = 363$.

5. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $p_B = 0.55$, $K = 70$, $p_2 = \%50$, $\mathbf{L = 80}$. Leialtasun-maila aldatuta. Zuzendu p_B: $p_B \cdot \sqrt{\frac{80}{100}} = 0.4919$.
   $N = 1.0 \cdot (A+B) \cdot (A+B-1) \cdot \frac{70}{100} = 6930$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.4919 \cdot (1-0.4919) = 384.06$,
   $n = 384.06 \cdot \frac{1}{1+\frac{384.06}{6930}} = 362.29$: $n = 363$.

6. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $\mathbf{p_B \leq 0.35}$, $K = 70$, $p_2 = \%50$, $L = 80$. Ele$B$idunen p_B proportzioa aldatuta. Zuzendu p_B: $p_B \cdot \sqrt{\frac{80}{100}} = 0.3130$.
   $N = 1.0 \cdot (A+B) \cdot (A+B-1) \cdot \frac{70}{100} = 6930$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.0.3130 \cdot (1-0.0.3130) = 330.45$,
   $n = 330.45 \cdot \frac{1}{1+\frac{330.45}{6930}} = 314.20$: $n = 315$.

7. $\alpha' = \%95$ eta $\delta = 0.10$. $A+B = 100$, $p_B \leq 0.35$, $K = 70$, $\mathbf{L = 40}$. Leialtasun-maila aldatuta. Zuzendu p_B: $p_B \cdot \sqrt{\frac{40}{100}} = 0.2214$.
   i $N = 1.0 \cdot (A+B) \cdot (A+B-1) \cdot \frac{70}{100} = 6930$,
   $n_0 = (\frac{1.96}{0.10})^2 \cdot 0.2214 \cdot (1-0.2214) = 264.88$,
   $n = 264.86 \cdot \frac{1}{1+\frac{264.86}{6930}} = 254.31$: $n = 255$.

Ondorengo orrialdeetan taula batzuk aurkezten ditugu. Taula horiek eredu honen ondorio dira, eta, datu zehatzetatik abiatuta, laginaren tamaina kalkulatzeko balio dute. Taula asko eta asko beharko lirateke gerta daitezkeen kasu guztiak jasotzeko. Taula horiek guztiak txosten honetatik aparte, osagarri gisa, aurkeztuko dira. Dena den, taulen adibide batzuk ikusgai daude segidan.

$$\frac{K/100}{p_2}$$

K \ p_2	0.33	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7
1	0.03	0.029	0.025	0.022	0.02	0.018	0.017	0.015	0.014
5	0.15	0.143	0.125	0.111	0.10	0.091	0.083	0.077	0.071
10	0.30	0.286	0.250	0.222	0.20	0.182	0.167	0.154	0.143
15	0.45	0.429	0.375	0.333	0.30	0.273	0.250	0.231	0.214
20	0.60	0.571	0.500	0.444	0.40	0.364	0.333	0.308	0.286
25	0.75	0.714	0.625	0.556	0.50	0.455	0.417	0.385	0.357
30	0.90	0.857	0.750	0.667	0.60	0.545	0.500	0.462	0.429
35	1.05	1.000	0.875	0.778	0.70	0.636	0.583	0.538	0.500
40	1.20	1.143	1.000	0.889	0.80	0.727	0.667	0.615	0.571
45	1.35	1.286	1.125	1.000	0.90	0.818	0.750	0.692	0.643
50	1.50	1.429	1.250	1.111	1.00	0.909	0.833	0.769	0.714
55	1.65	1.571	1.375	1.222	1.10	1.000	0.917	0.846	0.786
60	1.80	1.714	1.500	1.333	1.20	1.091	1.000	0.923	0.857
65	1.95	1.857	1.625	1.444	1.30	1.182	1.083	1.000	0.929
70	2.10	2.000	1.750	1.556	1.40	1.273	1.167	1.077	1.000
75	2.25	2.143	1.875	1.667	1.50	1.364	1.250	1.154	1.071
80	2.40	2.286	2.000	1.778	1.60	1.455	1.333	1.231	1.143
85	2.55	2.429	2.125	1.889	1.70	1.545	1.417	1.308	1.214
90	2.70	2.571	2.250	2.000	1.80	1.636	1.500	1.385	1.286
95	2.85	2.714	2.375	2.111	1.90	1.727	1.583	1.462	1.357
100	3.00	2.857	2.500	2.222	2.00	1.818	1.667	1.538	1.429

N / $\frac{K/100}{P_2}$ \ A+B	25	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200
0.02	6	9	16	24	35	48	63	80	99	120	143	168	195	224	254	287	322	359	398
0.1	30	44	78	122	177	242	316	400	495	600	714	838	973	1118	1272	1436	1611	1796	1990
0.2	60	87	156	245	354	483	632	801	990	1199	1428	1677	1946	2235	2544	2873	3222	3591	3980
0.4	120	174	312	490	708	966	1264	1602	1980	2398	2856	3354	3892	4470	5088	5746	6444	7182	7960
0.6	180	261	468	735	1062	1449	1896	2403	2970	3597	4284	5031	5838	6705	7632	8619	9666	10773	11940
0.8	240	348	624	980	1416	1932	2528	3204	3960	4796	5712	6708	7784	8940	10176	11492	12888	14364	15920
1	300	435	780	1225	1770	2415	3160	4005	4950	5995	7140	8385	9730	11175	12720	14365	16110	17955	19900
1.2	360	522	936	1470	2124	2898	3792	4806	5940	7194	8568	10062	11676	13410	15264	17238	19332	21546	23880
1.4	420	609	1092	1715	2478	3381	4424	5607	6930	8393	9996	11739	13622	15645	17808	20111	22554	25137	27860
1.6	480	696	1248	1960	2832	3864	5056	6408	7920	9592	11424	13416	15568	17880	20352	22984	25776	28728	31840
1.8	540	783	1404	2205	3186	4347	5688	7209	8910	10791	12852	15093	17514	20115	22896	25857	28998	32319	35820
2	600	870	1560	2450	3540	4830	6320	8010	9900	11990	14280	16770	19460	22350	25440	28730	32220	35910	39800
2.2	660	957	1716	2695	3894	5313	6952	8811	10890	13189	15708	18447	21406	24585	27984	31603	35442	39501	43780
2.4	720	1044	1872	2940	4248	5796	7584	9612	11880	14388	17136	20124	23352	26820	30528	34476	38664	43092	47760
2.6	780	1131	2028	3185	4602	6279	8216	10413	12870	15587	18564	21801	25298	29055	33072	37349	41886	46683	51740
2.8	840	1218	2184	3430	4956	6762	8848	11214	13860	16786	19992	23478	27244	31290	35616	40222	45108	50274	55720
3	900	1305	2340	3675	5310	7245	9480	12015	14850	17985	21420	25155	29190	33525	38160	43095	48330	53865	59700

N / $\frac{K/100}{P_2}$ \ A+B	210	220	230	240	250	260	270	280	290	300	310	320	330	340	350
0.02	439	482	527	574	622	673	726	781	838	897	958	1021	1086	1153	1222
0.1	2194	2409	2634	2868	3112	3367	3632	3906	4190	4485	4790	5104	5428	5763	6108
0.2	4389	4818	5267	5736	6225	6734	7263	7812	8381	8970	9579	10208	10857	11526	12215
0.4	8778	9636	10534	11472	12450	13468	14526	15624	16762	17940	19158	20416	21714	23052	24430
0.6	13167	14454	15801	17208	18675	20202	21789	23436	25143	26910	28737	30624	32571	34578	36645
0.8	17556	19272	21068	22944	24900	26936	29052	31248	33524	35880	38316	40832	43428	46104	48860
1	21945	24090	26335	28680	31125	33670	36315	39060	41905	44850	47895	51040	54285	57630	61075
1.2	26334	28908	31602	34416	37350	40404	43578	46872	50286	53820	57474	61248	65142	69156	73290
1.4	30723	33726	36869	40152	43575	47138	50841	54684	58667	62790	67053	71456	75999	80682	85505
1.6	35112	38544	42136	45888	49800	53872	58104	62496	67048	71760	76632	81664	86856	92208	97720
1.8	39501	43362	47403	51624	56025	60606	65367	70308	75429	80730	86211	91872	97713	103734	109935
2	43890	48180	52670	57360	62250	67340	72630	78120	83810	89700	95790	102080	108570	115260	122150
2.2	48279	52998	57937	63096	68475	74074	79893	85932	92191	98670	105369	112288	119427	126786	134365
2.4	52668	57816	63204	68832	74700	80808	87156	93744	100572	107640	114948	122496	130284	138312	146580
2.6	57057	62634	68471	74568	80925	87542	94419	101556	108953	116610	124527	132704	141141	149838	158795
2.8	61446	67452	73738	80304	87150	94276	101682	109368	117334	125580	134106	142912	151998	161364	171010
3	65835	72270	79005	86040	93375	101010	108945	117180	125715	134550	143685	153120	162855	172890	183225

$\frac{K/100}{p_2}$ \ N (A+B)	400	450	500	550	600	650	700	750	800	850	900	950	1000
0.02	1596	2020	2495	3020	3594	4218	4893	5618	6392	7216	8091	9016	9990
0.1	7980	10102	12475	15098	17970	21092	24465	28088	31960	36082	40455	45078	49950
0.2	15960	20205	24950	30195	35940	42185	48930	56175	63920	72165	80910	90155	99900
0.4	31920	40410	49900	60390	71880	84370	97860	112350	127840	144330	161820	180310	199800
0.6	47880	60615	74850	90585	107820	126555	146790	168525	191760	216495	242730	270465	299700
0.8	63840	80820	99800	120780	143760	168740	195720	224700	255680	288660	323640	360620	399600
1	79800	101025	124750	150975	179700	210925	244650	280875	319600	360825	404550	450775	499500
1.2	95760	121230	149700	181170	215640	253110	293580	337050	383520	432990	485460	540930	599400
1.4	111720	141435	174650	211365	251580	295295	342510	393225	447440	505155	566370	631085	699300
1.6	127680	161640	199600	241560	287520	337480	391440	449400	511360	577320	647280	721240	799200
1.8	143640	181845	224550	271755	323460	379665	440370	505575	575280	649485	728190	811395	899100
2	159600	202050	249500	301950	359400	421850	489300	561750	639200	721650	809100	901550	999000
2.2	175560	222255	274450	332145	395340	464035	538230	617925	703120	793815	890010	991705	1098900
2.4	191520	242460	299400	362340	431280	506220	587160	674100	767040	865980	970920	1081860	1198800
2.6	207480	262665	324350	392535	467220	548405	636090	730275	830960	938145	1051830	1172015	1298700
2.8	223440	282870	349300	422730	503160	590590	685020	786450	894880	1010310	1132740	1262170	1398600
3	239400	303075	374250	452925	539100	632775	733950	842625	958800	1082475	1213650	1352325	1498500

$\frac{K/100}{p_2}$ \ N (A+B)	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100	2200	2300	2400	2500
0.02	12089	14388	16887	19586	22485	25584	28883	32382	36081	39980	44079	48378	52877	57576	62475
0.1	60445	71940	84435	97930	112425	127920	144415	161910	180405	199900	220395	241890	264385	287880	312375
0.2	120890	143880	168870	195860	224850	255840	288830	323820	360810	399800	440790	483780	528770	575760	624750
0.4	241780	287760	337740	391720	449700	511680	577660	647640	721620	799600	881580	967560	1057540	1151520	1249500
0.6	362670	431640	506610	587580	674550	767520	866490	971460	1082430	1199400	1322370	1451340	1586310	1727280	1874250
0.8	483560	575520	675480	783440	899400	1023360	1155320	1295280	1443240	1599200	1763160	1935120	2115080	2303040	2499000
1	604450	719400	844350	979300	1124250	1279200	1444150	1619100	1804050	1999000	2203950	2418900	2643850	2878800	3123750
1.2	725340	863280	1013220	1175160	1349100	1535040	1732980	1942920	2164860	2398800	2644740	2902680	3172620	3454560	3748500
1.4	846230	1007160	1182090	1371020	1573950	1790880	2021810	2266740	2525670	2798600	3085530	3386460	3701390	4030320	4373250
1.6	967120	1151040	1350960	1566880	1798800	2046720	2310640	2590560	2886480	3198400	3526320	3870240	4230160	4606080	4998000
1.8	1088010	1294920	1519830	1762740	2023650	2302560	2599470	2914380	3247290	3598200	3967110	4354020	4758930	5181840	5622750
2	1208900	1438800	1688700	1958600	2248500	2558400	2888300	3238200	3608100	3998000	4407900	4837800	5287700	5757600	6247500
2.2	1329790	1582680	1857570	2154460	2473350	2814240	3177130	3562020	3968910	4397800	4848690	5321580	5816470	6333360	6872250
2.4	1450680	1726560	2026440	2350320	2698200	3070080	3465960	3885840	4329720	4797600	5289480	5805360	6345240	6909120	7497000
2.6	1571570	1870440	2195310	2546180	2923050	3325920	3754790	4209660	4690530	5197400	5730270	6289140	6874010	7484880	8121750
2.8	1692460	2014320	2364180	2742040	3147900	3581760	4043620	4533480	5051340	5597200	6171060	6772920	7402780	8060640	8746500
3	1813350	2158200	2533050	2937900	3372750	3837600	4332450	4857300	5412150	5997000	6611850	7256700	7931550	8636400	9371250

$\frac{K/100}{p_2}$ \ N (A+B)	2600	2700	2800	2900	3000	3100	3200	3300	3400	3500	3600	3700	3800	3900	4000
0.02	67574	72873	78372	84071	89970	96069	102368	108867	115566	122465	129564	136863	144362	152061	159960
0.1	337870	364365	391860	420355	449850	480345	511840	544335	577830	612325	647820	684315	721810	760305	799800
0.2	675740	728730	783720	840710	899700	960690	1023680	1088670	1155660	1224650	1295640	1368630	1443620	1520610	1599600
0.4	1351480	1457460	1567440	1681420	1799400	1921380	2047360	2177340	2311320	2449300	2591280	2737260	2887240	3041220	3199200
0.6	2027220	2186190	2351160	2522130	2699100	2882070	3071040	3266010	3466980	3673950	3886920	4105890	4330860	4561830	4798800
0.8	2702960	2914920	3134880	3362840	3598800	3842760	4094720	4354680	4622640	4898600	5182560	5474520	5774480	6082440	6398400
1	3378700	3643650	3918600	4203550	4498500	4803450	5118400	5443350	5778300	6123250	6478200	6843150	7218100	7603050	7998000
1.2	4054440	4372380	4702320	5044260	5398200	5764140	6142080	6532020	6933960	7347900	7773840	8211780	8661720	9123660	9597600
1.4	4730180	5101110	5486040	5884970	6297900	6724830	7165760	7620690	8089620	8572550	9069480	9580410	10105340	10644270	11197200
1.6	5405920	5829840	6269760	6725680	7197600	7685520	8189440	8709360	9245280	9797200	10365120	10949040	11548960	12164880	12796800
1.8	6081660	6558570	7053480	7566390	8097300	8646210	9213120	9798030	10400940	11021850	11660760	12317670	12992580	13685490	14396400
2	6757400	7287300	7837200	8407100	8997000	9606900	10236800	10886700	11556600	12246500	12956400	13686300	14436200	15206100	15996000
2.2	7433140	8016030	8620920	9247810	9896700	10567590	11260480	11975370	12712260	13471150	14252040	15054930	15879820	16726710	17595600
2.4	8108880	8744760	9404640	10088520	10796400	11528280	12284160	13064040	13867920	14695800	15547680	16423560	17323440	18247320	19195200
2.6	8784620	9473490	10188360	10929230	11696100	12488970	13307840	14152710	15023580	15920450	16843320	17792190	18767060	19767930	20794800
2.8	9460360	10202220	10972080	11769940	12595800	13449660	14331520	15241380	16179240	17145100	18138960	19160820	20210680	21288540	22394400
3	10136100	10930950	11755800	12610650	13495500	14410350	15355200	16330050	17334900	18369750	19434600	20529450	21654300	22809150	23994000

$\frac{K/100}{p_2}$ \ N (A+B)	4100	4200	4300	4400	4500	4600	4700	4800	4900	5000	5100	5200	5300	5400	5500
0.02	168059	176358	184857	193556	202455	211554	220853	230352	240051	249950	260049	270348	280847	291546	302445
0.1	840295	881790	924285	967780	1012275	1057770	1104265	1151760	1200255	1249750	1300245	1351740	1404235	1457730	1512225
0.2	1680590	1763580	1848570	1935560	2024550	2115540	2208530	2303520	2400510	2499500	2600490	2703480	2808470	2915460	3024450
0.4	3361180	3527160	3697140	3871120	4049100	4231080	4417060	4607040	4801020	4999000	5200980	5406960	5616940	5830920	6048900
0.6	5041770	5290740	5545710	5806680	6073650	6346620	6625590	6910560	7201530	7498500	7801470	8110440	8425410	8746380	9073350
0.8	6722360	7054320	7394280	7742240	8098200	8462160	8834120	9214080	9602040	9998000	10401960	10813920	11233880	11661840	12097800
1	8402950	8817900	9242850	9677800	10122750	10577700	11042650	11517600	12002550	12497500	13002450	13517400	14042350	14577300	15122250
1.2	10083540	10581480	11091420	11613360	12147300	12693240	13251180	13821120	14403060	14997000	15602940	16220880	16850820	17492760	18146700
1.4	11764130	12345060	12939990	13548920	14171850	14808780	15459710	16124640	16803570	17496500	18203430	18924360	19659290	20408220	21171150
1.6	13444720	14108640	14788560	15484480	16196400	16924320	17668240	18428160	19204080	19996000	20803920	21627840	22467760	23323680	24195600
1.8	15125310	15872220	16637130	17420040	18220950	19039860	19876770	20731680	21604590	22495500	23404410	24331320	25276230	26239140	27220050
2	16805900	17635800	18485700	19355600	20245500	21155400	22085300	23035200	24005100	24995000	26004900	27034800	28084700	29154600	30244500
2.2	18486490	19399380	20334270	21291160	22270050	23270940	24293830	25338720	26405610	27494500	28605390	29738280	30893170	32070060	33268950
2.4	20167080	21162960	22182840	23226720	24294600	25386480	26502360	27642240	28806120	29994000	31205880	32441760	33701640	34985520	36293400
2.6	21847670	22926540	24031410	25162280	26319150	27502020	28710890	29945760	31206630	32493500	33806370	35145240	36510110	37900980	39317850
2.8	23528260	24690120	25879980	27097840	28343700	29617560	30919420	32249280	33607140	34993000	36406860	37848720	39318580	40816440	42342300
3	25208850	26453700	27728550	29033400	30368250	31733100	33127950	34552800	36007650	37492500	39007350	40552200	42127050	43731900	45366750

$$p_B = \frac{B}{A+B}$$

p_B zuzendua	L 10	20	30	40	50	60	70	80	90	100
0.05	0.02	0.02	0.03	0.03	0.04	0.04	0.04	0.04	0.05	0.05
0.1	0.03	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.10
0.15	0.05	0.07	0.08	0.09	0.11	0.12	0.13	0.13	0.14	0.15
0.2	0.06	0.09	0.11	0.13	0.14	0.15	0.17	0.18	0.19	0.20
0.25	0.08	0.11	0.14	0.16	0.18	0.19	0.21	0.22	0.24	0.25
0.3	0.09	0.13	0.16	0.19	0.21	0.23	0.25	0.27	0.28	0.30
0.35	0.11	0.16	0.19	0.22	0.25	0.27	0.29	0.31	0.33	0.35
0.4	0.13	0.18	0.22	0.25	0.28	0.31	0.33	0.36	0.38	0.40
0.45	0.14	0.20	0.25	0.28	0.32	0.35	0.38	0.40	0.43	0.45
0.5	0.16	0.22	0.27	0.32	0.35	0.39	0.42	0.45	0.47	0.50
0.55	0.17	0.25	0.30	0.35	0.39	0.43	0.46	0.49	0.52	0.55
0.6	0.19	0.27	0.33	0.38	0.42	0.46	0.50	0.54	0.57	0.60
0.65	0.21	0.29	0.36	0.41	0.46	0.50	0.54	0.58	0.62	0.65
0.7	0.22	0.31	0.38	0.44	0.49	0.54	0.59	0.63	0.66	0.70
0.75	0.24	0.34	0.41	0.47	0.53	0.58	0.63	0.67	0.71	0.75
0.8	0.25	0.36	0.44	0.51	0.57	0.62	0.67	0.72	0.76	0.80
0.85	0.27	0.38	0.47	0.54	0.60	0.66	0.71	0.76	0.81	0.85
0.9	0.28	0.40	0.49	0.57	0.64	0.70	0.75	0.80	0.85	0.90
0.95	0.30	0.42	0.52	0.60	0.67	0.74	0.79	0.85	0.90	0.95
1	0.32	0.45	0.55	0.63	0.71	0.77	0.84	0.89	0.95	1.00

		$\left(\frac{z}{\delta}\right)^2$	δ = 0.01	0.02	0.03	0.04	0.05
0.90		1.645	27060.25	6765.06	3006.69	1691.27	1082.41
0.95	**z**	1.96	38416.00	9604.00	4268.44	2401.00	1536.64
0.99		2.575	66306.25	16576.56	7367.36	4144.14	2652.25

n_0 $\left(\frac{z}{\delta}\right)^2$

p_B zuzendua	1082.41	1536.64	1691.27	2401	2652.25	3006.69	4144.14	4268.44	6765.06	7367.36	9604	16576.56	27060.25	38416	66306.25
0.05	52	73	81	115	126	143	197	203	322	350	457	788	1286	1825	3150
0.1	98	139	153	217	239	271	373	385	609	664	865	1492	2436	3458	5968
0.15	139	196	216	307	339	384	529	545	863	940	1225	2114	3451	4899	8455
0.2	174	246	271	385	425	482	664	683	1083	1179	1537	2653	4330	6147	10609
0.25	203	289	318	451	498	564	778	801	1269	1382	1801	3109	5074	7203	12433
0.3	228	323	356	505	557	632	871	897	1421	1548	2017	3482	5683	8068	13925
0.35	247	350	385	547	604	685	943	972	1540	1677	2185	3772	6157	8740	15085
0.4	260	369	406	577	637	722	995	1025	1624	1769	2305	3979	6495	9220	15914
0.45	268	381	419	595	657	745	1026	1057	1675	1824	2377	4103	6698	9508	16411
0.5	271	385	423	601	664	752	1037	1068	1692	1842	2401	4145	6766	9604	16577
0.55	271	385	423	601	664	752	1037	1068	1692	1842	2401	4145	6766	9604	16577
0.6	271	385	423	601	664	752	1037	1068	1692	1842	2401	4145	6766	9604	16577
0.65	271	385	423	601	664	752	1037	1068	1692	1842	2401	4145	6766	9604	16577
0.7	271	385	423	601	664	752	1037	1068	1692	1842	2401	4145	6766	9604	16577
0.75	267	379	417	591	653	740	1020	1051	1665	1814	2364	4080	6660	9454	16318
0.8	250	355	390	554	612	693	955	984	1559	1698	2213	3820	6235	8852	15277
0.85	218	309	340	482	532	603	831	856	1357	1478	1926	3324	5426	7703	13294
0.9	167	237	261	370	409	463	638	657	1042	1134	1479	2552	4165	5913	10205
0.95	96	136	149	212	234	265	365	376	596	649	846	1459	2382	3381	5835

n **N**

n_0	100	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000
100	50	67	75	80	84	86	88	89	90	91	92	93	93	94	94	95	95	95	95	96
200	67	100	121	134	143	150	156	160	164	167	170	172	174	175	177	178	179	180	181	182
300	75	120	150	172	188	200	210	219	225	231	236	240	244	248	250	253	255	258	260	261
400	80	134	172	200	223	241	255	267	277	286	294	300	306	312	316	320	324	328	331	334
500	84	143	188	223	250	273	292	308	322	334	344	353	362	369	375	381	387	392	396	400
600	86	150	200	240	273	300	324	343	361	375	389	400	411	420	429	437	444	450	456	462
700	88	156	210	255	292	324	350	374	394	412	428	443	455	467	478	487	496	504	512	519
800	89	160	219	267	308	343	374	400	424	445	464	481	496	510	522	534	544	554	563	572
900	90	164	225	277	322	360	394	424	450	474	495	515	532	548	563	576	589	600	611	621
1000	91	167	231	286	334	375	412	445	474	500	524	546	566	584	601	616	630	643	656	667
1100	92	170	236	294	344	389	428	464	495	524	550	574	596	617	635	652	668	683	697	710
1200	93	172	240	300	353	400	443	480	515	546	574	600	624	647	667	686	704	721	736	750
1300	93	174	244	306	362	411	455	496	532	566	596	625	650	675	697	718	737	755	772	788
1400	94	175	248	312	369	420	467	510	548	584	617	647	675	700	725	747	768	788	807	824
1500	94	177	250	316	375	429	478	522	563	600	635	667	697	725	750	775	797	819	839	858
1600	95	178	253	320	381	437	487	534	576	616	652	686	718	747	775	800	825	848	869	889
1700	95	179	255	324	387	444	496	544	589	630	668	704	737	768	797	825	850	875	898	919
1800	95	180	258	328	392	450	504	554	600	643	683	720	755	788	819	848	875	900	925	948
1900	95	181	260	331	396	457	512	563	611	656	697	736	772	807	839	869	898	925	950	975
2000	96	182	261	334	400	462	519	572	621	667	710	750	788	824	858	889	919	948	975	1000

n_0 \ N	2000	2500	3000	3500	4000	4500	5000	5500	6000	6500	7000	7500	8000	8500	9000	9500	10000	10500	11000	11500	12000	12500	13000	13500	14000	14500	15000
100	96	97	97	98	98	98	99	99	99	99	99	99	99	99	99	99	100	100	100	100	100	100	100	100	100	100	100
200	182	186	188	190	191	192	193	193	194	195	195	195	196	196	196	196	197	197	197	197	197	197	197	198	198	198	198
300	261	268	273	277	280	282	284	285	286	287	288	289	290	290	291	291	292	292	293	293	293	293	294	294	294	294	295
400	334	345	353	359	364	368	371	373	375	377	379	380	381	383	383	384	385	386	386	387	388	388	389	389	389	390	390
500	400	417	429	438	445	450	455	459	462	465	467	469	471	473	474	476	477	478	479	480	480	481	482	483	483	484	484
600	462	484	500	513	522	530	536	541	546	550	553	556	559	561	563	565	567	568	569	571	572	573	574	575	576	577	577
700	519	547	568	584	596	606	615	621	627	632	637	641	644	647	650	652	655	657	659	660	662	663	665	666	667	668	669
800	572	607	632	652	667	680	690	699	706	713	718	723	728	732	735	738	741	744	746	748	750	752	754	756	757	759	760
900	621	662	693	716	735	750	763	774	783	791	798	804	809	814	819	823	826	829	832	835	838	840	842	844	846	848	850
1000	667	715	750	778	800	819	834	847	858	867	875	883	889	895	900	905	910	914	917	920	924	926	929	932	934	936	938
1100	710	764	805	837	863	884	902	917	930	941	951	960	968	974	981	986	991	996	1000	1004	1008	1012	1015	1018	1020	1023	1025
1200	750	811	858	894	924	948	968	986	1000	1013	1025	1035	1044	1052	1059	1066	1072	1077	1082	1087	1091	1095	1099	1103	1106	1109	1112
1300	788	856	907	948	982	1009	1032	1052	1069	1084	1097	1108	1119	1128	1136	1144	1151	1157	1163	1168	1173	1178	1182	1186	1190	1194	1197
1400	824	898	955	1000	1038	1068	1094	1116	1136	1152	1167	1180	1192	1203	1212	1221	1229	1236	1242	1249	1254	1259	1264	1269	1273	1277	1281
1500	858	938	1000	1050	1091	1125	1154	1179	1200	1219	1236	1250	1264	1275	1286	1296	1305	1313	1321	1327	1334	1340	1345	1350	1355	1360	1364
1600	889	976	1044	1099	1143	1181	1213	1240	1264	1284	1303	1319	1334	1347	1359	1370	1380	1389	1397	1405	1412	1419	1425	1431	1436	1441	1446
1700	919	1012	1086	1145	1193	1234	1269	1299	1325	1348	1368	1386	1403	1417	1430	1442	1453	1464	1473	1482	1490	1497	1504	1510	1516	1522	1527
1800	948	1047	1125	1189	1242	1286	1324	1357	1385	1410	1432	1452	1470	1486	1500	1514	1526	1537	1547	1557	1566	1574	1582	1589	1595	1602	1608
1900	975	1080	1164	1232	1289	1336	1377	1413	1444	1471	1495	1516	1536	1553	1569	1584	1597	1609	1621	1631	1641	1650	1658	1666	1673	1680	1687
2000	1000	1112	1201	1273	1334	1385	1429	1467	1500	1530	1556	1579	1600	1620	1637	1653	1667	1680	1693	1704	1715	1725	1734	1742	1750	1758	1765

n **N**

n_0	100	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000
2000	96	182	261	334	400	462	519	572	621	667	710	750	788	824	858	889	919	948	975	1000
2500	97	186	268	345	417	484	547	607	662	715	764	811	856	898	938	976	1012	1047	1080	1112
3000	97	188	273	353	429	500	568	632	693	750	805	858	907	955	1000	1044	1086	1125	1164	1200
3500	98	190	277	359	438	513	584	652	716	778	837	894	948	1000	1050	1099	1145	1189	1232	1273
4000	98	191	280	364	445	522	596	667	735	800	863	924	982	1038	1091	1143	1193	1242	1289	1334
4500	98	192	282	368	450	530	606	680	750	819	884	948	1009	1068	1125	1181	1234	1286	1336	1385
5000	99	193	284	371	455	536	615	690	763	834	902	968	1032	1094	1154	1213	1269	1324	1377	1429
5500	99	193	285	373	459	541	621	699	774	847	917	986	1052	1116	1179	1240	1299	1357	1413	1467
6000	99	194	286	375	462	546	627	706	783	858	930	1000	1069	1136	1200	1264	1325	1385	1444	1500
6500	99	195	287	377	465	550	632	713	791	867	941	1013	1084	1152	1219	1284	1348	1410	1471	1530
7000	99	195	288	379	467	553	637	718	798	875	951	1025	1097	1167	1236	1303	1368	1432	1495	1556
7500	99	195	289	380	469	556	641	723	804	883	960	1035	1108	1180	1250	1319	1386	1452	1516	1579
8000	99	196	290	381	471	559	644	728	809	889	968	1044	1119	1192	1264	1334	1403	1470	1536	1600
8500	99	196	290	383	473	561	647	732	814	895	974	1052	1128	1203	1275	1347	1417	1486	1553	1620
9000	99	196	291	383	474	563	650	735	819	900	981	1059	1136	1212	1286	1359	1430	1500	1569	1637
9500	99	196	291	384	475	565	652	738	823	905	986	1066	1144	1221	1296	1370	1442	1514	1584	1653
10000	100	197	292	385	477	567	655	741	826	910	991	1072	1151	1229	1305	1380	1453	1526	1597	1667
10500	100	197	292	386	478	568	657	744	829	914	996	1077	1157	1236	1313	1389	1464	1537	1609	1680
11000	100	197	293	386	479	569	659	746	832	917	1000	1082	1163	1242	1321	1397	1473	1547	1621	1693
11500	100	197	293	387	480	571	660	748	835	920	1004	1087	1168	1249	1327	1405	1482	1557	1631	1704
12000	100	197	293	388	480	572	662	750	838	924	1008	1091	1173	1254	1334	1412	1490	1566	1641	1715
12500	100	197	293	388	481	573	663	752	840	926	1012	1095	1178	1259	1340	1419	1497	1574	1650	1725
13000	100	197	294	389	482	574	665	754	842	929	1015	1099	1182	1264	1345	1425	1504	1582	1658	1734
13500	100	198	294	389	483	575	666	756	844	932	1018	1103	1186	1269	1350	1431	1510	1589	1666	1742
14000	100	198	294	389	483	576	667	757	846	934	1020	1106	1190	1273	1355	1436	1516	1595	1673	1750
14500	100	198	294	390	484	577	668	759	848	936	1023	1109	1194	1277	1360	1441	1522	1602	1680	1758
15000	100	198	295	390	484	577	669	760	850	938	1025	1112	1197	1281	1364	1446	1527	1608	1687	1765

n **N**

n_0	2000	2500	3000	3500	4000	4500	5000	5500	6000	6500	7000	7500	8000	8500	9000	9500	10000	10500	11000	11500	12000	12500	13000	13500	14000	14500	15000
2000	1000	1112	1201	1273	1334	1385	1429	1467	1500	1530	1556	1579	1600	1620	1637	1653	1667	1680	1693	1704	1715	1725	1734	1742	1750	1758	1765
2500	1112	1250	1364	1459	1539	1608	1667	1719	1765	1806	1843	1875	1905	1932	1957	1980	2000	2020	2038	2054	2069	2084	2097	2110	2122	2133	2143
3000	1200	1364	1500	1616	1715	1801	1875	1942	2000	2053	2100	2143	2182	2218	2250	2280	2308	2334	2358	2380	2400	2420	2438	2455	2471	2486	2500
3500	1273	1459	1616	1750	1867	1969	2059	2139	2211	2275	2334	2387	2435	2480	2520	2558	2593	2625	2656	2684	2710	2735	2758	2780	2800	2820	2838
4000	1334	1539	1715	1867	2000	2118	2223	2316	2401	2477	2546	2609	2667	2720	2770	2815	2858	2897	2934	2968	3000	3031	3059	3086	3112	3136	3158
4500	1385	1608	1800	1969	2118	2250	2369	2475	2572	2660	2740	2813	2880	2943	3000	3054	3104	3150	3194	3235	3273	3309	3343	3375	3406	3435	3462
5000	1429	1667	1875	2059	2223	2369	2500	2620	2728	2827	2917	3001	3077	3149	3215	3276	3334	3388	3438	3485	3530	3572	3612	3649	3685	3718	3750
5500	1467	1719	1942	2139	2316	2475	2620	2750	2870	2980	3081	3174	3260	3340	3414	3484	3549	3610	3667	3721	3772	3820	3865	3908	3949	3988	4025
6000	1500	1765	2000	2211	2400	2572	2728	2870	3000	3120	3231	3334	3429	3518	3601	3678	3750	3819	3883	3943	4000	4055	4106	4154	4200	4244	4286
6500	1530	1806	2053	2275	2477	2660	2827	2980	3121	3250	3371	3483	3587	3684	3775	3860	3940	4015	4086	4153	4217	4277	4334	4388	4440	4489	4535
7000	1556	1843	2100	2334	2546	2740	2917	3081	3231	3371	3500	3621	3734	3839	3938	4031	4118	4201	4278	4352	4422	4488	4550	4610	4667	4721	4773
7500	1579	1875	2143	2387	2609	2813	3000	3174	3334	3483	3621	3750	3871	3985	4091	4192	4286	4375	4460	4540	4616	4688	4757	4822	4884	4944	5000
8000	1600	1905	2182	2435	2667	2880	3077	3260	3429	3587	3734	3871	4000	4122	4236	4343	4445	4541	4632	4718	4801	4879	4953	5024	5091	5156	5218
8500	1620	1932	2218	2480	2720	2943	3149	3340	3518	3684	3839	3985	4122	4250	4372	4487	4595	4698	4795	4888	4976	5060	5140	5216	5289	5359	5426
9000	1637	1957	2250	2520	2770	3000	3215	3414	3600	3775	3938	4091	4236	4372	4500	4622	4737	4847	4950	5049	5143	5233	5319	5401	5479	5554	5625
9500	1653	1980	2281	2558	2815	3054	3276	3484	3678	3860	4031	4192	4343	4487	4622	4750	4872	4988	5098	5203	5303	5398	5489	5577	5660	5740	5817
10000	1667	2000	2308	2593	2858	3104	3334	3549	3750	3940	4118	4286	4445	4595	4737	4872	5000	5122	5239	5349	5455	5556	5653	5745	5834	5919	6001
10500	1680	2020	2334	2625	2897	3150	3388	3610	3819	4015	4200	4375	4541	4698	4847	4988	5122	5250	5373	5489	5600	5707	5809	5907	6000	6091	6177
11000	1693	2038	2358	2656	2934	3194	3438	3667	3883	4086	4278	4460	4632	4795	4950	5098	5239	5373	5500	5623	5740	5852	5959	6062	6161	6255	6347
11500	1704	2054	2380	2684	2968	3235	3485	3721	3943	4153	4352	4540	4718	4888	5049	5203	5349	5489	5623	5750	5873	5990	6103	6210	6314	6414	6510
12000	1715	2069	2400	2710	3000	3273	3530	3772	4000	4217	4422	4616	4800	4976	5143	5303	5455	5600	5740	5873	6000	6123	6240	6353	6462	6567	6667
12500	1725	2084	2420	2735	3031	3309	3572	3820	4055	4277	4488	4688	4879	5060	5233	5398	5556	5707	5852	5990	6123	6250	6373	6491	6604	6713	6819
13000	1734	2097	2438	2758	3059	3343	3612	3865	4106	4334	4550	4757	4953	5140	5319	5489	5653	5809	5959	6103	6241	6373	6500	6623	6741	6855	6965
13500	1742	2110	2455	2780	3086	3375	3649	3908	4154	4388	4610	4822	5024	5216	5400	5577	5745	5907	6062	6210	6353	6491	6623	6750	6873	6992	7106
14000	1750	2122	2471	2800	3112	3406	3685	3949	4200	4440	4667	4884	5091	5289	5479	5660	5834	6000	6161	6314	6462	6604	6741	6873	7000	7123	7242
14500	1758	2133	2486	2820	3136	3435	3718	3988	4244	4489	4721	4944	5156	5359	5554	5740	5919	6090	6255	6414	6567	6713	6855	6992	7123	7250	7373
15000	1765	2143	2500	2838	3158	3462	3750	4025	4286	4535	4773	5000	5218	5426	5625	5817	6000	6177	6347	6510	6667	6819	6965	7106	7242	7373	7500