

BUTLLETÍ

DE LA SOCIETAT CATALANA
DE MATEMÀTIQUES
Institut d'Estudis Catalans

Í n d e x

- ❑ *Carles M. Cuadras*
El llegat de Galton, Pearson, Fréchet i d'altres:
com mesurar i interpretar l'associació estadística 5

 - ❑ *Xavier Mora i Maria Oliver*
Eleccions mitjançant el vot d'aprovació.
El mètode de Phragmén i algunes variants 57

 - ❑ English summaries 103
-



Volum 30 • Número 1 • Any 2015



Institut
d'Estudis
Catalans

BUTLLETÍ

DE LA SOCIETAT CATALANA
DE MATEMÀTIQUES

Institut d'Estudis Catalans

Volum 30 • Número 1 • Juny 2015

BARCELONA
2015

© dels autors dels articles

Editat per la Societat Catalana de Matemàtiques
filial de l'Institut d'Estudis Catalans
Carrer del Carme, 47
08001 Barcelona

Text revisat lingüísticament
per la Unitat de Correcció del Servei Editorial de l'IEC.

Imprès a Limpergraf, SL
Polígon industrial Can Salvatella
Carrer de Mogoda, 29-31
08210 Barberà del Vallès

ISSN: 0214-316-X
Dipòsit Legal: B. 19272-1987

Són rigorosament prohibides, sense l'autorització escrita dels titulars del *copyright*, la reproducció total o parcial d'aquesta obra per qualsevol procediment i suport, incloent-hi la reprografia i el tractament informàtic, la distribució d'exemplars mitjançant lloguer o préstec comercial, la inclusió total o parcial en bases de dades i la consulta a través de xarxa telemàtica o d'Internet. Les infraccions d'aquests drets estan sotmeses a les sancions establertes per les lleis.

Índex

CARLES M. CUADRAS	
El llegat de Galton, Pearson, Fréchet i d'altres: com mesurar i interpretar l'associació estadística	5
XAVIER MORA I MARIA OLIVER	
Eleccions mitjançant el vot d'aprovació. El mètode de Phragmén i algunes variants	57
English summaries	103

El llegat de Galton, Pearson, Fréchet i d'altres: com mesurar i interpretar l'associació estadística

CARLES M. CUADRAS

Resum: Presentem en tres parts els conceptes de correlació i d'associació estadística, començant per la noció de correlació de Galton, millorada per Pearson. Utilitzem com a il·lustració les dades clàssiques de Galton i Pearson sobre heretabilitat de pares i fills respecte a l'estatura. La segona part explica com s'han d'estudiar les mateixes dades des d'una perspectiva multivariant (anàlisi de correlació canònica i de correspondències). Utilitzem també dades de Fisher. Mostrem com podem associar dades de tipus general mitjançant distàncies. La tercera part la dediquem a les distribucions bivariants. Presentem la teoria de funcions i valors propis per a dos nuclis, que s'aplica al desenvolupament diagonal d'una distribució bivariant, incloent-hi els desenvolupaments continus en termes d'integrals. Proposem una família de còpules canòniques, que permet generar distribucions bivariants.

Paraules clau: dades de Galton i Pearson, correlació intraclàssica, correlació canònica, anàlisi de correspondències, associació basada en distàncies, operadors integrals, distribucions amb marginals donades, funcions canòniques.

Classificació MSC2010: 62H20, 60E05.

Part I: Estadística clàssica

1 Introducció

Des de sempre hem volgut esbrinar la causa d'un fenomen. Per què s'ha produït un incendi? Què ha provocat una guerra? Quina és la principal causa del càncer de pulmó? Per què un cantant té èxit? En algunes parts de la física i la mecànica, el binomi causa-efecte és ben conegut. Sovint les relacions entre variables observables i mesurables són deterministes i es poden expressar amb una relació funcional $y = f(x)$, que lliga una variable dependent y amb una d'independent x . Aquesta relació és molt important i ben estudiada en matemàtiques i altres ciències exactes. Però en biologia, economia, sociologia, meteorologia, etc., a ningú se li acudiria imaginar que pugui existir una relació

tan exacta entre dues variables. Per exemple, la pressió arterial d'una persona té relació amb el pes, però la relació és «imprecisa».

Com podem relacionar dues variables? En principi podem suposar $y = f(x) + e$, on $f(x)$ representa la part determinista i e la part aleatòria de cada observació y . Altrament dit,

$$\text{observació} = \text{model} + \text{error},$$

entenent per «error» la desviació del model. Aleshores és natural que ens plantegem com cal mesurar aquesta relació mitjançant un coeficient entre 0 i 1, de manera que el valor 0 signifiqui manca de relació i el valor 1 es pugui interpretar com que hi ha relació perfecta, matemàticament parlant.¹

Aleshores es planteja la determinació de la funció f i la naturalesa de la desviació e , sovint interpretada com una variable aleatòria amb distribució normal.

2 Galton i el coeficient de correlació

L'any 1886 Francis Galton [37], contemporani de Mendel i cosí de Darwin, es va plantejar la tasca de mesurar objectivament la relació entre dues variables observades sobre una mateixa població, a fi de donar una prova científica de la teoria de l'evolució. Aleshores va tenir la idea genial de proposar el coeficient de correlació r , que va il·lustrar amb les dades de $n = 205$ famílies. Va anotar les variables següents:

$$\begin{aligned} X_1 &= \text{alçada del pare}, & Y_1 &= \text{alçada del fill}, \\ X_2 &= \text{alçada de la mare}, & Y_2 &= \text{alçada de la filla}. \end{aligned}$$

D'entrada Galton va trobar que els pares i els fills eren, en general, més alts que les mares i les filles, és a dir, hi havia dades de dues poblacions. Encertadament va intuir que el coeficient de correlació hauria de ser una mesura d'associació vàlida per a una sola població. De fet, quan barregem dues poblacions, es poden presentar paradoxes (vegeu la secció 7). Per tant, Galton va decidir introduir una correcció consistent a augmentar l'alçada de les dones. Concretament, va considerar les variables

$$X = (X_1 + 1.08X_2)/2, \quad Y = Y_1 \text{ o } Y = 1.08Y_2,$$

on, per a cada X , la variable Y pot prendre tants valors com fills i filles tenia el matrimoni, obtenint $n = 934$ parelles d'observacions. La variable X rep el nom de *mid-parent* i encara avui dia s'utilitza en antropologia.

¹ Gairebé mai sabrem si el model representat per la funció f és el correcte. Un model és una simplificació de la realitat. M. Chasles va dir: «La geometria és l'art de raonar bé sobre figures falses».

Amb la correcció de multiplicar per 1.08 l'alçada de cada dona, Galton va tractar les dades com si tingués només una població [41]. La taula 1 il·lustra algunes dades. La família 18 té 3 fills i cap filla. La 102 té 3 fills i 3 filles i la 198 té 4 fills i 1 filla.

Família	Pare	Mare	Gènere	Fill/Filla
18	78	73	1	66.5-64.5-64
102	69	66	1	70-68.5-68
102	69	66	2	65-63-62.5
198	65.5	60	1	68-68-67-67
198	65.5	60	2	62

TAULA 1: Exemple de dades sobre alçades (en polzades) de pares i fills obtingudes per Galton el 1886.

El coeficient de correlació és el quocient $r = S_{xy}/(S_x S_y)$, on S_{xy} és la covariància i S_x, S_y són les desviacions típiques de les variables X, Y . El valor absolut d'aquest coeficient és invariant per transformacions lineals de les variables i satisfà $|r| \leq 1$. Va ser proposat el 1895 per K. Pearson, amb la qual cosa millorà la invenció de Galton.²

Galton va obtenir $r = 0.50$, que indicava una certa relació entre pares i fills respecte a l'alçada. Aleshores va representar les dades en un diagrama (figura 1), va dibuixar a ull una recta d'ajust i va afirmar que l'alçada d'un fill era aproximadament 2/3 de l'alçada dels pares (variable mid-parents).

La recta $y = a + bx$ que millor s'ajusta a les dades, en el sentit dels mínims quadrats, on

$$a \text{ i } b \text{ són tals que } \sum_{i=1}^n (y_i - a - bx_i)^2 \text{ és mínima,}$$

és la recta de regressió, que és

$$y = \bar{y} + b(x - \bar{x}), \quad \text{on } b = S_{xy}/S_x^2, \quad (1)$$

essent $\bar{x}$ i $\bar{y}$ les mitjanes de les variables i b el coeficient de regressió o pendent de la recta de regressió. Aquesta recta descriu el model abans esmentat, és a dir, la funció $f(x)$ és lineal. A la secció següent parlarem de l'origen del terme *regressió*.

² El descobriment del coeficient de correlació és descrit per Galton (que tenia una bona visió matemàtica) a *Memories of My Life*. El coeficient r apareix implícitament en treballs de Gauss i Bravais sobre la distribució normal bivariant. K. Pearson, a qui hem d'atribuir la formulació actual, va proposar (com també ho proposaria Hilbert) una teoria de la relativitat. La seva definició de *coeficient de correlació* es basa en l'estructura d'espai vectorial de les variables aleatòries. Tot i la importància del coeficient de correlació, l'estadístic Sigmund Schott (1928) va afirmar: «És molt lamentable que la complicada fonamentació matemàtica del càlcul de la correlació i la seva prolíxa resolució aritmètica exclouin per sempre la possibilitat d'aplicar en algun cas concret aquest notable procediment».

	Galton	<i>n</i> = 934	Pearson	<i>n</i> = 1078
	Mid-parents	Fills	Pare	Fill
Mitjana	69.20	69.23	67.20	68.16
Desviació típica	1.80	2.58	2.72	2.74
Correlació/regressió	<i>r</i> = 0.50	<i>b</i> = 0.71	<i>r</i> = 0.51	<i>b</i> = 0.52

TAULA 2: Mitjanes, desviacions típiques i coeficient de correlació i regressió que resulten de les dades originals de Galton i de Pearson-Lee. Dades en polzades; 67.2 polzades són 170.7 cm., 68.16 polzades són 173 cm.

EXEMPLE. La taula 2 conté alguns resultats estadístics per a les dades de Galton (*n* = 934) i de Pearson-Lee (*n* = 1078), publicades l'any 1903 [53] i que comentarem a la secció 8.

3 La regressió a la mitjana

El coeficient de regressió *b* = 0.71 (taula 2) va donar peu a Galton per afirmar que hi havia una «regressió a la mitjana». Altrament dit, en predir *y* donat *x*, s'esperava que un pare alt, l'alçada del qual era superior a la mitjana, tindria fills també alts, però no tant com el pare. És a dir, el fill (o filla) estaria més a prop de la mitjana que el seu progenitor. De manera similar, un pare baix tindria tendència a tenir fills baixos però no tant com el pare.

Aquest notable fet, que dona nom al terme *model de regressió* (tot i que *regressió* és epistemològicament incorrecte), sembla paradoxal. Però no ho és en sentit biològic. Si els pares alts tinguessin fills més alts, a la generació següent els fills encara ho serien més, i al cap de moltes generacions hi hauria gegants. De la mateixa manera, si els pares baixos tinguessin fills més baixos, a la llarga hi hauria nans. Tampoc hi ha paradoxa estadística, atès que si ajuntem tots els fills, la mitjana es recupera i si hi ha regressió dins d'una família, no n'hi ha a la població.

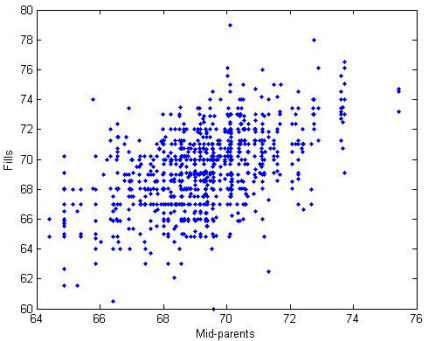


FIGURA 1: Diagrama que representa l'alçada combinada de pare i mare sobre la dels fills i filles, per a 934 observacions obtingudes per Galton.

4 Correlació múltiple

El coeficient de correlació en la forma que ara el coneixem va ser elaborat per K. Pearson. Podem generalitzar la correlació simple r , vàlida per a dues variables, definint la correlació múltiple R . Donada una variable resposta Y i p variables explicatives $X_1, \dots, X_p$, el coeficient R es defineix com la correlació simple entre Y i $\hat{Y}$, essent $\hat{Y} = b_0 + b_1X_1 + \dots + b_pX_p$ la combinació lineal que millor s'ajusta a Y en el sentit que els coeficients b_i verifiquen

$$E(Y - \hat{Y})^2 = E(Y - b_0 - b_1X_1 - \dots - b_pX_p)^2 \text{ és mínima.}$$

Es calcula mitjançant

$$R^2 = \mathbf{r}'\mathbf{R}^{-1}\mathbf{r},$$

on $\mathbf{R}$ és la matriu $p \times p$ de correlacions entre les variables X , i $\mathbf{r}$ és el vector amb les correlacions de la variable Y amb cadascuna de les X .

EXEMPLE. Per a les dades de Galton, si Y és l'alçada del pare (o de la mare), X_1 i X_2 són les alçades de fill i filla, obtenim:

$$\mathbf{r}_{\text{pare}} = \begin{bmatrix} 0.5457 \\ 0.5067 \end{bmatrix}, \quad \mathbf{r}_{\text{mare}} = \begin{bmatrix} 0.3650 \\ 0.3808 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & 0.5419 \\ 0.5419 & 1 \end{bmatrix}.$$

Els coeficients de correlació múltiple del pare sobre fill/filla i de la mare sobre fill/filla, són:

$$R_{\text{pare}} = 0.6067, \quad R_{\text{mare}} = 0.4249.$$

Curiosament, respecte a l'alçada, hi ha més influència del pare.

5 Un sofisma sobre la correlació

Plantegem una paradoxa aparent o sofisma sobre la correlació. Suposem que X i Y són dues variables correlacionades sobre la mateixa població, amb variàncies σ_x^2 i σ_y^2 , respectivament. Denotem per σ_{xy} la covariància i per ρ_{xy} el coeficient de correlació. En tots els casos suposem que els paràmetres són poblacionals.

Considerem ara una mostra aleatòria simple de X de mida n . És a dir, considerem $X_1, \dots, X_n$ variables independents, igualment distribuïdes amb la mateixa distribució que X .

La variància de la mitjana mostral $\bar{X}_n = (X_1 + \dots + X_n)/n$ és σ_x^2/n i la covariància entre $\bar{X}_n$ i Y és

$$\text{cov}(\bar{X}_n, Y) = \frac{1}{n} \text{cov}(X_1 + \dots + X_n, Y) = \frac{1}{n} n \sigma_{xy} = \sigma_{xy}.$$

Per tant, atès que $\rho_{xy} = \sigma_{xy}/(\sigma_x \sigma_y)$, el coeficient de correlació entre $\bar{X}_n$ i Y és

$$\text{cor}(\bar{X}_n, Y) = \frac{\sigma_{xy}}{(\sigma_x/\sqrt{n})\sigma_y} = \sqrt{n}\rho_{xy}.$$

Ens trobem que si n és prou gran, podríem tenir $\sqrt{n}\rho_{xy} > 1$. Per exemple, si $\rho_{xy} = 0.5$ i $n > 4$ aleshores $\sqrt{n}\rho_{xy} > 1$. Això contradiu la propietat que el coeficient de correlació mai supera el valor 1. Tornarem a comentar aquesta aparent irregularitat més endavant.³

6 El tot pot ser més gran que la suma de les parts

Si les variables X estan incorrelacionades dues a dues, és a dir, la matriu de correlacions és $\mathbf{R} = \mathbf{I}$ (identitat), aleshores $R^2 = r_1^2 + \dots + r_p^2$, essent $r_1, \dots, r_p$ les correlacions simples de Y amb cadascuna de les variables X . En general les X estan correlacionades i hom esperaria que $R^2 < r_1^2 + \dots + r_p^2$, atès que no hi ha redundància en la suma de quadrats però sí en R^2 . No obstant això, hi ha situacions reals on

$$R^2 > r_1^2 + \dots + r_p^2. \quad (2)$$

El cas $p = 2$ va ser estudiat per Hamilton i Routledge [40, 56]. En el cas general, Cuadras [9] va demostrar que la desigualtat (2), que podem escriure com a $R^2 = \mathbf{r}'\mathbf{R}^{-1}\mathbf{r} > \mathbf{r}'\mathbf{r}$, és equivalent a

$$\sum_{i=1}^p r_{Z_i}^2 (1 - \lambda_i) > 0,$$

essent r_{Z_i} , $i = 1, \dots, p$, les correlacions simples entre la variable resposta Y i les components principals $Z_1, \dots, Z_p$, i $\lambda_1, \dots, \lambda_p$, els valors propis de $\mathbf{R}$. Recordem que les components principals Z_i són les combinacions lineals de les variables X amb variància màxima condicionada a que siguin incorrelacionades. Aquestes variàncies són precisament els valors propis de $\mathbf{R}$.

Atès que els primers valors propis són més grans que 1 i els últims, més petits, tenim que (2) es verifica si Y té una correlació alta amb les components principals de menor variància (figura 2). Per tant, és erroni creure que variables correlacionades són redundants.

Molt relacionada amb (2) és la possibilitat d'augmentar les correlacions simples però que, sorprenentment, disminueixi la correlació múltiple. Considerem la matriu de correlacions $\mathbf{R}$ de quatre variables X i dos vectors $\mathbf{r}_1, \mathbf{r}_2$ que contenen les correlacions amb les X de dues variables dependents Y_1, Y_2 .

$$\mathbf{R} = \begin{bmatrix} 1 & 0.31 & 0.40 & 0.52 \\ 0.31 & 1 & 0.52 & 0.40 \\ 0.40 & 0.52 & 1 & 0.31 \\ 0.2 & 0.40 & 0.31 & 1 \end{bmatrix}, \quad \mathbf{r}_1 = \begin{bmatrix} 0.60 \\ 0.50 \\ 0.40 \\ 0.30 \end{bmatrix}, \quad \mathbf{r}_2 = \begin{bmatrix} 0.59 \\ 0.49 \\ 0.10 \\ 0.10 \end{bmatrix}.$$

³ Alguns estadístics destacats van necessitar hores, fins i tot dies, a trobar una explicació a aquest desconcertant sofisma sobre la correlació.

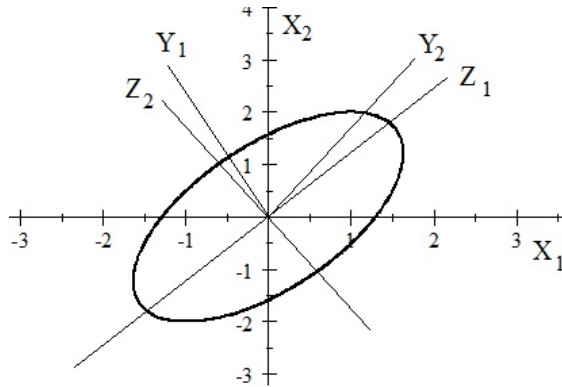


FIGURA 2: La variable dependent Y_1 segueix la direcció de la segona component principal Z_2 , perpendicular a la direcció principal de les dades, representada per la primera component principal Z_1 .

Com que les correlacions simples són més grans per a Y_1 , s'esperaria que també ho sigui la correlació múltiple. No obstant això, es verifica

$$R_1^2 = \mathbf{r}_1' \mathbf{R}^{-1} \mathbf{r}_1 = 0.483 < R_2^2 = \mathbf{r}_2' \mathbf{R}^{-1} \mathbf{r}_2 = 0.688.$$

L'explicació d'aquesta anomalia es la següent: $\mathbf{r}_1$ és un vector amb direcció (mesurada per l'angle) pròxima als últims vectors propis de $\mathbf{R}$, mentre que $\mathbf{r}_2$ segueix una direcció més semblant als primers vectors propis. Altrament dit, Y_1 estaria més influïda per les dues últimes components principals, mentre que Y_2 estaria més relacionada amb les dues primeres. Vegeu la figura 2, amb només dues variables X , però igualment il·lustrativa. La conseqüència de tot això és que el costum de descartar variables X poc correlacionades amb Y podria ser inadequat. Per a més detalls, vegeu [10, 59].

7 Fal·làcia ecològica

Examinem la figura 3, que representa el risc de càncer respecte al consum de calories per a diferents països. Considerant només les mitjanes de cada país (cercle negre), veiem que el risc augmenta amb el consum de calories. Però examinant cada país per separat (cercles blancs agrupats), veiem que el risc disminueix si augmenta l'alimentació. La correlació és, de fet, negativa, però si agrupem poblacions, apareix una falsa correlació positiva. L'anomenada *fal·làcia ecològica* es deu al fet d'ajuntar poblacions diferents i aplicar la correlació, ignorant que aquest coeficient només es pot utilitzar a cada població per separat.

EXEMPLES. Per a una malaltia hereditària, es va detectar una certa correlació positiva entre l'edat X d'aparició en el pare i l'edat Y d'aparició en el fill [42].

En realitat X i Y són independents, i l'edat d'aparició en el fill no té res a veure amb la del pare. La falsa correlació era deguda a que hi ha dos gens A i B , que es presenten amb probabilitats p_A i p_B i que causen la mateixa malaltia. És a dir, les edats observades provenen de dues poblacions diferents. Si X i Y (igualment distribuïdes condicionades a cada gen) tenen mitjanes $E(X|A) = E(Y|A) = \mu_A$ i $E(X|B) = E(Y|B) = \mu_B$ i anàlogament variàncies σ_A^2 i σ_B^2 , depenent del gen causant A o B , les esperances matemàtiques de les variables observades X i Y , els seus quadrats i el producte XY són:

$$\begin{aligned} E(X) &= E(Y) = p_A\mu_A + p_B\mu_B, \\ E(X^2) &= E(Y^2) = p_A(\mu_A^2 + \sigma_A^2) + p_B(\mu_B^2 + \sigma_B^2), \\ E(XY) &= p_A\mu_A^2 + p_B\mu_B^2. \end{aligned}$$

Operant, obtenim les variàncies:

$$\text{var}(X) = \text{var}(Y) = p_A\sigma_A^2 + p_B\sigma_B^2 + p_Ap_B(\mu_A - \mu_B)^2.$$

De manera similar, la covariància és $\text{cov}(X, Y) = p_Ap_B(\mu_A - \mu_B)^2$. El coeficient de correlació és, doncs,

$$\rho_{xy} = \left[1 + \frac{p_A\sigma_A^2 + p_B\sigma_B^2}{p_Ap_B(\mu_A - \mu_B)^2} \right]^{-1}.$$

Concretament, si $p_A = p_B = 1/2$, $\sigma_A^2 = \sigma_B^2 = \sigma^2$, i la diferència $|\mu_A - \mu_B|$ entre les dues mitjanes val 2σ , aleshores obtenim $\rho_{xy} = 0.5$, tot i que les variables X i Y són independents a cada població.

Un altre exemple apareix en el polèmic llibre *The Bell Curve*, de Berrstein i Murray, publicat el 1994. En aquest tractat s'afirma que els blancs són més intel·ligents que els negres. En conseqüència, argumenten els autors, els blancs també seran superiors en una altra habilitat mesurada per una variable H , que es pugui quantificar i que estigui correlacionada positivament amb la intel·ligència I . Per tant, a fi d'escollir entre un blanc i un negre per encarregar-se d'una feina (pilotar una aeronau, per exemple), que es faria millor si l'habilitat mesurada mitjançant H és alta, si ambdós candidats tenen el mateix coeficient d'intel·ligència, aleshores el valor de H en el blanc serà més alt que en el negre. Per tant, seria convenient proposar el candidat blanc. L'argumentació, fruit de barrejar dues poblacions, és falsa, com va denunciar Kaplan [46]. És més, per a una mateixa I , el negre tindria un valor H més alt que el blanc, com podem apreciar a la figura 4, que conté les dues rectes de regressió, i per tant el negre estaria més capacitat per a la feina. De manera més precisa, suposem la mateixa habilitat mitjana de H en blancs i negres, $E(H|B) = E(H|N) = 100$, però una certa superioritat en intel·ligència: $E(I|B) = 100$, $E(I|N) = 90$. Si les desviacions típiques són iguals a 15 i el coeficient de correlació entre H i I és 0.6, aleshores les prediccions (utilitzant les rectes de regressió) indicarien la superioritat del negre sobre el blanc:

$$\begin{aligned} \text{Blanc} \quad H &= 100 + 0.6 \times (105 - 100) = 103, \\ \text{Negre} \quad H &= 100 + 0.6 \times (105 - 90) = 109. \end{aligned}$$

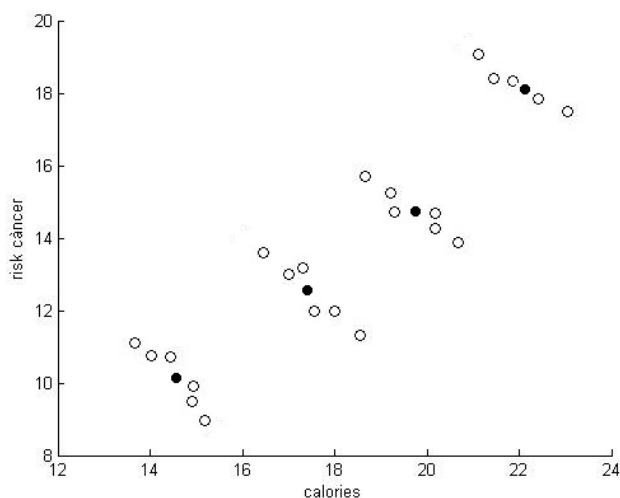


FIGURA 3: Fallàcia ecològica. La correlació considerant només les mitjanes de les poblacions (cercles negres) és positiva. Però la correlació dintre de cada població (cercles blancs agrupats) és negativa.

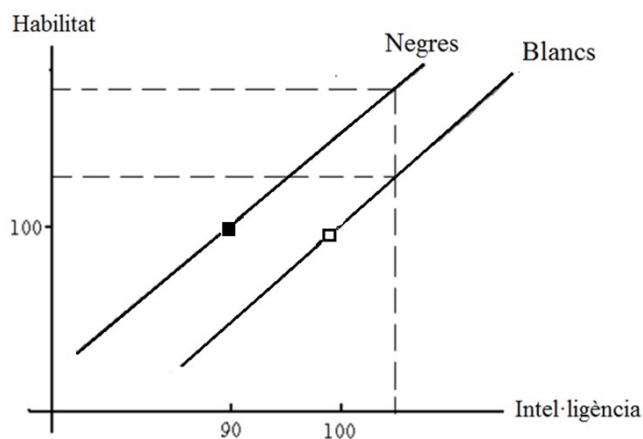


FIGURA 4: Suposant que la intel·ligència sigui superior en els blancs, si un negre té el mateix coeficient d'intel·ligència que un blanc, aleshores el superaria en una altra habilitat correlacionada amb la intel·ligència. Equivocadament, a *The Bell Curve* s'afirma el contrari.

8 Més sobre la correlació

8.1 Raó de correlació

Quan K. Pearson va definir el coeficient de correlació en la seva forma actual, de seguida deuria veure que les dades de Galton no responien perfectament a la definició. En efecte, per a cada valor de X (variable mid-parent), tenim 1, 2 o més valors de Y , perquè una parella pot tenir més d'un fill. A més, si bé podem suposar que les alçades dels pares són independents, les dels fills estan correlacionades dintre de cada família. De fet, el 1903, K. Pearson [53] va efectuar els càlculs amb noves dades i va considerar $n = 1078$ famílies amb només un o dos fills. Va obtenir un coeficient de correlació $r = 0.51$ entre pare i fill varó (vegeu la taula 2).

Podem estudiar les dades de Galton considerant un model lineal

$$y_{ij} = \mu + \beta_i x_i + e_{ij}, \quad i = 1, \dots, k, j = 1, \dots, n_i, \quad (3)$$

on y_{ij} és l'alçada del fill j de la família i , μ és una mitjana general, β_i és el coeficient de regressió per a la família i corresponent a un mid-parent x_i . El terme e_{ij} és la desviació del model, i reflecteix el fet que els fills d'una mateixa família no tenen la mateixa estatura. Sigui $\hat{\mu}_H$ l'estimació per mínims quadrats suposant que totes les β_i són 0. Aquesta estimació és la mitjana general de les y_{ij} . Denotem per $R_H^2 = \sum_{i,j} (y_{ij} - \hat{\mu}_H)^2$ la suma de quadrats residual. Siguin $\hat{\mu}, \hat{\beta}_i, i = 1, \dots, k$, les estimacions sense imposar restriccions i sigui $R_0^2 = \sum_{i,j} (y_{ij} - \hat{\mu} + \hat{\beta}_i x_i)^2$ la suma de quadrats residual. Tenim que $R_0^2 \leq R_H^2$, atès que R_H^2 és un mínim restringit. Aleshores

$$r^2 = 1 - \frac{R_0^2}{R_H^2} \quad (4)$$

és un coeficient que indica el grau de relació lineal entre X i Y . Veiem fàcilment que si el model és perfecte (no hi ha desviació e_{ij}), aleshores $r^2 = 1$. Per a les dades de Galton, r coincideix amb el coeficient de correlació, i per tant $r = 0.50$.

Plantegem ara un model general

$$y_{ij} = g(x_i) + e_{ij}, \quad i = 1, \dots, k, j = 1, \dots, n_i,$$

on g és una funció possiblement no lineal. Si suposem $g = 0$, l'estimació de μ i la suma de quadrats residual R_H^2 són les mateixes d'abans. L'estimació no restringida de $g(x_i)$ és la mitjana corresponent al grup i . Aplicant la mateixa mesura d'ajust, obtenim

$$\hat{\eta}^2 = 1 - \sum_{i=1}^k \frac{n_i S_i^2}{n S_Y^2},$$

on $n = n_1 + \dots + n_k$, S_i^2 és la variància per a la família i , S_Y^2 és la variància global. Es compleix $\hat{\eta}^2 \geq r^2$. Si no hi ha desviació del model no lineal, és a dir, $S_i^2 = 0$, aleshores $\hat{\eta}^2 = 1$. En les aplicacions cal comparar $\hat{\eta}$ amb r .

Considerem la coneguda descomposició de la variabilitat en l'anàlisi de la variància, ANOVA (vegeu [57]),

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^k n_i (\bar{y}_{i.} - \bar{y})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2, \quad (5)$$

essent $\bar{y}$ la mitjana general, $\bar{y}_{i.}$ la mitjana del grup i . Escriurem $Q_T = Q_E + Q_D$, on les quantitats Q_T , Q_E , Q_D són les sumes de quadrats totals, entre grups i dintre de grups. Aleshores també podem expressar la raó de correlació per

$$\hat{\eta}^2 = 1 - \frac{Q_D}{Q_T} = \frac{Q_E}{Q_T}.$$

EXEMPLE. Per a les dades de Galton obtenim $\hat{\eta} = 0.58$, un indicatiu clar de relació no lineal, ja que $\hat{\eta} = 0.58$ és significativament més gran que $r = 0.50$. Possiblement sigui degut a les repeticions dels valors dels pares que tenen més d'un fill. Però Pearson va comentar que era degut a una manera amateur d'obtenir les dades [58]. De fet, Pearson va treballar amb 1078 famílies i unes dades més ben observades i millor ajustades als conceptes de *correlació* i *regressió*. Va obtenir $\hat{\eta} = 0.52$, fet que indica una bona relació lineal, ja que la diferència entre $\hat{\eta} = 0.52$ i $r = 0.51$ no és significativa.

Si dividim per n els termes de (5) obtenim

$$\frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^k \frac{n_i}{n} (\bar{y}_{i.} - \bar{y})^2 + \sum_{i=1}^k \frac{n_i}{n} \frac{1}{n_i} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2, \quad (6)$$

que seria una versió estadística de la descomposició de la variància d'una variable aleatòria. És a dir, donat un parell de variables (X, Y) , la descomposició (6) és la versió ANOVA de

$$\text{var}(Y) = \text{var}[E[Y|X]] + E[\text{var}[Y|X]], \quad (7)$$

on $Y|X$ representa la variable Y condicionada a X . La corba de regressió de la mitjana és la millor corba que ajusta Y en funció de X , és a dir, $y = E[Y|X = x]$. El grau de concentració de les observacions (x, y) de (X, Y) al llarg d'aquesta corba és

$$\eta^2 = 1 - \frac{E[\text{var}[Y|X]]}{\text{var}(Y)}.$$

8.2 Correlació intraclàssica

Quina és la correlació entre germans per a una característica física com l'alçada? La resposta no s'obté mitjançant el coeficient de correlació ordinari, atès que no podem disposar dels valors dels germans en parelles ordenades (x, y) .

Suposem el model

$$y_{ij} = \mu + A_i + e_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i,$$

on A_i és una mena de variable aleatòria tal que $E(A_i) = 0$ i $\text{var}(A_i) = \sigma_A^2$ per a tota i . Suposem les condicions usals d'independència completa entre les A_i , e_{ij} i també que $\text{var}(e_{ij}) = \sigma^2$.

Es defineix la correlació intraclàssica ρ_I com la que hi ha entre dues observacions y_{ij} i $y_{ij'}$ dintre del mateix grup (o família). Es demostra fàcilment [57] que $\rho_I = \sigma_A^2 / (\sigma_A^2 + \sigma^2)$.

Siguin Q_A i Q_R les sumes de quadrats entre grups i dintre de grups (quantitats que abans hem indicat per Q_E i Q_D). Dividint pels graus de llibertat corresponents, considerem les mitjanes

$$\bar{Q}_A = \frac{Q_A}{k-1}, \quad \bar{Q}_R = \frac{Q_R}{n-k},$$

essent $n = n_1 + \dots + n_k$. Segons [33], una estimació de ρ_I és

$$\hat{\rho}_I = \frac{\bar{Q}_A - \bar{Q}_R}{\bar{Q}_A + (n_0 - 1)\bar{Q}_R}, \quad (8)$$

on $n_0 = (n - \sum_{i=1}^k n_i^2 / n) / (k - 1)$.

Relacionem ara $\hat{\rho}_I$ amb la raó de correlació $\hat{\eta}^2$. Amb la present notació, tenint en compte que $Q_A / Q_R = \hat{\eta}^2 / (1 - \hat{\eta}^2)$, és fàcil veure que la correlació intraclàssica és

$$\hat{\rho}_I = \frac{(n-1)\hat{\eta}^2 - (k-1)}{(n - kn_0 + n_0 - 1)\hat{\eta}^2 + (n_0 - 1)(k-1)}.$$

A diferència de $\hat{\eta}$, la correlació intraclàssica $\hat{\rho}_I$ és un coeficient estadístic del qual no hi ha versió probabilística.

EXEMPLE. Per a les dades (no balancejades amb $n_i > 1$) de Galton, aplicant (8), obtenim $n_0 = 5.23$ i $\hat{\rho}_I = 0.38$, que indica una certa correlació intraclàssica entre les alçades de germans i germanes.

8.3 La perspectiva bayesiana

És oportú comentar aquí la condició fonamental de la metodologia ANOVA: la independència estocàstica i la igualtat de variàncies del terme d'error. Així, en el model ANOVA d'un sol factor,

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i,$$

suposem la distribució normal $N_{n_i}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$ per a cada successió e_{ij} , $j = 1, \dots, n_i$. Aquesta hipòtesi és discordant amb l'enfocament típicament bayesià, que suposa les mostres amb distribució conjunta «intercanviable». Altrament dit, la distribució de cadascuna de les n_i observacions y_{ij} és (condicionalment a cada grup) la mateixa, però la distribució conjunta és simètrica.

Per alleugerir la notació, denotem n_i per n i els n_i errors per $x_1, \dots, x_n$. Aleshores:

$$(x_1, \dots, x_n) \sim (x_{j_1}, \dots, x_{j_n}),$$

on $(j_1, \dots, j_n)$ és una permutació de $(1, \dots, n)$. El símbol $\sim$ significa «mateixa distribució que».

Segons el teorema de Bruno de Finetti, la densitat conjunta pren la forma [5]

$$p(x_1, \dots, x_n) = \int_{\mathbf{I}_\alpha} \prod_{j=1}^n f(x_j | \alpha) \pi(\alpha) d\alpha,$$

on $f(x_j | \alpha)$ és una mena de model estadístic on el paràmetre α prové de l'observació d'una variable latent amb suport l'interval $\mathbf{I}_\alpha$ i densitat de probabilitat $\pi(\alpha)$. La distribució dels valors $x_1, \dots, x_n$ són independents si els condicionem a un valor fix de α . Suposant que la variància és constant σ^2 , indicant la mitjana condicionada per $\mu(\alpha) = \int_{\mathbb{R}} x f(x | \alpha) dx$, la mitjana global d'un error x és $a = \int_{\mathbf{I}_\alpha} \mu(\alpha) \pi(\alpha) d\alpha$. A més, si posem $A = \int_{\mathbf{I}_\alpha} \mu(\alpha)^2 \pi(\alpha) d\alpha$, el valor esperat del producte de dos errors x i x' és precisament

$$\int_{\mathbf{I}_\alpha} \int_{\mathbb{R}^2} x x' f(x | \alpha) f(x' | \alpha) \pi(\alpha) d\alpha dx dx' = A.$$

Com que la variància és $\sigma^2 + A - a^2$, fàcilment es demostra que la correlació entre dos errors és

$$\rho = \frac{A - a^2}{\sigma^2 + A - a^2}.$$

Si suposem α fix amb probabilitat 1 (enfocament clàssic), obtenim $A = a^2$ i $\rho = 0$.

Aplicant aquest enfocament bayesià, atès que no coneixem $\pi(\alpha)$, un model raonable seria suposar que $(x_1, \dots, x_n)$ segueix la distribució normal simètrica $N_n(\mathbf{0}, \Sigma)$, i la matriu de covariàncies és

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}.$$

Aquesta distribució és intercanviable i pressuposa (dintre d'un grup) una mateixa correlació ρ per a cada parella d'observacions (o errors) diferents. En general, un bayesià postularia $\rho \neq 0$ i un clàssic consideraria la restricció $\rho = 0$, que en el cas de normalitat implicaria independència completa entre les observacions. És correcte el model restringit? Afortunadament per a l'estadística clàssica, Wilks [60] va demostrar que el test F de Fisher-Snedecor és també vàlid amb dades normals equicorrelacionades, és a dir, en el cas que sigui $\rho \neq 0$.

Ara és el moment de resoldre el sofisme de la secció 5, on la correlació entre $\bar{X}_n$ i Y semblava que podia ser més gran que 1. El parany consistia a

prendre $X_1, \dots, X_n$ independents. L'acció d'obtenir una mostra aleatòria simple, tan usual en estadística, és incorrecta si X està correlacionada amb Y . Hem d'admetre, doncs, que $X_1, \dots, X_n$ no són independents.

Dit això, hi ha maneres diferents de veure que $|\text{cor}(\bar{X}_n, Y)| \leq 1$. Per exemple, si suposem el model d'equicorrelació, que implica que les n variables X són intercanviables i no pas independents, aleshores $\text{cov}(X_i, X_j) = \sigma_X^2 \rho$, $i \neq j$. Per tant,

$$\text{var}(\bar{X}_n) = [n\sigma_X^2 + n(n-1)\sigma_X^2\rho]/n^2,$$

i el coeficient de correlació és

$$\text{cor}(\bar{X}_n, Y) = \frac{\sqrt{n}\rho_{xy}}{\sqrt{1 + (n-1)\rho}},$$

que mai pot superar el valor 1.

Part II: Estadística multivariant

Els mètodes de les dues seccions següents es poden trobar amb detall a [8, 15].

9 Anàlisi de correlació canònica

La correlació múltiple entre Y i $X_1, \dots, X_p$ és la màxima correlació entre Y i una combinació lineal de $X_1, \dots, X_p$. La generalització, proposada per Hotelling,⁴ considera dos vectors aleatoris $\mathbf{X} = (X_1, \dots, X_p)$ i $\mathbf{Y} = (Y_1, \dots, Y_q)$, i determina les combinacions lineals [44]

$$U = \mathbf{X}\mathbf{a} = a_1X_1 + \dots + a_pX_p, \quad V = \mathbf{Y}\mathbf{b} = b_1Y_1 + \dots + b_qY_q,$$

de tal manera que la correlació entre U i V sigui màxima, on $\mathbf{a} = (a_1, \dots, a_p)'$ i $\mathbf{b} = (b_1, \dots, b_q)'$ són dos vectors.

Indiquem per $\mathbf{S}_{11}$ i $\mathbf{S}_{22}$ les matrius de covariàncies mostrals de $\mathbf{X}$ i $\mathbf{Y}$, respectivament (el cas poblacional té un tractament molt semblant). Sigui $\mathbf{S}_{12}$ la matriu $p \times q$ amb les covariàncies de les variables $\mathbf{X}$ amb les variables $\mathbf{Y}$. Tenim aleshores la supermatriu $\mathbf{S} = (\mathbf{S}_{ij})$,

	$\mathbf{X}$	$\mathbf{Y}$
$\mathbf{X}$	$\mathbf{S}_{11}$	$\mathbf{S}_{12}$
$\mathbf{Y}$	$\mathbf{S}_{21}$	$\mathbf{S}_{22}$

on $\mathbf{S}_{21} = \mathbf{S}_{12}'$. Si $\mathbf{S}_{11} = (s_{ij})$, aleshores $\text{var}(U) = \sum_{i,j=1}^p a_i a_j s_{ij} = \mathbf{a}'\mathbf{S}_{11}\mathbf{a}$, i anàlogament $\text{var}(V) = \mathbf{b}'\mathbf{S}_{22}\mathbf{b}$ i $\text{cov}(U, V) = \mathbf{a}'\mathbf{S}_{12}\mathbf{b}$.

⁴ H. Hotelling va introduir l'anàlisi de correlació canònica el 1936, en un intent de relacionar aptituds físiques i mentals. El mètode tindria una gran influència en anàlisi factorial, anàlisi de correspondències, anàlisi discriminant, anàlisi canònica de poblacions i altres mètodes multivariants.

Podem suposar que $\text{var}(U) = \text{var}(V) = 1$. Aleshores el problema equival a:

$$\text{maximitzar } \mathbf{a}'\mathbf{S}_{12}\mathbf{b} \quad \text{restringit a } \mathbf{a}'\mathbf{S}_{11}\mathbf{a} = 1, \quad \mathbf{b}'\mathbf{S}_{22}\mathbf{b} = 1.$$

La solució no és única i els vectors de coeficients $\mathbf{a}$ i $\mathbf{b}$ que compleixen aquesta condició són els vectors canònics. La màxima correlació entre U i V , combinacions lineals de $\mathbf{X}$ i $\mathbf{Y}$, rep el nom de primera correlació canònica r_1 . Clarament, $r_1 = 0$ si $\mathbf{X}$ és independent de $\mathbf{Y}$, i $r_1 = 1$ si hi ha una relació lineal entre els dos vectors. Per tant, r_1 té un paper destacat en la mesura de l'associació entre $\mathbf{X}$ i $\mathbf{Y}$.

Si $U_1 = \mathbf{X}\mathbf{a}_1$, $V_1 = \mathbf{Y}\mathbf{b}_1$ és la primera parella de variables canòniques, definim $U_2 = \mathbf{X}\mathbf{a}_2$, $V_2 = \mathbf{Y}\mathbf{b}_2$ com la parella de variables (també combinacions lineals de $\mathbf{X}$ i $\mathbf{Y}$) incorrelacionades amb U_1 i V_1 . Aleshores $r_2 = \text{cor}(U_2, V_2)$ és la segona correlació canònica. Anàlogament obtenim la tercera i següents variables i correlacions canòniques, que satisfan les condicions d'optimització esmentades a dalt.

Podem formular una expressió conjunta per als vectors canònics $\mathbf{a}_i$ i $\mathbf{b}_i$ utilitzant la descomposició singular d'una matriu. Suposant que $p \geq q$, sigui $m = \min\{p, q\} = q$ i considerem la matriu $p \times q$

$$\mathbf{Q} = \mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2}. \quad (9)$$

Calculem la descomposició singular $\mathbf{Q} = \mathbf{U}\mathbf{D}_s\mathbf{V}'$, on $\mathbf{U}$ és una matriu $p \times q$ amb columnes ortonormals, $\mathbf{V}$ és una matriu $q \times q$ ortogonal i $\mathbf{D}_s$ és una matriu diagonal amb els valors singulars de $\mathbf{Q}$. Altrament dit, $\mathbf{U}'\mathbf{U} = \mathbf{I}_q$, $\mathbf{V}'\mathbf{V} = \mathbf{I}_q$, $\mathbf{D}_s = \text{diag}(s_1, \dots, s_m)$. Aleshores els vectors canònics i les correlacions canòniques són

$$\mathbf{a}_i = \mathbf{S}_{11}^{-1/2}\mathbf{u}_i, \quad \mathbf{b}_i = \mathbf{S}_{22}^{-1/2}\mathbf{v}_i, \quad r_i = s_i.$$

Formalment: $U_i = \mathbf{X}\mathbf{a}_i$, $V_i = \mathbf{Y}\mathbf{b}_i$, $r_i = s_i = \text{cor}(U_i, V_i)$, $i = 1, \dots, m$, és a dir, les correlacions canòniques són els valors singulars de $\mathbf{Q}$.

Cada parella de variables canòniques té correlació màxima, i manté la condició d'ortogonalitat, és a dir,

$$\text{cor}(U_i, V_i) = r_i, \quad \text{cor}(U_i, U_j) = \text{cor}(V_i, V_j) = \text{cor}(U_i, V_j) = 0 \quad \text{si } i \neq j.$$

A més es compleix la relació recursiva següent entre els vectors canònics:

$$\mathbf{a}_i = s_i^{-1}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{b}_i, \quad \mathbf{b}_i = s_i^{-1}\mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{a}_i.$$

EXEMPLE. Per a les dades de Galton, podem considerar les quatre variables

$$\begin{aligned} X_1 &= \text{alçada del pare}, & Y_1 &= \text{alçada del fill}, \\ X_2 &= \text{alçada de la mare}, & Y_2 &= \text{alçada de la filla}. \end{aligned}$$

Volem relacionar (Y_1, Y_2) amb (X_1, X_2) . Atès que necessitem quaternes completes, hem seleccionat famílies amb almenys un fill i una filla, prenent la mitjana

quan hi ha més d'un descendent. Obtenim $n = 150$ i les matrius de covariàncies:

$$\mathbf{S}_{11} = \begin{pmatrix} 6.9654 & 0.4516 \\ 0.4516 & 5.3950 \end{pmatrix}, \quad \mathbf{S}_{12} = \begin{pmatrix} 3.1759 & 2.7013 \\ 1.8697 & 1.7857 \end{pmatrix},$$

$$\mathbf{S}_{21} = \begin{pmatrix} 3.1759 & 1.8697 \\ 2.7013 & 1.7857 \end{pmatrix}, \quad \mathbf{S}_{22} = \begin{pmatrix} 4.8624 & 2.4138 \\ 2.4138 & 4.0804 \end{pmatrix}.$$

Els dos vectors canònics i els valors singulars són:

$$\mathbf{a}_1 = (0.3050; 0.2312)', \quad \mathbf{b}_1 = (0.2742; 0.2642)', \quad s_1 = 0.71,$$

$$\mathbf{a}_2 = (0.2265; -0.3646)', \quad \mathbf{b}_2 = (0.4648; -0.5265)', \quad s_2 = 0.038.$$

La primera correlació canònica és $r_1 = 0.71$ i les primeres variables canòniques són:

$$U_1 = 0.3050X_1 + 0.2312X_2, \quad V_1 = 0.2742Y_1 + 0.2642Y_2.$$

La correlació ha augmentat de 0.50 a 0.71. De fet $r_1 = 0.71$ és una mesura global d'associació entre les estatures de pares i fills.

Resulta interessant constatar que el càlcul de la correlació entre $U_1 = 0.3050X_1 + 0.2312X_2$ i la variable mid-parent $(X_1 + 1.08X_2)/2$ (però considerant el conjunt complet de les 934 dades inicials) dona $r = 0.995$, un valor molt elevat. Veiem que la variable mid-parent emprada per Galton a fi de reduir diferències entre pares i mares és gairebé equivalent a la primera variable canònica.

Definim ara una mesura global d'associació entre els vectors aleatoris $\mathbf{X} = (X_1, \dots, X_p)$ i $\mathbf{Y} = (Y_1, \dots, Y_q)$. En situacions pràctiques interessa trobar una mesura d'associació entre dues matrius de dades $n \times p$ i $n \times q$. Una bona mesura es basa en la lambda de Wilks, que es fa servir per decidir si una hipòtesi és certa o hauria de ser rebutjada.

Sigui $f(\mathbf{x}, \theta)$ un model estadístic que consisteix en una densitat de probabilitat indexada per un paràmetre $\theta \in \Theta$. Donada una mostra aleatòria simple de mida n la funció de versemblança és

$$L(\mathbf{x}_1, \dots, \mathbf{x}_n; \theta) = f(\mathbf{x}_1, \theta) \times \dots \times f(\mathbf{x}_n, \theta).$$

L'estimació màxim-versemblant de θ (proposada per R. A. Fisher) és el valor $\hat{\theta}$ que maximitza $L(\mathbf{x}_1, \dots, \mathbf{x}_n; \theta)$. Plantegem una hipòtesi nul·la H_0 sobre el paràmetre θ , que es pot formular com una restricció sobre el conjunt Θ . Escrivem $H_0: \theta \in \Theta_0 \subset \Theta$. Sigui $\hat{\theta}_0$ l'estimació màxim-versemblant de θ dins Θ_0 . Tenim aleshores dues funcions de versemblança. La raó de versemblança és el quocient

$$\lambda = L(\mathbf{x}_1, \dots, \mathbf{x}_n; \hat{\theta}_0) / L(\mathbf{x}_1, \dots, \mathbf{x}_n; \hat{\theta}).$$

Atès que el numerador és un màxim restringit, clarament $0 \leq \lambda \leq 1$. Un valor de λ pròxim a 1 indicaria que les dues versemblances són semblants i aniria a

favor de H_0 . Sota un model regular i suposant H_0 certa, l'estadístic $\chi^2 = -2 \ln \lambda$ convergeix en llei a la distribució khi quadrat si $n \rightarrow \infty$, propietat que permet decidir sobre l'acceptació o el rebuig de H_0 . Sovint, en anàlisi multivariant, la raó de versemblança està relacionada amb la lambda de Wilks, que es defineix com el quocient de determinants $\Lambda = \det(\mathbf{A}) / \det(\mathbf{A} + \mathbf{B})$, on $\mathbf{A}$ i $\mathbf{B}$ són matrius de Wishart estocàsticament independents.

En el cas de dos vectors $\mathbf{X}$ i $\mathbf{Y}$ normals multivariants, la independència estocàstica entre ambdós vectors es formula plantejant la hipòtesi $H_0 : \Sigma_{12} = \mathbf{0}$, on Σ_{12} és la matriu $p \times q$ que conté les covariàncies poblacionals entre les variables de cada vector. Es demostra que la raó de versemblança λ està relacionada amb

$$\Lambda = |\mathbf{S}| / (|\mathbf{S}_{11}| |\mathbf{S}_{22}|) = (1 - r_1^2) \times \cdots \times (1 - r_m^2),$$

on $\mathbf{S}$ és la supermatriu de covariàncies $(p + q) \times (p + q)$. Sota H_0 , aquesta Λ segueix la distribució lambda de Wilks i verifica $\Lambda = \lambda^{2/n}$. Com que Λ pròxima a 1 afavoreix la hipòtesi nul·la d'independència entre $\mathbf{X}$ i $\mathbf{Y}$, resulta raonable definir la mesura d'associació global següent:

$$A_W = 1 - \Lambda = 1 - \prod_{i=1}^m (1 - r_i^2), \quad (10)$$

on $m = \min\{p, q\}$. Aleshores A_W val 0 en cas d'independència estocàstica (suposant normalitat) i val 1 si hi ha una relació lineal entre $\mathbf{X}$ i $\mathbf{Y}$. Per a les dades de Galton obtenim $A_W = 0.5060$.

10 Anàlisi de correspondències

Podem també estudiar l'associació entre les alçades de pares i fills definint intervals de classe per a les variables i aplicant anàlisi de correspondències,⁵ mètode multivariant que permet visualitzar les files i columnes d'una taula de contingència $\mathbf{N} = [n_{ij}]$, d'ordre $I \times J$. Aquest mètode, amb la contribució prèvia de R. A. Fisher i d'altres, va esdevenir popular amb l'obra de J.-P. Benzécri.⁶ Vegeu [4, 39].

Denotem per $\mathbf{P}$ la matriu de correspondències i per $\mathbf{r}$, $\mathbf{c}$ els totals marginals de les seves files i columnes:

$$\mathbf{P} = \frac{1}{n} \mathbf{N}, \quad \mathbf{r} = \mathbf{P} \mathbf{1}_J, \quad \mathbf{c} = \mathbf{P}' \mathbf{1}_I,$$

⁵ L'anàlisi de correspondències ha estat descoberta per diversos autors amb plantejaments diferents. H. O. Hirschfeld (1935) i R. A. Fisher (1940) van ser els pioners. Després K. Maung (1941) va relacionar els resultats de Fisher (donar valors a variables categòriques que maximitzen la correlació) amb l'anàlisi de correlació canònica de Hotelling. J.-P. Benzécri li donaria una interpretació geomètrica, continguda en els seus dos llibres sobre *L'analyse des données*, publicats el 1973. El mètode es difondria a partir de les obres de L. Lebart i d'altres, i en especial de M. Greenacre, publicades independentment el 1984.

⁶ Convidat per J. Torrens-Ibern, J. Robert, deixeble de Benzécri, va donar el 1974 a Barcelona un curs sobre anàlisi de correspondències.

essent $n = \sum n_{ij}$ i $\mathbf{1}_I, \mathbf{1}_J$ els vectors amb uns de dimensions I, J , respectivament. Podem interpretar $\mathbf{N}$ com el resum de les relacions entre dues matrius $\mathbf{X}$ i $\mathbf{Y}$ de dades binàries que relacionen dos conjunts de variables categòriques (possiblement ordinals) amb I i J categories. Si considerem les matrius diagonals $\mathbf{D}_r = \text{diag}(\mathbf{r})$ i $\mathbf{D}_c = \text{diag}(\mathbf{c})$, que contenen les freqüències relatives marginals per files i per columnes, es verifica

$$\mathbf{X}'\mathbf{X} = n\mathbf{D}_r, \quad \mathbf{Y}'\mathbf{Y} = n\mathbf{D}_c, \quad \mathbf{X}'\mathbf{Y} = n\mathbf{P} = \mathbf{N}.$$

Recordem que, donada una mostra $\mathbf{x} = (x_1, \dots, x_n)'$, podem expressar la variància per $s^2 = n^{-1}\mathbf{x}'\mathbf{x} - (n^{-1}\mathbf{1}'\mathbf{x})^2$, on $\mathbf{1}$ és el vector columna amb n uns. D'una manera similar, una matriu de covariàncies és $n^{-1}\mathbf{X}'\mathbf{X} - \bar{\mathbf{x}}\bar{\mathbf{x}}'$, on $\bar{\mathbf{x}} = n^{-1}\mathbf{X}'\mathbf{1}$ és el vector de mitjanes. En el nostre cas el vector (columna) de mitjanes de les I variables fila és $\mathbf{r} = n^{-1}\mathbf{X}'\mathbf{1}$. Anàlogament per a les J columnes. Per tant, les matrius de covariàncies per a les variables files, columnes, i entre files i columnes són

$$\mathbf{S}_{11} = \mathbf{D}_r - \mathbf{r}\mathbf{r}', \quad \mathbf{S}_{22} = \mathbf{D}_c - \mathbf{c}\mathbf{c}', \quad \mathbf{S}_{12} = \mathbf{P} - \mathbf{r}\mathbf{c}'.$$

Si ara volem donar valors a les categories de les variables, de manera que les correlacions siguin màximes, aplicant (9) i calculant la descomposició singular, obtenim

$$\mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{r}\mathbf{c}')\mathbf{D}_c^{-1/2} = \mathbf{U}\mathbf{D}_s\mathbf{V}', \quad (11)$$

on $\mathbf{D}_s$ és matriu diagonal amb els valors singulars, que coincideixen amb les correlacions canòniques. Els vectors canònics són

$$\mathbf{a}_i = \mathbf{D}_r^{-1/2}\mathbf{u}_i, \quad \mathbf{b}_i = \mathbf{D}_c^{-1/2}\mathbf{v}_i, \quad i = 1, \dots, K = \min\{I, J\},$$

que matricialment podem expressar per

$$\mathbf{A} = \mathbf{D}_r^{-1/2}\mathbf{U}\mathbf{D}_s, \quad \mathbf{B} = \mathbf{D}_c^{-1/2}\mathbf{V}\mathbf{D}_s.$$

La multiplicació a la dreta per $\mathbf{D}_s$ no altera les correlacions. Es compleixen les relacions següents entre les coordenades de files i de columnes

$$\mathbf{A} = \mathbf{D}_r^{-1}\mathbf{P}\mathbf{B}\mathbf{D}_s^{-1}, \quad \mathbf{B} = \mathbf{D}_c^{-1}\mathbf{P}'\mathbf{A}\mathbf{D}_s^{-1}. \quad (12)$$

Considerem ara els perfils de les files $(p_{i1}/r_i, \dots, p_{iJ}/r_i)$, que podem interpretar com les «probabilitats condicionades» de les J columnes a una fila i determinada. Definim una distància (al quadrat) entre els perfils de dues files i, i' ,

$$\delta_{ii'}^2 = \sum_{j=1}^J \frac{(p_{ij}/r_i - p_{i'j}/r_{i'})^2}{c_j}. \quad (13)$$

Si dues files tenen el mateix perfil aleshores $\delta_{ii'}^2 = 0$. Aquesta distància es coneix amb el nom de *khi quadrat*.

Es defineix la *inèrcia* o *variabilitat geomètrica* com la mitjana ponderada de les distàncies khi quadrat entre files:

$$V_{\delta} = \frac{1}{2} \sum_{i=1}^I \sum_{i'=1}^I r_i \delta_{ii'}^2 r_{i'}. \quad (14)$$

V_{δ} és una mesura de dispersió de les dades, versió multivariant de la variància. Es demostra que $V_{\delta} = \sum_{k=1}^K s_k^2 = \chi^2/n$, on s_i , $i = 1, \dots, K = \min\{I, J\}$, són els valors singulars de (11) i χ^2 és l'estadístic khi quadrat per contrastar la hipòtesi d'independència entre files i columnes de $\mathbf{N}$, és a dir, $H_0 : p_{ij} = p_{i\cdot} \times p_{\cdot j}$, on p_{ij} és la probabilitat conjunta i $p_{i\cdot}$, $p_{\cdot j}$ són les probabilitats marginals.

A més, les distàncies euclidianes (al quadrat) entre les files de la matriu $\mathbf{A} = \mathbf{D}_r^{-1/2} \mathbf{U} \mathbf{D}_s$ coincideixen amb les distàncies khi quadrat entre files. Les dues primeres coordenades (a_{i1}, a_{i2}) proporcionen una representació òptima en dimensió dos, en el sentit que la seva variabilitat geomètrica, considerant només dues coordenades, és màxima i val $V_{\delta}(2) = \sum_{k=1}^2 s_k^2$, que en general aporta una proporció alta de la variabilitat total $V_{\delta} = \sum_{k=1}^K s_k^2$.

Anàlogament podem definir el perfil de les columnes i una distància khi quadrat entre columnes. Aleshores la matriu $\mathbf{B} = \mathbf{D}_c^{-1/2} \mathbf{V} \mathbf{D}_s$ conté les coordenades euclidianes per a una representació de les columnes.

D'altra banda, les relacions (12), recíproques entre $\mathbf{A}$ i $\mathbf{B}$, demostren que les coordenades (a_{i1}) de les files són les mitjanes, ponderades pels perfils de les files, de les coordenades (b_{j1}) de les columnes. Anàlogament per a les coordenades de les columnes. Per exemple, cada primera coordenada de les files verifica

$$a_{i1} = \frac{1}{s_1} \left(b_{11} \frac{p_{i1}}{r_i} + b_{21} \frac{p_{i2}}{r_i} + \dots + b_{J1} \frac{p_{iJ}}{r_i} \right), \quad i = 1, \dots, I,$$

i anàlogament les columnes. Però aquesta mitjana baricèntrica no és perfecta (per raons geomètriques, els punts $\mathbf{a}_i$ no poden ser mitjanes dels $\mathbf{b}_j$ i recíprocament). Hi ha, per tant, un factor dilatador $1/s_1 > 1$. Una conseqüència important d'aquesta relació és que podem representar conjuntament, en un mateix gràfic, les files i columnes d'una taula de contingència $I \times J$.

EXEMPLES. Agrupem les dades de Galton i formem intervals de classe per a les alçades de mid-parents i fills. Per a les alçades dels pares es consideren els cinc intervals:

fins a 66, (66, 68], (68, 70], (70, 72], més de 72,

l'últim indicat a la figura 5 per «◦74». En el cas dels fills hi ha més variabilitat i es consideren els set intervals:

fins a 64, (64, 66], (66, 68], (68, 70], (70, 72], (72, 74], més de 74,

l'últim indicat en el gràfic per «□76». Creuant aquests intervals obtenim una taula de contingència que representem per anàlisi de correspondències simples.

Veiem clarament l'associació, ja que les alçades dels pares són pròximes a les dels fills, en el sentit, per exemple, que les alçades entre 68 i 70 dels fills són mitjanes ponderades de les alçades dels pares, i el valor que predomina és precisament la «probabilitat condicionada» a l'interval (68, 70] dels pares. La variabilitat geomètrica del gràfic (dimensió 2) representa el 96% de la variabilitat total (dimensió 4). El primer valor singular o primera correlació canònica és 0.49, valor molt semblant a la correlació $r = 0.50$ que s'obté amb les dades originals.

La taula 3 (esquerra) és una taula 4×4 que conté dades de freqüències obtingudes per Galton sobre $n = 1000$ individus, classificats segons el color dels ulls (1 blau, 2 verd o gris, 3 gris fosc, 4 castany fosc) de pares i fills. Convé tenir en compte que ara les variables són categòriques i no podem calcular el coeficient de correlació clàssic. La figura 6 (esquerra) és el resultat de l'anàlisi de correspondències simples. L'associació entre pares i fills respecte al color dels ulls és notable. La variabilitat geomètrica és el 86.7% de la variabilitat total. La primera correlació canònica és $r_1 = 0.39$.

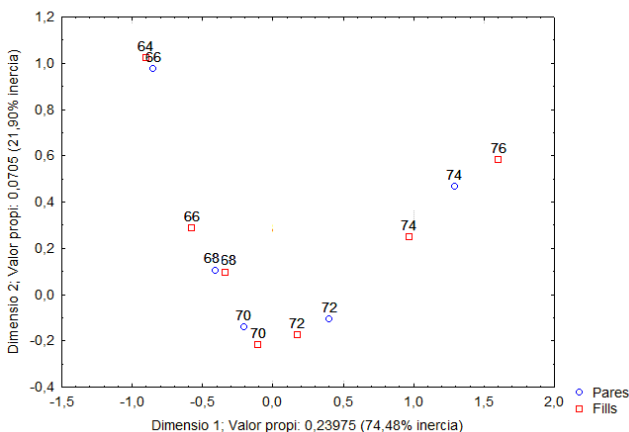


FIGURA 5: Representació per anàlisi de correspondències de les files (mid-parents) i columnes (fills), d'una taula de contingència 5×7 resultat de distribuir en intervals de longitud 2 les dades de Galton, que relacionen les estatures de pares i fills. El valor 70 representa l'interval de classe de 68 a 70.

10.1 Associació en taules d'ordre superior

La taula 3 (dreta) és una taula $2 \times 2 \times 2$ que conté dades de freqüències obtingudes per Galton sobre $n = 5008$ individus, classificats segons el color (clar o fosc) dels ulls, tenint en compte el color dels ulls dels pares i els avis. Galton va considerar 78 famílies amb molts fills, per tant hi ha relació (en el sentit

de correlació intraclàssica) entre germans i germanes. Ara tenim tres variables categòriques, que es poden representar aplicant anàlisi de correspondències ordinari a la taula de Burt, taula simètrica que conté les freqüències combinant les variables. Aquest tipus de representació es coneix per anàlisi de correspondències múltiples [15, 39]. La figura 6 (dreta) representa les freqüències de fills, pares i avis d'acord amb el color dels ulls. La separació entre color C i color F, que al seu torn estan agrupats, mostra que hi ha associació entre color i parentesc.

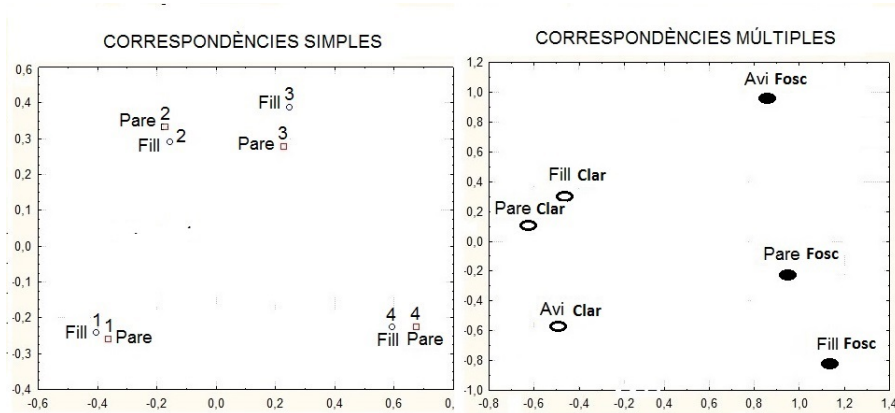


FIGURA 6: Anàlisi de correspondències simples (esquerra) i de correspondències múltiples (dreta), del colors dels ulls. Classificacions i freqüències obtingudes per Galton de 1000 individus (esquerra) i de 5008 individus (dreta).

		Pare				Avi				
Color		1	2	3	4	Clar		Fosc		
Fill	1	194	70	41	30	Pare Fill C Fill F	C	F	C	F
	2	83	124	41	36		1928	552	596	508
	3	25	34	55	23		303	395	225	501
	4	56	36	43	109					

TAULA 3: Classificació i freqüències obtingudes per Galton de 1000 individus (esquerra) i 5008 individus (dreta), segons el color dels ulls, considerant quatre colors (esquerra) i color clar o fosc (dreta), i amb el registre també del color dels ulls dels seus pares i avis.

Intentem ara mesurar l'associació estadística entre fills, pares i avis respecte del color dels ulls. En primer lloc hem d'especificar un model. Siguin p_{ijk} les probabilitats, on $i, j, k \in \{1, 2\}$ són els índexs per a fill, pare i avi. Per exemple,

p_{111} és la probabilitat que un fill tingui els ulls clars, així com el seu pare i el seu avi. L'estimació és $\hat{p}_{111} = 1928/5008$. Posem $p_{ij\cdot} = p_{ij1} + p_{ij2}$ i anàlogament $p_{i\cdot k}$, $p_{\cdot jk}$, $p_{i\cdot\cdot}$, etc. Especifiquem ara un model. En teoria, el més correcte seria suposar que hi ha vuit probabilitats p_{ijk} que sumen 1. Però estadísticament no podem decidir la seva validesa perquè hi ha tants paràmetres lliures com observacions. Les estimacions dels p_{ijk} donarien freqüències esperades $n\hat{p}_{ijk}$ que coincidirien amb les observades (l'estadístic khi quadrat seria $\chi^2 = 0$). Hem d'afinar una mica més. Suposarem que el color dels ulls dels fills és independent entre pares i fills, entre avis i fills i entre avis i pares. En termes de probabilitats:

$$p_{ij\cdot} = p_{i\cdot\cdot} \times p_{\cdot j\cdot}, \quad p_{i\cdot k} = p_{i\cdot\cdot} \times p_{\cdot\cdot k}, \quad p_{\cdot jk} = p_{\cdot j\cdot} \times p_{\cdot\cdot k}.$$

Per a les dades de la taula 3 (dreta), el test per a aquesta hipòtesi aplicant la raó de versemblança dona una khi quadrat de $\chi^2 = -2 \ln \lambda = 16.8$ amb un grau de llibertat, bastant significativa i, per tant, rebutgem que hi ha independència entre fills i pares, etc. Les dades tampoc s'ajusten a un model genètic combinant dos al·lels. De fet, el model correcte no el coneixem (vegeu la nota al peu de la pàgina 6). Acceptem, doncs, els paràmetres $p_{ij\cdot}$, $p_{i\cdot k}$, $p_{\cdot jk}$ com a probabilitats genèriques, que es desvien del model de treball format pels productes $p_{i\cdot\cdot} \times p_{\cdot j\cdot}$, etc.

Considerem ara el model d'independència completa: el color dels ulls de fills, pares i avis és independent. En termes de probabilitats:

$$p_{ijk} = p_{i\cdot\cdot} \times p_{\cdot j\cdot} \times p_{\cdot\cdot k}.$$

Aquest model és una restricció de l'anterior. Atès que la khi quadrat és $\chi^2 = -2 \ln \lambda = 868.3$ amb dos graus de llibertat, molt significativa, el model s'hauria de rebutjar. També podem afirmar que el model anterior ($\chi^2 = 16.8$) s'ajusta molt millor a les dades observades. Com cal mesurar aquest ajust, interpretat en termes d'associació?

Acceptem el valor khi quadrat dividit pels graus de llibertat com una distància entre freqüències observades i esperades, és a dir, una mesura de la desviació del model corresponent. Aplicant (4) obtenim

$$\theta = 1 - \frac{16.8}{868.3/4} = 0.9226.$$

El grau d'associació global (relatiu als dos models) que hi ha entre les tres generacions respecte al color (clar o fosc) dels ulls és alt.

El cas general de taules de contingència $J_1 \times \cdots \times J_q$ és semblant. La representació de les variables categòriques la podem fer aplicant l'anàlisi de correspondències múltiples. Aleshores cal establir dos models, un de general, l'altre més restringit, la desviació del qual expressi la dependència global. Aplicant el mateix procediment, obtindrem una mesura de dependència relativa. Usualment els càlculs es fan aplicant models log-lineals [15], atès que prenent logaritmes, els productes esdevenen sumes.

10.2 Alternatives a l'anàlisi de correspondències

Quan les freqüències no formen una taula $I \times J$ sino que provenen de I poblacions multinomials independents, aleshores és més adequada *la distància de Hellinger* entre perfils [23], seguint el plantejament degut a Rao [54]:

$$\delta_{ii'}^2 = \sum_{j=1}^J \left(\sqrt{p_{ij}/r_i} - \sqrt{p_{i'j}/r_{i'}} \right)^2.$$

Hi ha altres alternatives que responen al nom d'anàlisi no simètric, d'anàlisi log-ràtio, etc. Totes es poden unificar parametritzant la distància, que depèn de dos paràmetres $0 \leq \alpha, \beta \leq 1$, per

$$\delta_{ii'}^2 = \sum_{j=1}^J \left[\left(\frac{p_{ij}}{r_i c_j} \right)^\alpha - \left(\frac{p_{i'j}}{r_{i'} c_j} \right)^\alpha \right]^2 c_j^{2\beta},$$

vegeu [19]. Aleshores les coordenades de les files i de les columnes s'obtenen a partir de la descomposició singular:

$$\mathbf{D}_r^{1/2} \mathbf{H}_c \left\{ \frac{1}{\alpha} [(\mathbf{D}_r^{-1} \mathbf{P} \mathbf{D}_c^{-1})^{(\alpha)} - \mathbf{1} \mathbf{1}'] \right\} \mathbf{D}_c^\beta = \mathbf{U} \mathbf{D}_s \mathbf{V}', \quad (15)$$

on $\mathbf{M}^{(\alpha)} = [m_{ij}^\alpha]$, $\mathbf{H}_c = \mathbf{I} - \mathbf{1} \mathbf{r}'$. La taula 4 resumeix els diferents mètodes, amb els acrònims CA (anàlisi de correspondències), HD (anàlisi amb distància Hellinger [54]), LR (anàlisi log-ràtio [1]), NSCA (anàlisi de correspondències no simètriques). En general, les coordenades de les files són $\mathbf{A} = \mathbf{D}_r^{-1/2} \mathbf{U} \mathbf{D}_s$ i les de les columnes, $\mathbf{B} = \mathbf{D}_c^{\beta-1} \mathbf{V} \mathbf{D}_s$. La matriu centradora $\mathbf{H}_c = \mathbf{I} - \mathbf{1} \mathbf{r}'$, que premultiplica a (15), no apareix a (11) perquè no afecta les solucions CA i NSCA. La representació en dimensió 2 és una forma geomètrica d'explicar una part de la inèrcia $V_\delta = \sum s_i^2$, on $s_1, s_2, \dots$ són els valors singulars.

En el mètode CA, la inèrcia és el coeficient de contingència de Pearson i està lligat a la khi quadrat per a taules de contingència. Si considerem HD, aleshores la inèrcia té a veure amb el coeficient d'afinitat de Matusita, que mesura el grau de concordància entre dues densitats. La inèrcia en el mètode NSCA està relacionada amb el coeficient de Goodman-Kruskal, que mesura la predictibilitat de les columnes donades les files. Per tant, NSCA és adequat si les columnes fan el paper de variables resposta i les files, de variables predictors. Finalment, LA és útil per a dades composicionals [1] (els valors de cada fila són positius i sumen una quantitat fixa).

Hi ha encara més mètodes [38], alguns basats en freqüències acumulades [12]. Però el mètode CA és el més racional i el que millor respecta les propietats de les probabilitats. Vegeu [22] per a una perspectiva general.

Coordenades files, columnes		$A = D_r^{-1/2}UD_s$ $B = D_c^{\beta-1}VD_s$
SVD, inèrcia		$Q = UD_sV'$ $V_\delta = \text{tra}(Q'Q)$
Mètode	α β	Q
CA (Benzécri, Greenacre)	1 1/2	$D_r^{1/2}(D_r^{-1}PD_c^{-1} - 11')D_c^{1/2}$
HD (Hellinger, Rao)	1/2 1/2	$D_r^{1/2}H_c(D_r^{-1/2}P^{(1/2)}D_c^{-1/2} - 11')D_c^{1/2}$
LR (Aitchison, Greenacre)	0 1/2	$D_r^{1/2}H_c \ln(D_r^{-1}PD_c^{-1})D_c^{1/2}$
NSCA (Lauro, D'Ambra)	1 1	$D_r^{1/2}(D_r^{-1}PD_c^{-1} - 11')D_c$

TAULA 4: Quatre mètodes per representar la relació entre dos conjunts de variables categòriques.

11 Associació estadística amb dades generals

Fem una ullada a la taula 5, que combina un dels *Reports de la recerca a Catalunya (1996-2002)*, publicat per l'IEC, i una taula extreta de [7]. La diagonal conté el nombre d'articles de matemàtiques i estadística publicats per cada país sense col·laboració amb els altres països de la llista. Sobre la diagonal hi ha el nombre d'articles publicats amb autors d'altres països. La part per sota de la diagonal indica la relació comercial (1 si és important, 0 si no ho és). Per exemple, Espanya (incloent-hi Catalunya) ha publicat 10 560 articles, dels quals 8597 tenen autors espanyols i cap dels altres països de la llista, 692 amb coautors d'EUA, 473 amb França, etc.; i Espanya ha mantingut amb els EUA una relació comercial important (valor 1).

Hi ha associació estadística entre col·laboració científica i relació comercial? Vegem una manera general de mesurar-la.

	EUA	Esp.	Fran.	R. U.	Ital.	Alem.	Can.	Japó	Xina	Russ.
EUA	63 446	692	2281	2507	1642	2812	2739	1039	1773	893
Espanya	1	8597	473	347	352	278	163	69	104	177
França	1	1	17 155	532	916	884	496	269	167	606
Regne Unit	1	1	1	12 585	490	810	480	213	339	365
Itàlia	0	1	1	0	13 197	677	290	169	120	512
Alemanya	1	1	1	1	1	16 588	499	350	408	984
Canadà	1	0	0	1	0	0	7927	228	601	204
Japó	1	0	0	0	0	0	1	20 001	371	193
Xina	1	0	0	0	0	0	1	1	39 140	64
Rússia	0	0	0	0	0	0	0	0	1	18 213

TAULA 5: Nombre d'articles publicats amb autors d'un sol país (diagonal) i amb autors d'altres països (sobre la diagonal). Relació comercial significativa entre països (sota la diagonal).

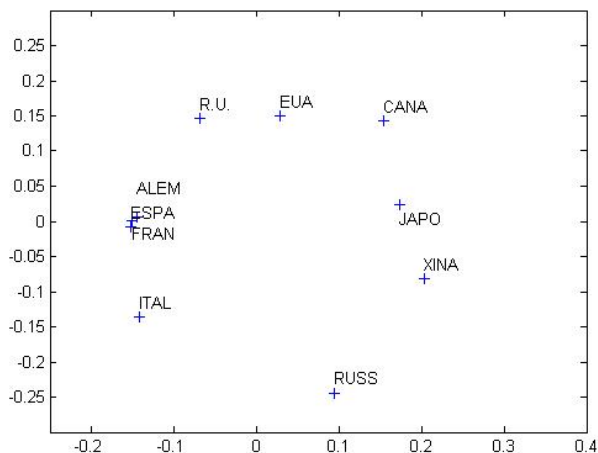


FIGURA 7: Representació RMDS de deu països tenint en compte la col·laboració científica i la relació comercial.

Sigui $\Omega = \{\omega_1, \dots, \omega_n\}$ un conjunt finit amb n objectes o individus. Suposem que per algun procediment podem definir una mesura de dissimilaritat o distància $\delta: \Omega \times \Omega \rightarrow \mathbb{R}_+$ entre cada parella d'objectes: $\delta_{ij} = \delta(\omega_i, \omega_j)$. Suposem que δ és simètrica i no negativa: $\delta_{ij} = \delta_{ji} \geq \delta_{ii} = 0$. Obtenim aleshores una matriu $n \times n$ de distàncies $\Delta_x = (\delta_{ij})$.

Suposarem que aquesta matriu de distàncies és euclidiana. És a dir, existeix una configuració de punts $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$, amb coordenades $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$, $i = 1, \dots, n$, de tal manera que $\delta_{ij}^2 = (\mathbf{x}_i - \mathbf{x}_j)'(\mathbf{x}_i - \mathbf{x}_j)$. Tenim, doncs, que les coordenades dels elements de Ω formen una matriu $n \times p$, que denotarem per $\mathbf{X} = [x_{ij}]$, de tal manera que la distància euclidiana entre cada parella de files i, j coincideix amb la distància inicial δ_{ij} .

A fi de saber si Δ_x és euclidiana, sigui $\mathbf{I}_n$ la matriu identitat i $\mathbf{1}_n$ el vector d'uns. És ben conegut que una manera d'obtenir $\mathbf{X}$ (que no és única) a partir de Δ_x , consisteix a trobar primer les matrius $\mathbf{A}_x = -\frac{1}{2}\Delta_x^{(2)}$ i $\mathbf{G}_x = \mathbf{H}_c\mathbf{A}_x\mathbf{H}_c$, on $\Delta_x^{(2)} = (\delta_{ij}^2)$ i $\mathbf{H}_c = \mathbf{I}_n - (1/n)\mathbf{1}_n\mathbf{1}_n'$ és la matriu de centrat. Seguidament calculem la descomposició espectral $\mathbf{G}_x = \mathbf{U}\Lambda_x^2\mathbf{U}'$, que proporciona la matriu de coordenades $\mathbf{X} = \mathbf{U}\Lambda_x$. Aleshores Δ_x és una matriu de distàncies euclidianes si i només si $\mathbf{G}_x$ és una matriu semidefinida positiva (els valors propis són no negatius).

Suposant que els valors propis continguts en la matriu diagonal Λ_x^2 estan disposats en ordre decreixent, les matrius $\mathbf{X}$ i $\mathbf{U}$ contenen les coordenades principals i estàndard, respectivament, dels n individus en relació amb les distàncies Δ_x . Aquest procediment, conegut per anàlisi de coordenades principals [8, 15], permet fer una representació gràfica dels n individus en dimensió reduïda (usualment 2) prenent les primeres coordenades principals, és a dir, les

dues primeres columnes de $\mathbf{X}$. El resultat és semblant a l'anàlisi de correspondències, que de fet seria una anàlisi de coordenades principals ponderat. Hem de dir, però, que aquí volem relacionar aquestes coordenades amb un segon conjunt de dades observades sobre el mateix conjunt Ω .

Considerem, doncs, un altre conjunt de dades amb observacions sobre els mateixos n individus, i que, mitjançant una distància o dissimilaritat apropiada, podem obtenir una segona matriu de distàncies Δ_y , de la qual calculem $\mathbf{A}_y$ i a continuació $\mathbf{G}_y = \mathbf{V}\Lambda_y^2\mathbf{V}'$, com hem fet abans. Si els valors propis a la diagonal de Λ_y^2 estan ordenats, les coordenades principals de Ω respecte a Δ_y són $\mathbf{Y} = \mathbf{V}\Lambda_y$. Les coordenades estàndard són les n files de $\mathbf{V}$.

Amb aquestes coordenades, l'associació entre els dos conjunts de dades es pot plantejar quantificant l'associació entre les matrius $\mathbf{X}(n \times p)$ i $\mathbf{Y}(n \times q)$. Tenint en compte que són matrius centrades (la mitjana de cada columna és 0), plantejem ara un model matricial que generalitzi (1). Seguint el plantejament iniciat a [14], el model és $\mathbf{Y} = \mathbf{XB} + \mathbf{\Xi}$, essent $\mathbf{B}$ una matriu $p \times q$ de paràmetres i $\mathbf{\Xi}$ un matriu $n \times q$ de desviacions aleatòries.

L'estimació de $\mathbf{B}$ pel criteri dels mínims quadrats és $\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ i la matriu de predicció és $\hat{\mathbf{Y}} = \mathbf{X}\hat{\mathbf{B}} = \mathbf{PY}$, on $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ és la matriu de projecció. Òbviament, no hi ha cap relació entre $\mathbf{X}$ i $\mathbf{Y}$ si $\hat{\mathbf{B}} = \mathbf{0}$, mentre que la relació és perfecta si $\hat{\mathbf{Y}} = \mathbf{Y}$.

En el test de la hipòtesi $H_0 : \mathbf{B} = \mathbf{0}$, sigui $\mathbf{E} = (\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}})$ la «matriu error». De $\mathbf{P}^2 = \mathbf{P}$ obtenim $\mathbf{E} = \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{Y}}'\hat{\mathbf{Y}}$. Denotem per $\mathbf{H} = \hat{\mathbf{Y}}'\hat{\mathbf{Y}}$ la «matriu hipòtesi». Clarament $\mathbf{Y}'\mathbf{Y} = \mathbf{H} + \mathbf{E}$. Atès que $\mathbf{Y}$ és una matriu centrada, tenim que $\mathbf{Y}'\mathbf{Y}$ és proporcional a una matriu de covariàncies. Tenim, doncs, la versió matricial de la descomposició de la variabilitat en dues parts: la deguda a la hipòtesi i la deguda a la desviació de la hipòtesi (error).

Apliquem ara el criteri de la raó de versemblança, en aquest cas equivalent a la lambda de Wilks, $W = \det(\mathbf{E}) / \det(\mathbf{E} + \mathbf{H})$. Valors de W propers a 0 indiquen que la hipòtesi H_0 hauria de ser rebutjada i, per tant, $\mathbf{B} \neq \mathbf{0}$.

En el nostre context de matrius de coordenades principals, tenim $\mathbf{Y} = \mathbf{V}\Lambda_y$ així com $\hat{\mathbf{Y}} = \mathbf{PY} = \mathbf{UU}'\mathbf{V}\Lambda_y$. Per tant, $\mathbf{E} = \Lambda_y(\mathbf{I} - \mathbf{V}'\mathbf{UU}'\mathbf{V})\Lambda_y$ la qual cosa implica $\mathbf{E} + \mathbf{H} = \mathbf{Y}'\mathbf{Y} = \Lambda_y\mathbf{V}'\mathbf{V}\Lambda_y = \Lambda_y^2$. Obtenim $W = \det(\mathbf{I} - \mathbf{V}'\mathbf{UU}'\mathbf{V})$.

D'altra banda, d'acord amb (9), les correlacions canòniques r_i entre les columnes de $\mathbf{X}$ i les de $\mathbf{Y}$ satisfan l'equació $\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{v}_i = r_i^2\mathbf{Y}'\mathbf{Y}\mathbf{v}_i$, és a dir, $\mathbf{H}\mathbf{v}_i = r_i^2(\mathbf{E} + \mathbf{H})\mathbf{v}_i$, on $\mathbf{v}_i$ és el corresponent vector propi. Això implica $\mathbf{E}\mathbf{v}_i = (1 - r_i^2)(\mathbf{E} + \mathbf{H})\mathbf{v}_i$. Per tant, podem expressar W en termes de correlacions canòniques: $W = (1 - r_1^2) \times \dots \times (1 - r_m^2)$, on $m = \min(p, q)$. Trobem la mateixa mesura d'associació $A_W = 1 - W$ que hem obtingut abans, vegeu (10), però ara per a dades generals. A_W val 0 si no hi ha relació entre els dos conjunts de dades, i val 1 si hi ha una relació lineal en el sentit que les matrius de distàncies defineixen espais equivalents.

EXEMPLE. Considerem la taula 5 i la matriu de similaritat $\mathbf{S} = (s_{ij})$, on $s_{ij} = 1$ si hi ha relació comercial significativa, $s_{ij} = 0$ en cas contrari. La relació comercial interna de cada país és molt intensa, per tant convindrem que $s_{ii} = 3$.

Considerem també la matriu $\mathbf{T} = (t_{ij})$, on per a dos països i i j definim $t_{ij} = n_{ij} / \min\{n_i, n_j\}$, essent n_{ij} el nombre d'articles publicats en col·laboració. Els nombres n_i, n_j són el total d'articles publicats pels països i i j . En el cas d'Espanya i EUA tenim $s_{12} = 1$, $t_{12} = 692/8597 = 0.0805$. Transformant les similaritats en distàncies (al quadrat)

$$\delta_{ij}^2(x) = 2(1 - s_{ij}), \quad \delta_{ij}^2(y) = 2(1 - t_{ij}),$$

tenim que $\mathbf{G}_x = \mathbf{H}_c \mathbf{S} \mathbf{H}_c = \mathbf{U} \Lambda_x^2 \mathbf{U}'$ i $\mathbf{G}_y = \mathbf{H}_c \mathbf{T} \mathbf{H}_c = \mathbf{V} \Lambda_y^2 \mathbf{V}'$. Considerem les coordenades principals $\mathbf{U} \Lambda_x$ i $\mathbf{V} \Lambda_y$ i prenem les dues primeres columnes de $\mathbf{U}$ i de $\mathbf{V}$. Denotem aquestes dues matrius de coordenades estàndard per $\mathbf{U}_2$ i $\mathbf{V}_2$. Fent una anàlisi de correlació canònica, aplicant (9) i atès que $\mathbf{S}_{11}$ i $\mathbf{S}_{22}$ són matrius identitat, obtenim $\mathbf{Q} = \mathbf{U}_2' \mathbf{V}_2$, que té dos valors singulars o correlacions canòniques $r_1 = 0.9182$, $r_2 = 0.6437$. Per tant, $A_W = 0.9081$, que indica una bona associació entre col·laboració científica i comercial.

Podem triar altres coeficients amb propietats semblants, com ara el producte $\prod r_i^2$, que en aquest cas donaria un valor bastant baix [14]. Però per diverses raons, A_W és millor, com es demostra en el tractament i la comparació d'imatges hiperespectrals [31].

La figura 7 és el resultat d'aplicar la *related metric scaling* [27], tècnica que permet representar els països tenint en compte les dues matrius de similaritats considerades. Consisteix a diagonalitzar la matriu

$$\mathbf{G} = \mathbf{G}_x + \mathbf{G}_y - \frac{1}{2}(\mathbf{G}_x^{1/2} \mathbf{G}_y^{1/2} + \mathbf{G}_y^{1/2} \mathbf{G}_x^{1/2}),$$

i calcular les coordenades principals conjuntes. Podem veure que $\mathbf{G} = \mathbf{G}_x + \mathbf{G}_y$ en el cas d'ortogonalitat de les configuracions i que $\mathbf{G} = \mathbf{G}_x = \mathbf{G}_y$ en el cas que les matrius de distàncies Δ_x, Δ_y coincideixin. La representació es pot fer també aplicant una tècnica coneguda per *multiple factor analysis* [35], que diagonalitza $\mathbf{G}_x + \mathbf{G}_y$ i, per tant, ignora la redundància entre les dues matrius, perquè prendre la suma implícitament suposa ortogonalitat, és a dir, $\mathbf{U}'\mathbf{V} = \mathbf{0}$. Considerant $\mathbf{G}_x^{1/2} \mathbf{G}_y^{1/2} + \mathbf{G}_y^{1/2} \mathbf{G}_x^{1/2}$ es tenen en compte les correlacions entre coordenades, contingudes en els productes $\mathbf{U}'\mathbf{V}$ i $\mathbf{V}'\mathbf{U}$, i es gaudeix d'algunes de les propietats de la distància de Mahalanobis (vegeu la secció 13.2).

12 MANOVA basat en distàncies

Suposem que tenim $k \geq 2$ conjunts de dades provinents de les poblacions $\Omega_1, \dots, \Omega_k$, obtinguts observant p variables quantitatives. En el model MANOVA d'una via, sigui $\mathbf{T} = \mathbf{B} + \mathbf{W}$ la descomposició de la matriu de productes creuats «total» en suma de «entre grups» i «dintre grups», vegeu (5), partint de n_i observacions provinents de Ω_i . A fi de contrastar la hipòtesi: $H_0 : \Omega_1 = \dots = \Omega_k$, en cas de normalitat multivariant i homogeneïtat de covariàncies, tenim que $\mathbf{W}$ i $\mathbf{B}$ fan el paper de $\mathbf{E}$ i $\mathbf{H}$, i l'estadístic clàssic per decidir sobre H_0 és la lambda de Wilks:

$$W = \det(\mathbf{W}) / \det(\mathbf{B} + \mathbf{W}).$$

Per tractar amb dades generals, suposem que mitjançant una funció distància δ entre observacions obtenim les matrius d'intradistàncies $\Delta_{11}, \dots, \Delta_{kk}$, i les matrius d'interdistàncies $\Delta_{12}, \dots, \Delta_{k-1k}$. La matriu de distàncies global és Δ . Calculant coordenades principals (no centrades) per a cada matriu de distàncies, i generalitzant la igualtat $(1/n) \sum_i (x_i - \bar{x})^2 = (1/2n^2) \sum_i \sum_j (x_i - x_j)^2$, obtenim les matrius $p \times p$ següents:

$$\begin{aligned}\mathbb{T} &= \sum_{g,h=1}^k \sum_{i,i'=1}^{n_g, n_h} (\mathbf{x}_{gi} - \mathbf{x}_{hi'}) (\mathbf{x}_{gi} - \mathbf{x}_{hi'})', \\ \mathbb{B} &= \sum_{g,h=1}^k n_g n_h (\bar{\mathbf{x}}_g - \bar{\mathbf{x}}_h) (\bar{\mathbf{x}}_g - \bar{\mathbf{x}}_h)', \\ \mathbb{W}_j &= \sum_{i,i'=1}^{n_j} (\mathbf{x}_{ji} - \mathbf{x}_{ji'}) (\mathbf{x}_{ji} - \mathbf{x}_{ji'})',\end{aligned}$$

que verifiquen $\mathbb{T} = \mathbb{B} + n \sum_{g=1}^k n_g^{-1} \mathbb{W}_g$. Per tant, tenim que $\text{tr}(\mathbb{T}) = \text{tr}(\mathbb{B}) + n \sum_{g=1}^k n_g^{-1} \text{tr}(\mathbb{W}_g)$.

Calculant les mitjanes de les distàncies (al quadrat), aplicant (14) amb pes $1/n$, podem obtenir la variabilitat geomètrica per a Δ i per a la Δ_{ii} de cada població Ω_i separatament. D'aquesta manera, per a dades generals i treballant només amb distàncies, és possible descompondre la variabilitat geomètrica

$$V_\delta(\text{total}) = V_\delta(\text{entre}) + \sum_{i=1}^k \frac{n_i}{n} V_\delta(\text{dintre } i), \quad (16)$$

que seria una generalització de (6). Aleshores un test per decidir si hi ha diferències significatives entre les k poblacions, podria estar basat en el quocient $y = V_\delta(\text{entre})/V_\delta(\text{total})$; vegeu [14].

Recentment, M. Anderson i d'altres, amb la tècnica anomenada PERMANOVA, han continuat el mateix enfocament amb dades depenent de dos o més factors (incloent-hi interaccions), tot i que la justificació teòrica és encara insuficient.

Hi ha una versió probabilística de (16), relacionada amb (7). Considerem un vector aleatori $\mathbf{X}$ que té per densitat la mixtura $f(\mathbf{x}) = w_1 f_1(\mathbf{x}) + \dots + w_k f_k(\mathbf{x})$, on totes les densitats tenen el mateix suport $\mathbb{S}$. Suposem que existeix una representació $\psi: \mathbb{S} \rightarrow \mathbb{E}$, on $\mathbb{E}$ és un espai euclidià (o de Hilbert separable). Aleshores la variabilitat geomètrica de $\mathbf{X}$, respecte a una distància δ , definida com a [30]

$$V_\delta(\mathbf{X}) = \frac{1}{2} \int_{\mathbb{S} \times \mathbb{S}} \delta^2(\mathbf{x}, \mathbf{y}) f(\mathbf{x}) f(\mathbf{y}) d\mathbf{x} d\mathbf{y} \quad (17)$$

verifica

$$V_\delta(\mathbf{X}) = V(\mu_1, \dots, \mu_k) + \sum_{i=1}^k w_i V_i,$$

essent $V(\mu_1, \dots, \mu_k) = \frac{1}{2} \sum_{i,j=1}^k w_i \delta^2(\mu_i, \mu_j) w_j = \sum_{i=1}^k w_i \delta^2(\mu_i, \mu)$, on $\mu_i = E_i[\psi(\mathbf{X})]$, $\delta^2(\mu_i, \mu_j) = \|\mu_i - \mu_j\|^2$, $\mu = w_1 \mu_1 + \dots + w_k \mu_k$. La variabilitat geomètrica V_i és la de (17) però canviant f per f_i ; vegeu [21].

13 Més conceptes multivariants

13.1 Correlació intraclàssica multivariant

Considerem de nou la correlació intraclàssica (8). De fet, la definició univariant de correlació entre dues observacions y_{ij} i $y_{ij'}$ és pràctica però convencional. Més aviat hem d'acceptar que $\hat{\rho}_I$ és una mesura d'homogeneïtat entre les mostres d'un mateix grup o família. Des d'aquesta perspectiva, resulta raonable definir una mesura de correlació intraclàssica multivariant entre les observacions d'una mateixa subpoblació quan disposem de k subpoblacions.

En el model MANOVA esmentat abans, sigui $\mathbf{T} = \mathbf{B} + \mathbf{W}$, i considerem $n_0 = (n - \sum_{i=1}^k n_i^2/n) / (k - 1)$. La mesura proposada, que generalitza (8), és

$$\hat{\Phi}_I = \frac{\text{tra}(\mathbf{B})/(k-1) - \text{tra}(\mathbf{W})/(n-k)}{\text{tra}(\mathbf{B})/(k-1) + (n_0-1) \text{tra}(\mathbf{W})/(n-k)}. \quad (18)$$

Més generalment, relacionant (16) amb les traces de $\mathbf{B}$ i $\mathbf{W}$, una definició de *correlació intraclàssica* per a dades generals, treballant amb una distància δ entre observacions, seria:

$$\hat{\Phi}_I = \frac{V_\delta(\text{entre})/(k-1) - \sum_{i=1}^k (\frac{n_i}{n}) V_\delta(\text{dintre } i)/(n-k)}{V_\delta(\text{entre})/(k-1) + (n_0-1) \sum_{i=1}^k (\frac{n_i}{n}) V_\delta(\text{dintre } i)/(n-k)}.$$

EXEMPLE. Per a les dades de flors iris, formades per les espècies Iris setosa, Iris versicolor i Iris virginica, emprades per R. A. Fisher per il·lustrar l'anàlisi discriminant [15], tenim $k = 3$, $n_1 = n_2 = n_3 = 50$. Per tant, $n_0 = 50$. Aplicant (18), obtenim $\hat{\Phi}_I = 0.907$. Hi ha, per tant, una bona homogeneïtat entre les flors dintre de cada espècie.

13.2 Fal·làcia ecològica multivariant

La fal·làcia ecològica amb dues variables (figura 3) es pot interpretar en el sentit que les dades globals segueixen una direcció, gairebé perpendicular a la direcció de les dades de cada subpoblació.

A fi de detectar la fal·làcia amb dades relativament complicades, suposem dues poblacions amb vectors de mitjanes $\mu_i = (\mu_{i1}, \dots, \mu_{ip})'$, $i = 1, 2$, i matriu de covariàncies (comuna) $\Sigma = [\sigma_{ij}]$. La distància de Mahalanobis entre les dues poblacions es defineix per $(\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$. Aquesta distància (al quadrat) és invariant per transformacions lineals de les variables, té en compte les correlacions i apareix de manera natural en molts models multivariants.

Si suposem que les variables estan incorrelacionades, Σ seria diagonal, i la distància de Mahalanobis esdevindria

$$(\mu_1 - \mu_2)' [\text{diag}(\Sigma)]^{-1} (\mu_1 - \mu_2) = \sum_{j=1}^p (\mu_{1j} - \mu_{2j}) / \sigma_j^2,$$

que es coneix per *distància de Pearson*.

Aleshores, tenint en compte que $R^2 = \mathbf{r}'\mathbf{R}^{-1}\mathbf{r}$ i $\mathbf{r}'\mathbf{r}$ són formalment similars a una distància de Mahalanobis i de Pearson, respectivament, entre $\mathbf{r}$ i $\mathbf{0}$, (2) suggereix la desigualtat

$$(\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2) > (\mu_1 - \mu_2)' [\text{diag}(\Sigma)]^{-1} (\mu_1 - \mu_2). \quad (19)$$

Cuadras i Fortiana [28] proven que (19) es presenta si la direcció del segment que uneix els punts μ_1 i μ_2 és essencialment ortogonal a la direcció principal (comuna) de les dades de cada subpoblació, que és donada per les primeres components principals (variables combinació lineal amb màxima variabilitat). Per tant, la fal·làcia ecològica estaria relacionada amb la desigualtat (19).

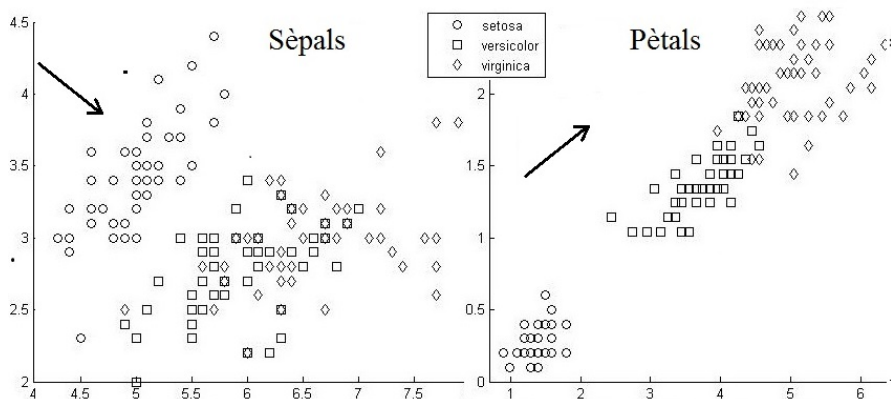


FIGURA 8: Representació de tres espècies de flors del gènere *Iris*, segons longitud i amplada de sèpals i de pètals. Per als sèpals, la distància de Mahalanobis predomina sobre la de Pearson, i els centres de les poblacions segueixen una direcció gairebé ortogonal a la principal de les dades de cada espècie. Per als pètals, la distància de Pearson predomina sobre la de Mahalanobis i els centres de les poblacions segueixen la mateixa direcció que la principal de les dades.

EXEMPLE. La figura 8 il·lustra la fal·làcia amb les dades sobre flors de les espècies *Iris* setosa, versicolor i virginica, considerades a la secció anterior. La figura 8 (esquerra) representa les tres espècies segons longitud i amplada de sèpals de la flor. La taula 6 (esquerra) conté les distàncies de Mahalanobis (a sobre de la diagonal), que predominen sobre les de Pearson (a sota de la diagonal).

Aleshores la direcció dels centres de les poblacions és gairebe ortogonal a la direcció principal de les dades, i, per tant, hi ha fallàcia ecològica. A més, les correlacions ajuntant les cent cinquanta observacions són negatives, però les correlacions dintre de cada espècie són positives. També a la figura 8 (dreta), utilitzant ara longitud i amplada dels pètals, observem el contrari, i per tant no hi ha fallàcia. Les distàncies de Mahalanobis, taula 6 (centre), són més petites que les de Pearson i la direcció dels centres de les espècies és gairebé la mateixa que la de les dades de cada espècie. Això concorda amb les correlacions, que resulten totes positives considerant les cent cinquanta observacions, i també calculant-les per a cada espècie per separat.

No resulta tan fàcil visualitzar la fallàcia considerant les quatre variables. Aleshores és quan resulta útil la desigualtat (19). Aquesta es compleix comparant *setosa* amb *versicolor* i amb *virginica*. D'acord amb la taula 6 (dreta), la suma de les distàncies de Mahalanobis i de Pearson, considerant longituds i amplades de sèpals i pètals, dona 286.7 i 277.0, respectivament. Mahalanobis predomina sobre Pearson i, per tant, hi ha fallàcia ecològica en sentit multivariant.

	Sèpals			Pètals			Sèpals i pètals		
<i>Iris</i>	<i>set.</i>	<i>vers.</i>	<i>virg.</i>	<i>set.</i>	<i>vers.</i>	<i>virg.</i>	<i>set.</i>	<i>vers.</i>	<i>virg.</i>
<i>setosa</i>	-	14.9	21.6	-	48.2	112	-	89.8	179
<i>versicolor</i>	7.0	-	1.61	70.1	-	14.0	77.1	-	17.2
<i>virginica</i>	11.2	1.96	-	166	20.7	-	177	22.7	-

TAULA 6: Distàncies de Mahalanobis (sobre la diagonal) i Pearson (sota la diagonal), entre tres espècies de flors del gènere *Iris*, considerant sèpals, pètals i les quatre mesures juntes.

13.3 Coeficients d'asimetria i curtosi

Siguin $\mathbf{x}_1, \dots, \mathbf{x}_n$ les observacions de p variables sobre una població. Denotem per $\bar{\mathbf{x}}$ el vector de mitjanes i per $\mathbf{S}$ la matriu $p \times p$ de covariàncies. Considerant els productes escalars de Mahalanobis $d_{ij} = (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})$, són mesures multivariants d'asimetria i curtosi (proposades per K. V. Mardia) les quantitats

$$b_1 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_{ij}^3, \quad b_2 = \frac{1}{n} \sum_{i=1}^n d_{ii}^2.$$

En el cas de distribució normal multivariant $N_p(\mu, \Sigma)$, els valors poblacionals són $\beta_1 = 0$ i $\beta_2 = p(p+2)$.

En la mateixa línia de la secció 12, suposem dades generals (possiblement mixtes), descrites per una matriu $n \times n$ de distàncies Δ . Calculem la corresponent matriu de productes interns $\mathbf{G} = \mathbf{U}\Lambda^2\mathbf{U}'$, que proporciona la matriu de coordenades principals $\mathbf{X} = \mathbf{U}\Lambda$. Aleshores, posant $(a_{ij})^{(m)} = (a_{ij}^m)$, la generalització, en termes de distàncies, de les mesures d'asimetria i curtosi és:

$$\hat{\beta}_1 = n \mathbf{1}'_n [(\mathbf{G}\mathbf{G}^-)^{(3)}] \mathbf{1}_n, \quad \hat{\beta}_2 = n \operatorname{tra}[(\mathbf{G}\mathbf{G}^-)^{(2)}],$$

on $\mathbf{G}^- = \mathbf{U}\mathbf{\Lambda}^{-2}\mathbf{U}'$ és la inversa generalitzada de $\mathbf{G}$. Els valors teòrics de referència serien $\beta_1 = 0$ i $\beta_2 = m(m+2)$, on m és el nombre de coordenades principals, és a dir, el nombre de columnes de $\mathbf{X}$.

EXEMPLES. Amb les dades de Galton s'obtenen $b_1 = 0.04$ i $b_2 = 9.0$, que són bastant semblants als valors teòrics $\beta_1 = 0$ i $\beta_2 = 8$ de la normal bivariant. Per a les cinquanta flors de l'espècie *Iris virginica*, s'obté $b_1 = 3.15$ i $b_2 = 24.3$, essent els valors teòrics $\beta_1 = 0$ i $\beta_2 = 24$ en cas de multinormalitat. Les dades tenen asimetria i una curtosi molt propera a la normal. Per a la taula 5 (part superior), si considerem les dues primeres coordenades principals calculades sobre la matriu $\mathbf{T}$ (secció 11), obtenim $\hat{\beta}_1 = 2.2$ i $\hat{\beta}_2 = 5.9$, que es desvien una mica de $\beta_1 = 0$ i $\beta_2 = 8$.

Part III: Distribucions bivariants amb marginals donades

14 Classes de Fréchet-Hoeffding

Quan K. Pearson va estudiar les dades de Galton sobre estatures de pares i fills, així com les seves pròpies dades (amb A. Lee) [53], va tenir molta cura d'ajustar-les a una distribució normal bivariant. Aquesta distribució va ser un model matemàtic que resumia correctament les dades observades. Per tant, una altra manera d'abordar l'associació estadística consisteix a utilitzar la teoria de distribucions bivariants amb marginals donades, que està relacionada amb la teoria de còpules.

Suposem que $H(x, y) = P[X \leq x, Y \leq y]$ és la funció de distribució conjunta del vector aleatori (X, Y) , definit sobre un espai de probabilitats $(\Omega, \mathcal{A}, P)$. Les distribucions marginals són F i G essent

$$F(x) = P[X \leq x] = H(x, \infty), \quad G(y) = P[Y \leq y] = H(\infty, y).$$

Considerarem que $H \in \mathcal{F}(F, G)$, essent $\mathcal{F}(F, G)$ la classe de funcions de distribució bivariants amb marginals F i G . Un exemple de distribució bivariant $H_0 \in \mathcal{F}(F, G)$ és $H_0(x, y) = F(x)G(y)$, que correspon al cas d'independència estocàstica.

Fréchet [36] va introduir les funcions H_- i H_+ següents:

$$H_-(x, y) = \max\{F(x) + G(y) - 1, 0\}, \quad H_+(x, y) = \min\{F(x), G(y)\},$$

i va demostrar⁷ la desigualtat

$$H_-(x, y) \leq H(x, y) \leq H_+(x, y).$$

⁷ Quan el 1977, a la biblioteca del Seminari Matemàtic de Barcelona, vaig consultar el número corresponent a 1951 de la revista on Fréchet va publicar el seu fonamental article, els plects estaven sense tallar. El document no havia estat consultat abans. Per cert, és una coincidència que els únics funcionaris per oposició del Seminari Matemàtic —l'eminent analista F. Sunyer-Balaguer (1967) i C. M. Cuadras (1974-1979)— eren de Figueres (Girona).

Aleshores les funcions $H_-, H_+ \in \mathcal{F}(F, G)$, anomenades *cotes inferior i superior de Fréchet-Hoeffding*, verifiquen:

$$H = H_- \Leftrightarrow F(X) = 1 - G(Y) \quad (\text{q.s.}),$$

$$H = H_+ \Leftrightarrow F(X) = G(Y) \quad (\text{q.s.}).$$

Aixó significa que si la distribució és H_- , hi ha dependència funcional negativa (Y és funció decreixent de X). Si la distribució és H_+ , hi ha dependència funcional positiva (Y és funció creixent de X).

EXEMPLE. Suposem que dues variables X i Y mesuren el mateix fenomen però seguint procediments diferents i que hi ha una relació no lineal entre X i Y . Si disposem de n parelles (x_i, y_i) , aleshores podem aplicar algun model de regressió. Però si només disposem de la distribució de cada variable per separat, la relació $F(x) = G(y)$ pot resoldre el problema. Com a il·lustració, suposem que $F(x) = (x - 1)^2$ si $1 \leq x \leq 2$, $G(y) = [1 - e^{-(y-2)}] / (1 - e^{-2})$ si $2 \leq y \leq 4$. De $y = G^{-1}(F(x))$ obtenim $y = 2 - \ln[1 - (1 - e^{-2})(x - 1)^2]$. Tanmateix en una situació pràctica haurem de treballar amb les funcions de distribució empíriques $F_n(x)$ i $G_n(y)$, com en el cas de la mesura del diàmetre d'un bacteri, seguint dos procediments diferents (analitzador de partícules i citometria). A [45], seguint aquest procediment, els autors obtenen polinomis de segon grau per a diverses classes de bacteris. Per exemple, per a *E. coli* obtenen $y = 8.74x^2 - 6.71x + 2.08$, on y (que va de 0.7×10^{-3} fins a 1.3×10^{-3}) és el diàmetre en micres i x (que va de 0.35×10^{-3} fins a 0.65×10^{-3}) és una mesura de flux citomètric que proporciona un analitzador de partícules.

14.1 Correlacions de Hoeffding

Hoeffding [43] va demostrar que la covariància, si existeix, es pot calcular en termes de les funcions acumulatives H, F i G , aplicant

$$\text{cov}(X, Y) = \int_{R^2} [H(x, y) - F(x)G(y)] dx dy. \quad (20)$$

Aleshores, si denotem per ρ el coeficient de correlació per a la distribució H , es compleix la desigualtat

$$\rho_- \leq \rho \leq \rho_+,$$

essent ρ_- i ρ_+ els coeficients de correlació corresponents a les cotes H_- i H_+ .

EXEMPLE. Si desconeixem la distribució conjunta de X i Y (com a l'exemple anterior), les correlacions màxima i mínima que podem obtenir, fixades les marginals, són ρ_+ i ρ_- . Per exemple, si X és uniforme sobre $(0, 1)$, $F(x) = x$, $0 < x < 1$, i Y és exponencial, $G(y) = 1 - e^{-y}$, $y > 0$, aleshores $\rho_+ = \sqrt{3}/2$ i $\rho_- = -\sqrt{3}/2$.

14.2 Desenvolupaments diagonals

D'acord amb (11), en un model bivariant discret, on el rang de les variables és numerable, podem expressar la densitat de probabilitat conjunta p_{ij} com una suma

$$p_{ij} = r_i c_j \left(1 + \sum_{n \geq 1} s_n a_{in} b_{jn} \right),$$

on $\{s_n\}$ són valors singulars, que fan el paper de correlacions canòniques. L'anàlisi de correspondències seria un cas finit del desenvolupament diagonal que presentem tot seguit.

Suposem que la derivada de Radon-Nikodym $dH(x, y)/dF(x) dG(y)$ existeix i que la mesura dH és absolutament contínua respecte a $dF \times dG$. És a dir, podem expressar H integrant respecte a $dF(x) dG(y)$. A més suposem que el coeficient de contingència de Pearson Φ^2 , definit com

$$\Phi^2 + 1 = \int_a^b \int_c^d [dH(x, y)]^2 / [dF(x) dG(y)],$$

és finit. Aleshores podem obtenir el desenvolupament:

$$dH(x, y) = dF(x) dG(y) \left[1 + \sum_{n \geq 1} \rho_n a_n(x) b_n(y) \right]. \quad (21)$$

Si les densitats conjunta i marginals (respecte de la mesura de Lebesgue) són $h(x, y)$, $f(x)$ i $g(y)$, llavors podem escriure el desenvolupament (21) com

$$h(x, y) = f(x)g(y) \left[1 + \sum_{n \geq 1} \rho_n a_n(x) b_n(y) \right]. \quad (22)$$

Aleshores $\{a_n(x)\}$, $\{b_n(y)\}$, anomenades funcions canòniques, són conjunts de funcions ortonormals, que són complets respecte de f i g . D'altra banda

$$E[a_n(X)] = E[b_n(Y)] = 0, \quad E[a_n(X)b_m(Y)] = \rho_n \delta_{mn},$$

essent ρ_n l'enèsima correlació canònica entre X i Y . Les $a_n(X)$, $b_m(Y)$ són les variables canòniques. El desenvolupament diagonal (22) va ser introduït per Lancaster [47].

Disposant les correlacions canòniques en ordre descendent, aleshores ρ_1 és la primera correlació canònica i és la màxima correlació entre una funció de X i una funció de Y

$$\rho_1 = \sup \text{cor}(\alpha(X), \beta(Y)).$$

Es verifica $0 \leq \rho_1 \leq 1$ i es pot provar que $\rho_1 = 0$ si i només si X i Y són estocàsticament independents. A més Rényi [55] va demostrar que ρ_1 és l'única mesura de dependència que verifica set postulats proposats per ell mateix. Per trobar ρ_1 , vegeu [52]. Si tenim n mostres (x_i, y_i) , un procediment pràctic per estimar ρ_1 consisteix a aplicar anàlisi de correlació canònica (secció 9) a potències de x_i i de y_i , és a dir, relacionar $(x_i, x_i^2, \dots, x_i^m)$ amb $(y_i, y_i^2, \dots, y_i^m)$.

EXEMPLE. Un element de $\mathcal{F}(F, G)$, que apareix sovint a les aplicacions i en la construcció de models bivariants, és la distribució FGM (deguda a Farlie, Gumbel i Morgenstern), que depèn del paràmetre θ ,

$$H_\theta(x, y) = F(x)G(y)\{1 + \theta[1 - F(x)][1 - G(y)]\}, \quad -1 \leq \theta \leq 1.$$

El coeficient de correlació és $\rho = \theta/3$. Aquest coeficient verifica $-1/3 \leq \rho \leq 1/3$ i és tal que $|\theta/3|$ és la primera correlació canònica. De fet, només n'hi ha una, i la FGM seria el cas més senzill de desenvolupament diagonal.

14.3 Còpules

Sigui (U, V) un vector aleatori amb valors sobre $\mathbf{I}^2$, essent $\mathbf{I} = [0, 1]$. Una còpula és una funció de distribució $C(u, v)$ amb marginals uniformes $(0, 1)$. Per tant, $C(u, v)$ verifica

$$C(0, v) = C(u, 0) = 0, \quad C(u, 1) = u, \quad C(1, v) = v.$$

El cas d'independència entre U i V correspon a la còpula $C_0(u, v) = uv$. Les cotes de Fréchet-Hoeffding són

$$C_-(u, v) = \max\{u + v - 1, 0\}, \quad C_+(u, v) = \min\{u, v\}.$$

Per a tota còpula C es verifica

$$C_-(u, v) \leq C(u, v) \leq C_+(u, v), \quad u, v \in \mathbf{I}. \quad (23)$$

Les còpules són importants pel teorema de Sklar, que diu que a tota distribució H li podem associar una còpula C_H tal que $H(x, y) = C_H(F(x), G(y))$. Per tant, n'hi ha prou a estudiar còpules per generar distribucions bivariants. Per exemple, la família de còpules

$$C_\theta(u, v) = uv[1 + \theta(1 - u)(1 - v)], \quad -1 \leq \theta \leq 1, \quad (24)$$

genera la distribució FGM. Només hem de substituir u i v per $F(x)$ i $G(y)$.

Les còpules, també anomenades *funcions de dependència* o *representacions uniformes*, capturen l'associació d'un parell (U, V) i en general, aplicant el teorema de Sklar, descriuen l'associació de qualsevol parell de variables aleatòries (X, Y) .

Per mesurar l'associació entre X i Y amb distribució conjunta H i còpula C , hom utilitza els coeficients de correlació ρ_S de Spearman i τ de Kendall, que es defineixen per

$$\rho_S = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3,$$

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1.$$

Tot i que ρ_S i τ estan definits en termes de còpules, tenen sentit per a qualsevol distribució conjunta $H(x, y)$. Ambdues mesures d'associació verifiquen:

1. Són invariants respecte de transformacions estrictament creixents de X i Y .
2. $\tau = \rho_S = 0$ si $H = H_0$ (independència estocàstica).
3. $\tau = \rho_S = -1$ si $H = H_-$ (cota inferior de Fréchet-Hoeffding).
4. $\tau = \rho_S = +1$ si $H = H_+$ (cota superior de Fréchet-Hoeffding).

Numèricament ρ_S i τ són semblants. S'han estudiat molt les relacions entre ρ_S i τ , la més coneguda de les quals és

$$-1 \leq 3\tau - 2\rho_S \leq 1.$$

Generalment es fa servir τ si es pot calcular fàcilment amb la còpula triada (cas de les anomenades còpules arquimedianes). Es fa servir ρ_S si el seu càlcul resulta més assequible que el de τ .

EXEMPLES. Per a la distribució FGM, amb còpula (24), s'obté $\rho_S = \theta/3$ i $\tau = 2\theta/9$. Es pot donar el cas que ρ_S valgui 0 però les variables siguin dependents. Per exemple:

$$C(u, v) = uv + (2u - 1)u(1 - u)(2v - 1)v(1 - v)/2,$$

no és la còpula independència, però $\rho_S = 0$. A més a més, en aquest exemple també és $\tau = 0$.

Un altre exemple ens el dona la distribució de Plackett, que s'aplica com a alternativa a la normal bivariant. Aquesta distribució apareix en el context de l'estudi de la dependència en taules de contingència 2×2 i es defineix com a $H(x, y)$ amb marginals $F(x)$ i $G(y)$, de tal manera que

$$\theta = \frac{H(x, y)[1 - F(x) - G(y) + H(x, y)]}{[F(x) - H(x, y)][G(y) - H(x, y)]}$$

és constant per a tot x, y . El paràmetre $\theta > 0$ mesura la dependència. Hi ha independència estocàstica si $\theta = 1$. Suposem ara que les distribucions marginals són de Cauchy, distribució que no té moments (la mitjana i la variància no existeixen). Aleshores no podem calcular el coeficient de correlació lineal de Pearson, però, emprant la còpula corresponent, podem calcular la rho de Spearman, que per a la distribució de Plackett és

$$\rho_\theta = \frac{\theta + 1}{\theta - 1} - \frac{2\theta}{(\theta - 1)^2} \log(\theta).$$

Fins aquí hem resumit els aspectes més bàsics de les distribucions amb marginals donades i les seves còpules. Per a un estudi ampli d'aquest tema, vegeu [51].

Si la modelització de les dades bivariants la podem descriure mitjançant una còpula, tenim un camí idoni per esbrinar el tipus d'associació entre les variables. Però hi ha moltes famílies de còpules: simètriques, arquimedianes, max-estables,

singulars, etc. Recentment es tendeix a identificar primer l'estructura d'una còpula partint de les dades [48] i després triar la còpula escaient.

Les seccions següents contenen resultats que resumeixen tres articles [13, 17, 20] sobre còpules.

15 La covariància entre dues funcions

Sigui (X, Y) un vector aleatori amb funció de distribució H i marginals F i G . Vegem primer dues generalitzacions de (20) i del desenvolupament (21). D'una banda tenim que

$$\text{cov}(\alpha(X), \beta(Y)) = \int_{R^2} [H(x, y) - F(x)G(y)] d\alpha(x) d\beta(y),$$

suposant que les funcions α i β verifiquen les condicions d'existència de la variància [3, 11]. D'altra banda tenim que

$$H(x, y) - F(x)G(y) = \sum_{n \geq 1} \rho_n \int_a^b M_1(x, s) da_n(s) \int_c^d M_2(t, y) db_n(t), \quad (25)$$

on $M_1(x, y) = \min\{F(x), F(y)\} - F(x)F(y)$, $M_2(x, y) = \min\{G(x), G(y)\} - G(x)G(y)$; vegeu [12].

Convé observar que (25) és un desenvolupament de la funció de distribució en termes de correlacions canòniques i cotes de Fréchet. De fet, estem interessats a trobar les correlacions i les funcions canòniques en el cas de còpules. Altrament dit, hem de trobar funcions ϕ_1 i ϕ_2 tals que

$$\rho = \sup \text{cor}(\phi_1(U), \phi_2(V)). \quad (26)$$

Per a aquest fi ens serà útil el resultat següent.

TEOREMA 1. *Siguin C, C^* dues còpules corresponents a dos vectors aleatoris (U, V) i (U^*, V^*) . Si $\alpha(u)$ i $\beta(v)$ són funcions de variació afitada definides sobre $\mathbf{I} = [0, 1]$, aleshores*

$$\text{cov}(\alpha(U), \beta(V)) - \text{cov}(\alpha(U^*), \beta(V^*)) = \int_{\mathbf{I}} \int_{\mathbf{I}} [C(u, v) - C^*(u, v)] d\alpha(u) d\beta(v).$$

En particular, la covariància entre $\alpha(U)$ i $\beta(V)$ és

$$\text{cov}(\alpha(U), \beta(V)) = \int_{\mathbf{I}} \int_{\mathbf{I}} [C(u, v) - uv] d\alpha(u) d\beta(v), \quad (27)$$

i la variància de $\alpha(U)$ és

$$\text{var}[\alpha(U)] = \int_{\mathbf{I}} \int_{\mathbf{I}} [\min\{u, v\} - uv] d\alpha(u) d\alpha(v). \quad (28)$$

A partir d'ara suposarem que (U, V) és intercanviable. Altrament dit, la còpula $C(u, v)$ és una funció simètrica. Aleshores podem suposar $\phi_1 = \phi_2$ a (26). El problema es redueix a trobar ρ i ϕ de tal manera que

$$\rho = \sup \text{cov}(\phi(U), \phi(V)) / \text{var}[\phi(U)].$$

16 Operadors integrals

Considerem una còpula C , la cota superior $\min\{u, v\}$, que denotem per M , i la còpula d'independència uv , que anomenem Π . Tenim que $C < M$ en el mateix sentit que a (23). Aleshores és clar que a les integrals de (27) i (28) hi intervenen els nuclis simètrics $K = C - \Pi$ i $L = M - \Pi$, que verifiquen $K < L$.

Definim a continuació els operadors integrals i els productes escalars relacionats amb una còpula C . Advertim que fem servir el mateix símbol K per al nucli i per a l'operador integral associat. Suposarem que totes les funcions pertanyen al conjunt $\mathbb{B}$ de funcions de variació afitada a l'interval $\mathbf{I} = [0, 1]$. Sabem que $\mathbb{B} \subset L^2(\mathbf{I})$.

DEFINICIÓ 2. L'operador integral K sobre $\mathbb{B}$ és

$$K\varphi(u) = \int_{\mathbf{I}} K(u, v) d\varphi(v).$$

De manera similar definim $L\varphi$ i $T\varphi$ essent $T = aK + bL$.

K i L proporcionen la covariància i la variància, respectivament. Els operadors integrals també serveixen per resoldre equacions diferencials.

EXEMPLE. Considerem l'equació $y'' + f(x) = 0$, $y(0) = y(1) = 0$, on $f(x)$ és una funció contínua a l'interval $(0, 1)$. Es pot provar que la solució és $y(x) = \int_{\mathbf{I}} L(x, s) dF(s)$, essent $F(x)$ una primitiva de $f(x)$ i $L(s, t) = \min\{s, t\} - st$.

DEFINICIÓ 3. El producte escalar generalitzat entre les funcions integrables ϕ , φ de $\mathbb{B}$ és

$$(\phi, K\varphi) = \int_{\mathbf{I}^2} K(u, v) d\phi(u) d\varphi(v).$$

De manera similar, definim $(\phi, L\varphi)$, $(\phi, T\varphi)$ i $(K\phi, L\varphi)$.

DEFINICIÓ 4. L'operador K és L -compacte si per a tota successió L -afitada $\{\phi_n\}$, en el sentit que $(\phi_n, L\phi_n)$ és finit, la successió $\{K\phi_n\}$ conté una subsuccessió convergent $\{K\phi_{n(i)}\}$. És a dir, $K\phi_{n(i)} \rightarrow \phi^*$, en el sentit que existeix ϕ^* tal que $\|K\phi_{n(i)} - \phi^*\|_L \rightarrow 0$ quan $i \rightarrow \infty$.

Són ben conegudes les propietats dels operadors a l'espai $L^2(\mathbf{I})$ quan K és general i L és l'operador identitat. Però aquí L , que permet trobar la variància, no és la identitat.

Les propietats següents es poden demostrar aplicant (27) i (28) i adaptant propietats conegudes en el cas que L sigui l'operador identitat.

1. Els operadors integrals K i L són lineals i satisfan

$$(\phi, L\phi) \geq \max\{(\phi, K\phi), 0\} \geq 0.$$

Per tant, L és definit positiu i el producte escalar $(\phi, L\varphi)$ defineix la norma $\|\phi\|_L = [(\phi, L\phi)]^{1/2}$.

2. K és de Hilbert-Schmidt, i. e., $\int_{\mathbf{I}^2} K(u, v)^2 du dv < \infty$.
3. L'operador integral K és L -compacte.

4. Si K és lineal i L -compacte, aleshores $T = aK + bL$ és també compacte per a qualsevol parella de constants $a, b \in \mathbb{R}$.

Clarament, si canviem $K = C - \Pi$ per $L = M - \Pi$, les propietats anteriors segueixen essent vàlides, atès que L és també la diferència de dues còpules.

17 Anàlisi generalitzada de funcions pròpies

Vegem ara alguns resultats sobre funcions pròpies i valors propis relacionats amb els operadors K i L .

DEFINICIÓ 5. Una funció pròpia generalitzada de K respecte de L és la parella (ϕ, λ) , on ϕ és una funció i λ és un nombre real, tals que

$$\int_{\mathbf{I}} K(u, v) d\phi(v) = \lambda \int_{\mathbf{I}} L(u, v) d\phi(v), \quad (29)$$

uniformement en $u \in \mathbf{I}$. Ho expressem per $K\phi = \lambda L\phi$.

EXEMPLE. Sigui la còpula $F_\lambda = \lambda M + (1 - \lambda)\Pi$. Aleshores $K = F_\lambda - \Pi = \lambda(M - \Pi)$. Per tant, $K\phi = \lambda L\phi$ on $L(s, t) = \min\{s, t\} - st$ i ϕ és qualsevol funció integrable sobre l'interval $(0, 1)$.

Si (ϕ, λ) és una parella pròpia (funció i valor propi) de K respecte de L , amb $\lambda \neq 0$, tenim que

$$\frac{(\phi, K\phi)}{(\phi, L\phi)} = \lambda. \quad (30)$$

No és fàcil provar el contrari: si (ϕ, λ) satisfà (30) uniformement en $u \in \mathbf{I}$, aleshores és una parella pròpia que satisfà (29).

Presentem a continuació algunes propietats de les parelles pròpies. Aquí $\|K\|_L$, $\|K\|$ i $\|T\|$ són normes dels operadors K i T .

TEOREMA 6. Sigui $\|K\|_L = \sup |(\xi, K\xi)/(\xi, L\xi)|$, $\|K\| = \sup |(\xi, K\xi)/(\xi, \xi)|$ i $\|T\| = \sup |(\xi, T\xi)/(\xi, \xi)|$, essent $\xi \neq \mathbf{0}$, on $T = aK + bL$, $a, b \in \mathbb{R}$. Si α i β són dues funcions de $\mathbb{B}$, aleshores:

1. Si K és definit positiu,

$$(\alpha, K\beta) \leq \min\{\|K\|_L \|\alpha\|_L \|\beta\|_L, \|K\| \|\alpha\| \|\beta\|\}. \quad (31)$$

2. En general,

$$|(\alpha, T\beta)| \leq \|T\| \|\alpha\| \|\beta\|. \quad (32)$$

Amb l'ajut d'aquest teorema es demostra que existeix almenys un valor propi, que en el nostre cas significa que existeix la màxima correlació canònica. El teorema següent és ben conegut quan L és l'operador identitat. En el nostre

context de dos operadors generats per còpules, la demostració canvia bastant i no és fàcil.⁸

TEOREMA 7. *K té almenys un valor propi respecte de L.*

Com en el cas clàssic (on L és la identitat), la demostració consisteix a provar que el valor propi, que sempre existeix, és

$$\lambda_1 = \sup_{\xi \neq 0} (\xi, K\xi) / (\xi, L\xi). \quad (33)$$

Però aquí L no és la identitat i la demostració necessita incorporar algunes innovacions [17].

Ara que sabem que hi ha valors propis, mencionem algunes propietats.

1. Si λ és un valor propi, aleshores $\lambda \leq \lambda_1 \leq 1$, on λ_1 es defineix a (33).
2. Si ψ i ψ^* són funcions pròpies amb valors propis $\lambda \neq \lambda^*$, aleshores $(\psi, K\psi^*) = (\psi, L\psi^*) = 0$.
3. Sigui $\{\psi_n, \lambda_n\}$ un conjunt de parelles pròpies tals que $(\psi_n, L\psi_n) > 0$. Si $\{\psi_n, \lambda_n\}$ és infinit numerable, aleshores $\lim_{n \rightarrow \infty} \lambda_n = 0$.
4. Sigui $\{\psi_n, \lambda_n\}$ un conjunt de parelles pròpies tals que $(\psi_n, L\psi_n) > 0$. Definim

$$\Phi_n(u) = \int_{\mathbf{I}} L(u, s) d\psi_n(s). \quad (34)$$

Es verifica $\Phi_n(0) = \Phi_n(1) = 0$, $\Phi'_n = -\psi_n$ i la funció contínua Φ_n té diversos zeros a $\mathbf{I}$.

Recordem ara el teorema de Mercer [2]. Si $A(s, t)$ és el nucli d'un operador integral continu, simètric i definit no negatiu, i $\{\psi_n, \lambda_n\}$ és el conjunt numerable de parelles pròpies, on els λ_n són positius i ψ_n és un sistema ortonormal, aleshores

$$A(s, t) = \sum_{n \geq 1} \lambda_n \psi_n(s) \psi_n(t), \quad (35)$$

on la sèrie convergeix absolutament i uniformement en s, t .

Si A és l'operador amb nucli $L(u, v) = \min\{u, v\} - uv$, i considerem les funcions $h_n(x) = \int_0^x \psi_n(s) ds$, podem obtenir una successió ortogonal de variables aleatòries, que permet desenvolupar U (i en general qualsevol variable X) en sèrie del tipus $X = \sum_{n \geq 1} d_n h_n(X)$, on la convergència és en mitjana quadràtica. Vegeu [16, 24, 26].

EXEMPLE. Suposem U amb distribució uniforme sobre $(0, 1)$. Aleshores $\min\{s, t\} - st = \sum_{n \geq 1} \lambda_n \psi_n(s) \psi_n(t)$, essent $\psi_n(x) = \sqrt{2} \sin(n\pi x)$, $\lambda_n = 1/(n\pi)^2$, $h_n(x) = [\sqrt{2}/(n\pi)][1 - \cos(n\pi x)]$.

En el cas de dos operadors integrals K i L , el desenvolupament en sèrie (35) presenta algunes modificacions.

⁸ Donades dues matrius simètriques $\mathbf{A}$ i $\mathbf{B}$, la segona definida positiva, no és immediat provar que hi ha almenys un valor propi de $\mathbf{A}$ respecte de $\mathbf{B}$ (un λ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}$), sense fer ús de matrius inverses ni determinants, conceptes que no podem aplicar (en general) treballant amb operadors integrals.

TEOREMA 8. *Sigui $\{\psi_n, \lambda_n\}$ el conjunt numerable de parelles pròpies de K respecte de L , tals que $(\psi_n, L\psi_n) = 1$. Suposem que K i L són dos operadors integrals continus, simètrics, essent L definit positiu. Aleshores és vàlid el desenvolupament*

$$K = \sum_{n \geq 1} \lambda_n L\psi_n \otimes \psi_n L, \quad (36)$$

en el sentit que $[K(u, v) - \sum_{i=1}^n \lambda_i \int_I L(u, s) d\psi_i(s) \int_I L(t, v) d\psi_i(t)^2]^2 \rightarrow 0$ si $n \rightarrow \infty$, uniformement en $u, v \in I$. A més, si $\{\lambda_n\}$ és infinit numerable, es compleix que $\lim_{n \rightarrow \infty} \lambda_n = 0$.

Ressaltem que els valors propis poden ser negatius i que per a certes còpules, el desenvolupament (36) podria ser una integral en lloc d'una sèrie.

Si fem ús de la funció (34), veiem fàcilment que

$$K(u, v) = \sum_{n \geq 1} \lambda_n \Phi_n(u) \Phi_n(v). \quad (37)$$

EXEMPLE. Considerem la còpula

$$C(u, v) = uv + u(1-u)v(1-v) + (2u-1)u(1-u)(2v-1)v(1-v)/2.$$

Si $K(u, v) = C(u, v) - uv$ i $L(u, v) = \min\{u, v\} - uv$, els valors propis de K respecte a L són $\lambda_1 = 1/3$ i $\lambda_2 = 1/10$. Es verifica

$$K = \frac{1}{3} L\psi_1 \otimes \psi_1 L + \frac{1}{10} L\psi_2 \otimes \psi_2 L,$$

essent $\psi_1(u) = \sqrt{3}(1-2u)$ i $\psi_2(u) = \sqrt{5}(6u^2 - 6u + 1)$ les corresponents funcions pròpies.

18 Anàlisi canònica

Apliquem ara aquest esquema a les còpules. El resultat següent relaciona les parelles pròpies amb les funcions i correlacions canòniques.

TEOREMA 9. *Sigui (U, V) un vector aleatori que té per distribució la còpula C . Suposem que $\phi: I \rightarrow \mathbb{R}$ és una funció tal que $\text{var}[\phi(U)]$ és finita. Considerem els operadors integrals $K = C - \Pi$ i $L = M - \Pi$. Es verifica:*

1. *El coeficient de correlació entre $\phi(U)$ i $\phi(V)$ és*

$$\text{cor}(\phi(U), \phi(V)) = \frac{(\phi, K\phi)}{(\phi, L\phi)}, \quad (38)$$

essent $(\phi, K\phi) = \text{cov}(\phi(U), \phi(V))$, $(\phi, L\phi) = \text{var}[\phi(U)]$.

2. *Si ϕ és una funció pròpia de K respecte a L , aleshores ϕ és una funció canònica i $(\phi, K\phi)/(\phi, L\phi) = \rho$ és la corresponent correlació canònica.*

Per tant, a partir de (36), podem expressar la còpula C per

$$C = \Pi + \sum_{n \geq 1} \rho_n L \phi_n \otimes \phi_n L, \quad (39)$$

que seria una variant de (25).

Tenint en compte propietats geomètriques basades en la distància khi quadrat (13), vegeu [16, 29], podem definir la dimensió d'una còpula com la cardinalitat del conjunt $\{\lambda_n\}$ de valors propis de K respecte a L .

Tenim aleshores:

1. Còpula de dimensió 0: és la còpula independència $\Pi(u, v) = uv$.
2. Còpula de dimensió 1: un exemple és la còpula FGM, definida a (24).
3. Còpula de dimensió 2: un exemple és

$$C_2(u, v) = uv + \theta_1 u(1-u)v(1-v) + \theta_2 (2u-1)u(1-u)(2v-1)v(1-v), \quad (40)$$

on (θ_1, θ_2) pertany a una regió continguda a $[-2, 2] \times [-1, 2]$.

4. Còpula de dimensió infinita numerable: un exemple és la còpula AMH (Ali, Mikhail i Haq):

$$AMH_\theta(u, v) = uv / [1 - \theta(1-u)(1-v)], \quad -1 \leq \theta \leq 1. \quad (41)$$

5. Còpula de dimensió contínua, és a dir, infinita no numerable: un exemple és la còpula de Cuadras-Augé:

$$CA_\theta(u, v) = (uv)^{1-\theta} (\min\{u, v\})^\theta, \quad 0 \leq \theta \leq 1. \quad (42)$$

Aquesta còpula presenta una novetat important en la teoria de les funcions pròpies, tradicionalment basada en el supòsit que els valors propis formen un conjunt numerable. Sobre això en parlarem a la secció següent.

18.1 Ajust d'una còpula per una altra de coneguda

Un problema de les còpules (i dels models en general), és que no sempre sabem quina és la que podem ajustar a les dades observades. Suposem que tenim dues còpules: C_M , la còpula que fem servir com un model, i C_T , la còpula «veritable», la que veritablement segueixen les nostres dades bivariants.

Suposem que el desenvolupament canònic

$$dC_M(u, v) = du dv + \sum_n \rho_n A_n(u) du B_n(v) dv$$

de C_M existeix i és coneguda, essent A_n i B_n funcions canòniques unitàries (mitjana 0 i variància 1). Ens interessa aproximar C_T per una combinació lineal finita de correlacions i funcions canòniques. Dit amb més precisió, volem obtenir l'aproximació (vegeu [25])

$$\frac{dC_M(u, v)}{du dv} \simeq 1 + \sum_{i=1}^k c_i A_i(u) B_i(v),$$

on $c_1, \dots, c_k$ són coeficients reals que minimitzen l'expressió

$$\int_I \int_I \left(\frac{dC_M(u, v) - du dv}{du dv} - \sum_{i=1}^k c_i A_i(u) B_i(v) \right)^2 du dv. \quad (43)$$

TEOREMA 10. *Suposem $(U, V) \sim C_T$. Els coeficients que minimitzen (43) són $c_i = \rho_i$, essent*

$$\rho_i = \text{cor}(A_i(U), B_i(V)), \quad i = 1, \dots, k.$$

L'avantatge d'aquest procediment és clar: coneixem les funcions canòniques A_i i B_i i, per tant, podrem calcular (mitjançant estimació estadística) les correlacions entre les parelles $A_i(U)$ i $B_i(V)$ respecte de la veritable còpula C_T , que és la que segueixen les nostres dades, però que en general desconeixem.

θ	$\rho_1 = \rho_S(C_2)$	ρ_2	a	$\rho_S(AMH)$	$\tau(AMH)$	$\tau(C_2)$
-1	-0.2711	0.0217	0.0055	-0.2710	-0.1817	-0.1815
-0.5	-0.1489	0.0080	0.0017	-0.1489	-0.0995	-0.0995
0.5	0.1924	0.0223	0.0032	0.1924	0.1288	0.1286
1	0.4783	0.2323	0.0261	0.4784	0.3333	0.3335

TAULA 7: Aproximació de la còpula AMH (model cert però desconegut) per una còpula coneguda que té dimensió dos. L'ajust és força bo i els coeficients de correlació rho i tau són gairebé els mateixos.

EXEMPLE. Una aproximació en dimensió 1 consisteix a utilitzar la còpula (24). La FGM és un model molt utilitzat en estadística, tot i que dona un ajust pobre. Una aproximació bastant millor seria la còpula (40) (o una versió amb dimensió superior). Suposem, per exemple, que la còpula AMH, definida a (41), és la C_T , i que l'aproximem per la còpula (40), que fa el paper de model C_M . Calculant

$$a = \max_{u, v \in I} |C_T(u, v) - C_M(u, v)|,$$

obtenim valors petits, és a dir, un bon ajust, depenent del paràmetre θ de la còpula correcta AMH. A més, les correlacions de Spearman i Kendall donen pràcticament el mateix. La taula 7 mostra els valors dels coeficients de l'aproximació, expressats en termes de correlacions canòniques, on ρ_1 coincideix amb la rho de Spearman per a la còpula C_2 .

Ara bé, si reduïm la dimensió d'una còpula eliminant termes del desenvolupament diagonal (36), o de (37), hem de tenir en compte que no sempre obtenim una nova còpula. De manera similar, en anàlisi de correspondències, quan representem les files i columnes en dimensió dos, el gràfic podria reflectir una taula amb valors negatius.

19 Dimensió contínua

Considerem el cas de parelles pròpies (ϕ, λ) on la funció ϕ té norma 0. Això significa, en el nostre context, tractar amb variables aleatòries no constants però amb variància 0. Però aquesta suposició fa trontollar la teoria dels valors propis. Abans de seguir, comentem algunes diferències amb la teoria ordinària dels operadors aplicats a funcions que no són constants.

1. Les funcions pròpies corresponents a un mateix valor propi generen un subespai de dimensió finita. Si la norma és 0 haurem de considerar valors propis amb multiplicitat infinita.
2. De $K\phi = \lambda\phi$, deduïm que les funcions pròpies ordinàries són contínues a $\mathbf{I} = [0, 1]$. Ara tenim $K\varphi = \lambda L\phi$ i podem trobar funcions contínues a trossos amb discontinuïtats de salt.
3. Les funcions es normalitzen segons $(\phi, L\phi) = 1$. Ara hem de considerar el cas $(\phi, L\phi) = 0$. Aleshores suposarem que existeix una família de funcions $\{\phi_\varepsilon\}$ de tal manera que $(\phi_\varepsilon, L\phi_\varepsilon) > 0$ i a més

$$\lim_{\varepsilon \rightarrow 0} \phi_\varepsilon = \phi, \quad \lim_{\varepsilon \rightarrow 0} \frac{K\phi_\varepsilon}{L\phi_\varepsilon} = \lambda > 0,$$

en el sentit que $K\phi_\varepsilon(u)/L\phi_\varepsilon(u) \rightarrow \lambda$ uniformement en $u \in \mathbf{I}$.

4. El conjunt $\mathbb{S}_\lambda$ de valors propis és finit o infinit numerable. En aquest nou escenari, els valors propis de K respecte a L poden constituir un conjunt $\mathbb{S}_\lambda$ no numerable. Per exemple, $\mathbb{S}_\lambda$ podria ser un interval que té la potència del continu.

EXEMPLE. Definim $\phi_\varepsilon(u) = 1$ si $u \in (1 - \varepsilon, 1]$, $\phi_\varepsilon(u) = 0$ en altre cas, on $\varepsilon > 0$ és arbitràriament petit. Sigui ϕ l'indicador de $\{u = 1\}$ i $L = M - \Pi$, $K = M^{1/2}\Pi^{1/2} - \Pi$. Es pot provar (vegeu [12, 16]) que $(\phi_\varepsilon, L\phi_\varepsilon) = \varepsilon(1 - \varepsilon)$, $(\phi_\varepsilon, K\phi_\varepsilon) = (1 - \varepsilon)^{3/2} - (1 - \varepsilon)^2$, per tant, $\lim_{\varepsilon \rightarrow 0} K\phi_\varepsilon/L\phi_\varepsilon = 1/2$. Tenim que $(\phi, 1/2)$ és una parella pròpia de K respecte a L de tal manera que $(\phi, K\phi) = (\phi, L\phi) = 0$.

Ressaltem que les funcions pròpies de norma 0 poden coexistir amb les de norma positiva.

20 Còpules amb part singular

Una família general que conté còpules amb dimensió finita, numerable o contínua, és

$$D_\theta(u, v) = \min\{u, v\}G_\theta(\max\{u, v\}),$$

on θ és un paràmetre i G_θ és una funció de la qual parlarem tot seguit. Un membre d'aquesta família és la CA_θ , definida a (42), que és la mitjana geomètrica de les còpules M i Π . Tenim que $CA_0 = \Pi$ i $CA_1 = M$. Així doncs θ és una mesura de dependència.

Aquesta família, introduïda per Cuadras i Augé [18], ha estat estudiada i aplicada a [32, 49, 51] i a d'altres treballs. Té una part singular, ja que: $P[U = V] = \theta/(2 - \theta)$. De fet, si δ_A representa la funció indicadora de l'esdeveniment A , la densitat és

$$f_\theta(u, v) = (1 - \theta) \max\{u, v\}^{-\theta} \delta_{\{u \neq v\}} + \theta u^{1-\theta} \delta_{[u=v]}, \quad (44)$$

respecte de la mesura $\mu = \lambda^2 + \lambda^1$, essent λ^2 la mesura de Lebesgue sobre $\mathbb{R}^2$ i λ^1 la mateixa mesura concentrada a $\{u = v\}$. La rho de Spearman i la tau de Kendall de CA_θ són: $\rho_S = 3\theta/(4 - \theta)$ i $\tau = \theta/(2 - \theta)$.

EXEMPLE. La família CA_θ s'ha utilitzat per definir un índex d'estabilitat financera, anomenat «Cuadras-Augé index»; vegeu [6]. Comentem, però, una situació més senzilla: la distribució dels sous de les parelles d'una població. Siguin (X, Y) els sous de l'home i la dona que són parella. Els dos sous poden ser independents, però hi ha la possibilitat que siguin iguals (cas freqüent en professors i funcionaris). Suposem que les funcions de distribució són de Pareto: $F(x) = 1 - (m_0/x)^\alpha$ si $x \geq m_0$, $G(y) = 1 - (m_0/y)^\beta$ si $y \geq m_0$, on m_0 representa el sou mínim i α i β són paràmetres tals que $2 < \alpha < \beta$. Observem que la mitjana de X és més gran que la de Y , i que si el paràmetre β tendeix a infinit, el sou mitjà tendeix al mínim m_0 . Podem prendre CA_θ com a còpula relacionada amb la distribució conjunta $H(x, y)$ de (X, Y) , on els sous mínim i màxim correspondrien als valors 0 i 1, respectivament. Però la densitat (44) de (U, V) indica que els sous iguals tenen més probabilitat a mesura que augmenten. Considerem, doncs, $(1 - U, 1 - V)$. És fàcil veure que la còpula és

$$C_S(u, v) = u + v - 1 + CA_\theta(1 - u, 1 - v).$$

Per a C_S també els sous mínim i màxim corresponen als valors 0 i 1. La densitat adopta l'expressió

$$g_\theta(u, v) = (1 - \theta) \max\{1 - u, 1 - v\}^{-\theta} \delta_{\{u \neq v\}} + \theta(1 - u)^{1-\theta} \delta_{[u=v]}.$$

Ara els sous alts tenen menys probabilitat de coincidir.

Aplicant el teorema de Sklar (secció 14.3), la funció de distribució conjunta, possible model per a (X, Y) , és $H(x, y) = C_S(F(x), G(y))$, és a dir, $H(x, y) = 0$ si $x, y < m$, i

$$H(x, y) = 1 - (m_0/x)^\alpha - (m_0/y)^\beta + CA_\theta((m_0/x)^\alpha, (m_0/y)^\beta) \text{ si } x, y \geq m_0.$$

Continuant amb la teoria, a fi de trobar els valors i les funcions pròpies per als nuclis $K = CA_\theta - \Pi$ i $L = M - \Pi$, definim per a $0 \leq y \leq 1$ i $\varepsilon > 0$ arbitràriament petit, la funció

$$\mathcal{H}_{y,\varepsilon}(x) = \begin{cases} 0 & \text{si } x < y, \\ 1 & \text{si } y \leq x < y + \varepsilon, \\ \varepsilon/(y + \varepsilon) & \text{si } x \geq y + \varepsilon, \end{cases}$$

que té per límit la funció indicadora $\phi_Y(x)$,

$$\lim_{\varepsilon \rightarrow 0} \mathcal{H}_{Y,\varepsilon}(x) = \phi_Y(x) = \begin{cases} 1 & \text{si } x = Y, \\ 0 & \text{si } x \neq Y. \end{cases}$$

Presentem ara la descomposició en valors i funcions pròpies de K respecte a L . El resultat següent es pot provar aplicant teoria de funcions generalitzades. Aquí fem l'anàlisi sobre els nuclis $K_{\theta_i} = M^{\theta_i} \Pi^{1-\theta_i} - \Pi$, $i = 1, 2$. Observem que $L = M - \Pi = K_1$.

TEOREMA 11. *Si $0 \leq \theta_1 < \theta_2 \leq 1$, el conjunt de parelles pròpies de K_{θ_1} respecte a K_{θ_2} està format per (ϕ_Y, λ_Y) on*

$$\phi_Y = \mathcal{H}_{Y,0}, \quad \lambda_Y = (\theta_1/\theta_2)Y^{\theta_2-\theta_1}, \quad 0 \leq Y \leq 1.$$

Tenint ara en compte que les parelles pròpies són funcions i correlacions canòniques, obtenim el resultat següent.

TEOREMA 12. *El conjunt de funcions i correlacions canòniques de la família CA_θ és $(\phi_Y, \theta Y^{1-\theta})$, $0 \leq Y \leq 1$.*

Per tant, el conjunt de correlacions canòniques és un interval que té la potència del continu. Les variables canòniques tenen, en el límit, variància 0, però les correlacions canòniques són positives. Una conseqüència del teorema anterior és el resultat següent.

COROL·LARI 13. *La màxima correlació per a la família CA_θ és el paràmetre θ , essent $\mathcal{H}_1$ la funció canònica, on $\mathcal{H}_1$ és la distribució de Heaviside.*

La prova que θ és la màxima correlació apareix primer a [12]. Recordem que la màxima correlació, introduïda per H. Gebelein el 1941, és l'única mesura de dependència que compleix els postulats de Rényi [55]. Per tant, θ és una bona mesura de dependència per a aquesta família.

Considerem a continuació la sèrie (36), construïda suposant que les funcions pròpies tenen norma 1. Però ara la norma és 0 i la sèrie esdevé una integral. La successió de correlacions canòniques és, en aquest context, una funció, en general contínua.

DEFINICIÓ 14. Una funció de correlació canònica és una funció integrable $f_\theta: \mathbf{I} \rightarrow \mathbf{I}$, que satisfà $\int_{\mathbf{I}} f_\theta(\rho) d\rho \leq 1$.

TEOREMA 15. *La còpula Cuadras-Augé es pot expressar en termes del desenvolupament diagonal continu*

$$CA_\theta(u, v) = uv + \int_{\max\{u,v\}}^1 \theta \rho^{1-\theta} (u/\rho)(v/\rho) d\rho.$$

En general, la família $D_\theta(u, v) = \min\{u, v\} G_\theta(\max\{u, v\})$ admet el desenvolupament integral

$$D_\theta(u, v) = uv + \int_{\max\{u,v\}}^1 f_\theta(u)(u/\rho)(v/\rho) d\rho, \quad (45)$$

on $f_\theta(\rho) = G_\theta(\rho) - \rho G'_\theta(\rho)$ és una funció de correlació canònica.

21 Còpules canòniques

El terme $\max\{u, v\}$ que apareix a (45) verifica $\max\{u, v\} = uv / \min\{u, v\}$. Doncs bé, si C és una còpula, en general

$$uv + \int_{uv/C(u,v)}^1 f_{\theta}(\rho)(u/\rho)(v/\rho) d\rho$$

també ho és. De fet Π/C es podria substituir per qualsevol quocient de dues còpules.

Encara amb més generalitat, podem obtenir la família (46), de la qual (47) seria un cas particular.

TEOREMA 16. *Sigui f_{θ} una funció de correlació canònica indexada per $\theta \in \mathbf{I}$. Definim les funcions $F_{\theta}(\rho) = \int [f_{\theta}(\rho)/\rho^2] d\rho$ i $G_{\theta}(\rho) = -\rho F_{\theta}(\rho) + \rho F_{\theta}(1) + \rho$. Per a quocients adequats Q de dues còpules, la descomposició integral (45) permet generar la família de còpules canòniques*

$$C_{\theta} = \Pi \times G_{\theta}(Q)/Q. \quad (46)$$

En particular, si $Q = \Pi/M$, obtenim la família

$$D_{\theta}(u, v) = \min\{u, v\} G_{\theta}(\max\{u, v\}), \quad (47)$$

on $G_{\theta}: \mathbf{I} \rightarrow \mathbf{I}$ és una funció creixent contínua de tal manera que $G_{\theta}(\rho)/\rho$ és decreixent dins $(0, 1]$.

f_{θ}	G_{θ}	Q	Família
0	ρ	C/Π	Independència
1	1	Π/C	Qualsevol còpula
θ	$\theta + \rho\bar{\theta}$	Π/M	Fréchet
$\theta\rho^2$	$-\theta\rho^2 + (1+\theta)\rho$	Π/AMH_1	Farlie-Gumbel-Morgenstern
$\theta\rho^2/[\bar{\theta} + \rho\theta]^2$	$-\rho/[\bar{\theta} + \rho\theta]$	FGM_{-1}/Π	Ali-Mikhail-Haq
$\theta\rho^{1-\theta}$	$\rho^{1-\theta}$	Π/M	Cuadras-Augé
$f_{\theta}(\rho)$	$G_{\theta}(\rho)$	Π/M	Durante
$\theta\rho$	$-\theta\rho \log(\rho) + \rho$	Π/W	Nova família (48)

TAULA 8: Algunes còpules generades per funcions canòniques i el desenvolupament diagonal en versió integral.

La taula 8 descriu algunes construccions, on $\bar{\theta} = 1 - \theta$. Les famílies Farlie-Gumbel-Morgenstern (FGM), Ali-Mikhail-Haq (AMH) han estat descrites abans, vegeu (24), (41). La família de Fréchet és la mitjana ponderada de la cota superior i la independència:

$$F_{\theta}(u, v) = \theta \min\{u, v\} + (1 - \theta)uv.$$

D'altra banda, D_θ de (47) va ser definida per Cuadras i Augé [18] i estudiada per Durante [34]. És interessant observar que si $Q = \Pi/M$, la densitat (44) adopta la forma simètrica

$$f_\theta(u, v) = (1 - \theta)(1/Q)^\theta \delta_{\{u \neq v\}} + \theta(Q)^{1-\theta} \delta_{\{u=v\}}.$$

Finalment, amb $f_\theta(\rho) = \theta\rho$ i $Q = \Pi/W$, on W és la còpula $W(u, v) = [u^{-2} + v^{-2} - 1]^{-1/2}$, obtenim la nova família de còpules

$$NC(u, v) = uv[1 - \theta \log(uv/W(u, v))]. \quad (48)$$

EXEMPLE. Sovint volem descriure un fet real que ha estat observat des de diverses perspectives. Com a il·lustració, sigui θ_1 la probabilitat d'un esdeveniment $\mathbb{A}$, com ara l'eficàcia d'un tractament mèdic. Realitzem n_1 proves independents i sigui $L_1(x_1, \theta_1) = \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1)^{n_1 - x_1}$ la funció de versemblança, on x_1 és la freqüència observada de $\mathbb{A}$. L'estimador màxim versemblant de θ_1 és la freqüència relativa x_1/n_1 .

Suposem ara que θ_1 és un paràmetre aleatori amb distribució uniforme $(0, 1)$. La versemblança, sota aquesta perspectiva bayesiana, és

$$L(x_1) = \int_0^1 \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1)^{n_1 - x_1} d\theta_1.$$

Considerem dos experiments independents sota els quals obtenim les freqüències relatives x_1/n_1 i x_2/n_2 per a estimar θ_1, θ_2 . Suposem que θ_1 i θ_2 són descripcions d'un metapàrmetre θ , també aleatori, que ens donaria la «veritable» probabilitat de $\mathbb{A}$ en condicions ideals. (Imaginem que un tractament mèdic s'ha estudiat en dos centres diferents.) Suposem que la distribució conjunta de (θ_1, θ) és una còpula (FGM, Plackett o una còpula canònica), i anàlogament la de (θ_2, θ) . La versemblança de θ és

$$L(x_1, x_2 | \theta) = \prod_{i=1}^2 \int_0^1 \binom{n}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i} p(\theta_i | \theta) d\theta_i,$$

on $p(\theta_i | \theta)$ és la densitat de θ_i condicionada a θ i que es calcula a partir de la còpula. La distribució *a priori* de θ és uniforme a $(0, 1)$. Aplicant el teorema de Bayes, la densitat *a posteriori* del metapàrmetre és

$$p(\theta | x_1, x_2) = \frac{L(x_1, x_2 | \theta)}{\int_0^1 L(x_1, x_2 | \theta) d\theta}.$$

Un cop tenim $p(\theta | x_1, x_2)$ podem prendre la moda (o la mitjana) com una estimació de la probabilitat θ ; vegeu [50].

Resumint, la teoria de còpules, i en particular les còpules canòniques, defineixen una classe àmplia de models probabilístics que, correctament utilitzats, descriuen i sintetitzen l'associació estadística d'un conjunt de dades bivariants.

Referències

- [1] AITCHISON, J.; GREENACRE, M. «Biplots of compositional data». *J. Roy. Statist. Soc. Ser. C*, 51 (4) (2002), 375–392.
- [2] ASH, R. B. *Information Theory*. Nova York: Dover Publications, Inc., 1990. Reedició corregida de l'original de 1965.
- [3] BEARE, B. K. «A generalization of Hoeffding's lemma, and a new class of covariance inequalities». *Statist. Probab. Lett.*, 79 (5) (2009), 637–642.
- [4] BENZÉCRI, J.-P. *L'analyse des données. II. L'analyse des correspondances*. 2a ed. París: Dunod, 1976.
- [5] BERNARDO, J. M. «Una introducció a l'estadística bayesiana». *Butlletí de la Societat Catalana de Matemàtiques*, 17 (1) (2002), 7–64.
- [6] CHERUBINI, U.; MULINACCI, S.; GOBBI, F.; ROMAGNOLI, S. *Dynamic copula methods in finance*. Nova York: Wiley, 2011.
- [7] COX, T. F.; COX, M. A. A. *Multidimensional Scaling*. Londres: Chapman & Hall, 1994. (Monographs on Statistics and Applied Probability; 59)
- [8] CUADRAS, C. M. *Métodos de Análisis Multivariante*. Barcelona: PPU, 1991.
- [9] CUADRAS, C. M. «Interpreting an inequality in multiple regression». *Amer. Statist.*, 47 (4) (1993), 256–258.
- [10] CUADRAS, C. M. «Increasing the correlations with the response variable may not increase the coefficient of determination: a PCA interpretation». A: *New Trends in Probability and Statistics*. Vol. 3. Utrecht: VSP, 1995, 75–83.
- [11] CUADRAS, C. M. «On the covariance between functions». *J. Multivariate Anal.*, 81 (1) (2002), 19–27.
- [12] CUADRAS, C. M. «Correspondence analysis and diagonal expansions in terms of distribution functions». *J. Statist. Plann. Inference*, 103 (1–2) (2002), 137–150.
- [13] CUADRAS, C. M. «Continuous canonical correlation analysis». *Research Letters in the Information and Mathematical Sciences*, 8 (2005), 97–103.
- [14] CUADRAS, C. M. «Distance-based association and multi-sample tests for general multivariate data». A: *Advances in Mathematical and Statistical Modeling*. Boston: Birkhäuser, 2008, 61–71. (Stat. Ind. Technol.)
- [15] CUADRAS, C. M. *Nuevos Métodos de Análisis Multivariante*. Barcelona: CMC Editions, 2014.
- [16] CUADRAS, C. M. «Nonlinear principal and canonical directions from continuous extensions of multidimensional scaling». *Open J. Stat.*, 4 (2) (2014), 132–149.
- [17] CUADRAS, C. M. «Contributions to the diagonal expansion of a bivariate copula with continuous extensions». *J. Multivariate Anal.*, 139 (2015), 28–44.

- [18] CUADRAS, C. M.; AUGÉ, J. «A continuous general multivariate distribution and its properties». *Comm. Statist. A—Theory Methods*, 10 (4) (1981), 339–353.
- [19] CUADRAS, C. M.; CUADRAS, D. «A parametric approach to correspondence analysis». *Linear Algebra Appl.*, 417 (1) (2006), 64–74.
- [20] CUADRAS, C. M.; CUADRAS, D. «Eigenanalysis on a bivariate covariance kernel». *J. Multivariate Anal.*, 99 (10) (2008), 2497–2507.
- [21] CUADRAS, C. M.; CUADRAS, D. «Partitioning the geometric variability in multivariate analysis and contingency tables». A: *Classification and Multivariate Analysis for Complex Data Structures*. Heidelberg: Springer, 2011, 237–244. (Stud. Classification Data Anal. Knowledge Organ.)
- [22] CUADRAS, C. M.; CUADRAS, D. «A unified approach for the multivariate analysis of contingency tables». *Open J. Stat.*, 5 (2015), 223–232.
- [23] CUADRAS, C. M.; CUADRAS, D.; GREENACRE, M. J. «A comparison of different methods for representing categorical data». *Comm. Statist. Simulation Comput.*, 35 (2) (2006), 447–459.
- [24] CUADRAS, C. M.; CUADRAS, D.; LAHLOU, Y. «Principal directions of the general Pareto distribution with applications». *J. Statist. Plann. Inference*, 136 (8) (2006), 2572–2583.
- [25] CUADRAS, C. M.; DÍAZ, W. «Another generalization of the bivariate FGM distribution with two-dimensional extensions». *Acta Comment. Univ. Tartu. Math.*, 16 (1) (2012), 3–12.
- [26] CUADRAS, C. M.; FORTIANA, J. «A continuous metric scaling solution for a random variable». *J. Multivariate Anal.*, 52 (1) (1995), 1–14.
- [27] CUADRAS, C. M.; FORTIANA, J. «Visualizing categorical data with related metric scaling». A: BLASIUS, J.; GREENACRE, M. (ed.). *Visualization of Categorical Data*. Nova York: Academic Press, 1998, 365–376.
- [28] CUADRAS, C. M.; FORTIANA, J. «The importance of geometry in multivariate analysis and some applications». A: *Statistics for the 21st Century*. Nova York: Dekker, 2000, 93–108. (Statist. Textbooks Monogr.; 161)
- [29] CUADRAS, C. M.; FORTIANA, J.; GREENACRE, M. «Continuous extensions of matrix formulations in correspondence analysis, with applications to the FGM family of distributions». A: HEIJMANS, R. D. H.; POLLOCK, D. S. G.; SATORRA, A. (ed.). *Innovations in Multivariate Statistical Analysis*. Dordrecht: Kluwer Ac. Publ., 2000, 101–116.
- [30] CUADRAS, C. M.; FORTIANA, J.; OLIVA, F. «The proximity of an individual to a population with applications in discriminant analysis». *J. Classification*, 14 (1) (1997), 117–136.
- [31] CUADRAS, C. M.; VALERO, S.; CUADRAS, D.; SALEMBIER, P.; CHANUSSOT, J. «Distance-based measures of association with applications in relating hyperspectral images». *Comm. Statist. Theory Methods*, 41 (13–14) (2012), 2342–2355.

- [32] DOBROWOLSKI, E.; KUMAR, P. «Some properties of the Marshall-Olkin and generalized Cuadras-Augé families of copulas». *Aust. J. Math. Anal. Appl.*, 11 (1) (2014), 13 p.
- [33] DONNER, A.; WELLS, G. «A comparison of confidence interval methods for the intraclass correlation coefficient». *Biometrics*, 42 (2) (1986), 401–412.
- [34] DURANTE, F. «A new family of symmetric bivariate copulas». *C. R. Math. Acad. Sci. Paris*, 344 (3) (2007), 195–198.
- [35] ESCOPIER, B.; PAGÈS J. *Analyses factorielles simples et multiples*. París: Dunod, 1990.
- [36] FRÉCHET, M. «Sur les tableaux de corrélation dont les marges sont données». *Ann. Univ. Lyon. Sect. A. (3)*, 14 (1951), 53–77.
- [37] GALTON, F. «Regression towards mediocrity in hereditary stature». *J. of the Anthropological Institute*, 15 (1886), 246–263.
- [38] GOODMAN, L. A. «Correspondence analysis, association analysis, and generalized nonindependence analysis of contingency tables: Saturated and un saturated models, and appropriate graphical displays». A: CUADRAS, C. M.; RAO, C. R. (ed.). *Multivariate Analysis: Future Directions 2*. Amsterdam: Elsevier, 1993, 265–294.
- [39] GREENACRE, M. J. *Theory and Applications of Correspondence Analysis*. Londres: Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], 1984.
- [40] HAMILTON, D. «Sometimes $R^2 > r_{yx_1}^2 + r_{yx_2}^2$. Correlated variables are not always redundant». *Amer. Statist.*, 41 (2) (1987), 129–132.
- [41] HANLEY, J. A. «“Transmuting” women into men: Galton’s family data on human stature». *Amer. Statist.*, 58 (3) (2004), 237–243.
- [42] HARRIS, H.; SMITH, C. A. B. «The sib-sib age of onset correlation among individuals suffering from a hereditary syndrome produced by more than one gene». *Ann. Eugenics*, 14 (1949), 309–318.
- [43] HOEFFDING, W. «Masstabinvariante Korrelationstheorie». *Schriften Math. Inst. Univ. Berlin*, 5 (1940), 181–233.
- [44] HOTELLING, H. «Relations between two sets of variates». *Biometrika*, 28 (3–4) (1936), 321–377.
- [45] JULIÀ, O.; COMAS, J.; VIVES-REGO, J. «Second-order functions are the simplest correlations between flow cytometric light scatter and bacterial diameter». *Journal of Microbiological Methods*, 40 (1) (2000), 57–61.
- [46] KAPLAN, J. «A statistical error in The Bell Curve». *Chance*, 10 (1997), 20–21.
- [47] LANCASTER, H. O. «The structure of bivariate distributions». *Ann. Math. Statist.*, 29 (1958), 719–736.
- [48] LI, B.; GENTON, M. G. «Nonparametric identification of copula structures». *J. Amer. Statist. Assoc.*, 108 (502) (2013), 666–675.
- [49] MAI, J.-F.; SCHERER, M. «Efficiently sampling exchangeable Cuadras-Augé copulas in high dimensions». *Inform. Sci.*, 179 (17) (2009), 2872–2877.

- [50] MORENO, E.; VÁZQUEZ-POLO, F. J.; NEGRÍN, M. A. «Objective Bayesian meta-analysis for sparse discrete data». *Stat. Med.*, 33 (21) (2014), 3676–3692.
- [51] NELSEN, R. B. *An Introduction to Copulas*. 2a ed. Nova York: Springer, 2006. (Springer Series in Statistics)
- [52] PAPADATOS, N.; XIFARA, T. «A simple method for obtaining the maximal correlation coefficient and related characterizations». *J. Multivariate Anal.*, 118 (2013), 102–114.
- [53] PEARSON, K.; LEE, A. «On the laws of inheritance in man: I. Inheritance of physical characters». *Biometrika*, 2 (1903), 357–462.
- [54] RAO, C. R. «A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance». *Qüestió* (2), 19 (1–3) (1995), 23–63.
- [55] RÉNYI, A. «On measures of dependence». *Acta Math. Acad. Sci. Hungar.*, 10 (1959), 441–451.
- [56] ROUTLEDGE, R. D. «When stepwise regression fails: correlated variables some of which are redundant». *Internat. J. Math. Ed. Sci. Tech.*, 21 (3) (1990), 403–410.
- [57] SCHEFFÉ, H. *The Analysis of Variance*. Nova York: John Wiley & Sons; Londres: Chapman & Hall, 1959.
- [58] WACHSMUTH, A.; WILKINSON, L.; DALLAL, G. E. «Galton's bend: a previously undiscovered nonlinearity in Galton's family stature regression data». *Amer. Statist.*, 57 (3) (2003), 190–192.
- [59] WALLER, N. G. «The geometry of enhancement in multiple regression». *Psychometrika*, 76 (4) (2011), 634–649.
- [60] WILKS, S. S. «Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution». *Ann. Math. Statistics*, 17 (1946), 257–281.

DEPARTAMENT D'ESTADÍSTICA
FACULTAT DE BIOLOGIA
UNIVERSITAT DE BARCELONA
AV. DIAGONAL, 643
08028 BARCELONA
ccuadras@ub.edu

Eleccions mitjançant el vot d'aprovació. El mètode de Phragmén i algunes variants

XAVIER MORA I MARIA OLIVER

Resum: S'estudien diverses variants d'un mètode de representació proporcional que Edvard Phragmén va proposar a finals del segle XIX. Cada elector expressa la seva opinió mitjançant un vot d'aprovació, on indica totes les opcions que li semblen adients. Les opcions poden ser candidats individuals, com suposava Phragmén, o bé candidatures de partit que ofereixen múltiples representants cadascuna. Donat un conjunt de vots d'aquest tipus i un nombre prefixat d'escons, es tracta de repartir aquests escons entre les diferents candidatures de manera que el resultat sigui el més representatiu possible. El present estudi presta una atenció especial a una propietat de monotonia que demana que l'addició d'aprovacions a favor d'una candidatura no pugui produir una disminució del nombre d'escons que li són assignats. En la variant bàsica, on els escons s'assignen mitjançant un procediment seqüencial, aquesta monotonia es compleix en el cas de candidats individuals, però no en el cas de candidatures de partit. També s'exploren breument unes variants directes, on els escons s'assignen tots d'una vegada mitjançant algun criteri adient d'optimització. Aquestes variants semblen millors des del punt de vista de la monotonia, però presenten un problema de multiplicitat de solucions que fa necessari algun criteri addicional de selecció.

Paraules clau: representació proporcional, vot d'aprovació, llistes obertes, Edvard Phragmén.

Classificació MSC2010: 91B12.

1 Introducció

Suposem que un grup de 1000 individus vol elegir quatre persones que representin tot el grup, incloent-hi no solament les opinions i els punts de vista majoritaris, sinó també, en la deguda proporció i en la mesura que sigui possible, totes les altres opinions i els altres punts de vista. Es compta amb vuit candidats, a, b, c, d, e, f, g, h, i es demana a cada membre del grup que indiqui

Aquest article és una versió més elaborada del treball de fi de grau de Matemàtiques realitzat el curs 2011–2012 a la Universitat Autònoma de Barcelona per Maria Oliver sota la direcció de Xavier Mora.

quins d'aquests candidats li semblen acceptables per a reflectir la seva opinió. Imaginem que les respostes dels 1000 individus es reparteixen de la manera següent:

250 : a, b, c, d; 250 : b, c, e; 200 : e, f, g, h; 200 : a, e, f; 100 : a, c, e, f. (1)

Quins són els quatre representants més adequats?

Aquest és el problema que es planteja en l'elecció de representants mitjançant llistes obertes no ordenades. En aquest article ens ocuparem d'ell junt amb una generalització en què les opcions sobre les quals s'expressa l'elector no són els candidats individuals, sinó unes candidatures col·lectives cadascuna de les quals pot proporcionar múltiples representants. En aquest escenari també pot passar que l'elector trobi acceptable més d'una d'aquestes candidatures. Suposem, per exemple, que un país vol elegir un parlament de 135 diputats i que es demana a cada elector que indiqui quines de les candidatures que presenten els partits li semblen prou oportunes des del seu punt de vista. Cada candidatura ofereix tants candidats com escons hi ha al parlament. Imaginem que els electors voten com segueix, on A, B, C, D, E, F, G denoten les candidatures corresponents als diferents partits i les xifres indiquen centenars de milers d'electors:

6 : A; 8 : A, B; 2 : B, C; 2 : B, C, D; 3 : C, D;
3 : C, E; 2 : E; 4 : F; 2 : F, G; 2 : G. (2)

Quants diputats s'han d'assignar a cada candidatura?

Aquesta forma de vot en què l'elector indica totes les opcions que li semblen adients s'anomena **vot d'aprovació**. Òbviament, un vot d'aquest tipus dona més informació que no pas un vot on només es permet indicar una sola opció. A conseqüència d'això, els resultats poden reflectir millor les veritables opinions dels electors (que ho aconsegueixin o no dependrà de l'algorisme utilitzat). Una altra conseqüència molt positiva és que s'evita en certa mesura el problema del vot útil: l'elector no té res a perdre si inclou en el seu vot les opcions que realment prefereix, possiblement acompanyades d'altres opcions menys preferides però amb més possibilitats d'èxit.¹

En la forma tradicional de vot que demana a cada elector una sola opció, habitualment s'accepta també com a vàlid el vot en blanc. Tanmateix, el significat d'aquest últim no és gens clar: en principi, tant pot significar el rebuig de totes les opcions plantejades, com l'acceptació de qualsevol d'elles. Amb el vot d'aprovació, en canvi, aquests dos significats oposats queden perfectament separats: el primer cas correspon a no aprovar cap opció, mentre que el segon correspon a aprovar-les totes.

En lloc del vot d'aprovació, es pot plantejar d'usar el **vot preferencial**, on a més d'indicar les opcions que li semblen adients, l'elector les ordena també per ordre de preferència. De fet, el vot preferencial en termes de candidats individuals s'utilitza per a l'elecció de representants en el mètode del **vot**

¹ Les notes les posem en paràgrafs sagnats amb un cos de lletra més petit, la qual cosa facilita de continuar la lectura.

únic transferible. Com veurem més avall, però, aquest mètode presenta certs problemes —de manca de monotonia respecte al contingut dels vots— que fan interessant d'estudiar unes altres vies.

En principi, també serien possibles altres formes de vot que no considerarem aquí, com ara que també es puguin expressar preferències entre les opcions no aprovades, o bé que només s'expressin preferències entre les diverses opcions, sense especificar quines d'aquestes compten amb l'aprovació de l'elector. En certs contextos, el terme «vot preferencial» es refereix a aquesta última forma de vot. Però en el context de l'elecció de representants és habitual entendre aquest terme en el sentit que hem dit, és a dir, com un vot d'aprovació junt amb una ordenació de les opcions aprovades.

Més concretament, aquest article girarà al voltant d'un mètode que va ser proposat a finals del segle XIX pel matemàtic suec Edvard Phragmén. El mètode en qüestió combina idees del vot únic transferible i del mètode de D'Hondt, la qual cosa ens portarà a revisar abans aquests antecedents.

Tal com ja hem dit més amunt, ens situarem en un marc general en què les opcions que aprova l'elector no són necessàriament candidats individuals, sinó que poden ser candidatures col·lectives a cadascuna de les quals es tracta d'assignar un cert nombre de representants. De fet, aquest treball està especialment motivat per aquesta possibilitat, la qual seria adient en eleccions d'àmbit extens, on els electors difícilment coneixen directament els candidats, i també en escenaris on domina la disciplina de partit. La proposta que les opcions sobre les quals s'expressi l'elector siguin les candidatures col·lectives presentades pels partits també ha estat feta recentment per altres autors en el context del vot preferencial [23].

L'article està estructurat de la manera següent: Les seccions 2–6 introdueixen les diverses idees i antecedents que ens interessen. En la secció 7 estudiem detalladament la variant bàsica del mètode de Phragmén, la qual treballa amb vots d'aprovació i assigna els escons mitjançant un procediment seqüencial. Tal com hem dit, però, ens situem en el marc general de candidatures col·lectives. Una mica inesperadament, veurem que en el cas general falla una propietat de monotonia que sí que es compleix en el cas de candidatures individuals. En la secció 8 explorem breument les variants directes, on els escons s'assignen tots d'una vegada mitjançant algun criteri adient d'optimització. Segons comentarem, hi ha indicis que aquestes variants sí que tenen la propietat de monotonia. A canvi, però, ens trobem amb un problema de multiplicitat de solucions que fa necessari algun criteri addicional de selecció. L'article se suplementa amb dos apèndixs que recullen, respectivament, una breu nota biogràfica sobre Edvard Phragmén i una classificació de les diverses variants que aquest autor va considerar del seu mètode.

2 Representació proporcional amb partits o sense

2.1 L'objectiu general de la representació proporcional és elegir un conjunt de representants on les diverses opinions i punts de vista estiguin presents en

la mateixa proporció que a tota la societat [24, 51]. El 1789, mesos abans de la Revolució Francesa, Honoré Gabriel Riquetti, comte de Mirabeau, ho comparava amb la representació que un mapa fa d'un territori, on certament convé que les longituds i àrees del mapa siguin tan proporcionals a les del territori com sigui possible [24, § 117].

Per intentar assolir aquest objectiu, sovint es demana als electors que escullin una, i només una, de les candidatures presentades pels partits polítics. Fet això, es tracta de repartir els escons del parlament entre aquestes candidatures assignant a cadascuna d'elles un nombre d'escons tan proporcional als vots obtinguts com sigui possible. Com que els nombres en qüestió són enters, les desviacions respecte a la proporcionalitat exacta són inevitables. Aquestes desviacions esdevenen especialment importants quan el nombre d'escons que es reparteixen és petit, com passa sovint com a resultat d'una divisió en circumscripcions territorials. Tot plegat dona lloc a una diversitat de regles de repartiment, entre les quals destaquen la de Jefferson i D'Hondt i la de Webster i Sainte-Laguë, que seran discutides amb més detall a § 5.

Tanmateix, tot això suposa que cada elector s'identifica amb una i només una de les propostes que presenten els partits. Aquesta hipòtesi pot estar molt allunyada de la realitat, la qual cosa pot comportar resultats desencertats.

2.2 En aquest article ens interessem en procediments alternatius basats en formes més obertes de vot que no requereixen una classificació prèvia dels candidats en classes disjunts. Tal com veurem, tot i això encara es pot mirar de respectar el principi de representació proporcional.

Tanmateix, no serveix pas qualsevol procediment. Per exemple, suposem que es volen elegir n representants a partir d'un cert conjunt prou ampli de candidats. Un possible procediment és demanar a cada elector que doni una llista dels n candidats que més li agraden i declarar elegits els n candidats que obtinguin més vots.

Aquest procediment pot semblar força natural, però el cert és que és molt lluny de l'ideal de la representació proporcional. Suposem, per exemple, que $n = 4$ i que els vots són com mostrem a continuació:

$$75 : a, b, c, d; \quad 25 : e, f, g, h. \quad (3)$$

És a dir, 75 electors coincideixen a seleccionar els candidats a, b, c, d , mentre que 25 electors coincideixen a seleccionar els candidats e, f, g, h . Aplicant el procediment que hem dit, resultarien elegits els candidats a, b, c, d . Tanmateix, per la manera en què s'ha votat és obvi que l'electorat està clarament dividit en dos sectors que tenen una presència respectiva del 75 % i 25 %. En canvi, en el conjunt escollit aquests dos sectors tenen una presència respectiva del 100 % i 0 %.

Certament, l'ideal de la representació proporcional demana elegir com a representants tres candidats del conjunt $X = \{a, b, c, d\}$ i un del conjunt $Y = \{e, f, g, h\}$, ja que $75/25 = 3/1$.

2.3 En l'exemple precedent, el fet que es reconegui la similitud entre els elements de X per una banda i els de Y per una altra és possible perquè els electors han seleccionat més d'un candidat. Si només n'haguessin seleccionat un, llavors no hi hauria cap pista sobre la relació de similitud entre candidats.

A més a més, per tal de capturar aquesta relació de la millor manera possible, és preferible no forçar l'elector a donar cap nombre concret de candidats. Per exemple, si hi ha 75 electors que s'identifiquen amb qualsevol dels candidats del conjunt $X' = \{a, b, c, d, e, f\}$ i cap altre, i 25 electors que s'identifiquen amb qualsevol dels de $Y' = \{g, h\}$ i cap altre, llavors el fet de demanar-los que seleccionin exactament quatre candidats no permetrà als partidaris de X' incloure tots els seus preferits, alhora que obligarà els partidaris de Y' a incloure algun element de X' en la seva llista, la qual cosa desdibuixarà la partició que hi ha realment entre X' i Y' . En canvi, si se'ls permet seleccionar un nombre arbitrari de candidats, llavors la votació pot posar de manifest aquesta partició. Evidentment, perquè quedi reflectida de la millor manera possible cal que cada elector inclogui tots els candidats que li semblen bé. És per això que *en el vot d'aprovació és essencial permetre que l'elector indiqui totes les opcions que li semblen adients*. Per descomptat, també cal que realment existeixi una partició del tipus que estem dient. Tot això és aplicable també a una partició en més de dues classes.

En l'exemple (3) és molt clara la divisió en dos punts de vista diferents. Però en general pot passar com a (1), on la relació de similitud entre candidats no és pas tan clara. Com ho hem de fer en un cas així?

3 El vot únic transferible

3.1 El mètode més conegut que aplica la idea de representació proporcional sense requerir als electors que es pronunciïn per un sol partit és el **vot únic transferible**. Aquest mètode utilitza el vot preferencial (en el sentit que hem dit a § 1, és a dir, un vot d'aprovació junt amb una ordenació de les opcions aprovades). En principi, les opcions que es consideren són els candidats individuals, independentment de si pertanyen a un partit o un altre, i de fet, independentment de si estan definits uns partits o no. Tot i això, s'aconsegueix un efecte de representació proporcional, la qual cosa vol dir que aquest mètode detecta d'alguna manera les similituds entre candidats a partir dels vots dels electors.

El mètode del vot únic transferible s'associa normalment a l'advocat anglès Thomas Hare, que el va proposar el 1857 i va obtenir un gran ressò a través del filòsof i aleshores membre del parlament anglès John Stuart Mill. De tota manera, l'any anterior havia estat utilitzat ja a Dinamarca, on el matemàtic danès Carl Christofer Georg Andræ, llavors ministre de Finances del seu país, l'havia proposat independentment. I encara abans, la mateixa idea l'havia utilitzat el 1840 a Austràlia Rowland Hill, i fins i tot el 1819 a Anglaterra Thomas Wright Hill, pare de l'anterior.

El mètode del vot únic transferible parteix del principi que cada diputat hauria de representar el mateix nombre d'electors. Aquest nombre d'electors per diputat s'anomena **quota** i nosaltres l'anomenem q . El seu valor es fixa des del principi i s'usa de la manera següent: mentre quedin més candidats que escons per assignar, el que determina que un candidat sigui elegit és que reuneixi —en el sentit que de seguida concretarem— un nombre de vots superior, o en tot cas igual, a la quota q .

D'ara endavant w i n representen, respectivament, el nombre d'electors i el nombre de representants que s'han d'elegir. Hare i Andraë prenen com a quota q el quocient $w/n =: q_0$ i l'objectiu era superar o igualar aquest valor. El 1868, però, l'advocat anglès Henry Richmond Droop va proposar de prendre q igual a $w/(n+1) =: q_D$ i requerir que el nombre de vots superi estrictament aquest valor. El valor q_0 rep el nom de **quota simple**, o **quota de Hare**, i q_D es coneix com a **quota de Droop**. Noteu que en el cas extrem $n = 1$, és a dir, l'elecció d'un sol representant comú per a tothom, la quota de Hare correspon a la noció d'unanimitat, mentre que la de Droop correspon a la noció de majoria absoluta.

La manera de reunir els vots requerits consisteix a figurar com a primera opció en la butlleta d'un elector, o alternativament, figurar-hi per sota d'altres candidats que ja han obtingut un escó o que no poden arribar a obtenir-ne cap. Tan bon punt un candidat assoleix la quota, se li assigna un escó i cadascun dels vots sobrants és transferit al candidat que el segueix en la butlleta. Quan no hi ha cap candidat que assoleixi la quota, llavors s'elimina el que en aquell moment té menys vots; si en aquest moment resten tants candidats com escons, aleshores aquests candidats resulten elegits independentment dels vots amb què comptin i finalitza el procediment; altrament, cadascun dels vots del candidat eliminat és transferit al candidat següent en la butlleta. En l'apartat que segueix ho il·lustrem amb un parell d'exemples.

Per ser exactes, Droop prenia com a quota a superar o igualar el primer enter N estrictament superior a $w/(n+1)$. Que un nombre enter (de vots) superi o iguali aquest valor N és exactament equivalent a superar estrictament el valor $w/(n+1)$. Tanmateix, amb els nombres fraccionaris de vots que sorgiran més avall, aquestes dues condicions no són exactament equivalents. D'altra banda, aquesta distinció també afecta la quantitat exacta de vots que es transferiran. Sobre aquestes qüestions, seguirem Irwin Mann en requerir que se superi estrictament la quota $w/(n+1)$ i al mateix temps transferir tot l'excés sobre aquesta quantitat [56, p. 271-272].

3.2 Vegem amb més detall un parell d'exemples. Considerem primer un cas similar a (3) però amb vots preferencials:

$$75 : a > b > c > d; \quad 25 : e > f > g > h. \quad (4)$$

Suposem que volem elegir $n = 4$ representants. La quota de Droop és $100/(4+1) = 20$. Per tant, resulten elegits directament tant a com e . En el cas de a sobren 55 vots, els quals passen a b ; aquest resulta elegit i encara sobren 35 vots, els

quals passen a c; aquest també resulta elegit, però els 15 vots que ara sobren i passen a d no són suficients perquè aquest últim candidat obtingui un escó. En el cas de e, només sobren 5 vots, els quals passen a f però també són insuficients per a donar-li un escó. En total, ha resultat elegit el conjunt de representants següent: {a, b, c, e}, que inclou tres membres del conjunt $X = \{a, b, c, d\}$ i un del conjunt $Y = \{e, f, g, h\}$, tal com demana la noció de proporcionalitat.

Considerem ara l'exemple següent, on hi ha 96 electors:

$$34 : a > b > c > d; \quad 20 : b > a > d > c; \quad 21 : c > d > a > b; \quad 21 : d > c > b > a. \quad (5)$$

Suposem que volem elegir $n = 3$ representants. La quota de Droop és $96/(3 + 1) = 24$. Per tant, resulta elegit directament el candidat a. Els 10 vots que li sobren a a passen a b, que resulta elegit i transfereix 6 vots; com que 10 dels 30 vots que havia obtingut b donaven c com a opció subseqüent i els altres 20 donaven a, ja elegit, i després d, és raonable mantenir la proporcionalitat i passar 2 dels 6 vots sobrants a c i els altres 4 a d. Com a resultat d'això queda elegit d, però no c. Per tant, resulta elegit el conjunt de representants següent: {a, b, d}. Noteu que aquest conjunt inclou dos membres del conjunt $X = \{a, b\}$ i només un del conjunt $Y = \{c, d\}$, la qual cosa està d'acord amb el fet que els vots (5) es poden veure en la forma

$$54 : X > Y; \quad 42 : Y > X. \quad (6)$$

En canvi, si ens haguéssim guiat només per les primeres opcions dels vots (5) haguéssim elegit només un membre de X i dos de Y.

En considerar altres exemples, o petites variacions dels precedents, se susciten de seguida certes qüestions secundàries que requereixen regles addicionals en les quals no entrarem i sobre les quals hi ha diverses variants.

3.3 En el cas especial $n = 1$ —un sol representant comú per a tots— el vot únic transferible pren la forma que descrivim a continuació, la qual es coneix com a **vot alternatiu** [56, p. 193–195].

Si un candidat figura com a primera opció en més de la meitat dels vots, aleshores resulta elegit. En cas contrari, s'elimina el candidat que té menys vots i cadascun d'aquests vots es transfereix al candidat que ocupa el lloc següent de la llista en el vot. Fet això, es torna a mirar si algun candidat reuneix més de la meitat dels vots, en el qual cas resulta elegit. Si no n'hi ha cap que compleixi aquesta condició, aleshores es procedeix novament a eliminar el que té menys vots i a transferir els seus vots. Aquest procés es repeteix successivament fins que resulti elegit un candidat. A tot això cal afegir que en cada volta es revisa el nombre total de vots per tal de comptar només els que encara contenen algun candidat no eliminat.

Aquest procediment es coneix també com a «doble volta instantània» (*instant runoff voting*), però aquest nom no és gaire encertat, ja que en general poden ser necessàries més de dues voltes.

3.4 Tot i que el procediment del vot únic transferible està clarament inspirat en la idea de representació proporcional, segons com es miri, la seva proporcionalitat no és gens clara: què és proporcional a què? La regla de Jefferson i D'Hondt i altres regles similars de repartiment responen clarament a un intent que el nombre d'escons assignats a un partit sigui tan proporcional al nombre de vots obtinguts com sigui possible. Però aquí ni tan sols estan definits els partits. I encara que ho estiguessin, tampoc estaria gens clar què vol dir el nombre de vots obtinguts.

Tanmateix, en certes situacions s'evidencia l'existència d'una propietat realment relacionada amb la proporcionalitat. Concretament, es tracta del cas en què un determinat nombre v d'electors coincideixen a encapçalar la seva llista de preferències per un mateix conjunt X de candidats, en un ordre o un altre. Aquests vots es poden interpretar com a favorables al conjunt X . Segons el **criteri de proporcionalitat de Droop** —vegeu, per exemple, [36, § 3.1]—, si $v > mq_D$ amb m enter, llavors haurien d'obtenir un escó almenys m candidats del conjunt X , sempre que aquest conjunt contingui efectivament tants candidats. El compliment d'aquest criteri per part del vot únic transferible el va posar en relleu el 1984 Michael Dummett [15, p. 282–283] (vegeu també [58, 55]).

L'exemple següent, tret de [24, § 204], mostra l'interès d'aquesta propietat: S'han d'elegir cinc representants a partir dels 100 vots següents: $51 : a > b > c > d > e$; $17 : f > g > h > i > j$; $17 : g > f > h > i > j$; $15 : h > f > g > i > j$. Aquests vots es poden veure en la forma $51 : X$; $49 : Y$ amb $X = \{a, b, c, d, e\}$ i $Y = \{f, g, h, i, j\}$, de manera que l'ideal de la representació proporcional demana tres representants del conjunt X i dos del conjunt Y . La quota de Droop val $q_D = 100/6 = 50/3$, de manera que es compleix $51 > 3q_D$ i $49 > 2q_D$. D'acord amb la propietat que acabem de comentar, el vot únic transferible basat en la quota de Droop dóna el resultat desitjat. En canvi, la quota simple $q_0 = w/n = 20$ dóna dos representants de X i tres de Y .

3.5 El vot únic transferible s'utilitza sobretot en països de la *Commonwealth*. En la seva versió més tradicional, que s'usa per exemple en les eleccions al parlament irlandès [27], les transferències de vot entre candidats es fan materialment, transportant les butlletes d'una banda a l'altra. Això posa en evidència el fet que cada elector queda assignat a un diputat en concret.

Tanmateix, té l'inconvenient que, en general, segons quines butlletes són transferides, pot sortir elegit un candidat o un altre (vegeu, per exemple, [56, p. 272–273]). Aquest problema s'arregla en certes versions més elaborades on la representació d'un elector pot quedar distribuïda entre diversos diputats. Un mètode d'aquest tipus és utilitzat des de 1984 per la *Royal Statistical Society* a l'hora d'elegir els membres del seu *Council* [22].

3.6 Malgrat tot això, el vot únic transferible té un defecte gairebé inacceptable. Suposem que amb uns vots concrets resulta seleccionat un determinat candidat x . Imaginem que aquests vots experimenten un canvi que només actua en el sentit d'afavorir aquest candidat x . Tothom estarà d'acord que després

d'aquest canvi hauria de continuar seleccionat x . D'aquesta propietat en direm **monotonia** (respecte al contingut dels vots). Doncs bé, el vot únic transferible no la compleix! Tot seguit en donarem un exemple amb $n = 1$. Segons el que hem vist més amunt, això significa que aquest defecte està present ja en el conegut mètode de la doble volta per a l'elecció d'un sol representant, el qual s'utilitza, per exemple, en les eleccions presidencials franceses.

Vet aquí un exemple. S'ha d'elegir un sol representant i hi ha tres candidats a , b , c . Suposem que els electors expressen les preferències següents:

$$36 : a > b > c; \quad 30 : c > a > b; \quad 24 : b > c > a; \quad 10 : b > a > c. \quad (7)$$

La quota de Droop és $100/2 = 50$. Com que no hi ha cap candidat que la superi amb les primeres preferències dels electors, es passa a eliminar el que en té menys, que resulta ser el candidat c , amb 30 vots. Aquests vots passen a a , que en reuneix 66 i, per tant, resulta elegit. Suposem ara que els darrers deu electors canvien d'opinió en sentit favorable a a : en lloc de $b > a > c$, ara voten $a > b > c$. Aquest canvi també porta a l'eliminació d'un candidat, però ara el que resulta eliminat és b , amb 24 vots, i aquests vots passen a c , que en reuneix 54 i, per tant, resulta elegit. Així doncs, el candidat a ha deixat de ser elegit tot i que els vots han variat clarament a favor seu.

Aquest defecte del vot únic transferible és assenyalat a [13] i [7, p. 38] (vegeu també [59]). Davant de mancances inesperades com aquesta, els defensors d'un mètode —en aquest cas el vot únic transferible— solen argumentar que es tracta de fets molt poc freqüents. De tota manera, el que acabem de veure no hauria de passar mai. Si una variació clarament favorable a un determinat candidat fa que aquest deixi de sortir elegit, vol dir que almenys un dels dos resultats —abans o després de la variació— és incorrecte.

Arran d'aquests fenòmens, Michael Dummett, el mateix que el 1984 havia posat en relleu el compliment del criteri de proporcionalitat de Droop, afirmava el 1992 que el vot únic transferible és el segon mètode més dolent després del vot uninominal, on només es permet indicar una sola opció [16, p. 110-111].

Respecte a aquesta qüestió és molt interessant la contribució feta el 1987 per Douglas Woodall [57]. Aquest autor va demostrar la impossibilitat matemàtica d'un mètode basat en el vot preferencial que combini les propietats de proporcionalitat i monotonia que estem considerant amb unes altres que també són presents en el vot únic transferible. La primera d'aquestes propietats addicionals demana simplement que si els electors es limiten a indicar una sola opció, resultin elegides les opcions més votades. L'altra requereix que les opcions menys preferides d'una butlleta mai no ajudin ni perjudiquin les més preferides. Aquesta última condició és prou desitjable, ja que altrament els electors es poden veure conduïts a ometre o falsejar les seves veritables preferències. La demostració que dona Woodall de la impossibilitat esmentada consisteix simplement a posar un exemple concret on l'aplicació de les diverses propietats permet arribar a una contradicció.

Per a més detalls sobre el mètode del vot únic transferible, remetem el lector a [56, p. 267-279].

4 Candidatures collectives

4.1 Tal com hem dit en la introducció, en el que segueix adoptem un marc general on les opcions sobre les quals s'expressa l'elector són conjunts de candidats individuals. Aquests conjunts, que anomenarem **candidatures**, són els mateixos per a tots els electors i poden estar reduïts a un sol candidat, en el qual cas es recupera el plantejament original.

Les candidatures collectives poden ser fixades en el plantejament de la votació. Aquest seria el cas de les llistes tancades que presenten els partits o coalicions electorals, llevat que aquí tenim la intenció que l'elector no hagi de limitar el seu vot a escollir-ne només una.

Tanmateix, les candidatures collectives també poden sorgir de manera espontània quan es vota en termes de candidats individuals. Així, en la votació (3) es poden veure com a candidatures collectives els conjunts $X = \{a, b, c, d\}$ i $Y = \{e, f, g, h\}$. En canvi, en la votació

$$65 : a, b, c, d; \quad 15 : a, f, h; \quad 20 : e, f, g, h \quad (8)$$

hauríem de prendre $X1 = \{a\}$, $X2 = \{b, c, d\}$, $Y1 = \{e, g\}$ i $Y2 = \{f, h\}$. En termes d'aquests conjunts, aquesta última votació es veu en la forma següent:

$$65 : X1, X2; \quad 15 : X1, Y2; \quad 20 : Y1, Y2. \quad (9)$$

Quan les candidatures sorgeixen de manera espontània tot i que es vota en termes de candidats individuals, els diversos elements d'una mateixa candidatura se solen qualificar de **clons**.

En el marc general que introduïm, una **assignació d'escons** es limita a especificar quants escons rep cada candidatura i . Aquest nombre l'anomenarem n_i . El seu valor no pot superar en cap cas el nombre de candidats individuals que formen la candidatura en qüestió. D'aquest valor màxim en direm la **capacitat** de la candidatura i , i l'anomenarem v_i . En el cas de candidatures reduïdes cadascuna a un sol candidat individual tenim $v_i = 1$ per a tota i . En el cas de llistes de partit s'acostuma a entendre que $v_i = n$ per a tota i .

Una assignació d'escons en el sentit que acabem de dir no acaba de determinar quins candidats concrets resulten elegits. A aquest efecte, hi ha diverses possibilitats. Una d'aquestes és simplement fer un sorteig. Una altra és utilitzar un ordre prefixat dins de cada candidatura, tal com fan els partits en el sistema de llistes tancades. Finalment, es pot considerar la possibilitat que l'elector també expressi preferències entre candidats individuals, en el qual cas correspondrà basar-se en aquesta informació.

De tota manera, en el que segueix ens ocupem només de la primera part del problema, és a dir, la determinació del nombre d'escons que rep cada candidatura.

4.2 Quan es vota en termes de candidats individuals però existeix una agrupació d'aquests en partits i uns líders reconeguts dins de cada partit, aleshores el

vot d'aprovació dona peu a les anomenades estratègies de **decapitació** [8, p. 25, 50, 56]. Si els seguidors d'un determinat partit aproven sistemàticament tots els seus candidats, aleshores uns pocs vots que aprovin la mateixa llista llevat dels líders poden aconseguir que aquests últims no resultin elegits.

En rigor, això dependrà del mètode que s'utilitzi per determinar els candidats elegits. Però ja es veu a venir que tots els mètodes raonables tindran aquest problema, i el mètode de Phragmén que utilitza el vot d'aprovació no n'és una excepció. Davant d'això, Phragmén va proposar unes certes variacions que utilitzen el vot preferencial (vegeu l'apèndix B).

Tanmateix, des del moment que tenen sentit les estratègies de decapitació vol dir que hi ha unes candidatures col·lectives ben definides, de manera que es pot plantejar la votació en termes d'aquestes candidatures col·lectives, i llavors desapareix la qüestió de la decapitació.

5 La regla de Jefferson i D'Hondt i les seves alternatives

5.1 La regla de Jefferson i D'Hondt i les seves alternatives corresponen al cas en què els electors voten en termes de candidatures col·lectives però es limiten a indicar-ne només una. D'altra banda, també suposen que cada candidatura pot arribar a proporcionar n representants. D'aquest escenari en direm a partir d'ara el **cas uninominal no limitat**.

En aquest escenari, el principi de proporcionalitat demana que els nombres d'escons assignats n_i siguin proporcionals als nombres de vots obtinguts, que d'ara endavant seran denotats per w_i .

La proporcionalitat exacta la donarien els valors següents, que anomenarem **quotes exactes**:

$$n_i = n w_i / w, \quad (10)$$

$$\text{on } w = w_1 + w_2 + w_3 + \dots + w_I, \quad (11)$$

i I denota el nombre de candidatures. Noteu que les quotes exactes s'obtenen a partir dels nombres de vots w_i tot dividint-los per la quantitat $q_0 = w/n$. Més amunt, en el context del vot únic transferible, ens referíem a aquesta última quantitat com a «quota simple». Aquí usarem una altra terminologia on aquesta quantitat, i certes alternatives que de seguida veurem, reben el nom de **divisor**.

Però en general les quotes exactes no seran nombres enters, de manera que cal decidir com arrodonir-les per tal de desviar-se de la proporcionalitat tan poc com sigui possible.

Una possibilitat molt natural és la regla de les **restes majors**, que consisteix a assignar d'entrada tants escons com digui la part entera de la quota exacta, i després donar-ne un de més als partits amb restes més grans.

5.2 Tanmateix, en l'esperit del principi de proporcionalitat és més propi procedir de la manera següent: trobar un divisor q de manera que, en formar els quocients w_i/q i arrodonir el seu valor per defecte, la suma dels nombres

*enters obtinguts sigui exactament n (la qual cosa requereix $q \leq q_0$). Aquesta regla va ser proposada el 1882 per l'advocat belga Victor D'Hondt. Noranta anys abans, Thomas Jefferson havia proposat exactament la mateixa regla amb motiu del repartiment territorial dels escons de la cambra baixa dels Estats Units d'Amèrica. D'acord amb això, d'ara en endavant en direm **regla de Jefferson i D'Hondt**.*

Encara que a la pràctica és molt poc freqüent, de vegades hi pot haver empats que donin lloc a indecisions. Per exemple, amb la regla de Jefferson i D'Hondt ens trobem amb problemes a l'hora de repartir dos escons entre dos partits que han obtingut, respectivament, $\mathbf{w} = (2000, 1000)$ vots. Per a $q \in (1000, 2000)$ les parts enteres de w_i/q són $(1, 0)$, mentre que per a $q \leq 1000$ són $(2, 1)$, i cap d'aquests dos repartiments dona un total de dos escons. Tanmateix, per a dades molt properes obtenim tant el repartiment $\mathbf{n} = (2, 0)$, per exemple per a $\mathbf{w} = (2001, 999)$, com el repartiment $\mathbf{n} = (1, 1)$, per exemple per a $\mathbf{w} = (1999, 1001)$. Davant d'això, en el cas problemàtic $\mathbf{w} = (2000, 1000)$ és natural admetre qualsevol d'aquests dos repartiments. En general, admetrem, doncs, la possibilitat de més d'una solució. En relació amb això i amb la regla de Jefferson i D'Hondt, resulta convenient redefinir la part entera d'un nombre real z de manera que quan z sigui enter s'admetin llavors tant el mateix valor z com $z - 1$. Amb aquesta finalitat farem ús de la notació següent:

$$\ll z \gg = \begin{cases} \text{l'enter més gran inferior a } z, & \text{si } z \text{ no és enter;} \\ \text{o bé } z \text{ o bé } z - 1, & \text{si } z \text{ és enter.} \end{cases} \quad (12)$$

Equivalentment, $\ll z \gg$ és qualsevol enter que satisfaci les desigualtats

$$\ll z \gg \leq z \leq \ll z \gg + 1. \quad (13)$$

Per parlar amb tota propietat hauríem de considerar $\ll z \gg$ com un conjunt i escriure $k \in \ll z \gg$ en lloc de $k = \ll z \gg$; tanmateix, per als nostres propòsits serà suficient suposar que s'ha fet una elecció concreta del valor de $\ll z \gg$ sempre que z sigui enter.

A més de la formulació que n'hem donat més amunt, la regla de Jefferson i D'Hondt es pot reformular de les diverses maneres equivalents que recull el resultat següent:

TEOREMA 5.1. *Siguin $\mathbf{w} = (w_1, \dots, w_I)$ uns nombres reals positius i $\mathbf{n} = (n_1, \dots, n_I)$ uns nombres enters no negatius restringits a sumar un valor donat n . Les condicions següents són totes equivalents:*

- Existeix un nombre real positiu q tal que $n_i = \ll w_i/q \gg$ per a tota i .*
- n_i és el nombre de vegades que w_i figura com a numerador en els n quocients més grans de la forma w_i/k amb $i = 1, 2, \dots, I$ i $k = 1, 2, \dots$*
- $\mathbf{n}$ s'obté recursivament procedint de la manera següent: Per a $n = 0$ es posa $n_i = 0$ per a tota i . El pas de n a $n + 1$ es fa incrementant n_i a $n_i + 1$ per a una i qualsevol que maximitzi la quantitat $w_i/(n_i + 1)$.*

(d) $\min_i w_i/n_i \geq \max_i w_i/(n_i + 1)$.

(e) *n maximitza la quantitat $\min_i w_i/n_i$, i a més, minimitza el nombre de valors de i que realitzen el mínim que apareix en aquesta última expressió.*

El quocient w_i/n_i que apareix als apartats (d) i (e) es pot veure com el preu, en vots, que li costa cada escó a la candidatura i . Des d'aquest punt de vista, les desviacions respecte a la proporcionalitat no són més que diferències en el preu de cada escó per a les diferents candidatures. La caracterització (e) del resultat precedent correspon, doncs, a dir que les candidatures més afavorides siguin tan poc afavorides com sigui possible, i que a més, el nombre d'aquestes candidatures sigui el més petit possible. En el cas habitual d'absència d'empats entre els quocients w_i/k , aquesta caracterització, que aleshores es redueix al primer criteri que hi apareix, va ser posada en relleu el 1901 per Léon Rouyer, i també el 1910 per Maurice Equer. La caracterització general, vàlida en presència d'empats, es demostra a [36, § 2].

En lloc del quocient w_i/n_i , és interessant considerar també el seu invers n_i/w_i . Aquest ens dona la fracció d'escó que correspon a cadascun dels electors que han votat la candidatura i . Dit d'una altra manera, ens mesura la quantitat de representació que obté cadascun d'aquests electors. Des d'aquest punt de vista, la caracterització (e) del teorema 5.1 ens diu que la regla de Jefferson i D'Hondt minimitza la representació dels electors més representats.

5.3 En lloc d'arrodonir els quocients w_i/q per defecte, com fa la regla de Jefferson i D'Hondt, també hi ha altres opcions. En particular, és molt natural l'arrodoniment a l'enter més proper (que admet dues possibilitats en el cas de valors semienters). El resultat és la regla de **Webster i Sainte-Laguë**: *trobar un divisor q de manera que, en formar els quocients w_i/q i arrodonir-los a l'enter més proper, la suma dels nombres enters obtinguts sigui exactament n* . Aquesta regla va ser proposada per Daniel Webster el 1832 en el context del repartiment territorial, i per André Sainte-Laguë el 1910 en el context que aquí ens ocupa d'assignar un nombre de representants a cada candidatura segons els vots obtinguts.

Més concretament, Sainte-Laguë hi va arribar com a resultat de buscar la manera de minimitzar la suma dels quadrats de les diferències de representació entre els electors. Segons hem dit més amunt, la representació que obté un elector que ha votat la candidatura i és el quocient n_i/w_i del nombre d'escons que rep aquesta candidatura pel nombre d'electors que l'han votat. Si la proporcionalitat fos perfecta, tots aquests quocients serien iguals, i valdrien, per tant, n/w . En conseqüència, s'anul·laria la quantitat següent, que expressem en dues formes que són exactament equivalents:

$$y = \sum_i w_i \left(\frac{n_i}{w_i} - \frac{n}{w} \right)^2 = \frac{1}{2w} \sum_{i,j} w_i w_j \left(\frac{n_i}{w_i} - \frac{n_j}{w_j} \right)^2. \quad (14)$$

En termes estadístics, n/w i y/w no són altra cosa que la mitjana i variància de les representacions dels electors, és a dir, dels nombres n_i/w_i repetits cadascun amb la freqüència w_i . Doncs bé, Sainte-Laguë va demostrar que el criteri de minimitzar la suma y portava a la regla que hem enunciat més amunt:

TEOREMA 5.2. *La regla de repartiment de Webster i Sainte-Laguë és l'única que minimitza la suma y dels quadrats de les diferències de representació entre els electors.*

La regla de Webster i Sainte-Laguë aconsegueix millors resultats que la de Jefferson i D'Hondt pel motiu que les desviacions respecte a la proporcionalitat depenen de la mida dels partits. Més concretament, amb la regla de Jefferson i D'Hondt ocorre que les desviacions $\delta_i = n_i - nw_i/w$ tendeixen a ser positives per als partits grans, és a dir, per a valors grans de w_i/w , i negatives per als petits. Aquest efecte, que té un caràcter probabilístic, adquireix especial importància quan el territori està dividit en moltes circumscripcions, ja que aleshores es produeix un efecte acumulatiu. En canvi, la regla de Webster i Sainte-Laguë és molt més equilibrada des d'aquest punt de vista.

Per a més detalls sobre la regla de Jefferson i D'Hondt i les seves alternatives, remetem el lector a [36] i [49].

6 La correspondència de representació

6.1 Els mètodes de representació proporcional no es limiten a decidir quins candidats es converteixen en diputats; darrere d'aquesta decisió hi ha també una especificació —si més no potencial— de la **correspondència de representació**, és a dir, la correspondència entre electors i representants, la qual s'ha d'amoïllar a les restriccions i possibles preferències expressades pels electors. La proporcionalitat que es voldria aconseguir es refereix a les quantitats de representants i d'electors que es posen en correspondència. La idea és que el quocient d'aquestes dues quantitats hauria de ser el mateix per a qualsevol sector de la societat. Una altra manera de veure-ho és que cada elector hauria de tenir assignada la mateixa quantitat de representants. Com que hi ha molts menys representants que electors, a la pràctica aquesta quantitat de representants per elector serà una fracció molt petita de la unitat. Però el que importa no és això, sinó que tots els electors tinguin uns valors «com més iguals millor» d'aquesta variable. D'ara endavant ens hi referirem com la (quantitat de) **representació** obtinguda per l'elector en qüestió.

Quan els vots es limiten a escollir una sola candidatura, com hem considerat en la secció precedent, llavors cada elector queda representat pels diputats que hagi obtingut la candidatura escollida, i aquests diputats es reparteixen a parts iguals entre tots els electors que han votat aquesta candidatura. Per tant, cada elector que ha optat per la candidatura i obté una representació igual a n_i/w_i , on recordem que n_i denota el nombre d'escons obtinguts per aquesta candidatura i w_i , el nombre d'electors que hi han optat.

Tal com ja hem dit, les versions més tradicionals del vot únic transferible assignen a cada elector un sol representant o bé cap. En aquest escenari, el principi d'igualtat entre electors també porta a repartir cada representant a parts iguals entre els electors als quals és assignat. La proporcionalitat (aproximada) del mètode deriva del fet que cada escó requereix un nombre d'electors superior a la quota i que els vots sobrants són transferits a altres candidats.

Ja hem dit també que en les versions més elaborades del vot únic transferible la representació d'un elector sol quedar distribuïda entre diversos diputats. A continuació considerarem amb més detall aquest tipus de correspondències entre electors i representants, ja que també apareixen en els mètodes que estudiarem. D'altra banda, les considerarem només en el marc del vot d'aprovació, que és el que utilitzen aquests mètodes. En aquest cas, la restricció que ha de complir la correspondència de representació és molt clara: l'assignació de representants a un elector només pot utilitzar els candidats que ell ha aprovat.

6.2 Tot seguit concretem la manera d'especificar una correspondència de representació.

Sovint descriurem els vots mitjançant una enumeració freqüencial. És a dir, que en lloc de donar els vots un a un, direm quants electors han votat d'una o una altra manera (de fet, és el que hem fet des del principi en donar exemples de votacions). Les maneres de votar les anomenarem **opinions**, i el nombre d'electors que han expressat l'opinió k l'anomenarem u_k . Si estem enumerant els vots un a un, llavors $u_k = 1$ per a qualsevol k . Per a referir-nos als electors que han expressat l'opinió k , tot sovint direm simplement «els electors k ».

Per a especificar el contingut dels vots d'aprovació, utilitzarem una variable binària indexada per i i k : a_{ik} val 1 si els electors k aproven la candidatura i ; en cas contrari val 0. Alternativament, ho expressarem també escrivint $k \checkmark i$ en el primer cas, i $k \not\checkmark i$ en el segon. El conjunt de candidatures aprovades pels electors k el denotarem A_k . Aquest conjunt el suposarem sempre no buit. Dit d'una altra manera, només tindrem en compte els vots que aproven alguna candidatura. El nombre total de vots en consideració és

$$w = \sum_k u_k. \quad (15)$$

El nombre d'aprovacions obtingudes per la candidatura i és

$$w_i = \sum_k a_{ik} u_k = \sum_{k \checkmark i} u_k. \quad (16)$$

D'altra banda, el nombre de candidatures aprovades pels electors k és

$$\sigma_k = \sum_i a_{ik}. \quad (17)$$

La notació que hem introduït fins aquí es referia a les dades del problema. Parlem ara de les incògnites. Tot gira al voltant de la correspondència de

representació, la qual especificarem mitjançant una variable real, positiva o nul·la, indexada també per i i k ; aquesta variable la denominarem x_{ik} i el seu valor dóna la fracció de la candidatura i que és assignada a cada elector de l'opinió k . El nombre d'escons obtinguts per la candidatura i és

$$n_i = \sum_k x_{ik} u_k. \quad (18)$$

D'altra banda, la suma

$$r_k = \sum_i x_{ik} \quad (19)$$

dóna la quantitat de representació que obté cada elector de l'opinió k .

Les restriccions que tenim són les següents. En primer lloc, les fraccions de representació han de ser positives o nul·les:

$$x_{ik} \geq 0, \quad \text{per a qualssevol } i, k. \quad (20)$$

En segon lloc, la representació de cada elector s'ha de restringir a les candidatures que ell aprova:

$$x_{ik} = 0, \quad \text{sempre que } a_{ik} = 0. \quad (21)$$

En tercer lloc, el nombre d'escons que obté una candidatura ha de ser enter, no negatiu, i inferior o igual a la seva capacitat v_i (§ 4):

$$n_i \in \mathbb{Z}, \quad 0 \leq n_i \leq v_i. \quad (22)$$

Finalment, cal que el nombre total d'escons sigui el desitjat:

$$\sum_i n_i = n. \quad (23)$$

Noteu que la restricció (21) permet limitar la suma (18) als electors k que aproven la candidatura i , i una cosa similar passa amb la suma (19):

$$n_i = \sum_{k \in i} x_{ik} u_k, \quad (24)$$

$$r_k = \sum_{i \in k} x_{ik}. \quad (25)$$

En vista de (18) o (24), és clar que si totes les quantitats u_k es multipliquen per una mateixa constant i totes les quantitats x_{ik} es divideixen per aquesta mateixa constant, aleshores s'obté el mateix repartiment (n_i). Per tant, el que importa no són els valors concrets de u_k i x_{ik} , sinó les seves proporcions respectives, és a dir, els quocients u_k/u_ℓ i $x_{ik}/x_{i\ell}$.

L'objectiu que es tracta d'aconseguir és que les representacions r_k siguin com més iguals millor per a totes les opinions k . Com és ben sabut —vegeu, per exemple, [36, § 5]— això admet diverses interpretacions, les quals donaran lloc a mètodes diferents.

6.3 La noció de correspondència de representació apareix també en els treballs de Monroe [35] i de Potthoff i Brams ([47], [5, ch. 6]). Entre les diverses possibilitats que exploren aquests autors, s'hi pot trobar una idea d'alguna manera dual a la precedent: minimitzar una mesura de la manca de representativitat (*misrepresentation*), és a dir, $\sum_{i,k} (1 - a_{ik}) x_{ik} u_k$, sota les restriccions (20), (22), (23) i una altra que imposa una mateixa quantitat predeterminada de representació r_k per a qualsevol k (a canvi d'això no s'imposa pas la restricció (21), i per tant no valen ni (24) ni (25), sinó solament (18) i (19)). En virtut de (18) i (23), aquest plantejament equival a maximitzar la representativitat $\sum_{i,k} a_{ik} x_{ik} u_k$ sota les condicions esmentades.

6.4 Suposem per un moment que no hi hagués la restricció (22), és a dir, que els nombres d'escons n_i poguessin ser fraccionaris i prendre qualsevol valor entre 0 i n . Suposem també que cada elector ha aprovat almenys una candidatura. En aquestes condicions no hi ha obstacle per a aconseguir que les representacions r_k siguin exactament iguals (i coincideixin, per tant, amb n/w). En efecte, per a cada opinió k es poden trobar fàcilment unes x_{ik} que compleixin les condicions (20)–(21) i tinguin com a suma el valor n/w .

Si els electors k han aprovat més d'una candidatura, aleshores aquesta assignació de valors a les x_{ik} es pot fer de moltes maneres diferents. Davant d'això, el més natural és tractar de la mateixa manera totes les candidatures aprovades per uns mateixos electors. És a dir, amb la notació que hem introduït a (17), prendre $x_{ik} = (n/w)/\sigma_k$ per a qualsevol i tal que $k \checkmark i$ (i $x_{ik} = 0$ per a qualsevol i tal que $k \not\checkmark i$). Els nombres (fraccionaris) d'escons que resulten per a cada candidatura són $n_i = (n/w) \sum_{k \checkmark i} u_k / \sigma_k$. Aquests nombres són exactament proporcionals a les quantitats

$$\widehat{w}_i = \sum_{k \checkmark i} \frac{u_k}{\sigma_k}, \quad (26)$$

les quals compten els vots que obté cada candidatura quan cada vot es reparteix a parts iguals entre totes les candidatures que hi són aprovades.

Així doncs, en el supòsit que n_i/n pogués prendre qualsevol valor fraccionari entre 0 i 1, aleshores posant $n_i/n = \widehat{w}_i/w$ s'aconsegueix alhora una representació igual de tots els electors i un tractament igual de les diverses candidatures. Tot i que el supòsit que considerem contradiu la restricció dels n_i a valors enters, hom s'hi acostarà cada vegada més en el límit $n \rightarrow \infty$ sempre que les candidatures tinguin capacitat màxima $v_i = n$. És raonable, doncs, demanar que en aquest límit el quocient n_i/n s'acosti al valor $\widehat{w}_i/w$.

Per a valors finits de n o candidatures individuals ($v_i = 1$) les dues condicions d'igualtat de què parlàvem fa un moment —entre electors i entre candidatures— fàcilment deixen de poder-se complir simultàniament.

Davant d'això, és natural pensar a implementar la restricció a nombres enters mitjançant alguna regla de repartiment enter, com ara la de Jefferson i D'Hondt o la de Webster i Sainte-Laguë, aplicada a les quantitats $\widehat{w}_i$.

De fet, D'Hondt proposava una regla d'aquesta mena per a tractar els vots «mixtos», és a dir, els que combinaven candidats de diferents partits. A diferència del que nosaltres considerem, un vot d'aquest tipus combina diversos partits en unes proporcions no necessàriament iguals, ja que pot contenir més candidats d'un partit que d'un altre. En aquest context és raonable procedir com proposava D'Hondt ([10, p. 25–30], [11, p. 32–33, 43–49]), és a dir, que cada vot d'aquest tipus sigui dividit entre els partits en qüestió en proporció als respectius nombres de candidats (la qual cosa correspon a substituir (26) per una expressió més general de la forma $\widehat{w}_i = \sum_{k \in I} \alpha_{ik} u_k$ amb $\sum_{i \in I} \alpha_{ik} = 1$). Aquesta regla també està clarament enunciada a [2, princip IV].

Tanmateix, el que vol expressar un vot d'aprovació no és pas una divisió en parts iguals, sinó simplement un conjunt de possibilitats acceptables, la qual cosa permet donar preferència a l'objectiu fonamental d'aconseguir que les representacions r_k siguin com més iguals millor. Aquesta és la idea essencial del mètode de Phragmén, el qual estudiarem detalladament en les properes seccions.

El malencert d'aplicar una regla de repartiment enter a les quantitats $\widehat{w}_i$ és especialment acusat en el cas de candidatures individuals, és a dir, amb la restricció $n_i \in \{0, 1\}$. En efecte, en aquest cas qualsevol regla raonable de repartiment enter d'acord amb unes quantitats donades es redueix a seleccionar els n valors més grans d'aquestes quantitats. Si ho apliquem a l'exemple (3) amb les quantitats (26) —i també amb les quantitats (16)— aleshores surten elegits els quatre candidats més votats, la qual cosa és totalment contrària a l'esperit de la representació proporcional.

6.5 En contra del plantejament que estem fent es pot objectar que voler especificar quin diputat representa cada elector és una utopia, i més encara si la representació d'un elector la repartim entre diversos diputats. A la pràctica, l'únic que importa és quins candidats són elegits diputats. Això és ben cert. Però cal recordar que parlem de vots on es poden aprovar diverses candidatures. En aquest context, el problema de decidir quin repartiment d'escons és més adient per a representar el conjunt de l'electorat no és pas senzill. En l'esperit de la representació proporcional, el que toca fer és considerar la correspondència detallada (x_{ik}) i buscar la que doni unes representacions (r_k) més equitatives, en algun sentit o un altre que caldrà especificar.

Un problema anàleg que pot ajudar a entendre el que ens ocupa és el següent: en lloc d'electors als quals cal donar representació, tenim un grup de persones que necessiten alimentació; en lloc de candidats, tenim aliments de diversos tipus, els quals mesurem amb una unitat comuna (que podria correspondre, per exemple, a un determinat contingut energètic). De cada aliment n'hi ha un determinat nombre enter d'unitats, i el consum que se'n faci també ha de ser un nombre enter d'unitats (si s'enceta una unitat llavors cal consumir-la tota). L'objectiu és repartir una quantitat total determinada n d'aliments d'una manera com més equitativa millor, tot respectant la restricció de cada persona als aliments que ella digui.

L'analogia esdevindria més completa si suposéssim que darrere cada aliment hi ha algú, el seu fabricant, que té especial interès que es consumeixi aquell aliment i no els altres, la qual cosa intenta aconseguir mitjançant publicitat i potser també mitjançant algun canvi en l'aliment que l'acosti més als gustos existents.

6.6 Una correspondència de representació pot ser mostrada sobre el paper mitjançant el que en direm un **diagrama de rectangles**, que és una variació del que fa Phragmén a [43, p. 89–90]. La figura 1 en dóna un exemple.

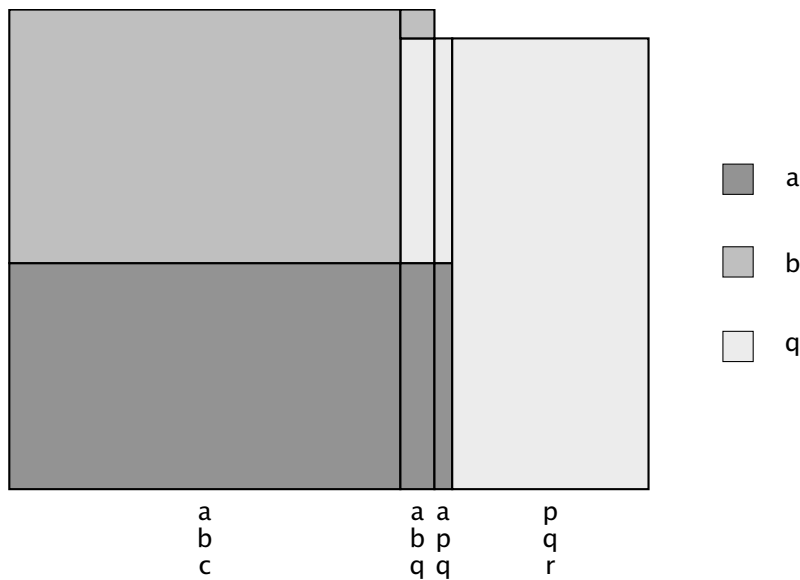


FIGURA 1: Diagrama de rectangles d'una correspondència de representació.

La base de la figura està dividida en tants segments com opinions k s'han manifestat. A sota de cada segment hem indicat la corresponent llista de candidatures aprovades. El segment k té una longitud proporcional a u_k , el nombre d'electors que han manifestat aquella opinió. A sobre del segment k hi ha tants rectangles com candidatures i donen representació a aquests electors. En cada rectangle s'indica a quina candidatura correspon; aquí ho fem mitjançant un codi de colors. L'altura del rectangle ik és proporcional a x_{ik} . Per tant, l'àrea d'aquest rectangle és proporcional a $x_{ik}u_k$, i l'àrea total dels rectangles de la candidatura i és proporcional al nombre d'escons obtinguts. D'altra banda, la suma de les altures dels rectangles que hi ha a sobre del segment k dóna la representació r_k que obté cada elector d'aquesta opinió.

7 El mètode de Phragmén. Variant seqüencial

En aquesta secció estudiem a fons la forma bàsica del mètode de Phragmén, que aquest autor va presentar en els anys 1894–1899 [39, 40, 42, 43]. Abans i després d'aquestes dates, Phragmén va considerar diverses variacions sobre el tema, de les quals donem una classificació en l'apèndix B. A diferència de Phragmén, que considerava només el cas de candidatures individuals, nosaltres ens situarem en el marc general en què les opcions sobre les quals s'expressa l'elector poden ser candidatures col·lectives, cadascuna de les quals ofereix un determinat nombre de candidats individuals.

7.1 La forma bàsica del mètode de Phragmén té un caràcter seqüencial: en cada pas s'assigna un escó a alguna candidatura i aquest escó es reparteix, en parts no necessàriament iguals, entre els electors que l'han aprovat.

Com veurem de seguida, la idea de buscar la màxima proporcionalitat possible s'entén en un sentit que és comú amb la regla de Jefferson i D'Hondt. De fet, i tal com veurem més avall, en el cas uninominal no limitat el mètode de Phragmén es redueix a la regla de Jefferson i D'Hondt.

Suposem, de moment, que les candidatures són individuals. Doncs bé, el primer escó s'assigna al candidat que ha reunit més aprovacions. Aquest primer escó sí que es reparteix a parts iguals; els receptors són tots els electors que han aprovat el dit candidat. Com que es tracta del candidat més aprovat, la representació que obté cada elector és la més petita possible. *El següent escó s'assignarà a un altre candidat i es repartirà de manera que tots els electors que l'aproven igualin la representació acumulada per cadascun d'ells fins aquell moment. El candidat que se selecciona és el que fa que aquesta representació acumulada sigui com més petita millor.* Noteu que els electors que no aproven aquest segon candidat s'han quedat amb la representació que tenien, que pot ser nul·la si tampoc havien aprovat el primer candidat. El següent escó se selecciona i reparteix d'acord amb el mateix principi, i així successivament.

En el cas de candidatures col·lectives, l'única diferència és que *cada candidatura pot continuar rebent escons mentre no arribi a la seva capacitat v_i .*

Vegem com es concreta això en un exemple (extret dels articles de Phragmén). Suposem que s'han d'elegir tres representants, que hi ha sis candidats a, b, c, p, q, r, i que els vots són els següents:

$$1034 : a, b, c; \quad 90 : a, b, q; \quad 47 : a, p, q; \quad 519 : p, q, r. \quad (27)$$

Segons el que hem dit, el primer escó el rep el candidat a, que és el que totalitza més aprovacions, és a dir, $1034 + 90 + 47 = 1171$. Cada elector que l'ha aprovat rep una part alíquota d'aquest escó, és a dir, $1/1171 = 0.854 \cdot 10^{-3}$ escons.

Vegem ara quin candidat obté el segon escó. El candidat b ha estat aprovat pels electors de les opinions abc i abq, els quals ja tenen representació, tots ells la mateixa, a través de a. Per tant, si el nou escó és assignat a b, es repartirà uniformement entre aquests $1034 + 90 = 1124$ electors, cadascun dels quals

elevaà la seva representació en $1/1124 = 0.890 \cdot 10^{-3}$ escons, i arribarà, per tant, a $0.854 \cdot 10^{-3} + 0.890 \cdot 10^{-3} = 1.744 \cdot 10^{-3}$ escons. De manera similar, si el nou escó és assignat a c, aleshores es repartirà a parts iguals entre els 1034 electors de l'opinió abc, cadascun dels quals elevaà la seva representació en $1/1034 = 0.967 \cdot 10^{-3}$ escons, i arribarà, per tant, a $0.854 \cdot 10^{-3} + 0.967 \cdot 10^{-3} = 1.821 \cdot 10^{-3}$ escons. Considerem ara la possibilitat d'assignar el segon escó al candidat p. Aquest és aprovat pels electors apq, que ja tenen representació a través de a, i també pels pqr, que encara no tenen representació. Segons hem dit, si l'assignem a p, el nou escó ha de ser repartit de manera que iguali la representació de tots aquests electors. Per a calcular aquesta representació final comuna ρ hem de resoldre l'equació $519\rho + 47(\rho - 0.854 \cdot 10^{-3}) = 1$, que dóna

$$\rho = \frac{47 \times 0.854 \cdot 10^{-3} + 1}{(519 + 47)} = 1.838 \cdot 10^{-3}. \quad (28)$$

Un càlcul similar per al candidat q ens diu que en aquest cas els electors abq, apq i pqr quedarien tots ells amb una representació igual a $1.703 \cdot 10^{-3}$ escons. Finalment, en el cas del candidat r, aprovat només pels electors pqr, s'obté que aquests electors obtindrien una representació de $1/519 = 1.927 \cdot 10^{-3}$ escons. Recopilant, les cinc possibilitats d'assignació del segon escó eleven la representació dels electors més afavorits als valors següents: b: $1.744 \cdot 10^{-3}$, c: $1.821 \cdot 10^{-3}$, p: $1.838 \cdot 10^{-3}$, q: $1.703 \cdot 10^{-3}$, r: $1.927 \cdot 10^{-3}$. Per tant, el segon escó és assignat al candidat q, que minimitza la representació acumulada fins ara pels seus electors.

Passem, finalment, a considerar el tercer escó. En aquest moment, els electors abc tenen cadascun $0.854 \cdot 10^{-3}$ escons, i els abq, apq i pqr tenen cadascun $1.703 \cdot 10^{-3}$ escons. Mitjançant uns càlculs anàlegs als del paràgraf precedent, s'obté que el tercer escó ha de ser assignat al candidat b, el qual deixa els electors abc i abq amb una representació de $1.812 \cdot 10^{-3}$ escons.

El diagrama de rectangles de la correspondència obtinguda és el que hem donat com exemple a la figura 1.

Tal com fa notar Phragmén [43, p. 299-300], cada vegada que es considera un candidat concret i la seva distribució entre els electors, el problema és anàleg a la distribució d'una unitat de volum d'un cert líquid entre uns recipients cilíndrics que potser ja contenen altres líquids de la mateixa densitat. Cada recipient k té una base d'àrea u_k i només admet els líquids i tals que $k \nmid i$. D'altra banda, tots els recipients que admeten el líquid i estan comunicats entre si, de manera que quan s'hi aboca una unitat de volum del líquid i , la gravetat fa que aquest es distribueixi amb un mateix nivell superior en tots aquests recipients.

7.2 En termes matemàtics, es parteix de $n = 0$ i la distribució nul·la, i el pas de n a $n + 1$ es fa de la manera següent: per a cada candidatura i que encara no ha esgotat la seva capacitat v_i , es considera la representació total que reunirien els electors que aproven aquesta candidatura en el supòsit d'assignar-li aquest

nou escó, és a dir, la representació que ja tenen més el nou escó, i es divideix pel nombre total d'aquests electors:

$$\rho_i^{(n+1)} = \frac{\sum_{k \nmid i} u_k r_k^{(n)} + 1}{w_i}. \quad (29)$$

Un valor més o menys elevat d'aquest quocient vol dir que els electors en qüestió resultarien més o menys beneficiats en el cas de seleccionar la candidatura i . En el mateix esperit que la regla de Jefferson i D'Hondt, és a dir, que els electors més afavorits ho siguin com menys millor, Phragmén proposa d'assignar el nou escó a una candidatura i que minimitzi aquest quocient (sota la condició que encara no hagi esgotat la seva capacitat v_i). Si n'hi ha més d'una, aleshores considerem admissible qualsevol d'aquestes. *En el que segueix ens referirem a aquesta candidatura mitjançant la notació $\bar{i}_{n+1}$ i escriurem ρ_{n+1} en lloc de $\rho_{\bar{i}_{n+1}}^{(n+1)}$; d'altra banda, també convindrem a posar $\rho_0 = 0$.* El nou escó de la candidatura escollida $\bar{i}_{n+1}$ es reparteix entre els seus electors de tal manera que tots ells assoleixin la representació ρ_{n+1} . Tal com veurem de seguida, això s'aconsegueix mitjançant l'actualització següent del repartiment x_{ik} :

$$x_{ik}^{(n+1)} - x_{ik}^{(n)} = \begin{cases} \rho_{n+1} - r_k^{(n)}, & \text{si } i = \bar{i}_{n+1} \text{ i } k \nmid i; \\ 0, & \text{en cas contrari.} \end{cases} \quad (30)$$

Multiplicant aquestes quantitats per u_k i sumant sobre k , veiem que efectivament els increments $x_{ik}^{(n+1)} - x_{ik}^{(n)}$ totalitzen exactament un escó. D'altra banda, sumant (30) sobre i s'obté que

$$r_k^{(n+1)} - r_k^{(n)} = \begin{cases} \rho_{n+1} - r_k^{(n)}, & \text{si } k \nmid \bar{i}_{n+1}; \\ 0, & \text{en cas contrari.} \end{cases} \quad (31)$$

D'això es dedueix el que havíem anunciat més amunt: *els electors k que han aprovat $\bar{i}_{n+1}$ assoleixen tots ells $r_k^{(n+1)} = \rho_{n+1}$.* Més generalment, una aplicació reiterada de (31) porta a la conclusió següent (on considerem l'índex n en lloc de $n + 1$):

PROPOSICIÓ 7.1. *Per a qualsevol $n \geq 0$ es compleix $r_k^{(n)} = \rho_m$, on m és el màxim índex, anterior o igual a n , tal que k ha aprovat la candidatura $\bar{i}_m$; si k no ha aprovat cap de les candidatures $\bar{i}_1, \dots, \bar{i}_n$, llavors $r_k^{(n)} = \rho_0 = 0$. És a dir:*

$$r_k^{(n)} = \rho_m, \quad \text{on } m = \max(\{0\} \cup \{p \in \mathbb{Z} \mid 1 \leq p \leq n, k \nmid \bar{i}_p\}). \quad (32)$$

En els paràgrafs precedents hem donat per suposat que la successió $(\rho_n \mid n = 0, 1, 2, \dots)$ no decreix en cap moment. Si tinguéssim $\rho_{n+1} < \rho_n$, llavors deixaria de ser cert que ρ_{n+1} fos la representació dels electors més afavorits en aquell moment, tal com suposàvem. D'altra banda, alguna de les quantitats que apareixen al segon membre de (30) i (31) llavors podria ser negativa, la qual cosa podria portar a valors negatius de $x_{ik}^{(n)}$, que no són admissibles. Tot això queda descartat amb el resultat següent.

PROPOSICIÓ 7.2. $\rho_{n+1} \geq \rho_n$ per a qualsevol $n \geq 0$. En el cas $n = 0$ la desigualtat és estricta.

PROVA. Procedirem per inducció. Per agilitzar el llenguatge, escriurem i en lloc de $\bar{i}_{n+1}$. Per a $n = 0$ tenim, efectivament,

$$\rho_1 = \rho_i^{(1)} = \frac{1}{w_i} > 0 = \rho_0. \quad (33)$$

Per a passar de n a $n + 1$ només cal notar que

$$\rho_{n+1} = \rho_i^{(n+1)} = \frac{\sum_{k \leq i} u_k r_k^{(n)} + 1}{w_i} \geq \frac{\sum_{k \leq i} u_k r_k^{(n-1)} + 1}{w_i} \geq \rho_n, \quad (34)$$

on la darrera desigualtat és conseqüència immediata de la definició de ρ_n , i l'anterior és conseqüència de la hipòtesi d'inducció a través de (32). $\square$

PROPOSICIÓ 7.3. $\rho_n = \max_k r_k^{(n)}$.

PROVA. Combinant la proposició 7.2 amb la igualtat (32) s'obté la desigualtat $r_k^{(n)} \leq \rho_n$ per a qualsevol k . D'altra banda, (32) també ens diu que $r_k^{(n)}$ és igual a ρ_n per als electors k que han aprovat $\bar{i}_n$. $\square$

Tot i que en general la desigualtat $\rho_{n+1} \geq \rho_n$ no és estricta, es compleix sempre el fet següent:

PROPOSICIÓ 7.4. Cada elector rep una fracció estrictament positiva de cada candidatura que ell ha aprovat i que ha resultat escollida.

PROVA. Comencem per observar que n'hi ha prou amb considerar candidats individuals, ja que una candidatura col·lectiva equival a un conjunt de candidats individuals que van sempre junts. La demostració serà per contradicció. Suposem el contrari del que volem demostrar: l'escó $(n + 1)$ -èsim és assignat a un cert candidat i amb la particularitat que hi ha uns electors ℓ que havien aprovat aquest candidat però no reben representació a través d'ell. Si passa això vol dir que, abans d'assignar aquest escó, els electors ℓ ja havien assolit la quantitat de representació que ara obtenen tots els que han aprovat i . És a dir, $r_\ell^{(n)} = \rho_{n+1}$. Suposem que els electors ℓ han assolit aquesta representació en ser assignat l'escó m -èsim, de manera que tenim $r_\ell^{(m-1)} < r_\ell^{(m)} = \rho_{n+1}$.

Combinant els resultats precedents, se'n dedueix que $r_\ell^{(m-1)} < r_\ell^{(n)}$. D'altra banda, per a una k arbitrària tenim la desigualtat no estricta $r_k^{(m-1)} \leq r_k^{(n)}$. Aquests fets ens permeten escriure que

$$\rho_m \leq \rho_i^{(m)} = \frac{\sum u_k r_k^{(m-1)} + 1}{w_i} < \frac{\sum u_k r_k^{(n)} + 1}{w_i} = \rho_i^{(n+1)} = \rho_{n+1}. \quad (35)$$

Tanmateix, la proposició 7.3 permet deduir també la desigualtat contrària $\rho_m \geq r_\ell^{(m)} = \rho_{n+1}$, de manera que efectivament hem arribat a una contradicció. $\square$

OBSERVACIÓ. En lloc de (29), Phragmén —vegeu, per exemple, [42, p. 191]— expressa $\rho_i^{(n+1)}$ mitjançant la fórmula següent, que deriva de (29) en introduir-hi (32):

$$\rho_i^{(n+1)} = \frac{w_i^{(n,1)} \rho_1 + w_i^{(n,2)} \rho_2 + \dots + w_i^{(n,n-1)} \rho_{n-1} + w_i^{(n,n)} \rho_n + 1}{w_i^{(n,1)} + w_i^{(n,2)} + \dots + w_i^{(n,n-1)} + w_i^{(n,n)} + w_i^{(n,0)}}, \quad (36)$$

on, per a cada $m = 1, 2, \dots, n$, $w_i^{(n,m)}$ representa el nombre total dels electors que aproven els candidats i i $\bar{i}_m$ però no cap dels candidats $\bar{i}_p$ amb $m < p \leq n$, i d'altra banda, $w_i^{(n,0)}$ representa el nombre total dels electors que aproven i però no cap dels $\bar{i}_p$ amb $p = 1, 2, \dots, n$.

7.3 A continuació considerem el nombre total d'escons que reuneix un conjunt X de candidatures. En relació amb això usarem la notació següent, on cal recordar que A_k representa el conjunt de candidatures aprovades pels electors k :

$$n_X = \sum_{i \in X} n_i, \quad (37)$$

$$u_X = \sum_{A_k = X} u_k, \quad (38)$$

$$w_X = \sum_{A_k \cap X \neq \emptyset} u_k. \quad (39)$$

PROPOSICIÓ 7.5. *Per a qualsevol conjunt X de candidatures es compleix la desigualtat següent: $n_X \leq \rho_n w_X$.*

PROVA.

$$\begin{aligned} n_X &= \sum_{i \in X} \sum_k x_{ik} u_k = \sum_k \sum_{i \in X} x_{ik} u_k = \sum_{A_k \cap X \neq \emptyset} \sum_{i \in X} x_{ik} u_k \\ &\leq \sum_{A_k \cap X \neq \emptyset} r_k u_k \leq \sum_{A_k \cap X \neq \emptyset} \rho_n u_k = \rho_n w_X, \end{aligned}$$

on hem partit de (18), hem tingut en compte que $x_{ik} = 0$ si no és que $i \in A_k$, i hem usat les desigualtats que es dedueixen de (19)-(20) i de la proposició 7.3. $\square$

COROL·LARI 7.6. *El nombre d'escons que rep una candidatura i compleix la desigualtat següent: $n_i \leq \rho_n w_i$.*

PROPOSICIÓ 7.7. *Sigui X un conjunt arbitrari de candidatures. Mentre n_X no arriba a la seva capacitat màxima $\sum_{i \in X} v_i$, es compleix sempre la desigualtat següent: $n_X + 1 \geq \rho_n u_X$.*

PROVA. Suposem la desigualtat contrària, és a dir,

$$n_X + 1 < \rho_n u_X. \quad (40)$$

Sigui i qualsevol candidatura de X amb $n_i < v_i$. Veurem que la desigualtat precedent implicaria $\rho_i^{(n+1)} < \rho_n$, en contradicció amb el fet que $\rho_n \leq \rho_{n+1} \leq \rho_i^{(n+1)}$. En efecte, partint de (29), podem escriure

$$\rho_i^{(n+1)} = \frac{\left(\sum_{k \in X, k \neq i} u_k r_k^{(n)}\right) + \left(\sum_{k \in X, k = i} u_k r_k^{(n)}\right) + 1}{\left(\sum_{k \in X, k \neq i} u_k\right) + \left(\sum_{k \in X, k = i} u_k\right)}. \quad (41)$$

Ara bé, el segon terme del denominador no és altre que u_X . D'altra banda, el segon terme del numerador està acotat superiorment per n_X ; en efecte,

$$\sum_{k \in X} u_k r_k^{(n)} = \sum_{j \in X} \sum_{k \in X} u_k x_{jk}^{(n)} \leq \sum_{j \in X} \sum_k u_k x_{jk}^{(n)} = \sum_{j \in X} n_j = n_X. \quad (42)$$

Finalment, les representacions $r_k^{(n)}$ que apareixen en el primer terme del numerador es poden acotar totes elles per ρ_n . Per tant, es compleix la desigualtat

$$\rho_i^{(n+1)} \leq \frac{\left(\sum_{k \in X, k \neq i} u_k \rho_n\right) + n_X + 1}{\left(\sum_{k \in X, k \neq i} u_k\right) + u_X}, \quad (43)$$

que combinada amb la hipòtesi (40) dóna $\rho_i^{(n+1)} < \rho_n$. □

7.4 En aquest apartat ens ocupem d'un parell d'escenaris especials. El primer és el cas uninominal no limitat. Recordi's que el caràcter uninominal vol dir que cada elector aprova una sola candidatura. Per tant, les possibles opinions es corresponen amb les diferents candidatures, i per a $X = \{i\}$ les quantitats u_X i w_X definides per (38) i (39) són iguals entre si i coincideixen amb la w_i de la secció 5. D'altra banda, el caràcter no limitat vol dir que cada candidatura i té capacitat $v_i = n$, és a dir, que pot arribar a proporcionar n representants. En aquestes condicions les proposicions 7.5 i 7.7 ens diuen que

$$\frac{w_i}{n_i + 1} \leq \rho_n \leq \frac{w_i}{n_i}. \quad (44)$$

D'acord amb la caracterització (d) del teorema 5.1, això ens permet arribar a la conclusió següent:

PROPOSICIÓ 7.8. *En el cas uninominal no limitat l'algorisme de Phragmén es redueix a la regla de Jefferson i D'Hondt.*

Val a dir que també és fàcil arribar directament a la caracterització (c) del teorema 5.1. En efecte, en les condicions que estem considerant, les desigualtats (42) i (43) per a $X = \{i\}$ es converteixen en igualtats i en la segona d'elles desapareixen els primers termes del numerador i del denominador. Això estableix la igualtat $\rho_i^{(n+1)} = (n_i + 1)/w_i$ per a qualsevol i . Per tant, l'escó $(n + 1)$ -èsim s'assigna a una candidatura i que minimitzi aquesta quantitat, és a dir, que maximitzi $w_i/(n_i + 1)$, tal com contempla la coneguda formulació recursiva de la regla de Jefferson i D'Hondt.

Considerem ara el cas que en podem dir uninominal totalment limitat: cada elector aprova una sola candidatura de capacitat 1. En altres paraules, cada elector aprova un sol candidat individual. En aquest cas, tenim $\rho_i^{(n+1)} = 1/w_i$ sempre que el candidat i encara no hagi estat elegit. Per tant, s'arriba a la conclusió següent:

PROPOSICIÓ 7.9. *En el cas en què cada elector es limita a aprovar un sol candidat individual, el mètode de Phragmén es redueix a seleccionar els n candidats més votats.*

7.5 A continuació ens ocupem de la important qüestió de la **monotonia respecte al contingut dels vots.** En el cas del vot d'aprovació es tracta simplement que, si no varia res més, l'addició d'aprovacions a favor d'una candidatura no pugui produir mai una disminució dels escons que li són assignats. Tot i que Phragmén es refereix diverses vegades a aquesta qüestió com a motivació del seu mètode, la veritat és que en cap moment no demostra detalladament que el seu procediment iteratiu tingui aquesta propietat. Tot seguit demostrarem que és efectivament així en el cas de candidatures individuals, però no en el cas de candidatures col·lectives.

PROPOSICIÓ 7.10. *En el cas de candidatures individuals, és a dir, $v_i = 1$ per a tota i , el procediment iteratiu de Phragmén té la propietat de monotonia respecte al contingut dels vots.*

PROVA. A continuació anomenarem i el candidat en qüestió i usarem una titlla per indicar les quantitats corresponents als nous vots. D'acord amb la hipòtesi que i rep aprovacions addicionals i que no varia res més, es compleix

$$\tilde{u}_k \geq u_k, \quad \text{sempre que } k \neq i. \quad (45)$$

Suposem que amb els vots originals el candidat en qüestió resultava elegit per a l'escó $(n + 1)$ -èsim, és a dir, que $i = \bar{i}_{n+1}$. Per a aconseguir el nostre objectiu, n'hi ha prou amb veure que, amb els nous vots, si aquest candidat no surt elegit per a un escó anterior, llavors surt elegit forçosament per a l'escó $(n + 1)$ -èsim. Com que suposem que les aprovacions dels altres candidats són exactament les mateixes que en els vots originals, la hipòtesi que el candidat i no ha sortit elegit per a cap escó anterior implica que no hi ha cap variació en les quantitats $\rho_1, \dots, \rho_n$ i per tant en $r_k^{(n)}$, i tampoc en $\rho_j^{(n+1)}$ per a $j \neq i$. En vista d'això, el nostre objectiu es redueix a demostrar que $\tilde{\rho}_i^{(n+1)} \leq \rho_i^{(n+1)}$.

Partint de (29) i de l'expressió anàloga amb els vots modificats, i tenint en compte que $w_i = \sum_{k \sqrt{i}} u_k$ i anàlogament per a $\tilde{w}_i$, se'n dedueix successivament que

$$\begin{aligned}\tilde{w}_i \tilde{\rho}_i^{(n+1)} - w_i \rho_i^{(n+1)} &= \sum_{k \sqrt{i}} (\tilde{u}_k - u_k) r_k^{(n)}, \\ \tilde{w}_i \tilde{\rho}_i^{(n+1)} - \tilde{w}_i \rho_i^{(n+1)} + \tilde{w}_i \rho_i^{(n+1)} - w_i \rho_i^{(n+1)} &= \sum_{k \sqrt{i}} (\tilde{u}_k - u_k) r_k^{(n)}, \\ \left(\sum_{k \sqrt{i}} \tilde{u}_k \right) (\tilde{\rho}_i^{(n+1)} - \rho_i^{(n+1)}) &= \sum_{k \sqrt{i}} (\tilde{u}_k - u_k) (r_k^{(n)} - \rho_i^{(n+1)}),\end{aligned}\quad (46)$$

on el signe del segon membre és negatiu o zero ja que $r_k^{(n)} = \rho_m$ per alguna $m \leq n$, i d'altra banda $\rho_i^{(n+1)} = \rho_{n+1} \geq \rho_m$. $\square$

En el cas de candidatures collectives l'addició d'aprovacions a favor d'una pot comportar una disminució del nombre d'escons que li són assignats. Per exemple, per a $n = 3$ i els vots

$$4 : A; \quad 7 : B; \quad 1 : A, B; \quad 16 : A, C; \quad 4 : B, C, \quad (47)$$

el mètode de Phragmén dóna dos escons a A, però si un dels vots que només aproven B passa a aprovar també A (de manera que els coeficients esdevenen respectivament 4, 6, 2, 16, 4) aleshores A obté només un sol escó (l'assignació successiva d'escons és ABAABA... en el primer escenari i ACBAAB... en el segon).

Un altre exemple interessant és el següent, on també considerem $n = 3$:²

$$10 : A; \quad 3 : B; \quad 12 : C; \quad 21 : A, B; \quad 6 : B, C. \quad (48)$$

En aquest cas, el mètode de Phragmén també dóna dos escons a A, però si els vots que aproven només la candidatura A s'incrementen en una unitat (passant a ser 11 en lloc de 10) aleshores A obté només un sol escó (l'assignació successiva d'escons és ACABAC... en el primer escenari i ABCABA... en el segon).

Tal com hem dit més amunt en relació amb el vot únic transferible, la manca de monotonía és un defecte poc menys que inacceptable. A la secció 8 veurem que aquest problema podria desaparèixer en certes variants del mètode de Phragmén.

7.6 El resultat següent estableix una propietat de l'estil del criteri de proporcionalitat de Droop. És una generalització d'un resultat que és ben conegut per a la regla de Jefferson i D'Hondt [36, prop. 3.1]. La demostració també serà anàloga. Recordi's de (38) que u_X vol dir el nombre d'electors que aproven exactament un determinat conjunt X de candidatures.

PROPOSICIÓ 7.11 (continguda a [28, sats 13.5, (ii)]). *Si $u_X > m q_D$, amb $q_D = w/(n+1)$ i $m \leq \sum_{i \in X} v_i$, aleshores $n_X \geq m$.*

² Aquest exemple ens ha estat assenyalat pel professor Svante Janson.

PROVA. La proposició 7.5, aplicada a X^c , el conjunt complementari de X , ens garanteix la desigualtat següent:

$$n - n_X \leq \rho_n w_{X^c} \leq \rho_n (w - u_X). \quad (49)$$

D'altra banda, l'acotació inferior que estem suposant sobre u_X i la identitat $w = (n + 1) q_D$ impliquen que

$$w - u_X < q_D (m + 1 - m). \quad (50)$$

A partir d'aquí procedirem per reducció a l'absurd. Suposem que no tinguéssim $n_X \geq m$ sinó la desigualtat contrària $n_X < m - 1$. Combinant-la amb (49) i (50) s'obté que $1/\rho_n < q_D$. Usant novament la hipòtesi d'acotació inferior de u_X , se'n dedueix que $1/\rho_n < u_X/m$. Tanmateix, la proposició 7.7 ens assegura que $\rho_n \leq (n_X + 1)/u_X$. Per tant, en resulta que $n_X + 1 > m$, és a dir $n_X \geq m$. $\square$

7.7 El cas de dues candidatures

Considerem dues candidatures A i B. Siguin α, β, ζ les fraccions de l'electorat que aproven respectivament només la candidatura A, només la B, o totes dues A i B (tenim, doncs, $\alpha + \beta + \zeta = 1$). Les quantitats de representació per elector que obtenen respectivament aquests tres tipus d'electors seran denotades aquí per r_n, s_n, t_n , on n indica el nombre d'escons en joc; els nombres d'escons que reben respectivament A i B seran denotats per p_n, q_n ($p_n + q_n = n$). Com que t_n correspon als electors que aproven tant A com B, la proposició 7.1 ens assegura que $t_n = \rho_n$, mentre que $r_n = \rho_{m_1}$ i $s_n = \rho_{m_2}$ amb $m_1, m_2 \leq n$ i $\max(m_1, m_2) = n$. En virtut de la proposició 7.2 se'n dedueix que $t_n = \max(r_n, s_n)$. D'altra banda, la regla que determina (r_{n+1}, s_{n+1}) a partir de (r_n, s_n) es pot escriure en la forma

$$(r_{n+1}, s_{n+1}) = \begin{cases} (r'_n, s_n), & \text{si } r'_n \leq s'_n; \\ (r_n, s'_n), & \text{si } r'_n \geq s'_n, \end{cases} \quad (51)$$

on $r_0 = s_0 = 0$ i r'_n i s'_n vénen donats per les fórmules

$$r'_n = \frac{\alpha r_n + \zeta t_n + 1}{\alpha + \zeta}, \quad s'_n = \frac{\beta s_n + \zeta t_n + 1}{\beta + \zeta}. \quad (52)$$

En el cas $r'_n = s'_n$ admetem les dues possibilitats, la qual cosa dóna lloc a múltiples solucions. El nombre d'escons obtinguts per cada candidatura segueix el procés següent:

$$(p_{n+1}, q_{n+1}) = \begin{cases} (p_n + 1, q_n), & \text{si } r_{n+1} > r_n; \\ (p_n, q_n + 1), & \text{si } s_{n+1} > s_n. \end{cases} \quad (53)$$

És natural preguntar-se com depenen p_n, q_n de α, β, ζ i n . Per a $\zeta = 0$, en què la regla de Phragmén es redueix a la de Jefferson i D'Hondt, la caracterització (a) del teorema 5.1 implica que $p_n/n \rightarrow \alpha$ i $q_n/n \rightarrow \beta$ quan $n \rightarrow \infty$.

PROVA. Tenint en compte (13), la caracterització (a) del teorema 5.1 ens permet escriure $n_i \leq w_i/q \leq n_i + 1$. I sumant aquestes desigualtats sobre i s'obté que $n \leq w/q \leq n + I$, on I denota el nombre de candidatures. D'aquestes desigualtats es dedueix d'una banda que $n_i \leq (n + I)w_i/w$ i d'altra banda que $n_i + 1 \geq nw_i/w$. Per tant, $w_i/w - 1/n \leq n_i/n \leq (n + I)w_i/nw$, d'on es dedueix el resultat desitjat: $\lim_{n \rightarrow \infty} n_i/n = w_i/w$.

Per a $\zeta > 0$ el comportament asimptòtic de $f_n = p_n/n$ quan $n \rightarrow \infty$ no és pas tan senzill. D'entrada, no és clar que $f_n = p_n/n$ convergeixi sempre cap a un límit. Dit això, les exploracions numèriques semblen indicar tal convergència per a «quasi tota parella» de valors de α i β . El límit f_∞ seria semblant al que indica la figura 2, on hem representat la dependència de f_{1200} respecte a α i β .

El valor de f_{1200} està indicat per la intensitat de gris: el color blanc correspon al valor 0 i el negre al valor 1. El valor de f_{1200} és constant en cadascun dels lòbuls que s'observen en la figura.³ El lòbul més extens, situat al mig, correspon al valor $1/2$. Els següents en extensió, situats a banda i banda de l'anterior, corresponen als valors $1/3$ i $2/3$. Més enllà, n'hi ha dos més que corresponen als valors $1/4$ i $3/4$. També se n'observen prou bé uns altres de més primers que corresponen als valors $1/5$, $2/5$, $3/5$ i $4/5$.

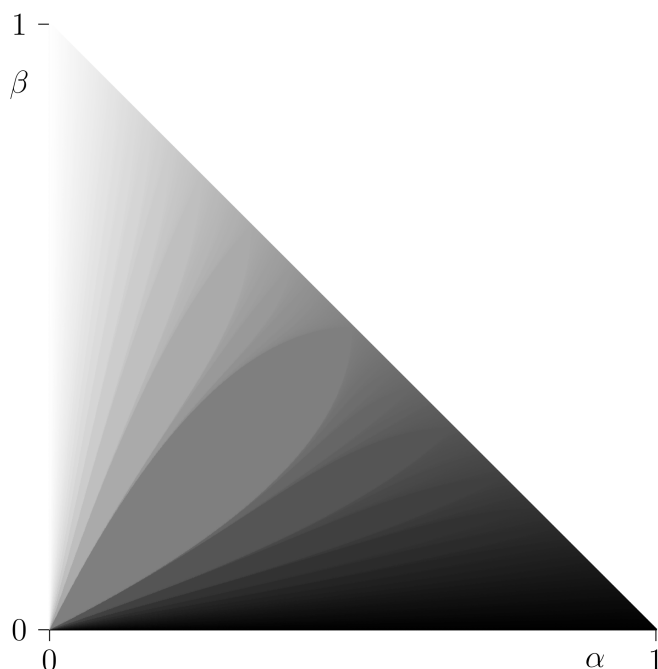


FIGURA 2: Comportament asimptòtic de la variant bàsica del mètode de Phragmén per a dues candidatures. Dependència de f_{1200} respecte de α i β .

³ Malauradament, es produeix una il·lusió òptica que exagera localment els contrastos. Vegeu, per exemple, Wiki Radiography, 2010, Mach bands and other optical illusions, <http://www.wiki-radiography.com/page/Mach+bands+and+other+Optical+Illusions>. Tanmateix, si es cobreix la figura llevat d'una part continguda en un sol lòbul, aleshores es percep una intensitat uniforme de gris.

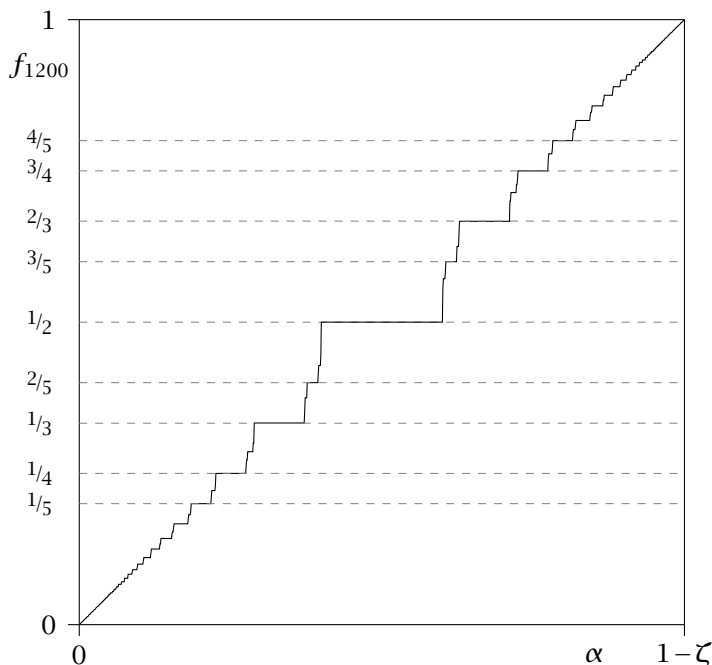


FIGURA 3: Dependència de f_{1200} respecte α per a $\zeta = 0.376$.

Tal com suggereix la figura 3, per a $\zeta > 0$ fixada, la dependència de f_∞ respecte de α tindria un caràcter similar a la funció de Cantor [14], amb la particularitat que aquí cada racional tindria com a antiimatge un interval de mesura positiva.

En qualsevol cas, el fet que f_∞ sigui constant en aquests lòbuls és una mica indesitjable. Fixem-nos, per exemple, en el lòbul central on $f_\infty = 1/2$. Tal com està delimitat, resulta que els escons es reparteixen uniformement entre A i B tant per a $(\alpha, \beta, \zeta) = (0.374, 0.25, 0.376)$ com per a $(\alpha, \beta, \zeta) = (0.25, 0.374, 0.376)$, tot i que en el primer cas el nombre d'electors que aproven només la candidatura A és superior gairebé un 50% al nombre dels que aproven només la candidatura B, i que en el segon cas passa exactament a l'inrevés. En lloc d'això, és més raonable, segons hem argumentat a l'apartat 6.4, repartir els escons proporcionalment a $\alpha + \zeta/2$ i $\beta + \zeta/2$ —que en la figura 3 correspon a una recta des de $(0, \zeta/2)$ fins a $(1 - \zeta, 1 - (\zeta/2))$.

Aquest fenomen no desapareix pas amb un nombre finit d'escons. Per exemple, per a qualsevol valor parell de n es continua obtenint un repartiment a parts iguals per als valors de (α, β, ζ) donats en el paràgraf precedent (i per a tots els valors intermedis).

8 Variants directes. Algunes perspectives

8.1 El procediment que hem estudiat en la secció anterior assigna els escons de manera successiva; en cada pas es mantenen les assignacions anteriors i el nou escó s'assigna de manera que es minimitzi la representació dels electors més afavorits. En lloc d'això, també es poden repartir tots els escons de cop prestant atenció a la quantitat de representació que obtenen els diferents electors i mirant de minimitzar la desigualtat entre ells. Aquests mètodes els anomenem **variants directes**.

Tal com recull el teorema 5.1 —quan afirma que són equivalents els seus apartats (c) i (e)—, en el cas uninominal no limitat el criteri de minimitzar la representació dels electors més afavorits dona el mateix resultat tant si es procedeix seqüencialment tot mantenint les assignacions anteriors com si es consideren directament tots els escons que hi ha per repartir.

Tanmateix, aquesta equivalència deixa de ser certa quan els electors no es limiten a aprovar una sola candidatura. En aquest cas, la restricció de mantenir les assignacions anteriors pot impedir d'accedir a altres repartiments que redueixen encara més la representació dels electors més afavorits. Considerem, per exemple, els vots d'aprovació següents referents a tres candidats individuals:

$$2 : b; \quad 2 : a, b; \quad 1 : c; \quad 3 : a, c. \quad (54)$$

Per a $n = 1$, l'assignació successiva es redueix a un sol pas, de manera que certament coincidirà amb l'assignació directa. L'objectiu de minimitzar la representació dels electors més afavorits porta a assignar l'escó al candidat més aprovat, en aquest cas a, i a repartir aquest candidat a parts iguals entre tots els electors que l'han aprovat. Per a $n = 2$, la variant d'assignació successiva busca quin altre candidat complementa millor la representació ja assignada a a, i arriba a la conclusió que és b. En canvi, l'assignació directa no manté necessàriament el candidat a, sinó que es planteja de nou com minimitzar la desigualtat de representació mitjançant dos representants. En aquest cas, és obvi que la millor elecció en aquest sentit és b, c, la qual permet donar la mateixa representació a tots els electors.

Tal com es veu de seguida amb aquest exemple, *les variants directes no seran pas monòtones respecte al nombre d'escons n*. És a dir, que un augment del nombre de representants que s'han d'elegir pot comportar que un candidat concret passi de ser elegit a no ser-ho. Aquest fenomen, del mateix tipus que la coneguda *paradoxa d'Alabama* [49, § 9.4–9.7], és força indesitjable. Tanmateix, en la pràctica el nombre de representants que s'han d'elegir és una quantitat prefixada, de manera que aquesta manca de monotonia no és tan greu.

En general no es pot aconseguir que tots els electors tinguin exactament la mateixa quantitat de representació. Davant d'això, és natural mirar de minimitzar algun índex que mesuri la desigualtat de les quantitats que obtenen els diferents electors.

També és cert, però, que no hi ha pas una sola manera de mesurar la desigualtat d'un repartiment.

Fins ara ens hem basat en el valor màxim de les fraccions que obtenen els diferents electors, és a dir, $\max_k r_k$. Aquesta quantitat no és ben bé una mesura de desigualtat, però basta restar-li el valor mitjà, és a dir, n/w , per obtenir un índex que s'anulla només en el cas d'un repartiment totalment equitatiu. Aquest índex ve donat, doncs, per

$$\mu = \max_k r_k - \frac{n}{w}. \quad (55)$$

De tota manera, com que els valors de n i w ens vénen donats, en lloc de minimitzar aquesta diferència μ podem minimitzar simplement el valor de $\max_k r_k$.

Alternativament, podem prendre com a mesura de desigualtat la variància de les representacions que obtenen els diferents electors. A l'apartat 5.3 ja hem considerat aquest punt de vista en el cas uninominal no limitat, el qual hem vist que dóna lloc a la regla de Webster i Sainte-Laguë. En el cas general, la variància ve donada per les fórmules següents, que generalitzen (14):

$$y = \sum_k u_k \left(r_k - \frac{n}{w} \right)^2 = \frac{1}{2w} \sum_{k,\ell} u_k u_\ell (r_k - r_\ell)^2 \quad (56)$$

(la igualtat entre ambdues expressions es comprova fàcilment desenvolupant els quadrats i usant el fet que $\sum_k u_k r_k = n$ i que $\sum_k u_k = w$).

La idea de minimitzar la variància de les representacions apareix ja en un treball de Phragmén publicat el 1896 [42] (bastant anterior, per tant, al treball de Sainte-Laguë, que data de 1910). Més concretament, Phragmén va considerar aquest criteri en el cas de candidatures individuals i sota una restricció especial, és a dir, que cada candidat elegit sigui repartit de la mateixa manera entre tots els electors que l'han aprovat. Mitjançant un exemple va mostrar que aquesta restricció comporta una manca de monotonia respecte al contingut dels vots, és a dir, que un candidat pot deixar de ser elegit a conseqüència de rebre més aprovacions [42, p. 186–188]. Tanmateix, també va fer veure que si es prescindeix d'aquesta restricció aleshores sí que hi ha monotonia respecte al contingut dels vots [42, p. 188].

Però Phragmén considerava només el cas de candidatures individuals. A l'apartat 7.5 hem vist com el pas de candidatures individuals a candidatures col·lectives feia desaparèixer aquesta propietat de monotonia en la variant seqüencial. Tanmateix, les exploracions numèriques semblen indicar que *les variants directes es comportarien de manera monòtona respecte al contingut dels vots no solament en el cas de candidatures individuals, sinó també en el cas general de candidatures col·lectives*. En relació amb això cal precisar que quan hi ha multiplicitat de solucions el que variaria de manera monòtona no és el nombre d'escons que rep la candidatura en qüestió en una solució concreta, sinó el seu valor màxim al llarg de totes les solucions.

Malauradament, aquesta possible bona propietat de les variants directes ve acompanyada per certs inconvenients. El primer és la dificultat de càlcul, ja que

som davant d'un problema d'optimització que requereix algorismes bastant més complexos que el d'assignació successiva.

El segon inconvenient és la multiplicitat de solucions. Considerem, per exemple, el següent cas de dues candidatures provinent de l'apartat 7.7:

$$374 : A; \quad 250 : B; \quad 376 : A, B. \quad (57)$$

Per a $n = 10$, l'algorisme seqüencial dóna com a resultat el repartiment $(n_A, n_B) = (5, 5)$. En canvi, les variants directes donen quatre possibilitats, és a dir: $(n_A, n_B) = (4, 6), (5, 5), (6, 4), (7, 3)$. En efecte, resulta que tots aquests repartiments permeten aconseguir una representació totalment equitativa. Això és així perquè hi ha un gran nombre d'electors que accepten ser representats tant per A com per B. Per poc que es prodiguin les aprovacions a diverses candidatures per part d'un mateix elector, el nombre de repartiments diferents que s'obtenen amb les variants directes creix molt fàcilment amb el nombre de candidatures i amb el nombre total d'escons.

8.2 En l'exemple precedent, el resultat seqüencial $(5, 5)$ no es correspon amb el fet que hi ha bastants més electors partidaris d'A que de B. En aquest sentit podria ser més indicat el repartiment $(6, 4)$ que trobem entre els resultats directes. No obstant això, la mateixa lògica que ens ha portat a aquest repartiment també ens porta al repartiment $(4, 6)$, el qual no s'adiu gens a l'esmentat predomini de partidaris d'A. D'alguna manera ens hem passat de llarg.

Davant d'això, es fa necessari algun criteri addicional per a seleccionar adequadament entre els diversos repartiments que proporciona el criteri d'optimització directa. Una manera de fer-ho podria ser la següent. Tal com hem comentat a l'apartat 6.4, els vots que aproven diverses candidatures també podrien ser dividits a parts iguals entre aquestes. Si es procedeix d'aquesta manera, aleshores anem a parar a regles com les de Jefferson i D'Hondt i de Webster i Sainte-Laguë, on la multiplicitat de repartiments es limita a casos molt singulars (d'igualtat exacta entre certs nombres de vots o els seus múltiples); a canvi d'això, s'està renunciant a minimitzar tot el possible la desigualtat de representació entre els diferents electors. Doncs bé, es tractaria de situar-se en un punt intermedi on encara es minimitzés tot el possible la desigualtat però s'obtingués només un repartiment (o els menys possibles). Més concretament, es podrien combinar els dos tractaments mitjançant un paràmetre $\theta \in [0, 1]$ (on entendrem que $\theta = 1$ correspon a dividir cada vot per complet entre les candidatures aprovades en aquest vot) i escollir un valor d'aquest paràmetre on s'assoleixin alhora els dos objectius esmentats.

Els mètodes que apliquen aquesta idea els anomenem **variants directes relaxades**.

En l'exemple (57) aquest procediment selecciona el repartiment $(n_A, n_B) = (6, 4)$ que més amunt assenyalàvem com a més plausible. Aquest repartiment és l'únic que resulta òptim quan θ supera el valor 0.670. D'altra banda, permet una representació totalment equitativa fins a $\theta = 0.798$. Per sobre d'aquest últim

valor ja no és possible una representació totalment equitativa, però l'assignació (6, 4) encara minimitza μ i γ fins a $\theta = 1$.

En general, però, el resultat de les variants directes relaxades pot dependre de quina mesura de desigualtat s'utilitza. D'altra banda, també pot diferir del repartiment que minimitza la desigualtat per a $\theta = 1$.

Així, en l'exemple (1) amb $n = 4$ resulten elegits els candidats a, c, e, f o bé a, b, c, e segons que s'utilitzi respectivament μ o bé γ . Tanmateix, per a $\theta = 1$ ambdues mesures de desigualtat donen com a òptim el segon d'aquests dos conjunts.

Pel que fa a l'exemple (2) amb $n = 135$, s'obté el repartiment $\mathbf{n} = (40, 22, 19, 8, 14, 20, 12)$ tant en el cas de μ com en el cas de γ . En canvi, per a $\theta = 1$ aquestes dues mesures de desigualtat donen, respectivament, els repartiments (40, 23, 18, 8, 14, 20, 12) i (40, 22, 18, 9, 14, 20, 12).

Aquests resultats s'han obtingut mitjançant el programari d'optimització CPLEX [26]. Tanmateix, la gran multiplicitat de solucions no deixa de ser una font d'incertesa, de manera que un estudi més conclusiu hauria de començar per millorar el procediment de càlcul.

9 Conclusions

El vot d'aprovació permet que els electors donin més informació que amb el vot uninominal. En conseqüència, els resultats poden reflectir millor les diverses opinions dels electors. Tanmateix, perquè sigui així calen procediments més laboriosos que els habituals.

Les opcions que són objecte d'aprovació poden ser candidats individuals o bé candidatures col·lectives. El cas de candidats individuals és apropiat en eleccions a petita escala, on els electors tenen una coneixença personalitzada dels candidats. En canvi, en eleccions d'abast ampli és més adient que les opcions per aprovar siguin les candidatures col·lectives presentades pels partits. En aquest cas, l'objectiu és determinar quants escons ha de rebre cada candidatura.

A finals del segle XIX Phragmén va proposar un mètode iteratiu que s'estén fàcilment al cas de candidatures col·lectives. Tanmateix, en aquest cas apareixen alguns fenòmens indesitjables, com ara la manca de monotonia respecte al contingut dels vots i una tendència a igualar més del compte els nombres d'escons que reben les diferents candidatures.

Hi ha indicis que aquests problemes es podrien resoldre mitjançant certes variants que s'han apuntat en aquest article i que convindria estudiar més detingudament.

Tots aquests mètodes són més complicats del que hom voldria. Tanmateix, el seu grau de complicació és comparable a certes variants del mètode de vot únic transferible que són utilitzades per alguns estats i institucions.

En relació amb això, Phragmén afirmava el següent: *tot el que és correcte es pot fer també fàcilment comprensible, encara que per aconseguir-ho pot fer falta temps i feina; a més, el públic en general aprecia millor els resultats que les teories* [41].

A Nota biogràfica [9, 6, 19, 21]

Edvard Phragmén va néixer l'any 1863 a Örebro (Suècia) i va morir el 1937 a Estocolm. El seu nom complet era Lars Edvard Phragmén. Des del segle XVII els seus avantpassats per línia masculina havien sigut pastors protestants, però el seu pare, Lars Phragmén (1832-1920), va trencar amb aquesta tradició i es va dedicar a les matemàtiques.

Seguint aquests passos, el 1882 Edvard Phragmén va ingressar a la Universitat d'Uppsala per cursar estudis de matemàtiques. L'any següent, però, va optar per l'escola superior (*högskola*) d'Estocolm, on no s'obtenia cap mena de títol però hi havia un ambient d'estudi molt autèntic i desinteressat. En particular, l'influent professor Gösta Mittag-Leffler liderava una recerca capdavantera en anàlisi matemàtica [12, 20]. A partir d'aquell mateix any 1883, Phragmén va contribuir a aquesta recerca amb diversos treballs [6].

A principis de l'any 1888, Mittag-Leffler va confiar a Phragmén el càrrec de secretari de redacció de la revista *Acta Mathematica*, una tasca que Phragmén va desenvolupar molt a consciència. Tant va ser així que a finals del mateix any va detectar certs errors en una memòria d'Henri Poincaré sobre el problema dels tres cossos la qual ja havia estat escollida guanyadora d'un important concurs, convocat pel rei Òscar II de Suècia i Noruega amb motiu del seu seixantè aniversari [9, 1]. Els errors en qüestió eren realment seriosos, fins al punt que Mittag-Leffler va decidir aturar l'edició i recuperar unes quantes còpies que ja s'havien distribuït en forma impresa. Poincaré es va veure obligat a fer canvis substancials en la memòria i a costejar les despeses originades, les quals superaven amb escreix l'import del premi. En la versió definitiva, que no va ser publicada fins al 1890, Poincaré expressava així el seu agraïment a Phragmén [46, p. 5]:

Je dois beaucoup de reconnaissance à M. Phragmén qui non seulement a revu les épreuves avec beaucoup de soin, mais qui, ayant lu le mémoire avec attention et en ayant pénétré le sens avec une grande finesse, m'a signalé les points où des explications complémentaires lui semblaient nécessaires pour faciliter l'entière intelligence de ma pensée. Je lui dois la forme élégante que je donne au calcul de S_i^m et de T_i^m à la fin du § 12. C'est même lui qui, en appelant mon attention sur un point délicat, m'a permis de découvrir et de rectifier une importante erreur.

El fet d'haver corregit la memòria de Poincaré va donar molt de prestigi a Phragmén, que el desembre de 1889 obtenia el grau de llicenciat (*filosofie licentiat*) a l'Universitat d'Uppsala. El febrer de 1890 va ser nomenat professor associat (*docent*) de matemàtiques i mecànica a l'escola superior d'Estocolm, i el setembre de 1892 va esdevenir-hi professor d'anàlisi matemàtica superior, succeint en aquesta plaça Sofia Kovalevskaya, que havia mort el febrer de 1891 d'una pneumònia.



Edvard Phragmén
(retrat procedent de l'Institut Mittag-Leffler;
fotògraf: Herman Hamnqvist).

En el període 1892–1895 Phragmén va formar part dels òrgans de govern de l'escola. En aquell moment hi havia molta polèmica ja que alguns professors, entre ells el rector Otto Pettersson, que era professor de química, proposaven que l'escola atorgués títols acadèmics, mentre que uns altres s'hi oposaven, entre ells Mittag-Leffler i Phragmén. Aquest tema va dominar unes accidentades eleccions a rector que van tenir lloc el desembre de 1894 i gener de 1895, en les quals es va barrejar també la qüestió de qui tenia dret de vot en el claustre. De fet, es va arribar a votar tres vegades. La segona hauria donat el càrrec a Phragmén. Tanmateix, el resultat definitiu va ser la reelecció d'Otto Pettersson. Amb ell i els seus successors, l'escola va evolucionar progressivament cap a l'atorgament de títols acadèmics.

Alhora que continuava treballant en els temes més clàssics d'anàlisi matemàtica, Phragmén es va interessar cada vegada més en altres temes més aplicats.

Entre els quals hi ha el que hem tractat en el present article: com elegir un conjunt de representants quan els electors s'expressen mitjançant vots d'aprovació o preferencials. La seva primera proposta sobre aquest problema data de març de 1893 [37] i va ser presentada en el marc de l'entitat *Studenten och Arbetare* —Estudiants i Treballadors—, la qual tenia com a objectiu la col·laboració entre els intel·lectuals i la classe obrera. Phragmén va continuar estudiant el problema en quatre publicacions que van aparèixer de l'any 1894 al 1899 [39, 40, 42, 43]. Posteriorment, hi va contribuir també des de dues comissions estatals de les quals va formar part els anys 1902–1903 i 1912–1913 [50, 48]. Entre aquestes dues comissions, l'any 1906 també hi va fer una nova aportació amb motiu d'un informe sobre el tema que havia demanat la comissió finlandesa per a la reforma electoral [44, 45].



La comissió reial sueca de 1902–1903 sobre els drets de vot en la seva primera reunió.

D'esquerra a dreta: Sixten G. von Friesen, Carl O. Moberg, Ivar Månsson i Trää, Axel Asker, Christofer Johan Rappe, Gustaf J. G. A. Berg, J. Elof Biesèrt, L. Edvard Phragmén i Emil Svensén

(Stockholms Stadsmuseum, imatge SSMD000156; fotògraf: Anton Blomberg [32]).

Cal notar que pocs dies després de la presentació de Phragmén esmentada en l'associació d'Estudiants i Treballadors, el mateix Gösta Mittag-Leffler donava una conferència sobre el tema de la representació proporcional davant l'associació sueca d'economistes [33]. A més, el 1906 va escriure l'informe per a la comissió finlandesa que ja hem esmentat [34]. Altres matemàtics de l'entorn de Phragmén que van prestar atenció al tema de la representació proporcional són Ivar Bendixson [2, 3, 4], Anders Lindstedt [30, 31] i Ernst Lindelöf [29].

Els noms de Phragmén i Lindelöf estan associats a un celebrat resultat d'anàlisi complexa i d'equacions en derivades parcials que generalitza a la vegada el principi del màxim i el teorema de Cauchy i Liouville. Aquest resultat data de 1908 i estén un treball previ de Phragmén de 1904 (vegeu [6]).

Un altre camp que va atraure l'atenció de Phragmén, i que finalment va absorbir-lo fortament, va ser la matemàtica actuarial, un camp on hi havia llavors molta demanda i en el qual també va estar ocupat el mateix Mittag-Leffler [9, 21]. Un moment crucial en la vida de Phragmén va ser l'inici de l'any 1904, quan, després d'haver treballat com a actuari i com a inspector estatal d'assegurances, va esdevenir director general d'un nou organisme d'inspecció d'assegurances (1904–1908) [21]. Phragmén s'havia imaginat que aquest càrrec podia ser compatible amb la seva dedicació acadèmica, però la realitat va ser una altra i el maig de 1905 va deixar definitivament la plaça de professor a l'escola superior d'Estocolm. Posteriorment, va presidir l'Associació Sueca d'Actuaris (1909–1934, com a successor de Mittag-Leffler), va formar part del consell directiu de la l'Associació Sueca d'Assegurances (1912–1935) i va presidir el consell directiu de la Societat Anònima Estatal de Reassegurances (1914–1937).

B Diverses variants considerades per Phragmén

Les propostes de Phragmén sobre representació proporcional inclouen una diversitat de variants. Aquestes es poden classificar en dues grans classes, 1 i 2, que corresponen essencialment als mètodes «primer» i «segon» de Phragmén en la terminologia de [8] (on aquests termes són aplicats concretament a les variants 1.2 i 2.1).

1. *Procediments de quota preestablerta.*

Aquests procediments es poden considerar com a variants del mètode de vot únic transferible. Com aquest, comencen per establir la quota, és a dir, el nombre de vots que cal reunir per assignar un escó. També coincideixen amb el vot únic transferible en la idea de transferir els excedents als altres candidats indicats en les butlletes. Els escons s'assignen successivament, fixant-se sempre en el candidat que reuneix més vots, originals o transferits. Tanmateix, si en algun moment s'escau que cap candidat arriba a la quota, aleshores *no* es procedeix a eliminar el candidat que té menys vots, sinó que es relaxa la condició de superar o igualar la quota inicial. En particular, en el cas de voler elegir un sol representant aquests procediments escullen sempre el més votat.

1.1. Variant preferencial inicial (1893) [37, 38]. Aquesta primera proposta considerava només vots preferencials. La quota utilitzada és inicialment la quota de Droop, però si en algun moment s'escau que cap candidat arriba a la quota, aleshores aquesta es disminueix com si es volgués un representant més. De manera molt similar a com ja havia proposat J. B. Gregory a Tasmània pels volts de 1880 ([56, p. 273], [28, § 12.1.4]), quan un candidat supera la quota i

obté un escó, l'excedent es reparteix per igual entre tots els vots que tenia aquell candidat i de cada vot es transfereix la fracció excedent al següent candidat.

1.2. Variant de tipus restes majors (introduïda segurament entre 1893 i 1894). S'exposa a [8, § A.b, p. 47–50]. A banda de considerar vots d'aprovació, es diferencia de la variant precedent pel fet que utilitza la quota simple de Hare i que quan cap candidat arriba a la quota aleshores se selecciona el candidat que en aquell moment reuneix més vots. En el cas de llistes de partit es redueix al mètode de les restes majors. Gustaf Eneström va arribar independentment a aquesta mateixa idea el 1896 arran de voler simplificar la variant 2.1 [17, 18].

1.3. Variant final (1906) [44, 45]. També utilitza la quota simple, però la quantitat de vots que es bescanvia per un escó sol ser inferior a la quota. Concretament, si v representa el nombre de vots en qüestió, la part que es bescanvia per un escó és $v/(k+1)$ on k és l'únic enter tal que $kq_0 < v \leq (k+1)q_0$. Aquesta especificació correspon a la versió definitiva [45]; en la versió anterior que es dona a [44] la quota simple es recalcula cada vegada en termes del nombre de vots que resten actius i el nombre d'escons encara no assignats.

1.4. Variants que distingeixen dos graus d'aprovació [8, 44, 45]. En prevenció de les estratègies de decapitació (apartat 4.2) Phragmén va suplementar les dues variants precedents amb la possibilitat que el votant distingeixi dos graus d'aprovació (la qual cosa es pot considerar com una forma especial de vot preferencial). Quan es vota així, llavors els candidats secundaris d'un vot no entren en joc fins que no hagin estat seleccionats tots els candidats primaris d'aquell vot.

2. Procediments d'optimització.

Aquests procediments miren de minimitzar la desigualtat de representació entre els diferents electors.

2.1. Variant seqüencial (1894–1899) [39, 40, 42, 43]. És la que nosaltres estudiem a la secció 7. Considera vots d'aprovació i assigna els escons de manera successiva. Cada escó s'assigna de manera que en aquell moment es minimitzi una certa mesura de la desigualtat de representació entre els diferents electors.

2.2. Variants directes (1896). Els escons no s'assignen de manera successiva, sinó que es distribueixen tots d'una vegada de manera que es minimitzi alguna mesura especificada de la desigualtat de representació entre els diferents electors. En general, la determinació d'aquesta distribució requereix algorismes d'optimització més complexos que els algorismes d'assignació successiva. Tot i això, Phragmén va adoptar aquest punt de vista directe o global en alguna ocasió, com ara a [42, p. 184, 188]. Nosaltres també l'adoptem a la secció 8.

2.3. *Variant preferencial relativa* (1895) [40, § 7]. Aquesta proposta consistia a substituir un vot preferencial de la forma $a > b > c > d > e$ per un vot d'aprovació de la forma a, b, c, d, e més una fracció especificada, com ara $1/10$, repartida a parts iguals entre les llistes a, b, c, d ; a, b, c ; a, b ; a .

2.4. *Variant preferencial absoluta* (1903–1913). Utilitza el procediment seqüencial, però els vots preferencials són tractats amb el caràcter «absolut» següent (que generalitza la idea de la variant 1.4): el candidat $(k + 1)$ -èsim d'un vot preferencial no entra en joc fins que no han estat seleccionats els k candidats anteriors. En una primera versió, la informació preferencial es reduïa a distingir dos graus d'aprovació [8, p. 56, 62–64], però després es va admetre una ordenació qualsevol dels candidats aprovats [48, p. 47–54]. Aquesta versió general havia estat considerada ja el 1910 per Nore Tenow [54]. L'algorisme final es descriu també a [28, § 12.3.5, 13]. A diferència dels mètodes de vot únic transferible, en el cas $n = 1$ no es redueix al vot alternatiu (apartat 3.3), sinó a elegir el candidat que rep més vots de primera opció. Tanmateix, coincideix amb una determinada variant del vot únic transferible —l'anomenat mètode de Gregory inclusiu— si s'escull adequadament la quota d'aquest últim i l'ordre de les transferències [28, sats 13.1].

Aquesta variant és monòtona respecte a variacions dels vots a favor d'un determinat candidat. Tanmateix, no compleix el criteri de proporcionalitat de Droop. Per tal de garantir aquest últim, l'any 2002 Olli Salmi va proposar una modificació mitjançant certs elements del vot únic transferible [52, 53]. Aquesta modificació ha estat estudiada detalladament per Douglas Woodall [60], que en distingeix dues versions i demostra que totes dues compleixen el criteri de proporcionalitat de Droop. Un altre estudi en aquesta direcció és el de Ross Hyman [25]. Tanmateix, en el cas $n = 1$ aquestes modificacions es redueixen al vot alternatiu, la qual cosa implica la indesitjable manca de monotonia que hem comentat a l'apartat 3.6.

La variant preferencial absoluta va acabar formant part del sistema electoral suec, on va ser introduïda el 1921 i encara hi és present actualment [28, § 13.10, C.2.3]. Tanmateix, no afecta el nombre d'escons que són assignats a cada candidatura, sinó només quins candidats concrets reben aquests escons. De fet, l'elector ha d'optar per una sola candidatura, tot i que pot variar fins a un cert punt la llista ordenada de candidats. El nombre d'escons que obté cada candidatura es determinava inicialment mitjançant la regla de Jefferson i D'Hondt, però el 1952 aquesta va ser substituïda per una certa modificació de la regla de Webster i Sainte-Laguë. Per a més detalls remetem el lector a [28].

Agraïments

Els autors volem manifestar el nostre agraïment al senyor Mikael Rågstedt, bibliotecari de l'Institut Mittag-Leffler, per haver-nos ajudat a localitzar diversos documents. També estem molt agraïts al senyor Mikael Abrahamsson, biblio-

tecari del Centre de Ciències Matemàtiques de la Universitat de Lund, pel seu ajut a obtenir còpies de diversos treballs de Phragmén i d'altres autors. Els professors Magnus Fontes i Joachim Kock ens han ajudat també en la traducció de textos nòrdics. Moltes gràcies a tots dos. També volem donar les gràcies a la professora Rosa Camps, la qual ens ha ajudat a entendre les idees de Phragmén i a resoldre diverses qüestions tècniques. Finalment, agraïm també els comentaris dels professors Aureli Alabert i Svante Janson, els quals ens han permès millorar el contingut en diversos punts.

Referències

- [1] BARROW-GREEN, J. *Poincaré and the three body problem*. Providence: American Mathematical Society; Londres: London Mathematical Society, 1997. (History of Mathematics; 11)
- [2] BENDIXSON, I. *Regeringens förslag till proportionellt valsätt och dess verkningar. Framställning och kritik*. Estocolm: Nordiska Bokhandeln, 1905.
- [3] BENDIXSON, I. «Rösträttsfrågan». Conferència donada el 2 de novembre de 1906 en l'associació Stockholms Högskolas Studentföreningen. Resum manuscrit de la conferència i de la discussió, segons notes d'Adolf Larsson: Estocolm, Kungliga Biblioteket, Gösta Mittag-Leffler papper, L62:55, núm. 6.
- [4] [BENDIXSON, I.] «Promemoria rörande regeringens proportionella valmetod, af en antiproportionalist». *Frisinnade Landsföreningens Meddelanden* (Estocolm), núm. 52 (8) (1907).
- [5] BRAMS, S. J. *Mathematics and democracy. Designing better voting and fair-division procedures*. Princeton: Princeton University Press, 2008.
- [6] CARLEMAN, T. «L. E. Phragmén in memoriam». *Acta Math.*, 69 (1) (1938), XXXI-XXXIII.
- [7] CARSTAIRS, A. M. *A Short History of Electoral Systems in Europe*. Londres: George Allen & Unwin, 1980.
- [8] CASSEL, C. G. *Proportionella Val. Systematisk framställning*. Apèndix I de [50]. Estocolm: Isaac Marcus' Boktryckeri-Aktiebolag, 1903.
- [9] CRAMÉR, H. «Lars Edvard Phragmén: Minnesteckning». *Kungliga Svenska Vetenskapsakademiens Årsbok* (1958), 279-301. Reproduït a *Levnadsteckningar över Kungliga Svenska Vetenskapsakademiens Ledamöter*, 9 (Estocolm: Almqvist & Wiksell, 1962), 255-277.
- [10] [D'HONDT, V.] *La Représentation Proportionnelle des Partis, par un électeur*. Brussel·les: Bruylant-Christophe, 1878.
- [11] D'HONDT, V. *Système Pratique et Raisonné de Représentation Proportionnelle*. Gant: J. S. Van Dooselaere; Brussel·les: C. Muquardt, 1882.

- [12] DOMAR, Y. «Matematisk forskning under Stockholms högskolas första decennier». Univ. Stockholm, 1981. Reproduït a <http://www2.math.su.se/matematik/Historia/Historia.pdf>. Traducció anglesa: «Mathematical research during the first decades of the University of Stockholm». <http://www2.math.su.se/matematik/Historia/Histeng.pdf>.
- [13] DORON, G.; KRONICK, R. «Single transferable vote: an example of a perverse social choice function». *American Journal of Political Science*, 21 (1977), 303–311.
- [14] DOVGOSHEY, O.; MARTIO, O.; RYAZANOV, V.; VUORINEN, M. «The Cantor function». *Expo. Math.*, 24 (1) (2006), 1–37.
- [15] DUMMETT, M. A. E. *Voting Procedures*. Oxford: Clarendon Press, 1984.
- [16] DUMMETT, M. A. E. «Towards a more representative voting system: The Plant Report». *New Left Rev.*, 1/194 (1992), 98–113.
- [17] ENESTRÖM, G. H. Carta a l'editor. *Aftonbladet* (23 febrer 1896).
- [18] ENESTRÖM, G. H. «Om aritmetiska och statistiska metoder för proportionella val». *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 53 (1896), 543–570.
- [19] ENGLUND, K. «L Edward Phragmén». *Svenskt Biografiskt Lexikon*, 29 (1995–1997), 303–306. Reproduït a <http://sok.riksarkivet.se/sbl/artikel/7271>.
- [20] GÅRDING, L. *Matematik och Matematiker. Matematiken i Sverige före 1950*. Lund: Lund University Press, 1994. Traducció anglesa: *Mathematics and mathematicians. Mathematics in Sweden before 1950*. Providence: American Mathematical Society; Londres: London Mathematical Society, 1998. (History of Mathematics; 13)
- [21] HAGBERG, J. «Svenska Aktuarietföreningen 100 år. Aktuarier i svensk försäkring». *Nordisk Forsikringstidskrift*, 4 (2005), 377–385.
- [22] HILL, I. D. «Election method for Council». Royal Statistical Society (1995). <http://www.rss.org.uk/site/cms/contentviewarticle.asp?article=1005>.
- [23] HILL, I. D. «Party lists and preference voting». *Voting Matters*, 29 (2011), 15–19.
- [24] HOAG, C. G.; HALLETT, JR., G. H. *Proportional Representation*. Nova York: Macmillan, 1926.
- [25] HYMAN, R. «Divisor method proportional representation in preference-ballot elections». *Voting Matters*, 28 (2011), 15–30.
- [26] IBM. «CPLEX Optimizer. High-performance mathematical programming solver for linear programming, mixed integer programming, and quadratic programming». <http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/index.html>.

- [27] IRELAND GOVERNMENT. Department of the Environment, Heritage and Local Government. *Guide to Ireland's PR-STV Electoral System*. <http://www.environment.ie/en/LocalGovernment/Voting/PublicationsDocuments/FileDownload,1895,en.pdf>.
- [28] JANSON, S. *Proportionella Valmetoder*. <http://www2.math.uu.se/~svante/papers/sjv6.pdf>, 2012–2014.
- [29] LINDELÖF, E. «Utkast till ett på proportionalitetsidéen grundadt valsätt, jämte förklaring och motivering af de principer som ligga till grund för detsamma». Manuscrit de 1906. Estocolm: Kungliga Biblioteket: Gösta Mittag-Leffler papper, L62:55, núm. 3.
- [30] LINDSTEDT, A. «Promemoria angående metoder för proportionella val». Estocolm: Kungliga Boktryckeriet, 1906.
- [31] LINDSTEDT, A. «Förslag till metod för proportionella val». Estocolm: Kungliga Boktryckeriet, 1906.
- [32] MILLQVIST, V. «Rösträttskommittén». *Idun. Illustrerad Tidning för Kvinnan och Hemmet*, 15 (45) (1902), 719–720.
- [33] MITTAG-LEFFLER, M. G. «Proportionella val». Conferència donada el 16 de març de 1893 en l'associació Nationalekonomiska Föreningen. Referència: *Nationalekonomiska Föreningens Förhandlingar*, 1893 (1893), p. 44.
- [34] MITTAG-LEFFLER, M. G. «Resultatet av en granskning av A. Lindstedts och E. Phragmén's förslag till proportionell valmetod». Informe dirigit al Ministeri de Justícia de Finlàndia. Manuscrit de 1906. Estocolm: Kungliga Biblioteket: Gösta Mittag-Leffler papper, L62:55, núm. 2.
- [35] MONROE, B. L. «Fully proportional representation». *American Political Science Review*, 89 (1995), 925–940.
- [36] MORA, X. «La regla de Jefferson-D'Hondt i les seves alternatives». *MatMat*, 2013 (4), 34 p. <http://mat.uab.cat/matmat/PDFv2013/v2013n04.pdf>.
- [37] PHRAGMÉN, E. «Om proportionella val». Conferència donada en l'associació Studenter och Arbetare (1893). Resum: *Stockholms Dagblad* (14 març 1893).
- [38] PHRAGMÉN, E. «Öfver proportionella val». Contingut: Quatre taules amb comentaris, 13 pàgines impreses, circa 1893, nom de l'autor i títol escrits a llapis. Institut Mittag-Leffler, Biblioteket, Sep. Coll., Sv. 9 [Phragmén].
- [39] PHRAGMÉN, E. «Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis». *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 51 (3) (1894), 133–137.
- [40] PHRAGMÉN, E. *Proportionella Val. En valteknisk studie*. Estocolm: Lars Hökerbergs, 1895. (Svenska Spörsmål; 25)
- [41] PHRAGMÉN, E. Carta a l'editor. *Aftonbladet* (23 febrer 1896).
- [42] PHRAGMÉN, E. «Sur la théorie des élections multiples». *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 53 (1896), 181–191.
- [43] PHRAGMÉN, E. «Till frågan om en proportionell valmethod». *Statsvetenskaplig Tidskrift*, 2 (1899), 297–308 (núm. 2, 88–98).

- [44] PHRAGMÉN, E. «Utkast till föreskrifter beträffande val av riksdagsmän i andra kammaren och deras suppleanter». Estocolm: Kungliga Boktryckeriet, 1906.
- [45] PHRAGMÉN, E. «Promemoria beträffande en förenklad form af den af under-teknad föreslagna valmetoden». Manuscrit de 1906. Estocolm: Kungliga Biblioteket: Gösta Mittag-Leffler papper, L62:55, núm. 4.
- [46] POINCARÉ, H. «Sur le problème des trois corps et les équations de la dynamique». *Acta Math.*, 13 (1890), A3–A270.
- [47] POTTHOFF, R. F.; BRAMS, S. J. «Proportional representation: Broadening the options». *Journal of Theoretical Politics*, 10 (1998), 147–178.
- [48] PROPORTIONSVALKOMMITTÉN [VON FRIESEN, S.; APPELBERG, G.; BENDIXSON, I.; PHRAGMÉN, E.]. *Betänkande angående ändringar i gällande bestämmelser om den proportionella valmetoden*. Estocolm: Isaac Marcus' Boktryckeri-Aktiebolag, 1913.
- [49] PUKELSHEIM, F. *Proportional representation. Apportionment methods and their applications*. Cham: Springer, 2014.
- [50] RÖSTRÄTTSKOMMITTÉN [ASKER, A.; BERG, G. J. G. A.; BIESÈRT, J. E.; VON FRIESEN, S. G.; MÅNSSON, I.; MOBERG, C. O.; PHRAGMÉN, L. E.; RAPPE, C. J.; SVENSÉN, E.; CASSEL, C. G.] *Betänkande med förslag till proportionellt valsätt vid val till Riksdagens andra kammare jämte vissa ändringar i regeringsformen och riksdagsordningen*. Estocolm: Hæggström, 1903.
- [51] ROVIRA I VIRGILI, A. *Els Sistemes Electorals*. Barcelona: Barcino, 1932. 2a ed: Barcelona, Ed. Undarius, 1977.
- [52] SALMI, O. «D'Hondt without lists». *Election Methods Mailing List*. <http://lists.electorama.com/pipermail/election-methods-electorama.com/2002-September/073943.html>, 2002. Vegeu també *ibidem* 2002-August/073862.html, 2002-November/074151.html, 2002-November/106914.html i 2003-August/108863.html.
- [53] SALMI, O. «Henkilökohtainen suhteellinen vaali». <http://www.uusikaupunki.fi/~olsalmi/vaalit/henkkoht.html>, 2004.
- [54] TENOW, N. B. «En proportionell rangmetod, afsedd att på de proportionella valens område utgöra en motsvarighet till den vanliga formen för majoritetsval i enmansvalkretsar». *Statsvetenskaplig Tidskrift*, 13 (1910), 317–325.
- [55] TIDEMAN, T. N. «The single transferable vote». *Journal of Economic Perspectives*, 9 (1995), 27–38.
- [56] TIDEMAN, T. N. *Collective Decisions and Voting*. Ashgate, 2006.
- [57] WOODALL, D. R. «An impossibility theorem for electoral systems». *Discrete Math.*, 66 (1-2) (1987), 209–211.
- [58] WOODALL, D. R. «Properties of preferential election rules». *Voting Matters*, 3 (1994), 8–15.

- [59] WOODALL, D. R. «Monotonicity. An in-depth study of one example». *Voting Matters*, 4 (1995), 5–7.
- [60] WOODALL, D. R. «QPQ, a quota-preferential STV-like election rule». *Voting Matters*, 17 (2003), 1–7.

XAVIER MORA
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT AUTÒNOMA DE BARCELONA
CATALONIA, SPAIN
xmora@mat.uab.cat

MARIA OLIVER
DEPARTAMENT DE TECNOLOGIES DE LA INFORMACIÓ I LES COMUNICACIONS
UNIVERSITAT POMPEU FABRA
CATALONIA, SPAIN
maria.oliverp@upf.edu

English summaries

Carles M. Cuadras

The legacy of Galton, Pearson, Fréchet and others: How to measure and interpret statistical association

Along three parts, we present the concepts of correlation and statistical association, starting with Galton's notion of correlation, improved by Pearson. As an illustration, we use the classic data of Galton and Pearson on heritability of parents and children regarding the stature. The second part explains how to study the same data from a multivariate perspective (canonical correlation analysis, correspondence analysis). We also use data from Fisher. We show how to associate general data sets with distances. The third part is devoted to the bivariate distributions. We present the theory of eigenfunctions and eigenvalues for two kernels, which is applied to the diagonal expansion of a bivariate distribution, including continuous expansions in terms of integrals. We propose a family of canonical copulas, which can generate bivariate distributions.

Keywords: Galton's and Pearson's data, intraclass correlation, canonical correlation, correspondence analysis, association based on distances, integral operators, distributions with given marginal, canonical functions.

MSC2010 Subject Classification: 62H20, 60E05.

Xavier Mora and Maria Oliver

Elections by means of approval voting. Phragmén's method and some variants of it

We study certain variants of a method for proportional representation that was proposed at the end of the nineteenth century by Edvard Phragmén. Every elector expresses his opinion by means of an approval ballot, where he indicates all the options that he deems suitable. The options can be individual candidates, as it was considered by Phragmén, or party candidatures which can provide several representatives each. Given a set of votes of this kind and a prefixed number of seats, the aim is to distribute these seats among the candidatures so as to achieve as much representativeness as possible. This work pays special attention to a condition of monotonicity which requires that adding approvals in favour of a certain candidature cannot result in its getting a smaller number of seats. In the basic variant, where seats are assigned by means of a sequential procedure, such a monotonicity holds in the case of individual candidates but not in the general case. We briefly explore also some direct variants, where seats are assigned all at once by means of a suitable optimization criterion. These variants seem to be better from the point of view of monotonicity. However, they suffer from a problem of multiplicity of solutions that calls for a supplementary criterion to choose among them.

Keywords: proportional representation, approval voting, open lists, Edvard Phragmén.

MSC2010 Subject Classification: 91B12.

Instruccions per als autors

Els articles sotmesos a publicació s'han d'enviar als editors o a qualsevol membre del comitè editorial, per correu electrònic, preferentment en format PDF. Els originals han de contenir la versió anglesa del títol, un resum breu en català i en anglès, paraules clau en català i en anglès i els codis de la classificació per matèries MSC2010.

Les versions definitives dels articles acceptats s'han de presentar en codi $\text{T}_{\text{E}}\text{X}$, preferentment en l'estil $\text{\LaTeX}$ propi del BUTLLETÍ. Aquest estil es pot obtenir a les pàgines web de la Societat Catalana de Matemàtiques (SCM). Fem notar que en aquesta publicació s'utilitza preferentment el punt per a separar decimals, en lloc de la coma recomanada per l'IEC, per poder facilitar la comprensió de les expressions matemàtiques. Per tal d'accelerar el procés de producció, es prega als autors que segueixin les indicacions contingudes en el document d'exemple.

La versió en paper del BUTLLETÍ s'imprimeix en blanc i negre. Quan un article contingui figures en color i es consideri convenient, l'autor proporcionarà una versió dels gràfics substituint el color per tons de grisos i línies de gruix variable. Així mateix modificarà els comentaris que facin referència al color de les figures. En qualsevol cas el BUTLLETÍ publicarà l'original en color en el seu format electrònic.

Els autors dels articles publicats al BUTLLETÍ en retenen el dret de còpia (*copyright*) i autoritzen l'IEC a difondre'ls, tant a través de la publicació impresa com mitjançant els portals digitals propis o d'altres amb què s'estableixin els convenis oportuns a aquest efecte. És responsabilitat dels autors assegurar que es disposa dels drets de reproducció dels gràfics i de les figures que hi apareguin. Cada autor rebrà una còpia en PDF d'alta qualitat de la versió digital del seu article i un exemplar imprès del número del BUTLLETÍ en el qual es publiqui.

La correspondència administrativa relacionada amb el BUTLLETÍ s'ha d'adreçar a la SCM.

Comitè editorial

Julia Cufí (editor en cap)
Departament de Matemàtiques
Universitat Autònoma de Barcelona
jcufi@mat.uab.cat

Bartomeu Coll
Dep. de Matemàtiques i Informàtica
Universitat de les Illes Balears
tomeu.coll@uib.cat

Núria Fagella
Dep. de Matemàtica Aplicada i Anàlisi
Universitat de Barcelona
fagella@maia.ub.es

Josep Maria Font
Dep. de Probabilitats, Lògica i Estadística
Universitat de Barcelona
jmfont@ub.edu

Armengol Gasull
Departament de Matemàtiques
Universitat Autònoma de Barcelona
gasull@mat.uab.cat

Gábor Lugosi
ICREA i Departament d'Economia
Universitat Pompeu Fabra
gabor.lugosi@upf.edu

Rosa Camps (editora adjunta)
Departament de Matemàtiques
Universitat Autònoma de Barcelona
rcamps@mat.uab.cat

Jorge Mateu
Departament de Matemàtiques
Universitat Jaume I
mateu@mat.uji.es

Marc Noy
Departament de Matemàtica Aplicada II
Universitat Politècnica de Catalunya
marc.noy@upc.edu

Francesc Planas
Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya
francesc.planas@upc.edu

Agustí Reventós
Departament de Matemàtiques
Universitat Autònoma de Barcelona
agusti@mat.uab.cat

Marta Sanz-Solé
Dep. de Probabilitats, Lògica i Estadística
Universitat de Barcelona
marta.sanz@ub.edu



Societat Catalana de Matemàtiques

La **Societat Catalana de Matemàtiques (SCM)** és una societat filial de l'Institut d'Estudis Catalans, que continua les activitats de la Secció de Matemàtiques de la Societat Catalana de Ciències, que fou fundada per l'Institut l'any 1931. Les finalitats de la **SCM** són: el conreu de les ciències matemàtiques, l'extensió del seu coneixement en la societat catalana, el foment del seu ensenyament i de la seva investigació teòrica i aplicada, així com la publicació de tota mena de treballs que s'adeqüin a aquests objectius. La **SCM** desenvolupa les seves activitats en les terres de llengua i cultura catalanes. El català és, doncs, la llengua pròpia de la **SCM** i la que és usada normalment en tots els seus actes i publicacions.

La **SCM** edita les publicacions periòdiques *SCM/Notícies* i *Butlletí de la Societat Catalana de Matemàtiques*. Els socis de la **SCM** reben, gratuïtament, aquestes dues publicacions.

La **SCM** té convenis de reciprocitat amb diverses societats matemàtiques d'arreu del món, mitjançant els quals els socis de la **SCM** obtenen una reducció en la quota de soci d'aquestes societats. Així mateix, els socis de la **SCM** poden fer-se socis de la Societat Matemàtica Europea pagant una quota complementària.

La Junta Directiva de la **SCM** està constituïda per les persones següents:

PRESIDENT: Xavier Jarque i Ribera

VICEPRESIDENT: Enric Ventura i Capell

ADJUNTA DE LA VICEPRESIDÈNCIA: Iolanda Guevara i Casanova

SECRETARI: Albert Ruiz i Cirera

TRESORERA: Natàlia Castellana i Vila

VOCALS: Núria Fagella i Rabionet, Josep Grané i Manlleu, Agustí Reventós i Tarrida, Carles Romero i Chesa, Oriol Serra i Albó, Esther Silberstein, Manel Udina i Abelló

DELEGAT DE L'IEC: Joan Girbau i Badó

L'adreça de la **SCM** és carrer del Carme, 47, 08001 Barcelona. Telèfon: 933 248 583. Fax: 932 701 180. Correu electrònic: scm@iec.cat. Adreça web: <http://scm.iec.cat>.

El Butlletí de la Societat Catalana de Matemàtiques publica, en llengua catalana, exposicions matemàtiques de qualitat, que puguin interessar a un nombre elevat de lectors. Es donarà prioritat a aquells treballs en què destaquin la claredat d'exposició i l'interès general del tema. El Butlletí està obert a tots els camps de la matemàtica i també als aspectes matemàtics de les ciències experimentals, la tecnologia, l'economia, etc., així com a altres àrees, com la història, la didàctica i la filosofia, sempre que els treballs tinguin un component matemàtic important. També tenen cabuda al Butlletí aquells articles que desenvolupin un aspecte significatiu de la problemàtica de la professió matemàtica al nostre país.

El Butlletí publica un volum a l'any, dividit en dos números, que es trameten gratuïtament a tots els socis. El Butlletí es publica també en format electrònic. L'edició electrònica del Butlletí pot obtenir-se des del portal de revistes científiques en línia de l'IEC o al servidor <http://scm.iec.cat>.

La correspondència administrativa s'ha d'adreçar a la Societat Catalana de Matemàtiques.

Editor en cap _____

Julia Cufí

Departament de Matemàtiques

Universitat Autònoma de Barcelona

Editora adjunta _____

Rosa Camps

Departament de Matemàtiques

Universitat Autònoma de Barcelona

Comitè editorial _____

Bartomeu Coll

Dep. de Matemàtiques i Informàtica

Universitat de les Illes Balears

Jorge Mateu

Departament de Matemàtiques

Universitat Jaume I

Núria Fagella

Dep. de Matemàtica Aplicada i Anàlisi

Universitat de Barcelona

Marc Noy

Departament de Matemàtica Aplicada II

Universitat Politècnica de Catalunya

Josep Maria Font

Dep. de Probabilitats, Lògica i Estadística

Universitat de Barcelona

Francesc Planas

Departament de Matemàtica Aplicada I

Universitat Politècnica de Catalunya

Armengol Gasull

Departament de Matemàtiques

Universitat Autònoma de Barcelona

Agustí Reventós

Departament de Matemàtiques

Universitat Autònoma de Barcelona

Gábor Lugosi

ICREA i Departament d'Economia

Universitat Pompeu Fabra

Marta Sanz-Solé

Dep. de Probabilitats, Lògica i Estadística

Universitat de Barcelona

Societat Catalana de Matemàtiques

Carrer del Carme, 47 - 08001 Barcelona

tel. 933 248 583 - fax 932 701 180

scm@iec.cat - <http://scm.iec.cat>