

# **Elemente de analiză comparativă**

Adrian Dușa  
*Elemente de analiză comparativă*

Copyright © 2014 Adrian Dușa  
Toate drepturile rezervate, inclusiv dreptul de a reproduce fragmente  
din carte.

Colecția Sociologie

Tritonic grup editorial  
București, România  
Str. Coacăzilor nr. 5  
Email: editura@tritonic.ro  
www.tritonic.ro

Descrierea CIP a Bibliotecii Naționale a României  
DUȘA, ADRIAN

Elemente de analiză comparativă / Adrian Dușa  
București, Tritonic, 2014.  
ISBN 978-606-749-004-6

Coperta colecției: Alexandra Bardan  
Editor: Bogdan Hrib  
Tehnoredactor: Adrian Dușa  
Comanda TA1 / octombrie 2014  
Bun de tipar: decembrie 2014  
Tipărit în România

Orice reproducere, totală sau parțială, a acestei lucrări, fără acordul  
scris al autorului, este interzisă și se pedepsește conform Legii  
dreptului de autor.

**Adrian Duşa**

**Elemente de analiză  
comparativă**





# Cuprins

<b>Prezentare generală</b>	<b>1</b>
<b>1 Calitativ versus cantitativ în cercetarea socială</b>	<b>5</b>
1.1 Cercetarea calitativă, orientată spre studii de caz . . . . .	9
1.2 Cercetarea cantitativă, orientată spre variabile . . . . .	12
1.3 Critica cercetării calitative . . . . .	17
1.3.1 Reprezentativitatea . . . . .	17
1.3.2 Selecția cazurilor . . . . .	19
1.3.3 Replicabilitatea rezultatelor . . . . .	23
1.4 Critica cercetării cantitative . . . . .	24
1.4.1 Cifrele nu pot substitui realitatea . . . . .	24
1.4.2 Baza de date nu înlocuiește munca de teren . . . . .	26
1.4.3 Problema asumpțiilor . . . . .	27
1.4.4 Complexitate și comprehensiune . . . . .	28
1.4.5 Constituirea cazurilor . . . . .	30
1.4.6 Studiul cazurilor uniforme . . . . .	31
1.4.7 Definirea cazurilor negative . . . . .	32
1.4.8 Examinarea cazurilor multiple și conjuncturale . . . . .	33
1.4.9 Tratarea cazurilor nonconforme . . . . .	33
1.4.10 Alte critici . . . . .	34
1.5 Complementaritatea celor două metode . . . . .	38

<b>2</b>	<b><i>Via media</i>: analiza calitativ-comparativă</b>	<b>45</b>
2.1	Comparația și analiza cauzală . . . . .	47
2.2	Condiții necesare, suficiente și contrafactice . . . . .	54
2.3	Canoanele inductive ale lui Mill . . . . .	60
2.4	Inducție, deducție, abducție . . . . .	65
2.5	Similaritate și diferență sistemică . . . . .	70
2.6	Critica cercetării comparative . . . . .	74
2.6.1	Problema cutiei negre . . . . .	75
2.6.2	Problema lui Galton . . . . .	78
2.6.3	Problema “variabile multe, cazuri puține” . . . . .	80
2.6.4	Problema determinismului . . . . .	83
2.6.5	Problema diversității limitate . . . . .	85
2.6.6	Problema contradicțiilor . . . . .	87
<b>3</b>	<b>Scopuri ale comparației</b>	<b>89</b>
3.1	Pentru a formula enunțuri sociologice . . . . .	90
3.2	Pentru a înțelege mai bine propria societate . . . . .	92
3.3	Pentru a infirma ipoteze . . . . .	94
3.4	Pentru validarea rezultatelor de cercetare . . . . .	96
3.5	Pentru a rafina concepte și modele teoretice . . . . .	97
3.6	Pentru a identifica șabloane sistematice . . . . .	100
<b>4</b>	<b>Noțiuni de teoria mulțimilor</b>	<b>101</b>
4.1	Sistemul binar și algebra booleană . . . . .	102
4.2	Tipuri de mulțimi utilizate în analiza comparativă . . . . .	103
4.2.1	Mulțimi distincte bivalente . . . . .	104
4.2.2	Mulțimi distincte multivalente . . . . .	106
4.2.3	Mulțimi vagi . . . . .	109
4.3	Operații cu mulțimi . . . . .	111
4.4	Diagramele Venn . . . . .	114

---

<b>5</b>	<b>Metodologia analizei calitativ-comparative</b>	<b>119</b>
5.1	Analiza cu mulțimi distincte bivalente . . . . .	119
5.1.1	Analiza de necesitate . . . . .	121
5.1.2	Analiza de suficiență . . . . .	125
5.1.3	Relevanța condițiilor cauzale: acoperirea . . . . .	127
5.1.4	Conjunctii și disjunctii cauzale . . . . .	132
5.1.5	Tabela de adevăr . . . . .	137
5.1.6	Minimizarea booleană . . . . .	140
5.1.7	Legile lui DeMorgan . . . . .	147
5.1.8	Problema codificărilor binare . . . . .	149
5.2	Analiza cu mulțimi distincte multivalente . . . . .	150
5.3	Analiza cu mulțimi vagi . . . . .	158
5.3.1	Cercetarea socială și mulțimile vagi . . . . .	159
5.3.2	Operațiuni cu mulțimi vagi . . . . .	162
5.3.3	Calibrarea mulțimilor vagi . . . . .	167
5.3.4	Analiza de necesitate . . . . .	173
5.3.5	Analiza de suficiență . . . . .	176
5.3.6	Includere și acoperire . . . . .	178
5.3.7	Tabela de adevăr . . . . .	181
5.3.8	Așteptări direcționale și asumptii simplificatoare . . . . .	192
5.4	Un exemplu demonstrativ . . . . .	197
	<b>Glosar</b>	<b>211</b>
	<b>Bibliografie</b>	<b>213</b>





# Prezentare generală

And what should they know of England who only England  
know?

Rudyard Kipling

În această carte este prezentată analiza calitativ-comparativă (din engleză QCA - Qualitative Comparative Analysis), așa cum este dezvoltată în prezent în cadrul dezbaterilor metodologice. O analiză aflată încă în plină dezvoltare, suscită discuții aprinse și extrem de interesante. De la crearea ei acum aproape trei decenii (Ragin, 1987), analiza continuă să primească o atenție ridicată din partea comunității academice.

Sociologia este o știință care utilizează foarte mult metoda comparației; aproape că nu există cercetare care să nu ia în considerare și alte cazuri similare sau diferite. Ne comparăm cu vecinii, ne comparăm cu cei la fel cu noi, ne comparăm cu cei mai dezvoltăți, ne comparăm cu cei din vest ș.a.m.d. În lumea socială au existat, există și vor exista întotdeauna diferențe: între indivizi, între grupuri, între comunități, între culturi, între religii și putem continua la nesfârșit pe diferite criterii de stratificare. Așa cum bine o definește Smelser (1976, pp.2-3), analiza comparativă este “studiul unităților sociale disimilare ... (în special ale societăților sau culturilor diferite)”.

Durkheim spunea că “sociologia comparată nu este o ramură particulară a sociologiei, este sociologia însăși” (Durkheim, 2002, p.181). Nu am putea să înțelegem societatea în care trăim până când nu ne comparăm cu alții, cu *ceilalți*. În acest sens, probabil că una dintre cele mai reprezentative propoziții scrise vreodată îi aparține lui Lipset (1994, p.154):

“Cine cunoaște doar o singură țară practic nu cunoaște niciuna”.

Sunt valabile aceste afirmații? Este sociologia într-o suprapunere perfectă față de metoda comparativă? Din alte puncte de vedere, afirmațiile par să constrângă sociologia la un set limitat de acțiuni de cercetare. De fapt, întrebarea la care se reduce totul este dacă poate exista sociologie prin studierea unui singur caz (să spunem o singură națiune).

Categoric, acest lucru este posibil, iar propozițiile emise de Durkheim și Lipset par a fi ușor de respins. Pe de altă parte, chiar și în interiorul unei națiuni sunt regiuni diferite care pot fi comparate pentru a înțelege mai bine structura întregului. Chiar și în interiorul unei regiuni există comunități diferite care pot fi comparate pentru a realiza un portret comprehensiv al regiunii. Chiar și în interiorul unei comunități sunt indivizi diferiți care pot fi comparați între ei pentru a descrie cât mai realist comunitatea. De fapt, binecunoscuta formulă a dispersiei pare să aplice tot o metodă comparativă, întrucât toți indivizii sunt comparați cu media.

La limită, chiar dacă se analizează un *singur* individ (ceea ce este un nonsens în sociologie), putem să-l înțelegem doar dacă-i vom studia mai multe instanțe comportamentale, pe care să le comparăm ulterior.

Imaginea pe care încerc să o produc este desigur o exagerare, întrucât sociologia nu poate fi redusă complet la o analiză comparativă (altminteri s-ar fi numit știința comparației). Cine urmează acest drum se supune riscului de a cădea în situația celui care, dacă are în mână un ciocan, vede toate lucrurile ca pe niște cuie. Chiar dacă această lucrare este puternic orientată comparativ nu voi limita sociologia doar la analiza comparativă, ci mă voi rezuma să afirm că acest tip de analiză formează un subset al analizelor sociologice, un subset în opinia mea foarte important care merită întreaga atenție în cadrul unui demers de cercetare.

De-a lungul cercetărilor la care am participat am avut timp suficient pentru a-mi lămuri aspectele metodologice ale comparației și opinia mea este că, deși această metodă este deja bine fundamentată în metodologia vest-europeană și americană, în cercetarea socială românească este relativ vag percepută și mai mult, că uneori această reprezentare este chiar falsă. De exemplu, prin comparație se înțelege de multe ori în mod eronat acțiunea de *diferențiere cantitativă* și așa merge chiar mai departe afirmând că scopul tacit acceptat al acestei diferențieri este acela de a alcătui simple *ierarhizări*.

S-ar mai putea afirma că cercetarea noastră comparativă se sprijină în mare măsură pe realizarea de clasamente, alcătuite aproape în totalitate din indicatori cantitativi; de aici și afirmații de tipul “suntem printre ultimele locuri în Europa în ceea ce privește PIB-ul pe cap de locuitor”.

Construirea indicatorilor cantitativi nu este o operațiune simplă. Nu procesul în sine este complicat, ci obținerea unor indicatori cu un grad înalt de validitate și fidelitate. Importantă este în special problema validității, întrucât dacă indicatorul nu măsoară ceea ce trebuie să măsoare atunci întreaga cercetare se află sub semnul întrebării. Foarte mulți indicatori sunt extrem de utilizați deși au o validitate scăzută.

Un exemplu foarte cunoscut este Indicele Dezvoltării Umane (utilizat chiar la “compararea” țărilor); în opinia multor specialiști, modelul operațional al acestui concept ar putea fi semnificativ îmbunătățit prin utilizarea unui alt set de indicatori, însă nu se realizează acest lucru pentru a putea face comparații cu anii din urmă.

Deși foarte multe aspecte de cercetare sunt legate în primul rând de definirea conceptuală (de exemplu ce înțelegem prin *dezvoltare umană*), o cantitate imensă de efort se concentrează spre dezvoltarea indicatorilor, ca și cum între concept și indicatorii utilizați ar exista un semn de egalitate. Mahoney și Goertz (2006) avertizează că este posibil ca o astfel de abordare să determine un efect cu totul neașteptat și contrar: conceptul ajunge să fie definit de către indicatorii utilizați pentru a-l măsura.

În plus, există o tendință accentuată în cercetările cross-naționale (de exemplu European Social Survey) de a traduce foarte exact *itemii* chestionarului, fără a ține cont că aceștia sunt importanți doar pentru a măsura concepte. Comparabilitatea trebuie asigurată la nivelul conceptelor, iar itemii pe baza cărora sunt măsurate acestea pot fi mai mult sau mai puțin diferiți.

În mod cert, cercetarea sociologică dispune astăzi de un instrumentar metodologic mult mai bogat decât acum o jumătate de secol; asistăm la o adevărată explozie a unor tipuri de analize din ce în ce mai sofisticate. Ne așteptăm ca explicația sociologică să avanseze pe măsura instrumentelor disponibile și totuși lucrul acesta nu se întâmplă. Marile teorii despre fenomenele sociale nu s-au construit (și aș îndrăzni să spun că nu se vor construi niciodată) utilizând programe de calculator.

Procesele sociale au o dinamică aparte, iar înțelegerea lor presupune o aplicare adecvată a analizei comparative. Metodologia acesteia presupune mai mult decât niște simple diferențieri numerice; presupune niște abilități mai subtile care combină atât deprinderi cantitative cât și ușurința de a lucra cu concepte și metode calitative. În esența ei, analiza comparativă este de natură calitativă (în fapt, analiza poartă numele de QCA, de la *Qualitative Comparative Analysis*). Acțiunea de comparație este o operațiune subiectivă: de exemplu, pentru un european este foarte dificil să diferențieze între chinezi și coreeni, însă în acea parte a lumii diferențele sunt percepute mult mai ușor. Același lucru este valabil și invers, diferențele între român tipic și un norvegian tipic sunt puțin sesizabile pentru un asiatic. Percepția diferenței este așadar o chestiune de proximitate spațială, în consecință subiectivă.

S-au construit instrumente software și pentru analiza comparativă, bazate pe proceduri algoritmico-matematice, însă acestea nu trebuie

confundate cu instrumentele statistice. Cele două tipuri de instrumente se referă la abordări diferite ale cercetării sociale, așa că o comparație directă a eficienței acestora nu-și are rostul. Dacă în cercetările transversale accentul cade pe recoltarea datelor de la cât mai mulți indivizi, cercetarea comparativă își dovedește utilitatea în special la unități macrosociale. A compara doi indivizi este o acțiune care nu poate fi generalizată, întrucât acțiunile individuale nu se supun unor regularități matematice (dimpotrivă, au o componentă haotică absolut impredictibilă). La nivel macrosocial însă (comunități, regiuni, națiuni), indicatorii agregați sunt mult mai stabili iar deplasarea acestora se face într-un ritm suficient de lent pentru a putea fi utilizați în comparație. De aceea, cele mai dese comparații se fac între țări, însă analiza comparativă nu este în nici un fel restricționată doar la acest nivel.

Analiza comparativă are un rol bine definit în cadrul cercetării sociale actuale, iar cartea de față încearcă să prezinte avantajele utilizării acestui tip de analiză.

Primul capitol este dedicat unei discuții asupra scopului analizei comparative, pentru a justifica utilitatea și locul acestui tip de analiză în cadrul cercetării sociale. În al doilea capitol voi prezenta specificul acestei analize, în contextul teoretic inevitabil. Deoarece acest capitol face multiple trimiteri către celelalte două mari “rivale” din cadrul cercetării sociale, următorul capitol face o descriere a acestora, cu sublinierea punctelor tari și slabe ale fiecăreia dar mai ales a complementarității dintre acestea.

Natura intrinsecă a analizei calitativ-comparative face apel la niște asumptii fundamentale asupra filozofiei de cercetare socială, de exemplu asupra naturii cauzale a relațiilor dintre fenomenele sociale. De aceea, după prezentarea algoritmului am simțit nevoia de clarificare a acestor asumptii, cu toate părțile ei bune și cu problemele metodologice asociate. Capitolele din partea a doua a cărții încearcă să stabilească legăturile dintre cele trei mari tipuri de cercetare socială (calitativă, cantitativă și comparativă), precum și să determine limitele posibile ale acestor abordări.

Limitele analizei comparative au fost dezbătute în mod consistent în literatura de specialitate, fapt care a condus în mod necesar la necesitatea dezvoltării acesteia. În fapt, au avut loc mai multe tipuri de dezvoltări, pe care le prezint în ultima secțiune teoretică.

# Capitolul 1

## Calitativ versus cantitativ în cercetarea socială

Există aspecte ale cercetării sociale de care nu se pot face abstracție atunci când tratăm problema comparației. Analizată din perspectiva cercetării calitative ori a celei cantitative, sau contestată din perspectiva asumțiilor ontologice la care știința în general nu a răspuns încă (de exemplu, problema cauzalității), analiza comparativă este strâns legată și se interpune între problemele metodologice specifice de tip calitativ și cantitativ.

De aceea, consider că este important să dedic un capitol acestor tipuri de întrebări metodologice, pe de o parte pentru că analiza comparativă nu poate exista în afara acestora, iar pe de altă parte pentru că încercarea de lămurire a acestora aduce un sprijin enorm încercării de a înțelege analiza comparativă însăși.

Științele sociale și politice sunt împărțite, în ce privește cercetarea, între două lumi: cercetarea calitativă și cercetarea cantitativă (Ragin, 1987; Rihoux, 2003; Tilly, 1997; Berg-Schlosser, 2002; King et al., 2000, printre alții).

Metodele de cercetare cantitative s-au dezvoltat foarte mult (în special după explozia puterii de calcul a computerelor moderne) și putem spune că în prezent domină peisajul metodologic al științelor sociale. Și totuși, oricât de sofisticate ar fi metodele cantitative de ultimă generație, există fenomene sociale foarte interesante din punct de vedere sociologic, care nu pot fi sub nici un chip explicate de acestea.

Să luăm de exemplu revoluția românească din 1989, indubitabil un fenomen social major, unul care suscită și astăzi nenumărate întrebări: Care au fost principalele cauze care au condus la declanșarea

evenimentului? Care este rolul contextului internațional în care se afla România în 1989? Întrebări pot fi formulate și în legătură cu evenimente ulterioare: Care au fost factorii declanșatori ai mineriadelor? Ce anume a provocat violențele inter-etnice din Târgu-Mureș?

Revoluția românească face parte însă dintr-un eveniment istoric mult mai larg, de colaps al întregului bloc comunist din Europa de Est. Cu toții suntem de acord că sociologia fost prinsă pe picior greșit în 1989, nimeni nu a reușit să prevadă acest lucru (nu există lucrări publicate înainte de 1989 care să anunțe iminența producerii acestui eveniment). Mai mult, nu s-a reușit încă o explicație general acceptată iar dezbaterile pe marginea acestui eveniment istoric continuă.

Evenimente care suscită interes pentru cercetarea comparativă există nu doar la nivel de țări, ci și la nivel de arii restrânse, până la comunități distincte. Newman (2002) relatează despre cercetarea întreprinsă de Academia Națională din SUA asupra unor evenimente pe cât de rare, pe atât de dramatice și bulversante pentru întreaga opinie publică: utilizarea armelor de foc în școlile americane, soldate uneori cu zeci de copii morți (celebrul caz al Liceului Columbine).

În mod cert, pentru astfel de evenimente abordarea cantitativă este inutilizabilă. Ar fi nevoie să culegem date despre foarte multe evenimente asemănătoare pentru a putea construi variabile și să testăm statistic relația dintre acestea. Or, astfel de evenimente sunt extrem de rare în istorie, unele chiar unice. Pe baza a doar câteva astfel de cazuri, statistica este cu totul neputincioasă, iar pe un astfel de teren teoretic înfloresc dezbaterile de natură calitativ-comparativă.

Deși în principiu este acceptat că împărțirea calitativ-cantitativ este artificială și că, de fapt, cele două tipuri de cercetare sunt complementare, în practică se poate ușor constata la fiecare dintre noi o tendință mai mult sau mai puțin accentuată spre unul sau altul dintre cei doi poli. Pe măsură ce crește specializarea, se adâncește și neîncrederea în metoda “cealaltă”, fiecare apărându-și tabăra cu o înflăcărare aproape religioasă.

Mahoney și Goertz (2006) numesc cele două tradiții “culturi alternative”, cu toate caracteristicile legate de valori, credințe și norme. Ei au identificat nu mai puțin de 10 aspecte care diferențiază cele două culturi (vom regăsi multe dintre acestea în descrierile de la secțiunile 1.1 și 1.2):

- abordări explicative;
- concepții despre cauzalitate;
- explicații multivariate;

- cauzalitate multiplă, conjuncturală (echifinalitate);
- acoperire și generalizare;
- practici de selectare a cazurilor;
- ponderarea observațiilor;
- abordarea cazurilor foarte importante;
- lipsa de adecvare;
- concepte și măsurare.

Dintr-o perspectivă simplistă, putem privi cercetarea calitativă ca fiind orientată spre studii de caz, cu scopul de a explica aspecte particulare pentru fiecare caz în parte, iar cercetarea cantitativă ca fiind orientată spre variabile, compuse din variația a foarte multe cazuri măsurate în același timp.

Distincția dintre cele două abordări mai este descrisă în literatură și sub forma distincției dintre abordarea *ideografică* și cea *nomotetică*. În căutarea cauzelor unui fenomen, un cercetător care utilizează metoda ideografică va căuta unul sau mai multe cazuri (în general puține la număr) în care se întâmplă fenomenul de interes și studiază acele cazuri în profunzime până când găsește toate circumstanțele care provoacă fenomenul; dimpotrivă, un cercetător care utilizează metoda nomotetică va aduna cât mai multe cazuri cu puțință, apoi va căuta șabloane comune care indică circumstanțele generale în care se produce fenomenul respectiv.

Diferența majoră dintre cercetarea calitativă și cea cantitativă constă în scopul cu care este efecturată. Ambele utilizează evidența empirică pentru a construi reprezentări ale câmpului social studiat, însă modul cum se utilizează această evidență diferă atât ca metodă cât și ca scop. Deși ambele doresc să explice realitatea, modul cum se realizează acest lucru este cu totul diferit. Dacă utilizarea cea mai eficientă a analizei cantitative urmărește avansarea științei prin testarea ipotezelor în sens Popperian, metodele calitative sunt îndreptate mai degrabă spre formarea conceptelor și spre explicarea cazurilor care prezintă o importanță teoretică ieșită din comun.

Disputele artificiale din jurul celor două tabere sunt datorate exact acestei neînțelegeri conceptuale; în realitate, cele două abordări sunt complementare: există cazuri în care nu se poate aplica decât o metodă de cercetare calitativă, după cum atunci când avem suficiente date pentru a testa o ipoteză este recomandată aplicarea metodelor cantitative.

Opinia mea este că diferența dintre abordarea calitativă și cea cantitativă este o reproducere a diferenței dintre științele social-umane și științele exacte. Cel mai înalt obiectiv al științei este căutarea universalilor, adică

a acelor legi valabile în orice punct din spațiu și timp. Mărul a căzut cu aceeași viteză atât în vremea lui Newton cât și pe vremea lui Einstein; atracția gravitațională este așadar o lege universal valabilă. Din acest motiv, un fizician poate să manifeste față de sociologie o superioritate distantă: la urma urmei, sociologia este o știință tânără în comparație cu fizica și (cu foarte puține excepții) încă nu a generat legi universale în spațiu și timp, care să acopere perfect toate diferențele culturale atât pentru civilizațiile cunoscute cât și pentru cele posibile.

Cele trei legi universale prezentate de Zamfir (1990) constituie mai degrabă excepția decât regula în ceea ce privește caracterul universal al legilor sociologice. Unul dintre autorii cei mai cunoscuți care au înaintat o astfel de teorie este Karl Marx (care susținea că dimensiunea economică este universală în societatea umană), însă evenimentele ulterioare au arătat că teoria lui nu este deloc universală deși are calități explicative certe. Argumente în ce privește eșecul teoriei acestuia sunt prezentate în secțiunea 1.5.

Atracția către abordarea cantitativă este în mod cert determinată de asemănarea cu științele exacte, deși opinia mea este că societățile umane nu vor putea fi niciodată incluse într-o ecuație matematică. Oamenii nu sunt nici previzibili nici predictibili, cu atât mai puțin societățile umane în ansamblul lor. Istoria poate lua proporții gigantice datorită influenței unei singure persoane (de exemplu apariția lui Napoleon), iar unele culturi sunt așezate pe un fundament non-logic și în consecință imprezicibil (cine ar fi putut să prevadă atacurile de la 11 septembrie din Statele Unite?).

Dacă sociologia nu are (în general) legi universale, atunci sarcina acesteia nu este alta decât studierea și explicarea fenomenelor locale particulare, atât în spațiu cât și în timp. O teorie din științele sociale va fi doar temporar valabilă, iar fenomenele locale pot fi studiate atât cu ajutorul metodelor cantitative, cât și cu (aș putea să spun mai ales cu) ajutorul celor calitative.

Scopul următoarelor două secțiuni nu este de a prezenta într-un mod exhaustiv cele două metode (există deja o literatură suficient de bogată pentru ambele), ci doar de a schița particularitățile fiecăreia. Prin urmare, nu voi dedica foarte mult spațiu acestor două prezentări, rezervând mai mult loc pentru prezentarea analizei comparative care propune o soluție de mediere.



## 1.1 Cercetarea calitativă, orientată spre studii de caz

În prezentarea abordărilor calitative, atenția se orientează (întemeiat) spre o denumire aproape sinonimă: studiile de caz. Între cele două denumiri există o relație de suprapunere aproape perfectă, întrucât analiza calitativă se referă la studierea în detaliu a unui număr restrâns de cazuri. De aici și denumirea de “studiu de caz”: un caz este studiat din multe puncte de vedere, pentru a spori șansele de a găsi o explicație pentru fenomenul studiat.

Ambiguitatea termenilor mai survine și din faptul că în științele sociale sintagma “studiu de caz” este neclară. În cele mai multe situații, reprezentarea conceptuală se referă la faptul că cercetătorul merge pe teren și observă în detaliu o unitate de analiză: o localitate, o comunitate, o familie, chiar și o persoană poate constitui un studiu de caz. Unitățile de analiză nu se limitează însă doar la cele mici: regiunile, țările, continente întregi pot constitui *cazuri*, în funcție de particularitățile studiului. Statele Unite reprezintă o singură entitate de analiză, deși este constituită din zeci de state și ocupă o jumătate de continent ca întindere, la fel Uniunea Europeană poate constitui un caz.

Prin contrast, o cercetare cantitativă pe un eșantion de peste 1000 de cazuri pare că utilizează cuvântul “caz” într-un mod diferit, însă în realitate nu se poate studia întregul Univers într-o singură cercetare. Sociologia este o știință regională, în sensul în care trebuie să răspundă unor probleme sociale specifice unui anumit loc într-un anumit moment. Problemele românești ale democrației în tranziție nu sunt asemănătoare cu problemele rasiale din Statele Unite. Când se studiază (chiar și printr-o cercetare cantitativă) anumite probleme *specifice*, nu este acest lucru de fapt un studiu de caz?

Dacă exemplul anterior nu este suficient de convingător, atunci să presupunem că un cercetător dorește să studieze eficiența comunicării într-o organizație și alege pentru analiză o firmă de o dimensiune relativ mare (câteva sute de angajați), cu compartimente bine definite și separate. Există două tipuri posibile de analiză: utilizând metode calitative (observație, interviu, focus-group) și metode cantitative (de exemplu, alegem la întâmplare să spunem 150 de angajați din toate nivelurile și compartimentele cărora le administrăm un chestionar).

Metoda cantitativă pare, nu-i așa, mai “științifică” întrucât lucrează cu un număr mare de cazuri (de la 120 de chestionare în sus începe să funcționeze așa numita lege a numerelor mari) și pot fi aplicate diverse analize și teste statistice. Însă este trecut cu vederea faptul cel mai important: în fapt, ambele metode studiază *o singură firmă*.

Privită din acest punct de vedere, între o firmă și o țară nu este decât o deosebire de scală. În ambele cazuri pot fi extrase eșantioane aleatoare suficient de mari pentru analiză statistică, însă tot în ambele cazuri este efectuat un singur studiu de caz pe o anumită problemă dată.

Dificultatea realizării distincției dintre limbajul cantitativ și cel calitativ provine și din ambiguitatea termenului de “caz”, pentru care este destul de dificilă acceptarea unei definiții comune între mai multe ramuri ale științelor. Cercetarea comparativă a stabilit o definiție convențională: “cazurile sunt definite de niște limite în jurul unor locuri și a unor perioade temporale” (Ragin și Becker, 1992, p.5).

Cu alte cuvinte, România în anul 2014 constituie un caz, Europa de Est în anul 1989 constituie un caz, la fel de bine cum *Sinuciderea* (Durkheim, 1993) constituie un studiu de caz. Studiile calitative și cele cantitative diferă într-adevăr după modul cum tratează cazurile. În cercetările bazate pe mărimi mari ale eșantioanelor, cazurile sunt empiric observabile, ele *există* în câmpul social investigat și constituie niște linii într-un set de date. În abordarea calitativă cazurile sunt o combinație de empiric și constructe teoretice; unul și același “caz” din realitate poate fi subiect de studiu pentru mai multe ipoteze teoretice.

Platt (1992) este de părere că, deși nu percepem întotdeauna, în practică facem apel la o considerabilă cantitate de convenționalism care ne influențează și ne orientează cercetarea. De exemplu, este *familia* o entitate direct observabilă? Cei mai mulți cercetători ar confirma fără ezitare, însă familia nu este nimic altceva decât o convenție socială. Grupurile familiale de acum câteva zeci de ani nu mai seamănă cu familia din ziua de astăzi, așadar familia este un concept în continuă schimbare și redefinire, cu alte cuvinte o convenție.

În cercetarea calitativă, cazurile “empirice” sunt analizate într-o manieră cât mai exhaustiv cu putință, din perspectiva mai multor aspecte iar cercetătorul se familiarizează cu fiecare caz în parte și depune un efort considerabil pentru a atinge această cunoaștere. De aceea, nu este posibilă decât studierea unui număr foarte limitat de cazuri. Înarmat cu un aparat conceptual și cu o ipoteză de lucru, cercetătorul trece cazurile prin mai multe filtre până când ajunge să le cunoască suficient de bine în așa fel încât, utilizând o combinație de încrucișări și metode inductive, ajunge să afirme dacă datele acumulate susțin sau nu ipoteza de plecare.

Rolul pe care îl joacă metodele calitative în științele sociale este nu doar important, cât mai ales expeditiv cu prea multă ușurință. Studiat în mod exhaustiv, chiar și un singur studiu de caz poate fi util pentru dezvoltarea științei. Dacă suntem de acord că sociologia este o știință regională și locală, atunci se poate concentra pe fenomene la orice nivel de agregare

socială: poate studia la fel de bine omenirea în ansamblul ei, un continent, o națiune, o zonă din cadrul unei țări, o localitate, o comunitate restrânsă și, de ce nu, chiar o familie. Atunci când unitățile de analiză sunt din ce în ce mai puține (în cazul unei familii,  $N = 1$ ), metodele cantitative nu pot fi aplicate.

În studiul unei singure entități, sociologia seamănă foarte mult cu istoria. Cazul respectiv ajunge să fie cunoscut în amănunțime, deși acest lucru nu servește la nimic mai departe (istoria unui caz nu poate fi prelungită și aplicată altor cazuri, deoarece acestea au propria lor istorie). Ragin (1987) arată că studiile calitative au două funcții principale: istoric interpretative și cauzal analitice.

Aplicate unei singure entități, ambele funcții vor avea o valabilitate restrânsă. Este ca și cum am cartografia foarte amănunțit un continent pentru a obține o hartă utilă în navigare, însă harta respectivă nu poate fi utilizată pentru alt continent, oricât de amănunțită ar fi (această metaforă îi aparține lui Barrington Moore, conform Skocpol și Somers, 1980, p.195).

Ne-am obișnuit să considerăm Terra ca fiind singura planetă locuită din univers, însă nu este nici o garanție că nu mai sunt și alte planete cu semne de viață ori chiar locuite de civilizații. Să presupunem că se descoperă o astfel de planetă: nu există niciun motiv să credem că viața de pe acea planetă seamănă cu a noastră. Așadar, nu există niciun motiv să credem că modul de organizare socială al acelei planete este la fel cu al nostru, iar sociologia pământeană (oricât de bine ar explica societatea *omenească*) nu poate fi universal valabilă pentru toate societățile planetelor locuite din univers.

Mai dificil de înțeles este faptul că nici măcar fizica nu este universală. Einstein a arătat acest lucru foarte limpede prin teoria relativității, iar exemplul ipotetic al descoperirii unei alte planete se poate aplica la fel de bine descoperirii unui alt univers. Nu există nici un fel de garanție că legile fizice ale acestui univers sunt valabile pentru toate universurile posibile, însă imaginea mărului căzând este atât de puternic înrădăcinată în mintea noastră încât ne vine foarte greu să acceptăm o reprezentare alternativă.

Concluzia este că fiecare unitate de studiu (univers, specie, cultură, regiune, națiune etc.) poate fi studiată ca atare și că studiul unui caz izolat este la fel de important ca și studiul unui eșantion de 1067 de cazuri. Obiectivele de cercetare asumate sunt cele care fac diferența dintre aplicarea unei metode sau alta. Dacă un caz particular are o poziție înaltă în agenda comună, atunci merită a fi studiat calitativ.

Chiar dacă la prima vedere par să nu dispună de o metodologie foarte structurată (cel puțin nu la fel de structurată ca în abordarea cantitativă), studiile de caz își găsesc adevărata lor valoare atunci când sunt așezate în perspective comparative. Cercetarea comparativă are o natură tot calitativă însă dispune de o metodologie apropiată de “așteptările” cantitativiștilor, în sensul că are un mecanism de producere a rezultatelor bazat pe un algoritm replicabil: aplicând aceeași metodă, cercetători diferiți vor ajunge la același rezultat.

## 1.2 Cercetarea cantitativă, orientată spre variabile

Deși în practică se implică reciproc, există o deosebire de nuanță între “cercetarea cantitativă” și “cercetarea orientată spre variabile”, prima incluzând-o pe cealaltă. Atât cercetătorii din științele sociale cât și oamenii în general gândesc cantitativ într-o măsură mai mare sau mai mică. Utilizarea atributelor “mult” sau “puțin” implică în mod necesar o așa-numită *cuantificare*, adică o operațiune de atribuire a unei măsuri numerice unor proprietăți care caracterizează un fenomen de interes.

Utilizarea variabilelor presupune un pas mai departe, prin culegerea sistematică a unor observații cantitative și aranjarea acestora într-o matrice de date în care pe linii sunt aranjate cazurile studiate iar pe coloane proprietățile fenomenului (sau variabilele). O astfel de aranjare a datelor deschide calea către multiple tipuri de analize, a căror începuturi le-am putea identifica în scrierile lui Mill (1970).

Printre cunoscutele lui canoane inductive (descrise pe larg la secțiunea 2.3) se află acela al “variațiilor concomitente”, probabil prima descriere sistematică a ceea ce în statistica modernă se numește coeficient de corelație.

Cercetarea bazată pe variabile nu presupune în mod necesar o analiză care să utilizeze doar măsuri numerice; în baza de date pot fi introduse și atribute calitative ale fenomenelor studiate, iar variabilele se disting așadar în *calitative* și *cantitative*. Aceasta este prima formă de împărțire a variabilelor pe categorii, fiecare dintre acestea incluzând încă două tipuri distincte de variabile care împreună formează cele patru niveluri de măsurare:

- nominal;
- ordinal;
- interval;
- raport.

Aceste niveluri de măsurare stau la baza oricărei analize statistice a datelor, iar modul cum sunt măsurate variabilele din baza de date influențează în mod dramatic tipurile posibile de analize ulterioare. Există foarte multe situații în care cercetătorul știe ce dorește să studieze, însă construiește instrumentul de măsurare într-un mod mecanic, fără a se întreba dacă variabilele respective pot fi utilizate ulterior în analiză. Instrumentul poate fi perfect construit, respectând toate regulile din metodologia cercetării sociale, însă în etapa de analiză a datelor există riscul ca variabilele respective să nu poată fi utilizate întrucât sunt măsurate altfel decât așteaptă analiza.

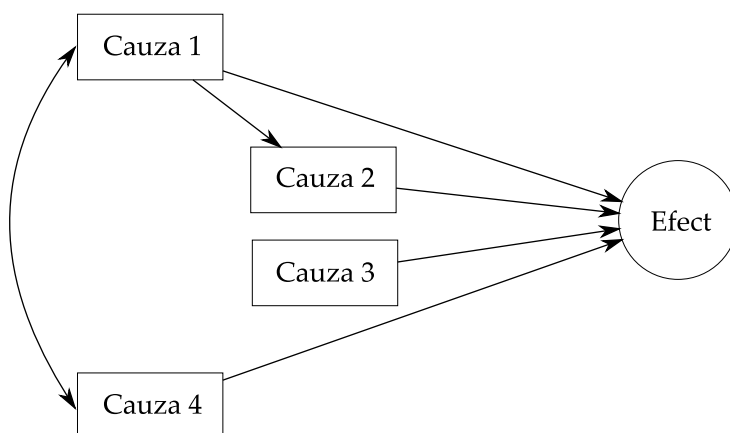
Strategia trebuie să fie cu totul inversă: cercetătorul pornește de la obiectivele cercetării, identifică și operaționalizează conceptele de interes, apoi decide ce fel de analize statistice sunt potrivite pentru testarea ipotezelor de studiu. Numai după identificarea analizelor se construiește apoi instrumentul de măsurare și se decid metodele de culegere a datelor, în așa fel încât la prelucrarea acestora totul să fie pregătit conform planului de cercetare.

În primele două niveluri de măsurare (nominal și ordinal) se regăsesc variabilele “calitative”: la prima vedere pare puțin ciudat ca analiza cantitativă să apeleze la variabile denumite la fel cu analiza “opusă”, însă totul se clarifică atunci când identificăm scopul în care sunt utilizate atributele calitative ale fenomenelor. Dacă în analiza calitativă cercetătorul utilizează aceste proprietăți pentru a construi reprezentări conceptuale despre cazurile studiate, analiza cantitativă utilizează variabile calitative pentru a calcula probabilitatea de apariție a uneia sau alteia dintre proprietăți.

De exemplu (luând în considerare individul ca unitate de analiză) una dintre proprietățile cu care putem caracteriza individul este etnia. Utilizând un eșantion *suficient de mare* de indivizi putem construi un tabel de frecvențe relative ale etniilor înregistrate, iar aceste frecvențe pot fi apoi utilizate ca aproximări pentru probabilitățile de apariție a unei etnii sau alta. Cu cât eșantionul este mai mare, cu atât aceste aproximări sunt mai aproape de realitate (în condițiile în care eșantionul a fost extras din întreaga populație respectând criteriile aleatoare).

Cercetătorul poate să utilizeze mai apoi această informație în legătură cu alte variabile, atât calitative cât și cantitative. Se pot construi tabele de contingență între două variabile calitative, testând asocierea dintre acestea cu ajutorul testului  $\chi^2$ , ori se pot verifica diferențele cantitative dintre grupurile determinate de variabila calitativă, fie cu testul  $t$  pentru două eșantioane independente fie cu ajutorul testului  $F$  din analiza de varianță, pentru mai mult de două grupuri.

Figura 1.1: Model de diagramă PATH



Analiza cantitativă este extrem de atrăgătoare pentru cercetarea socială întrucât sunt posibile nu doar cuantificări ale unor proprietăți, ci și cuantificări ale relațiilor dintre acestea. Orice analiză începe cu schițarea unei scheme conceptuale, păstrând în centrul atenției fenomenul de interes (denumit și *variabilă dependentă*). Măsurat metric, orice fenomen dispune de o anumită cantitate de variație cauzată de unul sau mai mulți factori (denumiți și *variabile independente*), fie direct, fie printr-un lanț causal, fie prin conexiuni reciproce. Este important de subliniat însă că o variabilă independentă nu reprezintă în mod obligatoriu o Cauză, ci poate reprezenta un alt efect al unei cauze latente, neobservabile în mod direct.

Figura 1.1 prezintă un model schematic de relații între anumite cauze (încadrate în dreptunghiuri) și un efect (înconjurat cu un cerc). Toate condițiile cauzale au câte o influență directă asupra efectului, însă cauza 1 mai are și un efect indirect prin intermediul cauzei 2 pe care o influențează, iar între cauza 1 și cauza 4 există o relație de influență reciprocă reflectată prin săgețile de la ambele capete ale arcului de cerc. Singura cauză care influențează direct și nemijlocit efectul este cauza 3, care nu prezintă nici un fel de legături cu celelalte cauze din model.

O astfel de diagramă este utilizată cu precădere în analiza de cale (PATH) cu ajutorul căreia se identifica măsurile numerice ale *influenței* uneia sau alteia dintre variabilele independente asupra variabilei dependente (deși o variabilă independentă nu se confundă întotdeauna cu o cauză). Pentru a ajunge la măsurile respective, trebuie făcută însă o legătură între nodurile diagramei (cauzele plus efectul) și variabilele din baza de date. Aceste

noduri nu sunt întotdeauna direct măsurabile, ci pot reprezenta concepte abstracte care necesită fiecare câte o operaționalizare și un instrument de măsurare dedicat.

Măsurarea conceptelor se poate realiza prin intermediul operaționalizării acestora, etapă în urma căreia sunt identificate dimensiunile și în final indicatorii direct observabili care sunt utilizați apoi ca itemi în instrumentul de măsurare (de obicei un chestionar). După etapa aplicării instrumentului, se pot agrega răspunsurile individuale pentru fiecare grup de itemi (fie printr-o scală, fie printr-un indice combinat) în așa fel încât cercetătorul obține scoruri metrice pentru fiecare nod în parte, la nivel de individ.

Rezultatul agregării mai multor itemi într-un singur scor este așadar o măsură numerică a cărei plajă de valori depinde, pe de o parte, de numărul itemilor pe baza cărora a fost calculat scorul și, pe de altă parte, de nivelul de măsurare ale fiecărui item în parte. Această măsură numerică agregată este un fel de aproximare a conceptului abstract pe care dorim să îl măsurăm, presupunând că acel concept dispune de un interval de variație cu un capăt minim și unul maxim. Dacă scorul rezultat este mic, înseamnă că individul prezintă un nivel scăzut pe conceptul respectiv, iar dacă scorul este mare (să spunem aproape de maxim) înseamnă că individul prezintă o implicare ridicată pe conceptul respectiv.

Un posibil exemplu este preponderența individului pentru implicarea în societatea civilă (fenomen de interes pentru țările post-comuniste, întrucât un efect pervers al comunismului este tocmai scăderea participării la activități *comune*). Spiritul civic este un concept abstract care poate fi identificat numai cu ajutorul unor indicatori direct observabili, iar instrumentul clasic prezintă itemi de participare în regim de voluntariat în diverse organizații non-guvernamentale apolitice, ori culturale, ori de interes comunitar, sau chiar participare la acțiuni de protest, greve etc. Se pot număra asociațiile ori activitățile comunitare la care un individ participă în mod voluntar, scorul cumulativ fiind cu atât mai mare cu cât implicarea individului în societatea civilă este mai mare, iar un scor de zero reprezintă o lipsă totală de implicare.

Scorul rezultat poate fi utilizat apoi fie ca variabilă dependentă (pentru a găsi factorii responsabili de implicarea în societatea civilă) fie ca variabilă independentă, dacă un alt efect depinde de implicarea individului în societatea civilă. Există o varietate de modele posibile, în funcție de imaginația și abilitatea cercetătorului de a le formula. Unele modele sunt realiste sau plauzibile în timp ce altele se pot dovedi fanteziste, însă avantajul major al analizei cantitative este acela că poate să testeze dacă un model este adecvat realității sau nu.

Presupunând că datele din eșantion sunt reprezentative pentru populație, că indivizii au fost extrași în mod aleator și că nu există erori de măsurare, atunci datele constituie un foarte bun substitut empiric al realității, iar dacă modelul cercetătorului rezistă testelor de semnificație pe baza datelor empirice înseamnă că este un model adecvat realității. Voi utiliza pentru exemplificare regresia multiplă, deși nu este nici pe departe cea mai bună formă de analiză a datelor însă este una dintre cele mai utilizate și mai cunoscute. Una dintre măsurile cele mai utilizate pentru stabilirea adecvării modelului este coeficientul de determinare multiplă care arată procentul de variație a variabilei dependente explicat de model. Coeficientul dispune de o variație între 0 și 1, unde 0 înseamnă 0% (modelul nu explică absolut nimic din variația dependentei) iar 1 înseamnă 100% (modelul explică perfect variația dependentei).

În practică, regresia este o analiză extrem de robustă, chiar dacă nu sunt îndeplinite întotdeauna toate condițiile de aplicare. Un eșantion suficient de mare poate să suplinească o parte dintre aceste neajunsuri, iar rezultatele oferite pot fi utilizate ulterior în scopuri de predicție a variabilei dependente în funcție de anumite valori ale independentelor.

Relațiile dintre variabilele metrice pot fi analizate cu ajutorul unui evantai extrem de larg de analize statistice. De la banalele teste de semnificație până la studii de segmentare a populației cu ajutorul unor tehnici mai complicate cum este analiza cluster, agregarea variabilelor cu ajutorul analizei factoriale plus o multitudine de alte tehnici aplicabile în situații speciale, analiza cantitativă dispune de o faimă deosebită în paradigma curentă de cercetare datorită exactității matematice și a replicabilității rezultatelor.

Măsurile numerice nu se rezumă doar la nivel de individ, baze de date cu indicatori existând și la nivel de localități, regiuni ori chiar națiuni. În lume există între 192 și 194 de țări (cazurile speciale fiind Vaticanul și Taiwan - nerecunoscută oficial din cauza presiunilor politice ale Chinei), însă acest număr este suficient de mare pentru a compila o bază de date cu indicatori numerici utilă pentru analize statistice, în special pentru scopuri comparative.

Sigur, există limite importante (prezentate la secțiunea 1.4) însă anumite tipuri de ipoteze, general valabile pentru întreaga omenire pot fi utilizate pentru prelucrare statistică. Ipotezele care utilizează termeni relativi cum este cel de *cultură* nu sunt potrivite pentru o astfel de bază de date, însă Przeworski și Teune (1970) au propus (spre marea satisfacție a susținătorilor analizei cantitative, de exemplu Goldthorpe, 1997a) înlocuirea numelor proprii ale țărilor cu nume ale variabilelor.



O ipoteză care privește durata medie a vieții în perspectivă comparativă poate să facă apel atât la concepte latente cum este cel de calitatea vieții, dar și la variabile direct măsurabile cum este consumul de grăsimi polisaturate. Japonia prezintă una dintre cele mai înalte durate ale speranței de viață la naștere, iar tot Japonia are unul dintre cele mai mici consumuri medii de grăsimi polisaturate, așa încât această variabilă este un candidat serios pentru explicarea fenomenului. Teoria actuală, ca și abilitatea cercetătorului de a formula modele statistice, pot să facă din analiza cantitativă o trusă de instrumente indispensabile care trebuie să facă parte din arsenalul oricărui sociolog.

### 1.3 Critica cercetării calitative

Multe dintre criticile aduse analizei calitative sunt deopotrivă valabile și pentru analiza comparativă. De exemplu, problema lui Galton sau cea asociată a cutiei negre sunt valabile pentru ambele. Mai mult, unele sunt valabile pentru toate tipurile de cercetări, inclusiv cea cantitativă. Am ales să prezint cele mai multe astfel de probleme în capitolul 2 întrucât în această lucrare atenția este îndreptată spre analiza comparativă și pe răspunsurile pe care le poate oferi aceasta.

Problemele expuse aici, împreună cu problemele prezentate la analiza comparativă formează un amestec echilibrat care contrabalansează multitudinea de probleme atinse în secțiunea 1.4, pentru analiza cantitativă.

#### 1.3.1 Reprezentativitatea

Problema reprezentativității este poate cea mai importantă critică adusă cercetării calitative în general. Întrucât cercetările calitative se desfășoară pe un număr limitat de cazuri, este posibil ca rezultatele să nu fie valabile pentru celelalte elemente din care este compusă populația. În abordarea cantitativă, un eșantion de 10 persoane are o eroare mult mai mare decât un eșantion de 1000 de persoane, întrucât eroarea standard se diminuează pe măsură ce mărimea eșantionului crește.

Analiza statistică are nevoie de multe cazuri pentru a putea utiliza distribuția normală: o regulă de bază stabilește faptul că cel mai mic eșantion mare are 30 de cazuri, întrucât de la această mărime distribuția  $t$  oferă aceleași rezultate ca distribuția  $z$  (distribuția normală).

Un eșantion de mărime mică nu poate să conțină la fel de multă diversitate ca în întreaga populație, astfel încât eșantionul utilizat este

în mod cert nereprezentativ. Într-un mod similar, atunci când nu avem la îndemână decât foarte puține cazuri (să spunem 10), nu sunt recomandate calcularea frecvențelor relative. De exemplu, un procent de 40% ar putea fi obținut fie prin raportarea a 400 de cazuri dintr-un total de 1000, fie prin raportarea a 4 cazuri dintr-un total de 10. Procentul este același, însă informația transmisă poate fi înșelătoare.

Privită din acest punct de vedere, problema pare insolubilă, iar cercetarea calitativă pare trimisă într-o poziție de o importanță secundară. Totuși, există motive puternice pentru care nu se analizează calitativ decât un număr foarte restrâns de cazuri, unul foarte important fiind acela al restricțiilor legate de timp și efort. Studiarea în profunzime a unor situații sociale necesită mult mai mult timp decât simpla aplicare a unui chestionar, iar efortul cercetătorului este pe măsură.

Prezentând mai multe studii calitative în care cazurile studiate sunt extrem de diverse, Platt (1992, p.32) amintește o carte celebră a lui Lazarsfeld, Berelson și Gaudet scrisă în 1944 *People's Choice*. Lazarsfeld și echipa sa au dorit să studieze comportamentul de vot al oamenilor și modul cum se schimbă intenția de vot pe parcursul campaniei electorale. Într-o perioadă în care metodele de eșantionare stratificate nu erau încă inventate, Lazarsfeld nu a optat pentru un eșantion național, ci a studiat un eșantion de mai multe sute de persoane extras din populația unui departament din statul Ohio.

Întrebarea care se impune aici este: câte cazuri a studiat Lazarsfeld? Câteva sute de persoane? Ori a studiat doar departamentul respectiv, care face parte din populația mai largă a celor câteva sute de departamente din toată America? În fapt, Lazarsfeld subliniază că nici nu a fost preocupat de problema reprezentativității, întrucât ceea ce a dorit el să studieze era simpla opțiune de vot.

Este acest studiu reprezentativ pentru toate Statele Unite? Privit statistic, nu este reprezentativ nici măcar pentru statul Ohio, cu atât mai puțin la nivelul SUA. Și totuși, Lazarsfeld a studiat un (singur) caz de comportament electoral, cu toate că a analizat un panel de câteva sute de persoane. Rezultatele acestui studiu nu au urmărit extrapolarea și generalizarea la nivelul tuturor statelor, ci să *explice* comportamentul de vot și mai ales schimbarea acestuia generată de stimulii din campania electorală. Privită din acest punct de vedere, studiul publicat de Lazarsfeld și echipa lui manifestă un înalt caracter de validitate și a constituit un punct de referință în cercetarea socială a timpului.

Faptul că a fost ales departamentul respectiv sau altul, ori statul Ohio în locul altuia, este puțin important. Ceea ce contează este faptul că studiul a produs un model teoretic al schimbării comportamentului de vot iar

acest lucru a împins cunoașterea științifică mai departe. Scopul comun al analizelor calitative și comparative este explicarea fenomenelor sociale, iar abordările calitative pot contribui foarte bine la acest scop.

Este utilă introducerea, în această secțiune, a distincției dintre cercetarea exploratorie și cercetarea confirmatorie. Cercetarea calitativă are un rol deosebit și este utilizată cu bune rezultate în explorarea câmpului social, pentru generarea de teorii. Un exemplu reprezentativ de studiu al unei singure unități sociale este monografia lui Stahl (1939) care, prin studiul satului Nerej, a reușit să producă o teorie a satelor devălmașe confirmată mai apoi prin trei volume de date statistice.

### 1.3.2 Selecția cazurilor

O altă critică adusă cercetării calitative se referă la modalitatea de selecție a cazurilor studiate. Dogma de bază a analizei cantitative este selecția aleatoare a cazurilor; studiile care nu respectă această regulă fundamentală sunt descalificate rapid, iar cercetările calitative sunt primele pe această listă. Privită prin paradigma cantitativ-statistică, o selecție nealeatoare a cazurilor crește eroarea de non-eșantionare, ceea ce mărește artificial distanța de la statisticile din eșantion până la parametrii din populație.

Apoi, cercetătorul ajunge la niște concluzii studiind cazurile A, B și C, însă nu cumva se ajungea la alte concluzii dacă erau studiate cazurile X, Y și Z? Nu cumva rezultatele studiului ar fi fost diferite dacă ar fi fost studiate *alte* cazuri? De ce acele cazuri și nu altele? Cum procedează analistul din cercetarea calitativă ca să aleagă cazurile pe care le studiază?

Ca orice lucru care nu este pe deplin înțeles, o selecție *convenabilă* provoacă anxietate și respingere din partea cantitativiştilor. Impresia generală despre selecția cazurilor în cercetarea calitativă este că depinde foarte mult de subiectivitatea cercetătorului.

King et al. (2000) arată că acest lucru produce distorsiuni de selecție, întrucât cercetătorul dorește să demonstreze ceea ce știe dinainte că trebuie să rezulte din cercetare, în același mod în care un partizan politic alege doar exemplele care susțin o anumită concluzie. Deși autorii recunosc faptul că selecția aleatoare nu este posibilă pentru cercetările cu eșantion mic, ei nu oferă nici o soluție practică pentru această situație decât că ar trebui să alegem toate cazurile posibile și să nu lăsăm pe dinafară cazuri într-un mod *convenabil*.

Sunt întru totul de acord cu această afirmație, însă discuția deviază partizan în paragraful următor când promovează selecția care "...permite

șansa obținerii unei variații pentru variabila dependentă” (King, Keohane și Verba, 2000, p.121). Cu alte cuvinte, dacă nu există variație nu putem face selecție, în consecință nu putem face cercetare socială. Problemele sunt așezate într-un cadru conceptual preponderent cantitativ, ca și cum fenomenele sociale nu pot fi analizate decât în acest fel. Autorii uită, sau nu vor să ia în considerare faptul că uneori fenomenele sunt atât de rare încât chiar nu există variație.

Apoi, depinde foarte mult ce înțelegem prin cazuri. Să luăm ca exemplu situația minerilor disponibilizați și să presupunem că aceștia sunt în număr de 10.000 (populația pe care dorim să o studiem). În principiu avem suficiente cazuri care să permită extragerea unui eșantion mare de mineri, însă adevăratele cazuri nu sunt indivizii, ci zonele miniere în care au fost făcute disponibilizările.

Cu alte cuvinte, nu avem 10.000 de cazuri ci doar câteva reprezentând zonele miniere unde au fost făcute disponibilizările. Reconversia profesională a acestora nu depinde doar de ei, ci și de oportunitățile locale de angajare precum și de programele naționale organizate în acest scop de stat. Un cercetător interesat de aceste probleme nu are la dispoziție decât un număr limitat de cazuri posibile pentru analiză, din care selectează pentru analiză câteva dintre acestea. Diversitatea cazurilor (așadar variația acestora) este mult mai limitată decât lasă King et al. (2000) să se înțeleagă.

Una dintre cele mai noi tipuri de analiză care pot ajuta în aceste cazuri este analiza multinivel, în care sunt modelate atât diferențele individuale cât și cele datorate grupurilor din care indivizii fac parte. Diferența față de regresia clasică (în care pot fi introduse variabile independente calitative sub formă de variabile dummy) este că eroarea este mult mai bine explicată, aceasta fiind împărțită și în funcție de variabilele independente la nivelurile mai mari de agregare. Totuși, analiza multinivel nu poate fi utilizată ca un substitut cantitativ pentru analiza comparativă, întrucât variabila dependentă este măsurată întotdeauna la cel mai mic nivel de agregare (cel mai adesea individul), ori analiza comparativă este interesată numai de nivelurile înalte de agregare. De aceea, în această lucrare prin cazuri voi înțelege numai unități agregate.

Pe ce criterii se alege totuși aceste cazuri? Aceasta este o întrebare la care nu există un răspuns simplu. Probabil că cel mai comod la îndemână este aplicarea eșantionării expert, în care cercetătorul cunoaște atât de bine populația studiată încât poate să puncteze imediat cazurile cele mai interesante din punctul de vedere al obiectivelor studiului, însă King et al. (2000) acuză bias-ul de selecție ca și cum toate studiile calitative sunt distorsionate din oficiu. Exemplele pe care ei le oferă în legătură cu acest

aspect utilizează într-un mod ciudat tot unități de analiză care permit o analiză cantitativă (indivizii și acțiunile acestora). Concluzia impusă de această lucrare este că orice studiu calitativ este evident inferior întrucât *ar putea* fi efectuat cu instrumente cantitative dacă cercetătorii ar dispune de competența necesară.

Un exemplu de selecție după variabila dependentă este lucrarea lui Skocpol (1979) care a selectat trei revoluții cu același rezultat, din zone și perioade istorice diferite în speranța că va găsi acele variabile care explică rezultatul. Este foarte dificil de imaginat cum ar fi putut Skocpol să considere revoluția ca o variabilă dependentă și să găsească o *variație* a revoluțiilor ca și cum fiecare caz (țară) ar prezenta mai mult sau mai puțin o situație revoluționară. Dimpotrivă, ceea ce se cunoaște sunt cazurile în care revoluția nu se petrece (cele mai multe cazuri, în perioadele de pace) și cazurile (foarte puține la număr) în care revoluțiile se întâmplă pur și simplu. Nu există jumătăți de revoluție, nici zecimi de revoluție, după cum nu există un concept hilar de revoluție *medie*.

Ebbinghaus (2005) este mai aproape de interpretarea modurilor posibile de considerare a cazurilor, în special a celor comparative. În cercetările cros-naționale, de exemplu, atât cercetarea calitativă cât și cea cantitativă suferă de exact aceeași problemă: națiunile sunt rezultatul unor procese istorice de formare extrem de diferite, iar a introduce în aceeași ecuație (bază de date) țări ca Statele Unite și Islanda doar pentru că sunt autonome politic este la fel de greșit ca selecția *convenabilă* din calitativ. Unele țări sunt incomparabile, ori analiza cantitativă nici nu ia în considerare o asemenea variantă.

Apoi, toate cercetările cantitative care utilizează în mod comparativ datele OECD ar trebui aruncate la gunoi întrucât, chiar în definiția dată de King et al. (2000) se bazează pe un eșantion nealeator și biasat, reprezentând doar o duzină de țări din cele 194 care constituie Națiunile Unite.

Mahoney și Goertz (2004) oferă trei reguli de bază pentru selecția cazurilor:

**Principiul posibilului.** Este evident că o analiză științifică trebuie să ia în considerare atât cazuri pozitive (în care fenomenul se întâmplă) cât și cazuri negative (în care fenomenul de interes nu se întâmplă). În general, numărul cazurilor negative este infinit; întrucât (de obicei) fenomenele studiate comparativ sunt rare, restul situațiilor reprezintă o paletă foarte largă de selecție. De aceea, alături de cazurile pozitive, se vor selecta acele cazuri negative unde fenomenul de interes este cât de cât *posibil*. Cazurile în care fenomenul nu se poate întâmpla sub nici o formă, deși sunt

cazuri negative prin definiție, sunt irelevante întrucât nu ajută cu nimic comparația.

**Regula de includere.** Cum decide un cercetător dacă fenomenul este posibil sau nu? Decizia nu este una simplă, însă poate fi ușurată dacă analizăm cazurile posibile prin valorile variabilelor independente. Un caz (negativ) este declarat relevant (posibil) dacă valoarea sa pe cel puțin o variabilă independentă este corelată pozitiv cu valoarea fenomenului de interes, chiar dacă celelalte variabile independente acționează în sensul inhibării fenomenului. Un astfel de caz poate fi inclus în analiză, dacă nu a fost cumva exclus prin aplicarea următoarei reguli.

**Regula de excludere.** Un caz este considerat irelevant dacă, indiferent de valorile variabilelor independente, fenomenul de interes nu se produce. Această regulă este mai importantă ca regula de includere, întrucât cazurile irelevante trebuie excluse.

Aceste reguli simple oferă soluții de ieșire din această problemă, iar situația este și mai clară în cadrul studiilor comparative. Deși am putea (la limită) să ne imaginăm o cercetare comparativă la nivel de individ, un astfel de demers nu ar avea un sens practic. Orice indivizi am selecta, prin orice metodă (aleatoare, probabilistă ori alta), concluziile la care am ajunge nu s-ar bucura de un mare interes teoretic. Este așadar rezonabil să afirm că un interes de cercetare comparativ nu poate fi atras decât pe baza unor cercetări la nivel agregat: grupuri, comunități, regiuni, națiuni.

La nivel agregat, cazurile studiate nu sunt pur și simplu echivalente pentru a putea fi extrase la întâmplare, ci au o semnificație aparte în funcție de obiectivele cercetării și de fenomenul social studiat (Ragin, 1997). Ca și în cazul analizei calitative, unul dintre scopurile analizei comparative este acela de a explica un fenomen social de foarte mare interes, prin studierea și compararea în profunzime a unui număr (de obicei restrâns) de cazuri care prezintă anumite particularități. Se aleg deci acele cazuri (foarte puține la număr) care satisfac obiectivele de cercetare; este important de subliniat că cerința de maximizare a variației este imposibilă într-o asemenea cercetare întrucât numărul cazurilor disponibile este foarte mic.

Voi încheia această secțiune prin sublinierea importanței selecției cazurilor; în ciuda aparenței de subiectivitate, analiza calitativă acordă mult mai multă atenție acestei probleme decât analiza cantitativă; în afara cazurilor în care poate fi utilizată o selecție aleatoare, analiza cantitativă este aceea care se află de fapt într-un permanent pericol de confuzie asupra cazurilor utilizate și a selecției acestora.

### 1.3.3 Replicabilitatea rezultatelor

Problema replicabilității rezultatelor este legată de cea a selecției cazurilor, ambele fiind de fapt niște efecte ale subiectivității cercetătorului. Dacă în privința selecției cazurilor pentru analiză există metode de a preveni și a controla această subiectivitate, în privința interpretării datelor de cercetare lucrurile nu sunt nici pe departe la fel de simple. De aceea, problema replicabilității rezultatelor este (printre altele) una într-adevăr serioasă.

Fiecare cercetător poate avea propria viziune despre fenomenul studiat și este bine să fie așa. Dacă toți oamenii ar gândi la fel nu am mai putea face știință, nu am mai putea avansa noi teorii. Ideile noi sunt binevenite, însă în cercetarea calitativă există riscul unei percepții individuale asupra realității: cu alte cuvinte câți cercetători, atâtea reprezentări.

Replicabilitatea unei analize presupune obținerea acelorași rezultate atunci când cercetători diferiți studiază aceleași cazuri. Este o condiție de la sine înțeleasă pentru o cercetare serioasă, iar cercetarea cantitativă are în privința aceasta un ascendent serios: calculatorul va arăta aceeași valoare medie, indiferent de cercetătorul care comandă tastatura. Lucrând pe aceeași bază de date, un cercetător din România va obține aceleași rezultate ca un cercetător din altă țară.

Problema nu este specifică doar științelor sociale. De exemplu, interpretările istorice ale unor evenimente pot să fie diferite dacă istorici diferiți le studiază. Există numeroase cazuri de reinterpretare istorică, unele bazate pe dovezi nou descoperite. Științele sociale se află într-o poziție avantajată față de istorie întrucât datele sociale pot fi colectate în prezent iar analiza nu depinde neapărat de arhivele de date istorice.

O posibilă soluție pentru problema replicabilității este aplicarea constantă a metodologiei de cercetare. La modul ideal, doi cercetători cu experiență similară vor ajunge la concluzii asemănătoare dacă vor aplica aceleași metode de cercetare. Este puțin probabil ca rezultatele culese să fie total diferite, doar interpretările acestora pot să difere în funcție de imaginația și de experiența de cercetare acumulată.

O soluție care se apropie de standardele “cantitative” este oferită de abordarea comparativă a cercetărilor calitative, prin aplicarea algoritmului de minimizare booleană QCA, discutat la secțiunea 5.1. Utilizând aceleași date de cercetare, rezultatul va fi întotdeauna la fel, indiferent de experiența cercetătorului care aplică algoritmul.

Este important de subliniat faptul că, deși este simplu de aplicat, QCA nu trebuie efectuată în mod mecanic. Ca și în cazul analizei cantitative,

metode specifice au nevoie de date de intrare specifice, iar construirea datelor de intrare este operațiunea cea mai sensibilă din tot procesul comparativ. Apoi, analiza are unele opțiuni care corespund unor echivalente ale asumțiilor din analiza cantitativă, iar tratarea acestora cu superficialitate este la fel de dăunătoare și poate modifica substanțial rezultatele, amenințând pretenția de replicabilitate a analizei comparative.

QCA nu rezolvă problema decât la nivel de analiză a datelor. Culegerea acestora (strâns legată de selecția cazurilor discutată în secțiunea anterioară) rămâne o problemă serioasă: cazuri diferite pot genera concluzii diferite, chiar dacă acestea pot fi replicate. De aceea, QCA cere nu doar o atenție sporită pentru alegerea cazurilor relevante, ci și argumentarea alegerii acestora.

## 1.4 Critica cercetării cantitative

Dacă ții o mână în congelator iar cealaltă pe o plită încinsă, la medie este bine însă pentru niciuna dintre mâini situația nu este fericită.

Aceasta este probabil cea mai cunoscută formulare care exprimă neîncrederea în capacitatea cifrelor de a exprima o stare de fapt. Cercetătorii de orientare preponderent calitativă (spun “preponderent” pentru că de fapt toată lumea apelează mai mult sau mai puțin la cuantificări, fie și sub forma cuvintelor “mult”, “puțin”, “puternic” etc.) au și ei, la rândul lor, anumite rezerve mai mult decât îndreptățite vis-a-vis de utilitatea analitică a metodelor cantitative.

Unele dintre aceste rezerve exprimă o neîncredere mai mult emoțională, în timp ce altele (prezentate la sfârșitul acestei secțiuni) punctează niște dificultăți metodologice majore în fața cărora analiza cantitativă nu poate să răspundă.

### 1.4.1 Cifrele nu pot substitui realitatea

O primă rezervă față de analiza cantitativă/statistică este faptul că cifrele nu spun totul despre o populație. România anului 2007 a înregistrat creșteri economice importante, însă nivelul de trai al populației nu a crescut semnificativ (unii ar putea să spună că a scăzut). Privită doar din perspectiva PIB-ului, am putea să spunem că România merge într-o direcție bună. Mulți analiști care lucrează cu baze de date



rulează teste de semnificație statistică fără să-și pună întrebări majore asupra datelor pe care le utilizează. Chiar dacă datele sunt corecte, chiar dacă ele reprezintă realitatea la nivel agregat, cifrele nu spun totul despre o populație.

Utilizând date statistice, putem spune ceva despre o populație doar la nivel agregat însă viața concretă a oamenilor se realizează la nivel micro (familie, comunitate) și nu la nivel macro (societate). Pentru a înțelege concret modul de manifestare a unui fenomen social într-o societate, este necesară studierea societății respective prin mijloace mai întâi calitative (observație, interviuri, studii de caz) și apoi prin metode cantitative.

Agregarea într-o singură cifră a unei bogății fantastice de informație presupune renunțarea în mod brutal la dinamica și unicitatea cazurilor particulare. Faptul că avem un PIB înalt nu trebuie să ne facă indiferenți la condițiile dure de viață ale multor oameni simpli a căror singură vină este că s-au născut în grupuri defavorizate. Există segmente întregi de populație complet diferite de imaginea pictată de statisticile agregate.

Apoi, există rezerve serioase cu privire la modul de producere a statisticilor (chiar și a celor oficiale). În funcție de metodologia de calcul aplicată, cifrele oficiale pot să arate foarte diferit. Să luăm de exemplu rata șomajului: cifrele Institutului Național pentru Statistică arată diferit de cifrele Ministerului Muncii, ambele fiind instituții cât se poate de oficiale. În care dintre aceste două cifre să avem mai multă încredere?

Un alt exemplu este numărul de persoane care au emigrat din România în ultimii ani. Deși estimările se situează undeva în jur de 2 milioane de persoane, este foarte dificil de produs cifre oficiale exacte cu privire la acest fenomen. Doar emigrările definitive pot să fie raportate cu certitudine, însă este discutabil câți dintre actualii migranți vor mai dori să se întoarcă definitiv în țară, așa încât cifrele oficiale asupra fenomenului migrației sunt incerte, oricâtă încredere am avea în statisticile oficiale.

Una dintre cele mai utilizate informații despre veniturile populației este salariul mediu pe economie. Această statistică este cât se poate de inutilă (aș spune chiar că este utilizată exclusiv în scopuri de manipulare și propagandă politică) întrucât, deși cele mai multe salarii se află mult sub valoarea medie, o proporție mică de salarii exorbitante reușesc să ridice media până la o valoare ce nu mai reprezintă marea parte a populației.

Este binecunoscut faptul că distribuția veniturilor este alungită la dreapta, iar media este o măsură sensibilă la acțiunea salariilor extreme

din acea coadă a distribuției. O valoare mai potrivită în această situație este venitul median, însă diferența dintre cel mediu și cel median este probabil atât de mare încât decidenții politici nu sunt pregătiți să o admită în mod public. Iată cum “realitatea” poate fi manipulată prin atât de credibilele cifre statistice oficiale.

În fine, există o legătură strânsă între nivelul de dezvoltare economică și calitatea cifrelor oficiale. Tindem să avem mai multă încredere în cifrele prezentate de instituții oficiale din Germania (o țară cu tradiție în disciplină și corectitudine) decât de instituții oficiale din Africa. Același lucru este valabil între statisticile produse în cadrul democrațiilor stabile (corelate de altfel cu nivelul de dezvoltare economică) și statisticile produse în cadrul unor state totalitariste. Privind retrospectiv, câtă încredere am mai putea avea în cifrele oficiale prezentate de regimul comunist din România dinaintea de 1989?

Totuși, în cercetările internaționale comparative, cifrele sunt introduse în aceeași bază de date ca și cum ar fi fost produse de către aceeași instituție statistică. Investim niveluri diferite de încredere în cifre separate, însă acordăm aceeași încredere unei baze de date care a adunat cifrele respective într-un singur loc. Acestea sunt doar câteva exemple care arată că cifrele trebuie privite și utilizate cu rezervă.

#### 1.4.2 Baza de date nu înlocuiește munca de teren

După 1990, diverse studii au fost realizate în România de către specialiști ai Băncii Mondiale, ori ai Fondului Monetar Internațional, ori ai diverselor organisme internaționale care au investit în această direcție. Rezultatul unor studii de acest gen sunt, de obicei, publicate în vreo revistă de specialitate din străinătate, iar autorul articolului este felicitat.

Dar cine este acest autor? Este cumva un bun și fin cunoscător al societății românești? În afară de utilizarea datelor puse la dispoziție de vreo instituție statistică românească (chiar pe baza unei metodologii foarte moderne), cunoaște respectivul autor și altceva despre societatea românească, despre cutumele și situațiile particulare din diversele zone ale țării?

Desigur, există cazuri de rapoarte întocmite de specialiști locali ai Băncii Mondiale, însă există probabil cazuri de autori care scriu despre România fără să o fi vizitat vreodată. În această situație, cu toată încrederea ce o acordăm mecanismului matematic din spatele analizei statistice, câtă încredere putem să acordăm articolului ca întreg? Rareori se întâmplă ca o teorie despre un fenomen social să se potrivească perfect în toate

societățile, ori autorii care testează ipoteze doar din birou, pe baza unor date cantitative culese de alții, comit exact această eroare de a considera că pot studia o societate urmărindu-i doar variabilele.

Sociologia este și va rămâne întotdeauna o știință națională (uneori chiar regională); cu alte cuvinte, teoriile din “vest” nu se potrivesc întotdeauna cu cele din “est”; este nevoie de o sociologie a “estului” așa cum există o sociologie a muncii, a familiei etc. Sigur, există aspecte comune în toate societățile (în special în epoca globalizării, a fenomenelor transnaționale, în general a fenomenelor de difuziune culturală tratate la secțiunea 2.6.2), însă fenomenele sociale se manifestă diferit în diverse zone ale globului, iar teoriile despre acestea trebuie să fie enunțate într-o deplină cunoaștere ce nu poate fi obținută doar prin cifre, ci și calitativ, prin muncă de teren.

Metodologia cantitativă a explodat în ultimele decenii, în paralel cu dezvoltarea computerelor. Paradoxul observat însă este că, deși avem mai multe instrumente teoretice ca niciodată până acum, deși avem o încredere imensă în capacitatea statisticii de a valida modele, deși capacitatea de procesare a calculatoarelor a crescut imens, teoriile marilor gânditori nu se nasc din analiza cantitativă.

Nici o teorie nu “sare” din calculator odată cu prelucrarea datelor. Dimpotrivă, cercetătorii care utilizează eșantioane mari de date se lovesc de piedici din cele mai diverse în legătură cu baza de date utilizată, iar marile teorii pornesc dintr-o atitudine *reflexivă* asupra problemei studiate. Datele nu construiesc o teorie, ci pot doar să o confirme.

### 1.4.3 Problema asumpțiilor

Modelele utilizate pentru testarea statistică a ipotezelor sunt de cele mai multe ori relativ simple; cerința primordială a unui bun model este să conțină numărul cel mai mic de variabile independente care să explice cantitatea cea mai mare de variație a dependentei. Problema de care se lovesc cercetătorii care utilizează (ori abuzează de) metodele statistice este că acestea se bazează pe anumite asumpții, uneori greu de îndeplinit. Probabil cea mai răspândită practică în legătură cu asumpțiile este ignorarea acestora; la urma urmei, calculatorul poate să genereze rezultate matematic corecte, chiar dacă nu sunt respectate toate asumpțiile. Cu alte cuvinte, sub umbrela de încredere oferită de o formulă matematică se ascund erori calitative nevăzute întrucât aproape nimeni nu le mai verifică.

Un exemplu extrem de comun este calcularea mediei pe o scală de răspuns Likert cu 5 trepte. Deși teoria specifică foarte clar că pentru calcularea mediei este nevoie de un nivel de măsurare metric (asumpție

puternică), tratarea unui nivel de măsurare ordinal cu cinci trepte ca și cum ar fi metric a devenit un fapt comun pentru majoritatea instituțiilor de cercetare socială din România. Aproape fără excepție, motivația perpetuării acestei practici este: “pentru că așa se face”.

Printre cele mai importante asumții de la cea mai cunoscută analiză “avansată” (regresia multiplă) este asumția de normalitate a distribuiri erorilor. În analizele multivariate și mai complicate, asumția cere ca toate variabilele să aibă o distribuție multivariat normală. Dintre toți utilizatorii de metode cantitative însă, foarte puțini sunt aceia care înțeleg aceste asumții, iar numărul celor care țin cont de acestea este chiar și mai mic. Respectarea asumțiilor este însă importantă, în caz contrar rezultatele obținute nu pot fi utilizate cu încredere.

Soluții metodologice există pentru aproape orice problemă. Deoarece se întâmplă destul de des ca anumite asumții să nu fie atinse de către datele analizate, cercetătorii au la dispoziție un întreg evantai de alte metode statistice care așteaptă să fie utilizate. Pentru orice problemă de analiză a datelor, oricât de complexă ar fi, există câte un antidot analitic.

Norul de puncte nu este liniar? Nici o problemă, se pot aplica logaritmi în diverse baze pentru asigurarea liniarității, cu singura restricție că interpretarea trebuie făcută nu pe variația variabilei ci pe variația logaritmului acesteia. Chiar dacă artificul aplicat este absolut banal, slaba înțelegere a acestuia pune în dificultate marea majoritate a celor care analizează datele cantitative.

Unele soluții statistice la problema asumțiilor sunt relativ simple, altele mai complicate. Atunci când soluțiile devin din ce în ce mai complexe, se renunță la aplicarea acestora iar asumțiile sunt pur și simplu trecute cu vederea. La urma urmei *regresia este suficient de robustă ca să reziste la nerespectarea unor asumții*; nu există însă nici un argument pentru a justifica o astfel de afirmație. Problema asumțiilor merge așadar în tandem cu problema complexității.

#### 1.4.4 Complexitate și comprehensiune

Există situații în care puternica regresie liniară nu poate fi utilizată. Dacă variabila dependentă este categorială, utilizarea unei alte analize devine obligatorie. Regresia logistică ori cea multinomială poate să răspundă acestei provocări, însă subtilitatea mecanismului scapă multor cercetători. Iar acestea nu sunt printre cele mai complicate. Există analize de corespondență, există scalare multidimensională; există unele analize ori tehnici care nici nu au o traducere serioasă în limba română (ex. repeated measures ANOVA sau spatial autocorrelation), iar pentru

amatorii de senzații tari există statistică bayesiană ori analiză pe clase latente cu programe software dedicate (cum sunt LatentGold ori LISREL).

Complexitatea acestor analize a explodat într-atât de mult încât chiar și pentru persoane cu pregătire temeinică în matematică utilizarea lor este dificilă. În sociologia contemporană un bun cercetător nu este doar un sociolog sau doar un matematician ci o combinație fericită din ambele. Or, puțini matematicieni înțeleg subtilitatea conceptelor sociologice, după cum și mai puțini sociologi sunt capabili să mânuiască instrumente matematice indispensabile în analize complicate. Cei care reușesc într-adevăr să le utilizeze în mod corect sunt atât de puțini încât de multe ori nu au cu cine comunica.

Chiar și pentru un cercetător cu pregătire cantitativă este dificilă urmărirea unei demonstrații complexe. Cercetătorul cu o pregătire preponderent calitativă, oricât de mult ar încerca să se convingă de faptul că există o rațiune matematică în spatele analizei, aceasta este atât de abstractă și ezoterică încât exclamă îngrozit: “voi cantitativiști *torturați* datele până când acestea *urlă* ceea ce vreți să aflați”.

În opinia mea, chiar dacă o asemenea afirmație este în mod evident forțată, nu este extrem departe de adevăr. La orice analiză există excepții care deviază în afara mersului normal dictat de asumții; chiar și analizele care rezolvă excepțiile au excepțiile lor, iar analizele care le rezolvă și pe acelea devin din ce în ce mai complicate. Cu cât nișa excepției devine mai mică, cu atât analiza care o rezolvă explodează în complexitate. Cantitatea de efort necesară înțelegerii modului de funcționare al unei astfel de analize complexe este cu mult mai mare decât beneficiul potențial de a fi rezolvată o excepție de nișă.

În fapt, cu cât cercetătorul se avântă în înțelegerea mecanismelor matematice, cu atât se îndepărtează mai mult de la munca de analiză socială. Cercetătorul începe să semene mai mult cu un matematician ori un statistician decât cu un sociolog, iar publicul capabil să înțeleagă un discurs presărat cu jargoane abstracte este redus ca dimensiune. În cele din urmă, statistica este învinsă chiar de propria ei capacitate de a rezolva probleme.

Dacă în stadiul actual al dezvoltării științelor sociale statistica are o astfel de problemă, să presupunem că undeva în viitor cercetătorii vor reuși să o depășească și că metodele cantitative vor fi la îndemâna celor mai mulți analiști (printr-o investiție mult mai mare în abilitățile matematice). Să presupunem mai apoi, tot prin absurd, că aceștia vor putea modela matematic orice fel de fenomen social la fel de simplu cum cercetătorii de astăzi utilizează testele de semnificație.

La ce ar folosi tot acest efort, în condițiile în care societatea nu rămâne niciodată la fel? La ce ar folosi un model statistic extrem de precis dacă el oricum nu este valabil decât pentru o foarte mică durată de timp, raportată la vechimea omenirii? Apoi, statistica nu va putea să prevadă decât fenomenele raționale, niciodată pe cele haotice. De exemplu, cine ar fi putut vreodată să prevadă că teroriștii vor dărâma turnurile gemene din New York cu niște avioane?

În majoritatea celorlalte științe, statistica este un instrument nu doar necesar, ci aproape obligatoriu. Medicina modernă și mai ales producerea medicamentelor se face pe fundamente statistice (dacă pentru un eșantion suficient de mare de pacienți nu apar efecte secundare, medicamentul poate fi eliberat). Biologia și științele pământului pot să dispună de un număr virtual nelimitat de eșantioane, mai nou până și exacta fizică se bazează pe statistică în studiul corpurilor cerești. Toate aceste științe pot face uz nelimitat de statistică, însă nu și științele sociale unde unitățile de analiză se schimbă și se transformă într-un ritm mai mare decât pot cercetătorii să le modeleze.

Dintr-un anumit punct de vedere, acestea pot să apară ca niște speculații emoționale, însă există rezerve metodologice de necontestat, argumente logice foarte pertinente care pun în dificultate analiza comparativă. De exemplu, ca răspuns la unele atacuri metodologice (King et al., 2000; Goldthorpe, 1997a, ș.a.m.d.) îndreptate atât spre analiza calitativă cât și spre cea comparativă, Ragin (1997, pp.27-42) a găsit cinci tipuri de probleme în fața cărora analiza cantitativă este pusă într-o vizibilă dificultate, iar următoarele secțiuni sunt dedicate prezentării acestora.

#### 1.4.5 Constituirea cazurilor

Una dintre slăbiciunile analizei cantitative este faptul că tratează toate cazurile la fel, acestea fiind nici mai mult nici mai puțin decât niște linii care se adaugă în baza de date. Cel puțin, așa a fost până la apariția analizei multinivel și a modelelor ierarhice, însă este discutabil cât de mult a reușit să rezolve acest tip de analiză, în special în cadrul comparațiilor internaționale. În analiza cantitativă clasică, baza de date este cu atât mai bună cu cât are mai multe linii (cazuri).

Prin contrast, analiza calitativă tratează toate cazurile prin specificul și particularitățile fiecăruia. Includerea ori excluderea cazurilor din analiză nu se face aleator (ca selecția elementelor din statistică) ci pe baza unor decizii bine argumentate din perspectiva unor teorii, sau pentru că sunt atât de interesante încât pot să ajute la conturarea unei teorii.

Cazurile sunt alese tocmai pentru că prezintă un interes teoretic înalt. De obicei însă, cazurile care prezintă un asemenea interes sunt extrem de puține. Sfatul pe care îl oferă analiza cantitativă în asemenea situații este includerea unor noi cazuri în analiză prin modificarea cuprinderii conceptuale; Cu alte cuvinte, în loc să studiem anumite “mere”, putem studia un concept mai larg de “fructe” întrucât, nu-i așa, fructele sunt mai multe.

Ragin nu este de acord cu această practică, din simplul motiv pentru că schimbarea acoperirii conceptuale modifică automat și scopul cercetării în sine. De exemplu, cineva dorește să analizeze Revoluția Franceză și să afle care au fost circumstanțele sociale care au făcut posibil apariția unui asemenea eveniment. Momentul acelei revoluții este unic în timp și este puțin probabil că se va mai repeta vreodată; este tocmai această unicitate care deopotrivă atrage și respinge interesul științific. Tocmai pentru că este unic, nu poate fi analizat prin metode statistice. Schimbarea acoperirii conceptuale (să studiem revoluții în general) aduce într-adevăr mai multe cazuri în analiză însă nu aduce nici un plus de cunoaștere cu privire la problema teoretică inițială.

Ragin arată foarte bine că în analiza calitativă cazurile nu sunt niște simple bile pe care le eșantionăm dintr-o urnă, ci pentru unele interese teoretice unitățile pot fi, iar pentru altele nu pot fi considerate cazuri. Analiza calitativă își construiește cazurile pe măsură ce se desfășoară, iar această abordare este de-a dreptul înspăimântătoare pentru un statistician. Nu doar Ragin este de această părere; de exemplu, Przeworski subliniază faptul că “nu putem face analiză comparativă dacă nu ne îngrijim de selecție, ori altfel spus până când nu întrebăm de fiecare dată cum sunt produse observațiile” (conform Ebbinghaus, 2005, p.134)

#### 1.4.6 Studiul cazurilor uniforme

Întreaga filozofie a analizei cantitative se bazează pe conceptul de variație, iar în statistică totul pornește de la abaterea standard. Atunci când cazurile studiate nu prezintă variație, statistica este pusă într-o situație extrem de dificilă întrucât nu poate opera cu constante. Totuși, există situații în care apariția unui fenomen să fie însoțită *întotdeauna* de prezența unor cauze. Într-o bază de date, corelația dintre cele două variabile ar fi egală cu 1 iar metodele statistice se opresc. Ori dimpotrivă, atunci când există o asemenea potrivire între cauză și efect, ar fi o greșeală să ignorăm această relație pentru simplul motiv că nu poate fi de folos într-o analiză statistică.

Mai mult, obiectivul principal al statisticii este acela de a explica *variația* variabilei dependente (fenomenul de interes), ceea ce presupune că există cazuri în care fenomenul se produce și cazuri în care nu se produce, ba mai mult că se produce în diverse grade (ca are o plajă valori pe un interval de variație larg). Prin contrast, analiza calitativă selectează cazurile în funcție de *producerea* fenomenului, ceea ce înseamnă că se elimină variația întrucât fenomenul este constant prezent în toate cazurile studiate. Fie pentru că numărul de cazuri în care fenomenul se produce este foarte limitat, fie pentru că analiștii construiesc cazurile în funcție de obiectivul cercetării, statistica nu este capabilă să răspundă în fața unei asemenea situații: nu există variație, nu avem ce explica.

Această abordare transformă problema selecției cazurilor (reproșată analizei calitative) într-o virtute, întrucât reușește acolo unde analiza cantitativă nu funcționează.

#### 1.4.7 Definirea cazurilor negative

Ragin își continuă prezentarea seriei de situații problematice prin definirea cazurilor negative. În analiza cantitativă, acestea sunt echivalentul unor “outliers” ori a unor cazuri “izolate” care nu se potrivesc cu marea majoritate a eșantionului.

Soluția cea mai la îndemână și probabil cea mai utilizată este eliminarea cazului respectiv din analiză ori pur și simplu ignorarea acestuia. La urma urmei, după toate criteriile probabilistice este suficient dacă frecvența relativă a cazurilor care concordă cu ipoteza să fie ridicată, nu contează că o proporție minusculă dintre cazuri contravin ipotezei. Este exact situația pe care analiza calitativă nu o ocolește ci dimpotrivă, încearcă să reconstruiască atât explicația conceptuală cât și constituirea cazurilor analizate. Pentru a putea spune care dintre cazuri sunt negative, analiza calitativă procedează mai întâi la definirea și constituirea cazurilor pozitive.

Pentru analizele de tip calitativ, un caz negativ nu este o simplă aberație care contrazice marea majoritate a celorlalte cazuri studiate, ci are o semnificație aparte care poate să răstoarne întregi teorii. În secțiunea 1.5 voi discuta două astfel de cazuri care contrazic teorii economice contemporane și aproape au obligat comunitatea științifică la rediscutarea paradigmatelor.

Fiecare caz negativ este deci important, iar analiza calitativă le tratează cu cel puțin atâta importanță cât acordă cazurilor pozitive. Acest tratament lipsește în analiza cantitativă, unde nu există diferențierea între cazuri negative sau pozitive ci o poziționare mai slabă sau mai



puternică pe o scală a variației. Dacă la nivel micro un individ care iese din tipar nu înseamnă o răsturnare de situație, în analiza calitativă fiecare element studiat are o semnificație aparte.

#### 1.4.8 Examinarea cazurilor multiple și conjuncturale

Aceasta este o situație pe care metodele cantitative nu o pot detecta. În mod normal, un fenomen poate să apară dintr-o multitudine de condiții cauzale, a căror efect nu este întotdeauna pur și simplu aditiv ca într-o ecuație de regresie. Cuvântul cheie aici este “combinație” cauzală, deoarece este perfect posibil ca într-o anumită combinație fenomenul să fie produs prin prezența cauzei A iar în altă combinație cauzală fenomenul poate să apară dacă aceeași cauză A este absentă.

Raționamentul este total străin analizei statistice, unde variabila independentă are sau nu o influență semnificativă asupra dependentei, cu o anumită valoare (mai mică sau mai mare) a coeficientului de regresie asociat. Forța acestui coeficient (impactul variabilei independente asupra variației dependentei) este ușor de observat prin analizarea valorii lui standardizate, însă este știut (dar și ignorat) faptul că orice coeficient își schimbă valoarea în diverse combinații cauzale; cu alte cuvinte, este perfect posibil ca într-un anumit model de regresie un coeficient să fie puternic semnificativ în timp ce în alt model același coeficient să fie complet nesemnificativ.

Desigur, majoritatea analizelor statistice moderne prevăd așa numitele efecte de interacțiune (exact corespondentele combinațiilor cauzale), însă introducerea acestor efecte în ecuație solicită și mai multe cazuri în setul de date iar acest lucru este, de obicei, imposibil. În schimb, analiza calitativ-comparativă poate să urmărească toate combinațiile care produc fenomenul studiat și mai mult decât atât poate să spună care dintre combinații este mai importantă decât altele (prin simpla frecvență relativă a diverselor combinații). Cu cât o combinație cauzală apare mai des, cu atât probabilitatea ca acea combinație să producă efectul este mai mare astfel că putem asocia niveluri diferite de importanță diverselor combinații cauzale.

#### 1.4.9 Tratarea cazurilor nonconforme

Există situații în care unele instanțe prezintă fenomene care contrazic o mare cantitate de evidență empirică. În cercetarea calitativă acestea sunt denumite cazuri nonconforme, iar în cercetarea cantitativă acestea sunt echivalentul “erorii”. De obicei, tot ce nu poate explica modelul

(cantitatea de variație rămasă neexplicată:  $1 - R^2$ ) intră automat în cantitatea difuză numită eroare. Pe cercetătorul cantitativ nu-l interesează eroarea, cât mai ales cantitatea de variație explicată de model.

Dimpotrivă, în analiza calitativă toate cazurile care nu merg pe linia modelului inițial sunt analizate în detaliu pentru a fi explicate ulterior. Această încercare este iarăși străină metodelor cantitative, întrucât este echivalentă încercării de a explica *întreaga* variație a dependentei (ori cel mai bun model nu este cel care explică totul, ci acela care explică cel mai mult cu cele mai puține variabile).

Judecând tratamentul aplicat cazurilor nonconforme, abordarea cantitativă pare incompletă: pe de o parte nu are o soluție pentru această problemă, iar pe de altă parte, nici măcar nu o consideră o problemă în sine. Cazurile sunt analizate probabilistic, în contradicție cu încercarea într-un fel deterministă de a găsi o explicație pentru cazurile care nu se potrivesc cu teoria.

Într-o lucrare de referință, Popper (2000) prezintă diverse strategii care formează metoda științifică de cercetare. Unul dintre exemplele oferite de acesta este celebra lui cioară albă, a cărei apariție răstoarnă toate clișeele teoretice conform cărora toate ciorile sunt negre. Este suficient un singur specimen de cioară albă ca să anuleze o întreagă teorie bazată pe milioane de ciori negre.

Abordarea cantitativă nu reușește să răspundă acestui paradox popperian, întrucât pentru baza de date un singur caz neconform nu pune niciun fel de problemă statistică (câteodată cazurile aberante sunt pur și simplu eliminate), un fapt care este problematic din punct de vedere științific.

De aici și necesitatea complementarității dintre cercetarea cantitativă și cea calitativă (discutată în secțiunea 1.5), acestea completându-se reciproc.

#### 1.4.10 Alte critici

Ebbinghaus (2005) are și el unele păreri critice la adresa analizei cantitative, care merită a fi luate în considerație. Într-o bază de date, toate cazurile au exact aceeași importanță: ele nu sunt nimic altceva decât niște linii într-o matrice, pe coloane fiind variabilele la care fiecare țară-caz afișează câte o valoare. Statistica inferențială cere ca toate cazurile să aparțină aceluiași grup (în caz contrar nu putem calcula de exemplu o medie), ori această asumție este extrem de suspectă: țările nu sunt toate la fel (nu aparțin toate aceluiași grup). Unele sunt extrem

de mici (cum este cazul Islandei) iar altele sunt foarte mari prin comparație (cazul SUA). Decizia de a oferi aceeași importanță tuturor liniilor (cazurilor) din baza de date este cât se poate de greșită.

O posibilă soluție ar fi ponderarea cazurilor cu mărimea populației din fiecare țară, însă eu aș adăuga faptul că acest lucru este posibil numai dacă vom considera țările ca straturi separate, ceea ce duce la o altă problemă întrucât condiția primordială în construirea straturilor este că acestea trebuie să fie omogene în interior și eterogene între ele. Această condiție nu poate fi îndeplinită, întrucât națiunile nu sunt perfect omogene în interior și nici nu sunt toate diferite între ele, dat fiind faptul că există societăți care aparțin de aceeași familie culturală.

O strategie alternativă, arată Ebbinghaus, este reducerea eterogenității țărilor (sub aspectul mărimii) prin selecția unor cazuri asemănătoare, ceea ce este întocmai strategia aplicată în analiza calitativă: cazurile nu sunt selectate la întâmplare, ci după niște criterii foarte bine definite.

O altă problemă identificată de Ebbinghaus este legată de disponibilitatea datelor. De exemplu, multe cercetări comparative de economie politică utilizează (doar) datele OECD, ceea ce este un exemplu cât se poate de clar de eșantionare biasată întrucât cele 24 de țări OECD care dispun de date nu reprezintă cele aproape 200 de țări membre ale Națiunilor Unite.

O analiză interesantă a diferenței dintre calitativ și cantitativ este făcută de către Coppedge (1999), care separă conceptele în *dense* (complexe, multidimensionale, specifice analizei calitative) și *subțiri* (reducționiste, simpliste, specifice indicatorilor din analiza cantitativă). De fapt ambele tipuri de cercetare utilizează concepte dense, numai că analiza cantitativă operaționalizează conceptele până când ajunge la indicatori direct măsurabili. Asupra acestor indicatori își îndreaptă Coppedge atenția, arătând că modul de construire a variabilelor are o influență directă asupra măsurării conceptelor dense și că de obicei măsurarea cantitativă este extrem de discutabilă.

Nici în analiza calitativă nu este simplu de utilizat un concept dens. De exemplu, democrația are o mulțime de definiții iar analiza trebuie să utilizeze una sau alta dintre aceste definiții. Pe măsură ce se înaintează cu analiza, conceptul devine din ce în ce mai îngust ori din ce în ce mai elaborat, însă nu ajunge niciodată la nivel de variabilă în sens cantitativ. Din acest punct de vedere, un citat al lui Coppedge (1999, p.471) este sugestiv:

Problema indicatorilor cantitativi nu este că ei măsoară cantități, ci faptul că aceștia sunt *calitativ diferiți* de conceptele categoriale pe care le măsoară. Adevărata

problemă a variabilelor continue este că măsoară numai versiuni subțiri și reduționiste ale unor concepte mai dense de care sunt interesați cercetătorii noncantitativiști<sup>1</sup>.

În opinia mea, Coppedge are dreptate când spune că măsurarea conceptelor este un lucru dificil. Cine a încercat vreodată să construiască o scală de măsurare a unei dimensiuni știe cu siguranță ce probleme insurmontabile trebuie depășite atunci când itemii nu corelează bine între ei sau când scala are o slabă consistență internă. A construi o scală bună este o operațiune extrem de dificilă, iar instrumentele prezente de obicei în chestionare sunt rareori fundamentate și verificate din acest punct de vedere. Ori, conceptele subțiri rezultate din modul de construire a variabilelor ridică întrebări asupra validității instrumentului de măsurare, problema cea mai mare a indicatorilor fiind faptul că nu reușesc să surprindă decât o mică parte a unui concept mai dens.

Apoi, să presupunem că un indice este bine construit și că prezintă în mod simultan atât validitate cât și fidelitate. Un astfel de indice nu dispune decât de o validitate temporară, întrucât viața socială este în continuă schimbare. Dacă ne gândim la un indice al comunicării, de exemplu, am fi utilizat anumite variabile acum 50 de ani și cu totul alte variabile în secolul XXI, odată cu explozia mijloacelor de comunicare. Un indice din prezent nu poate fi valabil peste 100 de ani întrucât este imposibil de anticipat în ce fel vor comunica oamenii peste un secol.

Criticile analizei cantitative pot continua cu foarte multe argumente, însă mă voi apropia de final cu digresiunea lui Dogan (1994) asupra limitelor cuantificării. Dogan arată că atunci când există un decalaj între metodele statistice, pe de o parte, și cantitatea și calitatea datelor, pe de altă parte, apare pericolul supracuantificării. Metodologia statistică s-a dezvoltat spectaculos însă eforturile nu au fost la fel de susținute în privința calității datelor și de foarte multe ori chiar în privința cantității acestora. Este adevărat că se fac foarte multe studii sociale în urma cărora sunt produse muniți de date însă, de exemplu, România are goluri imense în statisticile oficiale internaționale.

Dogan mai arată că cercetătorii preferă anumite modele statistice pentru simplul motiv că au indicatori cuantificabili, ori strategia corectă este mai întâi de formare a conceptelor și abia apoi de încercarea de cuantificare a acestora. O observație foarte interesantă este îndreptată în privința utilizării valorilor medii în comparațiile cros-naționale. Media presupune

---

<sup>1</sup>literele înclinate nu sunt la fel în original; prin înclinare am încercat să subliniez importanța lor

un grup relativ omogen, însă orice țară prezintă o eterogenitate mai mare sau mai mică. În plus, dacă distribuția unei variabile este deplasată de la normalitate, atunci media nu reflectă tendința centrală.

Deși cercetătorii utilizează în cercetările lor indicatori oficiali cum este PIB-ul, trebuie subliniat faptul că acesta este o măsură puternic subestimată în țările care prezintă o economie subterană. În primii zece ani după căderea regimului comunist, România a fost caracterizată cu un procent substanțial de economie gri, iar în cazul Indiei procentul este probabil și mai mare.

Cât privește relația dintre metodele statistice și scopul cercetării sociale, Dogan citează o afirmație a lui Tukey: “Mai bine un răspuns aproximativ la o întrebare corectă, decât un răspuns exact la o întrebare greșită”.

Tone de analize sofisticate nu pot să înlocuiască un raționament logic printr-o tehnică simplă și nesofisticată. Așa cum computerul nu poate înlocui omul în activități de rațiune și sinteză, analizele statistice nu vor putea niciodată să înlocuiască capacitatea critică a cercetătorului de a utiliza concepte și de a face analogii bazate pe experiență. Ideea este surprinsă într-un alt context și de către Ragin (1997), care face o diferență între conceptul de probabilitate utilizat în statistică și conceptul de plauzibilitate utilizat în analizele calitative. Ceea ce este probabil nu este întotdeauna și plauzibil, iar plauzibilitatea nu se poate determina statistic, ci numai prin studierea în profunzime a cazurilor studiate.

Interesantă, pentru viitor, ar fi explorarea aplicabilității conceptului de probabilitate bayesiană în studiile calitative. Statistica bayesiană diferă semnificativ față de cea clasică, pornind exact de la ideea de probabilitate, iar în varianta bayesiană aceasta este determinată simultan atât de experiența cercetătorului cât și de evidența empirică. În plus, statistica bayesiană poate să lucreze foarte bine pe cazuri puține și oferă soluții la fel de exacte chiar dacă multe dintre date lipsesc. Într-un mod surprinzător, problema cazurilor puține și cea a cazurilor lipsă (ori a diversității limitate) ar putea să-și găsească o posibilă rezolvare în metode mai puțin “ortodoxe” însă la fel de științifice.

În legătură cu erorile studiilor cantitative, Dogan mai arată că cercetătorii uită mult prea des că eroarea maximă de sondaj (de regulă  $\pm 3\%$ ) este garantată doar la nivel național, ori de cele mai multe ori ei se avântă la comparații între straturi ca și cum ar avea aceeași eroare. Apoi mai există și erorile de non-eșantionare, care ar putea fi controlate însă pe care puțini le iau în calcul foarte serios. De exemplu, sondajele de opinie românești au devenit, din varii motive, suspect de ieftine.

Concurența acerbă de pe piața de cercetare românească a determinat o scădere dramatică a costului de cercetare, însă aceasta nu are un efect vizibil asupra erorii declarate. Eșantioanele au tot o eroare maximă de  $\pm 3\%$ , însă calitatea și încrederea în cifrele rezultate lasă mult de dorit. Diferența de preț la care se renunță are cel puțin două efecte: primul, o scădere a salariilor operatorilor de teren (iar dacă aceștia nu sunt suficient de motivați nu vor completa chestionarele în mod corect, ca să nu spunem că le pot chiar falsifica), iar al doilea se reflectă în controlul post-cercetare. În mod normal, o cercetare serioasă beneficiază de un audit care să controleze munca operatorilor de teren (în mod ideal realizat de către o altă firmă de cercetare). Renunțarea la acest audit oferă libertate operatorilor de teren să încalce instrucțiunile, iar toate aceste lucruri nu se manifestă în cifrele declarate, vizibile, ci în calitatea datelor, invizibilă însă infinit mai importantă. Pe aceeași lungime de undă, Dogan demonstrează că statisticile oficiale, la nivel agregat, sunt cu atât mai slabe cu cât nivelul de dezvoltare socio-economică a unei țări este mai scăzut.

În concluzie, putem spune că o analiză cantitativă bine făcută are merite certe și incontestabile, însă a considera metoda cantitativă ca având răspunsul corect la toate întrebările este o eroare. Multe dintre aspectele cercetării sociale nu pot fi mânuite cu instrumente cantitative, la fel cum un calculator care poate efectua milioane de operații matematice pe secundă nu poate înlocui un minut de reflexie critică a cercetătorului.

## 1.5 Complementaritatea celor două metode

În introducerea acestui capitol am arătat că relația dintre metoda calitativă și cea cantitativă este o reproducere a diferenței între modurile de a face cercetare în științele socio-umane și în științele exacte. Istoria dezvoltării acestor două metode surprinde diferitele paradigme prin care a trecut sociologia până la faza actuală. Qnwuegbuzie și Leech (2005) au observat în acest sens câteva lucruri interesante care merită a fi menționate.

În faza ei inițială, sociologia a trecut prin faza pozitivistă în care metodologia de cercetare promova neimplicarea valorică și neutralitatea axiologică. Teza pozitivismului este că realitatea este obiectivă, în afara simțurilor noastre și că poate fi măsurată utilizând metode științifice, prin intermediul unor instrumente valide și fidele. Într-o mare măsură, este principiul după care se ghidează analiza cantitativă în toate formele ei actuale.

Începutul secolului 20 aduce ideea că realitatea socială este construită, deci în esență subiectivă, fapt care a dat naștere paradigmei cercetării calitative și utilizarea abordărilor interpretativ-hermeneutice. În directă opoziție cu neutralitatea axiologică a pozitiviștilor, în analiza calitativă cercetarea este influențată într-o mare măsură de valorile cercetătorului, care înțelege mult mai bine fenomenul (asumându-și aceste valori) decât un alt cercetător situat în afara realității sociale investigate.

Acesta este momentul în care metodele s-au despărțit, iar diviziunea încă mai scindează metodologia actuală de cercetare socială. Prin anii '50-'60 apare post-pozitivismul, un compromis între calitativ și cantitativ: deși erau de acord că realitatea este construită, s-a observat totuși existența unor relații stabile între fenomenele sociale, indiferent de modul cum erau interpretate subiectiv. În această etapă s-a întărit metoda științifică de cercetare (Popper, 2000) și importanța deosebită a metodologiei.

În același timp, s-au întărit și teoriile radicale, cum ar fi post-structuralismul și post-modernismul, care afirmă că nu există o realitate socială obiectivă, ci mai multe realități văzute din perspective diferite. Înarmați cu sentimentul superiorității metodei calitative, adepții Tezei Incompatibilității susțin că paradigmele calitative și cantitative nu pot coexista.

Tot în anii '60 apare o mișcare pragmatică care a generat Teza Compatibilității: abordările calitative și cantitative nu sunt mutual exclusive, relația dintre cele două constând în evenimente izolate de-a lungul unui continuum de cercetare științifică, unde rolul teoriei este central. Adepții acestei teze cred în existența ambelor orientări, atât subiectivă cât și obiectivă, utilizând atât logica inductivă cât și logica deductivă.

Qnwegbuzie și Leech (2005) au făcut chiar efortul de a enumera asemănările dintre cele două metode:

- ambele merg pe o abordare teoretică;
- ambele utilizează ipoteze de cercetare;
- ambele colectează date;
- ambele interpretează într-un fel sau altul datele culese;
- ambele încearcă verificarea datelor (validarea acestora);
- ambele încearcă reducerea (sintetizarea) datelor la o concluzie compactă, care utilizează cât mai puține variabile care să explice cât mai mult;
- reiau ciclul amendării teoriei în funcție de datele empirice culese.

Cu alte cuvinte, cele două metode sunt complementare, iar cei care nu văd avantajele mutuale sunt incapabili de “cercetare bilingvă” și se transformă în așa-numiții “cercetători pragmatici” (Qnwuegbuzie și Leech, 2005, p.268).

Cercetarea calitativă și cea cantitativă sunt, în fapt, două fețe ale aceleiași monede. Ambele încearcă să descrie într-un mod cât mai sintetic aceeași realitate studiată, prin metode diferite. Există situații în care metoda cantitativă nu are aplicabilitate, după cum în alte situații metoda calitativă de cercetare nu este cea mai potrivită. În funcție de situația concretă de cercetare, se utilizează una sau cealaltă, însă nu se poate spune că una este net superioară celeilalte. În fapt, cunoscătorii metodologiei de cercetare socială consideră, în mod just, că aceste două metode sunt complementare, rezultatele uneia fiind utilizate ca date de intrare pentru cealaltă.

Există un studiu sociologic cu ajutorul căruia voi arăta că există situații în care analizele cantitative nu numai că sunt în general imposibil de aplicat, însă chiar în situațiile în care pot fi aplicate nu pot genera concluzii de substanță: este vorba despre studiul lui Seymour Martin Lipset *American exceptionalism - Japanese uniqueness* (1994).

În lumea vestică (Europa de Vest și Statele Unite), marile teorii care au încercat să explice dezvoltarea economică au făcut apel la valori orientate predominant spre un comportament economic rațional, iar unele dintre aceste valori și caracteristici culturale se regăsesc în spiritul protestant răspândit cu precădere în Statele Unite. Alături de alți factori sociali și politici, explicația succesului economic constă în capacitatea individului de a se dezvolta prin efort personal și prin competiția dură care obligă individul să lupte în permanență pentru a urca dintr-o clasă inferioară într-una superioară.

Deși această explicație este perfect adaptată modelului cultural american, explicând într-o mare proporție și succesul economic vest-european, acest tip de teorii sunt puse într-o mare dificultate atunci când sunt supuse comparației cu o singură națiune asiatică: Japonia. După al doilea Război Mondial, Japonia a demonstrat un miracol de dezvoltare economică, concurând de la egal la egal cu Statele Unite.

Studiul comparativ al lui Lipset analizează doar aceste două națiuni, încercând, pe de o parte, să sublinieze diferențele culturale dintre cele două națiuni, iar pe de altă parte, să arate volatilitatea teoriilor economice bazate exclusiv pe o literatură vestică. Este un exemplu perfect de studiu a două cazuri extrem de diferite care însă prezintă același fenomen rezultat: o puternică dezvoltare economică.



Dacă modelul american este bazat pe individualism și modernitate (înțelesă ca renunțare la modelul social feudal și adoptarea unui model liberal de *laisse-faire*), acumularea de capital în Japonia nu are nimic în comun cu niciuna dintre condițiile de bază din lumea vestică. Japonezii sunt orientați într-o direcție opusă: dacă în vest individul este puternic valorizat și condiționat social să iasă în evidență prin propria lui capacitate, în Japonia individul are o puternică reținere în a ieși în evidență, modelul social fiind centrat spre grupul din care face parte. Sistemul social japonez prezintă un sistem ierarhic cu puternice caracteristici feudale, așa încât teoriile vestice asupra succesului economic sunt complet invalidate de acest unic caz din lumea estică.

“Cuiul care iese trebuie bătut la loc”, proverbul japonez citat de Lipset (1994, p.160), descrie perfect orientarea non-individualistă a japonezilor: orice individ care iese în evidență este tratat ca deviant și pedepsit/marginalizat de grupul din care face parte. Anulând interesul personal, individul japonez își concentrează toate eforturile în slujba interesului colectiv. La limită, interesul național primează atât în fața interesului individual cât și în fața intereselor companiilor private, chiar dacă între companii se manifestă o competiție puternică (strategiile de afaceri fiind întocmite exact pe principiile alianțelor din războaielor multiplilor Shoguni din perioada feudală).

Lipset utilizează o multitudine de studii atitudinale comparative între cele două națiuni, prezentând distribuții de procente ale răspunsurilor la itemi cu o scală de răspuns cel mult ordinală. Aproape niciuna dintre variabilele prezentate nu este cantitativă, iar studiul nu face apel la nici o analiză cantitativă avansată de tipul regresiei ori analizei factoriale.

Singura modalitate prin care se pot aplica analize cantitative este aplicarea unui instrument de cercetare pe un eșantion reprezentativ din fiecare țară, apoi prin testarea paralelă a unui model statistic. Dacă modelul generează rezultate asemănătoare atunci putem spune că teoria pe care se bazează este validă, devenind un punct de reper pentru studiile ulterioare. Din păcate însă, un model statistic bazat pe teoriile vestice asupra atitudinilor este în mod invariabil invalidat de atitudinile contrare (opuse chiar) ale japonezilor, datorate unui model cultural diferit.

Modelele statistice nu reușesc să spună absolut nimic despre cauzele acestor diferențe, întrucât ceea ce măsoară cercetările cantitative din prezent nu sunt decât efectele unor decizii din trecut. Măiestria lui Lipset este în mod excepțional demonstrată prin claritatea cu care descrie transformarea socială din Japonia începută pe la mijlocul secolului al XIX-lea, când Japonia se afla încă într-o perioadă

tradițional feudală. Elitele aristocratice japoneze au înțeles faptul că lumea vestică (industrializată) începea să depășească mult capacitatea Japoniei de creștere economică, așa că au decis că industrializarea este un proces obligatoriu.

Fenomenul economic japonez constă exact în deciziile acestei perioade și în extraordinara abilitate de predicție ale elitelor japoneze care au reconstruit din temelii societatea, păstrând în același timp tradițiile și stratificarea specific feudală. În acea perioadă, capitalul economic era deținut în cea mai mare parte de către elitele aristocratice (shogunii), însă industrializarea presupune o relație economică între agenți din clasa mijlocie. Singura pătură dintre clasa de jos și aristocrație, care deținea un status suficient de ridicat pentru a îndeplini această funcție erau samuraili, cărora le-au fost acordate sarcinile de a înființa companii private. Într-o societate puternic tradițională, în care schimbările macrosociale erau de neînchipuit, elitele japoneze au suspendat războaiele feudale interne și au utilizat atât prestigiul samurailor cât și legitimitatea împăratului Meiji pentru a restructura din temelii întreaga societate japoneză.

În decursul a numai o jumătate de secol, Japonia a devenit o societate puternic industrializată și urbanizată, fenomen incredibil datorat unei planificări minuțioase gândite de sus în jos. Transformările sociale produse de planificatorii Meiji sunt asemănătoare cu încercările comuniste din multe privințe: sunt centralizate, planificate, subsumate interesului comun, societatea urmărind interesul tuturor mai degrabă decât competiția dintre indivizi. Japonezii au reușit construirea unei societăți de consum, a unui sistem capitalist bazat pe competiție. Prin comparație, succesul japonez (explicat prin canalizarea tradițiilor în direcția dorită) reduce efortul ideologiei comuniste la o încercare patetică.

Nici unul din postulatele teoriilor vestice asupra circulației capitalului ori a luptei dintre clase nu se potrivește cu realitatea japoneză. În Japonia nu există luptă de clasă, întrucât individul nu se consideră asuprit de capitaliști ci consideră munca o datorie morală față de angajator. Planificarea centralizată încă mai există, însă nu are nimic în comun cu ideologia comunistă; Ministerul Comerțului Internațional și Industriei din Japonia nu face nimic altceva decât să orienteze economia japoneză în direcția maximizării probabilității de profit pe baza unor proiecții prospective. Dacă Marx ar fi trăit în ziua de astăzi, este foarte probabil să fi fost profund descumpănit de funcționarea societății japoneze.

Japonia nu este singurul caz în care societatea a fost transformată de sus în jos, în absența ideologiei comuniste. Într-un studiu comparativ cel puțin la fel de relevant, Ali Kazancigil (1994) arată că Turcia este un alt exemplu de națiune care a urmărit transformări sociale radicale, fiind singura țară din lumea musulmană care prezintă o traiectorie puternic orientată spre o democrație participativă. Spre deosebire de Japonia însă, Turcia nu prezintă același succes. Kazancigil citează o multitudine de studii comparative între Japonia și Turcia, a căror scop principal a fost acela de a explica pe de o parte succesul Japoniei și pe de altă parte insuccesul Turciei (ori mai bine spus succesul limitat al acesteia).

Ca și în cazul Japoniei, condițiile de plecare ale Turciei spre un final democrat includ un imperiu (extins până la porțile Europei de Vest) și o tradiție de supunere față de Sultan. Există însă și o diferență fundamentală: în timp ce în Japonia împăratul era considerat zeu, în teritoriile musulmane tradiția pleacă de la o supunere totală doar față de Allah.

Din acest motiv, lumea musulmană este și astăzi dezbinată de rivalități locale, fără politici unitare și coerente cum sunt cele din partea vestică a lumii. Țările musulmane prezintă așadar un nivel extrem de scăzut al implicării statului; singură Turcia face excepție de la această regulă, fapt explicabil dacă se ia în calcul necesitatea menținerii unui imperiu atât de extins cum a fost cel Otoman. Un astfel de imperiu nu putea să reziste decât printr-o centralizare puternică realizată prin instaurarea unei tradiții de supunere față de Sultan, conducătorul absolut al imperiului. Dacă dogma religioasă prevedea că oamenii sunt sclavii lui Allah, în Imperiul Otoman oamenii erau sclavii Sultanului. Rivalii și oponentii acestuia au fost fie eliminați fie le-au fost confiscate pământurile (sursa lor principală de venit și de putere).

Încă de la începutul formării sale, Imperiul Otoman a dezvoltat un puternic aparat birocratic, ridicând nivelul centralizării statale la un nivel foarte înalt. Această formulă de succes i-a permis stabilitatea în timp și a prevenit o decădere specifică societăților musulmane. De-a lungul a sute de ani de existență această tradiție a reușit o deplasare uimitoare, unică în lumea islamică, a trăirii profund religioase a turcilor: deși rămân în continuare un popor religios, ei acceptă acum legitimitatea organizării laice alături (nu mai prejos, nici mai presus) de organizarea religioasă.

Această tradiție a permis unor elite reformatoare școlite în țările din Vest (începând cu a doua jumătate a secolului XIX) să împingă transformarea socială și mai departe: în loc de sclavii Sultanului erau de acum sclavii “statului”, promovând o economie națională, iar grupul lui

Mustafa Kemal Atatürk a înlăturat pur și simplu organizarea religioasă de la putere. Nu întâmplător, Atatürk a dobândit ulterior el însuși o faimă aproape religioasă, devenind părintele noii națiuni. Ideologia promovată de acesta este cel puțin la fel de fervent apărată ca mișcarea fundamentalistă islamică, deși obiectivele finale ale celor două mișcări sunt puternic antagonice.

Aflată într-un trend economic puternic ascendent, Turcia se zbate între curentul modernist și cel tradițional islamic. Rămâne de văzut în ce măsură va reuși Turcia să-și consolideze eforturile, întrucât evenimentele din Piața Taksim în 2014 nu sunt extrem de încurajatoare, iar presiunea kurdă (printre altele) împiedică deocamdată aderarea Turciei la Uniunea Europeană. Deși nu prezintă același succes economic ca al Japoniei, se poate spune totuși că Turcia reprezintă un caz extrem de interesant al unei transformări sociale planificate de sus în jos, într-o direcție democratică.

Japonia rămâne singurul caz care contrazice teoriile economice și sociale din vest; iată cum este posibil ca un singur caz contradictoriu să răstoarne întregi teorii, după cum sunt posibile comparații extrem de substanțiale (fără a face apel la analize statistice complicate) utilizând doar două sau trei cazuri în perspectiva evoluției lor istorice.

## Capitolul 2

# *Via media*: analiza calitativ-comparativă

Obiectivul principal al acestui capitol este acela de a descrie logica și modul de funcționare al unei noi metode de a analiza comparativ anumite cazuri studiate. Studiarea cazurilor are un echivalent conceptual în așa numitele *studii de caz*, care sunt operațiuni specifice calitative. De aceea, analiza comparativă (așa cum va fi descrisă în acest capitol) este ea însăși o operațiune calitativă, iar denumirea ei completă este de analiză *calitativ-comparativă*. Pentru ușurința în exprimare însă, voi utiliza uneori termenul mai simplu și mai scurt de *analiză comparativă* (multe dintre articolele de specialitate fac aceeași asociere între cele două denumiri).

Analizele publicate de obicei în științele sociale și politice se centrează pe unul sau celălalt dintre polurile calitativ sau cantitativ. Sunt nenumărate publicații bazate pe un număr mare de cazuri (cu analize statistice mai mult sau mai puțin complicate), după cum sunt multe alte publicații care analizează în mod extensiv un studiu de caz ori cel mult două pentru comparație.

În foarte puține publicații se constată studierea unui număr mediu de cazuri (mai multe decât în studiile în profunzime, însă mai puține decât în studiile cantitative); dacă am produce un grafic în care pe orizontală sunt numărul de cazuri studiate iar pe verticală numărul de publicații, curba rezultată are o formă de U: foarte multe publicații la extremități (cercetări pe cazuri puține și pe cazuri multe) și foarte puține în mijloc (pe un număr mediu de cazuri).

În continuare voi sublinia acest al treilea design de cercetare posibil, unul care să fie suficient de calitativ încât să culeagă informații în

profunzimea fiecărui caz în parte dar care să dispună în același timp de o metodologie suficient de riguroasă care să asigure replicarea rezultatelor dacă un alt cercetător ar aplica aceeași metodă. De Meur și Rihoux (2002) descriu această analiză ca pe o “via media” între primele două design-uri de cercetare, unul care să combine în mod fructuos calitățile ambelor. De esență calitativă, însă desfășurată conform unui algoritm matematic, analiza comparativă dobândește încrederea utilizatorilor fiind pe de o parte ușor de înțeles și pe de altă parte replicabilă. Așa cum o vede Rihoux (2003), analiza comparativă este un pod care unește prăpastia dintre cele două margini opuse.

Deși acceptate după toate standardele (există o comunitate științifică, există teorii și școli concurente pe diverse ramuri), există o percepție mai mult sau mai puțin difuză că științele sociale reprezintă partea moale a științei în timp ce științele naturii, fizica, chimia și celelalte reprezintă nucleul dur al științei. Acest lucru se datorează dificultății cu care se găsesc legități cu caracter general, precum și complexității vieții sociale și a stadiului relativ incipient în care se află științele sociale în comparație cu tradiția centenară a științelor exacte.

Fenomenele sociale sunt în mod evident greu de abordat întrucât pentru studierea lor sunt necesare cunoștințe multiple, aproape enciclopedice. În timp ce unele fenomene sociale își au originea în decizii economice, altele sunt produse de politicile sociale și publice; cutumele și tradițiile sunt diferite în funcție de regiunea geografică și de condițiile climaterice; altele sunt explicate de factori culturali. Sunt o mulțime de factori care pot acționa, de cele mai multe ori în combinații greu de imaginat. Selecția condițiilor cauzale relevante presupune o muncă de Sisif și o înțelegere a fenomenelor sociale dublată de intuiție, întrucât nu toate cauzele se manifestă la fel și nu toate influențează fenomenul studiat în aceeași măsură.

Întrucât analiza comparativă se desfășoară pe un număr relativ restrâns de cazuri, o decizie cel puțin la fel de importantă se referă la selectarea cazurilor relevante pentru analiză. Aspectele metodologice ale analizei comparative se află într-o strânsă relație unele cu celelalte; de exemplu, cazurile studiate pot fi descrise atât din perspectiva evoluției lor temporale cât și din punctul de vedere al diverselor concepte în termeni de cauze și efect, însă cauzalitatea (așa cum voi arăta la secțiunea 2.1) are o strânsă legătură cu dezvoltarea temporală a evenimentelor (mai exact cu precedența temporală a cauzelor asupra efectului). În aceeași legătură se află tratarea condițiilor necesare și suficiente pentru producerea unui efect, subiect aflat într-o legătură organică cu analiza cauzală.

Se pot construi secțiuni separate atât pentru analiza cauzală cât și pentru analiza temporală însă ele sunt legate una de cealaltă în mod organic și recursiv. Structura unei lucrări nu poate fi decât liniară (capitole și secțiuni succesive) însă lectura acestora trebuie înțeleasă recursiv, la modul ideal multe secțiuni ar trebui să aibă aceeași poziție ierarhică față de capitolul de care aparțin (pe modelul unei organigrame).

Voi începe, în prima secțiune, cu o discuție comparației din perspectiva analizei cauzale, apoi o încadrare a acesteia în diverse cadre analitice (mai multe moduri de a privi relația dintre cauze și efecte) bazate pe canoanele lui Mill, după care voi descrie metodologia analizei comparative așa cum este prezentată în literatura de specialitate actuală. În final voi puncta cele mai importante critici care au dus ulterior la dezvoltarea acestei analize prin înglobarea teoriei mulțimilor vagi.

## 2.1 Comparația și analiza cauzală

Analiza comparativă nu este un scop în sine, ci s-a dezvoltat ca o metodologie robustă în urma unor nevoi de analiză sistematică a unor fenomene de interes. Probabil cea mai des întâlnită întrebare în știință (în oricare știință) se referă la cauzele apariției fenomenelor. De aceea, pentru a înțelege în profunzime analiza comparativă, o să tratez în această secțiune câteva perspective asupra relației cauzale dintre fenomene.

Cauzalitatea este un concept care a ridicat dintotdeauna dispute, mai întâi în filozofie apoi în știință în general (chiar și în știința de astăzi noțiunea de cauzalitate este utilizată cu precauție doar în cadrul unor experimente controlate). Zamfir (1990) a identificat o multitudine de probleme ce stau în calea stabilirii cauzalității (multe dintre acestea se suprapun și peste problemele tratate în secțiunile 1.4 și 2.6, în timp ce altele vor fi reluate pe parcursul lucrării în secțiuni specifice):

- problema sensului cauzării;
- problema duratei și a distanței;
- problema cauzelor aparente;
- riscul de a fi corect din rațiuni false;
- problema numărului<sup>1</sup>;
- problema interdependenței;
- problema interacțiunii;
- variația contextuală a cauzalității;
- problema empirismului cauzelor și a efectelor.

---

<sup>1</sup>o enunțare completă ar fi “problema numărului de cazuri studiate”

Acestea sunt câteva exemple (la care se pot adăuga problemele din secțiunile amintite mai sus) care demonstrează dificultățile cu care se confruntă cercetătorii în analizele sociologice și limitele inerente pe care este bine să le conștientizăm în activitatea de cercetare.

În această lucrare, distincția terminologică pe care o consider de o importanță critică este cea dintre “cauză” și “condiție cauzală”. Dintre toți factorii care contribuie la apariția unui fenomen, nu toți sunt cauze clare ale acestuia, unii fiind mai mult condiții cauzale favorizante, care contribuie la producerea fenomenului. De exemplu, vântul nu este o cauză a apariției unui incendiu, ci factorul care a provocat inițial focul, însă vântul este o condiție cauzală importantă într-un incendiu întrucât îl întreține și îl răspândește.

Întrucât cauzalitatea este atât de dificil de stabilit, în analiza comparativă nici nu se folosește cuvântul “cauză”, toți factorii fiind considerați “condiții cauzale”; fenomenul este produs de o combinație de condiții cauzale dintre care unele sunt *necesare* pentru producerea efectului în timp ce altele sunt *suficiente* prin ele însele (distincția este prezentată în secțiunea 2.2); Zamfir (1990) numește ultimele condiții “factori determinanți”.

Nu există o definiție exactă (sau cel puțin agreată) a conceptului de cauză ori a celui de efect. În opinia mea, filozofii utilizează acești termeni doar cu referire la exemple concrete. Plecând de la Aristotel (cu un model filozofic teleologic), până la Galileo în timpul Iluminismului (care a fost foarte aproape de un model coerent), Hume este printre primii care s-au aplecat în mod științific asupra acestor noțiuni, susținând că nu putem percepe relațiile cauzale în mod direct, astfel că nu putem cunoaște cauzalitatea în afara regularităților observate în cadrul unor asocieri de evenimente.

În lipsa unei definiții coerente, Hume (1748) a arătat doar că “*obiecte similare sunt legate întotdeauna de obiecte similare*” și definește cauza ca fiind:

...un obiect, urmat de un altul, unde toate obiectele similare cu primul sunt urmate de obiecte similare cu al doilea.

Viziunea lui Hume este depășită astăzi, întrucât lucrează cu “cauze” nu cu “condiții cauzale” și nu tratează interdependența cauzală, considerând cauzele de sine stătătoare. Totuși, consider că este important să o prezint întrucât își are originea în experiența empirică și a dat naștere mai târziu așa numitului model deductiv-nomologic.

În cadrul unui astfel de model, sunt asumate presupunerile că obiectul cauză există independent de existența obiectului efect, iar obiectul cauză



îl precede temporal pe obiectul efect. Hume adaugă la aceste ingrediente și faptul că “*dacă primul obiect nu ar fi, al doilea nu ar fi existat*”. Cu alte cuvinte, nu se cunoaște mecanismul exact prin care cauza generează efectul, însă știm din experiență că dacă nu s-ar fi produs cauza, nu s-ar fi produs nici efectul. Cauza înghețării apei este temperatura scăzută; dacă temperatura nu ar scădea la 0 grade Celsius, apa nu ar îngheța.

Definiția cauzalității a lui Hume evidențiază trei elemente care pot fi verificate prin observație:

*X este cauză a lui Y dacă:*

*Precedența:* X precede temporal pe Y

*Adiacența:* ambele sunt prezente în același cadru spațio-temporal

*Concomitența:* X și Y sunt întotdeauna concomitente

Zamfir (1990) arată că modelul lui Hume este complet neadecvat explicației sociologice, întrucât rareori o condiție cauzală poate să genereze un efect în mod direct. Realitatea este mult mai complicată decât un simplu model Cauză - Efect, procesul determinativ al fenomenelor fiind dispus sub forma unui lanț cauzal în care condițiile cauzale se află într-o relație de succesiune temporală, cauzând-se ele însele reciproc. Identificarea unui asemenea model nu poate fi făcută printr-o simplă inspecție empirică a instanțelor studiate, ci printr-un proces de explicație deductivă (o concepție cât se poate de modernă, rar întâlnită în literatura sociologică actuală).

Substanța discuției depășește în unele aspecte știința sociologiei, tratând unele concepte pur filozofice. De exemplu, a vorbi de cauzalitate înseamnă a presupune existența atât a condițiilor cauzale cât și a efectelor. Dintr-un anumit punct de vedere, tocmai existența acestora poate fi pusă în discuție întrucât realitatea, fiind construită social, este un concept discutabil. Anumiți factori cauzali pot fi puși în discuție doar pentru că ei au fost definiți, însă acest lucru nu înseamnă nici pe departe că factorii respectivi *există cu adevărat*. Mai mult, definirea factorilor poate fi făcută în mod diferit în culturi diferite, în funcție de modul cum a fost construită realitatea.

Construcția realității este o temă care ține mai mult de filozofie decât de cercetarea socială obișnuită. În mod curent, cercetătorii nu problematizează realitatea ci o iau ca pe un fapt dat. Deși este evident că fiecare individ percepe realitatea într-un mod propriu, se poate constata o ciudată convergență a tuturor realităților individuale percepute, prin intermediul acordurilor în comunicare. Astfel, cercetătorii au convingerea (îndreptățită din acest punct de vedere) că

în cadrul realității percepute există legături cauzale între fenomene și că acestea pot fi studiate prin metode științifice.

Uneori, cauzele sunt ușor de văzut și de cuantificat; probabil cel mai celebru exemplu este mărul lui Newton: cauza căderii mărului este atracția gravitațională. Aceste tipuri de raționamente sunt relativ stabile în timp și spațiu, fiind specifice mai ales științelor naturale (acele științe denumite “dure”). În alte științe însă (cele la care mă refer aici sunt desigur cele sociale și umaniste) prezintă o fluctuație a tuturor factorilor care interacționează, precum și a fenomenelor studiate. Dacă în fizică este foarte simplu de prezis ce se va întâmpla lăsând un pahar cu apă la minus 20 grade, în sociologie este aproape imposibil de prezis ce va face un om, în special aflat într-o situație limită.

Aproape toate fenomenele studiate se compun dintr-o cantitate pe care o denumim “explicabilă” (în anumite circumstanțe oamenii prezintă comunalități acționale), însă mai există o cantitate complet haotică pe care teoriile actuale nu o pot explica. În analiza cantitativă, prima se numește “variație explicată” și se măsoară în cazul regresiei cu coeficientul de determinație, iar a doua se numește pur și simplu “eroare”, este cantitatea complementară din total variație rămasă neexplicată.

A doua cantitate este cea mai interesantă, fiind compusă la rândul ei din două părți:

- o parte care ar putea fi explicată dacă teoriile cu care lucrăm ar fi mai avansate;
- o parte complet neexplicabilă, haotică, ce caracterizează comportamentul ființelor umane, inerent impulsiv și dependent de o componentă emoțională necontrolabilă

Într-adevăr, în regresie se lucrează cu modele de variabile independente (cauzele), pe care cercetătorul presupune că au o contribuție semnificativă la explicarea variației variabilei dependente (a efectului). Dacă în modelul construit de cercetător intră toate variabilele importante care au o influență majoră asupra efectului, atunci cantitatea de variație neexplicată este minimă și va fi compusă doar din acțiunea haotică. O posibilă problemă constă în presupunerile pe care le fac cercetătorii asupra realității construite, alte culturi construind alte realități (implicit alte modele cauzale). Presupunând că “realitatea” este corect construită și că modelul este adecvat acestei realități, eroarea va fi într-adevăr minimă iar valoarea coeficientului de determinație va atinge un maxim posibil.

Dacă însă realitatea nu este corect construită ori modelul cercetătorului nu este adecvat acestei realități, o cantitate importantă de variație pleacă din zona explicată înspre eroare, zona neexplicată. Eroarea va avea o valoare mult mai mare (ori invers, cantitatea de variație explicată va fi mult mai mică decât în realitate), însă acest fapt nu este cunoscut apriori de către cercetător. Nu cunoaștem și nu vom cunoaște niciodată care este cantitatea totală de variație care poate fi explicată; se poate foarte bine ca modelul nostru să nu conțină toate variabilele relevante.

Partea cealaltă, neexplicată, se referă la componenta complet impredictibilă a naturii umane: o persoană poate avea o preferință acum și una complet diferită peste două minute. Schimbarea preferințelor este unul din factorii care determină comportamentul acțional (subiect extrem de suculent în marketing și studierea paternurilor de consum), iar comportamentele acționale imprevizibile agregate la nivel de grup, comunitate, regiune etc. se transformă în factori sociali haotici ai căror efect este foarte greu de prezis ori estimat.

Este adevărat că indicatorii sociali sunt mai stabili la nivel agregat decât măsurătorile la nivel individual, însă nici măcar aceștia nu sunt feriți complet de acțiunea haotică. Între anumite limite, cercetătorul se poate baza pe stabilitatea indicatorilor agregați însă această stabilitate este numai temporară, pe termen scurt, întrucât nimeni nu poate să prevadă ce se va întâmpla cu valorile lor în viitor.

Cauzele care pot determina schimbarea valorilor acestor indicatori sunt multiple; în științele sociale am putea spune că numărul virtual de cauze (de condiții cauzale, mai exact) este foarte mare. Sigur că unele dintre acestea sunt irelevante, însă combinațiile posibile de factori cauzali sunt potențial infinite, în ideea că nicio cauză nu are un efect semnificativ de una singură, ci numai în combinație cu alți factori cauzali.

Se configurează astfel un mix cauzal extrem de subtil, pentru care Mill (1843) a găsit o excelentă comparație cu amestecul substanțelor chimice: în componența cauzalității intră o multitudine de compuși care interacționează în mod subtil unii cu alții, uneori anulându-se reciproc iar alții provocând adevărate explozii sociale.

Zamfir (1990) a denumit această situație “problema interacțiunii”, ridicând o limită de principiu a analizei sociologice. Fenomenele sociale nu sunt produse de acțiunea separată a unei cauze sau alta, ci de influența comună a mai multor cauze care interacționează și se influențează reciproc până la producerea efectului.

În aceste condiții, identificarea unei cauze anume dintr-o combinație cauzală complexă este o sarcină asemănătoare cu depistarea unui singur

compus chimic dintr-o substanță cu o formulă complexă. Această demers nu este deloc unul simplu; studiind lucrările lui Mill asupra cauzalității, Aickin (2004, p.533) identifică interpretarea acestuia într-unul din canoanele discutate, “metoda variațiilor concomitente”:

Dacă un fenomen variază într-o măsură oarecare în timp ce un alt fenomen variază într-o manieră particulară, primul este fie o cauză fie un efect al aceluși fenomen, ori este conectat de acesta printr-o legătură cauzală

Această perspectivă a lui Mill asupra cauzalității pare să fie mai mult cantitativă decât calitativă, întrucât se referă la noțiunea specific cantitativă de *variație*, în ciuda faptului că lucrările sale sunt utilizate ca bază de plecare pentru analiza comparativă (calitativă în esență). Spun discutabil întrucât în cercetarea cantitativă variabilele independente nu trebuie să coreleze unele cu altele (asumpția statistică de non-multicolinearitate), în timp ce în cauzalitatea chimică factorii cauzali interacționează unii cu alții, transformându-se întocmai cum două substanțe chimice combinate dau naștere unei substanțe noi.

Mill s-a concentrat pe operaționalizarea conceptului de cauzalitate; argumentând că aceasta nu poate fi demonstrată decât în cadrul unor experimente, a produs binecunoscutele lui canoane (descrise pe larg la secțiunea 2.4) arătând că, în timp ce canoanele diferenței și a concordanței se bazează pe o potrivire perfectă între cauză și efect, utilizarea variațiilor concomitente nu presupune acest lucru pentru a face inferențe cauzale. O singură excepție este suficientă pentru eliminarea unui factor cauzal cu primele două canoane, însă variațiile concomitente admit și astfel de excepții în analiza cauzelor. În fapt, funcția principală a primelor două canoane este aceea de a *elimina* cauze, în timp ce o analiză complexă analizează cauzele și sub aspectul posibilității sau chiar probabilității acestora, ceea ce duce mai mult spre partea cantitativă a cercetării.

Deși puternică (dacă judecăm după numărul utilizărilor și răspândirea ei), analiza cantitativă are totuși un neajuns fundamental în ce privește studiul analizei cauzale: cele mai multe cercetări cantitative sunt transversale, fixate în timp. Cutuma cercetării cantitative este aceea a studierii unui eșantion reprezentativ suficient de mare pentru o eroare suficient de mică. Chiar dacă astfel de cercetări se repetă destul de des (fie de către aceeași instituție, fie de instituții diferite interesate de același subiect), eșantioanele sunt aproape întotdeauna diferite (extrase în mod aleator din populație).

Or, analiza cauzală își dovedește adevăratul potențial prin studierea cauzelor și a efectelor acestora *în timp*. Sigur că există eșantioane panel, însă acestea sunt extrem de rare în comparație cu cercetările transversale (fiind și mult mai costisitoare) și de cele mai multe ori sunt utilizate doar pentru a raporta o stare de fapt și schimbarea acesteia în timp; studierea temporală a cauzelor și efectelor este extrem de rară, deși Coppedge (1999) este de părere că studierea longitudinală a manifestărilor unui caz este cel mai bun design de cercetare disponibil pentru testarea ipotezelor referitoare la cauzele unor evenimente specifice. Bartolini (1993) are o opinie similară, arătând că cea mai indicată strategie de cercetare comparativă este aceea care combină un design cros-cultural cu unul cros-temporal: cazurile nu se modifică doar în spațiu și nu sunt diferite doar între culturi, ci se modifică și odată cu trecerea timpului în cadrul aceluiași spațiu și aceleași culturi.

Mahoney (2000) observă un lucru foarte interesant în legătură cu modul diferit de înțelegere a noțiunii de cauzalitate dintre abordarea calitativă și cea cantitativă: în vreme ce analiștii pe N-mic (centrați pe comparații nominale) se zbat să caute acele situații în care relația dintre condițiile cauzale și efect este perfect predictibilă (conform canoanelor lui Mill), analiștii pe N-mare admit și includ în ecuație factorul aleator (eroarea, zgomotul, așa-numitul *random noise*) și interpretează cauzalitatea în termeni de efecte parțiale și probabiliste (King et al., 2000).

Într-unul din articolele sale, Ragin (2003) punctează încă o diferență între cele două abordări. În studiile de caz, cauzalitatea se examinează în mod holistic iar aspectele observate nu sunt nimic altceva decât părți ale unui întreg; toate elementele analizate converg într-o structură unitară care nu este întotdeauna ușor de observat (de aceea, doar cercetătorii cu experiență pot observa structuri), iar sarcina cercetătorului este de a identifica modul cum s-au combinat cauzele pentru a produce fenomenul studiat (efectul). În cercetările orientate spre număr mare de cazuri, sarcina cercetătorului este de a găsi cel mai mic model posibil care să explice o cantitate cât mai mare din variația variabilei dependente. Variabilele independente se află în competiție unele cu celelalte, fiind oprite doar acelea care au un coeficient de corelație ridicat cu variabila dependentă însă un coeficient de corelație scăzut cu celelalte variabile independente. Modelul cauzal derivat în acest fel se aplică tuturor cazurilor, chiar dacă în esență fiecare caz este unic și la o studiere atentă spune propria poveste.

Există și în statistică metode de analiză bazate pe serii temporale, însă utilizarea lor este îngreunată de culegerea informațiilor care este dificilă în timp. Este nevoie de informații pe zeci sau sute de variabile, toate extrase pe mii de cazuri pentru a face inferențe statistice valide.

Cauzalitatea are și o interpretare probabilistică; una dintre problemele interpretării lui Hume (care se aplică într-o anumită măsură și la interpretarea lui Mill) se referă la proprietatea concomitenței. Se poate întâmpla ca, deși X să fie într-adevăr o cauză a lui Y, să nu se întâmple întotdeauna în mod concomitent. În teoria probabilităților există formule de calcul pentru așa-numitele probabilități condiționate; de exemplu, putem calcula probabilitatea ca Y să se întâmple *dat fiind* faptul că X s-a întâmplat deja (altfel spus probabilitatea ca Y să fie cauzat de X).

În măsura în care criteriul concomitenței a lui Hume este adevărat, atunci:

$$\begin{aligned} P(Y|X) &= 100\% && \text{probabilitatea lui Y dat fiind X este de 100\%, iar} \\ P(Y|\!X) &= 0\% && \text{probabilitatea lui Y dat fiind } nonX \text{ (dat fiind faptul} \\ &&& \text{că X nu se întâmplă) este zero} \end{aligned}$$

În mod evident această regulă este prea restrictivă pentru științele sociale dominate într-o mare măsură de partea haotică și imprevizibilă a ființelor umane, așa că din punct de vedere probabilistic criteriul lui Hume poate fi relaxat până la a spune că X este o cauză a lui Y dacă probabilitatea ca Y să se întâmple este mai mare dacă X se întâmplă decât dacă X nu se întâmplă:

$$P(Y|X) > P(Y|\!X)$$

Deși modele statistice se pot construi pentru studiul unităților micro-sociale, conceptul de cauzalitate capătă un cu totul alt înțeles atunci când ne referim la unități macro-sociale. Nu doar pentru că sunt puține la număr și nu se pot aplica analize statistice, însă stabilitatea acestor unități este diferită, transformarea lor se face cu totul altfel. Producerea unui fenomen, la un astfel de nivel, se poate realiza printr-o multitudine de combinații secvențiale care se desfășoară în strânsă legătură unele cu celelalte. În acest sens, Rueschmeyer și Stephens (1997, p.57) punctează direct: “cauzalitatea este o chestiune de secvențialitate”; ideea este într-un totu similară conceptului de lanț causal discutat de Zamfir (1990). De aceea, analiza istorică a cazurilor studiate nu trebuie trecută cu vederea în cadrul unei cercetări comparative. Acest punct de vedere are legătură și cu problema cutiei negre discutată la secțiunea 2.6.1.

## 2.2 Condiții necesare, suficiente și contrafactice

Această scurtă incursiune într-un univers causal social de înaltă complexitate a punctat dificultatea inerentă în identificarea cauzelor unui fenomen, iar mai dificilă este identificarea cauzelor *importante*, care influențează puternic fenomenul studiat. Ambele obiective au legătură

cu teoria pe care cercetătorul dorește să o testeze și există instrumente construite pentru atingerea acestora. În statistică există analizele de regresie (liniară și nonliniară, multiplă, logistică, multinomială etc.), analiza path, analiza cluster și altele, unde cauzele importante se depistează prin modelele testate, iar cele puternice sunt cauzele ale căror coeficienți sunt pe de o parte semnificativ diferiți de zero și pe de altă parte au o contribuție puternică la creșterea variației explicate ( $R^2$ ).

Mai există însă și alte metode de a testa ipoteze despre cauze și modul cum influențează acestea fenomenul de interes. În afară de metodele calitative care nu sunt acoperite în această lucrare mai există metodologia analizei calitativ-comparative, care va fi discutată în cele ce urmează.

Ambele tradiții (calitativă și cantitativă) utilizează noțiuni și ipoteze despre cauze și efecte. În timp ce în statistică acestea sunt discutate sub aspectul corelațiilor dintre variabile, analiza calitativă bazată pe studii în profunzime a înaintat conceptul de “condiție causală”: în urma studiului de caz, se constată că efectul nu se produce decât în anumite condiții (cauzale).

La fel ca în cercetarea cantitativă, studiile în profunzime arată că nu toate condițiile cauzale sunt relevante. Unele sunt mai importante decât altele, însă acest lucru este discutat într-un mod cu totul diferit. Dacă în analiza cantitativă este suficient să repetăm un experiment de un număr suficient de mare de ori (același lucru cu a eşantionă suficient de multe cazuri), în analiza calitativă acest lucru nu este posibil. Fenomenele studiate sunt atât de rare încât cauzele acestora nu pot fi testate cu ajutorul modelelor statistice ci trebuie deduse în mod logic.

De exemplu, care sunt condițiile care au determinat cel de al doilea Război Mondial? În mod cert nu putem repeta/experimenta războaiele mondiale doar pentru a le găsi condițiile cauzale, însă putem să emitem ipoteze asupra acestora într-o modalitate sistematică și logică. Dintre toate condițiile posibile, trebuie păstrate doar acelea fără de care al doilea război mondial *nu ar fi avut loc*.

Acest raționament implică mai departe în două alte linii de gândire: pe de o parte, sunt relațiile dintre ceea ce există (evidența) și modelul construit de cercetător (inferența), iar pe de altă parte, sunt relațiile dintre ceea ce există și ceea ce ar fi putut să fie dacă istoria ar fi avut un alt mers (condițiile contrafactice, de exemplu soarta Europei ar fi fost cu totul alta dacă Anglia și Franța ar fi acționat mai devreme și mai decis atunci când Germania a invadat Polonia).

Prima linie de gândire a relațiilor dintre evidență și inferență este discutată de Dion (1998), care ajunge la o definiție simplă a condițiilor

necesare și suficiente. Atunci când studiem un fenomen Y:

- Condiția cauzală X este *necesară* dacă și numai dacă pentru producerea fenomenului Y este necesar ca X să se producă. Un exemplu este analiza lui Skocpol (1979) asupra revoluțiilor, care poate fi extinsă cu ușurință și la experiența românească din 1989: “O revoluție socială se produce *numai dacă* statul este în criză”. Fenomenul de interes este revoluția socială (asupra cauzelor căreia facem inferențe), iar una dintre cauzele posibile este criza statului. Dacă statul nu ar fi în criză atunci revoluția nu ar avea loc.
- Condiția cauzală X este *suficientă* pentru producerea fenomenului Y dacă producerea lui X *implică* producerea lui Y. Un exemplu interesant de astfel de condiție poate fi: “Criza statului conduce la revoluție socială”

Diferența dintre o condiție cauzală necesară și una suficientă este că, în timp ce condiția necesară antrenează efectul dar numai în combinație cu alte condiții cauzale, cea suficientă poate antrena efectul indiferent dacă mai sunt prezente sau nu și alte condiții cauzale. Suficiența nu este neapărat singulară, adică nu este obligatoriu să existe o singură condiție cauzală suficientă; pot exista mai multe condiții cauzale suficiente, fiecare dintre acestea antrenând efectul indiferent dacă sunt prezente și alte condiții (suficiente sau nu). În prima afirmație, criza statului este o condiție necesară iar în al doilea exemplu este o condiție suficientă: deși cele două afirmații par echivalente, o analiză atentă arată diferența fundamentală dintre acestea.

Aceeași linie este prezentă și la Mahoney (2004, p.84), iar diferența dintre cele două tipuri de condiții cauzale devine mai clară: absența unei condiții necesare va antrena întotdeauna absența efectului, pe când prezența condiției suficiente va antrena întotdeauna prezența efectului. În exprimări simple, acestea sunt:

Condiție *necesară*: “Y se întâmplă numai dacă se întâmplă X”

Condiție *suficientă*: “Dacă se întâmplă X, atunci se întâmplă și Y”

La producerea unui fenomen participă un număr mare de condiții cauzale, dintre care doar unele sunt necesare. Nu toate dintre aceste condiții sunt relevante, așa încât cercetătorul trebuie să elimine condițiile cauzale triviale chiar dacă acestea sunt necesare. De exemplu, existența aerului și a apei este o condiție necesară pentru declanșarea unei revoluții. În mod cert fără aer nu se poate face nimic însă existența aerului ori a apei însă această condiție cauzală (chiar dacă necesară) este irelevantă



pentru producerea revoluției din 1989 și nu vor fi luate în considerare în cercetarea socială. Atunci când sunt punctate toate condițiile considerate necesare pentru producerea unui fenomen, este necesară o inspecție atentă a acestora pentru identificarea și eliminarea celor triviale.

Greutatea cea mai mare constă în depistarea acestor condiții. Dion (1998) arată că o condiție trivială este una care apare întotdeauna, indiferent dacă efectul se produce sau nu. Din acest punct de vedere, o condiție poate fi declarată ca necesară dacă are o relație *probabilă* în raport cu efectul (dacă se produce cu o anumită probabilitate, cu alte cuvinte dacă are o anumită variație).

Variația cauzei necesare este una dintre cele mai simple modalități de a depista o cauză necesară, sau invers constanța unei cauze este o bună modalitate de a elimina acea cauză pe motive de trivialitate. Dion mai arată că nici măcar variația nu poate spune cu certitudine absolută dacă o cauză este necesară sau nu; în unele cazuri, poate exista variație și totuși cauza să fie trivială: este extrem de probabil ca un scrutin electoral să fie televizat (deși pot exista cazuri în care să nu fie televizat), însă televizarea evenimentului nu este o condiție necesară a producerii scrutinului.

Condițiile suficiente nu pot fi triviale întrucât acestea conduc întotdeauna la producerea efectului; existența aerului nu este suficientă pentru producerea unui fenomen social. Chiar și în rândul acestor tipuri de condiții cauzale pot exista unele care să fie eliminate. În opinia lui Mahoney (2004), unele condiții suficiente sunt tautologice. Acestea sunt un set de factori prezenți chiar în definiția efectului; de exemplu, un PIB de 100.000 euro este o condiție suficientă pentru a avea o economie dezvoltată, însă creșterea PIB-ului este determinată chiar de dezvoltarea economiei, astfel că această condiție este tautologică.

Un alt mod de a conceptualiza condițiile necesare și suficiente este unul care utilizează teoria mulțimilor. Dacă vom considera mulțimea tuturor evenimentelor în care este prezentă condiția cauzală X, atunci aceasta este o condiție necesară pentru fenomenul de interes Y dacă X este prezentă întotdeauna atunci când apare Y. Mahoney și Goertz (2006, p.9) utilizează acest raționament și arată că:

- O condiție necesară X este o supramulțime în care este inclusă mulțimea cazurilor când se întâmplă Y (altfel spus, ultima fiind o submulțime a lui X). Fiind o supramulțime, X se întâmplă în absolut toate cazurile în care se produce Y; este adevărat că X (ca o mulțime mai întinsă) este prezentă și în situații când Y nu se produce, însă pentru producerea lui Y este într-adevăr necesară.

- O condiție suficientă X este o submulțime a unei mulțimi mai mari a cazurilor când se produce fenomenul de interes Y (toate situațiile în care este prezentă condiția cauzală X sunt conținute în mulțimea mai largă a cazurilor când se produce fenomenul de interes Y). Desigur, producerea lui Y este determinată de o mulțime de alți factori însă dacă Y se produce de fiecare dată când este prezent X, atunci condiția cauzală X este suficientă (ori poate fi una dintre condițiile suficiente) pentru producerea lui Y.

Schneider și Wagemann (2012) au o abordare similară pentru a descriere logica elementară a necesității și suficienței. O condiție X este necesară pentru un fenomen Y dacă fenomenul nu se poate întâmpla fără condiția X. Atunci când fenomenul Y se întâmplă, este imposibil ca X să lipsească.

Logica suficienței este descrisă în oglindă: o condiție cauzală X poate fi considerată suficientă pentru fenomenul Y dacă, atunci când condiția X se întâmplă prezentă, fenomenul Y se întâmplă și acesta.

Identificarea condițiilor necesare este un proces care pornește de la fenomenul studiat și caută condițiile cauzale, în timp ce pentru identificarea condițiilor suficiente se pornește de la condiția cauzală pentru a verifica dacă produce de fiecare dată fenomenul studiat.

În limbaj formal, cele două se notează astfel:

$$\begin{aligned} X \text{ este o condiție necesară: } & X \Leftarrow Y \\ X \text{ este o condiție suficientă: } & X \Rightarrow Y \end{aligned}$$

A doua linie de gândire (a condițiilor cauzale contrafactice) este puțin diferită însă nu mai puțin interesantă. Contrafactualii sunt condiții cauzale care apar în urma inversării conceptuale a stării de fapt (a condițiilor *factice*, a celor care s-au întâmplat), de aceea ei se referă în special la trecut. În cercetarea calitativă experimentele bazate pe condiții contrafactice sunt mai mult teoretice și pot părea speculative însă au un rol important în testarea teoriilor.

Puțini știu că acest tip de analiză este nu numai posibil, ci chiar foarte utilizat în cercetarea cantitativă. Analiza de regresie liniară oferă această posibilitate, una dintre funcțiile principale ale acesteia fiind aceea de predicție. Având valorile coeficienților de regresie și a constantei, cercetătorul poate să experimenteze obținând diferite valori ale variabilei dependente manipulând valorile variabilelor independente. De exemplu, dacă variabila dependentă este mortalitatea infantilă iar una dintre independente este un indice de dezvoltare a sistemului sanitar al localității/regiunii, atunci se poate vedea diferența dintre

mortalitatea infantilă la o dezvoltare scăzută și una pentru o dezvoltare ridicată. Strategia este identică cu obiectivul cercetării calitative, fiind echivalentul întrebării: “Ce s-ar întâmpla cu mortalitatea infantilă dacă am crește dezvoltarea sistemului sanitar?”

Printre primele utilizări ale condițiilor contrafactice au fost cele realizate de Weber (1978), prin ceea ce el numea “experimente imaginare”. Ca și Durkheim, Weber a intuit în mod corect faptul că teoriile despre societate nu pot fi testate cu ajutorul experimentelor (ca în științele naturale). De aceea, el a propus modificarea imaginată a unei condiții cauzale după care să ne întrebăm în ce sens s-ar fi modificat efectul, în acest fel obținând un model cauzal.

În afară de condițiile contrafactice, metodele de cercetare cantitative nu dispun de instrumente pentru identificarea condițiilor necesare și suficiente; mulți cercetători nici nu-și orientează cercetarea într-o astfel de manieră, iar cei care sunt conștienți de problemă tind să nu o considere relevantă întrucât, la urma urmei, aceste condiții sunt foarte greu de găsit într-un noian de condiții cauzale. Această abordare nu este lipsită de temei întrucât realitatea nu prezintă aproape niciodată o linie dreaptă de la o cauză la un efect: modelele cauzale sunt rareori atât de simple încât să implice o singură cauză pentru un fenomen complex.

Galileo (1632) a înțeles acest lucru și a definit cauzalitatea prin toate cauzele care sunt necesare și suficiente pentru producerea efectului; modelul lui nu se bazează pe o cauză separată a unui efect, ci consideră cauza ca fiind întregul set de condiții cauzale responsabile pentru producerea efectului:  $Y$  nu se produce decât dacă toate condițiile  $X_1 \dots X_n$  se produc concomitent. În viziunea lui Galileo,  $Y$  nu se produce dacă vreuna dintre condițiile cauzale  $X$  nu se produce. Chiar dacă a fost abandonată întrucât nu era falsificabilă (nu putem găsi niciodată toate cauzele unui fenomen, în special cele mai slabe), această idee punctează ceea ce în teoria modernă se conturează în asumția complexității cauzale (Ragin, 1987).

Cele mai multe cauze sunt limitate în putere și cuprindere empirică, fiind incapabile să declanșeze efectul de unele singure. Diverse condiții cauzale, sau mai mult diverse *combinații* de condiții cauzale pot să producă același efect. Dacă în analiza cantitativă nu putem găsi acest lucru, există o metodologie în cadrul analizei comparative care este capabilă să identifice aceste combinații. Următoarele secțiuni vor descrie fundamentele acestei metodologii precum și algoritmul matematic pe care se bazează.

## 2.3 Canoanele inductive ale lui Mill

Scriverile lui Mill sunt larg considerate a fi primele formulări sistematice ale metodei comparative moderne (Ragin, 1987; Rihoux, 2003; Aickin, 2004, printre alții). Pentru a identifica circumstanțele care influențează în mod invariabil un fenomen de interes, există două metode puternice (“canoane” așa cum au rămas ele încetățenite în teorie). Prima metodă compară toate instanțele în care se întâmplă fenomenul de interes, iar a doua compară instanțe în care fenomenul se întâmplă cu instanțe în care fenomenul nu se întâmplă, acestea sunt identificate de către Mill (1843, cartea III, cap.VIII) ca “metoda concordanței” și “metoda diferenței”.

**Metoda concordanței** este de departe cea mai utilizată metodă în analiza comparativă, fiind utilizată și în cadrul algoritmului de minimizare booleană prezentat în secțiunea 5.1.6. În pofida acestei utilizări răspândite, această metodă este considerată inferioară întrucât poate să conducă la false generalizări empirice (Ragin, 1987):

“Dacă două sau mai multe cazuri în care se produce fenomenul supus investigației au o singură circumstanță comună, acea unică circumstanță prin care toate cazurile concordă este cauza (sau efectul) fenomenului studiat”.

Tabelul 2.1 este o ilustrare a principiului concordanței a lui Mill și prezintă studiarea a trei instanțe diferite în care fenomenul de interes (Y) se întâmplă. În primul caz, se constată să spunem condițiile cauzale A, B, C și X; în al doilea caz se constată condițiile cauzale D, E, F și X

Tabelul 2.1: Ilustrarea metodei concordanței

	Instanțe cauzale				Efect
Caz 1	A	B	C	X	Y
Caz 2	D	E	F	X	Y
Caz n	G	H	I	X	Y
	<span style="display: inline-block; width: 100px; border-top: 1px solid black;"></span>				
	Diferențe globale				<span style="display: inline-block; width: 100px; border-top: 1px solid black;"></span>
					Similarități cruciale

Sursa: Adaptare după Skocpol și Somers (1980, p.184)

iar în al treilea caz se constată condițiile cauzale G, H, I și X. După cum se poate vedea, singura condiție cauzală care este prezentă în mod invariabil alături de Y este X. Conform metodei concordanței, avem dovezi pentru a spune că X este o cauză a lui Y.

Metoda concordanței presupune studierea cazurilor similare, obiectivul cercetătorului fiind acela de a căuta în toate circumstanțele în care se produce efectul pentru a identifica acea singură condiție cauzală care precede fenomenul în toate instanțele studiate. Se poate întâmpla ca nici o instanță să nu concorde prin vreo condiție cauzală, după cum se poate întâmpla să existe mai multe condiții cauzale prin care instanțele să concorde (în ultimul caz, fiecare dintre condițiile concordante pot fi considerate cauze). Circumstanțele care pot fi excluse fără a modifica efectul nu sunt legate de acesta prin vreo legătură cauzală.

Metoda concordanței procedează așadar prin excludere, prin eliminarea condițiilor cauzale care nu sunt prezente în toate instanțele în care este prezent efectul studiat. Chiar și în aceste condiții, avertizează Mill (1843, cap. 10.2), pot exista situații în care același fenomen poate fi produs de cauze diferite. Dacă în urma metodei concordanței găsim o cauză A în toate instanțele, aceasta nu înseamnă în mod obligatoriu că A este singura condiție cauzală relevantă a fenomenului, ci este posibil ca într-o altă instanță fenomenul să fie produs de o cauză B, în altă instanță să fie produs de o cauză C iar cauza A nu are de fapt nici o legătură cauzală cu fenomenul (deși nici B, nici C nu sunt prezente în toate instanțele).

Fenomenul este cunoscut sub aspectul relațiilor cauzale aparente: nici una dintre condițiile cauzale observate nu este o adevărată cauză ci o alta, ascunsă. Această situație este dublată de posibilitatea cauzalității multiple care face metoda concordanței aproape imposibil de aplicat. Singura soluție pe care o vedea Mill era încercarea de reconstruire a situației experimentale prin stimuli controlați (ceea ce în domeniul social este desigur imposibil, în cele mai multe situații).

**Metoda diferenței.** O altă soluție propusă de Mill la această problemă a fost formularea metodei diferenței, mai puternică decât prima întrucât nu numai că ia în calcul mai multe instanțe dar oferă posibilitatea experimentării indirecte:

“Dacă o instanță în care fenomenul investigat se produce și o altă instanță în care fenomenul nu se produce, au în comun toate circumstanțele mai puțin una care se întâmplă doar în prima instanță; circumstanța în care cele două instanțe sunt diferite este efectul, sau cauza, sau o parte indispensabilă a cauzei fenomenului”.

Tabelul 2.2: Ilustrarea metodei diferenței

	Instanțe cauzale			Efect
Caz 1	A	B	C	$\sim X$   $\sim Y$
Caz 2	A	B	C	$\sim X$   $\sim Y$
Caz n	A	B	C	X   Y

$\underbrace{\hspace{10em}}$ Similarități globale	$\underbrace{\hspace{10em}}$ Diferențe cruciale
---	---

Sursa: Adaptare după Skocpol și Somers (1980, p.184)

Cu alte cuvinte, condiția cauzală care produce efectul este cea care diferă odată cu efectul, în timp ce toate celelalte rămân similare.

Tabelul 2.2 este o ilustrare a metodei diferenței; spre deosebire de tabelul 2.1, aici este prezentată o situație inversă în care unele condiții cauzale sunt prezente atât atunci când fenomenul se produce cât și atunci când fenomenul nu se produce iar unele se conformează tiparului de manifestare al fenomenului. Sunt studiate tot trei instanțe, primele două în care fenomenul nu se produce, este absent ( $\sim Y$ ; semnul ‘ $\sim$ ’ este notația pentru cazurile negative, “not Y” în logică și algebra booleană) iar ultima în care fenomenul se produce (Y). Singura condiție cauzală care are un pattern identic cu al manifestărilor fenomenului este X, absentă în primele două instanțe și prezentă în ultima.

Instanțele în care Y este prezent (se produce fenomenul de interes) sunt numite cazuri pozitive, iar instanțele în care Y este absent (nu se produce) sunt numite cazuri negative. Logica celor două metode este similară însă deosebirea dintre acestea este că ultima face apel atât la cazurile pozitive cât și la cele negative pentru a găsi cauzele unui fenomen. Din acest motiv, metoda diferenței a mai fost denumită și o *dublă metodă a concordanței*.

Condițiile A, B și C nu pot fi cauze ale fenomenului Y întrucât acestea sunt prezente și în instanțe în care fenomenul este absent. Dacă am fi studiat doar primele două cazuri, toate condițiile ar fi fost considerate cauze ale lui Y, însă studierea celui de al treilea caz elimină toate condițiile cauzale în afară de una singură. Ca o practică generală, se împart cazurile în instanțe de prezență și absență a fenomenului de interes, iar condițiile cauzale care sunt legate cauzal de fenomen trebuie să fie invariabil prezente în primul grup de instanțe și invariabil absente în cel de al doilea. Este așadar o

aplicare dublă a metodei concordanței, odată pentru instanțele pozitive și apoi pentru cele negative, în acest fel fiind filtrate doar condițiile care sunt într-adevăr legate cauzal de fenomenul de interes.

Cazurile concordă atât în instanțele pozitive cât și în cele negative ale fenomenului: o cauză adevărată este prezentă în toate instanțele în care este prezent fenomenul de interes și este absentă în toate instanțele în care fenomenul de interes nu se produce. La fel ca la metoda concordanței, în metoda diferenței se procedează tot prin eliminarea cauzelor concurente însă la diferență se elimină acele cauze care nu urmează același tipar cu al fenomenului de interes. Desigur, dacă nu avem la dispoziție decât cazuri pozitive, atunci metoda diferenței nu se poate aplica.

Ragin (1987) observă în mod corect că un cercetător cantitativist ar putea face o paralelă simplă între această metodă și coeficientul de corelație. Dacă există concordanță atât în instanțele pozitive cât și în instanțele negative, atunci metoda diferenței arată de fapt un coeficient de corelație perfect egal cu 1. Ragin arată însă că există o deosebire fundamentală: în timp ce coeficientul de corelație se aplică în statistică și se bazează de fapt pe relații probabilistice (nu atinge niciodată valoarea 1), metoda diferenței se bazează pe comparații ale unor instanțe pe bază de invarianță.

În interpretarea originală a lui Mill, deosebirea dintre cele două metode este faptul că metoda concordanței are la bază supoziția că toate cauzele care pot fi eliminate nu au nici o legătură cu fenomenul, în timp ce metoda diferenței are la bază o supoziție opusă, aceea că toate cauzele care nu pot fi eliminate sunt legate de fenomenul studiat.

Deși metoda diferenței este preferabilă metodei concordanței, are aceeași problemă în sensul că poate să ne inducă în eroare atunci când există cauzalitate multiplă. Studiind răscoalele țărănești, Ragin (1987) identifică patru cauze posibile: foametea, comercializarea rapidă a produselor agricole, o clasă țărănească puternică și tradiționalismul. Interesant este faptul că Ragin a fost preocupat de problema răscoalelor țărănești cu mult timp înainte, dedicând un articol întreg răscoalei țăranilor români din 1907, scris împreună cu Chirot (1975). Există cazuri în care răscoala este produsă numai de foamete iar altele în care este produsă doar de comercializarea rapidă a produselor. În această situație spunem că există cauzalitate multiplă, care nu poate să fie identificată nici de metoda concordanței nici de cea a diferenței.

În plus, metoda diferenței utilizează instanțe negative ale fenomenului, iar acestea sunt practic nelimitate (instanțele pozitive sunt mai rare, cele negative pot fi găsite peste tot). A cros-valida rezultatele obținute prin metoda concordanței cu ajutorul metodei diferenței înseamnă o selecție corectă a cazurilor negative. Slaba selecție a acestora poate să

compromită un model cauzal altminteri corect.

Pentru a controla deficiențele inerente ale acestor metode, Mill a oferit mai multe alte metode: cea indirectă a diferenței, ori a reziduurilor, ori o alta care a precedat coeficientul de corelație, anume metoda variațiilor concomitente:

“Un fenomen care variază într-un fel oarecare ori de câte ori un alt fenomen variază într-o manieră particulară, este fie o cauză fie un efect al acelu fenomen, fie este legat de acesta printr-un fapt de cauzare”.

Titlurile secțiunilor 2, 3 și 4 ale capitolului VII din cartea lui Mill *A system of logic* se numesc cât se poate de sugestiv:

2. In the Social Science experiments are impossible
3. — the Method of Difference inapplicable
4. — and the Method of Agreement, and of Concomitant Variations, inconclusive.

Niciuna dintre metodele avansate de Mill nu a rezistat propriilor sale critici; metoda variațiilor concomitente este neconcludentă pe aceleași rațiuni ca și metoda concordanței, din cauza pluralității cauzale. Ultima însă este mai rezistentă întrucât are calitatea (pe care celelalte nu o au) de a nu se prăbuși la primul caz negativ care apare ori care este indicat. În general, primele metode sunt sensibile la acțiunea exemplurilor negative: orice caz care contrazice teoria o anulează. Dar, în special în științele sociale, un singur caz negativ nu ar trebui să conteze la fel de mult ca 99 de cazuri pozitive. Un alt motiv pentru care Mill își pune metoda sub semnul întrebării este faptul că este practic imposibil în științele sociale să găsim cazuri suficient de similare ori suficient de diferite.

Durkheim (2002, p.175) are o opinie similară cu cea a lui Mill, însă ridică metoda variațiilor concomitente la un nivel înalt în sociologie, interpretând-o undeva între metoda statistică și cea comparativă, ajungând până la a o considera “instrumentul prin excelență al cercetării sociologice”.

Indiferent de obiecții, concluzia care trebuie trasă este faptul că atunci când căutăm cazuri comparabile, cercetarea trebuie să respecte cât se poate de mult standardele acestor metode (Lijphart, 1971).

Canoanele inductive ale lui Mill reprezintă doar un punct de plecare pentru metodologia cercetării comparative din prezent. Luate separat, acestea nu pot să răspundă unor întrebări și probleme fundamentale,



însă nici nu se poate imagina metodologia actuală fără ele. În secțiunea următoare voi arăta cum se poate extinde inducția lui Mill cu deducția și abducția, pentru a forma o imagine mai clară asupra unui proces de cercetare modern.

## 2.4 Inducție, deducție, abducție

În știință în general și în științele sociale în particular, inducția a jucat un rol central. În această secțiune prezint diferențele dintre inducție, deducție și abducție, cu accent pe prezentarea orientărilor divergente dintre inducția analitică și deducție, utilizate în științele sociale.

**Inducția** se referă la un proces de gândire prin care se încearcă explicarea unui fenomen prin intermediul experiențelor trecute. Cu cât numărul de instanțe concordante (și direct observabile) crește, avem cu atât mai multe motive să credem că fenomenul se va produce din nou în condițiile observate.

În știință, inducția este metoda cea mai des utilizată pentru avansarea teoriilor întrucât cunoașterea este cumulativă, se bazează pe toată cunoașterea anterioară. Dacă teoriile noi nu s-ar baza pe succesul ori eșecul vechilor teorii, ar trebuie să reinventăm totul în mod ciclic. Inducția presupune așadar un proces de acumulare a experiențelor, în așa fel încât știința din trecut devine fundamentul pe care se pot clădi noi teorii în viitor.

În filozofie și logică, inducția este un proces contestat. Începând de la Hume (1748), inducția este desființată ca proces științific, deși este încă utilizată în mod extensiv. În esență, Hume a contestat faptul că prin acumularea experiențelor am putea vreodată să facem predicții asupra viitorului. Această abordare a lui Hume nu a fost rezolvată în totalitate, și continuă să aibă o influență asupra gândirii filozofice actuale. Faptul că de milioane de ani soarele a răsărit în fiecare dimineață nu este o dovadă suficient de puternică pentru a putea prezice că va răsări și mâine (desigur, noi putem “prezice”, însă aceasta nu înseamnă că se va și întâmpla).

Problema inducției este într-un fel similară cu cea a cauzei și efectului; faptul că observăm de milioane de ani un fenomen cauză (pământul se învâрте în jurul soarelui), prin inferență am putea să spunem că în viitor se va produce în mod inevitabil și fenomenul efect (soarele răsare în fiecare dimineață), însă acumularea experiențelor în acest fel, oricât de multe ar fi acestea, nu pot înlătura un alt fenomen C (neșteptat) să se producă în viitor; soarele se va răci la un moment dat, deci nu va mai răsări. Pentru a răspunde la această problemă, Hume a pus bazele sistemului

deductiv-nomologic.

**Deduția**, pe de altă parte, se referă la un proces de gândire opus: în loc să se prezică un efect pe baza evidențelor, se utilizează efectul ca bază de plecare și se deduc cauzele acestuia. Deduția analizează cauze posibile și le selectează doar pe cele care duc la producerea efectului.

Deși este preferata logicienilor (va oferi întotdeauna răspunsuri adevărate și valide), deducția nu poate înlocui inducția în avansarea teoriilor. Hume însuși spune că o persoană care respinge complet inducția în favoarea exclusivă a deducției va muri de foame, întrucât a nu avea încredere niciodată în ceea ce se va întâmpla pe baza experiențelor trecute duce spre blocaj acțional. În schimb, Hume propune o inducție “septică”.

**Abducția**, un concept aproape necunoscut în științele sociale, este situat undeva între inducție și deducție. Încercând să preia aspectele tari din ambele, abducția se referă la un proces de gândire prin care se alege cea ipotetă care, dacă ar fi adevărată, ar explica cel mai bine un fenomen observat. Este, cu alte cuvinte “o inferență prin cea mai bună explicație posibilă” (Harman, 1965).

În secolul XX, un alt mare filozof al științei s-a preocupat de problema inducției a lui Hume și a ajuns la o concluzie similară numai că soluția propusă este diferită. Popper (2000) demonstrează logic de ce știința nu poate fi (n.b. doar) inductivă, deoarece pentru a cunoaște cauzele unui fenomen ar trebui să cunoaștem și cauzele acelor cauze, mergând mai departe până la infinit. Din punct de vedere logic, argumentația lui Popper este adevărată, însă apelând la conceptul de *infinit* argumentația lui seamănă oarecum cu celebrul paradox cu Ahile și broasca țestoasă: Ahile (un mare alergător) nu va depăși niciodată broasca țestoasă care se află cu un metru în fața lui, întrucât pentru a o ajunge din urmă va trebui să treacă mai întâi de prima jumătate; pentru a parcurge cea primă jumătate va trebui să parcurgă iarăși în avans o jumătate din cea jumătate, raționamentul mergând până la infinit. Din punct de vedere strict logic, Ahile nu va reuși niciodată să traverseze infinitul pentru a depăși broasca, însă în realitate o depășește fără probleme.

În mod similar, așa cum logica nu poate să găsească o soluție acceptabilă pentru paradoxul lui Ahile, cred că argumentul logic al lui Popper nu este complet aplicabil în mersul științei. Sigur, este clar că știința nu se poate limita doar la un proces strict inductiv; simpla adunare de date și experiențe nu garantează generarea de noi teorii, însă ajută la testarea ipotezelor particulare care compun teoriile.

Popper are dreptate când susține că știința este și trebuie să fie în primul rând deductivă; la urma urmei, marile teorii (din orice domeniu) nu au fost generate în urma unui proces inductiv, prin acumulare de experiență directă, ci printr-un proces de gândire, de multe ori logic însă de cele mai multe ori teoriile sunt rezultatul unor sclipiri de geniu. Nu trebuie ignorate însă nici soluțiile practice găsite de-a lungul timpului de oameni care au observat pur și simplu că fenomene similare se întâmplă în situații similare (un raționament pur inductiv).

Soluția lui Popper la problema inducției a lui Hume a fost o răsturnare de perspectivă: în loc să demonstreze faptul că soarele va răsări mâine, el demonstrează doar credința că soarele va răsări, sub forma unei ipoteze falsificabile. Ipoteza este că soarele va răsări; dacă mâine acest lucru nu se va întâmpla, atunci această ipoteză va fi infirmată și o alta va trebui să-i ia locul. Aceasta, arată Popper, este calea pe care știința trebuie să o aplice: concurența teoriilor, dintre care numai una joacă rolul de paradigmă dominantă (aceea care trece de cele mai multe probe falsificabile).

Propunerea lui Popper a rămas în istoria filozofiei sub denumirea de model ipotetico-deductiv: știința nu trebuie să caute probe care să confirme o teorie, ci invers, probe care să o infirme. Cu fiecare astfel de probă infirmativă demonstrată, teoria își va pierde încet din putere și va fi în cele din urmă abandonată, lăsând locul uneia și mai puternice.

În științele sociale, spre deosebire de logică și filozofia științei, inducția ocupă un loc foarte important mai ales după ce Znaniecki a introdus în anul 1934 termenul de “inducție analitică” (Taylor și Bogdan, 1998). De orientare antipozitivistă, Znaniecki considera că sociologia trebuie să studieze fenomene sociale din perspectivă culturalistă. Individul este perceput nu ca un simplu receptor al mesajelor din domeniul social, ci purtător de sens și mai ales de creativitate. Znaniecki a intrat în istoria sociologiei cu celebra sa lucrare “Țăranul polonez în Europa și America”, analizând jurnale ori autobiografiile ale imigranților polonezi în America, precum și scrisorile dintre aceștia și familiile lor rămase în Europa.

Abordarea sa empiricistă seamănă foarte mult cu un proces inductiv: se studiază foarte multe instanțe ale unor fenomene, însă spre deosebire de procesul logic ori filozofic al inducției, meritul imens al lui Znaniecki este acela de a fi arătat prin lucrarea o modalitate de cercetare socială prin care se caută pattern-uri de similaritate în toate instanțele analizate.

Spre deosebire de științele naturale, unde toate instanțele analizate seamănă (soarele răsare la fel, frunzele cad la fel etc.), în sociologie fiecare actor social este unic. Mai mult, fiecare fenomen social studiat are particularități unice, fapt care face studiul acestora un proces extrem

de anevoios. Este nevoie de foarte multă experiență și intuiție pentru a depista comunalitățile dintre toate instanțele analizate. Sigur, știința modernă se bazează pe analiza statistică a fenomenelor, însă atunci când se studiază doar câteva cazuri metodele statistice nu pot fi aplicate.

Chiar și în cazul în care există foarte multe cazuri studiate (cum au fost documentele lui Znaniecki), acestea nu puteau fi analizate cantitativ, ci doar calitativ utilizându-se un proces sistematic inductiv. Nu se știe dacă Znaniecki a început să studieze scrisorile respective având deja format conceptul de culturalism ori a format acest concept pe măsură ce a citit scrisorile. Înclin să cred că este vorba despre a doua variantă, începând cercetarea cu intuiția descoperirii unui nou concept și a unei noi metodologii.

Ragin (1994) lămurește foarte elegant relația dintre procesul de inducție analitică și cadrele analitice construite în urma unei analize calitative. În fapt, inducția analitică nu este un proces inductiv propriu-zis, întrucât nu se pornește de la o simplă acumulare a datelor de la care se stabilesc ipotezele, ci se analizează datele cu o anumită reprezentare conceptuală, apoi în funcție de ceea ce spun datele reprezentarea se poate modifica, transforma, până când conceptul rezultat acoperă foarte bine datele calitative analizate. Cercetătorul nu studiază cazuri pur și simplu, ci poartă un veritabil dialog cu datele sale; cazurile sunt reanalizate în bucle iterative până când se stabilește o imagine clară asupra detaliilor fenomenului studiat, în fiecare iterație utilizându-se atât raționamente bazate pe experiență cât și veritabile deducții.

Reprezentarea conceptuală rezultată în urma dialogului cu datele formează așa numitul “cadru analitic”. Acesta este corespondentul modelului utilizat în statistică, cu diferența majoră că în cercetarea cantitativă cadrul analitic este fix (fiind utilizat doar la respingerea ipotezei de nul), în timp ce în cercetarea calitativ-comparativă cadrul analitic este flexibil.

Hicks (1994) arată că inducția analitică nu este nimic altceva decât o aplicație a canonului concordanței al lui Mill. Citându-l pe Robinson, el evidențiază faptul că inducția analitică este în mod inerent limitată întrucât poate specifica doar condițiile necesare, nu și pe cele suficiente (deoarece se iau în calcul doar instanțele în care evenimentul se întâmplă, nu și cele în care nu se întâmplă). Aceeași concluzie o subliniază și Ebbinghaus (2005, p.143): “...este adevărat că logica lui Mill bazată pe canoane inductive nu poate fi utilizată pentru a face inferențe asupra relațiilor cauzale în afara simplelor condiții necesare”.

De aceea, Hicks argumentează faptul că inducția analitică s-a dezvoltat în așa-numita “inducție neo-analitică”, prin analizarea instanțelor atât

pozitive cât și negative. Ciclul dialogului cu datele se reia (ajustându-se atât condițiile cauzale cât și explicația teoretică) până când cadrul analitic dobândește o trăsătură universală iar teoria devine atât de rafinată încât explică în mod corect fenomenul studiat.

Inducția (neo)analitică are, printre altele, rolul de a fi utilă în generarea de noi teorii prin formularea de cadre analitice universale. În acest context, termenul “universal” nu trebuie confundat cu legile universale din știință ori cu generalizarea rezultatelor de la eșantion la populație (ca în abordarea cantitativă), ci mai degrabă ar trebui să fie interpretat ca “universal valabile pentru toate cazurile studiate”. Până la apariția analizei calitativ-comparative, inducția neoanalitică era cea mai sistematică metodă de contruire a teoriilor, prin dezvoltarea progresivă a unui set inițial de ipoteze.

Ebbinghaus (2005, p.143) este oarecum sceptic cu privire la potențialul analizei comparative, arătând că “... metoda comparativă este mai degrabă utilizată în scopuri explorative, în ciuda imposibilității epistemologice a aplicării inducției și a limitelor din lumea reală care stau în calea testării deductive a ipotezelor”. O astfel de aventură explorativă poate cel mult să genereze o teorie, însă este nevoie de testarea riguroasă a acesteia în raport cu alte teorii concurente.

Deși mulți cercetători au afirmat răspicat că cercetările comparative se realizează exclusiv prin metoda inducției (utilizând canoanele lui Mill), Goldstone (1997) demonstrează fără echivoc faptul că analiza comparativă nu se poate baza pe metodele de tip Mill și nu se limitează la o natură pur inductivă. Mai bine spus, că nu are o natură *doar* inductivă. Cum anume ajung cercetătorii la descoperiri teoretice și la generarea de ipoteze rămâne un mister, însă procesul poate fi urmărit retrospectiv urmărind logica unor lucrări consacrate. Goldstone arată că procesul respectiv implică o combinație de raționamente atât inductive cât și deductive.

În examinarea comparativă a mai multor cazuri, sarcina cercetătorului este aceea de a examina cu foarte mare atenție, în detaliu, fiecare caz în parte pentru a identifica pattern-uri de similaritate. Acest proces este într-adevăr inductiv, însă formularea teoriei nu se poate opri aici; cazurile pozitive trebuie analizate comparativ cu cazuri negative, atât în ceea ce privește absența efectului cât și cu alte cazuri în care condițiile cauzale din ipoteză lipsesc. Analiza circumstanțelor particulare care duc la producerea unui efect este o altă etapă a analizei, una care implică un raționament deductiv întocmai ca munca unui detectiv care investighează o crimă ori un accident.

Zamfir (1990) are o abordare similară, afirmând că explicația sociologică nu se limitează la enunțarea listei de cauze și nici la observarea inductivă

a pattern-urilor de similaritate empirică. Dimpotrivă, întrucât realitatea socială este complexă, structura explicației sociologice este de asemenea un proces în care abilitatea cercetătorului de a face deducții logice este esențială. Dovezile empirice inductive nu reprezintă soluția explicației, ci doar niște informații pe care cercetătorul le potrivește, ca într-un joc de puzzle, până când reușește să avanseze (prin deducție logică) o explicație a producerii fenomenului de interes.

## 2.5 Similaritate și diferență sistemică

În metodologia cercetării comparative pot fi găsite și alte strategii; două dintre acestea au o logică similară canoanelor inductive ale lui Mill, fiind identificate și utilizate de către Przeworski și Teune (1970):

- **MSSD** – Most Similar System Design – studierea sistemelor similare: se selectează cazuri cât mai similare unele cu celelalte, plecând de la asumția că similaritatea crescută crește șansele identificării factorilor responsabili pentru diferențele dintre acestea
- **MDS** – Most Dissimilar System Design – studierea sistemelor diferite: se studiază cazuri cât mai diferite cu puțință pentru a demonstra că relația dintre cauză și efect este suficient de puternică pentru a rezista în cele mai diverse condiții.

Studierea sistemelor similare se aseamănă întrucâtva cu controlarea variabilelor (atât control experimental cât și control statistic): ținând toate variabilele constante, sunt depistate condițiile care fac diferența. În metoda comparativă constanța se manifestă sub forma similarității cazurilor, prin experiment se verifică dacă un stimul este responsabil pentru apariția efectului, iar în statistică se verifică dacă influența independentei este sau nu semnificativă. În cadrul experimentelor, se compară două grupuri: cel experimental care este expus unui stimul și cel de control care nu este expus stimulului (iată așadar că și metoda experimentului recurge de fapt tot la comparație).

Experimentul este cea mai înaltă formă a studierii sistemelor similare: cele două grupuri sunt perfect asemănătoare cu excepția unei singure variabile care este declarată ca fiind responsabilă pentru declanșarea efectului. Dacă în științele exacte experimentele sunt perfect posibile, în științele care studiază societăți umane posibilitatea experimentării este extrem de redusă. Pe de o parte, nu cunoaștem toate efectele intervenției într-un sistem social complex, iar pe de altă parte există

restricții etice și deontologice. Soluția la această dilemă constă în selectarea aleatoare a elementelor studiate în cele două grupuri, care este posibilă doar în cazul analizelor statistice nu și în analizele comparative întrucât numărul cazurilor studiate este prea mic pentru o pretenție de selecție aleatoare (cu atât mai mult pentru calcularea unor corelații semnificative).

Combinat cu metodele concordanței și diferenței ale lui Mill, reies patru strategii extrem de interesante, identificate de Faure (1994) și adaptate în mică măsură în această lucrare:

	Metoda concordanței	Metoda diferenței
MSSD studierea sistemelor similare	<b>MSSO</b> ( <i>Most Similar, Same Outcome</i> ) – condiții cauzale similare, același efect. Se studiază instanțele aflate la intersecția dintre asemănarea condițiilor cauzale și prezența uniformă a fenomenului	<b>MSDO</b> ( <i>Most Similar, Different Outcome</i> ) – condiții cauzale similare, efect diferit. Se studiază un grup de cazuri în care toate condițiile cauzale sunt asemănătoare mai puțin apariția fenomenului de interes
MDSD studierea sistemelor diferite	<b>MDSO</b> ( <i>Most Disimilar, Same Outcome</i> ) – condiții cauzale diferite, același efect. Se studiază un grup de cazuri eterogen din punctul de vedere al condițiilor cauzale, dar care manifestă împreună apariția fenomenului de interes	<b>MDDO</b> ( <i>Most Disimilar, Different Outcome</i> ) – condiții cauzale diferite, efect diferit. Se studiază cazuri în care condițiile inițiale sunt diferite iar efectul este de asemenea diferit

Exemplele oferite de Faure pentru fiecare celulă sunt următoarele:

**MSSO** Studiul lui Barrington Moore (*Social Origins of Dictatorship and Democracy*) asupra drumurilor către societăților moderne, în care a studiat societăți similare care au avut un parcurs asemănător, ajungând la un sistem democratic

**MSDO** Același studiu a lui Barrington Moore comparând de această dată societăți similare care au avut unele dintre ele un parcurs democratic

iar altele un parcurs nedemocratic

**MDSO** Pentru rezultate asemănătoare în sisteme diferite este citat studiul lui Skocpol asupra revoluțiilor (*States and Social Revolutions*): societăți extrem de diferite (Franța, Rusia și China) au experimentat toate același fenomen al revoluției sociale.

Aș adăuga un alt exemplu celebru oferit de Lipset (1994), al comparației dintre societatea americană și cea japoneză care prezintă ambele o stare de puternică dezvoltare economică însă o diferență culturală mai mult decât clară: America prezintă valori puternic individualiste în timp ce Japonia prezintă valori grupale puternice. Acest studiu a fost precedat de o altă cercetare foarte îndrăzneță la vremea ei, care lega stratificarea socială de valori (prin intermediul condițiilor de muncă), într-o serie de studii comparative între Statele Unite și Polonia (Slomczynski et al., 1981, comparând o societate capitalistă cu una socialistă) și între Statele Unite și Japonia (Naoi și Schooler, 1985, comparând o societate din vest cu una din est, pentru confirmarea comparației dintre Statele Unite și Polonia).

**MDDO** Pentru rezultate diferite în societăți diferite exemplul citat este studiul lui de Tocqueville (*Democrația în America*) prin care se compară societatea americană cu cea franceză din punctul de vedere al parcursului spre democrație.

Canoanele lui Mill și cele două strategii analitice se împletesc la controlarea situației experimentale: obținerea unui set relevant de cazuri similare ori diferite presupune o strategie corectă de selecție a cazurilor. De cele mai multe ori cercetarea calitativă și comparativă se confruntă cu un număr foarte restrâns de cazuri, însă există și situații în care, pentru testarea unei ipoteze, trebuie selectate cazurile cele mai substanțiale.

Atât MSSD cât și MDSD au avantaje evidente. De exemplu, (Coppedge, 1999, p.472) este de părere că studierea cazurilor longitudinale este în mod clar cel mai bun design de cercetare pentru testarea ipotezelor despre cauzele unui eveniment, fiind “forma perfectă” de MSSD. El se bazează pe ideea că societatea la momentul  $t$  este aproape identică cu societatea la momentul  $t + 1$ , singura diferență dintre acestea fiind cauza care a generat fenomenul cercetat. Toate celelalte condiții cauzale dintre cele două momente temporale nu se schimbă deci pot fi considerate constante (respectând în acest fel cerințele dure ale controlului experimental direct). Orice caz este mai similar cu el însuși la două momente temporale decât poate fi vreodată cu oricare alt caz.



Nici MDSD nu a fost ocolită de cercetări celebre; Skocpol (1979) a realizat (inițial în cadrul tezei sale de doctorat) o foarte interesantă comparație a trei societăți diferite care au manifestat revoluții sociale: revoluția franceză dintre 1787-1800, revoluția rusă dintre 1917-1921 și revoluția chineză dintre 1911-1949. În pofida numeroaselor critici ulterioare, analiza lui Skocpol este și rămâne un exemplu clasic de cercetare comparativă utilizând un număr extrem de redus de cazuri. Cu doar trei societăți o analiză statistică este imposibil de realizat, singurele instrumente rămase disponibile pentru analiză fiind cele calitativ-comparative. Prin intersecția unor cadre analitice bine gândite, Skocpol a reușit să propună un model care, în liniile sale principale, este foarte solid.

Analiza ei este mai mult decât un exemplu general de analiză comparativă; în mod specific, a utilizat analiza comparativă istorică pentru a testa ipoteze macrosociale într-o manieră sistematică. Toate cele trei revoluții sociale analizate sunt niște instanțe comparabile ale unui singur pattern; într-o schemă foarte simplistă, criza vechiului regim dublată de efervescența mișcărilor țărănești au generat revoluțiile sociale care s-au concretizat apoi în noi regimuri stabile: de la vechiul regim al Bourbonilor la regimul lui Napoleon în Franța, de la vechiul regim țarist la noul sistem stalinist în Rusia și de la vechiul regim imperial la noua ordine Sino-comunistă în China.

Ceea ce face analiza lui Skocpol nu doar interesantă ci aproape obligatorie pentru orice student în științele sociale, este tratarea comparativă a unor evenimente atât de importante încât au marcat nu doar societățile în care au apărut ci întreaga umanitate. Este dificil de imaginat cum ar fi arătat lumea fără revoluția franceză, ori cum ar fi arătat Europa în absența regimului comunist instaurat vreme de decenii, după revoluția rusă. Ori ce sistem s-ar fi stabilit în China dacă în Rusia nu ar fi câștigat Stalin, pentru că (în opinia mea) nu e întâmplător faptul că sfârșitul revoluției chineze din 1949 coincide cu succesul sovietic, atât teritorial cât și ideologic, după cel de-al doilea Război Mondial.

Analizate individual, fiecare dintre cele trei revoluții au aspecte și cauze particulare (cele trei culturi sunt indubitabil diferite din toate punctele de vedere). Pot fi identificate însă și aspecte comune: fiecare caz prezintă o societate predominant agrară, dominată despotic de regimuri absolutiste, și toate sunt regimuri istorice de lungă durată (spre deosebire deci de revoluțiile coloniale) cu importanță deosebită în arena politică internațională. Unele aspecte teoretice sunt compatibile cu teoriile lui Moore (1966) care a evidențiat importanța istorică și influența pe care au avut-o conflictele agrare asupra societăților moderne.

Într-o descriere ulterioară a analizei, Skocpol și Somers (1980) arată că există nu doar o singură metodă generică de comparație istorică, ci nu mai puțin de trei instrumente (tipuri) diferite: *analiza macro-cauzală* (comparațiile au rolul principal de a genera inferențe cauzale despre macrostructuri și macroprocesse), *demonstrația paralelă a istoriei* (sunt formulate teorii despre anumite evenimente și sunt prezentate în paralel evenimente istorice care susțin teoriile) și *contrastele contextuale* (unde nu se încearcă nici testarea unor teorii nici formularea de generalizări, ci sunt evidențiate aspectele divergente ale cazurilor studiate cu scopul invers de a limita inferența cauzală fără cunoașterea amănunțită a contextelor particulare).

Motivul pentru care Skocpol și Sommers fac distincția între aceste trei instrumente este datorat criticilor pe care analiza comparativ-istorică le-a primit și pe care cele două susțin că se datorează celorlalte două instrumente de comparație istorică, mai slabe din punct de vedere metodologic. Autoarele promovează virtuțile analizei macro-cauzale care primește critici nedrepte din cauza confuziei cu celelalte două. Pe de altă parte, între cele trei instrumente există un ciclu tranzitiv: analizele contrastuale pot sugera ipoteze testabile, care trec printr-un proces analitic macro-cauzal pentru a se dezvolta în teorii mai generale, care pot fi analizate apoi cu ajutorul demonstrațiilor paralele.

Comparația istorică este doar una dintre multiplele aplicații ale analizei comparative; după Rueschmeyer și Stephens (1997), importanța ei derivă din faptul că multe dintre fenomenele observate au rădăcini în evenimente istorice mai mult sau mai puțin îndepărtate. Mai importante decât aceste aplicații sunt însă modul de acumulare a informațiilor, de formulare a teoriilor și de testare a acestora. Aceste aspecte au suscitât controverse în legătură cu validitatea analizei comparative, unele dintre acestea prezentate în următoarele secțiuni.

## 2.6 Critica cercetării comparative

Cercetarea (calitativ-) comparativă încearcă să răspundă, în cea mai mare parte, la aceleași probleme specifice analizei calitative. De exemplu, opinia că cercetările calitative nu sunt reprezentative, este la fel de puternică și pentru cercetarea comparativă. Opiniile asupra modului cum aceasta din urmă reușește să răspundă la aceste probleme sunt (încă) împărțite (Goldthorpe, 1997a,b; Ragin, 1997; Goldstone, 1997), ceea ce demonstrează două lucruri: faptul că cercetarea în științele sociale se lovește de probleme pe care analiza cantitativă nu reușește să le rezolve (altfel spus, există probleme de natură calitativă

care nu pot fi rezolvate prin instrumente cantitative) și faptul că metodele aplicate în analiza comparativă încearcă să rezolve aceste probleme, chiar dacă instrumentele prin care se încearcă acest lucru încă nu sunt suficient de dezvoltate.

Toate tipurile de probleme pe care le voi prezenta în continuare au o legătură strânsă între ele, o legătură aproape organică. De exemplu, nu se poate concepe problema determinismului fără a face legătura cu problema cazurilor puține și cea a reprezentativității, toate fiind o consecință directă a dualității nomotetic-ideografic, despre care am amintit la introducerea din capitolul 1. La fel, problema determinismului nu poate fi despărțită decât artificial de problema cazurilor negative, exemplele valabile pentru una fiind valabile și pentru cealaltă.

În cele ce urmează voi prezenta sumar aceste probleme, cu precizarea că ordinea lor nu reflectă importanța acestora și că despărțirea lor în secțiuni separate este făcută mai mult pentru clarificarea conceptelor, unele secțiuni făcând trimiteri recursive de la una la cealaltă.

### 2.6.1 Problema cutiei negre

Aceasta este un tip de problemă pe care metodologii analizei calitative o îndreaptă către slăbiciunile analizei cantitative. În esență, observarea unor corelații între variabile (variații concomitente, cum le numea Mill) nu este în nici un caz o dovadă a corectitudinii unui model. Orientarea discursului dinspre inferențe descriptive spre inferențe cauzale nu este un lucru simplu de făcut, întrucât cercetătorul trebuie nu doar să găserască regularități empirice ci să înțeleagă cum sunt generate regularitățile respective.

Mai întâi, înțelegerea nu poate fi realizată decât în măsura specificării unui model corect și (mai ales) complet; nimic nu poate să ferească cercetătorul cantitativ să nu cadă în capcana propriului model, întrucât eroarea de specificare a modelului este cea mai frecventă în acest tip de analiză. Există cazuri în care analistul comite o eroare și mai mare: aceea de a lăsa calculatorul să decidă care este “cel mai bun model”.

De exemplu, există o tehnică în analiza de regresie numită “Stepwise”: cercetătorul introduce în analiză o serie de variabile independente (multe la număr), iar sarcina calculatorului este aceea de a ordona variabilele în funcție de contribuția acestora la mărimea lui  $R^2$  (cu alte cuvinte, cu cât de mult contribuie la explicarea variației dependentei). O asemenea tactică nu poate fi decât greșită (și nu se folosește decât în scopuri explorative) deoarece eșantionul, extras în mod aleator, poate să prezinte anumite corelații conjuncturale între variabile care să fie

valabile doar pentru eșantionul acela și nu pentru altele, iar valoarea corelației să fie, de fapt, departe de corelația din populație.

Calculatorul nu poate să ofere altceva decât niște calcule matematice și statistice. Cercetătorii calitativiști văd în acest punct o slăbiciune fundamentală a analizei statistice, întrucât a cunoaște anumite corelații între variabile nu oferă explicații complete ale modului de funcționare a unui fenomen social. Fie pentru că cercetătorul a introdus în analiză niște variabile eronate, fie pentru că programul de analiză statistică a considerat importante niște variabile altminteri neinteresante, este foarte posibil ca modelul final să nu conțină variabile în realitate importante.

Apoi, chiar dacă toate variabilele importante sunt până la urmă identificate și introduse în model, modul concret în care se împletesc acestea pentru a produce fenomenul analizat rămâne o necunoscută: cunoaștem datele de intrare (variabilele), cunoaștem datele de ieșire (variabila dependentă) însă analiza statistică nu poate să spună ce se întâmplă mai exact cu acele variabile, care este modul în care se combină, care este chimia socială prin care condițiile de intrare determină efectul. Între condițiile de intrare și efect se află așa numita "cutie neagră"; este procesul pe care utilizatorii de analiză statistică îl pot observa și măsura dar nu-l pot explica.

În contrast, cercetătorii care aplică metode calitative susțin că această problemă este ușor rezolvată în studiile de caz, întrucât nu se analizează variabile abstracte ci desfășurări concrete ale evenimentului, în toată bogăția de informații asociată cu observațiile și munca de teren.

Criticând analiza calitativă, Goldthorpe (1997a) arată că deși problema cutiei negre pare mai evidentă la cercetările cantitative, în fapt nici analiza în profunzime a cazurilor nu este ocolită de această problemă. Simpla descriere a cazului studiat (oricât de detaliată ar fi) nu poate fi un bun substitut al unei teorii cauzale întrucât (pentru a fi considerată o teorie) ar trebui să aibă măcar o palidă pretenție de generalitate și să poată fi testată empiric. Or, de cele mai multe ori analiza calitativă nu face nimic altceva decât să explice excelent un fenomen particular, ceea ce nu servește cu nimic avansului științei.

Goldthorpe vede analiza istorică într-o situație și mai dificilă: dacă o analiză calitativă recentă are avantajul informațiilor proaspăt culese, încercarea de introducere în ecuație a relatărilor istorice este viciată de calitatea și cantitatea acestor informații. De multe ori sursele istorice sunt nu doar contradictorii, ci au origini uneori suspecte.

În replică, Rueschmeyer și Stephens (1997) punctează și ei faptul că, deși nici analiza cantitativă nici analiza istorică nu pot stabili explicații cauzale clare (cu alte cuvinte nu pot rezolva satisfăcător problema cutiei

negre), fenomenele macrosociale nu pot fi în nici un caz înțelese cu ajutorul unei analize statistice transversale care utilizează doar câțiva indicatori comparabili pentru un anumit moment temporal. Doar înțelegând istoria unui caz și desfășurarea acestuia în timp putem spera să deschidem cutia neagră a mecanismului de producere a unui fenomen.

Problema cutiei negre se referă așadar la posibilitatea oferirii unei *explicații* a producerii unui fenomen, nu doar a condițiilor de intrare și de ieșire. Din acest punct de vedere, analizele calitative oferă explicații mult mai bune decât cele cantitative. Cu tot dezavantajul unor informații incomplete ori de o calitate îndoielnică, analiza istorică longitudinală se află într-o poziție net superioară analizei cantitative care nu studiază fenomenele decât transversal.

Sigur, nu se poate vorbi despre superioritatea unei analize asupra celeilalte, ci numai despre adecvarea metodei în funcție de situația concretă de cercetare. În cazul de față însă, slăbiciunea analizei cantitative constă chiar în pretenția lui Goldthorpe asupra generalității.

Pretenția de generalitate nu este cu adevărat o pretenție dacă se referă doar la cazurile din același spațiu temporal. O teorie cu adevărat general valabilă trebuie să răspundă cu aceeași promptitudine în oricare moment în timp, comparând cazuri chiar și la o distanță de mii de ani. Durkheim (1995) considera în mod just că studierea triburilor este o modalitate foarte simplă și directă de a testa acele caracteristici fundamentale care formează structura unei societăți, însă cum am putea să testăm o teorie despre globalizare studiind un trib?

Trebuie să recunoaștem că societățile și omenirea în ansamblul ei se află într-un proces de schimbare continuă. O teorie despre societatea globală din prezent ar putea să nu prezinte nici un fel de relevanță pentru o societate viitoare să spunem peste 10.000 de ani. Teune (1997, p.80) arată foarte clar că “problema majoră a științelor sociale este faptul că lucrează cu sisteme vagi, încurcate și schimbătoare”.

O teorie asupra societății prezente (chiar presupunând că ar putea fi testată cu mijloace și instrumente cantitative) nu este nimic altceva decât o teorie despre *acest* moment temporal și nu poate fi generalizată nici pentru trecut, nici pentru viitor. Acestea fiind spuse, întrebarea care se naște este: cu ce diferă o astfel de cercetare de explicația calitativă asupra unui singur studiu de caz?

Marele reproș care se aduce analizelor calitative este acela că explică anumite cazuri particulare însă nu pot fi generalizate. Utilizând metoda reducerii la absurd, același reproș se poate aduce oricărei teorii care încearcă să explice un anumit fenomen dintr-un moment temporal

specificat. Asemănarea este mai mult decât evidentă: o teorie în momentul  $T_i$  este echivalentul unui studiu de caz în mulțimea tuturor cazurilor posibile în timp.

Din acest motiv, împărtășesc opinia cercetătorilor cu abordare calitativ-comparativă: explicația (rezolvarea cutiei negre) are mult mai multe șanse prin studierea atentă a unui sistem particular (fie acela omenirea în ansamblul ei), luând în calcul și evoluția lui istorică. Dacă această abordare întâmpină dificultăți în rezolvarea acestei probleme, analiza cantitativă care utilizează doar niște corelații de indicatori cantitativi nu va reuși niciodată să o rezolve.

Sigur, există asocieri de variabile care vor rezista întotdeauna în timp. Între educație și venit va exista întotdeauna o corelație pozitivă, însă aceste tipuri de variabile nu pot constitui fundamentul explicației tuturor fenomenelor sociale posibile. De exemplu, nici un model cantitativ nu va reuși vreodată să explice evenimentul de la 11 septembrie din Statele Unite, iar acest eveniment unic și excepțional a marcat istoria globală a ultimilor ani. Cutia neagră a explicației acestor fenomene nu poate fi deschisă decât cu ajutorul unor metode de comparație calitativă.

Fenomenele sociale sunt complexe. Este puțin probabil ca științele sociale să reușească o explicație completă a acestora, mai ales că sunt foarte volatile și schimbătoare. Din acest motiv, atât cercetarea cantitativă cât și cea calitativă sunt supuse la fel de mult problemei cutiei negre, însă (cel puțin până la dezvoltarea unui sistem foarte solid de indicatori statistici globali), comparațiile calitative sunt cele care pot oferi o soluție mai bună ori cel puțin acceptabilă din punct de vedere teoretic.

### 2.6.2 Problema lui Galton

Problema lui Galton este strâns legată de problema cutiei negre, ambele fiind derivate din încercarea de a explica producerea unui fenomen. Originea problemei poate fi datată cu certitudine: conform lui White (2004), renumitul om de știință Francis Galton a participat în noiembrie 1888 la o prezentare făcută de către Edward Tylor asupra a ceea ce astăzi se numește cercetare cros-culturală.

Tylor era interesat de evoluția societăților, afirmând că pe măsură ce devin tot mai complexe trec de la o structură matrilineară la una patrilineară. Într-o cercetare remarcabilă pentru sfârșitul secolului XIX, Tylor a cules date statistice despre 350 de culturi pe care le-a utilizat pentru a-și demonstra teoria, publicată în anul 1889 sub titlul *On a Method of Investigating the Development of Institutions, Applied to Laws of Marriage and Descent*.

Galton, interesat și el de relațiile familiale (făcuse cercetări anterioare asupra transmiterii ereditare a geniului), a contestat imediat validitatea teoriei lui Tylor sub pretextul că societățile comparate de acesta nu reprezintă cazuri diferite ci fac parte din aceeași “familie” formată prin difuziunea socio-culturală provocată de o structură ereditară comună.

Această afirmație a lui Galton a devenit mai târziu problema la care se referă această secțiune: cazurile studiate nu sunt independente unele de celelalte, astfel că validitatea unor concluzii comparative sunt puse imediat sub semnul întrebării. Începând cu prima jumătate a secolului XX, cercetătorii care au încercat să compare societăți și culturi au fost permanent interesați să ofere explicații care să nu poată fi atinse de problema lui Galton, străduindu-se în mod constant să o rezolve.

Murdock et al. (2006) au preluat de la Tylor ideea unei baze de date cu indicatori comparabili. Pornind de la o încercare inițială de a afla dacă societățile predominant masculine sunt mai predispuse la război, ei s-au confruntat cu exact problema lui Galton și au testat această ipoteză construind un eșantion de 186 provincii culturale mai mult sau mai puțin independente unele de celelalte în ceea ce privește originea istorică și difuziunea culturală. Efortul lor s-a concretizat în așa numitul Eșantion Standard Cros-Cultural, care conține aproape 2000 de indicatori comparabili care pot fi utilizați pentru a testa diverse soluții pentru a depăși problema lui Galton.

Același tip de problemă a fost identificată de Zamfir (1990) în ceea ce el numește problema interdependenței, care tratează în esență același fenomen al difuziunii. Întrucât țările se influențează reciproc în strategiile de dezvoltare (exemplul oferit fiind cel al procesului de industrializare), variația naturală a acțiunilor la nivel de societate este limitată în mod artificial, rezultatul fiind restrângerea bazei empirice a sociologiei. Acolo unde nu există suficiente cazuri diferite (pentru a le putea compara), sociologia nu poate oferi în mod facil explicații asupra fenomenelor de interes.

Deși problema lui Galton a apărut pentru prima oară în antropologie (și este în continuare una dintre principalele preocupări ale antropologilor comparatiști), ea și-a făcut simțită prezența și într-o știință exactă cum este statistica.

În cadrul anumitor analize (cum este cea de regresie), problema este denumită “autocorelație”: atunci când elementele unui eșantion nu sunt independente (de exemplu, printre respondenți sunt membri din aceeași gospodărie), estimările statistice devin supra-ponderate deoarece mărimea reală a eșantionului nu este  $n$  ci  $n - k$  unde  $k$  este numărul de elemente care nu sunt independente.

Sistemele sociale sunt atât de complexe încât testarea unei teorii pe cazuri noi este practic imposibilă deoarece fiecare caz nou este strâns legat de toate celelalte, atât prin conexiuni temporal-istorice cât și prin interdependențe subtile în același spațiu temporal. Un caz “nou” nu este de fapt un caz ci o parte a macrosistemului, fapt confirmat de recente teorii asupra globalizării. Cu alte cuvinte, orice încercare epistemologică de testare a unei teorii sociale se lovește de riscul autoconfirmării (ori invers, de imposibilitatea respingerii deoarece fiecare caz este o parte a sistemului).

Goldthorpe (1997a) este mai optimist cu privire la această problemă, demonstrând că anumite caracteristici culturale rămân total diferite chiar dacă politicile aplicate sunt asemănătoare. Citând o cercetare a lui Shavit și Blossfeld asupra inegalităților de clasă în educație, el arată că (deși politicile naționale asupra educației sunt foarte omogene și construiesc o curiculă similară) variația pattern-urilor de inegalitate nu se schimbă odată cu schimbarea curiculei iar inegalitățile persistă în timp, ceea ce demonstrează o rezistență la difuziunea culturală<sup>2</sup>. Mai mult, în statistică (dar și în cercetările antropologice) au fost dezvoltate metode de a rezolva problema lui Galton (de exemplu statistica I a lui Moran prezentată de Eff, 2001), în timp ce alte curente nici măcar nu încearcă să rezolve problema, considerând difuziunea culturală o oportunitate de cercetare în sine.

### 2.6.3 Problema “variabile multe, cazuri puține”

Am arătat în secțiunile anterioare că în cercetarea comparativă, numărul de cazuri studiate este relativ restrâns. Chiar dacă pot fi studiate sensibil mai multe cazuri decât într-o analiză pur calitativă, sunt foarte comune situațiile în care numărul de ipoteze cauzale depășește foarte ușor numărul de cazuri disponibile pentru analiză. De exemplu, în cercetările privind țările OECD, ori cele care studiază țările lumii a treia, ori chiar analize comparative asupra țărilor din Europa de Est, numărul de cazuri este mic.

În cercetarea cantitativă, atunci când numărul de variabile depășește numărul de cazuri spunem că există o problemă de *supra-determinare*, iar analizele statistice multivariate nu mai pot fi aplicate deoarece sunt prea puține grade de libertate. Într-un mod mai mult sau mai puțin similar, analiza comparativă împărtășește această dificultate.

---

<sup>2</sup>eu cred însă că aceasta este doar o demonstrație a rezistenței la schimbare și că este doar o chestiune de timp până când politicile comune vor determina patternuri asemănătoare, chiar și într-un domeniu atât de sensibil cum este cel al inegalităților



Începând cu Skocpol, cercetătorii care aplică metode comparative susțin aplicarea robustă a metodelor inductive de tip Mill, care se pretează la studiul unui număr limitat de cazuri. Alți cercetători (Nichols, 1986; Goldthorpe, 1997a) atacă ferm această aventură afirmând că metodele milliene, dezvoltate inițial pentru cercetări experimentale, nu se pot aplica pentru studiul fenomenelor sociale deoarece este imposibil să găsim și să includem în model toți factorii cauzali relevanți (așa cum cer canoanele lui Mill).

În replică, Goldstone (1997) arată că, deși criticii au dreptate să afirme că metodele lui Mill nu se pot aplica în cercetările macro-sociologice, în fapt ambele tabere înțeleg greșit aplicarea acestora. Goldstone afirmă nici mai mult nici mai puțin că metodele bazate pe număr mic de cazuri nu sunt de tip Mill și că logica acestora este bazată mai degrabă pe o combinație de raționamente deductive și inductive denumită de acesta “trasarea proceselor” (în engl. *process tracing*).

Pentru utilizarea canoanelor diferenței ori a concordanței, două cazuri trebuie fie să difere sub toate aspectele mai puțin unul, fie să semene sub toate aspectele mai puțin unul. Este limpede că o astfel de situație nu poate fi obținută în realitatea socială, astfel că metodele lui Mill nu pot fi aplicate decât în situații experimentale controlate (ceea ce în domeniul științelor societale este indezirabil).

În accepțiunea lui Goldstone, este extrem de nefericit faptul că o autoare atât de rafinată ca Skocpol își asociază una dintre cele mai reprezentative lucrări comparative scrise vreodată cu canoanele inductive ale lui Mill. El demonstrează de altfel, studiind foarte atent lucrarea lui Skocpol, că metodele aplicate de aceasta nu au nici o legătură cu Mill, ci au o cu totul altă natură. Dacă Skocpol afirmă cu tărie contrariul, atunci fie se află în eroare fie pur și simplu nu înțelege natura inductivă a canoanelor lui Mill.

Nici analiza cantitativă, nici cea comparativă nu pot să răspundă satisfăcător la problema cazurilor puține. Chiar dacă nici o soluție nu este în final perfectă, există totuși propuneri. De exemplu, Lijphart (1971) menționează nu mai puțin de patru direcții de acțiune:

- Probabil că cea mai rezonabilă soluție (menționată și de alți autori) ar fi încercarea de mărire a numărului de cazuri cât mai mult cu putință, atât în termeni geografici cât și în analize longitudinale și istorice. Extinderea geografică se poate realiza fie prin includerea în analiză a unor cazuri depărtate geografic însă substanțiale din punct de vedere teoretic, fie prin împărțirea cazurilor existente în mai multe zone geografice care le compun (în

cazul în care există suficiente date statistice despre toate regiunile respective). Prin analizele longitudinale, se pot analiza fie aceleași cazuri în mai multe conjuncturi temporale, fie introducerea în analiză a unor cazuri istorice de notorietate (deci importante din punct de vedere teoretic), din nou dacă există suficiente informații.

- Acolo unde nu putem să mărim numărul de cazuri, am putea să micșorăm numărul de condiții cauzale din ipoteză, iar acest lucru ar putea fi realizat prin reducerea “spațiului de proprietăți” (așa cum îl definește Lazarsfeld, 1937). Tehnica este similară analizei factoriale din statistică: dacă anumite condiții cauzale reprezintă mai mult sau mai puțin aceeași caracteristică, atunci am putea să le combinăm pentru a reduce numărul de variabile.
- Dacă nici una dintre soluțiile inițiale nu poate fi realizată, Lipjhart propune concentrarea atenției doar spre cazurile strict comparabile, deoarece acestea vor afișa foarte multe condiții cauzale similare: acestea vor juca rolul factorilor de control din analiza statistică (rămânând constante de-a lungul cazurilor). Analiza se poate concentra apoi spre restul variabilelor, tehnica fiind o pseudo-reducere a numărului de variabile (neavând nici un efect asupra fenomenului, constantele vor putea fi excluse fără probleme din model)
- Concentrarea atenției spre variabilele cheie este o altă posibilă soluție. Într-o știință foarte departe de calcule exacte ca în matematică ori fizică, în care numărul de cazuri care pot fi studiate nu este foarte mare, nu putem avea pretenția să construim modele perfecte în care să includem toate variabilele care ar putea să aibă o influență asupra producerii unui fenomen. Ca și în statistică de altfel, cel mai bun model este cel care conține cele mai puține variabile, însă acelea trebuie să aibă un rol cheie în desfășurarea evenimentelor. Societățile primitive, arată Lipjhart, au o funcționare mult mai simplă decât cele moderne; de aceea, numărul de variabile necesare unui model este suficient de mic pentru a putea realiza o analiză comparativă. De aceea, cercetările antropologice au înregistrat primele succese în aplicarea metodelor comparative, iar societățile primitive ar putea juca rolul unui adevărat laborator cvasi-experimental pentru testarea și extragerea variabilelor importante dintr-un model.

Alerti la problema cazurilor puține, Dogan și Pelassy (1990) au analizat propunerile lui Lipjhart și au propus concentrarea analizei spre unități macrosociale la un nivel mai scăzut de agregare, ca o metodă suplimentară

de a mări numărul de cazuri: în loc să facem comparații între un număr mic de țări, am putea să analizăm diferite regiuni din fiecare țară. De exemplu în România sunt câteva provincii istorice relativ diferite între ele, iar astfel de regiuni există în fiecare țară.

O astfel de abordare duce la problema lui Galton, motiv pentru care problema din această secțiune nu are o soluție definitivă; în final, abilitatea cercetătorului este factorul decisiv care poate asigura succesul unei cercetări. În funcție de caracteristicile particulare ale unui studiu, se pot încerca una sau mai multe soluții prezentate. Atât problema cutiei negre, cât și cea a lui Galton sunt potențate de problema cazurilor puține în special în societatea de astăzi, întrucât urmarea procesului de globalizare are loc o scădere a cazurilor reale. Prin urmare, o posibilă opinie este aceea că nu mai avem o mulțime de societăți diferite pe care să le comparăm, ci o singură societate globală.

#### 2.6.4 Problema determinismului

Problema determinismului este de fapt o problemă a dualității dintre cele două abordări: cea deterministă (specifică analizei calitativ-comparative, care în accepțiunea lui Mill afirmă că prezența unei cauze va determina întotdeauna un fenomen efect) și cea probabilistă (specifică analizei cantitative: prezența unei cauze crește probabilitatea de apariție a fenomenului efect).

Unii cercetători, de exemplu Lieberson (1991), văd în constrângerea determinismului o slăbiciune fundamentală a analizei: o cauză poate să nu producă un efect datorită altor cauze care acționează în sens contrar. Exemplul pe care îl oferă Lieberson este extrem de ilustrativ; să spunem că o ipoteză afirmă: “dacă șoferul este beat la volan, se produc accidente”. Și totuși, există cazuri în care se produc accidente când șoferul nu este beat, după cum există cazuri în care un șofer beat nu produce un nici accident.

Dacă am privi totul în mod determinist, identificarea cazurilor în care nu se produc accidente (așa numitele cazuri negative, discutate la secțiunea 2.6.6), ar trebuie să anulăm ipoteza noastră inițială. Din contră, așa cum raționăm în viața de zi cu zi “ne așteptăm” ca un șofer beat să producă un accident. Prezența alcoolului crește probabilitatea producerii unui accident; cu alte cuvinte, dacă privim fenomenul în sens probabilist, nu renunțăm pur și simplu la ipoteză dacă găsim un caz negativ. Vom continua să credem că alcoolul poate să producă accidente, chiar dacă un șofer beat conduce o mașină fără nici un incident. Sunt și alți factori care concură la producerea de accidente: experiența șoferului, aglomerația

de pe străzi ori pur și simplu întâmplarea. Lieberson susține că analiza calitativ-comparativă (în starea ei metodologică de atunci) este incapabilă să trateze astfel de cazuri.

Această demonstrație este atacată de Savolainen (1994), care vine în apărarea metodelor de tip Mill arătând că aceste canoane pot fi utilizate în sens probabilist. De exemplu, în *Etica protestantă și spiritul capitalismului*, Weber (2003) a utilizat metoda diferenței: preocupat de apariția capitalismului, care s-a produs doar în anumite țări dar nu în altele, Weber a căutat factorul decisiv care nu se găsește decât în primele, ajungând la concluzia că acel factor este etica protestantă.

Savolainen este de părere că metoda aplicată trebuie să fi fost probabilistă: cu siguranță, unii protestanți nu au aplicat acumularea de capital iar unii catolici dimpotrivă, au acumulat și ei capital, însă ceea ce contează este efectul agregat al comportamentului tuturor protestanților versus celui al tuturor catolicilor: faptul de a fi protestant crește semnificativ probabilitatea de a acumula capital.

Replica lui Lieberson (1994) este promptă: Mill însuși (1970) avertizează că atât metoda concordanței cât și metoda diferenței nu sunt potrivite pentru studiul fenomenelor sociale (cartea VI, capitolul 7, conform lui Goldthorpe, 1997). În primul rând, pentru că nu avem posibilitatea de a face experimente pe scară largă, socială. Apoi, chiar dacă am avea posibilitatea de a face nenumărate experimente, până când am ajunge să testăm toate cauzele o anumită instanță se va fi schimbat deja, pentru că instanțele sociale sunt într-o perpetuă stare de schimbare.

O clarificare binevenită în această dispută sterilă este adusă mai târziu de către Goldstone (1997, p.109): într-adevăr, metodele de tip Mill pot fi utilizate în sens probabilist. Deși pare să fie de aceeași părere cu Savolainen, în fapt arată că și acesta greșește atunci când afirmă că metodele lui Mill ar putea fi aplicate la macro-sociologie.

Exemplul pe care îl oferă Goldstone implică în mod cert raționamente probabiliste, deși logica de bază este de tip Mill: un câmp de porumb este împărțit în două parcele, din care pe una se aplică un fertilizator (echivalentul grupului experimental din științele sociale) iar pe cealaltă parcelă nu se aplică nici un fel de fertilizator (grup de control).

Presupunând că este un fertilizator eficace, este *probabil* ca aplicarea acestuia să mărească recolta; nu se va obține *întotdeauna* o recoltă bogată (ca în interpretarea deterministă), întrucât în anii secetoși nu se va obține nici un fel de recoltă indiferent dacă se aplică sau nu un fertilizator.

Mergând mai departe cu logica exemplului, într-un an cu ploi normale se va obține o diferență semnificativă de recolte între cele două parcele, această diferență fiind atribuită unui singur factor: aplicarea fertilizatorului, ceea ce este perfect compatibil cu aplicarea metodei diferenței a lui Mill. Prin acest excelent exemplu, Goldstone demonstrează așadar că metodele de tip Mill nu sunt limitate la o analiză deterministă ci pot fi utilizate și în sens probabilist.

Goldstone lămurește definitiv această polemică: toată lumea greșește. Atât Savolainen (1994), care apără metoda, cât și Lieberman (1991, 1994) și Goldthorpe (1997a), care o atacă, nu analizează în mod corect situația. Plecând de la premisa că macrofenomenele sociale sunt stabile (fapt demonstrat de altfel încă din pe vremea lui Durkheim), afirmația “Condiția macrosocială A cauzează macrofenomenul B” nu are o natură deterministă ci, în fapt, probabilistă (chiar dacă sub o formă subtilă și uneori ascunsă).

Chiar dacă afirmația pare de origine deterministă, în fapt lucrurile stau cu totul diferit: macrofenomenele nu există în eter, ci sunt cuprinse din acțiunile agregate ale indivizilor, singurii actori reali care acționează probabilist (chiar haotic). Dacă eforturile a *suficient de mulți* actori individuali se combină însă, fenomenele macrosociale se produc. O condiție cauzală macrosocială cauzează un fenomen macrosocial, însă nu într-un mod determinist ci prin agregarea probabilităților de acțiune ale actorilor individuali.

Nu este corectă afirmația “A cauzează B cu o probabilitate P”, întrucât a vorbi despre o singură probabilitate este similar cu a vorbi despre un singur individ, ori societatea nu funcționează la unison cu toți indivizii; mai degrabă am putea spune: “atunci când există condiția macrosocială A, este foarte probabil ca indivizii să reacționeze în așa fel încât se produce macrofenomenul B”.

### 2.6.5 Problema diversității limitate

Din multe puncte de vedere, științele fizice sunt aproape deterministe: la urma urmei, un măr cade întotdeauna pe pământ datorită forței de atracție gravitațională. Mai mult, se poate calcula viteza cu care va cădea un măr, pe baza masei acestuia, iar toate aceste lucruri permit fizicienilor să emită legi de funcționare a sistemelor fizice și chiar predicții asupra funcționării acestora până la nivel de atom.

Prin contrast, științele sociale sunt mult mai complexe. În lumea socială, sunt mai multe posibilități de a produce un fenomen, există mai multe moduri posibile de a obține același rezultat. Nu există o singură

“rețetă” magică prin care să putem produce un anumit fenomen social, dimpotrivă este extrem de dificil să surprindem toate modalitățile de combinare a factorilor cauzali pentru a obține un anumit efect. În fapt, lucrurile pornesc chiar de la acest fapt: mai multe condiții cauzale pot contribui (într-o măsură mai mare sau mai mică, la un anumit moment în timp) la apariția unui fenomen.

Cercetătorii pot să inventarieze un set posibil de condiții cauzale, însă adevărata dificultate este cea de a identifica nu doar o listă de cauze, ci modul cum acestea se combină unele cu altele pentru a produce un anumit rezultat. Dacă ne referim doar la prezența sau absența acestor condiții cauzale (există și posibilitatea existenței acestora în diverse grade de intensitate, un aspect care va fi discutat în secțiunea 5.3), numărul total de modalități prin care aceste condiții cauzale se pot combina poate fi calculat cu o formulă foarte simplă, egală cu  $2^k$ , unde  $k$  este numărul de condiții cauzale incluse în analiză. Acest număr crește exponențial cu fiecare nouă condiție cauzală inclusă în analiză.

Chiar și pentru 5 condiții cauzale nu ar fi nici un fel de problemă pentru un cercetător să studieze  $2^5 = 32$  cazuri, pentru a acoperi toate combinațiile posibile. Când numărul condițiilor cauzale crește foarte mult însă, această strategie devine nu doar inefficientă ci chiar imposibilă. Pentru 10 condiții cauzale numărul total de combinații cauzale este egal cu 1024, ori niciodată o cercetare calitativă nu va acoperi un asemenea număr de studii de caz. Mai mult, chiar și cele (maxim) câteva zeci de studii de caz analizate calitativ nu sunt toate complet diferite unele de celelalte. O veche instrucțiune metodologică spune, cu privire la numărul optim de studii de caz necesare pentru o analiză calitativă, să fie studiate atâtea cazuri până când informația începe să se repete.

Acesta este punctul de maximă importanță pentru științele sociale: oricât de diversă este viața socială, până la urmă informația începe să se repete. Chiar dacă sunt mai multe modalități de a obține un anumit rezultat (ex. sunt mai multe modalități de a deveni bogat, sau mai multe modalități de a deveni fericit), până la urmă cele mai multe cazuri încep să semene unele cu celelalte. Vom observa, de exemplu, că oamenii care muncesc mult tind (de regulă) să aibă un venit mai mare decât al celor care muncesc mai puțin. Numărul tuturor modalităților de a combina un set de condiții cauzale este extrem de mare, însă în practică nu vom întâlni decât un număr destul de restrâns de configurații cauzale empirice. Aceasta este ceea ce Ragin (1987, p.104) numește “diversitate limitată”.

Sigur, o mare însemnătate o are definiția cercetătorului asupra fenomenului studiat (care poate să restrângă sau să lărgască numărul

cazurilor posibile) însă această situație este cât se poate de realistă. Reluând exemplul lucrării lui Skocpol (1979), definiția cu care s-a lucrat a identificat doar trei asemenea cazuri: Franța, Rusia și China. Termenul destul de restrictiv de “revoluție socială” ar fi putut fi lărgit la cel de “revoluție” pur și simplu, iar în analiză ar mai fi putut intra probabil și revoluțiile militare din America de Sud (ori, mai nou, revoluțiile din Europa de Est). Skocpol a preferat însă un termen mai precis, fapt care a restrâns cazurile posibile la cele trei studiate.

Chiar și în alte tipuri de studii, numărul de cazuri rămâne foarte mic. De exemplu, o analiză calitativ-comparativă pe țările din Uniunea Europeană nu va avea mai mult de 28 de cazuri. La un asemenea număr de cazuri observate, este logic imposibil de acoperit să spunem 1024 de combinații posibile pentru 10 condiții cauzale. Multe dintre deja puținele cazuri existente vor manifesta combinații cauzale asemănătoare, așadar nimic nu garantează acoperirea completă a tuturor combinațiilor posibile de cauze. În fapt, doar anumite combinații sunt observate în practică, cele mai multe fiind doar *logic* posibile dar niciodată întâlnite în procesul de cercetare.

### 2.6.6 Problema contradicțiilor

Așa cum este arătat la secțiunile 5.1.5 și 5.3.7, construirea tabeli de adevăr presupune ordonarea tuturor combinațiilor *posibile* de valori ale condițiilor cauzale și atribuirea unei valori de 1 sau 0 pentru variabila rezultat, în funcție de valorile combinațiilor *observate* pe acea variabilă. Este foarte probabil ca mai multe cazuri studiate să prezinte exact aceeași combinație de condiții cauzale (întrucât realitatea are o diversitate limitată). Dacă toate aceste cazuri sunt la fel de omogene și în privința fenomenului studiat atunci pe coloana rezultat din tabela de adevăr se trece valoarea observată a fenomenului.

Există însă situații în care, deși cazurile studiate manifestă aceeași combinație cauzală, ele manifestă valori diferite în ceea ce privește producerea efectului. În cazul analizei comparative pe mulțimi distincte (care utilizează algebra booleană cu doar două valori: 0 pentru absența fenomenului și 1 pentru prezența acestuia), această situație se traduce printr-o contradicție.

Cazurile studiate sunt contradictorii în ceea ce privește producerea fenomenului de interes, iar cercetătorul este pus în dificultate atunci când dorește să utilizeze mecanismul de minimizare booleană (prezentat la secțiunea 5.1). Mecanismul cere introducerea în analiză doar a combinațiilor cauzale în care fenomenul de interes este prezent (valoarea

1 la variabila efect). Cum diversitatea condițiilor cauzale este oricum limitată, orice combinație cauzală este importantă; excluderea uneia din analiză scade drastic eficiența procesului de minimizare.

Teoria clasică spune că orice contradicție infirmă ipoteza cauzală și nu poate fi inclusă în analiză. Să spunem însă că am studiat 10 cazuri, dintre care 9 se manifestă prin prezența fenomenului de interes iar al zecelea contrazice primele nouă cazuri prin manifestarea absenței fenomenului. Este al zecelea caz o dovadă suficientă pentru a contrazice primele nouă? Dar dacă am fi studiat 100 de cazuri, dintre care 99 ar fi manifestat prezența fenomenului, ar fi ultimul caz suficient pentru a contrazice complet primele 99?

Sigur, exemplul utilizat este unul extrem, dar se poate întâmpla. Datorită specificului analizei (calitative în esența ei), există constrângeri inerente cu privire la numărul de cazuri care *pot fi* studiate; în plus, diversitatea cazurilor este oricum limitată așa încât de obicei nu există mai mult de 4 sau 5 cazuri asociate cu o combinație cauzală. Cu un număr atât de mic de cazuri studiate, frecvența relativă a celor două stări prin care se manifestă fenomenul nu mai are nici un fel de relevanță iar orice caz care prezintă o contradicție cântărește la fel de mult ca 4 sau 5 cazuri opuse.

Din punct de vedere teoretic, problema este încă aprins discutată. La apariția analizei calitativ-comparative (pe acea vreme doar pe mulțimi distincte), unii cercetători (Ragin, 1987; Savolainen, 1994) erau de părere că ar trebui urmată abordarea lui Mill și că orice combinație contradictorie trebuie eliminată din analiză, indiferent de frecvențele de apariție ale prezenței și absenței fenomenului de interes. Teoria pe care se bazează o asemenea strategie spune că apariția unei situații contradictorii este un indicator al insuficienței cadrului teoretic în care se desfășoară cercetarea.

Soluția propusă în astfel de cazuri este revizuirea modelului teoretic; trebuie găsită o condiție cauzală suplimentară care să clarifice situația contradictorie. Este posibil ca situația contradictorie să fie generată nu pentru că teoria este greșită, ci pentru că este doar incompletă; o altă condiție cauzală nu a fost luată în calcul deși este necesară pentru explicarea fenomenului de interes. Uneori, contradicțiile sunt cauzate nu neapărat de teoria utilizată, ci de operațiunile prin care cercetătorul a efectuat dihotomizarea datelor de intrare (a se vedea secțiunea 5.1.8).

Contradicțiile la nivel agregat merită foarte multă atenție, întrucât se pot manifesta ca evidențe pentru modificarea teoriei. În fapt, la acest nivel orice contradicție este similară falsificării unei teorii, moment în care atât în accepțiunea lui Kuhn cât și în cea a lui Popper teoria trebuie modificată, o nouă paradigmă devenind dominantă.



## Capitolul 3

# Scopuri ale comparației

Analiza comparativă nu reprezintă un tip cu totul special de analiză, rupt de celelalte tipuri de metode și tehnici de analiză a câmpului social. Dimpotrivă, se înscrie într-un bagaj metodologic pe care orice sociolog ar trebui să-l stăpânească, alături de celelalte metode pe care le completează și le extinde.

În mare parte, scopurile cercetării cantitative se suprapun peste scopurile mai largi ale cercetării sociologice. Studiul vieții sociale, al grupurilor și al societăților este atât un scop al sociologiei cât și al analizei comparative. Deși scopurile sociologiei sunt mai largi, aș putea să spun că scopurile analizei comparative pot fi regăsite în totalitate printre acestea. O excelentă listă de asemenea scopuri a fost prezentată de către Dogan și Pelassy (1990) în celebra lor carte *How to compare nations. Strategies in comparative politics*. Întrucât găsesc această listă atât explicită cât și cuprinzătoare, am organizat secțiunile acestui capitol în jurul informațiilor prezentate de către cei doi autori, la care am adăugat și altele, pe care le consider importante în cercetarea socială.

Lista scopurilor pe care o voi prezenta nu este nici pe departe exhaustivă, însă cred că demonstrează faptul că în știință comparația este indispensabilă. Chiar dacă există situații în care se poate cerceta și în lipsa unei comparații directe, utilizând un singur caz, știința în general utilizează mai multe cazuri pentru analiză. În mod explicit sau implicit, cazurile sunt analizate, în cea mai mare parte, comparativ.

### 3.1 Pentru a formula enunțuri sociologice

În lucrarea originală (Dogan și Pelassy, 1990, p.15) secțiunea respectivă se numește “Pentru a formula legi sociologice”, însă am găsit această perspectivă puțin restrictivă întrucât nu toate cercetările sociologice care utilizează metoda comparației se încheie prin formularea unor legi.

Zamfir (1990, p.29) are o abordare mai practică, observând foarte devreme faptul că: “...greu putem găsi legi sociale universale”. În consecință, și-a orientat eforturile de cercetare în direcția formulării de *explicații* sociologice a fenomenelor sociale, utilizând multiple perspective asupra mecanismelor cauzale. În această direcție, Zamfir este de părere că raționamentul deductiv joacă un rol mai important decât observarea regularităților empirice prin metode inductive. În ceea ce mă privește, am încercat să acopăr prin termenul mai general de “enunț sociologic” ambele finalități ale cercetării sociale.

Analiza cauzală, analiza comparativă și analiza sociologică împart multiple finalități de cercetare, fără a se confunda unele cu celelalte. În secțiunea 2.1 am tratat aceste aspecte, consemnând această suprapunere: întrucât analiza comparativă se utilizează cu precădere pentru a explica apariția unor fenomene sociale, este normal să apeleze la metodologia specifică analizei cauzale.

O anumită doză de neclaritate apare nu atât din suprapunerea celor două, cât mai ales în modul prin care ele se diferențiază. În opinia mea, aceste diferențe apar atât prin intermediul mecanismului de analiză (predominant inductiv versus predominant deductiv), cât și prin finalitatea urmărită (formularea de legi versus formularea de explicații sociologice).

Scopul oricărei științe este găsirea universalilor, însă în sociologie acest lucru este extrem de dificil, datorită eterogenității unor amalgame sociale formate în diverse colțuri ale lumii și în momente temporale diferite:

- unele societăți au o suprafață extrem de redusă (Liechtenstein, Vatican) în timp ce altele se întind pe continente întregi (Statele Unite, Rusia, China); aceleași exemple ar putea fi contrastate din punctul de vedere al mărimii populației;
- unele societăți sunt extrem de sărace (de exemplu cele africane) în timp ce altele sunt foarte bogate (Luxemburg, Elveția etc.);
- din punct de vedere cultural, societatea europeană este cât se poate de diferită de societatea japoneză (cel puțin până la instaurarea regimului Meiji);

- unele societăți se caracterizează printr-o densitate a locuirii extrem de mare (să spunem o localitate cum este Mexico City) în timp ce în altele densitatea este extrem de scăzută (zone în care predomină deșertul, ori ferme întinse din Australia);
- unele societăți au o istorie de mii de ani (China) în timp ce altele abia se formează ori s-au format foarte recent (Republica Moldova).

Este aproape imposibilă găsirea unor enunțuri cu un caracter atât de general încât să acopere toate aceste tipuri de societăți, iar caracterul de universalitate nu se referă doar la societățile prezente, ci la toate societățile din toate timpurile, atât din trecut cât și cele viitoare.

În sociologie sunt extrem de puține exemple de legi universale, unele dintre acestea fiind excelent prezentate de Zamfir (1990). Legea amenințare/coeziune, legea ierarhie/diferențiere a satisfacției muncii, legea productivitate/stratificare constituie exemple foarte concludente. Dintre acestea, consider că prima (legea amenințare/coeziune) este capabilă să treacă testul timpului. Și celelalte două legi par destul de plauzibile pentru societatea de astăzi, rămâne însă de văzut cum va arăta societatea viitorului, atât sub aspectul stratificării cât și sub aspectul ierarhiilor artificiale.

În condiții de eterogenitate, formularea unor enunțuri cu caracter universal nu se poate face decât prin lărgirea ariei de cuprindere a conceptelor utilizate, fapt care generează însă un efect pervers: enunțurile vor fi atât de generale încât nu vor mai fi specifice pentru nici o societate, rămânând în sfera unor teorii abstracte fără nici o aplicabilitate practică de natură să incite interesul particular al unei societăți sau alta.

În absența legilor universale, sociologiei nu-i rămâne de făcut altceva decât să se concentreze pe fenomene locale, încercând să producă enunțuri utile pentru decidenții societăților particulare. Aceste enunțuri însă nu se realizează prin simpla analiză a unei singure societăți, ci prin efectuarea de comparații repetate. În fapt, cele mai multe dintre teoriile sociologiei actuale sunt enunțate pentru contexte particulare, pe diverse criterii: diferențiere geografică (există teorii specifice pentru America Latină, dar și pentru Europa de Est), diferențiere de gen, diferențe de abordare teoretică (funcționalist, structuralist, etc.). Nu se pot analiza întotdeauna un număr suficient de cazuri pentru a trage concluzii statistice, însă analiza comparativă a cazurilor similare sau diferite este una dintre cele mai bune strategii de cercetare socială, întrucât ceea ce este constant în producerea fenomenului va apărea împreună cu acesta.

### 3.2 Pentru a înțelege mai bine propria societate

Fără a avea o pretenție de generalitate, aş putea să afirm că în facultățile de științe sociale în general și de sociologie în particular, studentul primește o formare specială în modul de analiză a societății, în care o primă învățatură se centrează în jurul cuvântului “*depinde*”. Studentul este învățat să renunțe la cunoașterea comună, iar multe dintre preconcepțiile cu care ajunge în facultate sunt demontate bucată cu bucată, unele demascate ca fiind nici mai mult nici mai puțin decât false, ori simple prejudecăți. Lucruri pe care le consideră sfinte devin peste noapte relative; interpretarea acestora încetează să mai fie naturală dintr-o singură perspectivă, percepția acestora depinzând de conceptele prin care este filtrată realitatea. Cum se poate explica fenomenul X? Răspuns: *depinde* din ce punct de vedere îl analizăm.

Studentul mai învață apoi că istoria este relativă și că mai mult, nu reprezintă întotdeauna ceea ce s-a întâmplat cu adevărat ci doar o interpretare asupra evenimentelor importante, marcante. Întrucât istorici diferiți pot găsi interpretări diferite, rezultatul este o permanentă reconstrucție a istoriei, de cele mai multe ori în funcție de interesele politice ale momentului.

Următorul pas în formare este faptul că sociologul trebuie să fie obiectiv în legătură cu interpretarea fenomenelor sociale, întrucât explicațiile pot fi diferite în funcție de unghiul de vedere adoptat; mai târziu, neutralitatea axiologică se dovedește a fi nu doar imposibilă, dar și indezirabilă întrucât studierea unei societăți nu poate fi făcută decât utilizând sistemul de valori acceptat al societății studiate. În măsura în care studentul mai beneficiază și de o bursă în străinătate, mediul natural de viață i se schimbă pentru o perioadă; încercarea de adaptare socială în noul mediu aproape că îl obligă la analiză, comparând fiecare aspect cu cele cunoscute din țară.

Multe dintre noile experiențe i se vor părea ciudate și îl vor intriga; probabil că nu va reuși să înțeleagă perfect societatea pe care o vizitează, însă atunci când revine în țară începe să perceapă altfel relativitatea reprezentărilor pe care le avea înainte despre societatea natală. Încercând să înțeleagă *cealaltă* societate, ajunge să înțeleagă mai bine propria societate. Experiențele prin care trece studentul ipotetic sunt similare cu cele prezentate, într-un alt context, de către Dogan și Pelassy (1990), care au orientat discuția, pe lângă înțelegerea propriei societăți, și către evitarea etnocentrismului generat de amenințarea inconștientă reprezentată de expansiunea altor societăți.

Nu toți oamenii beneficiază de experiența expunerii îndelungate la o altă societate; noul filtru prin care studentul ipotetic privește ulterior

realitatea se dovedește a fi diferit, de exemplu, de reprezentările părinților săi, la care începe să descopere stereotipii cu privire la politică, istorie, relațiile cu țările vecine sau cele de altă religie, în general despre relațiile cu ceilalți, cu “străinii”. Același mecanism este valabil pentru majoritatea populației din orice țară. Începând de la gintă și triburi, toate societățile s-au format prin respingerea celui străin, a celui necunoscut, de care ne temem atât de mult încât aproape toate momentele importante ale istoriei mondiale sunt marcate de violență.

Formarea societăților este centrată pe viziunea *celuilalt*; poate că tocmai de aceea legea “amenințarea externă crește coeziunea internă” este una dintre foarte puținele legi universale din sociologie. Fără excepție, o societate oricât de eterogenă manifestă o creștere a omogenității atunci când amenințarea externă devine vizibilă. Schmitt (1996) a formulat o întregă teorie care stipulează că trebuie să existe un inamic pentru ca o societate să poată exista. Inamicul nu trebuie neapărat să fie o altă țară ci poate fi o idee (cum este terorismul în societatea de astăzi) sau chiar o persoană (de exemplu, Osama Bin Laden pentru Statele Unite).

Nu ne putem înțelege propria societate decât după ce o comparăm cu altele. Înțelegerea și acceptarea diversității sunt unele dintre cele mai sigure metode de a combate etnocentrismul și respingerea automată a celorlalți. Modalitatea *noastră* de a vedea lucrurile nu este nici singura modalitate posibilă, nici cea mai corectă dintre toate. Lucrul acesta nu înseamnă neapărat negarea, anularea propriei societăți; înțelegerea faptului că nu avem întotdeauna dreptate nu înseamnă neapărat că *ceilalți* au dreptate, ci doar faptul că punctele de vedere sunt diferite. Percepțiile sociale se reproduc de la o generație la alta, însă unele dintre acestea ne limitează câmpul vizual și implicit capacitatea de a ne schimba punctul de vedere.

Punctele de vedere diferite sunt benefice pentru cunoaștere; dacă toți oamenii ar gândi la fel, nu ar exista nici un fel de progres. Se spune că un punct de vedere diferit este un punct în care nu am fost încă; munții se văd altfel de la bază și din vârful. Comparația ne ajută să evităm false percepții, printr-o simplă operațiune a schimbării de optică, a schimbării punctelor de vedere.

Această perspectivă este fundamentală pentru studiul sociologic, iar în opinia mea trebuie completată cu studiul istoric al acelorași societăți, în timp. Dimensiunea temporală este rareori atinsă în cercetarea sociologică, însă transformarea istorică a unui singur caz poate să spună mult mai multe lucruri decât am putea crede. Din acest motiv, cercetarea sociologică trebuie aproape obligatoriu combinată cu cercetarea istorică.

### 3.3 Pentru a infirma ipoteze

În secțiunea 3.1 am arătat că analiza comparativă poate fi utilizată pentru formularea unor enunțuri sociologice. Unele dintre aceste enunțuri se pot transforma în legi, altele pot să rămână la nivelul de ipoteze care trebuie verificate empiric.

În fapt, orice enunț sociologic poate fi privit ca pe o ipoteză. Orice mod de gândire dominant, orice paradigmă se dovedește până la urmă inadecvată în urma numeroaselor testări empirice, iar teoria se modifică pentru a se acomoda cu noile dovezi. Rămâne valabilă acea teorie, acea paradigmă care rezistă la cele mai multe verificări empirice. Așadar, nici o ipoteză nu este pe deplin dovedită (nici măcar în științele dure cum este fizica) ci așteaptă doar să fie infirmată de noi dovezi.

În cercetarea socială, demonstrarea teoriilor este cu atât mai dificilă. Nici o metodologie utilizată în prezent (cantitativă ori calitativă) nu poate avea pretenția unei rezolvări deterministe (care să ajungă întotdeauna la același rezultat, vezi secțiunea 2.6.4) a enigmelor din viața socială.

Nici analiza comparativă nu are o astfel de pretenție, cu alte cuvinte este foarte dificil de utilizat pentru *confirmarea* completă a ipotezelor, însă își găsește o excelentă utilizare în sensul contrar, al *infirmării* ipotezelor. Metoda de infirmare este foarte simplă de înțeles iar utilizarea este facilă, fără a fi nevoie de cunoștințe metodologice avansate. În secțiunea 2.3 sunt prezentate canoanele inductive ale lui Mill, iar aici mă voi referi la metoda concordanței.

Pe scurt, o condiție cauzală poate fi eliminată dacă nu concordă cu producerea efectului. Tabelul 2.1 prezintă în mod schematic această idee: în toate instanțele analizate (caz 1, caz 2, caz n) fenomenul Y se produce, însă condiția cauzală A este prezentă doar în cazul 1 și nu în celelalte. Cu alte cuvinte, dacă nu este prezentă de fiecare dată când se produce fenomenul, nu poate fi considerată o cauză a fenomenului respectiv. O altă ilustrare a acestei metode este prezentată în tabelul 5.6, în care condiția cauzală B nu este prezentă în ambele cazuri analizate, fiind eliminată în soluția finală.

Mutând acest raționament din sfera teoreticului în cea practică, poate fi luat ca exemplu situația României după căderea sistemului comunist. De-a lungul a mai mult de 10 ani după revoluție, România a înregistrat puternice căderi economice marcate de inflație și de scăderea nivelului de trai. În această perioadă de tranziție către o economie de piață, segmente întregi de populație s-au bazat pe ajutorul statului, în a cărui atribuții intrau atât necesarele reforme economice cât și protecția socială a populației, puternic afectată de recesiunea economică.

Aceste două atribuții ale statului sunt situate în mod tradițional la poli opuși privind politicile aplicate, protecția socială fiind poziționată în polul stâng, socialist al spectrului politic iar reformele economice spre polul drept, liberal. Protecția socială a avut mult de suferit, în condiții de inflație galopantă, ajutoarele oferite de stat dovedindu-se incapabile să mențină un standard acceptabil de viață. Explicația oficială a acestui fapt a fost centrată în jurul căderii economice: cu alte cuvinte, scăderea protecției sociale a fost datorată căderii economice.

Conform acestei ipoteze, toate țările care manifestă o cădere economică trebuie să manifeste în același timp și o scădere proporțională a protecției sociale. Testarea ipotezei poate fi făcută prin comparația cu o țară aflată în condiții similare (încercând așadar, pe cât posibil, controlarea altor factori cauzali). Bulgaria este un exemplu comparativ foarte bun: a ieșit în același moment cu România din spațiul comunist și a suferit, ca și România, o masivă cădere economică în următorii ani, recesiunea din Bulgaria fiind chiar mult mai dură decât cea din România.

Dacă explicația oficială românească ar fi adevărată, Bulgaria ar fi trebuit să manifeste de asemenea o cădere masivă a protecției sociale, fapt care nu s-a întâmplat în aceeași măsură ca în România. Deși recesiunea economică a fost mai dură în Bulgaria, nivelul protecției sociale a rămas la un nivel ridicat, fapt care infirmă explicația românească. Dimpotrivă, nivelul protecției sociale este un factor care se decide la nivel politic, fiind o opțiune exclusivă a Guvernului.

Așadar, factorul politic este adevărata explicație pentru căderea protecției sociale, într-o măsură mult mai mare decât căderea economică. Acesta este un exemplu simplu și intuitiv pentru utilizarea comparației în testarea unei ipoteze care se referă la un fenomen social, fără a utiliza metode cantitative bazate pe mii de cazuri ci un simplu raționament logic utilizând doar două cazuri: România și Bulgaria.

Este o aplicare sugestivă a strategiei MSDO (*Most Similar, Different Outcome*), discutate în secțiunea 2.5. Ambele sunt țări din Europa de Est, ambele cu o istorie recentă de aproape 50 de ani de regim comunist, ambele cu o istorie mai îndepărtată de ocupație otomană, etc. Studiarea a două cazuri asemănătoare poate să fie profund relevantă pentru explicarea unor traiectorii diferite din prezentul apropiat.

În acest fel, analiza comparativă poate fi utilizată confirmarea, dar și pentru infirmarea unor ipoteze din cercetarea socială, eliminând astfel preconcepții și prejudecăți cu privire la funcționarea unui anumit sistem social.

### 3.4 Pentru validarea rezultatelor de cercetare

Operațiunea inversă celei de infirmare este cea de validare a unei ipoteze. Orice teorie se testează empiric (având așa numitul atribut de falsificabilitate). De obicei, testarea unei ipoteze se realizează prin intermediul unei cercetări sociale.

O verificare extrem de severă a rezultatelor de cercetare se referă la validitatea și fidelitatea instrumentului utilizat. Un instrument este valid dacă măsoară ceea ce trebuie să măsoare, în timp ce fidelitatea arată dacă instrumentul măsoară bine ceea ce trebuie să măsoare. Dacă instrumentul de cercetare nu este valid, rezultatele obținute sunt în mod necesar invalide și invers.

Operațiunea de verificare a validității rezultatelor de cercetare se realizează de cele mai multe ori prin comparații, atât sistematice cât și nesistematice. De exemplu, în analiza cantitativă una dintre operațiunile obișnuite în procesul de analiză este validarea eșantionului, care trebuie să verifice dacă eșantionul are o distribuție asemănătoare cu cea din populație, pe variabilele socio-demografice.

Cu alte cuvinte, eșantionul este comparat cu populația pentru a decide dacă are proporții asemănătoare. Mărimea diverselor straturi din populație trebuie să se regăsească în aceeași proporție în eșantion. De exemplu, proporțiile diverselor etnii din populație trebuie să fie asemănătoare în eșantion, în caz contrar unele dintre acestea fiind suprareprezentate iar altele subreprezentate. Același lucru pentru structura pe vârste a populației, ori pentru distribuția urban-rural, în general pentru toate variabilele pentru care există date oficiale provenite din recensăminte. În cazul în care structura eșantionului nu seamănă cu cea din populație, spunem că eșantionul nu este reprezentativ (nu reprezintă întreaga populație, unele categorii fiind subreprezentate).

Validarea eșantionului se poate face așadar prin comparații sistematice cu ajutorul testelor de semnificație; se testează severitatea deplasării eșantionului de la structura populației, unele devieri fiind semnificative în timp ce altele pot fi considerate minore. În cazul variabilelor cantitative (cum este de exemplu vârsta măsurată în ani împliniți) se poate utiliza familia testelor  $t$  iar pentru variabile calitative (cum este de exemplu structura pe etnii) se poate utiliza familia testelor  $\chi^2$ .

Validarea eșantioanelor este simplă dacă există puncte clare de referință (parametrii din populație, prezentați în anuarele statistice). În cazul altor variabile, cum sunt de exemplu cele de atitudine, lucrurile nu mai sunt atât de simple întrucât situația din populație este necunoscută.



În aceste cazuri, se utilizează comparația cu alte cercetări: dacă cercetări sociale similare au generat rezultate similare, atunci rezultatele se confirmă reciproc. Cu alte cuvinte, una dintre metodele preferate de validare a rezultatelor este comparația cu rezultatele altor cercetări.

În analiza calitativă lucrurile se petrec în mod similar. Dacă două sau mai multe analize calitative ajung la concluzii de cercetare asemănătoare, există motive puternice pentru validarea reciprocă. Este perfect posibil ca toate cercetările să aibă bias-uri (atât ale instrumentelor cantitative cât și ale interpretărilor calitative) însă acestea vor fi demonstrate mai târziu, atunci când alte cercetări vor produce o schimbare de optică. Este suficient însă de subliniat faptul că validarea rezultatelor de cercetare se poate realiza cu ajutorul comparațiilor.

Există mai multe metode de validare utilizate în măsurarea socială, însă pentru scopul acestei prezentări subliniez utilizarea comparației. Cea mai simplă formă de validare, chiar în absența unor cercetări suplimentare este validitatea aparentă (în engl. *face validity*), în care experți din domeniu apreciază corectitudinea rezultatelor unei cercetări în raport cu întreaga lor expertiză: rezultatele *par* să fie valide (într-o comparație implicită cu întreaga experiență a expertului). Validarea mai poate fi obținută cu metoda triangulației: un singur studiu (fie cantitativ sau calitativ) nu este suficient pentru a demonstra niște rezultate valide, însă dacă mai multe studii ajung în mod sistematic la aceleași concluzii, se validează reciproc.

Comparația este esențială pentru stabilirea validității unor rezultate. Nimeni nu poate să spună cu precizie matematică dacă anumite rezultate sunt valide sau nu, iar în absența unor formule precise de evaluare, singura soluție disponibilă este compararea rezultatelor din cercetări diferite.

### 3.5 Pentru a rafina concepte și modele teoretice

Viața socială nu poate fi percepută în mod direct, doar prin observație. Unul și același fenomen este interpretat în mod diferit de actori diferiți, de parcă realitatea nu ar fi una singură. Martori oculari ai aceluiași eveniment pot să povestească în moduri total diferite, în funcție de trăiri personale și percepții.

Chiar și obiecte fizice pot să fie percepute în diverse moduri, atunci când intră în joc iluziile optice. Ceea ce vedem nu este exact ceea ce vedem; nimeni nu poate să spună dacă ochii tuturor oamenilor filtrează lungimea de undă în exact aceeași manieră, așa încât culoarea este percepută în mod infinitezimal diferit de către fiecare persoană în parte. Culoarea este

așadar o convenție, după cum este orice percepție asupra realității.

Dacă acest lucru este valabil pentru obiecte fizice, atunci este cu atât mai valabil pentru perceperea fenomenelor sociale, atât de relativă de la o persoană la alta, de la o comunitate la alta, de la o națiune la alta. Pentru a cunoaște realitatea așa cum este, ne-ar trebui o pereche de “ochi sociali”, ori din acest punct de vedere suntem ca niște nevăzători, care se bazează pe alte simțuri pentru a percepe mediul înconjurător. Unele dintre modalitățile frecvent utilizate pentru a cunoaște realitatea sunt descrierile abstracte sau modelele teoretice. Acestea îndeplinesc rolul unei hărți mentale care ne ajută să navigăm prin realitate, să o cunoaștem, la fel cum liliicii sau delfinii utilizează ecolocația pentru a genera hărți ale locurilor prin care se mișcă. Pentru aceste vietăți, harta este constituită din ecouri ale undelor trimise înainte, însă pentru oameni hărțile mentale sunt construite cu ajutorul cuvintelor, mai exact cu ajutorul conceptelor.

Din acest punct de vedere “ochii” noștri sunt net inferiori, întrucât hărțile mentale pot arăta extrem de diferit nu atât în funcție de persoana care construiește harta, cât mai ales în funcție de conceptele utilizate pentru a construi modelul teoretic. Construirea conceptelor este un fel de design, fiecare încercând să fie cât mai original și să inventeze acea modalitate de a descrie realitatea care să fie acceptată de către comunitate pe post de paradigmă.

Conceptele utilizate într-o perioadă sau alta sunt așadar relative, unele rezistând în timp iar altele fiind abandonate imediat cum apar. Important este faptul că acestea sunt mediate, regândite, rafinate într-un proces fundamental comparativ. Dogan și Pelassy (1990, pp.24-35) tratează aceste aspecte într-un mod absolut remarcabil prin claritatea expunerii și finețea argumentării. Plecând de la expresia “*verbum dat esse rei*” a filozofiei scolastice, autorii prezintă o multitudine de exemple ale conceptelor care sunt sau au fost utilizate pentru a descrie realitatea într-un anumit mod.

Cercetarea socială și rezultatele acesteia depind într-o modalitate foarte înaltă de cadrul teoretic prin care se desfășoară. De exemplu, într-un fel arată o analiză a mediului politic dintr-o țară în tranziție utilizând conceptul de “dezvoltare politică” și într-un mod subtil diferit utilizând conceptul de “schimbare politică”.

Într-un fel este concepută o cercetare care se axează în jurul “combaterii sărăciei”, într-un cu totul alt fel este concepută în jurul conceptului de “integrare socială” a persoanelor sărace și într-un mod iarăși diferit în jurul conceptului de “inclusiune socială” a persoanelor excluse de la accesul la unele servicii sociale fundamentale.

Pe măsură ce scanarea realității are loc din perspectiva unui model sau altul, conceptele sunt gradual rafinate. În funcție de conceptul dominant utilizat, modelul teoretic (cadru analitic al cercetării) se rafinează la rândul său. Științele sociale din perioada actuală prezintă schimbări permanente a modelelor teoretice, determinate de posibilitatea unei cunoașteri fără precedent a realității empirice. Un model teoretic care se potrivește doar într-o anumită societate nu este acceptat de comunitatea academică extinsă, globală. Modelele concepute în “eprubetele” școlilor puternice de sociologie vestică sunt testate pe rând în societăți diverse, fiind abandonate în momentul în care nu pot să explice realități diverse în favoarea altora, mai ingenioase.

Conceptul de subdezvoltare, arată Dogan și Pelassy, este esențialmente de natură comparativă, fiind născut din analiza mai multor societăți aflate pe poziții diferite într-o axă a dezvoltării. Anomia socială a lui Durkheim este născută la fel din comparație, întrucât nimeni nu poate să spună dacă o societate este anomică dacă nu o compară cu alta.

Conceptele sunt importante, ele pot fi utilizate în crearea acelor hărți mentale, dar și pentru a crea o întreagă realitate dacă sunt utilizate la nivel de masă. Mass-media au un rol esențial în promovarea unei anumite imagini despre lumea înconjurătoare, riscul cel mai înalt fiind prezentat de manipularea opiniei publice, întrucât pot fi promovate false adevăruri. Sunt de preferat acele concepte atât neutre cât și precise în descrierea realității, care să evite stereotipurile sau prejudecățile.

Similar cu situația conceptelor, cadrele analitice sunt supuse aceluiași proces de reconstruire graduală, pe măsură ce noi comparații arată utilitatea unui model sau a altuia. Într-un fel arată societatea analizată prin teoria structuralismului, în alt fel prin teoria funcționalismului și în cu totul alt fel prin teoria (neo)instituționalismului. Aceste teorii sunt cu atât mai importante din perspectivă comparativă, cu cât structuri diferite pot îndeplini funcții asemănătoare și vice-versa, structuri asemănătoare pot îndeplini funcții diferite. Din punct de vedere comparativ, importante sunt echivalențele funcționale pe care diverse structuri le prezintă și care nu pot fi identificate decât în urma unor analize comparative foarte atente.

Oricare ar fi natura modelului teoretic, acesta trebuie să suporte multiple testări prin comparații empirice, altfel riscă să fie ancorat într-o dimensiune conceptuală atât de abstractă încât să nu mai aibă nici o legătură cu realitatea. Dimpotrivă, fiecare nouă testare empirică aduce o nouă interpretare, prin comparație cu celelalte instanțe analizate, astfel încât modelul teoretic se schimbă în mod perpetuu pentru a se acomoda schimbării continue a mediului social.

### 3.6 Pentru a identifica șabloane sistematice

Studierea fenomenelor sociale dispune rareori de suficiente date pentru a elabora concluzii statistice solide. În cele mai multe cazuri, în special dacă unitățile de analiză se află la un nivel înalt de agregare, numărul de cazuri disponibil pentru analiză este limitat. Mai mult, în analiza calitativă pură numărul unităților cercetate este limitat, întrucât nu se pot face studii în profunzime pe un număr ridicat de cazuri.

Sarcina celui care analizează un fenomen este aceea de a lega informația teoretică de evidența empirică; este important de subliniat că această legătură nu trebuie confundată cu *testarea* ipotezelor teoretice, atât de des menționată în știință în general și în științele exacte în special. Testarea ipotezelor, în special în ultimele decenii, se face cu instrumente statistice, ori în analiza calitativă pe cazuri puține acest lucru este imposibil. Dimpotrivă, strategia cercetătorului este aceea de a studia în amănunțime anumite cazuri relevante, iar în etapa de analiză a datelor culese strategia adoptată este aceea de a identifica (precum un detectiv) anumite similarități între cazuri care apar în mod sistematic, presupuse a fi determinante pentru apariția fenomenului.

Dacă anumiți factori sunt într-adevăr responsabili pentru producerea fenomenului (fie singular, fie – așa cum se întâmplă cel mai des – în combinații cauzale), atunci ei ar trebui să se regăsească în forme mai mult sau mai puțin asemănătoare în evidența empirică. Strategia este una inductivă, plecând de la datele culese spre concluzii teoretice, însă analistul nu se bazează doar pe evidența empirică ci mai degrabă pe raționamente deductive care îl ajută să construiască o reprezentare coerentă a mecanismului de producere a fenomenului de interes.

Există un principiu valabil, conform căruia cauzele importante apar în mod sistematic. Este una dintre asumțiile de bază a analizei (calitativ)-comparative, care funcționează pe baza observării apariției sistematice a unor combinații de condiții cauzale, eliminând pe acelea care nu apar constant odată cu producerea fenomenului de interes. Această secțiune este doar o introducere către nucleul lucrării, discutat cu amănunțime în secțiunea 2.1 dedicat analizei cauzale și mai ales în secțiunea 5.1 dedicată metodologiei analizei calitativ-comparative, utilizată în exact acest scop: pentru a identifica șabloane (*pattern-uri*) de similaritate prin eliminarea diferențelor.

## Capitolul 4

# Noțiuni de teoria mulțimilor

Analiza calitativ-comparativă se dovedește a fi stabilă și solidă, atât din punct de vedere teoretic/metodologic dar și prin soluțiile software existente (mai exact a algoritmilor pe baza cărora sunt construite aceste programe software). Aflați într-un plin proces de dezvoltare inovativă, cercetătorii QCA din științele sociale și politice se întrec în a oferi inovații și soluții la probleme particulare pentru care analiza clasică nu poate oferi răspunsuri clare și echivoce. Acest capitol prezintă direcțiile majore în care se dezvoltă această analiză bazată pe minimizare, de la varianta clasică la cele moderne.

Într-o versiune inițială a analizei, Ragin (1987) a intitulat entitățile cu care operează drept “binare” întrucât a codificat prezența și absența acestora cu cifrele 1 și 0. Utilizând aceste codificări, Ragin a introdus pe arena științelor sociale o utilizare originală a algebrei boolene.

Ca urmare a criticilor ulterioare, comunitatea metodologică a diferențiat între așa-numitele “mulțimi distincte” (în engleză *crisp sets* – csQCA) și “mulțimi vagi” (în engleză *fuzzy sets* – fsQCA), cu trecere prin mvQCA (*multi-value QCA*). Metodologia operează masiv și aproape exclusiv cu conceptul de “mulțimi”.

În matematică, denumirea de “mulțime” este foarte des utilizată, de exemplu  $\mathbb{R}$  - mulțimea numerelor reale, sau  $\mathbb{N}$  - mulțimea numerelor naturale etc.

În științele sociale se utilizează mai degrabă denumirea de “categorii”, organizate pe variabile (de unde și denumirea de variabile categoriale). O astfel de variabilă este mediul de rezidență, iar din perspectiva teoriei mulțimilor categoria “Urban” ar putea fi reinterpretată drept “mulțimea oamenilor care trăiesc în mediul urban”, iar categoria “Rural” ar putea fi reinterpretată drept “mulțimea oamenilor care trăiesc în mediul rural”.

Metodologia de cercetare din cadrul științelor sociale se bazează foarte mult pe comparația dintre aceste categorii (de exemplu, venitul mediu din urban cu venitul mediu din rural), iar o metodologie similară se poate orienta către compararea mulțimilor.

## 4.1 Sistemul binar și algebra booleană

... pentru a găsi soluții la întrebări din teoria probabilităților, nici o metodă generală nu poate fi stabilită fără a recunoaște în mod explicit [...] acele legi universale de gândire care se află la baza oricărei rațiuni

George Boole

Citatul de mai sus face referire la niște legi logice, care în combinație unele cu altele pot să transpună o conversație dintr-un limbaj natural într-un limbaj matematic. Acest tip special de limbaj utilizează doar două valori: *adevărat* și *fals* (în engleză *true* și *false*).

Istoria acestor valori este mult mai veche, datând încă din începuturile culturii chinezești. Dualitatea naturii, exprimată prin faimoșii termeni Yin și Yang, a fost urmărită încă de pe vremea chinezilor antici care au perceput lumea ca pe o continuă îmbinare a acestor două elemente, toate lucrurile existând prin influența mai mare sau mai mică a unuia dintre acestea (însă întotdeauna în combinație cu celălalt).

O gândire similară a avut-o filozoful și matematicianul Gottfried Leibniz (a cărui biografie a fost scrisă recent de către Aiton, 1985), care credea atât de mult în puterea simbolurilor pentru înțelegerea umană încât a inventat o matematică specială, binară, care nu utiliza decât două valori: 1 și 0. Leibniz și-a dedicat întreaga viață acestui sistem, transformându-l în ultimii lui ani într-un sistem mistico-religios în care valoarea 1 reprezintă binele iar valoarea 0 reprezintă răul. Într-un mod similar cu gândirea chinezească, în sistemul filozofic al lui Leibniz lumea era o continuă luptă între bine și rău, însă termenii chinezești de Yin și Yang au o complexitate mult mai mare și nu pot fi reduși doar la aceste două interpretări (ei mai semnifică și alte dualități cum sunt: da și nu, pornire și oprire, întuneric și lumină, masculin și feminin etc).

Filozofia lui Leibniz l-a determinat să creadă că matematica binară are o origine divină. El a creat (la începutul secolului XVIII) metode speciale de transformare a numerelor din sistemul zecimal în sistemul binar, fără să aibă nici un fel de aplicație concretă. A construit chiar o mașinărie specială care făcea această conversie în mod automat, însă comunitatea

academică a vremii i-a ignorat munca iar aplicațiile au trebuit să aștepte peste două sute de ani până la apariția calculatoarelor, sistemul binar având astăzi o răspândire la care doar Leibniz putea să viseze.

La aproape 150 de ani după Leibniz, pe la mijlocul secolului XIX, un alt mare matematician pe nume George Boole a preluat sistemul binar și l-a rafinat până la un nivel utilizabil atât în logică cât și în matematică. Comunitatea academică a vremii a neglijat această matematică iar primele ei aplicații au apărut abia câteva decenii mai târziu la MIT - Massachusetts Institute of Technology în Statele Unite.

Algebra specială pe care a creat-o îi poartă numele, utilizând (la fel ca sistemul binar al lui Leibniz) doar două valori: 1 și 0. Spre deosebire de Leibniz, algebra booleană utilizează valorile binare 1 și 0 ca substitut pentru *adevărat* și *fals*.

## 4.2 Tipuri de mulțimi utilizate în analiza comparativă

Spre finele secolului XIX, interesat de anumite proprietăți ale seriilor infinite, matematicianul și filozoful german Georg Cantor a creat o teorie a mulțimilor abstracte care avea să schimbe în mod radical fundațiile matematicii (Dauben, 1979).

Versiunea inițială (denumită și teoria naivă a mulțimilor) a fost completată și extinsă ulterior, iar varianta modernă a teoriei specifică mai multe tipuri de mulțimi și axiome decât cele menționate de Cantor. Pe scurt însă, orice mulțime poate fi definită ca o colecție de obiecte care împărtășesc o proprietate comună.

Obiectele din interiorul unei mulțimi se numesc elemente, iar fiecare asemenea obiect este unic (un obiect nu poate să facă parte dintr-o mulțime mai mult de o singură dată). În matematică au fost introduse un număr de concepte și proprietăți legate de mulțimi: finite, infinite, intersectabile, disjuncte, egale, echivalente, mulțimea vidă etc.

Fără pretenția unei utilizări formale a acestor diferențieri, poate fi observat faptul că metodologia de cercetare din științele sociale (în general) și cea comparativă (în special) utilizează un vocabular semi-formal cu mulți termeni din teoria mulțimilor. În cele mai multe cazuri nici nu se menționează explicit utilizarea acestor termeni, însă există sinonimii destul de clare, în special prin utilizarea categoriilor în acțiunile de segmentare a populației (urban / rural, bărbați / femei, studii primare / medii / superioare, etc.). Împărțirea populației pe categorii este una dintre cele mai frecvente activități de cercetare descriptivă din științele sociale.

În mod formal, se menționează în mod explicit două mari familii:

1. mulțimi distincte;
  - bivalente
  - multivalente
2. mulțimi vagi.

Prima familie, cea a mulțimilor distincte, este creația originală a lui Cantor. Un element este fie în interiorul, fie în exteriorul unei astfel de mulțimi. O persoană se află în localitatea X, ori în afara ei (nu poate fi în același timp și în interiorul și în exteriorul localității); o persoană este fie în viață fie decedată; un buton este apăsat sau nu este apăsat etc.

Această familie a mulțimilor distincte conține la rândul ei două subcategorii: cele bivalente conțin doar două valori de adevăr, în timp ce mulțimile multivalente conțin mai mult de două valori (un echivalent în științele sociale ar fi variabilele nominale sau ordinale cu mai multe categorii).

Familia mulțimilor vagi se diferențiază față de cele distincte prin faptul că poate avea un infinit de valori posibile, orice element având o includere mai mare sau mai mică în orice mulțime, *pornind* de la valoarea 0 (deloc inclus în mulțime), *până* la maxim valoarea 1 (complet inclus în mulțime).

#### 4.2.1 Mulțimi distincte bivalente

Mulțimile distincte bivalente sunt colecții de elemente bine definite (distincte), în sensul că acestea fie aparțin fie nu aparțin unei mulțimi (au sau nu au o anumită proprietate). Într-un sistem formal de notație, orice astfel de mulțime poate fi reprezentată prin enumerarea tuturor elementelor distincte, componente ale acesteia:

$$A = \{a_1, a_2, \dots, a_n\}$$

Sunt tipul de mulțimi pe care oamenii le utilizează în mod natural atunci când poartă o conversație: un eveniment fie s-a întâmplat, fie nu s-a întâmplat. În accepțiunea biologică a conceptului de *mamă*, o femeie este mamă doar dacă a dat naștere unui copil; dacă nu a dat naștere niciunui copil atunci nu este mamă.

Mulțimea tuturor elementelor (din toate mulțimile posibile) este denumită *mulțimea universală* (notată cu U), iar elementele acesteia sunt  $x_{1\dots n} \in U$ .



Toate celelalte mulțimi sunt incluse în acest univers, sunt submulțimi ale acestuia (altfel spus, toate sunt mulțimi ale universului U).

Pentru fiecare element al universului U, se poate răspunde cu “da” (sau “adevărat”) dacă elementul aparține unei mulțimi A, și cu “nu” (sau “fals”) dacă elementul respectiv nu aparține mulțimii A. În teoria clasică a mulțimilor, elementele dobândesc doar două valori de adevăr: 0 (“fals”) și 1 (“adevărat”). În notația formală, poate fi definită o funcție de apartenență care atribuie aceste două valori:

$$\mu_A(x) = \begin{cases} 0 & \text{dacă } x \notin A \\ 1 & \text{dacă } x \in A \end{cases}$$

Să presupunem că întregul univers U este constituit din toate numerele de pe cele șase fețe ale unui zar:

$$U = \{1, 2, 3, 4, 5, 6\}$$

Dacă P este mulțimea numerelor pare, atunci pentru fiecare dintre elementele universului U se poate atribui o valoare de adevăr:

$$\mu_U(x) = \{0, 1, 0, 1, 0, 1\}$$

asa încât:

$$P = \{2, 4, 6\}$$

Mulțimile distincte sunt ceea ce în științele sociale se numesc *categorii* mutual exclusive. De exemplu, un trandafir face parte din categoria “flori”, în timp ce un cal nu face parte din această categorie.

Unul și același element poate să facă parte în mod simultan din mai multe (sub)mulțimi. De exemplu, dacă A este mulțimea mamelor, iar B este mulțimea femeilor, atunci o mamă face parte atât din categoria A cât și din categoria B. Mulțimea mamelor este o submulțime a persoanelor care fac parte din categoria femeilor.

Pentru oricare două mulțimi A și B ale universului U, putem spune că A este o submulțime a lui B dacă și numai dacă atunci când un element aparține lui A rezultă că acel element este și un element a lui B:

$$A \subset B : \{x \in A \Rightarrow x \in B\}$$

O mulțime care nu conține nici un element se numește *mulțimea vidă* (notată cu semnul  $\emptyset$ ). Mulțimea vidă este un subset al oricărei alteia.

Logica fondată de Aristotel a fost concepută, așadar, într-o manieră bivalentă. O propoziție declarativă are o singură valoare de adevăr, este fie adevărată fie falsă. Acest sistem se bazează pe trei principii (legi), care stau la baza constituirii mulțimilor distincte bivalente:

1. Principiul identității: un obiect este ceea ce este (sau: ceea ce este este). Acest principiu este important pentru că face distincția dintre un obiect și celelalte obiecte (într-o perspectivă psihologică, distincția dintre *sine* și *ceilalți*).
2. Principiul non-contradicției: este imposibil ca un obiect să existe și să nu existe în același timp, sau un fenomen să se întâmple și să nu se întâmple. Ori unul, ori celălalt este adevărat, dar nu ambele.
3. Principiul terțului exclus: o propoziție este fie adevărată, fie negația acesteia este adevărată, nu există o a treia posibilitate în afară de acestea.

#### 4.2.2 Mulțimi distincte multivalente

Logica tradițională, scolastică, este în mod inerent bivalentă; principiul terțului exclus afirmă clar că nu poate exista o a treia posibilitate. Totuși, bivalența a fost pusă sub semnul întrebării de Aristotel însuși, care a formulat un paradox utilizând logica propozițională cu o combinație de expresii logice, testând bivalența în cadrul unor afirmații temporale.

Cu privire la trecut, se pot face afirmații cu valoare de adevăr. Despre un fenomen care s-a întâmplat nu se mai poate spune că nu s-a întâmplat, întrucât valoarea sa de adevăr a devenit atemporală: de îndată ce s-a întâmplat, acest lucru este valabil în orice moment ulterior (a fost adevărat imediat după aceea, în prezent este adevărat, în viitor va fi adevărat).

Logica bivalentă este perfect compatibilă cu trecutul (cu ceea ce cunoaștem deja), însă aplicarea acesteia pentru viitor este dificilă și a condus la paradoxul formulat de Aristotel. Cu privire la viitor, se pot aplica două sisteme de gândire:

1. Determinismul: dacă un lucru trebuie să se întâmple, atunci el se va întâmpla indiferent ce facem noi. Este un sistem similar cu fatalismul, întrucât dacă soarta decide ceea ce se va întâmpla, atunci noi nu avem nici un control asupra propriului destin.

2. Liberul arbitru: noi decidem ceea ce se va întâmpla. Acest sistem admite abilitatea de a alege ceea ce se întâmplă (precum și ceea ce nu se întâmplă), în absența unor constrângeri externe. În acest sistem, este posibil că vom alege să facem un lucru, însă la fel de posibil este să nu facem acel lucru.

Orice eveniment are o succesiune de condiții cauzale (care au și ele propriile lor condiții cauzale, la infinit, până la începutul timpului), și se poate spune că dat fiind acest lanț causal era de așteptat ca evenimentul să se întâmple, ori altfel spus lanțul causal a *determinat* producerea evenimentului.

Dacă ne situăm în viitor, ca și cum s-ar fi întâmplat deja, vom putea spune același lucru despre evenimentul care va fi fost produs, că va fi fost determinat de o serie de condiții cauzale. În această viziune deterministă, producerea evenimentului este atemporală întrucât date fiind condițiile cauzale care s-au întâmplat, se întâmplă și se vor întâmpla, evenimentul se va produce în mod necesar.

Aristotel se referea la o bătălie navală, față de care a făcut următoarea afirmație: “mâine va avea loc o bătălie”. Este imposibil să atribuim o valoare de adevăr acestei propoziții, întrucât fenomenul încă nu a avut loc. După ziua de mâine, această propoziție va fi ori adevărată, dacă bătălia va avea loc, ori falsă, dacă nu va avea loc. Astăzi însă, propoziția nu poate fi nici adevărată nici falsă.

Paradoxul lui Aristotel constă în proiectarea logicii deterministe pentru evenimentele din viitor, fapt care atribuie o valoare de *fals* afirmației că “există liberul arbitru”, ori acest lucru contrazice experienței din prezent care atribuie acestei afirmații o valoare de *adevărat*, încălcând astfel principiul non-contradicției (este imposibil ca un fenomen să fie în același timp și adevărat și fals).

Preocupat de aspectul determinismului, filozoful polonez Łukasiewicz a creat la începutul secolului XX un sistem logic (Borkowski, 1970) care depășește filozofia clasică bivalentă și oferă o soluție pentru paradoxul lui Aristotel. Acest sistem (notat cu  $\mathbb{L}_3$ ) conține nu două ci trei valori de adevăr:

$$\mu_A(x) = \begin{cases} 0 & \text{fals} \\ \frac{1}{2} & \text{nedeterminat (nici adevărat, nici fals), parțial adevărat} \\ 1 & \text{adevărat} \end{cases}$$

Contestată la început, filozofia lui Łukasiewicz a fost acceptată în cele din urmă și a condus la generalizarea sistemului său trivalent și apariția

unor sisteme *multi-valente* (cu  $n$  valori). Mulțimea valorilor de adevăr este obținută prin divizarea uniformă a intervalului  $[0, 1]$  în  $n$  valori de adevăr distincte:

$$\left\{ 0 = \frac{0}{n-1}, \frac{1}{n-1}, \frac{2}{n-1}, \dots, \frac{n-1}{n-1} = 1 \right\}$$

Cu aplicare pe științele sociale, un exemplu potrivit este cel al culorilor semaforului, cu următoarele valori: {roșu, galben, verde}. În logica bivalentă, ar fi fost nevoie de trei funcții de apartenență, corespunzătoare celor trei mulțimi:

- Semafoare de culoare roșie:  $\mu_{SR}(x) = \begin{cases} 0 & \text{nu, } (x \notin SR) \\ 1 & \text{da, } (x \in SR) \end{cases}$
- Semafoare de culoare galbenă:  $\mu_{SG}(x) = \begin{cases} 0 & \text{nu } (x \notin SG) \\ 1 & \text{da, } (x \in SG) \end{cases}$
- Semafoare de culoare verde:  $\mu_{SV}(x) = \begin{cases} 0 & \text{nu, } (x \notin SV) \\ 1 & \text{da, } (x \in SV) \end{cases}$

Într-o logică multivalentă însă, nu este nevoie decât de o singură funcție de apartenență la mulțimea culorilor semafoarelor:

$$\mu_S(x) = \begin{cases} 0, & \text{dacă semaforul este verde} \\ 1, & \text{dacă semaforul este galben} \\ 2, & \text{dacă semaforul este roșu} \end{cases}$$

Mulțimile de mai sus sunt asemănătoare cu variabilele ordinale, valorile fiind echivalente categoriilor ordonate ale acestora. Totuși, mulțimile distincte multivalente nu sunt specifice doar variabilelor ordinale, ci pot fi aplicate în egală măsură și variabilelor nominale, pentru care ordinea categoriilor nu este importantă.

De exemplu, regiuni (R) mutual exclusive pot juca rolul de valori, ordinea acestora fiind puțin importantă:

$$\mu_R(x) = \begin{cases} 0, & \text{regiunea de est} \\ 1, & \text{regiunea de vest} \\ 2, & \text{regiunea de nord} \\ 3, & \text{regiunea de sud} \end{cases}$$

Mulțimile distincte bivalente sunt denumite, în limba engleză, *crisp sets*. În literatura de specialitate QCA (Qualitative Comparative Analysis),

analiza bazată pe acestea a fost notată cu *csQCA*, iar atunci când au fost introduse mulțimile distincte multivalente, acestea au căpătat denumirea de *multi-value sets* și notate cu *mvQCA*.

În fapt, atât *csQCA* cât și *mvQCA* lucrează cu unități distincte (*crisp*). În ambele tipuri de mulțimi valențele sunt mutual exclusive, ceea ce înseamnă că mulțimile distincte bivalente sunt doar un caz particular al mulțimilor distincte multivalente.

Unica diferență dintre acestea este faptul că primul tip de mulțimi are doar două valori distincte, în timp ce al doilea are “cel puțin” două valori, așadar *mvQCA* este o generalizare a *csQCA* (iar *csQCA* este un caz particular al *mvQCA*).

### 4.2.3 Mulțimi vagi

Una dintre particularitățile mulțimilor distincte este faptul că sunt, în general, finite. Cel puțin în științele sociale, variabilele calitative nu au o infinitate de categorii. Pentru cvasi-majoritatea situațiilor de cercetare, universul studiului este împărțit într-un număr relativ restrâns de categorii, pentru a le putea diferenția ușor unele față de celelalte.

Explicația este relativ simplă: conștiința umană este relativ limitată atunci când trebuie să distingă între mai multe categorii simultan. Într-unul dintre cele mai citate articole din istoria psihologiei, Miller (1956) scria despre “Magicul număr șapte, plus sau minus doi: limite ale capacității noastre de a procesa informație”.

Mulțimile distincte bivalente au exact două valori, iar cele multivalente au, de obicei, până la 4 sau 5 valori distincte, mutual exclusive. Variabile cu foarte multe categorii există totuși, de exemplu nivelul de educație ori ocupația respondentului, însă unitatea de analiză pentru cercetarea comparativă este regăsită la un nivel mai înalt de agregare (comunitate, regiuni, națiuni etc.), unde numărul de categorii în care acestea se împart este relativ restrâns.

După generalizarea teoriei lui Łukasiewicz de la 3 ( $\mathbb{L}_3$ ) la  $n$  valori de adevăr ( $\mathbb{L}_n$ ), nu a mai fost decât o chestiune de timp până la extinderea completă a teoriei către o infinitate de valori posibile. În fapt, modelul  $\mathbb{L}_n$  este foarte aproape de mulțimile vagi, întrucât  $n$  poate fi un număr oricât de mare (am putea spune, un număr aproape infinit de mare).

În loc de împărțirea în  $n$  valori distincte, o altă soluție a fost introdusă câteva decenii mai târziu de către matematicianul Lotfi Zadeh, care a introdus în știință conceptul de *mulțimi vagi*. În definiția originală (Zadeh, 1965, p.338), acestea sunt prezentate drept:

... o clasă de obiecte cu un continuum de grade de apartenență. O astfel de mulțime este caracterizată de o funcție de apartenență [...] ce atribuie fiecărui obiect un grad de apartenență variabil între 0 și 1.

Așadar, între orice minim (deloc, zero) și maxim (complet, unu), pot exista oricâte, un continuum, o infinitate de grade de apartenență la o mulțime.

Aproape toate conceptele cu care se operează în științele sociale pot fi exprimate în mulțimi vagi, foarte rare sunt proprietățile care sunt perfect albe sau perfect negre. De obicei, așa cum se spune, “adevărul este undeva la mijloc”.

Un element poate să aparțină “mai mult sau mai puțin” unei mulțimi. De exemplu, o populație nu poate fi considerată pur și simplu bogată ori săracă, ci mai mult sau mai puțin una dintre acestea. În funcție de criteriul de definire a bunăstării, o populație aparține *cu preponderență* uneia sau alteia dintre mulțimi (a celor sărace sau a celor bogate), însă este greu de imaginat că poate să aparțină cu totul doar uneia dintre acestea.

În vorbirea curentă, cuvinte ca “bun”, “fragil”, “cunoscător” etc. reprezintă exemple de noțiuni vagi, întrucât depinde foarte mult de definiția și reprezentarea abstractă a celor care le utilizează. O persoană poate să pară cunoscătoare în fața unui public neavizat, însă dacă discută cu specialiști aceeași persoană poate să se dovedească aproape ignorantă.

Denumirea de mulțimi “vagi” poate fi uneori confundată cu ceva neclar, greu de definit. Termenul apare atunci când nu poate fi stabilit un prag clar pentru stabilirea apartenenței la o categorie sau alta.

Întrucât poate lua orice valoare în intervalul  $[0, 1]$ , este important de subliniat faptul că apartenența la o mulțime nu este același lucru cu probabilitatea ca un eveniment să se întâmple, probabilitate care ia valori în exact același interval. Apartenența la o mulțime (care mai este denumită și *incluere* în mulțime) reprezintă un fapt, un *datum*, chiar dacă este dificilă stabilirea unui prag clar de includere, pe când probabilitatea se referă la ceva care nu s-a întâmplat încă.

Ambele iau valori în intervalul  $[0, 1]$ , motiv pentru care foarte mulți le confundă, însă probabilitatea și includerea într-o mulțime sunt concepte foarte diferite, iar următorul experiment imaginar va face lucrurile mai clare. Să spunem că trebuie să atingem un obiect, despre care știm că are o probabilitate de 1% să fie extrem de fierbinte. Atingând acel obiect de 100 de ori în momente diferite, există o probabilitate (mică, dar există),

să ne provocăm o arsură severă la nivelul mâinii. De cealaltă parte, dacă știm că obiectul are o includere de doar 1% în mulțimea obiectelor extrem de fierbinți, putem să-l atingem de nenumărate ori fără niciun fel de risc, întrucât o includere atât de mică înseamnă că obiectul este aproape rece, așadar nu vom provoca niciodată o arsură gravă prin atingerea aceluiași obiect.

### 4.3 Operații cu mulțimi

Boole (1854) a adus în plus și trei operațiuni de bază care au rămas și astăzi la baza funcționării calculatoarelor. Utilizarea celor trei operațiuni poate fi explicată prin exemple foarte simple, caracterizând rezultatul unei expresii lingvistice fiind adevărat sau fals:

**SAU logic:** “Avuția poate fi obținută prin moștenire sau prin câștig”

Rezultatul acestei expresii este bogăția; dacă aceasta este obținută atunci rezultatul este o valoare de adevăr, în caz contrar rezultatul fiind o valoare de fals. Dacă oricare dintre condițiile “moștenire” și “câștig” sunt prezente (au valoare de adevăr) atunci rezultatul este o valoare de adevăr; dacă niciuna dintre condiții nu este prezentă (nu este adevărată) atunci rezultatul este fals. Alocarea valorilor de adevăr se face prin interogarea fiecărei condiții în parte și a rezultatului; dacă răspunsul este afirmativ avem o valoare de adevăr, iar dacă răspunsul este negativ avem o valoare de fals.

S-a primit moștenire? S-a obținut câștig? S-a obținut bogăție?

DA = Adevărat, NU = Fals

Fals	SAU	Fals	=	Fals
Fals	SAU	Adevărat	=	Adevărat
Adevărat	SAU	Fals	=	Adevărat
Adevărat	SAU	Adevărat	=	Adevărat

Pentru două condiții cauzale, nu există decât patru combinații posibile de condiții cauzale. În algebra booleană, cuvântul SAU poate fi înlocuit cu semnul adunării ‘+’. Înlocuind *Adevărat* cu valoarea 1 și *Fals* cu valoarea 0 obținem următoarele două tabele echivalente, simplificate:

0 SAU 0 = 0	0 + 0 = 0
0 SAU 1 = 1	0 + 1 = 1
1 SAU 0 = 1	1 + 0 = 1
1 SAU 1 = 1	1 + 1 = 1

În cazul operației SAU logic, se poate observa că valoarea de adevăr apare dacă cel puțin una dintre condițiile cauzale este adevărată, indiferent de valoarea celeilalte. Doar în măsura în care ambele condiții cauzale sunt false, valoarea rezultată este falsă.

**ȘI logic:** “În știință, performanțele se obțin prin inteligență și muncă”

În această expresie lingvistică, prima condiție cauzală a performanțelor (rezultatul expresiei logice) este inteligența iar a doua condiție cauzală este munca. Acolo unde există inteligență, există premise pentru performanțe însă nu este de ajuns întrucât acestea nu pot fi obținute decât prin depunerea unui efort susținut (cu alte cuvinte prin muncă). Simpla inteligență nu garantează performanță, după cum o persoană poate să muncească enorm însă fără un anumit nivel de inteligență nu poate performa.

La fel ca la operațiunea precedentă, alocarea valorilor de adevăr și fals se face prin interogarea condițiilor cauzale și a rezultatului:

Există inteligență? Există muncă? Există performanță?

DA = Adevărat, NU = Fals

Fals	ȘI	Fals	=	Fals
Fals	ȘI	Adevărat	=	Fals
Adevărat	ȘI	Fals	=	Fals
Adevărat	ȘI	Adevărat	=	Adevărat

Acolo unde nu există nici inteligență nici muncă, performanțele nu apar. Același rezultat îl avem și în următoarele două situații, în prezența doar a uneia dintre cele două condiții cauzale (muncă fără inteligență, ori inteligență nesusținută de muncă). Doar în ultimul caz, când există atât inteligență cât și muncă, se poate observa prezența performanțelor (valoarea de adevăr pentru rezultatul expresiei lingvistice).

În algebra booleană, cuvântul ȘI poate fi înlocuit cu semnul înmulțirii ‘·’. Înlocuind *Adevărat* cu valoarea 1 și *Fals* cu valoarea 0 obținem următoarele două tabele echivalente, simplificate:

0 ȘI 0 = 0	0 · 0 = 0
0 ȘI 1 = 0	0 · 1 = 0
1 ȘI 0 = 0	1 · 0 = 0
1 ȘI 1 = 1	1 · 1 = 1

În cazul operației ȘI logic, se poate observa că valoarea de adevăr apare doar în măsura în care ambele condiții cauzale sunt adevărate.



Este suficient ca una dintre ele să fie falsă, iar rezultatul dobândește o valoare de fals indiferent de valoarea celeilalte condiții cauzale.

**NEGAȚIA:** “În lumea mamiferelor, toate speciile care nu sunt masculi sunt femele”

Negația este cea mai simplă și mai intuitivă operațiune dintre cele trei, rolul acesteia fiind inversarea valorii de adevăr sau fals. Expresia lingvistică prezentată a făcut apel la un caz special de populație în care nu există decât două tipuri de animale posibile: masculi și femele. Sigur că în alte regnuri animale (chiar și la oameni) există specii la care nu există această împărțire: de exemplu, căluții de mare au toți același sex (schimbându-și permanent rolul), după cum la oameni există cazuri de hermafroditism. Pentru cazurile în care nu pot exista decât două categorii posibile, sistemul binar este un bun înlocuitor, una dintre categorii fiind reprezentată de cifra 1 iar cealaltă de cifra 0.

Ca și în limbajul curent, în algebra booleană tot ce nu este adevărat este fals și invers. Pentru orice element extras la întâmplare, putem afla valoarea acestuia de adevăr printr-o singură întrebare: este mascul?

Dacă răspunsul este da, atunci avem o valoare de adevăr; dacă nu, o valoare de fals și implicit cealaltă categorie va dobândi valoarea de adevăr. De aceea nici nu contează ce categorie formează întrebarea, putem la fel de bine să întrebăm: este femelă?

Tot ce nu este mascul este femelă, tot ce nu este femelă este mascul:

$\text{NOT}(\text{Mascul}) = \text{Femelă}$

$\text{NOT}(\text{Femelă}) = \text{Mascul}$

Termenul pentru negație este NOT, preluat ca atare din limba engleză (nu există un echivalent în limba română). La fel ca la primele două operații, cuvântul NOT se poate înlocui printr-un semn; spre deosebire de acestea (pentru care semnele au devenit deja un standard), pentru negație nu există un semn standardizat. În diverse cărți și tratate de matematică se pot întâlni semnele ! sau  $\sim$  sau  $\neg$  (printre altele).

$\sim \text{Adevărat} = \text{Fals}$

$\sim \text{Fals} = \text{Adevărat}$

Înlocuind *Adevărat* cu valoarea 1 și *Fals* cu valoarea 0 obținem următorul tabel simplificat:

$\sim 1 = 0$

$\sim 0 = 1$

Pentru simplitate, exemplele prezentate conțin doar două condiții cauzale însă este evident că atât operațiunea SAU cât și operațiunea ȘI pot opera cu orice număr de condiții. Numărul posibil de combinații posibile crește însă exponențial pentru fiecare nouă condiție cauzală introdusă în expresie, după o formulă matematică prezentată în secțiunea 5.1.5.

Cele trei operațiuni pot fi utilizate în combinație unele cu celelalte; un exemplu de astfel de operațiune combinată este “SAU negativ”, care este negația rezultatului operațiunii SAU logic:

$$\sim(1 + 1) = 0$$

Modalitatea de calcul este următoarea: se evaluează mai întâi expresia din paranteză (a cărei rezultat este una dintre valorile adevărat sau fals), după care se inversează rezultatul obținut. Dacă în operațiunea SAU logic expresia “adevărat sau adevărat” are ca rezultat un adevăr, negația acesteia este chiar rezultatul de mai sus ( $0 = \text{fals}$ ).

Un exemplu similar este operațiunea “ȘI negativ”, care este negația rezultatului operațiunii ȘI logic:

$$\sim(1 \cdot 0) = 1$$

Expresia din paranteză este falsă, iar negația acesteia este o expresie adevărată. Înțelegerea acestor noțiuni este surprinsă perfect de următoarea afirmație umoristică:

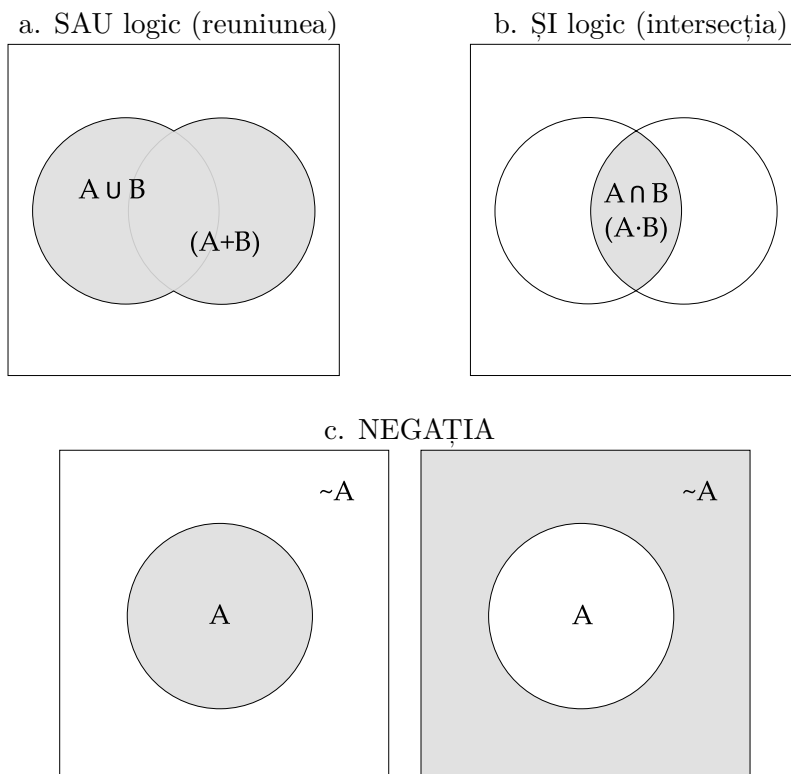
Există 10 tipuri de oameni în lume: cei care înțeleg sistemul binar și cei care nu-l înțeleg.

## 4.4 Diagramele Venn

Studiind teoria mulțimilor, matematicianul și filozoful englez John Venn a introdus (pe la sfârșitul secolului XIX) câteva metode inovatoare pentru a ilustra grafic relațiile dintre mulțimi.

Mai puțin cunoscut este faptul că Venn a dus mai departe ceea ce un alt mare matematician elvețian, Leonhard Euler, a inventat cu un secol înaintea sa. Euler a fost interesat și el de relațiile dintre mulțimi, însă a dorit să arate doar suprapunerea mulțimilor nu toate combinațiile posibile ca în diagramele lui Venn. Întrucât unele mulțimi sunt imposibile (de exemplu mulțimea oamenilor nu se intersectează cu mulțimea păsărilor), o diagramă Euler arată două cercuri separate, în timp ce o diagramă Venn arată și intersecția dintre acestea, chiar dacă această intersecție este logic imposibilă.

Figura 4.1: Operațiunile de algebră booleană prin diagrame Venn



Pentru a onora eforturile ambilor matematicieni, acest tip de diagrame se mai numesc (uneori) în matematică și diagrame Euler-Venn, însă au rămas cu denumirea mai răspândită și mai simplă de diagrame Venn, probabil și pentru faptul că Euler era deja consacrat în multiple ramuri ale matematicii cu sute de publicații în timp ce Venn nu a rămas în istorie decât cu diagramele sale.

Aceste diagrame sunt utile atât pentru o introducere vizuală în explicarea tabelii de adevăr de la secțiunea 5.1.5, cât și pentru vizualizarea soluției algoritmului de minimizare booleană de la secțiunea 5.1.6.

Cele trei operațiuni booleene pot fi vizualizate în figura 4.1. Prima operațiune prezentată (SAU logic) este afișată în diagrama 4.1.a prin reuniunea dintre două mulțimi; dacă mulțimea  $A$  înseamnă *moștenire* iar mulțimea  $B$  înseamnă *câștig*, atunci bogăția poate fi obținută fie când  $A$  este adevărată (când este prezentă) fie când  $B$  este adevărată (prezentă), fie când ambele sunt adevărate. Suprafața din afara celor două cercuri reprezintă mulțimea celor care nu au primit nici moștenire, nici nu au câștigat suficienți bani ca să se considere bogați. Suprafața plină care

acoperă cele două cercuri se poate nota prin ' $A \cup B$ ' sau prin ' $A + B$ '.

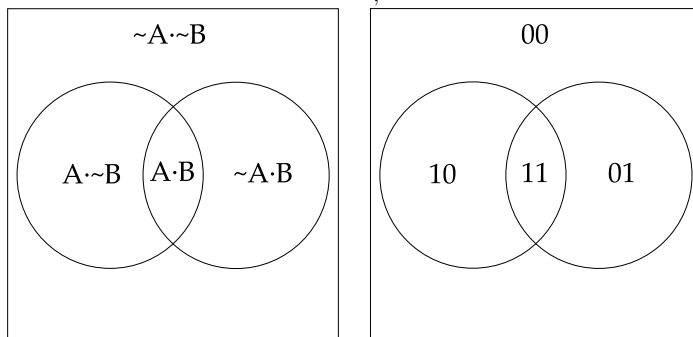
Diagrama 4.1.b prezintă cea de a doua operațiune (ȘI logic) prin intersecția dintre două mulțimi: dacă A este mulțimea oamenilor inteligenți iar B este mulțimea oamenilor care muncesc mult, atunci performanța în știință se întâlnește numai la intersecția dintre cele două mulțimi unde se regăsesc oamenii inteligenți care în același timp muncesc mult. Suprafața colorată dintre cele două mulțimi se poate nota prin ' $A \cap B$ ' sau prin ' $A \cdot B$ '.

Diagrama 4.1.c prezintă cea de a treia operațiune, a negației. În partea stângă este prezentată situația când A este adevărată iar în partea dreaptă este situația opusă, când  $\sim A$  este mulțimea adevărată (prin negarea unui fals, în partea stângă, se obține un adevăr în partea dreaptă). Această diagramă, care utilizează notația pentru negație ' $\sim$ ', face o introducere către adevărata utilizare a diagramelor Venn, pentru două sau mai multe mulțimi.

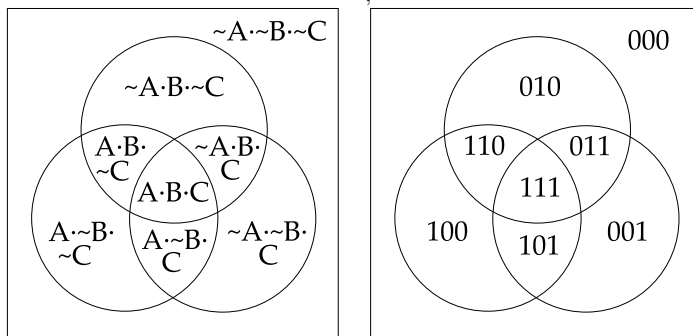
În următoarea figură prezentată, diagrama 4.2.a indică patru zone distincte:

Figura 4.2: Diagrame Venn

a. cu două mulțimi distincte



b. cu trei mulțimi distincte



- în afara celor două cercuri este zona  $\sim A \cdot \sim B$ , adică mulțimea elementelor care nu se află nici în A, nici în B;
- în cercul din stânga, în afara intersecției se află zona  $A \cdot \sim B$ , adică toate elementele care se află în A dar nu se află în B;
- la mijloc se află zona intersecției  $A \cdot B$ ;
- în cercul din dreapta, în afara intersecției se află zona  $\sim A \cdot B$  adică toate elementele care se află în mulțimea B dar nu și în mulțimea A.

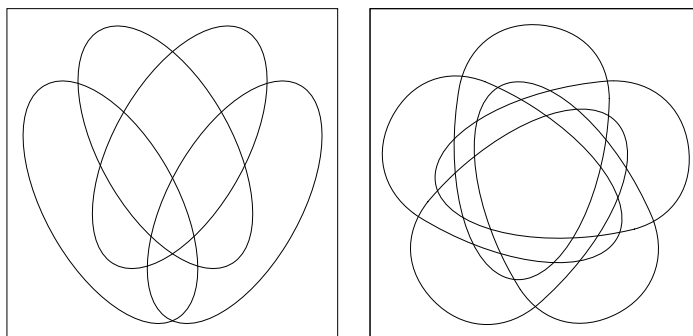
Apartenența la o mulțime este echivalentă cu prezența unei caracteristici: dacă se studiază un element și se constată că acesta aparține mulțimii oamenilor inteligenți, atunci spunem că inteligența este o caracteristică *prezentă* pentru elementul respectiv. Prezența unei caracteristici se notează cu cifra 1, iar absența acesteia cu cifra 0. Notăția este echivalentă cu răspunsul de adevărat sau fals la întrebarea: “este prezentă caracteristica?”. Dacă răspunsul este ‘adevărat’ se notează cu 1 iar dacă este ‘fals’ se notează cu cifra 0.

Combi-națiile de cifre sunt așadar la fel de informative ca și utilizarea literelor, avantajul constând în simplitate. De exemplu, mulțimea elementelor care nu se află nici în A nici în B se notează cu  $\sim A \cdot \sim B$  (not A și not B), însă aceeași informație este transmisă prin utilizarea combinației binare ‘00’.

Același tip de informație este valabil și pentru trei (ori prin generalizare la orice număr de) mulțimi: în diagrama 4.2.b, combinația  $\sim A \cdot B \cdot \sim C$  este echivalentă cu mult mai simpla și mai vizibilă combinație binară 010.

Deși Venn a susținut inițial că aceste diagrame nu pot fi desenate decât pentru trei mulțimi (prin intermediul a trei cercuri), el a recunoscut în

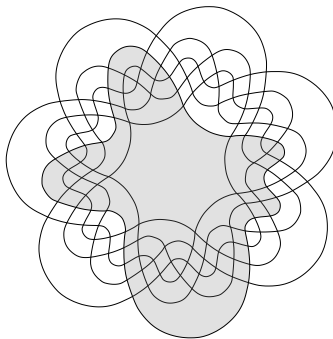
Figura 4.3: Diagrame Venn cu 4 și 5 mulțimi



cele din urmă că există o diagramă și pentru patru mulțimi; secretul constă în renunțarea la forma cercului și utilizarea unei elipse. Figura 4.3 arată și versiunea unei diagrame cu 5 mulțimi, pe care Venn nu a văzut-o niciodată.

Pentru multă vreme se credea că nici nu există diagrame pentru mai mult de 5 mulțimi, însă matematicienii dintr-o nouă ramură numită combinatorică au reușit să găsească soluții extrem de ingenioase. Ruskey și Weston (2005) au dedicat o pagină de web specială pentru diagramele Venn, arătând că în fapt aceste diagrame sunt posibile pentru virtual orice număr de mulțimi, iar dacă acest număr este prim atunci diagramele sunt chiar simetrice. Figura 4.4 demonstrează acest lucru printr-o superbă diagramă simetrică cu 7 mulțimi denumită “Adelaide”.

Figura 4.4: Diagramă Venn cu intersecțiile dintre 7 mulțimi



Se observă însă că numărul de intersecții crește în mod exponențial odată cu introducerea fiecărei noi mulțimi, ceea ce limitează aria de aplicabilitate practică a diagramei cu mai mult de patru sau cinci mulțimi.

## Capitolul 5

# Metodologia analizei calitativ-comparative

### 5.1 Analiza cu mulțimi distincte bivalente

Informațiile prezentate în secțiunile anterioare reprezintă abordări care se află la granița dintre o metodologie clar conturată și analiza subiectivă specifică cercetărilor calitative. Deși mulți dintre cercetătorii formați în tradiția cantitativă (bazată pe studierea unui număr mare de cazuri) văd în aceste abordări multiple slăbiciuni, alții le privesc ca pe o virtute. Una dintre încercările analizei calitative de a contura o metodologie clar definită este QCA (Qualitative Comparative Analysis – analiza calitativ-comparativă) promovată de Ragin în celebra lui carte *The Comparative Method. Moving beyond qualitative and quantitative strategies* din 1987.

Ideile prezentate atunci de Ragin au fost substanțial modificate ori îmbunătățite ulterior (detalii în secțiunea 5.3), însă cartea este considerată un punct de reper important în științele sociale. Metoda utilizată de Ragin este definită tot de acesta într-o publicație ulterioară: “studiul comparativ al diversității unui număr moderat de cazuri” (Ragin, 1994, p.33)

Analiza cu mulțimi distincte bivalente se suprapune foarte precis peste analiza cu mulțimi binare, iar acestea au condus în cele din urmă la apariția calculatoarelor, fapt care a revoluționat modul de a face știință în prezent. Începând cu Aristotel și Platon, marile descoperiri științifice au fost mai întâi apanajul filozofiei și apoi transpuse în experimente. Cu precădere în domeniul științelor socio-umane, marii teoreticieni au fost în primul rând mari filozofi. Doar după inventarea computerelor știința a dobândit o puternică latură cantitativă, deși este discutabil dacă

abordarea cantitativă duce (sau poate să ducă) la formarea de mari teorii. Aș înclina să spun că cercetările cantitative sunt foarte potrivite pentru testarea de ipoteze, însă pentru formularea teoriilor sunt necesare mult mai multe lucruri decât analiza statistică a datelor disponibile.

Un exemplu celebru și sugestiv este teza de master a lui Shannon (1938) a cărei fundament teoretic vine nu din matematica inginerescă ci dintr-un curs de filozofie pe care acesta l-a urmat pe când era student la MIT. Acel curs de filozofie i-a marcat cariera ulterioară, făcând conexiuni importante între algebra booleană și modul de închidere și deschidere a circuitelor electrice. Shannon a aplicat apoi principiile algebrei booleene în teza sa, care este considerată acum materialul de la care a plecat toată știința calculatoarelor. Este binecunoscut faptul că funcționarea acestora se bazează pe combinațiile de biți 1 și 0, semnificația acestora însemnând “deschis” și “închis” (cu referire la așa numitele *porți* din circuitele electrice).

Principiile stabilite de Shannon au fost utilizate mai apoi în crearea unor circuite electrice de control, un fel de mini-calculatoare care erau destinate unor operații foarte precise și limitate. Creșterea în complexitate a circuitelor electrice a generat o altă problemă: același rezultat final putea fi obținut prin mai multe circuite, altfel spus prin mai multe combinații binare, unele mai eficiente decât altele (cu un consum electric mai mare sau mai mic). Sarcina pe care au avut-o de rezolvat inginerii era de a găsi o metodă prin care rezultatul să fie obținut cu cel mai mic consum de energie posibil.

Această teză a avut o influență dramatică asupra științei și tehnologiei moderne, fundamentând bazele teoretice ale construirii calculatoarelor și mai departe a unor mii de alte aplicații. Dacă suntem de acord că în societatea actuală calculatoarele joacă un rol foarte important, teza lui Shannon a fost una dintre cele mai importante ale secolului trecut, ale cărei efecte se simt în toate colțurile oricărei societăți dezvoltate.

După aproape 20 de ani de la articolul lui Shannon a apărut o altă teză celebră care a influențat această știință: este vorba de teza de doctorat a lui McCluskey (1956) care a reușit să inventeze un algoritm prin care să fie găsită o soluție minimă la orice problemă de combinații binare. Odată cu această teză s-au pus bazele calculatoarelor moderne, în care erau posibile nu doar închiderea și deschiderea unor circuite electrice, dar și anumite operațiuni matematice utile în rularea programelor software. Algoritmii același a rămas în istorie sub denumirea “Quine-McCluskey”, pentru a onora și munca lui Quine (1952, 1955), un alt mare teoretician al circuitelor logice.



Calculatoarele din generația actuală au evoluat mult față de cele de pe vremea lui Shannon și McCluskey, însă principiile de bază rămân aceleași, toate funcționând în același sistem binar. În fapt s-au dezvoltat atât de mult încât permit circulara informației cu viteze suficient de mari pentru a permite o comunicare fără precedent nu doar între oameni de știință din același domeniu, ci și între domenii științifice diverse.

Algebra lui Bool, tezele lui Shannon și a lui McCluskey, adăugate la principiile comparative a lui Mill au fost utilizate de către Ragin la sfârșitul anilor '80 pentru a formula prima metodologie consistentă de analiză comparativă. Ragin a preluat algoritmul lui McCluskey din științele ingineresti și l-a adaptat la științele sociale, urmând un raționament asemănător: un fenomen poate fi obținut/declanșat de combinații cauzale multiple; sarcina pe care o are cercetătorul comparativist este de a găsi combinația minimă de cauze necesare și suficiente pentru a declanșa fenomenul. Paralela cu motivația lui McCluskey este evidentă, meritul lui Ragin fiind acela de a fi primul care a transpus un algoritm destinat științelor ingineresti pentru obiective de cercetare socială comparativă.

Există însă o deosebire între rezultatul urmărit de McCluskey și cel urmărit de Ragin: în timp ce în circuitele logice de control nu este necesar decât cel mai scurt traseu posibil (ori de unul dintre cele mai scurte trasee, pentru că pot exista mai multe trasee minime de aceeași lungime), în cercetarea socială este nevoie de toate combinațiile cauzale minime care declanșează fenomenul de interes. De aceea, algoritmii din științele ingineresti nu oferă decât o singură soluție minimă în timp ce algoritmii de analiză comparativă (pentru că există mai mulți) oferă toate soluțiile minime.

### 5.1.1 Analiza de necesitate

Ideile de bază ale condițiilor necesare pentru producerea unui fenomen au fost prezentate în secțiunea 2.2. Această secțiune se va concentra pe derularea unei analize de necesitate pe mulțimi distincte bivalente și pe identificarea situațiilor în care putem spune cu certitudine dacă o condiție este sau nu necesară.

În mulțimile bivalente (binare, booleene), atât fenomenul de interes, cât și condiția cauzală sunt măsurate cu ajutorul a două valori: 0 pentru absență și 1 pentru prezență. În consecință, orice set de date poate de acest tip poate fi redus la un simplu tabel încrucișat:

Tabelele încrucișate sunt des utilizate în cercetarea cantitativă, pentru determinarea gradului de corelație dintre două variabile. Figura 5.1.a

Tabelul 5.1: Încrucișări 2×2 specifice pentru:

		X				X	
		0	1			0	1
Y	1	0	37	Y	1	0	37
	0	35	0		0	35	23

reprezintă un exemplu de corelație pozitivă perfectă: toate cazurile în care X este absent sunt asociate cu absența lui Y, în timp ce cazurile în care X este prezent sunt asociate cu prezența lui Y. În restul celulelor, nu există niciun caz care să modifice valoarea de +1 a coeficientului de corelație.

Coeficientul de corelație este unul simetric, atât între variabile ( $r_{xy} = r_{yx}$ ) cât și cu privire la absența / prezența concomitentă a celor două variabile: ambele contribuie în aceeași măsură la calcularea corelației.

Analiza calitativ-comparativă are însă o diferență fundamentală față de corelația cantitativă, prin faptul că relația dintre condiția cauzală și fenomen este asimetrică.

Cu privire la relația dintre dezvoltarea economică (condiția X) și democrație (fenomenul Y), Goertz și Starr (2003, p.17) au făcut următoarea distincție între ipoteza bazată pe corelație și cea sugerată de condiția necesară:

- Cu cât mai mare este nivelul de dezvoltare economică, cu atât mai probabil este ca o țară să dezvolte o democrație
- Un nivel minim de dezvoltare economică este necesar pentru existența unei democrații

Faptul că prezența unei condiții cauzale este necesară pentru producerea unui fenomen nu spune absolut nimic despre absența fenomenului. Prezența unei condiții poate explica producerea fenomenului, însă absența condiției nu este nicidecum un factor explicativ pentru absența fenomenului de interes.

În cadrul corelației pozitive, cu cât o variabilă crește cu atât crește și cealaltă. Cu cât una din variabile este mai slabă, cu atât mai slabă este și

cealaltă. În QCA, există o anumită condiție cauzală (sau un set de condiții cauzale) care explică producerea fenomenului, însă este perfect posibil ca o cu totul altă condiție (set de condiții) să explice absența fenomenului.

De aceea, analiza de necesitate se concentrează doar pe cazurile în care fenomenul se produce (linia în care  $Y = 1$ ). Cazurile în care  $Y = 0$  sunt irelevante, indiferent dacă se indică prezența sau absența condiției cauzale.

Dacă  $X$  este o condiție necesară pentru producerea lui  $Y$ , atunci ar trebui să observăm prezența lui  $X$  atunci când fenomenul  $Y$  se produce. Practic, este imposibil să observăm producerea fenomenului în absența lui  $X$ .

În același timp, nu ar trebui să fie nicio instanță în care  $X$  să lipsească, atunci când  $Y$  este prezent. Aceasta este o regulă foarte importantă în determinarea necesității: dacă fenomenul este prezent, iar condiția cauzală este absentă, putem concluziona ca acea condiție cauzală nu este necesară pentru producerea fenomenului.

Logica necesității spune că ori de câte ori fenomenul este prezent, condiția necesară este deasemenea prezentă. De aceea, trebuie examinate doar cazurile în care fenomenul este prezent.

În analiza pe mulțimi distincte bivalente, pe tabelul încrucișat, o condiție cauzală este necesară pentru producerea fenomenului ( $Y = 1$ ) dacă sunt îndeplinite în mod cumulativ următoarele două condiții:

- există cazuri când  $X = 1$
- nu există cazuri când  $X = 0$

Orice distribuire de cazuri pe celelalte celule ale tabelului este irelevantă pentru stabilirea necesității condiției cauzale. În tabelul 5.1.b sunt 37 de cazuri unde ambele  $X$  și  $Y$  sunt prezente și niciun caz în care  $Y$  să fie prezent și  $X$  să lipsească. În consecință, putem spune că  $X$  este o condiție necesară pentru producerea fenomenului  $Y$ .

O astfel de distribuire perfectă a cazurilor este posibilă doar în teorie. În practică este foarte rar ca o celulă să nu conțină absolut niciun caz când  $X = 0$ , de aceea regula poate fi relaxată puțin (făcând deja un pas spre mulțimile vagi):

- o proporție (foarte) mare a cazurilor când  $X = 1$  (iar  $Y = 1$ )
- o proporție (foarte) mică a cazurilor când  $X = 0$  (iar  $Y = 1$ )

Este complet irelevant faptul că există 23 de cazuri în care prezența condiției cauzale este asociată cu absența fenomenului, întrucât relația de necesitate este asimetrică: ceea ce este necesar pentru producerea fenomenului nu trebuie neapărat să fie necesar pentru absența acestuia.

Din tabelul respectiv nu putem concluziona nimic despre corelația dintre X și Y, însă în mod cert X este o condiție necesară pentru Y. În termeni formali:

$$X \Leftarrow Y$$

Direcția săgeții nu implică niciun fel de relație cauzală, întrucât  $X \Leftarrow Y$  nu înseamnă că un fenomen Y conduce la o condiție cauzală X. Săgeata se traduce doar printr-o implicație logică, de tipul: atunci când Y se produce, sigur este prezentă și condiția X.

Data fiind relația asimetrică dintre X și Y, analiza de necesitate se efectuează separat pentru prezența fenomenului și pentru absența acestuia. Dacă putem demonstra că o condiție X este necesară pentru producerea fenomenului Y, inversa nu este neapărat valabilă.

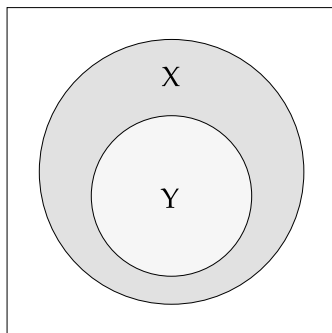
Absența lui X nu este neapărat o condiție necesară pentru absența lui Y ( $\sim X \Leftarrow \sim Y$ ).

Dimpotrivă, este destul de probabil să găsim că o altă condiție cauzală Z este necesară pentru *absența* fenomenului Y, iar atunci:

$$Z \Leftarrow \sim Y$$

Chiar dacă obiectivul cercetării este acela de a studia condițiile în care fenomenul se produce, studiul cazurilor negative se poate dovedi la fel de substanțial din punct de vedere teoretic, cu implicații de sine stătătoare.

Figura 5.1: Condiția X este necesară pentru Y ( $X \Leftarrow Y$ )



În termeni de teoria mulțimilor, relația de necesitate se traduce prin faptul că elementele componente ale lui Y reprezintă un subset al elementelor componente a lui X. Mulțimea X este mai mare decât mulțimea Y, care este inclusă în mulțimea X.

Ragin (2000; 2008b) a fost printre primii care s-au aplecat cu atenție asupra relației dintre fenomene utilizând teoria mulțimilor, și a introdus în metodologia de cercetare socială două noi măsuri numite incluziune și acoperire. Acestea vor fi prezentate pe larg la analiza cu mulțimi vagi, în secțiunea 5.3.6.

Privită grafic prin diagrame Venn, o mulțime  $X$  este necesară pentru o mulțime  $Y$  dacă mulțimea  $Y$  este complet inclusă în mulțimea  $X$ . Nu există niciun caz în care  $Y$  să fie prezent iar  $X$  să lipsească, întrucât toate cazurile în care  $Y$  se întâmplă (mulțimea  $Y = 1$ ) sunt complet incluse în mulțimea în care  $X$  se întâmplă ( $X = 1$ ).

### 5.1.2 Analiza de suficiență

Necesitatea și suficiența sunt două concepte complementare. Spre deosebire de analiza necesității, o condiție cauzală  $X$  este suficientă pentru producerea unui fenomen  $Y$  dacă atunci când  $X$  este prezentă,  $Y$  se produce.

Dacă la necesitate mulțimea  $Y$  este un subset al mulțimii  $X$ , la analiza de suficiență relația dintre cele două mulțimi este exact inversă:  $X$  este un subset al mulțimii  $Y$ . Dacă  $X$  este o condiție necesară pentru  $Y$ , atunci  $Y$  este suficientă pentru  $X$ .

De exemplu, prezența la examen este o condiție absolut necesară, dar nu și suficientă pentru a lua nota 10. În figura 5.1, condiția  $X$  este prezența la examen iar fenomenul rezultat  $Y$  este obținerea notei 10.

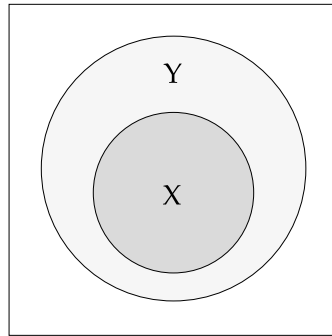
Dacă știm însă că nota obținută la examen este 10, această informație este *suficientă* pentru a deduce faptul că persoana s-a prezentat la examen.

În următoarea figură 5.2, avem relația inversă în care  $X$  (nota 10 la examen) este o condiție suficientă pentru  $Y$  (prezența la examen). Urmând același raționament, dacă  $X$  se întâmplă și este suficientă pentru  $Y$ , atunci este obligatoriu (în mod *necesar*) ca  $Y$  să se întâmple.

În termeni formali, relația dintre cele două se notează cu:

$$X \Rightarrow Y$$

Faptul că mulțimea  $X$  este mai mică decât (și inclusă în) mulțimea  $Y$  spune că  $X$  nu explică toate cazurile în care se produce  $Y$ . Mai sunt și alte situații în care  $Y$  se produce, iar  $X$  nu este responsabilă pentru producerea acestora. În această context, spunem că  $X$  este o condiție suficientă, însă nu este și o condiție necesară pentru producerea fenomenului  $Y$ .

Figura 5.2: Condiția X este suficientă pentru Y ( $X \Rightarrow Y$ )

Pentru ca X să fie o condiție în același timp necesară și suficientă pentru Y, cele două mulțimi ar trebui să fie perfect suprapuse (niciuna să nu fie mai mare decât cealaltă).

Necesitatea și suficiența mai sunt complementare și dintr-un alt punct de vedere. Dacă X este o condiție *suficientă* pentru Y, atunci absența condiției X este o condiție *necesară* pentru absența fenomenului Y.

Acest lucru poate fi explicat simplu utilizând aceeași teorie a mulțimilor: dacă X este o condiție suficientă pentru Y, înseamnă că mulțimea X este inclusă (un subset) în mulțimea Y (ca în figura 5.2). De aici rezultă că mulțimea  $\sim X$  (adică  $X = 0$ ) este mai mare decât mulțimea  $\sim Y$ , iar dacă  $\sim Y$  este un subset al lui  $\sim X$  înseamnă că  $\sim X$  este o condiție necesară pentru  $\sim Y$  (tabelul 5.2.b).

Tabelul 5.2: Încrucișări  $2 \times 2$  în care:

a. X este suficientă pentru Y

b.  $\sim X$  este necesară pentru  $\sim Y$ 

		X	
		0	1
Y	1	49	58
	0	23	0

		X	
		0	1
Y	1	49	58
	0	23	0

Într-un tabel  $2 \times 2$ , pentru a verifica dacă o condiție X este suficientă, se aplică o tehnică similară: selecția cazurilor când X se produce. Celelalte

două celule ale tabelului sunt irelevante, iar condițiile suficienței sunt:

- există cazuri când fenomenul  $Y$  se produce ( $Y = 1$ );
- nu există cazuri când fenomenul  $Y$  este absent ( $Y = 0$ ).

În tabelul  $2 \times 2$  de mai sus, se poate observa că mulțimea  $X$  este un subset al mulțimii  $Y$ , pe baza următoarelor aspecte:

- sunt  $49 + 58 = 107$  cazuri în care fenomenul  $Y$  se produce;
- în doar 58 din cazurile în care  $Y$  se produce (așadar o submulțime a lui  $Y$ ) este prezentă și condiția cauzală  $X$ , ceea ce înseamnă că (cel puțin) o parte a mulțimii  $X$  este inclusă în mulțimea  $Y$ ;
- mulțimea cazurilor în care condiția cauzală  $X$  este prezentă nu se extinde în afara mulțimii  $Y$  (nu există cazuri când  $X = 1$  și  $Y = 0$ ), ceea ce înseamnă că mulțimea  $X$  este *complet* inclusă în mulțimea  $Y$ , deci formează un subset al acesteia.

### 5.1.3 Relevanța condițiilor cauzale: acoperirea

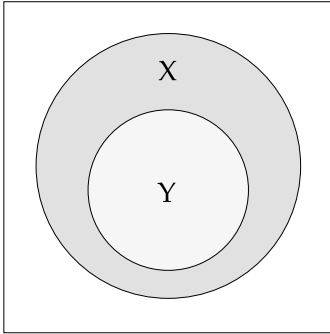
În științele sociale, este extrem de dificilă identificarea acelor condiții necesare și/sau suficiente pentru producerea unui fenomen. Pentru fiecare concept se scriu tomuri întregi de argumente, exemple și contra-exemple iar fiecare autor încearcă să demonstreze un anumit punct de vedere. Unele sunt valabile în anumite contexte, altele sunt valabile în general, în timp ce altele sunt demonstrate mai târziu ca fiind pur și simplu false.

Există totuși o metodă simplă pentru a determina dacă o condiție cauzală este sau nu relevantă. Un exemplu de relație necesară este legătura dintre prezența oxigenului și izbucnirea unui incendiu. Prezența oxigenului este în mod evident o condiție necesară pentru întreținerea focului, însă nu este deloc o condiție suficientă pentru producerea incendiilor. Avem așadar o condiție necesară dar nu și suficientă, următorul pas fiind determinarea relevanței acestei condiții.

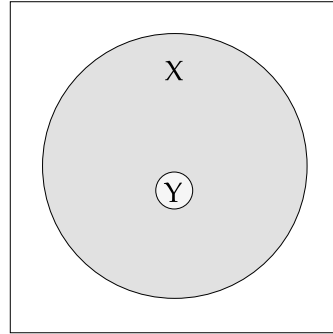
Este adevărat că observăm prezența aerului ( $X$ ) de fiecare dată când focul se produce ( $Y$ ), deci  $X$  este o condiție necesară pentru  $Y$ . Pe de altă parte, putem observa în mod intuitiv că prezența aerului este o condiție irelevantă pentru producerea unui incendiu. În cele mai multe cazuri există aer, dar nu se produce foc.

Figura 5.3: X este necesar pentru Y, unde:

a. X este o condiție relevantă



b. X este o condiție irelevantă



De aici se poate construi o măsură a relevanței unei condiții cauzale, calculată ca un fel de procent al cazurilor în care se produce fenomenul, din total cazuri în care este prezentă și condiția cauzală.

În figura 5.3.a, X este o condiție relevantă pentru producerea lui Y pentru că mulțimea  $Y = 1$  acoperă o parte destul de mare din mulțimea X. Pe de altă parte, în figura 5.3.b mulțimea Y acoperă doar o mică parte din mulțimea X.

Putem privi aceste două figuri în termeni de proporții: în prima situație, proporția cazurilor în care Y se produce, din total cazuri în care X se produce, este destul de mare, și foarte mică în a doua.

Utilizând o încrucișare  $2 \times 2$  generală (cum este cea prezentată în tabelul 5.3), pentru ipoteza că X este necesar pentru Y ( $X \Leftarrow Y$ ), putem verifica incluziunea mulțimii Y în mulțimea X care este dată de proporția cazurilor în care X și Y se întâmplă (celula c, intersecția dintre X și Y) supra total cazuri în care Y se întâmplă (suma dintre celulele a și c):

$$incl_{X \Leftarrow Y} = \frac{c}{a + c}$$

Relevanța condiției X pentru fenomenul Y, în aceeași ipoteză a necesității lui X pentru Y ( $X \Leftarrow Y$ ), este dată de proporția cu care submulțimea Y acoperă mulțimea X, adică proporția cazurilor în care X și Y se întâmplă (celula c, intersecția dintre X și Y) supra total cazuri în care X se întâmplă, adică suma dintre celulele c și d:

$$cov_{X \Leftarrow Y} = \frac{c}{c + d}$$

În termeni de probabilități, relevanța condiției X pentru fenomenul Y în ipoteza necesității are o formulă aproape identică denumită o



Tabelul 5.3: Încrucișare generală  $2 \times 2$ 

		X	
		0	1
Y	1	a	c
	0	b	d

“probabilitate condiționată”, mai exact este probabilitatea ca Y să se întâmple, dat fiind că X s-a întâmplat deja:

$$P(Y|X) = \frac{P(Y \cap X)}{P(X)} = \frac{c}{c + d}$$

Scorul de acoperire este un număr între 0 și 1. Cu cât acoperirea este mai mare (mai aproape de 1) cu atât mai relevantă este condiția cauzală pentru producerea lui Y, în ipoteza că X este necesară.

Pentru ipoteza suficienței lui X pentru producerea lui Y, avem formule similare. Includerea mulțimii X în mulțimea Y este calculată cu formula:

$$incl_{X \Rightarrow Y} = \frac{c}{c + d}$$

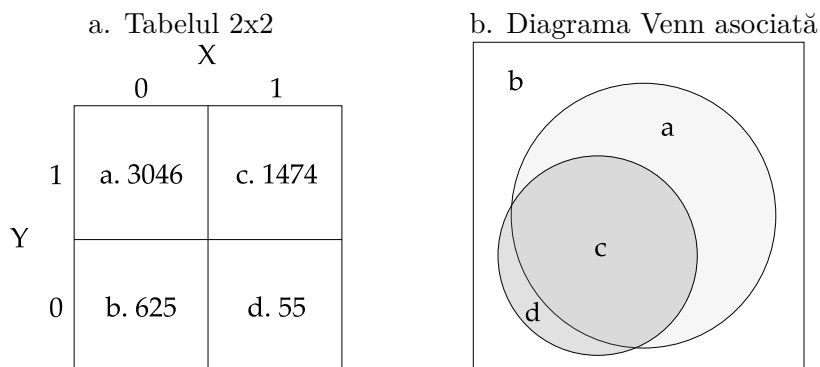
În mod similar, relevanța condiției X pentru ipoteza suficienței este dată de gradul de acoperire a submulțimii X în mulțimea Y:

$$cov_{X \Rightarrow Y} = \frac{c}{a + c}$$

Se poate observa ușor că formulele pentru suficiență pot fi derivate direct din formulele pentru necesitate, deoarece:

- includerea pentru necesitate este egală cu acoperirea pentru suficiență
- includerea pentru suficiență este egală cu acoperirea pentru necesitate

Figura 5.4: Includere și acoperire, adaptare după Ragin (2008b)



Utilizând un tabel  $2 \times 2$ , Ragin (2008b, pp.56-57) oferă un exemplu demonstrativ prin analiza relației dintre nivelul de educație ( $X = 1$  înalt,  $X = 0$  scăzut) și evitarea sărăciei ( $Y = 1$  ieșit din sărăcie,  $Y = 0$  în sărăcie), plecând de la ipoteza că un nivel înalt al educației este o condiție suficientă pentru ieșirea din sărăcie.

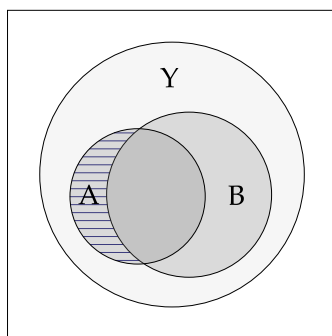
Mulțimea  $X$  cuprinde 1529 persoane, din care 1474 sunt incluse în mulțimea  $Y$ . Avem așadar o includere aproape totală de  $1474/1529 = 96,4\%$ . Din punctul de vedere al analizei cu mulțimi distincte, chiar și un singur caz în celula “d” poate să elimine ipoteza suficienței ca fiind nefondată, motiv pentru care analiza cu mulțimi vagi lucrează cu niște praguri de includere.

Pentru moment însă, dată fiind o includere atât de mare, putem spune că o educație înaltă este o condiție (aproape total) suficientă pentru ieșirea din sărăcie. Desigur, sunt mai multe modalități posibile de a ieși din sărăcie, educația înaltă fiind doar una dintre acestea. Educația înaltă este suficientă, dar nu neapărat necesară pentru ieșirea din sărăcie.

Dintre toate modalitățile posibile, unele se aplică la un număr restrâns de persoane (situație în care condiția este suficientă dar irelevantă), iar altele pot fi observate la un număr mare de persoane.

În exemplul lui Ragin, educația înaltă acoperă o parte semnificativă din totalul persoanelor ieșite din sărăcie. Din totalul de 4550 persoane ieșite din sărăcie, nu mai puțin de  $1474/4550 = 32,4\%$  au educație înaltă. După toate criteriile, această condiție cauzală este una extrem de relevantă pentru explicarea ieșirii din sărăcie. Printr-o paralelă cu analiza de regresie din statistică, dacă o variabilă independentă explică un procent atât de mare din variația dependentei, atunci este o variabilă care trebuie în mod obligatoriu să facă parte din modelul explicativ.

Figura 5.5: Acoperire brută și acoperire unică



Cu cât acoperirea se apropie mai mult de 100%, cu atât putem spune că acea condiție cauzală devine din ce în ce mai necesară pentru producerea fenomenului. La limită, dacă o condiție  $X$  este suficientă (inclusă în mulțimea  $Y$ ) și acoperă 100% mulțimea respectivă, atunci în plus față de suficiență avem o dovadă a faptului că mulțimea  $X$  este necesară pentru producerea fenomenului. În cazul în care mulțimea  $X$  se extinde peste mulțimea  $Y$ , atunci rămâne 100% o condiție necesară pentru  $Y$ , însă nu va mai fi neapărat o condiție suficientă.

Acoperirea pe care am calculat-o este denumită până acum este denumită “acoperire brută” (în engl. *raw coverage*). Se poate întâmpla însă ca o parte a acestei acoperiri să fie deja acoperită de o terță condiție cauzală, sau de o suită de alte condiții cauzale.

În figura 5.5, acoperirea brută a mulțimii  $A$  este dată întreaga suprafață a cercului  $A$  supra total suprafață mulțimea  $Y$ , însă este vizibil faptul că o parte a “explicației” oferite de condiția cauzală  $A$  este oferită și de condiția cauzală  $B$ . De aceea, ne interesează cantitatea “unică” a explicației oferite de mulțimea  $A$ , în acest caz suprafața hașurată. În termeni generali, acoperirea unică este calculată ca acoperire brută minus intersecția cu orice altă condiție cauzală, în interiorul mulțimii  $Y$ . Analiza cu mulțimi vagi prezintă niște formule de calcul pentru toate aceste măsuri.

Determinarea necesității sau suficienței, pe baza unui tabel 2x2 generic (cum este cel din figura 5.3), se face prin compararea celulelor  $a$ ,  $b$  și  $d$ :

- dacă celula  $a > celula d$  atunci mulțimea  $Y$  este inclusă în mulțimea  $X$ , deci  $X$  este necesară pentru  $Y$  ( $X \Leftarrow Y$ )
- dacă celula  $a < celula d$  atunci mulțimea  $X$  este inclusă în mulțimea  $Y$ , deci  $X$  este suficientă pentru  $Y$  ( $X \Rightarrow Y$ )

- dacă celula b este egală cu zero, atunci mulțimea X este complet în afara mulțimii Y, în consecință nu este nici necesară, nici suficientă pentru Y

Mai sunt situații în care celula b este mai mare decât zero, însă nici una dintre mulțimi nu este complet inclusă în cealaltă. De exemplu, în ipoteza suficienței, dacă celula b este aproximativ egală cu celula d atunci mulțimea X este doar pe jumătate inclusă în mulțimea Y, deci nu este “pe deplin” suficientă. Doar dacă includerea trece de un anumit prag, putem spune că X este suficientă pentru Y. Celula b trebuie să fie mult mai mare decât celula a (respectiv d) pentru ca incluziunea să fie îndeajuns de mare.

Includerea și acoperirea sunt așadar două instrumente care determină, pe de o parte, “cât” de necesară și/sau suficientă este o condiție cauzală, iar pe de altă parte determină relevanța (importanța) acelei condiții pentru explicarea fenomenului de interes. Deși pot fi calculate în mod independent una de cealaltă, procedura de lucru implică o succesiune logică:

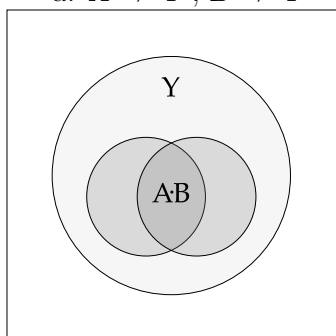
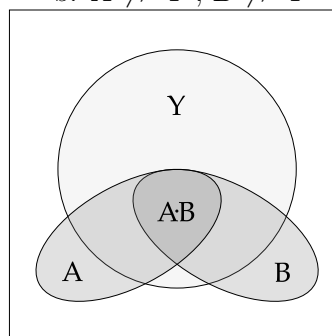
- calcularea includerii: este condiția cauzală necesară / suficientă?
- calcularea acoperirii: dacă răspunsul este afirmativ, este condiția cauzală relevantă (importantă)?

În situația în care răspunsul la prima întrebare este negativ (condiția cauzală nu este necesară, sau nu este suficientă), atunci nu mai are sens calcularea relevanței acesteia. Doar dacă poate fi demonstrată includere, are sens calcularea acoperirii.

#### 5.1.4 Conjunții și disjunții cauzale

Sunt foarte puține situațiile în care o singură condiție cauzală poate influența producerea unui fenomen. De cele mai multe ori, acest lucru se întâmplă doar în conjuncție cu o altă condiție cauzală, ori în conjuncție cu un întreg set de condiții cauzale. Desigur, dacă două mulțimi A și B sunt individual suficiente pentru producerea fenomenului Y, atunci și conjuncția (intersecția) dintre acestea este suficientă, așa cum poate fi văzut în figura 5.6.a.

Singura problemă este faptul că intersecția acoperă doar o parte mică din mulțimea Y, așa încât (deși este suficientă), nu pare să fie foarte relevantă. De aici se poate emite o regulă: cu cât se utilizează mai multe condiții

Figura 5.6: Suficiența unei conjuncții cauzale:  $A \cdot B \Rightarrow Y$ a.  $A \Rightarrow Y$  ;  $B \Rightarrow Y$ b.  $A \not\Rightarrow Y$  ;  $B \not\Rightarrow Y$ 

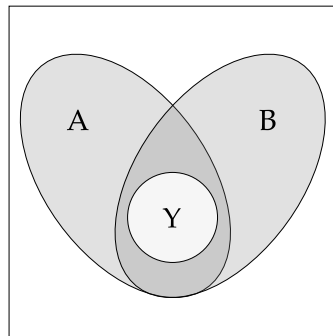
cauzale pentru construirea unei conjuncții, cu atât mai irelevantă devine acea conjuncție. De exemplu, o conjuncție ipotetică poate suna așa: dacă o persoană este din România, și este din județul Suceava, și locuiește în Gura Humorului, este suficient pentru a fi bucovinean. Această conjuncție  $A \cdot B \cdot C \Rightarrow Y$  este corectă, însă este atât de restrictivă încât nu se poate aplica decât la un număr extrem de mic de cazuri, ceea ce înseamnă că este complet irelevantă.

Construirea unei conjuncții poate fi realizată doar în măsura în care conjuncția rezultată acoperă o suprafață îndeajuns de mare din mulțimea Y pentru a fi în același timp suficientă și relevantă.

Prin contrast cu figura 5.6.a, în figura 5.6.b niciuna dintre cele două condiții cauzale A și B nu este suficientă pentru producerea fenomenului Y, întrucât niciuna nu este complet inclusă în mulțimea Y. Atât mulțimea A, cât și mulțimea B arată prezență în interiorul mulțimii Y, însă o proporție semnificativă se află în afara lui Y. Cu o incluziune imperfectă în mulțimea Y, niciuna nu este individual suficientă pentru producerea fenomenului respectiv.

O cu totul altă poveste o spune intersecția dintre A și B, care îndeplinește cele două condiții obligatorii: este complet inclusă în mulțimea Y și acoperă o parte destul de mare din Y pentru a fi relevantă pentru producerea fenomenului. Prezența individuală a lui A sau a lui B nu sunt suficiente pentru apariția lui Y, însă dacă A este prezentă în conjuncție cu prezența lui B (dacă sunt prezente în mod simultan), atunci această combinație cauzală este suficientă pentru producerea lui Y.

Să presupunem că fenomenul Y este bunăstarea materială, condiția cauzală A este inteligența iar condiția cauzală B este munca. Deși sunt mulți oameni inteligenți care produc bunăstare materială, sunt poate și

Figura 5.7: Necesitatea unei conjuncții cauzale:  $A \cdot B \Leftarrow Y$ 

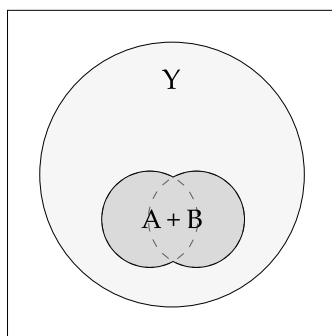
mai mulți care eșuează în acest sens. La fel și munca, sunt destui oameni care produc bunăstare prin muncă, dar sunt poate și mai mulți care muncesc foarte mult și se află la limita sărăciei. Luată individual, niciuna nu este suficientă însă oamenii inteligenți care muncesc mult produc întotdeauna bunăstare materială.

Combinăția dintre inteligență și muncă este suficientă pentru bunăstare, iar figura 5.6.b poate fi completată prin observația adițională că această conjuncție nu este neapărat necesară pentru producerea bunăstării. După cum se poate vedea în figură, o parte destul de mare din Y nu este acoperită de intersecția dintre A și B, ceea ce înseamnă că sunt și alte motive pentru care oamenii dobândesc bunăstare materială (o moștenire, ori câștig la Loto etc.)

Figura 5.7 ilustrează perspectiva cealaltă, atunci când o conjuncție cauzală  $A \cdot B$  este necesară (dar nu și suficientă) pentru producerea fenomenului Y. Spre deosebire de situația precedentă, în care nici A nici B nu sunt suficiente însă conjuncția dintre cele două este într-adevăr suficientă pentru Y, în acest caz conjuncția  $A \cdot B$  este necesară doar dacă atât A cât și B sunt necesare pentru producerea Y.

Y este inclusă în intersecția  $A \cdot B$  dacă și numai dacă Y este inclusă A și în B luate separat. De aici poate fi emisă următoarea regulă: dacă Y este inclusă în A și în același timp Y este inclusă în B, atunci A se intersectează cu B iar Y este inclusă în această intersecție.

Conjuncția dintre două condiții cauzale este echivalentă cu operațiunea de ȘI logic (intersecția dintre două mulțimi, discutată la secțiunea 4.3). Operațiunea conexă este cea de SAU logic (uniunea dintre două mulțimi), care este echivalentă cu termenul de “disjuncție” cauzală.

Figura 5.8: Suficiența unei disjuncții cauzale:  $A + B \Rightarrow Y$ 

Spre deosebire de conjuncție (acțiunea simultană a două condiții cauzale asupra fenomenului de interes), efectul produs de disjuncția a două mulțimi asupra efectului necesită mai multe explicații. În primul rând, o disjuncție este suficientă pentru producerea unui fenomen dacă și numai dacă fiecare dintre cele două condiții cauzale sunt individual suficiente pentru producerea acestuia.

În figura 5.8 se poate observa că atât A cât și B sunt incluse în mulțimea Y, așadar ambele sunt individual suficiente pentru producerea lui Y. Deși mulțimea A este suficientă, ea nu acoperă decât o mică parte din Y deci este o condiție suficientă dar irelevantă. În schimb, disjuncția dintre A și B se dovedește a fi suficientă.

Algoritmul de minimizare booleană (discutat în secțiunea 5.1.6) generează în mod curent soluții disjunctive de tipul 'A + B'. Soluția nu menționează pe niciuna în mod individual, ci doar pe ambele în disjuncție (efectul cumulat al ambelor, prin uniunea dintre cele două). Numai dacă-și unesc efectele pot genera producerea lui Y cu o relevanță (acoperire) îndeajuns de mare, altfel luate separat cele două nu contează foarte mult. Ambele, luate individual, sunt suficiente pentru a produce fenomenul de interes însă niciuna nu este necesară întrucât sunt mai multe căi posibile de a produce fenomenul.

O situație puțin mai complexă este reprezentată de necesitatea unei disjuncții cauzale, unde sunt posibile toate situațiile posibile privind relația individuală dintre fiecare condiție cauzală și fenomenul rezultat:

- atât A cât și B sunt individual necesare pentru Y
- doar una dintre A sau B este necesară pentru Y
- nici A, nici B nu sunt necesare pentru Y

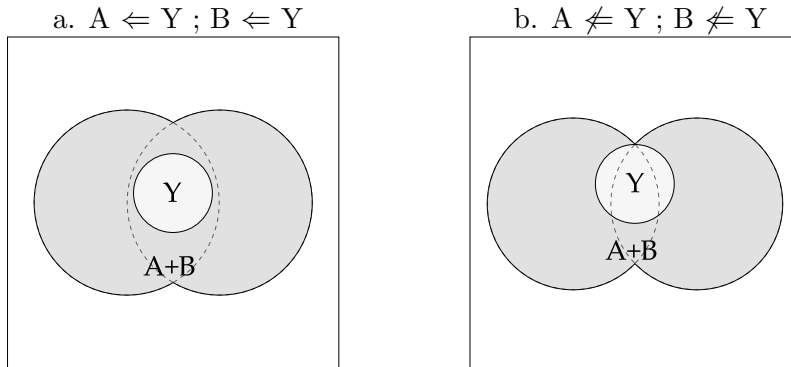
Figura 5.9: Necesitatea unei disjuncții cauzale:  $A + B \Leftarrow Y$ 

Figura 5.9.a reprezintă cea mai simplă posibilitate, în care atât A cât și B sunt individual necesare pentru Y, întrucât mulțimea Y este inclusă atât în A, cât și în B. În această situație, conjuncția dintre A și B este și ea necesară pentru Y. De exemplu, pentru a obține o diplomă universitară este necesară plata taxei anuale, dar este necesară și frecventarea cursurilor. Ambele sunt necesare, după cum este și disjuncția dintre acestea: este necesară “plata taxei sau frecventarea cursurilor”.

Interpretarea figurii 5.9.b este una particulară, întrucât nici A nici B nu sunt necesare pentru producerea lui Y, însă disjuncția dintre cele două este necesară. Mulțimea Y nu este complet inclusă nici în mulțimea A nici în mulțimea B, în schimb este complet inclusă în uniunea dintre mulțimea A și mulțimea B. Faptul că disjuncția  $A + B$  este necesară pentru Y poate reprezenta o dovadă a existenței unui concept de un nivel mai înalt care să fie necesar pentru Y, caz în care se recurge la teorie pentru analizarea relației dintre A și B.

Acest proces este cumva similar analizei de regresie din statistică, unde variația fiecărei variabile independente trebuie să fie unică (asumpția de non-multicolinearitate). Dacă două variabile împart o cantitate de variație comună, atunci se presupune că acele variabile măsoară același concept, recomandarea fiind de construire a unui scor agregat (eventual, în urma aplicării unei analize factoriale). Faptul că variabilele respective variază la fel poate fi explicat de prezența unui factor latent care le influențează pe ambele.

Crearea unei disjuncții (uniuni) de condiții cauzale pentru a crea o supra-mulțime necesară pentru Y este un proces simplu, în legătură cu care Schneider și Wagemann (2012) avertizează asupra pericolului unei aplicări mecanice deoarece se pot obține disjuncții necesare care să nu fie explicabile teoretic. De aceea, dacă niciuna dintre condițiile cauzale



nu este necesară pentru  $Y$ , se recomandă construirea unei disjuncții doar în situația în care supra-mulțimea obținută este interpretabilă teoretic și poate fi încadrată într-una din paradigmele acceptate.

În plus, introducerea unei noi mulțimi într-o disjuncție face ca suprafața acoperită de  $Y$  în interiorul disjuncției să fie din ce în ce mai mică. Cu cât disjuncția este mai mare, cu atât mai irelevantă devine relația de necesitate a acesteia pentru fenomenul de interes.

La limită, putem introduce toate condițiile cauzale imaginabile în disjuncție, însă suprafața acoperită de  $Y$  va fi atât de mică încât vor exista foarte multe cazuri în care se va întâmpla disjuncția dar nu și fenomenul, așa încât disjuncția nu va mai avea absolut nicio relevanță.

### 5.1.5 Tabela de adevăr

Algebra booleană are o aplicabilitate directă în analiza comparativă, prin utilizarea celor două simboluri în termeni de absență ori prezență a cauzelor și efectelor. Canoanele inductive ale lui Mill cât și strategiile prezentate de Przeworski și Teune utilizează exact această interpretare simbolică: două instanțe pot să se deosebească printr-o singură cauză (ceea ce înseamnă că în prima instanță cauza este *prezentă* iar în cea de a doua este *absentă*); două cazuri sunt similare prin apariția efectului, ceea ce înseamnă că efectul este *prezent* în ambele.

Să presupunem că se studiază un fenomen  $Y$ , despre care presupunem că este declanșat de condițiile cauzale  $A$  și  $B$ . Se efectuează un studiu de caz (în care se produce fenomenul) și se constată că în acel caz particular este prezentă condiția cauzală  $A$  însă condiția  $B$  este absentă. Putem spune așadar că acel caz:

- aparține mulțimii cazurilor în care fenomenul se produce; mai sunt și alte cazuri în care fenomenul nu se produce, iar acestea aparțin mulțimii negative  $\sim Y$
- aparțin mulțimii cazurilor care prezintă condiția cauzală  $A$ ; cazurile care nu prezintă această condiție cauzală aparțin mulțimii negative  $\sim A$
- nu aparține mulțimii cazurilor care prezintă condiția cauzală  $B$ , ci aparține mulțimii negative  $\sim B$

Prezența unei condiții cauzale indică apartenența la mulțimea elementelor cu aceeași caracteristică, iar absența unei condiții cauzale indică apartenența la mulțimea complementară, negativă.

Tabelul 5.4: Combinațiile pentru două, respectiv trei condiții cauzale

A	B	A	B	C
0	0	0	0	0
0	1	0	0	1
1	0	0	1	0
1	1	0	1	1
		1	0	0
		1	0	1
		1	1	0
		1	1	1

Analiza comparativă utilizează un număr limitat de cazuri pentru analiză, fiecare dintre acestea prezentând o anumită combinație de prezențe/absențe a condițiilor cauzale. Unele cazuri pot fi identice din acest punct de vedere, altele pot fi total diferite, iar metodologia analizei presupune aranjarea cazurilor în grupuri de combinații cauzale.

Numărul combinațiilor posibile este un număr finit, calculat cu formula  $2^k$ , unde  $k$  este numărul de condiții cauzale.

Pentru două combinații cauzale, numărul combinațiilor posibile de prezență/absență a acestora este de  $2^2 = 4$ ; pentru trei condiții cauzale, numărul combinațiilor posibile este de  $2^3 = 8$ . Tabelul 5.4 prezintă în partea stângă cele patru combinații pentru condițiile cauzale A și B, iar în partea dreaptă cele 8 combinații pentru condițiile cauzale A, B și C. Pentru fiecare situație în parte, numărul combinațiilor este același cu numărul intersecțiilor din figura 4.2: patru în partea de sus și opt în partea de jos.

În cele ce urmează, construirea tabelii de adevăr se bazează pe contribuția inițială a lui Ragin (1987) la analiza calitativ comparativă, însă odată cu apariția analizei pe mulțimi vagi acest algoritm a fost semnificativ modificat și îmbunătățit. Totuși, este utilă urmărirea procedurii originale pentru întărirea conceptelor cu care se lucrează în prezent.

În afară de prezentarea tuturor combinațiilor posibile, rolul tabelii de adevăr este acela de a arăta, pe de o parte, câte cazuri există în fiecare combinație cauzală și, pe de altă parte, dacă există conflicte în ce privește prezența ori absența fenomenului de interes. De exemplu, tabelul 5.5 prezintă distribuția a 12 cazuri ipotetice, din care în 9 dintre acestea fenomenul de interes este prezent iar în restul de 3 fenomenul este absent.

Tabelul 5.5: Tabelă de adevăr ipotetică pentru trei condiții cauzale

A	B	C	Y	~Y	Rezultat
0	0	0	3	0	1
0	0	1	0	0	?
0	1	0	0	2	0
0	1	1	1	1	C
1	0	0	2	0	1
1	0	1	1	0	1
1	1	0	0	0	?
1	1	1	2	0	1
Total			9	3	

Două cazuri în care fenomenul nu se produce prezintă combinația de pe linia a treia ‘0 1 0’, ceea ce înseamnă că nu există nici un alt caz studiat, cu acea combinație, în care fenomenul să fie prezent: în dreptul coloanei Y (prezența fenomenului) este cifra 0 iar în dreptul coloanei ~Y (absența fenomenului) este cifra 2. Aceste coloane reprezintă niște distribuții de frecvențe absolute, în care cifra 0 înseamnă “zero cazuri”, cifra 1 înseamnă “un caz” ș.a.m.d.

Ultima coloană a tabelului este coloana rezultat, reprezintă *decizia* cu privire la consistența relației dintre combinațiile cauzale ale cazurilor studiate. În prima linie, toate cele trei cazuri studiate prezintă fenomenul de interes, ceea ce înseamnă că pe ultima coloană se trece cifra 1 (care mai înseamnă *adevărat* sau *prezență*).

În dreptul liniei a treia este cifra 0, pentru absența fenomenului de interes: toate cele două cazuri studiate sunt consistente în această privință.

În dreptul liniei a doua este un semn de întrebare, ceea ce înseamnă că nu putem lua o decizie pentru acea combinație cauzală: nu putem spune dacă fenomenul de interes este prezent ori absent întrucât combinația respectivă nu a fost întâlnită în nici unul dintre cele 12 cazuri studiate (0 cazuri atât în coloana Y cât și în coloana ~Y). Același semn de întrebare este notat și în dreptul combinației de pe a șaptea linie.

O linie specială este a patra, în care se poate observa o contradicție a combinației ‘0 1 1’ pentru care într-unul din cele două cazuri fenomenul este prezent iar în celălalt este absent. Despre această combinație cauzală nu se poate spune cu siguranță dacă declanșează sau nu fenomenul de interes, fapt pentru care pe ultima coloană a tabelului de adevăr este trecută

litera C (de la rezultat ‘C’ontradictoriu).

Într-o tabelă de adevăr există așadar patru decizii cu privire la rezultatul unei combinații cauzale: prezența fenomenului (1), absența acestuia (0), situație contradictorie (C) și lipsa completă a rezultatului (?). Toate aceste coduri au un rol special în procesul de minimizare booleană, prezentat în secțiunea următoare.

În finalul acestei secțiuni este util de menționat că exemplele menționate sunt în mod intenționat rudimentare, pentru claritatea expunerii. Ele se bazează de fapt pe indicațiile lucrării originale a lui Ragin (1987), însă odată cu apariția analizei pe mulțimi vagi construirea tabelii de adevăr se bazează pe un algoritm modern puțin mai complicat, care atribuie decizia pentru fiecare linie a tabelii, ținând cont de toate celelalte linii.

### 5.1.6 Minimizarea booleană

Procesul de minimizare booleană are rolul de a găsi o combinație minimă de condiții responsabile pentru producerea unui fenomen. Numărul potențial de condiții, dar mai ales numărul combinațiilor posibile este extrem de mare, de aceea trebuie adus la un minim prin eliminarea acelor care nu influențează producerea fenomenului în mod semnificativ. În statistica pe eșantioane mari, procedura se numește “adecvarea modelului” (în engleză *model fit*), care presupune eliminarea variabilelor care nu au o influență semnificativă asupra variabilei dependente, până când se obține un model suficient de adecvat. Se urmărește găsirea unui număr cât mai mic de variabile independente, care să explice o proporție cât mai mare din variația dependentei.

Analiza calitativ-comparativă urmărește deasemenea eliminarea variabilelor redundante (ori contextuale) și păstrarea doar a combinațiilor cauzale care declanșează întotdeauna fenomenul studiat. Acest proces poartă numele de *minimizare booleană*.

Există un număr finit de combinații cauzale posibile; așa cum este prezentat în secțiunea precedentă, acesta este egal cu  $2^k$  unde  $k$  este numărul de condiții cauzale. Întrucât unele dintre aceste condiții cauzale sunt irelevante cu privire la declanșarea unui fenomen, în urma procesului de minimizare se obține una sau mai multe combinații cauzale mai mici decât oricare dintre combinațiile cauzale inițiale.

Un exemplu simplu care utilizează două condiții cauzale este reprezentarea binară a canonului concordanței al lui Mill, prezentat în tabelul 5.6. După cum se poate vedea, efectul este prezent în ambele instanțe cauzale, care concordă prin prezența cauzei A și absența cauzei

Tabelul 5.6: Minimizarea unei perechi de condiții cauzale

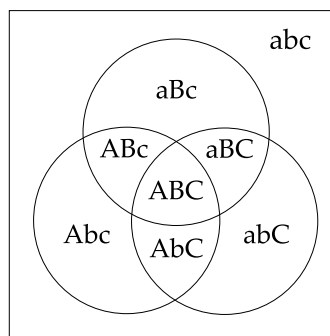
A	B	C	Fenomen
1	0	0	1
1	1	0	1
1	-	0	

C, în timp ce cauza B este într-o instanță absentă iar în cealaltă prezentă. Dacă aplicăm canonul concordanței, circumstanțele prin care cele două cazuri concordă sunt cauze posibile ale fenomenului, ceea ce înseamnă că a doua cauză este irelevantă întrucât nu concordă cu prezența efectului. Cauza B poate fi minimizată, ceea ce înseamnă că soluția minimă conține doar două condiții cauzale din cele trei inițiale: prezența cauzei A, simultan cu absența cauzei C, produce efectul.

Exemplul poate fi mai ușor înțeles dacă se utilizează o a treia notație pentru algebra booleană: prezența unei cauze cu literă mare iar absența unei cauze cu literă mică. În literatura de specialitate, reprezentările de prezență/absență a cauzelor prin litere se numesc “literali” (în engleză *literals*). Figura 5.10 prezintă intersecțiile de mulțimi prin literali, care sunt chiar mai simplu de înțeles: de exemplu, spațiul din afara celor trei mulțimi (negația simultană a tuturor mulțimilor) este mai ușor de înțeles utilizând notația ‘abc’ decât notația ‘ $\sim A \sim B \sim C$ ’.

În tabelul 5.6 pot fi identificate două combinații:  $Abc$  pe prima linie și  $ABc$  pe a doua linie, a căror expresie logic echivalentă este  $Abc + ABc$ .

Figura 5.10: Mulțimile în notația cu literali



Tabelul 5.7: Tabela de adevăr în notația cu literali

A	B	C	Literali
0	0	0	abc
0	0	1	abC
0	1	0	aBc
0	1	1	aBC
1	0	0	Abc
1	0	1	AbC
1	1	0	ABc
1	1	1	ABC

Algoritmul de minimizare Quine-McCluskey este capabil de a simplifica (de a minimiza) această expresie utilizând două proprietăți:

- Orice expresie logică poate fi rescrisă utilizând legea distributivă;
- Reuniunea dintre  $X$  și  $x$  (cu alte cuvinte  $X + x$ ) este întotdeauna egală cu 1 (sau este întotdeauna o expresie adevărată). De exemplu, expresia lingvistică “un specimen este mascul sau femelă” este întotdeauna adevărată întrucât cuprinde toate cazurile posibile, atât pozitive cât și negative.

Utilizând legea distributivă, expresia  $Abc + ABc$  poate fi rescrisă (prin scoaterea factorului comun  $Ac$ ) sub forma  $Ac(b + B)$ . Deoarece suma (reuniunea) dintre paranteze ( $b + B$ ) este egală cu 1, rezultatul este  $Ac$  iar concluzia este aceeași: efectul este produs de acțiunea concomitentă a *prezenței* condiției  $A$  și *absenței* condiției  $C$ . În comparație cu combinațiile inițiale care cuprindeau trei literali, expresia obținută este minimă întrucât conține doar doi literali. Acesta este o altă modalitate foarte simplă de prezentare a minimizării unei ecuații booleene, utilizând metoda concordanței a lui Mill.

Pentru comparația dintre două studii de caz (cum este exemplul de mai sus) este suficientă o inspecție vizuală, însă pentru un număr mai mare de cazuri studiate este nevoie de ajutorul unui computer pentru a analiza și minimiza toate perechile de cazuri posibile.

În acest scop, de-a lungul timpului au fost adoptate mai multe soluții software concurente. În general, prin programare se pot obține aceleași rezultate prin coduri diferite, acestea diferențiindu-se între ele doar prin

viteză și consum de resurse. Unele soluții pot fi extrem de lente și să utilizeze resurse enorme de calcul, iar altele pot fi extrem de rapide, consumând resurse puține. Resursele calculatoarelor (atât de memorie cât și de calcul) prezintă o traiectorie puternic ascendentă. Atât spațiul de stocare al hard-disk-urilor, memoria RAM cât și viteza micro-procesoarelor au cunoscut o creștere de-a dreptul spectaculoasă iar previziunile sunt și mai optimiste odată cu adoptarea nano-tehnologiilor. De aceea, se disting acei algoritmi care reușesc să aducă un plus de viteză, deși pentru algoritmul generic de minimizare booleană nici o resursă nu este prea mare întrucât complexitatea este exponențială cu fiecare nouă condiție cauzală introdusă în analiză. În genere, fiecare nou algoritmul încearcă să-i depășească pe cei precedenți, aducând fie un plus în ce privește viteza, fie o scădere în ce privește consumul de resurse, fie (cazul cel mai fericit) ambele.

Creatorul fiecărui nou algoritmul îl înglobează de obicei într-un pachet software. Metodologia care urmează a fi prezentată în următoarele paragrafe este general valabilă pentru oricare dintre soluțiile software existente, toate având la bază tezele lui Shannon și McCluskey. În științele sociale, algoritmul standard utilizat este QCA - Qualitative Comparative Analysis dezvoltat de Ragin (1987), care a adaptat procedura Quine-McCluskey în sociologie și științele politice. În fapt, fundamentul ideii lui Shannon vine tot din științele sociale, el fiind inspirat de un curs de filozofie pe care l-a urmat când era student, mai târziu arătând cum se aplică logica booleană la circuitele electrice.

Algoritmul Quine-McCluskey procedează în mod iterativ la compararea tuturor perechilor posibile de cazuri, minimizând progresiv acele perechi care diferă doar printr-un singur literal, în final rezultând un set unic de combinații cauzale care nu mai pot fi minimizate mai mult.

Pentru exemplificare voi utiliza chiar tabelul 5.5: din cele opt combinații cauzale posibile, doar patru sunt asociate fără nici un dubiu cu prezența fenomenului de interes; restul sunt fie cazuri negative, fie cazuri contradictorii ori în care nu se cunoaște starea fenomenului de interes. Întrucât obiectivul analizei este de a explica *aparitia* fenomenului (altfel spus prezența acestuia), vom selecta ca date de intrare doar cele patru combinații cauzale asociate (tabelul 5.8).

Pentru patru cazuri, există exact 12 perechi posibile de câte două cazuri (de câte două numere, dacă notăm cazurile selectate cu 1, 2, 3 și 4). Aceste perechi sunt: (1,2), (1,3), (1,4), (2,1), (2,3), (2,4), (3,1), (3,2), (3,4), (4,1), (4,2) și (4,3).

Se observă însă că toate perechile sunt dublate, întrucât a compara perechea (1,2) este același lucru cu a compara perechea (2,1). De aceea,

în prima iterație a algoritmului vor fi comparate doar cele șase perechi unice: (1,2), (1,3), (1,4), (2,3), (2,4) și (3,4). Acestea vor fi supuse unui proces de minimizare similar cu cel din tabelul 5.6:

- perechea (1,2) este compusă din combinațiile '0 0 0' și '1 0 0', care (întrucât diferă printr-un singur literal) poate fi minimizată în combinația '– 0 0' (sau prin literali 'bc')
- perechea (1,3) este compusă din combinațiile '0 0 0' și '1 0 1', care nu poate fi minimizată întrucât diferă prin doi literali
- perechea (1,4) este compusă din combinațiile '0 0 0' și '1 1 1', care nu poate fi minimizată întrucât diferă prin trei literali
- perechea (2,3) este compusă din combinațiile '1 0 0' și '1 0 1', care (întrucât diferă printr-un singur literal) poate fi minimizată în combinația '1 0 –' (sau prin literali 'Ab')
- perechea (2,4) este compusă din combinațiile '1 0 0' și '1 1 1', care nu poate fi minimizată întrucât diferă prin doi literali
- perechea (3,4) este compusă din combinațiile '1 0 1' și '1 1 1', care (întrucât diferă printr-un singur literal) poate fi minimizată în combinația '1 – 1' (sau prin literali 'AC')

Nici una dintre cele trei combinații rămase în analiză (bc, Ab și AC) nu mai pot fi minimizate mai mult, întrucât fiecare este diferită de celelalte cu mai mult de un singur literal. În limbajul minimizării booleene a lui McCluskey, cele trei combinații cauzale minimizate se numesc *implicanți primari*.

Acest exemplu s-a încheiat într-o singură iterație. Pentru un exemplu mai complicat am putea presupune că cele două instanțe în care nu se cunoaște dacă fenomenul apare (liniile 2 și 7 din tabelul 5.5) sunt în fapt instanțe pozitive ale apariției fenomenului (asumpție tratată mai pe larg

Tabelul 5.8: Cele patru combinații cauzale analizate

	A	B	C	Literali
1	0	0	0	abc
2	1	0	0	Abc
3	1	0	1	AbC
4	1	1	1	ABC



în secțiunea 2.6.5). Analiza mai câștigă astfel două cazuri care pot intra în procesul de minimizare, iar datele de intrare sunt:

A	B	C	Literali
0	0	0	abc
0	0	1	abC
1	0	0	Abc
1	0	1	AbC
1	1	0	ABc
1	1	1	ABC

Pentru șase cazuri, numărul total posibil de comparații pe perechi este de  $C_6^2 = 30$ , care împărțit la jumătate (întrucât perechile se repetă) ajunge la 15 comparații posibile. Urmând același fir logic ca în exemplul precedent obținem:

- perechile (1,4), (1,5), (1,6), (2,3), (2,5), (2,6), (3,6), (4,5) nu pot fi comparate
- perechea (1,2): '0 0 0' și '0 0 1' rezultă în '0 0 -' (sau 'ab')
- perechea (1,3): '0 0 0' și '1 0 0' rezultă în '- 0 0' (sau 'bc')
- perechea (2,4): '0 0 1' și '1 0 1' rezultă în '- 0 1' (sau 'bC')
- perechea (3,4): '1 0 0' și '1 0 1' rezultă în '1 0 -' (sau 'Ab')
- perechea (3,5): '1 0 0' și '1 1 0' rezultă în '1 - 0' (sau 'Ac')
- perechea (4,6): '1 0 1' și '1 1 1' rezultă în '1 - 1' (sau 'AC')
- perechea (5,6): '1 1 0' și '1 1 1' rezultă în '1 1 -' (sau 'AB')

În prima iterație au rezultat așadar 7 implicați:

A	B	C	Literali
0	0	-	ab
-	0	0	bc
-	0	1	bC
1	0	-	Ab
1	-	0	Ac
1	-	1	AC
1	1	-	AB

Se poate observa că unii dintre aceștia pot fi minimizați mai departe, de exemplu perechea 'bc' și 'bC' care va rezulta în 'b'. Este necesară o a doua iterație, de această dată cu 21 de comparații posibile pe perechi, din care singurele care pot fi mai mult minimizate sunt:

- perechea (1,4): '0 0 -' și '1 0 -' rezultă în '- 0 -' (sau 'b')
- perechea (2,3): '0 0 -' și '1 0 -' rezultă în '- 0 -' (sau 'b')
- perechea (4,7): '1 0 -' și '1 1 -' rezultă în '1 - -' (sau 'A')
- perechea (5,6): '1 - 0' și '1 - 1' rezultă în '1 - -' (sau 'A')

Din această a doua iterație a rezultat un număr de doi implicații care nu pot fi minimizați mai departe, așa încât aceștia sunt denumiți implicații primari, prezentați în tabelul 5.9.

Tabelul 5.9: Implicații primari rezultați în urma procesului de minimizare

A	B	C	Literali
-	0	-	b
1	-	-	A

Utilizând aceste date, putem concluda că soluția finală este 'A + b', care înseamnă că fenomenul de interes este declanșat de prezența cauzei A sau de absența cauzei B. Am obținut așadar nu doar o concluzie minimă și simplu de înțeles ci mai mult am reușit să identificăm faptul că a treia condiție cauzală (C) nu are nici o legătură cu producerea fenomenului (fapt care la prima vedere nu era chiar atât de evident).

Ragin a fost primul care a intuit utilitatea acestei tehnici pentru reprezentarea cauzelor și efectelor în științele sociale; sigur, munca lui este puternic influențată de canoanele inductive ale lui Mill (1843) însă are și meritul de a fi reușit o legătură între științele ingineresti și cele sociale. De natură calitativă, QCA reușește în acest fel să elimine perspectivele individuale subiective prin adoptarea acestui algoritm de minimizare care generează întotdeauna aceeași soluție. Rezultatele devin așadar replicabile, iar analiza dobândește acel tip de încredere specific cercetării cantitative.

Cercetătorii cantitativiști utilizează un număr mare de cazuri pentru a face inferențe bazate în primul rând pe corelații, însă atenția lor este

îndreptată exclusiv spre diferite grade de *prezență* a cauzelor unui fenomen. Utilizând logica binară, QCA ia în calcul atât prezența cât și absența condițiilor cauzale. Într-adevăr, un rezultat poate fi produs și de absența unei cauze. De exemplu am putea cu ușurință să ne imaginăm că anarhia sau haosul unui sistem poate fi declanșat (printre alți factori, și) de absența unui sistem legislativ coerent.

Soluțiile QCA sunt prezentate în combinații minime de prezență/absență a unor condiții cauzale, necesare și/sau suficiente pentru producerea unui fenomen. Pornind de la un număr relativ mic de combinații *observate*, analiza parcurge un proces de minimizare pentru a obține combinațiile minime. Pentru un număr mic de condiții cauzale (factori), procesul de minimizare poate fi parcurs și pe hârtie însă fiecare nouă condiție cauzală introdusă în analiză mărește numărul combinațiilor posibile în mod exponențial.

### 5.1.7 Legile lui DeMorgan

Pe vremea când calculatoarele nu erau atât de puternice iar algoritmi de minimizare booleană nu erau așa de avansați, obținerea unei soluții consuma extrem de mult timp. Pentru 10 condiții cauzale calculatoarele de acum 15 ani rulau algoritmul nopți întregi, în timp ce calculatoarele și algoritmi actuali obțin o soluție aproape instantaneu.

În analiza calitativ-comparativă, este câteodată necesar să se compare soluțiile pentru prezența fenomenului cu soluțiile pentru absența acestuia. Cu alte cuvinte, pot fi obținute soluții studiind cazurile în care fenomenul de interes se produce, însă idei la fel de interesante pot fi extrase studiind cazurile opuse, în care fenomenul nu se produce. Soluția obținută ar trebui să fie și aceasta inversă: dacă o condiție cauzală importantă se manifestă atunci când fenomenul se produce, ne așteptăm ca aceasta să lipsească în situațiile în care fenomenul nu se produce.

În loc să ruleze încă odată algoritmul pe cazurile opuse, cercetătorii preferau aplicarea unor reguli ale logicii, observate și formulate pentru prima dată de către Augustus DeMorgan (Ragin, 1987, p.98). Studiind logica propozițională, DeMorgan a observat că algebra booleană urmează un principiu tranzitiv:

$$\begin{aligned} \text{NOT}(A \text{ \texttt{ȘI}} B) &= (\text{NOT } A) \text{ SAU } (\text{NOT } B) \\ \text{NOT}(A \text{ SAU } B) &= (\text{NOT } A) \text{ \texttt{ȘI}} (\text{NOT } B) \end{aligned}$$

Utilizând pentru consistență aceleași notații ('~' pentru negație, '+' pentru SAU exclusiv și '.' pentru ȘI exclusiv), observațiile lui de Morgan pot fi rescrise astfel:

$$\begin{aligned}\sim(A \cdot B) &= \sim A + \sim B \\ \sim(A + B) &= \sim A \cdot \sim B\end{aligned}$$

Utilizând aceste reguli, transformarea unei soluții în inversul acesteia este relativ simplă. Ragin oferă un exemplu al unor posibile cauze care pot să ducă la grevă:

A = piața produsului în dezvoltare puternică

B = posibilitatea unei greve sectoriale în solidaritate

C = existența unor fonduri suficient de mari pentru organizarea unei greve

Procesând datele disponibile, soluția găsită cu ajutorul algoritmului boolean este:  $AC + Bc$

Cu alte cuvinte, o grevă se poate produce în două situații:

$AC$  = atunci când există o dezvoltare puternică a pieței și există fonduri suficient de mari pentru organizarea unei greve

sau

$Bc$  = chiar dacă nu există fonduri suficient de mari pentru organizarea unei greve ('c'), greva se poate produce dacă există posibilitatea unei greve simultane pe mai multe ramuri industriale ('B')

Pentru a afla condițiile în care greva *nu* se produce, putem aplica legile lui de Morgan:

$$\sim(AC + Bc) = \sim(AC) \cdot \sim(Bc) = (a + c)(b + C) = ab + aC + bc$$

Soluția finală a fost obținută prin aplicarea unor reguli algebrice elementare. Se poate observa însă că produsul 'c·C' lipsește din soluția finală, fapt explicabil dacă îl traducem în cuvinte: este intersecția dintre o mulțime și negația acesteia, care este mulțimea vidă.

Pentru soluții mai complicate, aplicarea legilor lui DeMorgan cer o atenție sporită, însă timpul de obținere a soluției era oricum mult mai scurt decât aplicarea întregului algoritm pe cazurile negative. Computerele și algoritmi moderni reduc acest timp la câteva zecimi de secundă, așa încât aplicarea manuală a acestor legi a devenit inefficientă în comparație cu rapiditatea computerelor.

Cu aplicare la procesul de minimizare booleană, Schneider și Wagemann (2012, p.114) avertizează asupra faptului că legile lui DeMorgan nu pot fi aplicate decât în situația în care tabela de adevăr este completă (toate combinațiile cauzale să existe în realitate, să fie observate în procesul culegerii datelor). Cum însă acest deziderat este dificil de îndeplinit în practică (din cauza problemei diversității limitate), legile lui deMorgan au o aplicabilitate redusă în procesul de minimizare booleană.

### 5.1.8 Problema codificărilor binare

După apariția cărții originale a lui Ragin (1987) în care a introdus metoda calitativ comparativă, aceasta a suferit nenumărate atacuri metodologice prin care s-a încercat demonstrarea inutilității acesteia, în special din partea cercetătorilor din științele cantitative.

De exemplu, un atac a lui Goldthorpe (1997a) se referă la specificul analizei calitativ-comparative (QCA) în care codificarea se face în doar două categorii: 0 pentru absență și 1 pentru prezență. Goldthorpe observă că cele mai multe condiții cauzale (ori chiar fenomene sociale) au o natură continuă, cu un interval de variație foarte larg și că recodificarea variabilelor în doar două valori (absență și prezență) este una dintre slăbiciunile fundamentale care viciază acest tip de analiză.

Problema atribuirii codurilor de 1 și 0 (prezență ori absență) este într-adevăr una delicată, întrucât lucrurile nu sunt întotdeauna foarte clar delimitate între acești doi poli. Goldthorpe are dreptate când spune că fenomenele sociale nu au o manifestare discretă, ci una continuă, cu o infinitate de valori posibile. Codificarea în prezență ori absență depinde în totalitate de experiența și intuiția cercetătorului care conduce ancheta. La un număr mic de cazuri, dacă o anumită condiție cauzală este codificată în 0 în loc de 1 (ori invers) atunci când situația este incertă, rezultatul minimizării booleene poate fi în totalitate diferit.

La această problemă, Ragin (1987, 1994) avertizează că procesul de codificare nu trebuie făcut în mod mecanic, ci numai în urma unui dialog permanent între teorie și datele disponibile. Codurile nu se atribuie după subiectivitate, ci în funcție de jaloanele prin care teoria ghidează cercetarea. Dacă teoria nu este suficient de adecvată realității empirice, va fi înlocuită în cele din urmă cu o alta, iar poziția jaloanelor se va schimba, însă codificarea trebuie întotdeauna să țină seama de aceste jaloane pentru a elimina riscul subiectivității. De exemplu, cineva ar putea să se întrebe care este gradul de implicare în societatea civilă în România; pentru acest lucru poate fi construit un indice cuprins din mai mulți itemi care vizează voluntariatul ori înscrierea în diverse asociații care acoperă activități ale societății civile (mai puțin partidele politice ori sindicatele), iar scorul final poate fi rescalat pe un interval între 0 și 100% (ori între 0 și 1).

Din cadrul a mai multe rapoarte al Barometrului de Opinie Publică, se poate arata faptul că în România post-decembristă implicarea în societatea civilă este extrem de scăzută. În mod normal, poate fi stabilit un prag la jumătate (50%) în funcție de care stabilim dacă în România există ori nu societate civilă: dacă procentul de populație implicată

depășește acest prag atunci codificăm cu valoarea 1, în caz contrar cu valoarea 0. Totuși, o astfel de strategie nu va da niciodată rezultate, întrucât în România nu s-au înregistrat niciodată astfel de procente (mai mult, puține sunt țările în care procentul de populație implicată în societatea civilă depășește acest prag). În fapt, ca și în cazul măsurării sărăciei, pragul respectiv nu poate și nici nu trebuie să fie unul absolut și fix, ci trebuie specificat în funcție de contextul particular al fiecărei țări (ori al unei comunități, dacă analiza comparativă se desfășoară la un nivel de agregare mai scăzut).

Așadar, stabilirea pragului în funcție de care se face codificarea cazurilor este un proces care cere multă atenție din partea cercetătorului. Mai mult, este un proces care trebuie din plin justificat teoretic așa încât oricine dorește replicarea analizei să ajungă la aceleași rezultate. Deși Goldthorpe (1997a) avertizează că o codificare atât de restrictivă obligă la raționamente deterministe, raționamentul contrar al lui Goldstone (1997), prezentat la secțiunea 2.6.4, este în opinia mea mult mai convingător.

O posibilă soluție la această problemă este mutarea analizei de la mulțimile distincte spre mulțimile vagi, perspectivă care va fi discutată în secțiunea 5.3.

## 5.2 Analiza cu mulțimi distincte multivalente

Cronqvist și Berg-Schlosser (2009) descriu mulțimile multivalente ca pe o extensie a mulțimilor bivalente (în engleză mvQCA - multi-value QCA, față de csQCA - crisp sets QCA). În fapt, dintr-un alt punct de vedere, poate fi ușor demonstrat că analiza pe mulțimi bivalente nu este nimic altceva decât un caz particular al analizei pe mulțimi multivalente.

Deși în literatura de specialitate s-a impus sintagma “multi-value”, în această carte este preferată denumirea de “multivalent”, pentru cel puțin câteva motive:

- entitățile cu care lucrăm în științele sociale nu reprezintă întotdeauna valori numerice, uneori sunt simple categorii verbale;
- chiar și în situația în care entitatea cu care operăm este numerică, uneori este dificil de împărțit în exact două categorii, cu exemple concrete în această secțiune;
- algoritmul de minimizare prezentat de Dușa (2010) utilizează numere în mai multe baze (de exemplu trece întreaga tabelă de adevăr din baza 2 în baza 3) iar sintagma “multi-value” nu face nicio referire la baza de calcul în care se lucrează.

Una dintre problemele analizei pe mulțimi bivalente, menționată în secțiunea anterioară, este legată de procesul de dihotomizare în prezență / absență a condițiilor cauzale. Împărțirea pe valorile 1 și 0 este similară cu crearea variabilelor dummy din analiza cantitativă și poate fi efectuată pe orice variabilă categorială, cu prezența și absența categoriei de interes.

Uneori însă, condițiile cauzale introduse într-un model pot avea mai mult de două categorii. În secțiunea 4.2.2 am prezentat câteva exemple, printre care cel al semaforului cu cele trei categorii (roșu, galben și verde) pentru care este dificil de interpretat crearea unei condiții cauzale bivalente.

Desigur, ca și la crearea variabilelor dummy, este posibilă crearea de variabile binare pentru fiecare din cele trei categorii, însă pentru analiza calitativ-comparativă crearea mai multor condiții cauzale este problematică întrucât numărul total de combinații cauzale din tabela de adevăr crește exponențial cu numărul de condiții cauzale. În plus, unele combinații cauzale rezultate sunt pur și simplu imposibile, întrucât culorile semaforului nu apar niciodată simultan.

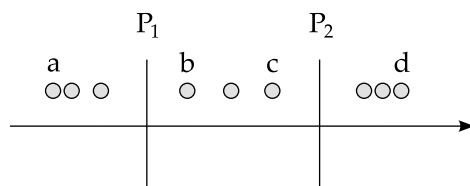
Motivul cel mai important pentru care comunitatea științifică a început să exploreze posibilitatea utilizării condițiilor multivalente este problema dihotomizării artificiale a condițiilor cauzale numerice. Prin utilizarea unui prag (oricât de obiectiv), cazuri asemănătoare pot fi recodificate în poli diferiți de prezență/absență, iar cazuri complet diferite pot fi atribuite aceluiași pol. Sunt situații în care distribuția observațiilor este în mod empiric grupată pe mai mult de două zone numerice, caz în care trasarea unui singur prag devine problematică.

Tabelul 5.10: Condiții bivalente și multivalente

Mediu	Bivalent	Semafor	Multivalent
Urban	1	Roșu	0
Rural	0	Verde	2
Urban	1	Roșu	0
Urban	1	Galben	1
Rural	0	Galben	1
Urban	1	Verde	2
Rural	0	Verde	2
Rural	0	Galben	1
Urban	1	Roșu	0

Din acest motiv, Cronqvist și Berg-Schlosser (2009, p.71) recomandă analiza prealabilă a datelor inițiale pentru detectarea grupărilor care apar în mod natural. Ei oferă și un exemplu sugestiv de date care nu permit o simplă dihotomizare, prezentat mai jos în figura 5.11.

Figura 5.11: Gruparea datelor pe mai mult de două categorii



Autorii avertizează asupra următoarelor posibilități într-un proces artificial de dihotomizare:

- dacă se utilizează doar pragul  $P_1$ , cazurile foarte diferite b și d ar intra în aceeași categorie din dreapta;
- pe de altă parte, dacă se utilizează doar pragul  $P_2$  ne vom afla într-o situație similară întrucât cazurile a și c vor fi atribuite aceleiași categorii din stânga, deși sunt foarte diferite;
- orice trasare artificială a unui alt prag, să spunem undeva între cazurile b și c ar despărți cazuri foarte asemănătoare în categorii diferite.

Acesta este un exemplu de date numerice care nu pot fi despărțite foarte ușor în doar două categorii. În situația prezentată, procedeul corect este gruparea cazurilor în trei categorii, în acest fel intrând în aria de acțiune a mulțimilor multivalente.

Înainte de apariția unui algoritm stabil de analiză a mulțimilor vagi, abordarea intermediară multivalentă a fost concepută de către Cronqvist (2003), care susținea încă de atunci că este o rezolvare a situației în care se înregistrează mai multe ordine de mărime ale unui fenomen, cu ajutorul unei scale intermediare (măsurată la nivel ordinal), între cea binară a mulțimilor distincte și cea continuă a mulțimilor vagi.

Această scală pornește tot din valoarea 0 (absența completă a fenomenului studiat), însă poate avea oricâte ordine de mărime în sus, în funcție de gradul în care un anumit fenomen este prezent. Baza de date crește în complexitate deoarece numărul posibil de combinații cauzale crește proporțional cu complexitatea fiecărei condiții cauzale în parte.



Orice valență adăugată unei condiții cauzale multiplică numărul total de combinații posibile în tabela de adevăr. De exemplu, dacă se studiază trei condiții cauzale dintre care prima are 3 valori, a doua cinci valori iar a treia două valori, numărul total de combinații este egal cu  $3 \times 5 \times 2 = 30$  de combinații cauzale posibile.

Stabilirea numărului optim de grupări se face în funcție de două criterii egal importante:

- să corespundă grupărilor naturale din distribuția valorilor inițiale;
- fiecare grupare să conțină un număr suficient de mare de cazuri.

Primul criteriu a fost discutat mai sus. Urmând același exemplu cu stabilirea a două praguri (deci se formează trei grupări), dacă la limită una dintre grupări conține un singur caz, atunci explicația influenței acelei grupări la producerea fenomenului de interes este complet particularizată și se pierde caracterul de generalitate al unei ipoteze științifice.

În opinia Cronqvist și Berg-Schlosser (2009), trei sau patru grupări sunt de obicei arhi-suficiente pentru satisfacerea tuturor criteriilor. Acest punct de vedere arată experiența lor de lucru cu datele multivalente, opinia acestora fiind logică întrucât mărirea numărului de grupări duce automat la scăderea numărului de cazuri pe fiecare grupare. Cum analiza calitativ-comparativă lucrează, de obicei, cu un număr relativ restrâns de cazuri, este puțin probabil să avem suficiente observații pentru foarte multe grupări.

Beneficiul major la împărțirea unei condiții cauzale numerice în mai mult de două grupări este posibilitatea de a efectua comparații la un nivel de măsurare ordinal, spre deosebire de analiza pe mulțimi distincte bivalente unde nu se pot face comparații decât la nivel nominal (prezență / absență a condiției cauzale).

Soluția minimă finală este obținută cu ajutorul aceluiași principiu din algoritmul clasic Quine-McCluskey: minimizarea unei perechi de combinații cauzale se poate realiza doar dacă diferă printr-un singură condiție cauzală, cu toate valorile acesteia.

Tablelul 5.11 arată un exemplu teoretic de minimizare a unei perechi de condiții cauzale cu valori multiple: se poate observa că, în afară de condiția cauzală A, toate celelalte concordă cu prezența fenomenului. Prin contrast, condiția A prezintă în mod simultan toate cele trei valori odată cu producerea fenomenului studiat, așa încât putem trage concluzia că nu participă la producerea acestuia.

Tabelul 5.11: Minimizarea cu valori multiple

A	B	C	Fenomen
0	3	0	1
1	3	0	1
2	3	0	1
–	3	0	

Aceste perechi de combinații cauzale pot fi minimizată întrucât nu diferă decât prin condiția cauzală A, iar implicantul primar rezultat este ‘– 3 0’. În această situație nu se mai pot înlocui cifrele cu litere mari și mici (ca în varianta binară); dacă se doresc totuși utilizarea literelor care înlocuiesc condițiile cauzale, combinația minimizată poate fi scrisă ‘– B{3} C{0}’ (sau mai simplu ‘B{3}C{0}’).

În termeni formali, ecuația poate fi scrisă astfel:

$$A\{0\}B\{3\}C\{0\} + A\{1\}B\{3\}C\{0\} + A\{2\}B\{3\}C\{0\} \Rightarrow Y$$

Toate valorile condiției A se manifestă odată cu producerea fenomenului, așadar condiția A nu are nicio legătură cu prezența acestuia, deci ecuația se poate rescrie cu:

$$B\{3\}C\{0\} \Rightarrow Y$$

Dacă doar două dintre valorile lui A ar fi fost prezente, minimizarea nu se poate efectua:

$$A\{0\}B\{3\}C\{0\} + A\{1\}B\{3\}C\{0\} \Rightarrow Y$$

În această situație, lipsește valoarea A{2}, deci este posibil ca A să contribuie la producerea fenomenului Y, prin intermediul valorilor A{0} și A{1}.

Din această notație poate fi dedus și faptul că analiza pe mulțimi bivalente este un caz particular al analizei pe mulțimi multivalente. În tabelul 5.9 din secțiunea precedentă, în loc de notația ‘A + b’ cei doi implicați primari pot fi rescriși sub forma: ‘A{1} + B{0}’. Mulțimile distincte bivalente nu sunt altceva decât niște mulțimi distincte multivalente cu doar două valori (categorii, grupări).

Ca și în cazul celor bivalente, condițiile cauzale multivalente pot fi utilizate pentru a *elimina* condiții cauzale irelevante. În plus, Mahoney (2000) este

de părere că metoda ordinală reprezintă trecerea naturală de la analiza mulțimilor distincte, binare, la metoda variațiilor concomitente a lui Mill. Utilizarea nivelului de măsurare ordinal este așadar o primă formă de testare a unor ipoteze probabilistice.

Mahoney reia exemplul lui Lieberson cu șoferul beat (secțiunea 2.6.4) și afirmă că această ipoteză ar putea fi testată prin metodologia comparativă studiind diverse cazuri în care diferă *gradul* în care șoferul este beat. În combinație cu alte variabile (necesare ori suficiente), putem afirma în acest cadru dacă ipoteza lui Lieberson este sau nu adevărată.

Iată cum, utilizând o extensie a unei analize considerată deterministă, poate fi testată o ipoteză pe care până acum se credea că nu poate fi testată decât cu ajutorul instrumentelor cantitativ-probabiliste.

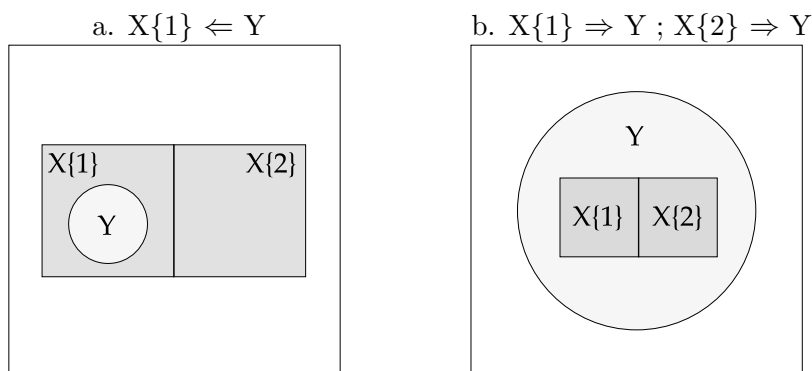
Deși variabilele ordinale nu sunt echivalente variabilelor metrice (în consecință nu poate fi calculat, de exemplu, un coeficient de corelație între o condiție cauzală și un efect), există metode de calculare a asocierii dintre două variabile calitative utilizând corelație de ranguri. În aceste condiții, metoda comparativă cu variabile ordinale poate fi utilizată pentru a *elimina* cauze ori combinații cauzale. Dacă o variabilă ordinală cauză (sau o combinație cauzală) nu este asociată cu o variabilă ordinală efect (ori dacă cele două nu au o corelație puternică de ranguri) atunci putem elimina cauza din setul de ipoteze.

În analiza pe mulțimi multivalente, fenomenul de interes nu poate lua decât maxim două valori: 1 și 0 (prezență / absență). Este singura variabilă care nu poate avea decât două valori, întrucât analiza calitativ-comparativă nu face altceva decât să caute combinații cauzale minime care să explice ‘producerea’ fenomenului, nu o stare intermediară a acestuia. De fapt, acest lucru este valabil pentru toate tipurile de analize, de la mulțimile bivalente până la mulțimile vagi, incluzând aici mulțimile multivalente.

Reprezentarea grafică a mulțimilor multivalente este similară diagraamelor Venn de la mulțimile bivalente. Pentru fiecare valență a condiției cauzale  $A$  vom avea câte o mulțime specifică (fiecare valoare este o mulțime în sine), singura diferență față de mulțimile bivalente fiind aceea că mulțimile corespunzătoare valorilor nu se vor putea intersecta niciodată întrucât acestea sunt mutual exclusive.

În figura 5.12 este prezentată relația dintre o condiție cauzală  $X$  cu trei valori (0, 1, 2) și un fenomen de interes  $Y$ . Fiecare valoare a condiției cauzale  $A$  are propria incluziune și propria acoperire în mulțimea  $Y$ . Sunt afișate doar două dintre cele trei valori (anume  $X\{1\}$  și  $X\{2\}$ ), întrucât categoria  $X\{0\}$  este reprezentată de spațiul din afara celor două.

Figura 5.12: Necesitate și suficiență multivalentă



Deși aparțin aceleiași condiții cauzale, valorile se comportă ca și cum ar fi mulțimi diferite. Chiar dacă reprezentarea grafică le arată lipite una de cealaltă, ideea de bază rămâne aceeași: mulțimile corespunzătoare valorilor sunt mutual exclusive deci nu se intersectează niciodată.

Din reprezentarea grafică a relației de necesitate din figura 5.12.a se poate vedea faptul că doar una dintre cele trei valori poate fi necesară, niciodată mai mult de una. Aceasta întrucât mulțimea  $Y$  nu poate fi inclusă în mod simultan în mai multe mulțimi mutual exclusive: dacă una dintre valori este necesară, înseamnă că restul valorilor nu sunt necesare.

În ce privește relația de suficiență, pot fi mai multe posibilități:

- nicio valoare suficientă (nici o mulțime nu este inclusă în  $Y$ )
- o singură valoare este suficientă (celelalte două fiind în afara lui  $Y$ )
- două valori suficiente (ambele fiind incluse în mulțimea  $Y$ )

Figura 5.12.b prezintă situația din urmă, în care ambele valori  $X\{1\}$  și  $X\{2\}$  sunt suficiente pentru producerea fenomenului. Rămâne valabil faptul că mulțimile celor două valori sunt mutual exclusive: ele sunt ambele incluse în  $Y$  însă nu se intersectează. Nu este posibil ca toate valorile să fie suficiente pentru  $Y$ , în mod simultan (în figură, mulțimea  $X\{0\}$  nu este inclusă în mulțimea  $Y$ ).

Relevanța oricărei valori pentru producerea fenomenului, dacă este o valoare suficientă, se calculează după aceleași criterii ca la mulțimile bivalente. Întrucât condițiile multivalente sunt tot distincte, se pot face tabele încrucișate doar că dimensiunea acestora nu va mai fi de  $2 \times 2$  ci de  $2 \times V$ , unde ' $V$ ' este numărul de valori ale condiției cauzale.

Tabelul 5.12: Tabel încrucișat 2×3

		X		
		0	1	2
Y	1	a	c	e
	0	b	d	f

Tabelul 5.12 conține 2 rânduri și 3 coloane, unde fenomenul studiat are 2 valori (prezență / absență) iar condiția cauzală X are 3 valori (0, 1 și 2). Valoarea 0 poate însemna “absență” dacă valorile inițiale erau de tip ordinal, sau poate însemna orice altă categorie dacă variabila inițială era măsurată la nivel nominal.

În rest, calcularea includerii și acoperirii au formule de calcul cu totul similare ca la analiza cu mulțimi distincte bivalente. De exemplu, incluziunea pentru ipoteza suficienței valorii  $X\{1\}$  pentru producerea fenomenului Y este:

$$incl_{X\{1\} \Rightarrow Y} = \frac{c}{c + d}$$

Similar, incluziunea pentru ipoteza suficienței valorii  $X\{2\}$  pentru producerea fenomenului Y este:

$$incl_{X\{2\} \Rightarrow Y} = \frac{e}{e + f}$$

Scorul de acoperire pentru ipoteza suficienței valorii  $X\{2\}$  pentru producerea fenomenului Y este:

$$cov_{X\{2\} \Rightarrow Y} = \frac{e}{a + c + e}$$

Urmând aceeași regulă care spune că includerea pentru suficiență este egală cu acoperirea pentru necesitate și că includerea pentru necesitate este egală cu acoperirea pentru suficiență, formula pentru includere în ipoteza necesității valorii  $X\{2\}$  pentru producerea fenomenului Y, formula de calcul este:

$$incl_{X\{2\} \Leftarrow Y} = \frac{e}{a + c + e}$$

În final, formula de calcul pentru acoperirea în ipoteza necesității valorii  $X\{1\}$  pentru producerea fenomenului  $Y$  este:

$$cov_{X\{2\} \leftarrow Y} = \frac{e}{e + f}$$

Algoritmul de minimizare a condițiilor multivalente este cu totul similar celui de la mulțimile bivalente, la final fiind produși niște implicații primari (combinații cauzale care nu mai pot fi reduse mai mult), iar prin combinarea acestora se ajunge la soluția minimă (sau soluțiile minime) care explică producerea fenomenului.

### 5.3 Analiza cu mulțimi vagi

În ceea ce privește condițiile necesare și suficiente, întreaga expunere din secțiunile anterioare se bazează pe asumția implicită a dihotomiei. O valoare fie este necesară, fie nu este necesară (se aplică principiul terțului exclus).

Toată această logică este direct descendentă din linia filozofică a lui Aristotel (secțiunea 4.2.1), apoi urmând viziunea cauzală a lui Hume (secțiunea 2.1) și culminând cu utilizarea algebrei boolene. Chiar și reprezentarea grafică prin diagramele Venn sunt o reflectare a filozofiei deterministe: un element este fie în interiorul mulțimii, fie în afara acesteia.

Acum mai bine de 25 ani, Charles Ragin a introdus QCA în științele sociale și politice publicând faimoasa carte *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies* (1987), un material care este încă o referință importantă în domeniul analizei calitativ-comparative.

În urma numeroaselor critici metodologice (prezentate în secțiunea 1.3), Ragin a înțeles că metoda bazată pe mulțimi distincte (în engleză *crisp sets*) este mult prea restrictivă pentru a putea fi acceptată ca atare și a extins dezvoltarea analizei spre mulțimile vagi.

Tot Ragin este cel a revenit cu o extensie a acestei analize de la mulțimile distincte la mulțimile vagi, publicând o altă carte de referință numită *Fuzzy-set social science* (2000), și ulterior rafinând instrumentele de măsurare în următorul reper teoretic *Redesigning social inquiry* (2008b), ca răspuns la un alt faimos reper teoretic pe linia cercetării cantitative, așa-numitul “KKV” (King, Keohane și Verba, 1994).

### 5.3.1 Cercetarea socială și mulțimile vagi

În această secțiune, mulțimile vagi sunt prezentate ca o extensie a analizei calitativ-comparative clasice (a mulțimilor distincte), însă teoria mulțimilor vagi are o aplicabilitate mult mai largă. Chiar dacă analiza calitativ-comparativă își găsește aplicabilitatea la nivel de macro-comunități (compararea indivizilor, a micro-unităților de analiză nu poate oferi rezultate generalizabile întrucât acestea nu sunt suficient de stabile), exemplele pe mulțimi vagi oscilează permanent între nivelul macro și cel micro deoarece este mai importantă explicarea utilizării corecte a mulțimilor vagi decât păstrarea unei linii de analiză la nivel macro. La final însă, combinarea mulțimilor vagi cu analiza calitativ-comparativă se va referi doar la unități agregate.

Există cazuri în care mulțimile nu pot fi decât distincte (de tip “da sau ba”). Putem exemplifica situația vizelor de intrare într-o țară: fie există un acord care înlătură vizele, fie un astfel de acord nu există, iar vizele sunt aplicate. Nu poate exista o situație de mijloc, o viză nu se poate aplica mai mult sau mai puțin: viza se aplică ori nu se aplică.

Alte situații însă nu sunt la fel de clare. Nu se poate spune, de exemplu, că o țară este pur și simplu democrată sau nu. Există țări care prezintă anumiți indicatori ai democrației (sistem electoral, posibilitatea de asociere, separarea puterilor în stat, garantarea proprietății private etc.), care pot prezenta însă în același timp și alți indicatori contrari (influența politicului asupra exprimării libere a opiniei, controlarea puterii judecătorești, corupția la nivel înalt etc.). În aceste cazuri, nu se mai poate vorbi despre o apartenență completă în mulțimea țărilor democratice, ci o apartenență mai mult sau mai puțin redusă.

Putem spune că avem democrație în România? Probabil că da însă oricum am defini democrația, în România se pot întâlni numeroase fenomene nedemocratice, unele dintre ele chiar anti-democratice. Pe de altă parte însă, nici nu putem spune că România este o țară cu totul nedemocrată. Luată per ansamblu, România o țară este mai mult sau mai puțin democrată.

Și totuși, *cât* de democratizată este România? Cu siguranță, undeva între aceste două limite:

- 0 (sau 0%) cu totul nedemocrată
- 1 (sau 100%) cu totul democrată

În limbajul mulțimilor vagi, România nu este pur și simplu democrată sau nu, ci are o apartenență *parțială* la mulțimea țărilor democratice (unde “parțială” poate însemna oriunde între apartenență foarte scăzută ori deloc și apartenență foarte mare ori chiar totală).

Dacă am putea să calculăm (ori să apreciem) un scor al democrației pentru România, dacă acest scor ar fi mai mare ca 0.5 atunci România ar putea fi considerată o țară mai mult democrată decât nedemocrată, și invers. Cu cât scorul se apropie mai mult de 1, cu atât putem afirma cu o mai mare certitudine că sistemul funcționează pe niște fundamente perfect democratice; cu cât scorul scade sub 0.5, putem spune că există o slabă democrație, iar atunci când scorul respectiv devine 0 putem afirma că democrația lipsește cu desăvârșire.

Între 0 și 1 este un spațiu vag, căreia orice factor îi aparține mai mult sau mai puțin. Un scor perfect egal cu 0 ori cu 1 nu este imposibil, însă este destul de rar întâlnit (de exemplu, cei mai mulți oameni au o apartenență egală cu 0 la mulțimea oamenilor cu avere de peste 1 miliard de Euro).

În cele mai multe cazuri, orice fenomen studiat aparține unei mulțimi, oricât de puțin. De exemplu, un scor de 0.13 nu înseamnă o excludere totală dintr-o mulțime, ci faptul că apartenența este extrem de slabă: este mai mult în afara mulțimii decât înăuntrul ei.

Apartenența la o clasă este o chestiune de interpretare subiectivă, în consecință orice afirmație în acest sens este după toate standardele "vagă" (fuzzy). De exemplu, o persoană cu o înălțime de 1.70 metri poate fi considerată înaltă într-o populație cu o înălțime medie scăzută (să spunem în China), dar poate fi considerată de o înălțime scăzută într-un club al oamenilor cu o înălțime de peste 2 metri. Alocarea cazurilor în clasa oamenilor înalți este așadar o operațiune care ține de subiectivitatea celui care face aprecierea și de punctul de reper în funcție de care se face această apreciere.

Orice studiu de caz poate fi alocat unei clase sau alta, iar gradul de apartenență la o clasă poate fi alocat fie printr-o procedură subiectivă fie prin intermediul unei proceduri matematice obiective. În ambele situații însă scorul de apartenență este cuprins între 0 și 1. Pentru replicabilitatea rezultatelor, este important ca modul de alocare a scorului să fie amănunțit explicat și să se bazeze pe interpretări teoretice solide.

Operațiunea poate să pară foarte complicată atunci când se utilizează scoruri matematice, însă, în realitate, teoria mulțimilor vagi a fost construită tocmai pentru a puncta diferența dintre științele exacte (matematica, fizica etc.) și cele umaniste. Dacă științele dure se bazează pe exactitatea calculelor utilizând formule specifice, științele social-umane nu vor prezenta niciodată astfel de exactități. Societatea umană este inexactă și nu se mișcă în acord cu anumite legi specifice; oamenii acționează în general mai mult sau mai puțin haotic, uneori chiar împotriva propriilor convingeri și preferințe. Toate fenomenele



sociale și umane se caracterizează printr-o cantitate mai mare sau mai mică de acțiune haotică, fapt care face imposibilă cuantificarea exactă și transpunerea acestora în legități matematice.

În schimb, oamenii utilizează într-un mod foarte confortabil expresii lingvistice care înlocuiesc, pe de o parte, cantitățile exacte, iar pe de altă parte, transmit o informație care (la modul ideal) este înțeleasă de către ceilalți. În procesul de comunicare pot fi transmise informații suficient de precise pentru a fi înțelese, fără detalii de înaltă exactitate.

De exemplu, dacă am caracteriza înălțimea unei persoane (să o numim Andrei), există două tipuri de răspuns:

- Andrei are 1.86 metri înălțime
- Andrei este *destul de înalt*

Într-o conversație obișnuită, utilizarea repetată a exactităților este ciudată: “Andrei este o persoană de 1.86 m., are cam 92.4 kg. și are un venit de 440 lei. Atenția îi scade după 3.7 minute și muncește în medie cam 8.2 ore pe zi”. Puțină lume așteaptă atâta exactitate într-o discuție, în schimb sunt foarte comune caracterizări de tipul: “Andrei este un tip destul de înalt, un pic rotofei și nu e foarte înstărit”. Aceasta cu atât mai mult cu cât anumite caracteristici nici nu pot fi cuantificate în mod direct: frumusețea, bunătatea, afecțiunea etc.

Atributele “destul de”, “un pic” și “nu foarte” sunt specifice științelor social-umane care se bazează în primul rând pe astfel de inexactități, iar teoria mulțimilor vagi utilizează aceste atribute prin transformarea lor în scoruri de apartenență la o mulțime sau alta:

- “destul de înalt” ar putea reprezenta un scor de 0.7 la clasa oamenilor înalți;
- “un pic rotofei” ar putea fi echivalent cu un scor de 0.4 la clasa oamenilor supraponderali;
- “nu foarte înstărit” ar putea reprezenta un scor de 0.2 la clasa oamenilor înstăriți.

Aceleași principii sunt ușor de aplicat pentru orice atribut: un om poate fi “frumușel”, “bunătatea întruchipată” și “foarte afectuos”. Uneori limbajul poate fi mult mai precis în comunicare decât orice cuantificare numerică. Caracterizarea bunăstării unei persoane poate fi făcută cu ajutorul științelor exacte: “Andrei are un venit de 440 lei, în condițiile

în care salariul mediu pe economie este de 890 lei”, sau cu ajutorul atributelor vagi din limbajul comun: “Andrei este destul de sărac”.

Deși se folosesc atribute vagi, caracterizarea poate fi extrem de precisă; o afirmație de tipul: “Andrei e sărac lipit pământului, trăiește ca vai de el” crează în mintea oricui o imagine foarte clară a unei sărăcii extreme, cu lipsuri puternice în toate dimensiunile existenței sociale.

Urmând procesul de transformare a atributelor calitative în cantități numerice, se poate stabili nivelul de apartenență la diverse clase; Ragin (2000, p.160) arată că logica fuzzy se potrivește perfect cu studiul științelor sociale, fiind o dezvoltare a matematicii care combină logica formală (extrem de precisă) cu formulările verbale (de obicei extrem de imprecise). Mulțimile vagi reușesc să combine așadar atribute calitative și cantitative într-un singur instrument: apartenența completă la o clasă ori faptul de a fi complet în afara unei clase sunt atribute calitative (“Andrei este frumos” sau “Andrei nu e frumos”), iar variația dintre cele două extreme (specifică variabilelor continue) reprezintă un atribut cantitativ.

Chiar în debutul articolului original din 1965, Zadeh punctează faptul că limbajul operează cu structuri abstracte: “Mai des decât ne dăm seama, clasele de obiecte întâlnite în lumea reală nu au criterii de apartenență precis definite”. Este exact ceea ce Fuhrmann (1998) înțelege prin neclaritatea conceptelor, ori altfel spus conceptul de “vag” a apărut tocmai datorită faptului că cele mai multe dintre conceptele utilizate în științele sociale sunt vagi.

Acest lucru însă nu împiedică deloc metodologia cercetării sociale, ci dimpotrivă reprezintă unul dintre punctele forte ale acesteia. În secțiunile următoare vor fi prezentate modalitățile concrete cu care științele sociale pot opera cu mulțimile vagi pentru a deriva concluzii bazate pe evidențe empirice.

### 5.3.2 Operațiuni cu mulțimi vagi

Întocmai ca la operațiunile din algebra booleană, mulțimile vagi prezintă un set specific de operațiuni utile pentru analiza datelor de cercetare. Dacă în mulțimile distincte sunt posibile doar trei operațiuni principale (negarea, intersecția și reuniunea, restul fiind combinații ale acestora), mulțimile vagi permit mai multe, printre care și acestea trei. Fiecare operațiune are particularități în funcție de tipul de mulțime cu care se lucrează.

Operațiunile pe mulțimi vagi permit atribuirea unor scoruri de apartenență a unui caz la diferite mulțimi, însă scopul acestora nu se oprește aici. Să considerăm apartenența etnică, un subiect foarte discutat în sociologia modernă, în special cea din America. Numărul mare de rase (caucazieni, afro-americani, indieni americani, hispanici, asiatici) au creat un mix rasial greu de descris. Aproape că nu mai există trăsături “pure”, oameni care să facă parte doar dintr-o singură rasă. În schimb, sunt foarte mulți oameni a căror rasă aproape că nu poate fi identificată (ori cel puțin nu au fost recunoscute ca etnii în mod oficial): ce rasă are un copil provenit dintr-un tată afro-american și o mamă pe jumătate asiatică și jumătate hispanică?

Așadar, atribuirea scorurilor de apartenență și utilizarea acestor scoruri în operațiuni complexe sunt necesare în construirea și testarea unei teorii (calitative sau comparative) asupra unui fenomen. În continuare sunt prezentate operațiunile posibile pe mulțimi vagi, iar primele operațiuni sunt comparate cu cele corespondente pe mulțimile distincte.

**Negația.** Atât în mulțimile distincte cât și în mulțimile vagi, negația este complementul unei mulțimi: tot ce nu este în mulțimea respectivă. Notând o mulțime oarecare cu litera  $A$ , complementul acesteia are mai multe notații: Zadeh îl notează cu  $A'$ , în alte lucrări se întâlnește semnul  $\neg A$ , iar Ragin îl notează cu  $\sim A$ .

În mulțimile distincte, negația schimbă valoarea 1 cu valoarea 0 și invers: Franța percepe taxe de autostradă (scor distinct 1) în timp ce Germania nu percepe astfel de taxe (scor distinct 0):

$$\sim 1 = 0$$

$$\sim 0 = 1$$

În mulțimile vagi, complementul lui  $A$  (cu alte cuvinte *notA*) este:

$$\sim A = 1 - A$$

Presupunând că România aparține de mulțimea țărilor democratice ( $A$ ) cu un scor de 0.7 atunci aparține în același timp la mulțimea țărilor nedemocratice ( $\sim A$ ) cu un scor de  $1 - 0.7 = 0.3$ .

Negația este o atitudine des întâlnită în cadrul cercetărilor sociale. Similar cu exemplul apartenenței rasiale, poate fi utilizat un exemplu al apartenenței etnice. În România există multe încercări de a estima numărul de locuitori de etnie roma, iar statisticile oficiale publicate de INSSE sunt foarte diferite de cifrele publicate de organizațiile de etnie roma. Zamfir și Zamfir (1993) au cercetat acest aspect și au publicat o altă metodologie. Singurul lucru cert este faptul că o persoană nu este pur și simplu roma decât dacă simte, acceptă și se autodeclară. Teoriile genetice au fost

abandonate odată cu al doilea Război Mondial. O persoană poate fi născută din părinți de etnie diferită, ceea ce nu înseamnă că este pe jumătate din fiecare etnie. Contează ceea ce simte o persoană, iar fiecare persoană simte mai puternic sau mai slab. Acest tip de autopercepție nu poate fi măsurat decât prin intermediul mulțimilor vagi.

Recensământul populației numără doar persoanele autodeclarate. Stigmatizarea populației roma a declanșat o reticență ridicată în a declara această etnie, iar în absența unui scor cert pentru o etnie sau alta, oamenii preferă un scor ridicat pe *negarea* etniei.

Un caz special îl constituie negația în cadrul studiilor de gen. Astfel, în abordarea clasică, toți oamenii care nu sunt bărbați sunt femei și invers. Sigur că în societatea modernă se poate pune în discuție o anumită ambiguitate datorată autopercepției, însă există persoane a căror sex chiar nu poate fi determinat (hermafrodiții). În aceste condiții, negarea nu duce pur și simplu către ‘celălalt’ sex; faptul că o persoană nu este femeie nu înseamnă neapărat că este bărbat. În cadrul mulțimilor vagi, un scor foarte scăzut de 0.2 pentru apartenența la mulțimea femeilor nu înseamnă neapărat un scor de 0.8 pentru apartenența la mulțimea bărbaților, ci pur și simplu o apartenență de 0.8 la mulțimea oamenilor care nu sunt femei.

**Intersecția.** De multe ori, în cadrul unei cercetări sunt căutate cazuri care îndeplinesc *în mod simultan* mai multe caracteristici. Dacă ne-ar interesa studierea criminalității urbane, atunci vom studia doar localitățile urbane care prezintă un nivel înalt de criminalitate. În mulțimile distincte, acest lucru este foarte simplu: codând în baza de date cu 1 toate localitățile urbane (și 0 localitățile rurale), apoi cu 1 localitățile unde se întâlnește criminalitate înaltă (și 0 pentru toate celelalte). Vom selecta pentru analiză doar acele cazuri care prezintă cifra 1 la ambele caracteristici.

O altă metodă posibilă este crostabularea (formarea unui tabel  $2 \times 2$  în urma căreia se obțin 4 celule) selectându-se doar localitățile care se află în celula corespondentă cifrei 1 la ambele caracteristici: este prezentă ȘI o caracteristică, ȘI cealaltă. Intersecția dintre două mulțimi este echivalentă cu înmulțirea acestora, iar operațiunea se mai numește “ȘI logic” și este indicată de semnul “.”:

$$0 \cdot 0 = 0$$

$$0 \cdot 1 = 0$$

$$1 \cdot 0 = 0$$

$$1 \cdot 1 = 1$$

Intersecția dintre două mulțimi distincte A și B este o altă mulțime C, mai mică decât fiecare dintre cele două; este o sub-mulțime formată doar într-o singură celulă din patru posibile. În mulțimile vagi, nu se poate obține un tabel întrucât există o infinitate de valori pentru fiecare variabilă în parte. Presupunând că o localitate poate fi mai mult sau mai puțin urbană, apartenența acesteia la intersecția dintre urban și criminalitate se calculează prin minimul scorurilor de la cele două caracteristici:

$$C = A \cap B = \min(A, B)$$

Dacă o localitate are un scor al urbanizării de 0.8 și un scor al criminalității de 0.3, atunci aparține de intersecția caracteristicilor cu cel mai mic dintre cele două scoruri:

$$A \cap B = \min(0.8, 0.3) = 0.3$$

Scorul înalt pe caracteristica urbanizare nu poate compensa scorul scăzut de criminalitate; cu alte cuvinte, nu are nici o importanță că localitatea este foarte dezvoltată, întrucât criminalitatea locală este destul de scăzută: se constată o apartenență scăzută în cadrul mulțimii orașelor cu criminalitate înaltă.

Zadeh (1965, p.341) denumește intersecția “cea mai mare mulțime care este conținută atât în A cât și în B”.

**Reuniunea.** Alte situații de cercetare se pot concentra pe studierea unor cazuri care prezintă *cel puțin* o caracteristică. Să presupunem că dorim să cercetăm dacă este adevărată expresia “Fericii cei săraci cu duhul”. Există două mulțimi de interes: cea a persoanelor fericite și cea a persoanelor sărace cu duhul (ori am putea să considerăm mulțimea complementară a persoanelor care nu sunt inteligente). Unii oameni pot fi fericiți chiar dacă sunt inteligenți, alții pot fi săraci cu duhul dar nu neapărat fericiți, în timp ce alții pot prezenta ambele caracteristici. Trebuie studiate ambele mulțimi, chiar dacă unele persoane fac parte din ambele mulțimi simultan. Doar cei care nu fac parte din nici una dintre cele două mulțimi nu intră în universul cercetării.

Ca și în cazul intersecției, reuniunea dintre două mulțimi distincte A și B este o altă mulțime C, mai mare decât fiecare dintre primele două; este o supra-mulțime formată din trei celule ale tabelului  $2 \times 2$ . Singurele persoane care nu fac parte din universul cercetării sunt cele care nu sunt nici fericite, nici sărace cu duhul. Operațiunea de reuniune este opusă celei de intersecție.

În cazul mulțimilor vagi, operațiunea opusă calculării minimului dintre două scoruri este calcularea maximului dintre acestea:

$$C = A \cup B = \max(A, B)$$

Dacă un individ are un scor de 0.7 pentru fericire și un scor de 0.2 pentru mulțimea celor săraci cu duhul, va avea un scor de 0.7 la mulțimea oamenilor fericiți SAU săraci cu duhul (maximul dintre cele două scoruri):

$$A \cup B = \max(0.7, 0.2) = 0,7$$

Nu contează că are un scor scăzut la una dintre mulțimi, întrucât are un scor ridicat la cealaltă; dacă are un scor ridicat la cel puțin una dintre mulțimi, atunci va avea un scor la fel de ridicat la reuniunea dintre cele două mulțimi. Același principiu se aplică și pentru mai mult de două mulțimi: scorul pentru reuniunea dintre trei mulțimi se calculează prin maximul dintre scorurile de apartenență la fiecare dintre cele trei mulțimi.

Zadeh denumește reuniunea “Cea mai mică mulțime care le conține atât pe A cât și pe B”.

**Concentrarea.** Dacă primele trei operațiuni pot fi efectuate și pe mulțimi distincte, următoarele sunt specifice mulțimilor vagi. Preocupat de ambiguitatea limbajului, Zadeh a dorit să surprindă într-o formulă contrastul dintre o caracteristică și superlativul acesteia. O persoană poate avea un scor ridicat în mulțimea oamenilor inteligenți, însă un scor sensibil mai scăzut în mulțimea oamenilor *foarte* inteligenți.

Rezultatul concentrării scorului de apartenență la o mulțime se obține prin ridicarea la pătrat a scorului respectiv. Acest artificiu se bazează pe observația pertinentă a faptului că un număr subunitar ridicat la pătrat generează tot un număr subunitar însă mai mic. Ridicarea la pătrat (concentrarea scorului) este echivalentul matematic al aplicării atributului “foarte” din cadrul limbajului.

De exemplu, dacă o persoană are un scor de 0.8 în cadrul mulțimii oamenilor inteligenți, va avea un scor de  $0.8^2 = 0.64$  în cadrul mulțimii oamenilor foarte inteligenți.

**Dilatarea.** În alte situații dorim să relaxăm oarecum criteriile de atribuire a apartenenței la o mulțime. Uneori, aceste criterii pot fi atât de restrictive încât nici unul dintre cazurile studiate ori cunoscute nu poate să atingă scoruri înalte de apartenență. De exemplu, puține persoane se pot lăuda cu un comportament perfect moral, dacă prin moralitate înțelegem respectarea unor standarde foarte înalte de conduită și comportament. Totuși, foarte multe persoane au un comportament aproape de moralitate, chiar dacă uneori sunt departe de standardele sociale acceptate.

De aceea, este nevoie de relaxarea acestor criterii: în loc să atribuim scoruri de apartenență la mulțimea oamenilor (perfect) morali, am putea să atribuim scoruri de apartenență la mulțimea oamenilor *aproape* morali. Ragin utilizează expresia “mai mult sau mai puțin” (morali), însă în opinia mea se induce o confuzie cu teoria generală a mulțimilor vagi, unde “mai mult sau mai puțin” înseamnă apartenență parțială la o mulțime.

Cuvântul “aproape” relaxează restricțiile într-un mod complementar, opus modului în care cuvântul “foarte” întărește restricțiile de apartenență la o mulțime. De aceea, rezultatul dilatării unui scor se obține prin extragerea radicalului din scorul respectiv (operațiunea inversă ridicării la pătrat): dacă o persoană are un scor de 0.4 în cadrul mulțimii oamenilor inteligenți, va avea un scor de  $\sqrt{0.4} = 0.63$  în cadrul mulțimii oamenilor aproape inteligenți.

Teoria mulțimilor vagi se află într-o continuă dezvoltare, nu doar în științele social-umane, ci în toate ramurile științei. De exemplu, ideile avansate de Zadeh au prins rădăcini puternice în comunitatea științifică japoneză, care le utilizează pentru dezvoltarea unor algoritmi de inteligență artificială. Aparatele electro-casnice au devenit acum inteligente: unele frigider monitorizează stocul unui produs și au funcții de comandă prin Internet pentru o nouă aprovizionare atunci când stocul este pe final. Interpretarea comenzilor vocale de către calculator sunt deja o realitate, însă aplicarea mulțimilor vagi permit computerelor să interpreteze corect *nuanțe* ale limbajului uman.

### 5.3.3 Calibrarea mulțimilor vagi

În general, obținerea scorurilor de apartenență la mulțimile vagi (de la 0 la 1) pornește de la o variabilă numerică inițială (datele brute), apoi se aplică niște simple transformări pentru obținerea scorurilor fuzzy.

Operațiunea de calibrare este fundamental diferită de măsurarea utilizată în analiza cantitativă. Ragin (2008a,b) arată foarte clar diferența dintre cele două, utilizând niște exemple simple de înțeles. Pornește de la valorile venitului per capita pentru 24 de țări, în ordine descrescătoare, începând cu Elveția (cu un venit per capita de 40110\$) până la Burundi (cu un venit per capita de 110\$).

La o transformare matematică naivă, luând cel mai mare scor drept reper pentru includerea în “mulțimea țărilor dezvoltate”, Elveția ar avea un scor fuzzy de  $40110/40110 = 1$  (o apartenență perfectă la această

mulțime), iar Burundi ar avea un scor fuzzy de  $110/40110 = 0.003$  (practic o apartenență egală cu zero, sau o excludere totală din această mulțime).

A doua țară ca mărime a venitului per capita ar fi Statele Unite, cu un scor de apartenență de  $34400/40110 = 0.858$  adică o apartenență foarte mare (dar nu completă). Utilizând acest procedeu artificial, se poate observa că vom avea întotdeauna doar o singură țară cu o apartenență totală la această mulțime, restul având scoruri de apartenență proporțional mai mici.

Această transformare mecanică este însă incorectă, din mai multe motive. Pe de o parte, pentru că sunt multe țări (nu doar una) pe care le-am putea considera ca fiind “dezvoltate”: după toate criteriile, Statele Unite ar trebui să aibă o apartenență totală la această mulțime, nu doar un scor de 0.858. Pe de altă parte, foarte important, efectul produs de o unitate de măsură al venitului per capita pentru dezvoltarea socială este complet diferit la țările cu un venit înalt, față de țările cu un venit scăzut. Efectul nu este nici pe de parte algebric, matematic, în sensul în care diferența dintre 40000 și 39000 este de 1000, aceeași diferență fiind și între 2000 și 1000. Deși diferența matematică este aceeași, nu se schimbă aproape nimic în traiul de zi cu zi al cetățenilor, dacă venitul per capita crește de la 39000 la 40000, însă o creștere de la 1000 la 2000 reprezintă o dezvoltare socială explozivă, întrucât venitul per capita se dublează ca mărime.

În măsurarea cantitativă a conceptelor, întreaga cantitate de variație este considerată la fel de importantă: și cea care crește venitul per capita de la 39000 la 40000, și cea care crește de la 1000 la 2000. Pentru instrumentele de lucru cantitative, ambele creșteri sunt la fel de mari însă în lumea reală cele două produc efecte complet diferite, în zone geografice diferite. Pentru o țară săracă din Africa, o creștere de 500\$ produce efecte sociale mult mai importante decât o creștere de 10000\$ în Elveția.

Un alt motiv important este legat de validitatea utilizării matematice a venitului per capita, ca indicator al dezvoltării sociale. Țările exportatoare de petrol afișează un venit per capita extrem de înalt, ceea ce ar putea să conducă la impresia că și nivelul de dezvoltare al țărilor respective este la fel de înalt, însă realitatea este că nivelul de dezvoltare este disproporționat de mic în comparație cu mărimea venitului per capita. Realitatea “nu pare” să concorde cu matematica proporțională, ceea ce conduce la un paradox al lipsei de concordanță între mărimea indicatorului de dezvoltare (venitul per capita) și nivelul real de dezvoltare al unei țări. Din acest motiv, în cercetarea cantitativă se apelează la indici compuși din mai multe componente, pe baza operaționalizării conceptului și despărțirea acestuia în dimensiuni și (de



obicei) mai mulți indicatori. Indicii compuși (agregați după o metodă sau alta) repară întrucâtva problema validității, însă au aceeași senzitivitate la tratamentul egal pentru întreaga cantitate de variație rezultată.

Toate aceste argumente conduc la ideea că ar trebui să existe niște “puncte de reper” în funcție de care să putem aprecia dacă o țară aparține sau nu la mulțimea țărilor dezvoltate. De exemplu, am putea considera că țările dezvoltate au un venit per capita de minim 20000\$, astfel încât toate țările care au o valoare cel puțin egală (sau mai mare) decât aceasta vor fi complet incluse în mulțimea țărilor dezvoltate.

Valoarea de 20000\$ nu a rezultat din analiza variației vreunui eșantion, nici din analiza cantitativă a tuturor țărilor cunoscute. Este un punct de reper complet subiectiv, obținut din experiența cercetătorului sau prin consensul comunității academice, în legătură cu care se consideră că o țară este în mod sigur dezvoltată.

Este un proces de standardizare a unor valori, similar celui care a stabilit valoarea de  $0^{\circ}$  Celsius când apa îngheață, sau valoarea de  $100^{\circ}$  Celsius când apa fierbe. Acestea nu sunt niște valori fundamentale, matematic universale, ci doar niște puncte de reper, niște standarde aprobate în cadrul comunității academice din fizică: valorile la care apa își schimbă starea de agregare.

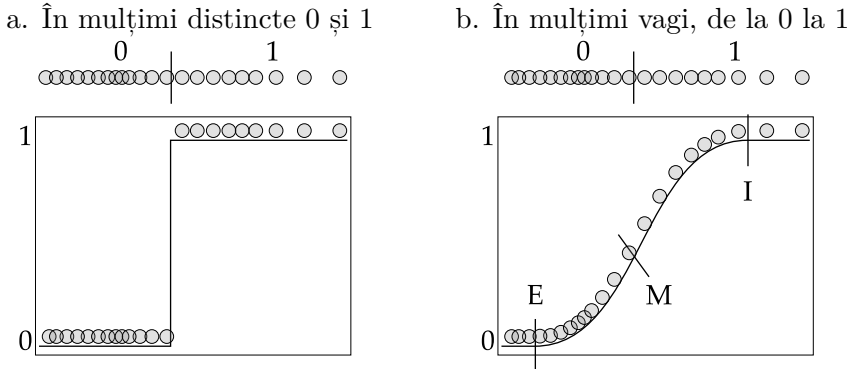
Ragin le numește “ancore calitative”, iar pentru mulțimile vagi acestea definesc trei praguri: cel de includere completă, cel de excludere completă și punctul de mijloc. Acest proces de calibrare reprezintă o punte de legătură între cercetarea calitativă și cercetarea cantitativă, întrucât:

- stabilirea ancorelor de calibrare este un proces în esență calitativ;
- derivarea scorurilor de includere în mulțime este un proces cantitativ.

Figura 5.13 arată două procese diferite de calibrare. Prima diagramă, cea din stânga, împarte valorile condiției cauzale în două grupuri, prin stabilirea unui prag cu ajutorul căruia valorile mai mici decât pragul vor fi recodificate cu valoarea 0, iar cele mai mari decât pragul vor fi recodificate cu valoarea 1.

Se poate observa ușor faptul că pragul respectiv nu diferențiază foarte bine între valorile din imediata vecinătate. Dacă ar fi mutat puțin mai la stânga, valoarea imediată ar fi recodificată cu 1 în loc de 0. Praguri aproape egale ca valoare vor determina recodificări diferite ale unor valori, ceea ce înseamnă că valorile nu se grupează în mod natural în două grupuri distincte.

Figura 5.13: Calibrarea cu mulțimea “țări dezvoltate”



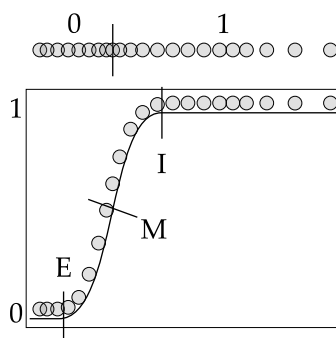
Aceleași valori, în diagrama b. din partea dreaptă, sunt recodificate cu ajutorul unei funcții logistice unde au fost specificate trei praguri:

- cel din stânga determină punctul de excludere completă (E), iar valorile mai mici decât pragul vor fi automat recodificate cu valoarea 0;
- cel din mijloc (M), valorile mai mari decât acesta fiind mai mult în interiorul mulțimii decât în afara ei, iar valorile mai mici fiind mai mult în afara mulțimii decât în interiorul acesteia;
- cel din dreapta determină punctul de includere completă în mulțime (I), iar valorile mai mari decât pragul vor fi automat recodificate cu valoarea 1.

Toate valorile dintre pragurile E și I sunt recodificate cu valori continue între 0 și 1. Acestea sunt aranjate în ordine crescătoare, de la aproape (dar puțin mai mari) ca 0 imediat în dreapta pragului E, până la aproape (dar puțin mai mici) ca 1 în stânga pragului I. În loc de o funcție brusc crescătoare ca în diagrama a. (care desparte valorile în două grupuri distincte), funcția din diagrama b. este ușor crescătoare, specifică pentru o funcție logistică.

Orice condiție cauzală numerică (de exemplu venitul per capita) poate fi transformată în scoruri de mulțimi vagi, obținându-se valori între 0 și 1. Indiferent de unitatea de măsură a condiției numerice inițiale (ani, kilograme, bani etc.) pot fi obținute scoruri vagi între 0 și 1, unde 0 înseamnă că valorile inițiale nu intră deloc în mulțimea respectivă, iar 1 înseamnă că sunt complet incluse în mulțime. Orice valori între 0 și 1 arată “cât de mult” sau “cât de puțin” sunt incluse valorile respective în mulțime.

Figura 5.14: Calibrarea cu mulțimea “țări cel puțin mediu dezvoltate”



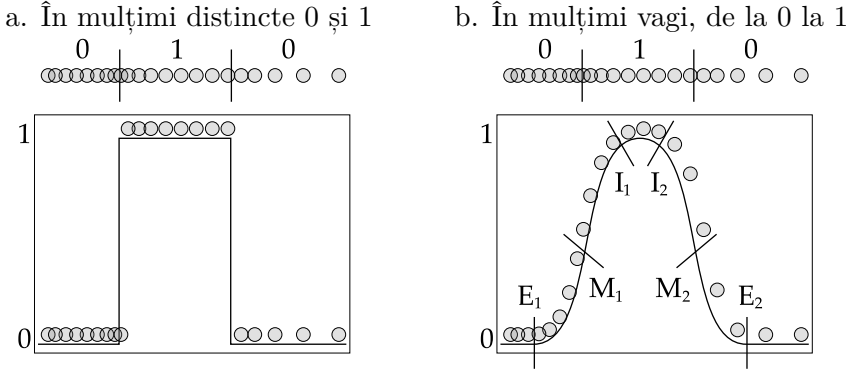
Valorile pragurilor E, M și I sunt stabilite în mod calitativ de către cercetător, nu există nicio formulă matematică capabilă să stabilească aceste valori în mod automat. Acest lucru se datorează definiției cu care lucrează cercetătorul atunci când descrie mulțimea în care vor fi calibrate valorile inițiale. De exemplu, dacă mulțimea se numește “țări dezvoltate” vom avea un set de praguri, însă dacă mulțimea se numește “țări cel puțin mediu dezvoltate” atunci toate pragurile trebuie mutate mai la stânga, în funcție de definiția cercetătorului cu privire la ceea ce este comun acceptat ca fiind o țară “mediu dezvoltată”. Figura 5.14 arată un posibil astfel de scenariu.

În genul acesta de calibrări se lucrează cu cuvinte (“dezvoltată”, sau “mediu dezvoltată” etc.) și trebuie subliniat faptul că niciun calculator din lume și nicio formulă matematică magică nu va putea înlocui judecata cercetătorului cu privire la definirea și interpretarea acestor cuvinte. De aceea, stabilirea pragurilor (procesul de calibrare în sine) este o etapă în esență calitativă, iar scorurile obținute sunt în esență cantitative.

De aceea, procesul de calibrare face o demonstrație excelentă de fuziune și complementaritate între cercetarea calitativă și cea cantitativă. Niciuna dintre acestea nu ar fi reușit să obțină scorurile finale fără cealaltă. Pe de o parte, stabilirea calitativă a unor praguri fără cunoștințele de funcție logistică, specific cantitative, nu putea duce nicăieri, dar nici funcția logistică, matematică, nu putea deriva în mod corect pragurile respective fără interpretarea calitativă a mulțimilor cu care se lucrează în definițiile teoretice.

Alegerea de a numi mulțimea țărilor “cel puțin mediu dezvoltate” poate să pară ciudată însă nu este întâmplătoare. Este singura formulare verbală

Figura 5.15: Calibrarea cu mulțimea “țări mediu dezvoltate”



care permite o aranjare crescătoare a valorilor dintre pragurile  $M$  și  $I$ , iar toate valorile din dreapta pragului  $I$  să fie recodificate cu 1 (cu alte cuvinte și țările “dezvoltate” fac parte din mulțimea țărilor “cel puțin mediu dezvoltate”).

O cu totul altă poveste o spune formularea mulțimii țărilor “mediu dezvoltate”. În această situație, țările foarte sărace (cele mai din stânga) nu fac parte din mulțime, la fel ca și în cazul țărilor dezvoltate. Pe de altă parte însă, o țară ca Elveția sau una ca Statele Unite sunt țări în mod evident dezvoltate, deci nici acestea nu fac parte din mulțimea țărilor “mediu” dezvoltate.

La fel ca în cazul precedent, funcția de apartenență la mulțimea țărilor “mediu” dezvoltate are o pantă ascendentă în partea din stânga, însă pe măsură ce venitul per capita crește tot mai mult (pe măsură ce se apropie de mulțimea țărilor dezvoltate) funcția de apartenență scade pe măsură.

Diagrama 5.15.a arată calibrarea în mulțimi distincte a țărilor mediu dezvoltate. Este nevoie de două praguri pentru a delimita zona “medie” de celelalte țări. În această situație, țările sărace vor fi codificate cu 0, iar țările bogate (dezvoltate) vor fi codificate și acestea cu 0 întrucât nu sunt țări mediu dezvoltate. Doar țările din mijloc pot fi codificate cu 1.

Diagrama 5.15.b arată un lucru similar pe mulțimi vagi, cu cele două pante ascendentă și descendentă. La calibrarea în mulțimi vagi a țărilor dezvoltate, pentru un singur prag pe mulțimi distincte au fost necesare trei praguri pe mulțimi vagi, la o singură pantă ascendentă. În mod similar, pentru fiecare dintre cele două pante la mulțimea țărilor mediu dezvoltate sunt necesare câte trei praguri, rezultând un total de șase:

- două de excludere completă ( $E_1$  și  $E_2$ ): toate cazurile din stânga lui  $E_1$  și din dreapta lui  $E_2$  sunt automat codificate cu 0;

- două de mijloc ( $M_1$  și  $M_2$ ): toate valorile din stânga lui  $M_1$  și din dreapta lui  $M_2$  sunt mai mult în afara mulțimii, iar cele dintre  $M_1$  și  $M_2$  sunt mai mult în interiorul mulțimii decât în afara ei;
- două de includere completă în mulțime ( $I_1$  și  $I_2$ ): toate cazurile dintre acestea sunt automat codificate cu 1.

Funcțiile de apartenență pot avea foarte multe forme: în formă de S, în formă de S inversat, pătratică, trapeziodale (cu baza în jos, cu baza în sus) etc. Alegerea formei corecte pentru calibrarea mulțimilor este direct dependentă de cunoașterea domeniului studiat, precum și de interpretarea calitativă a denumirii mulțimii. Pentru mulțimile vagi, este nevoie de un minim de 3 praguri, iar pentru funcțiile care prezintă pante atât ascendente cât și descendente este nevoie de specificarea a 6 praguri. Compararea mărimii acestora, în ordinea cronologică a specificării lor, determină ascendența sau descendența pantelor asociate.

Calibrarea mulțimilor are și alte forme. Ceea ce a fost prezentat în această secțiune reprezintă așa-numita “atribuire directă”: odată stabilite pragurile, toate valorile care sunt în afara lor vor fi recodificate fie cu 0 dacă sunt mai mici decât pragul de excludere, fie cu 1 dacă sunt mai mari decât pragul de includere completă.

Ragin (2008b) demonstrează și o altă metodă, care utilizează aproximarea pe o funcție logistică, cu o etapă intermediară care utilizează estimarea logaritmului șanselor de succes pentru apartenența deplină la o mulțime. Exprimarea pare puțin complicată, însă este perfect naturală pentru utilizatorii care au cunoștințe de bază în regresia logistică.

Marea majoritate a calibrărilor de mulțimi vagi sunt efectuate, în prezent, utilizând funcția logistică propusă de Ragin (există desigur programe software specializate care calculează valorile calibrate). Această secțiune are rolul de a clarifica necesitatea “calibrării” mulțimilor, înțelegerea deplină a diferitelor metode de calibrare și a diferențelor dintre acestea necesită studiul aprofundat al referințelor de specialitate.

### 5.3.4 Analiza de necesitate

Înainte de a analiza datele cu ajutorul tabelii de adevăr, unde rezultatele obținute se referă la suficiență, primul pas în analiza calitativ-comparativă este analizarea relației de necesitate între fiecare condiție cauzală (luată individual) și fenomenul studiat.

Ca și în cazul mulțimilor distincte, relația de necesitate este una în care o condiție cauzală este o supramulțime a fenomenului studiat. În mulțimile vagi, aceasta înseamnă că toate scorurile de apartenență ale unei condiții cauzale  $X$  trebuie să fie mai mari sau egale cu scorurile de apartenență la fenomenul studiat  $Y$ :

$$X \Leftarrow Y$$

Pentru ca  $X$  să fie o condiție necesară pentru  $Y$ , scorurile de apartenență în mulțimea  $X$  trebuie să fie în mod consistent mai mari sau egale cu scorurile de apartenență în mulțimea  $Y$ :

$$X_i \geq Y_i$$

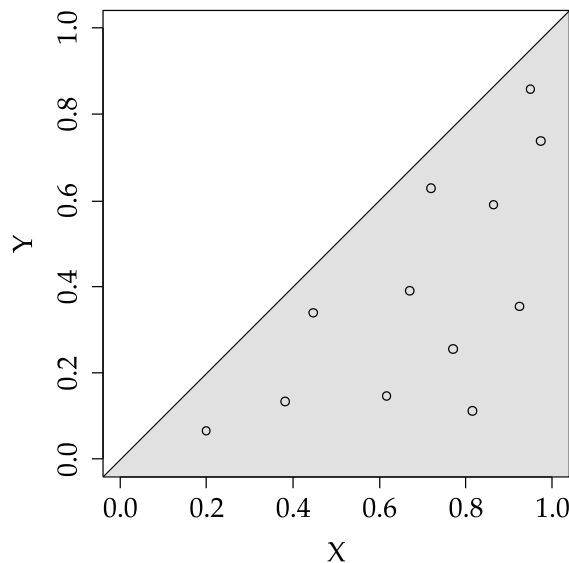
Spre deosebire de mulțimile distincte, unde relația de necesitate se poate observa ușor cu ajutorul unui tabel încrucișat, în cazul mulțimilor vagi această ultimă inegalitate se poate observa cu ajutorul unei diagrame de împrăștiere. La mulțimile distincte există un număr finit de celule ale tabelului încrucișat, iar oricare dintre cazuri aparține uneia dintre aceste celule, însă la mulțimile vagi oricare combinația dintre două scoruri fuzzy este fi reprezentată cu ajutorul unor coordonate într-un sistem cartezian, iar punctul respectiv poate să fie poziționat oriunde în cadrul acestui plan  $XoY$ .

Atunci când scorurile pe condiția cauzală  $X$  sunt în mod sistematic mai mari decât scorurile pe fenomenul studiat  $Y$ , punctele reprezentate de aceste coordonate ar trebui să se regăsească în triunghiul din partea din dreapta jos (sub diagonala principală) a sistemului cartezian  $XoY$ , așa cum este prezentat în figura 5.16.

În această figură, se poate observa că toate punctele se află sub diagonala principală, ceea ce înseamnă că toate coordonatele pe axa orizontală  $oX$  sunt mai mari decât coordonatele pe axa verticală  $oY$ .

Exemplul oferit de Ragin (2000, p.215) este perfect asupra relației de necesitate: jucătorii de basket care se antrenează mult vor înscrie și multe coșuri. Din această afirmație rezultă că antrenamentul este o condiție necesară (dar nu și suficientă) pentru a înscrie coșuri într-un meci de baschet. Vor fi desigur multe cazuri (în fapt cele mai multe cazuri) în care jucătorii care se antrenează mult nu vor înscrie foarte multe coșuri, ceea ce înseamnă că scorurile pe axa  $oX$  vor fi în mod sistematic mai mari decât scorurile pe axa  $oY$ .

În termeni de mulțimi, spunem că mulțimea jucătorilor care înscriu multe coșuri ( $Y$ ) este o submulțime a jucătorilor care se antrenează mult ( $X$ ).

Figura 5.16: X este o condiție necesară:  $X \Leftarrow Y$ 

Sau invers, toți jucătorii care se antrenează înscriu de obicei și coșuri, însă nu la fel de multe coșuri pe cât de mult se antrenează (altfel spus, mulțimea jucătorilor care se antrenează este o supramulțime a jucătorilor care înscriu coșuri).

Orice caz poziționat deasupra diagonalei duce la întreruperea imediată a relației de necesitate completă a lui X pentru Y. De exemplu, pot fi și jucători care se antrenează puțin și totuși înscriu destul de mult (datorită unui talent înnăscut), ceea ce înseamnă că antrenamentul nu este întotdeauna o condiție necesară pentru a înscrie.

Totuși, într-o situație extremă în care toate cazurile se află sub diagonală mai puțin unul, nu se poate anula complet afirmația de necesitate, ci s-ar putea afirma că X este în foarte mare măsură o condiție necesară pentru producerea fenomenului Y.

Această sintagmă “în foarte mare măsură” poate fi transformată, în fapt, într-o măsură numerică a cărei mărime ne va spune cu exactitate în “câtă” măsură este X necesară pentru Y:

$$\frac{\sum \min(X_i, Y_i)}{\sum Y_i} \in [0, 1]$$

În operațiile cu mulțimi vagi, minimumul dintre două valori arată intersecția dintre acestea, ceea ce înseamnă că formula se poate traduce în acest fel:

cât de mare este intersecția dintre X și Y, raportată la întreaga mărime a lui Y (sau cât de mult acoperă această intersecție din mulțimea Y). Ținând cont că mulțimea X este mai mare decât mulțimea Y (întrucât X este necesară), această formulă arată cât de mult este inclusă mulțimea Y în mulțimea X: cu cât mai inclusă este Y, cu atât mai necesară este se dovedește a fi X pentru Y.

Dacă mulțimea Y este complet inclusă în mulțimea X, înseamnă că intersecția este la fel de mare ca Y (altfel spus Y este complet intersectată cu X) iar formula va genera un scor egal cu 1, fapt care demonstrează, la fel ca în cazul mulțimilor distincte, caracterul de necesitate a lui X pentru Y.

În cazul în care Y nu este complet inclusă în mulțimea X, intersecția dintre X și Y va fi mai mică decât totalul mulțimii Y iar formula va arăta cât de consistentă este relația de necesitate dintre X și Y: cu cât scorul va fi mai apropiat de 1, cu atât mai consistentă este relația de necesitate.

În diagrama XoY, orice puncte deasupra diagonalei principale au coordonate în care scorul pe verticală (pe mulțimea Y) este mai mare decât scorul pe orizontală (pe mulțimea X), iar minimumul dintre acestea scade din maximumul posibil pe mulțimea Y. Orice astfel de puncte scad așadar din caracterul de necesitate a lui X pentru Y.

### 5.3.5 Analiza de suficiență

Ca și în cazul mulțimilor distincte, conceptele de necesitate și suficiență sunt complementare și explicabile prin intermediul conceptelor de includere și acoperire. Dacă în cazul relației de necesitate, mulțimea Y este inclusă în mulțimea X (este o submulțime a acesteia), în cazul relației de suficiență, mulțimea X este inclusă în mulțimea fenomenului studiat Y.

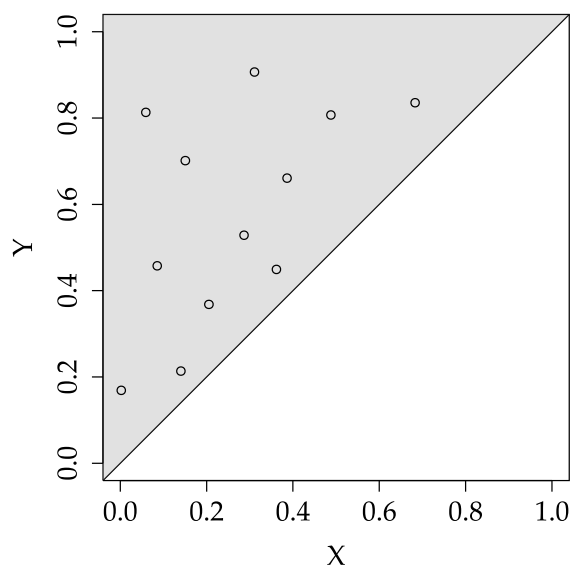
În cazul mulțimilor vagi, o condiție cauzală X este inclusă într-o mulțime Y (este o submulțime a acesteia) dacă toate scorurile de apartenență la mulțimea X sunt mai mici decât scorurile de apartenență la mulțimea Y:

$$X \Rightarrow Y \quad (X_i \leq Y_i)$$

În această situație, se poate spune că X este o condiție suficientă pentru producerea fenomenului Y. Într-o diagramă XoY, așa cum este prezentată în figura 5.17, toate punctele se află poziționate *deasupra* diagonalei principale, exact invers ca în cazul relației de necesitate.

Pentru orice punct aflat deasupra diagonalei principale, coordonata pe axa orizontală oX este un număr mai mic decât cel aferent coordonatei



Figura 5.17: X este o condiție suficientă pentru Y, adică  $X \Rightarrow Y$ 

pe axa verticală  $oY$ , fapt care demonstrează relația de includere a mulțimii  $X$  în mulțimea  $Y$ .

Orice punct aflat sub această diagonală scade din consistența relației de suficiență a lui  $X$  pentru  $Y$ . Pentru o proporție mică de astfel de cazuri, scorul de consistență a relației de suficiență va fi cu puțin mai mică decât 1, însă dacă multe astfel de cazuri apar în diagrama  $XoY$ , atunci scorul de consistență poate să scadă foarte mult

Într-un mod similar cu formula pentru necesitate, se poate deriva o formulă asemănătoare pentru relația de suficiență:

$$\frac{\sum \min(X_i, Y_i)}{\sum X_i} \in [0, 1]$$

În această formulă, partea de la numărător este identică și arată intersecția dintre  $X$  și  $Y$  (minimumul dintre cele două valori pentru fiecare punct), însă această sumă rezultată este raportată la suma scorurilor pe mulțimea  $X$ . Cu alte cuvinte, formula arată cât de consistentă este relația de suficiență a mulțimii  $X$  pentru mulțimea  $Y$ : cu cât scorul final va fi mai aproape de 1, cu atât mai consistentă va fi această relație.

### 5.3.6 Includere și acoperire

Ambele relații (necesitate și suficiență) dintre o condiție cauzală X și un fenomen Y se bazează, așa cum am văzut în secțiunile anterioare, pe conceptul de *includere* a unei mulțimi în cealaltă mulțime. Cele două formule pentru incluziune se folosesc în mod intermitent, una fiind reversul celeilalte pentru incluziune și acoperire, întocmai ca două fețe ale unei monede.

Determinarea relației dintre o condiție cauzală și un fenomen de interes începe cu analiza de incluziune pentru necesitate, după formula arătată mai devreme:

$$incl_{X \Leftarrow Y} = \frac{\sum \min(X_i, Y_i)}{\sum Y_i}$$

Interpretarea mărimii coeficientului de incluziune este întrucâtva similară cu interpretarea coeficientului de corelație: orice scor sub 0.5 arată o slabă consistență a acestei relații, iar un scor de 1 arată o consistență perfectă. Un scor acceptabil nu poate fi mai mic decât 0.75, preferabil peste 0.8.

Odată stabilită consistența relației de necesitate (a relației de incluziune a mulțimii Y în mulțimea X), pasul următor este stabilirea relevanței acestei incluziuni. Mulțimea Y poate fi inclusă complet în mulțimea X (cu un scor de consistență perfectă a relației de necesitate), însă este posibil ca acea condiție cauzală X să fie irelevantă pentru producerea lui Y.

De exemplu, prezența oxigenului este necesară pentru obținerea performanțelor sportive. O astfel de propoziție are logică, întrucât un sportiv generează foarte multe arderi în cadrul unui efort competitiv, iar acele arderi se bazează pe prezența oxigenului. Totuși, este clar că prezența oxigenului este irelevantă, întrucât acesta este prezent în extrem de multe cazuri în care performanța sportivă nu se produce.

Numărul de cazuri în care performanța este asociată cu prezența oxigenului este foarte mic, în raport cu numărul total de cazuri în care oxigenul este prezent. În termeni de mulțimi, acest lucru se traduce prin faptul că submulțimea Y acoperă o foarte mică parte dintr-o mulțime X foarte mare. Acest scor de acoperire este dat de următoarea formulă:

$$cov_{X \Leftarrow Y} = \frac{\sum \min(X_i, Y_i)}{\sum X_i}$$

Formula nu arată altceva decât gradul în care intersecția dintre X și Y acoperă totalul cazurilor în care X este prezentă. Este ușor de observat că

Tabelul 5.13: Scoruri de apartenență la două mulțimi distincte bivalente

a. Datele brute			b. Încrucișarea dintre X și Y	
X	Y	min(X , Y)		
0	0	0		
1	0	0		
0	1	0		
1	1	1		
1	1	1		
1	1	1		
1	0	0		
1	1	1		
1	1	1		
0	0	0		
6		5		

		X	
		0	1
Y	1	1	5
	0	2	2

această formulă este exact aceeași cu formula de incluziune pentru relația de suficiență.

Cu cât mai mare este scorul de acoperire a lui Y peste X, cu atât mai relevantă va fi condiția X în calitate de condiție necesară pentru producerea fenomenului Y.

Aceste formule sunt valabile atât pentru mulțimile vagi cât și pentru mulțimile distincte. În cazul celor distincte bivalente, scorurile de apartenență iau doar două valori: 0 și 1. Minimul dintre orice astfel de perechi nu poate fi decât fie 0 (dacă cel puțin una dintre cele două prezintă valoarea 0), fie 1, atunci când ambele sunt prezente.

Suma acestor intersecții (a minimului dintre toate perechile de valori) se dovedește a fi egală cu exact celula c din tabelul 5.3). Pentru exemplificare, să luăm datele din tabelul 5.13.a. Din cele 10 cazuri prezentate, în șase dintre acestea este prezentă fenomenul Y, iar dintre acestea în cinci cazuri este prezentă și condiția cauzală X, încrucișarea fiind prezentată în tabelul 5.13.b.

În cazul mulțimilor distincte bivalente, formula pentru incluziunea în cazul relației de necesitate a lui X pentru Y este:

$$incl_{X \leftarrow Y} = \frac{c}{a + c} = \frac{5}{1 + 5} = \frac{5}{6}$$

Rezultatul acestei formule este identic cu acela din formula pentru mulțimi vagi, întrucât suma pentru coloana min(X , Y) este egală cu 5 iar suma

pentru coloana corespunzătoare mulțimii  $Y$  este egală cu 6:

$$incl_{X \Leftarrow Y} = \frac{\sum \min(X_i, Y_i)}{\sum Y_i} = \frac{5}{6}$$

Așadar, formula pentru mulțimi vagi este universal valabilă, atât pentru mulțimile vagi cât și pentru mulțimile distincte (fie acestea bivalente sau multivalente). Spre deosebire de analiza cantitativă (statistică și probabilistică), relația de necesitate nu este simetrică pentru prezența și absența fenomenului: dacă o condiție cauzală este necesară pentru producerea fenomenului de interes, această afirmație nu spune absolut nimic despre cazurile în care fenomenul *nu* se produce.

În analiza statistică, un coeficient de corelație spune aceeași poveste despre toate cazurile (fie că variabila dependentă se produce sau nu), în schimb în analiza comparativă trebuie făcute analize separate pentru prezența și pentru absența fenomenului.

În ce privește relația de suficiență, includerea mulțimii  $X$  în mulțimea  $Y$  este dată de formula:

$$incl_{X \Rightarrow Y} = \frac{\sum \min(X_i, Y_i)}{\sum X_i}$$

Dacă scorul rezultat din această formulă este suficient de aproape de 1, atunci se în pasul următor se determină relevanța condiției  $X$  în relația de suficiență pentru producerea lui  $Y$ , cu ajutorul formulei reversate din incluziunea pentru necesitate:

$$cov_{X \Rightarrow Y} = \frac{\sum \min(X_i, Y_i)}{\sum Y_i}$$

Cu cât mai mult acoperă  $X$  din mulțimea  $Y$ , cu atât mai relevantă va fi această condiție pentru relația de suficiență în producerea fenomenului de interes  $Y$ .

Rareori se va întâmpla ca o singură condiție cauzală să fie necesară și/sau suficientă pentru producerea unui fenomen  $Y$ . Fenomenele sociale au o structură cauzală mult mai complexă, în cele mai multe din cazuri mai multe combinații (conjuncții sau disjuncții) de condiții cauzale fiind necesare pentru obținerea unui scor înalt de consistență a relației respective.

Pentru relația de suficiență, faptul că toate scorurile de apartenență la mulțimea  $Y$  sunt mai mari decât cele pentru mulțimea  $X$  înseamnă că

mai sunt și alte condiții (sau combinații de condiții) cauzale care contribuie la producerea fenomenului Y. Din acest motiv, analiza de suficiență este foarte importantă în metodologia analizei calitativ-comparative. Prin utilizarea acesteia, se poate observa dacă modelul explicativ construit pe baza condițiilor cauzale selectate este sau nu suficient (vom vedea în secțiunile care urmează că analiza de suficiență poate fi aplicată nu doar pentru o singură condiție cauzală, ci și pentru o combinație de condiții cauzale).

### 5.3.7 Tabela de adevăr

În secțiunea 5.1.5 au fost prezentate noțiunile de bază pentru construirea unei tabele de adevăr bazate pe mulțimi distincte. În esență, o tabelă de adevăr este o formă de reprezentare a ceea ce Lazarsfeld (1937) numea faimosul său concept de “spațiu de proprietăți”, bazat pe o tehnică a construcției de tipologii.

Tehnica lui Lazarsfeld identifică anumite compuneri (combinații unice) de atribute ale unui concept, iar fiecare dintre acestea generează un concept separat, compus. Atributele componente se pot aranja într-un tabel de contingență, procedura fiind demonstrată recent de Sandu (2011).

Construirea spațiilor de proprietăți a fost reluată de Elman (2005), care arată că fiecare celulă din acest tabel captează câte o grupare unică de de atribute ale unui concept, iar sarcina cercetătorului care construiește tipologii este aceea de a clasifica un caz într-una sau alta dintre aceste combinații de atribute. În abordarea lui Elman (p.5), fiecare tipologie se bazează pe o teorie care poate fi derivată în mod inductiv din observații sau în mod deductiv utilizând metode formale. Dimensiunile spațiului de proprietăți sunt oferite de variabilele pe care cercetătorul le utilizează în teoria sa, iar conținutul celulelor din acest spațiu provine din logica teoriei și din implicațiile logice ale acesteia.

Între tabelele de contingență și tabelele de adevăr există o corespondență exactă, singura diferență majoră fiind aceea că cele de contingență afișează informația într-un număr de celule, în timp ce tabelele de adevăr afișează aceeași informație într-un număr de linii. Tabelele 5.14 afișează ambele posibilități, atât pentru încrucișări  $2 \times 2$  cât și pentru  $2 \times 3$ . Acestea sunt două exemple foarte simple, însă tabelele încrucișate devin din ce în ce mai dificil de afișat la fiecare introducere a unei noi proprietăți.

În esență, dificultatea vine din faptul că un tabel încrucișat este limitat la afișarea în doar două dimensiuni (poate să crească numai pe orizontală sau numai pe verticală), așa încât introducerea unei a treia proprietăți C ar însemna încă un set de linii pe orizontală, de exemplu. Lucrurile

Tabelul 5.14: Spații de proprietăți:

a.  $2 \times 2$ 

	0	1
1	a	c
0	b	d

b.  $2 \times 3$ 

	0	1	2
1	a	c	e
0	b	d	f

c. tabelă de adevăr  $2 \times 2$ 

A	B	
0	0	b
0	1	d
1	0	a
1	1	c

b. tabelă de adevăr  $2 \times 3$ 

A	B	
0	0	b
0	1	d
0	2	f
1	0	a
1	1	c
1	2	e

se complică până când afișarea devine aproape imposibilă, la 4 sau 5 proprietăți. În schimb, această problemă este foarte simplă pentru tabela de adevăr, întrucât aceasta crește pe o singură dimensiune: numărul de rânduri, pentru fiecare nouă proprietate introdusă pe coloane.

Pentru oricât de multe proprietăți introduse în analiză, tabela de adevăr va avea un număr finit de linii, egal cu numărul de celule ale unui tabel încrucișat multidimensional. Acestea pot fi calculate foarte simplu, numărul total fiind chiar produsul numărului de stări (categorii) ale fiecărei proprietăți. Pentru două proprietăți cu două categorii fiecare (0 - prezență, 1 - absență), numărul total va fi egal cu  $2 \times 2 = 4$ .

Pentru trei proprietăți cu câte două categorii fiecare, numărul total este  $2 \times 2 \times 2 = 8$ .

Dacă numărul de categorii este egal pentru toate proprietățile, atunci formula se simplifică la  $2^3 = 8$ , iar pentru cazul general este:

$$c^k$$

unde  $c$  este numărul de categorii iar  $k$  este numărul de proprietăți.

Cazul mai general este însă acela în care numărul de categorii este (sau poate fi) diferit pentru fiecare proprietate în parte, astfel încât formula universal valabilă este:

$$c_1 \times c_2 \times \cdots \times c_k$$

În lucrarea *Fuzzy Sets Social Science*, Ragin (2000) definește mulțimile vagi ca un spațiu vectorial multidimensional cu un număr de colțuri egal cu  $2^k$ , unde  $k$  este numărul de condiții cauzale din QCA (același lucru cu numărul de atribute descriptive, sau proprietăți ale unui caz). Într-o relație similară atât cu tabelele de contingență cât și cu tabelele de adevăr, fiecare colț al acestui spațiu reprezintă o combinație unică de proprietăți, fapt care dovedește că Ragin reia sub o altă formă ceea ce Lazarsfeld avansa încă de la începutul secolului XX.

Numărul de colțuri este de  $2^k$ , întrucât pentru mulțimile vagi nu există mai multe “categorii” (ca în cazul mulțimilor distincte), ci doar două limite, minimă și maximă, reprezentate de valorile 0 și 1.

Construirea și utilizarea tabelului de adevăr pe mulțimi vagi necesită o introducere și o explicare amănunțită a relației dintre construcția de tipologii și diversele reprezentări ale tabelului de contingență asociate.

Cel mai simplu exemplu de tabel de contingență este cel  $2 \times 2$  din figura 5.18.a, reluat în figura 5.18.b într-o formă suplimentară, grafică.

După cum se poate observa, există o corespondență exactă între cele patru celule ale tabelului de contingență (diagrama a.) și cele patru colțuri ale spațiului bidimensional (diagrama b.), iar ambele pot fi mai departe reprezentate într-o tabelă de adevăr ca în tabelul 5.14.c

Aceasta este o reprezentare fidelă a spațiului de proprietăți a lui Lazarsfeld, care în utilizarea lui Ragin se numește “spațiu vectorial” (în

Figura 5.18: Tabel încrucișat:

a. formă tabelară

b. spațiu vectorial bidimensional

	0	1
1	a	c
0	b	d

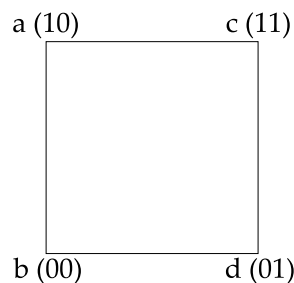
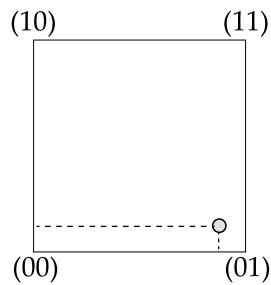


Figura 5.19: Poziționarea unui caz într-un spațiu vectorial bidimensional



acest exemplu bidimensional). Așa cum am văzut, “proprietățile” din acea tehnică a construcției de tipologii poartă numele de “condiții cauzale” în analiza calitativ-comparativă. În această perspectivă, este important de menționat faptul că tabela de adevăr se construiește numai pentru configurația de valori ale condițiilor cauzale, nu și pentru valorile de prezență sau absență ale fenomenului rezultat.

Atunci când aceste două proprietăți sau condiții cauzale au valori distincte (*crisp*), construirea tabelii de adevăr este extrem de facilă. Problema este puțin mai complicată atunci când valorile condițiilor cauzale sunt măsurate pe mulțimi vagi, cu scoruri de apartenență oriunde între 0 și 1.

Rezolvarea este însă relativ simplă cu ajutorul spațiului vectorial (bi)dimensional, în care orice caz (combinație de valori fuzzy) are niște coordonate pe orizontală și verticală. De exemplu, în figura 5.19, un caz prezintă valoarea de 0.11 pentru proprietatea A și 0.92 pentru proprietatea B, iar combinația respectivă este reprezentată de punctul albastru, poziționat foarte aproape de colțul (01) al spațiului vectorial.

În mulțimi vagi, cazul respectiv nu are exact coordonatele (0, 1) pentru a fi poziționat exact în colțul din dreapta jos, însă este suficient de aproape de acel colț pentru a fi poziționat acolo. Scorul de apartenență la acel colț este de 0.89 (metoda de calcul este prezentată mai jos) și ne așteptăm ca scorurile de apartenență la celelalte trei colțuri să fie foarte mici, întrucât punctul respectiv se află foarte departe de toate acele colțuri.

Metoda de calcul este extrem de simplă și face apel la operațiunile cu mulțimi vagi prezentate în secțiunea 4.3, în special cele de negație (NOT) și intersecție (ȘI logic). Pentru oricare dintre colțuri, valoarea 0 înseamnă absența proprietății respective (așadar scorul de apartenență trebuie negat), iar valoarea 1 înseamnă prezența proprietății (iar scorul de apartenență rămâne neschimbat).



Pentru colțul (00), ambele scoruri trebuie negate, apoi intersectate:

$$\min(\sim A, \sim B) = \min(1-0.11, 1-0.92) = \min(0.89, 0.08) = 0.08$$

Pentru colțul (10), doar scorul pentru proprietatea B trebuie negat:

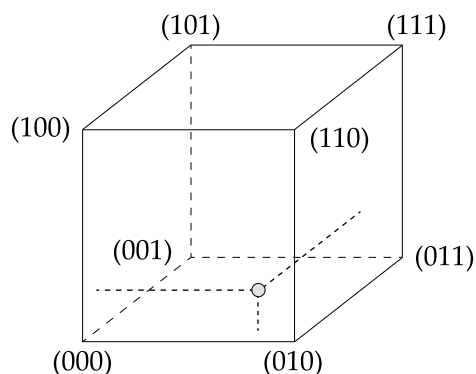
$$\min(A, \sim B) = \min(0.11, 1-0.92) = \min(0.11, 0.08) = 0.08$$

În fine, pentru colțul (11) se calculează doar intersecția dintre cele două scoruri:

$$\min(A, B) = \min(0.11, 0.92) = 0.11$$

Spațiul vectorial este un spațiu continuu, pentru a putea reprezenta orice posibile valori fuzzy ale condițiilor cauzale. Exemplul anterior poate fi complicat mai departe prin adăugarea unei a treia condiții cauzale, rezultând un spațiu vectorial tridimensional cu 8 colțuri, ca în figura 5.20 unde cazul reprezentat de punctul cu coordonate fuzzy 0.2, 0.7 și 0.1, foarte aproape de colțul 'aBc' (010).

Figura 5.20: Spațiu vectorial tridimensional



Pentru oricât de multe proprietăți ale unui fenomen, sau în QCA pentru orice număr de condiții cauzale  $k$ , se pot construi spații vectoriale cu un număr de colțuri egal cu  $2^k$ :

- pentru două condiții cauzale ( $k = 2$ ) vom avea  $2^2 = 4$  colțuri;
- pentru trei condiții cauzale ( $k = 3$ ) vom avea  $2^3 = 8$  colțuri;
- pentru patru condiții cauzale ( $k = 4$ ) vom avea  $2^4 = 16$  colțuri;
- pentru orice număr arbitrar de condiții cauzale  $k$  vom avea  $2^k$  colțuri.

Tehnica reducerii booleene pe baza comparației în perechi, prezentată în secțiunea 5.1.6, presupune existența unor mulțimi distincte, cu un număr finit de valori. Pe de altă parte, mulțimile vagi pot avea o infinitate de valori posibile, iar numărul combinațiilor posibile de valori fuzzy se multiplică odată cu creșterea numărului de condiții cauzale.

Întrucât nu este posibilă analizarea unui număr infinit de valori (combinațiile dintre acestea fiind infinite), problema trebuie redusă la transformarea valorilor fuzzy în valori distincte finite. Mai întâi se determină colțul cel mai apropiat față de coordonatele punctului, care va fi asociat mai departe cu colțul respectiv.

Un punct (mai exact un caz care prezintă o combinație de scoruri fuzzy pe mai multe condiții cauzale) poate avea orice astfel de coordonate în spațiul vectorial, apropiindu-se mai mult sau mai puțin de oricare dintre colțuri, însă apropierea cea mai mare va fi asociată numai cu unul singur colț.

Se obține un scor de apartenență parțială în oricare dintre colțuri, cu valori între 0 și 1 (cu cât este mai apropiat de un anumit colț, cu atât scorul de apartenență va crește spre valoarea 1). Pentru un singur colț, apartenența parțială va avea un scor mai mare decât 0.5, iar cazul va fi asociat cu colțul respectiv.

Tabelul 5.15 prezintă același spațiu vectorial tridimensional (corespondent celor trei condiții cauzale A, B și C), iar a patra coloană reprezintă apartenența parțială a punctului la fiecare dintre cele 8 colțuri. După cum se poate vedea, cazul respectiv are o singură apartenență mai mare decât 0.5, la colțul (010).

Tabelul 5.15: Apartenența cazului (0.2, 0.7, 0.1) la colțurile spațiului vectorial

A	B	C	Apartenență
0	0	0	0.3
0	0	1	0.1
0	1	0	0.7
0	1	1	0.1
1	0	0	0.2
1	0	1	0.1
1	1	0	0.2
1	1	1	0.1

Aceste colțuri pot reprezenta ceea ce Weber (2003) numește “tip ideal”. Birocrația este un tip ideal de organizare, iar fiecare țară are un sistem mai mult sau mai puțin birocratic (din nou, a cărui cuantificare nu poate fi decât vagă). Dacă proprietatea A ar însemna sărăcie, proprietatea B birocratie, iar proprietatea C ar însemna democrație, atunci un caz care prezintă combinația (011) se caracterizează printr-un sistem simultan birocratic, democrat și care nu prezintă deloc o stare de sărăcie.

Deși există susținători ai ceea ce se numește “ideal-type QCA”, opinia care pare să fie totuși unanim acceptată este că tipurile ideale nu pot fi utilizate ca atare în analiză întrucât acestea nu există niciodată în realitate, fapt care împiedică un proces corespunzător de calibrare a mulțimilor vagi observate. Dacă toate ancorele calitative ar fi stabilite la un nivel ideal mult prea departe de realitate, se pierde din puterea de discriminare între scorurile observate.

Din acest motiv, se preferă înlocuirea conceptului de tipuri ideale cu cele de “cazuri tipice”, acestea fiind așezate în fiecare colț al spațiului vectorial, iar cazurile observate sunt poziționate înăuntrul acestui spațiu prin raportare la cazurile tipice din colțuri.

Calcularea scorurilor de apartenență ale cazurilor la colțurile spațiului vectorial este doar un prim pas pentru derivarea întregii table de adevăr. Pasul al doilea este chiar mai important, întrucât mai rămâne de stabilit dacă acele  $2^k$  combinații de valori distincte au sau nu o legătură cu producerea fenomenului de interes. Dacă în primul pas au fost analizate scorurile de apartenență pe condițiile cauzale și poziționarea cazurilor în spațiul vectorial, în pasul al doilea se stabilește scorul distinct pentru fenomenul de interes, în dreptul fiecărui colț al spațiului vectorial.

Pentru unele dintre colțuri, vom putea stabili dacă fenomenul de interes se produce (valoarea 1), pentru alte colțuri vom stabili dacă fenomenul nu se produce (valoarea 0), iar pentru cele multe dintre colțuri nu vom putea spune nimic despre legătura cu prezența sau absența fenomenului, din lipsă de evidențe empirice.

Întregul algoritm de transformare a mulțimilor vagi în mulțimi distincte, pentru obținerea tablei de adevăr complete, a fost publicat tot de către Ragin (2005), care dovedește o imaginație incredibilă, dublată de un spirit constructiv demn de un întemeietor de școală.

Pentru a exemplifica întreaga procedură, voi face apel la un set de date ipotetic prezentat în tabelul 5.16. Acolo se regăsesc trei condiții cauzale A, B și C, precum și un fenomen de interes. Pentru fiecare dintre acestea sunt prezentate scorurile fuzzy ale unui număr de 8 cazuri numerotate de

Tabelul 5.16: Set de date ipotetic cu trei condiții cauzale și un fenomen

	A	B	C	Y
R <sub>1</sub>	0.2	0.7	0.1	0.8
R <sub>2</sub>	0.8	0.1	1.0	0.3
R <sub>3</sub>	0.0	0.0	1.0	0.2
R <sub>4</sub>	0.7	0.9	0.4	0.8
R <sub>5</sub>	0.9	0.6	1.0	0.7
R <sub>6</sub>	0.7	0.7	0.3	0.4
R <sub>7</sub>	1.0	0.2	0.9	0.9
R <sub>8</sub>	0.9	0.7	0.6	0.7

la R<sub>1</sub> la R<sub>8</sub>. Într-o cercetare reală, numărul cazurilor analizate poate fi mult mai mare, însă, pentru scopuri de prezentare a procedurii, este util un set de date de dimensiuni reduse.

Aceste cazuri ar putea reprezenta, de exemplu, cele 8 regiuni de dezvoltare ale României și este o pură întâmplare faptul că numărul acestor regiuni este perfect egal cu numărul celor  $2^3 = 8$  colțuri ale spațiului vectorial.

Procedura de construcție a tabelului de adevăr continuă cu următoarea observație: dacă un caz are câte o apartenență parțială la fiecare dintre colțurile spațiului vectorial, rezultă că fiecare colț are câte o distribuție de scoruri de apartenență de-a lungul tuturor cazurilor analizate.

Tabelul 5.17 prezintă întreaga matrice de scoruri de apartenență, pe linii fiind prezentate cazurile (cele 8 regiuni de dezvoltare) iar pe coloane cele

Tabelul 5.17: Matricea scorurilor de apartenență la colțuri

Reg.	abc	abC	aBc	aBC	Abc	AbC	ABc	ABC
	000	001	010	011	100	101	110	111
R <sub>1</sub>	0.3	0.1	0.7	0.1	0.2	0.1	0.2	0.1
R <sub>2</sub>	0.0	0.2	0.0	0.1	0.0	0.8	0.0	0.1
R <sub>3</sub>	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
R <sub>4</sub>	0.1	0.1	0.3	0.3	0.1	0.1	0.6	0.4
R <sub>5</sub>	0.0	0.1	0.0	0.1	0.0	0.4	0.0	0.6
R <sub>6</sub>	0.3	0.3	0.3	0.3	0.3	0.3	0.7	0.3
R <sub>7</sub>	0.0	0.0	0.0	0.0	0.1	0.8	0.1	0.2
R <sub>8</sub>	0.1	0.1	0.1	0.1	0.3	0.3	0.4	0.6

8 colțuri ale spațiului vectorial. Se poate observa că prima linie a acestei matrici conține exact scorurile din coloana cu apartenența din tabelul 5.15, specifice cazului  $R_1$  cu coordonatele (0.2, 0.7, 0.1).

În interiorul acestei matrici, sunt evidențiate celulele care arată spre care colțuri sunt distribuite cazurile: spre anumite colțuri pot fi distribuite mai multe cazuri, în timp ce alte colțuri pot să fie complet goale, niciunul dintre cazuri să fie poziționate aproape de acestea.

Atribuirea scorurilor distincte pentru prezența sau absența fenomenului se face pentru fiecare colț al spațiului vectorial în parte, pe baza tehnicii de incluziune prezentată într-o secțiune anterioară. Regula de bază este următoarea: dacă mulțimea scorurilor de apartenență la un colț este inclusă în mulțimea scorurilor de apartenență la fenomenul observat, atunci putem concluziona că acea combinație cauzală specifică colțului respectiv este suficientă pentru producerea fenomenului, așadar vom codifica scorul distinct 1.

De exemplu, primul colț față de care cel puțin un caz are un scor de apartenență mai mare decât 0.5 este abC (001). Distribuția scorurilor de apartenență a cazurilor la acest colț este prezentată în oglindă față de distribuția scorurilor de apartenență ale acelorași cazuri la fenomenul observat (tabelul 5.18), pentru a stabili dacă prima mulțime este sau nu inclusă în a doua, pe baza formulei de incluziune pentru suficiență:

$$incl_{abC \rightarrow Y} = \frac{\sum \min(abC, Y)}{\sum abC} = \frac{1.1}{1.9} = 0.579$$

Tabelul 5.18: Testarea incluziunii colțului abC (001) în Y

	abC	Y	min(abC, Y)
$R_1$	0.1	0.8	0.1
$R_2$	0.2	0.3	0.2
$R_3$	1.0	0.2	0.2
$R_4$	0.1	0.8	0.1
$R_5$	0.1	0.7	0.1
$R_6$	0.3	0.4	0.3
$R_7$	0.0	0.9	0.0
$R_8$	0.1	0.7	0.1
	1.9		1.1

Incluziunea acestei combinații cauzale în fenomenul Y este destul de scăzută. În mod normal, ar trebui ca mulțimea scorurilor de apartenență la colțul  $abC$  (001) să fie complet inclusă în mulțimea Y pentru a putea spune că acea combinație este suficientă pentru producerea fenomenului.

Totuși, în practică nu se observă aproape niciodată o relație de suficiență perfectă (cu un scor de incluziune egal cu 1), de aceea se acceptă scoruri puțin mai mici decât 1 însă ideal cât mai aproape de scorul maxim. Într-un fel, aprecierea mărimii acestui scor este similară cu interpretarea coeficientului de corelație din analiza cantitativă: nu se va observă niciodată o corelație perfectă, egală cu 1. De aceea, ideal ar fi ca acest coeficient să fie cât mai mare: cu cât mai aproape de 1, cu atât mai bine.

Din rațiuni care țin o practică empirică similară, un scor de incluziune egal cu 0.9 este considerat un prag acceptabil. Preferabil ar fi un scor de incluziune egal cu 0.95, iar în partea minimă nu se va accepta un scor de incluziune mai mic decât 0.75. În cazul combinației  $abC$  (001), scorul de incluziune este de numai 0.579 de unde putem concluziona că această combinație cauzală nu este suficientă pentru producerea fenomenului de interes, drept pentru care acesta va fi codificat cu valoarea 0.

Un caz contrar este prezentat în tabelul 5.19 de următoarea combinație cauzală  $aBc$  (010), pentru care incluziunea este perfectă cu un scor egal cu 1.0 în mulțimea fenomenului observat.

$$incl_{aBc \Rightarrow Y} = \frac{\sum \min(aBc, Y)}{\sum aBc} = \frac{1.4}{1.4} = 1.000$$

Tabelul 5.19: Testarea incluziunii colțului  $aBc$  (010) în Y

	aBc	Y	min(aBc, Y)
R <sub>1</sub>	0.7	0.8	0.7
R <sub>2</sub>	0.0	0.3	0.0
R <sub>3</sub>	0.0	0.2	0.0
R <sub>4</sub>	0.3	0.8	0.3
R <sub>5</sub>	0.0	0.7	0.0
R <sub>6</sub>	0.3	0.4	0.3
R <sub>7</sub>	0.0	0.9	0.0
R <sub>8</sub>	0.1	0.7	0.1
	1.4		1.4

Toate scorurile de apartenență la acest colț sunt mai mici decât scorurile de apartenență la fenomenul studiat, ceea ce înseamnă că acea combinație cauzală este complet inclusă în mulțimea  $Y$  deci arată o relație de suficiență perfectă pentru producerea acestuia.

Calculând în acest fel câte un scor de incluziune pentru fiecare dintre celelalte combinații cauzale care au cel puțin câte un caz poziționat în apropiere ( $AbC$ ,  $ABc$  și  $ABC$ ), rezultă astfel o tabelă de adevăr completă, prezentată în tabelul 5.20.

În coloana rezultat “OUT” se pot observa alocările de scoruri distincte pentru fenomenul studiat. Pe coloana “n” sunt numărul de cazuri în apropierea fiecăruia dintre cele 8 colțuri ale spațiului vectorial, iar pe coloana “incl” sunt prezentate scorurile de incluziune ale combinațiilor cauzale respective în mulțimea fenomenului studiat, pe ipoteza suficienței în producerea acestuia.

De remarcat faptul că liniile de la 1 la 8 în tabela de adevăr nu se confundă cu cele 8 regiuni, ci reprezintă totalul celor  $2^3 = 8$  combinații posibile. Regiunile sunt prezentate pe ultima coloană, de exemplu a șasea combinație (101) se regăsește în două regiuni, anume  $R_2$  și  $R_7$ .

Rezolvarea acestei tabele și obținerea soluției urmează apoi aceeași procedură ca în procesul de minimizare booleană pentru mulțimi distincte, pe linia exemplului din secțiunea 5.1.6: se selectează combinațiile cauzale asociate cu prezența fenomenului de interes (în acest exemplu combinațiile 3 și 8), apoi se aplică mecanismul de minimizare booleană până la obținerea implicanților primari (mai multe detalii în secțiunea următoare, pentru includerea în analiză a combinațiilor contrafactice, în scopul obținerii celei mai simple soluții minime cu putință).

Tabelul 5.20: Tabela de adevăr pentru mulțimi vagi

	A	B	C	OUT	n	incl	cazuri
1	0	0	0	?	0		
2	0	0	1	0	1	0.579	$R_3$
3	0	1	0	1	1	1.000	$R_1$
4	0	1	1	?	0		
5	1	0	0	?	0		
6	1	0	1	0	2	0.821	$R_2, R_7$
7	1	1	0	0	2	0.850	$R_4, R_6$
8	1	1	1	1	2	1.000	$R_5, R_8$

În acest fel, se pot realiza operațiuni boolene specifice mulțimilor distincte, pe baza unor valori provenite din mulțimi vagi. Analiza dobândește astfel un caracter reproductibil, aceleași rezultate fiind obținute pe baza acelorași date de intrare, indiferent de cercetătorul care le analizează.

### 5.3.8 Așteptări direcționale și asumptii simplificatoare

În tabelul 5.20 se pot observa câteva combinații cauzale în dreptul cărora codificarea pentru rezultat (OUT) este un semn de întrebare “?”. Așa cum am văzut în secțiunea anterioară, acelea sunt colțuri ale spațiului vectorial pentru care nu există niciun caz observat empiric în apropiere (nu avem evidențe pentru a stabili dacă fenomenul studiat se produce).

Pentru doar trei condiții cauzale, numărul combinațiilor pentru care nu avem evidențe empirice este relativ mic însă pe măsură ce se introduc în analiză mai multe astfel de condiții, numărul acestor combinații începe să crească dramatic. Este important de menționat faptul că această lipsă de evidențe empirice nu se datorează unei slabe culegeri a datelor de pe teren, ci dintr-un motiv mult mai subtil, discutat în secțiunea 2.6.5: în lumea socială, diversitatea cauzală este extrem de limitată.

Există desigur mai multe modalități de producere a unui fenomen, însă numărul acestora este de fapt extrem de mic în comparație cu mărimea numărului matematic posibil de combinații ale condițiilor cauzale. De exemplu, pentru 6 condiții cauzale numărul tuturor combinațiilor posibile este de  $2^6 = 64$ , însă în realitate nu vor fi întâlnite mai mult de 10-15 astfel de combinații. Chiar dacă se vor analiza sute sau chiar mii de cazuri, toate vor fi aglomerate în aceste 10-15 de combinații observate, restul fiind doar teoretic posibile însă aproape niciodată întâlnite în practică.

Acestea din urmă se numesc “combinații rămase” (în engl. *remainders*) întrucât sunt niște combinații pentru care nu există dovezi empirice pentru a stabili dacă vor contribui sau nu la producerea fenomenului de interes.

Atât în literatura de specialitate, cât mai ales în general în filozofie, aceste combinații cauzale se mai numesc cazuri ipotetice sau “contrafactice” (discutate în secțiunea 2.2), iar oamenii de știință au lucrat dintotdeauna cu aceste experimente mentale imaginându-și ce s-ar întâmpla dacă aceste condiții cauzale s-ar întruni în realitate. Utilizarea contrafactualilor este una dintre cele mai vechi și mai plăcute experimente mentale din lume. De exemplu, pe baza acestora se bazează o întreagă industrie de publicații SF, ori pe baza acestora au fost inventate toate poveștile din lume. În știință, școala filozofică grecească, antică, se baza aproape exclusiv pe utilizarea experimentelor mentale



întrucât nu aveau nici posibilitățile tehnice, nici teoriile moderne care să le permită experimentele avansate care sunt posibile astăzi.

Spre deosebire de analiza cantitativă, care încearcă delimitarea efectului net al fiecărei variabile independente asupra celei dependente (ținând sub control toate celelalte variabile), analiza comparativă funcționează pe un principiu mai aproape de simțul comun. În realitate, fenomenul de interes nu se produce aproape niciodată pe baza unei singure variabile independente (a unei singure condiții cauzale), ci doar în prezența mai multor variabile simultan.

Într-o lucrare recentă, Rihoux și Ragin (2009) au contribuit la întărirea concluziei că nu este deloc întâmplător faptul că numai anumite configurații cauzale se întâlnesc în practică, în timp ce altele nu. Anumite condiții cauzale se întrunesc numai în anumite conjuncturi, numai în anumite configurații cauzale.

În lumea fizică, prezența focului nu poate fi realizată decât în strânsă legătură cu prezența oxigenului. Pe un principiu similar în lumea socială, menținerea unui sistem de cercetare avansat este strâns legată de prezența unei economii dezvoltate. Aceste două condiții avansate merg mână în mână, fiind puțin probabil ca o economie dezvoltată să nu aibă și un sistem de cercetare avansat, după cum este și mai puțin probabil ca o economie slab dezvoltată să mențină un sistem de cercetare avansat.

Așadar, faptul că realitatea tinde să prezinte condiții cauzale numai sub formă de configurații contribuie la diversitatea empirică foarte limitată de care cercetătorii se lovesc în practică, iar unele combinații rămase sunt de-a dreptul imposibile. De exemplu, dacă o condiție cauzală A este genul, iar o altă condiție cauzală B este prezența sarcinii, atunci combinația dintre categoria “bărbat” și “gravid” este imposibilă și nu va fi întâlnită niciodată în practică.

Pentru situațiile în care combinațiile observate sunt foarte puține, soluțiile rezultate din minimizarea booleană sunt foarte complexe, până la un punct de maxim în care găsim câte o explicație pentru fiecare caz în parte, ceea ce contravine nevoii de generalitate specific unei teorii.

Pentru a obține soluții cât mai simple (cu un grad de generalitate cât mai înalt), o tehnică acceptată în analiza comparativă este includerea combinațiilor rămase (*remainders*) în procesul de minimizare booleană, pe ideea experimentelor mentale și a utilizării configurațiilor contrafactice: ce s-ar întâmpla dacă aceste combinații ar exista în realitate.

Desigur, acestea sunt combinații cauzale fără suport empiric (observat), însă rolul acestora este de a determina o soluție minimă, un fel de chintesență cauzală care să fie asociată cu producerea fenomenului, în

orice context plauzibil. Includerea contrafactualilor are rolul de a transforma soluțiile complexe, cu explicație contextuală, în soluții mai simple dar cu o putere de explicație mai largă.

Includerea combinațiilor rămase are acest rol, însă utilizarea nediscriminată a acestora poate să ducă la concluzii nevalide. Pot exista combinații imposibile, pe exemplul bărbatului gravid, însă chiar și dintre combinațiile plauzibile trebuie făcută o selecție atentă. După eliminarea combinațiilor imposibile, această a doua etapă de selecție se face pe baza așa numitor “așteptări direcționale”.

În esență, selectarea contrafactualilor plauzibili se face utilizând teoriile acceptate. De exemplu, dacă o teorie spune că există o legătură (la nivel comunitar) între un nivel înalt de educație și producerea bunăstării, atunci relația dintre cele două are o anumită direcție: cu cât mai mare educația, cu atât mai mare bunăstarea. Ne “așteptăm” ca bunăstarea să se producă în prezența unui nivel înalt de educație și nu în absența acestora.

Utilizând această așteptare direcțională, dintre combinațiile rămase vor fi eliminate acelea care nu prezintă un nivel de educație înalt, fiind păstrate pentru includerea în procesul de minimizare doar pe acelea care sunt conforme cu teoriile acceptate.

În finalul acestei secțiuni, trebuie reamintit faptul că analiza care implică prezența unui fenomen este asimetrică de analiza care implică absența acestuia. Același lucru se aplică și în situația includerii unor contrafactuali plauzibili: dacă teoria spune că *prezența* lui X conduce la producerea lui Y, cu totul altceva este să afirmăm ceva despre efectele produse de *absența* lui X. În teoria de specialitate, primul exemplu se referă la “contrafactuali ușori”, iar al doilea exemplu se referă la “contrafactuali dificili”.

Recomandarea este să se utilizeze fără probleme contrafactualii ușori, iar cei dificili să fie introduși în procesul de minimizare doar în urma unui proces foarte minuțios de raportare teoretică. Tocmai de aceea sunt denumiți dificili (nu reprezintă cazuri imposibile), pentru că introducerea lor în analiză se poate face numai după un proces foarte dificil de inspecție atentă.

Acest lucru este justificabil din punct de vedere logic, întrucât este relativ simplu să facem ipoteze despre ceva care se întâmplă (uneori prin comparație cu alte lucruri care se întâmplă), însă este cu mult mai dificil să elaborăm teorii despre ceva care nu se întâmplă (pe aceeași idee din logică asupra imposibilității de a demonstra ceva imposibil).

După cum putem vedea, utilizarea contrafactualilor este o etapă relativ complexă în analiza comparativă, iar dezvoltarea teoretică în acest domeniu este în plină dezvoltare. Există o mulțime de instrumente

practice pentru identificarea diferitelor tipuri de contrafactuali, de la utilizarea programelor software până la analiza plotului XoY pe diverse zone ale acestuia (Schneider și Wagemann, 2012; Thiem și Dușa, 2013).

Tratând o teorie ipotetică despre cauzele rășcoalelor țărănești (merită amintit faptul că Ragin a fost interesat și chiar a publicat articole despre rășcoale, inclusiv despre cea românească din 1907), Ragin (1987) a arătat care este tratamentul care poate fi aplicat liniilor din tabela de adevăr pentru care nu se cunoaște rezultatul în variabila efect.

Condițiile cauzale din modelul său arată astfel: A = Tradiționalism; B = Comercializare a agriculturii; C = Țărani cu situație medie și D = Abstenteismul elitelor.

Dintre combinațiile posibile de asemenea condiții cauzale, unele conduc la rășcoală, iar altele nu conduc la pornirea unei rășcoale. Pentru restul de combinații cauzale, nu se cunoaște care ar putea fi rezultatul. Pentru a putea spune ceva despre acestea, o primă operațiune este minimizarea combinațiilor care nu există. Ragin are o abordare diferită, mai întâi calculând soluția minimizată pentru societățile care *există* (atât cele care produc cât și cele care nu produc rășcoale), după care aplică legile lui de Morgan pentru a calcula rezultatul invers pentru societățile care *nu există*. Am arătat însă că o aplicare (manuală) a legilor lui de Morgan nu poate fi mai rapidă decât utilizarea algoritmilor moderni de minimizare booleană.

După încheierea acestei operațiuni, Ragin a obținut două seturi de soluții:

$$\begin{array}{ll} ac + aD + BD + Abd & \text{pentru societățile țărănești care } \textit{există} \\ ABd + aCd + AbD + BCd & \text{pentru societățile țărănești care} \\ & \textit{nu există} \end{array}$$

Cu alte cuvinte, cele mai întâlnite tipuri de societăți țărănești sunt cele în care:

$$\begin{array}{ll} ac & = \text{nivelul de tradiționalism este scăzut și există puțini țărani} \\ & \text{de condiție medie} \\ & \text{sau} \\ aD & = \text{nivelul de tradiționalism este scăzut și elitele sunt absente} \\ & \text{sau} \\ BD & = \text{agricultura se comercializează iar elitele sunt absente} \\ & \text{sau} \\ ABd & = \text{tradiționalismul este înalt, există puțini țărani de condiție} \\ & \text{medie, iar elitele locale sunt prezente} \end{array}$$

În a doua etapă, Ragin selectează din cadrul combinațiilor existente doar pe cele care produc o răscoală țărănească, aplicând o nouă minimizare booleană pentru a obține soluția:  $R = ABD + aCD$

Conform datelor existente, răscoala este provocată în două mari combinații de condiții cauzale. Se poate observa că una dintre condițiile cauzale (D) este prezentă în ambele combinații, așa încât absenteismul elitelor este o condiție necesară (nu și suficientă) pentru provocarea răscoalelor.

Mai departe, Ragin este interesat de cazurile lipsă, întrebându-se ce s-ar fi întâmplat dacă celelalte combinații cauzale ar fi existat. Dacă se permite asumția că aceste combinații *ar fi produs* răscoale în cazul în care ar fi existat, atunci combinațiile respective pot fi introduse în procesul de minimizare booleană pentru obținerea unei soluții mai simple. În urma unor asemenea “asumții simplificatoare”, așa cum le denumește Ragin (1987, p.112), soluția poate fi și mai mult minimizată:  $R' = AB + AC$

Acesta este un exemplu relativ simplu de tratare a cazurilor lipsă pentru rezolvarea problemei diversității limitate. Sigur, toate asumțiile simplificatoare trebuie justificate teoretic și sunt expuse criticilor metodologice. Cu toată fragilitatea aplicării unui asemenea procedeu, trebuie ținut seama că (într-un mod mai mult sau mai puțin explicit) astfel de asumții sunt utilizate frecvent în cercetarea socială la toate nivelurile.

În plus, o asumție simplificatoare este analogul experimentului imaginar al lui Weber: “ce s-ar întâmpla dacă...”. Așa cum am văzut, combinațiile rămase sunt echivalentul condițiilor cauzale contrafactice, așadar o gamă întreagă de unelte metodologice care converg în aceeași direcție pentru realizarea unei analize comparative.

Așteptările direcționale și asumțiile simplificatoare sunt, așa cum am văzut, două aspecte strâns legate de problema diversității limitate și de nevoia de a include în analiză cazuri contrafactice plauzibile pentru a obține o soluție minimă, la un nivel de generalitate cât mai înalt.

În științele sociale, faptul că diversitatea este limitată trebuie tratat ca pe un dat și nu ca pe o situație neplăcut problematică. Este un simplu aspect care caracterizează societățile umane, iar acest lucru se întâmplă deoarece situațiile sociale sunt relativ omogene sub aspectul organizării. În fond, fiecare societate are un sistem juridic, un sistem de organizare ierarhică, o stratificare și o distribuție inegală a veniturilor. Deși în mod logic ar putea exista și o societate în care veniturile să fie distribuite perfect egal, în care să nu fie nevoie de un sistem juridic întrucât toți indivizii se înțeleg perfect (eliminând astfel și necesitatea unei organizări

ierarhice), în realitate un astfel de sistem este doar ipotetic și nu va putea fi studiat niciodată.

Spre deosebire de fizică, unde experimentele presupun (și chiar obligă cercetătorul) la măsurători bazate pe diverse manipulări ale variabilelor experimentale, în științele sociale acest lucru este imposibil: nu putem interveni asupra unei societăți doar pentru a experimenta reacția la diverși stimuli. Nu putem face altceva decât să observăm ceea ce există, și să facem asumții asupra ceea ce s-ar putea întâmpla dacă anumite condiții cauzale ar fi îndeplinite.

## 5.4 Un exemplu demonstrativ

Este important de subliniat, de la bun început, că acest exemplu trebuie privit doar din punct de vedere demonstrativ. Nu este vorba de o cercetare în sensul ei clasic și nu urmărește testarea vreunei ipoteze specifice, iar pentru scopul acestei secțiuni nu sunt importante întrebările fundamentale teoretice din spatele acestuia și nici ipotezele pe care se bazează, ci doar demonstrația practică a aplicării algoritmului QCA.

În această secțiune voi demonstra cum se pot utiliza datele din cercetarea cantitativă, la nivel agregat, pentru efectuarea de comparații la nivel structural. Desigur, QCA nu este singurul tip de analiză care se poate aplica pe astfel de date, însă accentul cade pe diversitatea metodologică și multitudinea de metode de cercetare posibile, care se pot valida reciproc dacă produc concluzii similare.

Au fost utilizate datele din European Values Survey, valul 3 din 1999 și valul 4 din 2008, baza de date pe România (datele sunt disponibile pe pagina de web a proiectului <http://www.europeanvaluesstudy.eu/>).

Fenomenul de interes este implicarea în societatea civilă, studiată de Pralong (2008). Coordonatoare a Fundației pentru o Societate Deschisă la începutul anilor 1990, Pralong a fost interesată de emergența și dezvoltarea societății civile, cu implicare pe oferirea de finanțare pentru proiecte în care indivizii se asociază și urmăresc un interes comun.

Observând că numărul de proiecte este mai mare în Transilvania decât în celelalte zone istorice, Pralong a emis ipoteza conform căreia implicarea în societatea civilă are determinanți culturali (diversitatea etnică fiind mai mare în acea zonă a țării) și a urmărit să o testeze în teza sa de doctorat.

Pentru măsurarea dimensiunilor culturale ale unei populații, Pralong a utilizat teoria lui Hofstede (1997), care a studiat peste 100.000 de angajați ai IBM în nu mai puțin de 50 de țări de pe 5 continente. Prin

selectarea de angajați pe poziții similare, Hofstede a reușit să controleze factorii educaționali ori cei de specific instituțional, interesându-se numai de valorile culturale. Hofstede a enumerat cinci mari dimensiuni:

1. **Distanța față de putere** (PDI - Power Distance Index) măsoară nivelul perceput de egalitate existent în societate. Hofstede a măsurat această distanță din perspectiva indivizilor (nu a elitelor), din perspectiva celor lipsiți de putere. Cu cât distanța percepută față de elite (față de influențarea deciziilor acestora) este mai mare, cu atât societatea este percepută ca inegalitară; invers, cu cât această distanță este mai mică, societatea este percepută ca având inegalități minime între indivizi.
2. **Masculinitate vs. Feminitate** (MASC - Masculinity). Observând că rolurile asociate cu fiecare gen sunt definite cultural, Hofstede punctează faptul că societățile cu valori masculine sunt orientate către obținerea succesului și a foloaselor materiale, asumarea unor activități de putere și forță, în timp ce orientarea către valori feminine caracterizează societățile care sunt mai degrabă preocupate de grija față de membrii acesteia, în vederea îmbunătățirii calității vieții lor.
3. **Individualism vs. Colectivism** (IND - Individualism). Această dimensiune (de fapt o dualitate polară) urmărește diferențierea între culturile centrate pe individ și cele orientate spre valorile de grup. În societățile individualiste, fiecare individ trebuie să se descurce în viață mai mult sau mai puțin pe cont propriu, spre deosebire de societățile colectiviste în care fiecare membru este strâns integrat în grupul din care face parte. Două exemple tipice sunt Statele Unite (centrate vizibil spre valorile individuale) și Japonia (unde individul nu poate exista în afara unui grup).
4. **Evitarea incertitudinii** (UAI - Uncertainty Avoidance Index). Prin intermediul acestei dimensiuni, Hofstede a încercat să măsoare toleranța față de neprevăzut, pe ideea că “ceea ce este diferit, este periculos”. Ideea i-a venit atunci când călătorea cu trenul în Germania, observând exactitatea cu care veneau și plecau acestea, spre deosebire de alte țări în care câteva minute de întârziere sunt un fapt cotidian. Germanii așteaptă exactitate, își consultă permanent ceasul și intră imediat în alertă dacă lucrurile întârzie, fapt care demonstrează o toleranță scăzută față de situațiile neprevăzute, față de necunoscut.
5. **Orientarea pe termen lung sau scurt** (LTO - Long Term Orientation). Această ultimă dimensiune a fost utilizată de

Hofstede pentru a diferenția culturile asiatice de cele europene. Există ceva la asiatici, fundamental diferit de culturile europene, imposibil de surprins cu ajutorul primelor patru dimensiuni (care conțin un bias al valorilor vestice).

Desigur, teoria lui Hofstede este criticabilă atât sub aspectul ipotezelor de plecare cât și din punct de vedere metodologic, mai ales că acesta a făcut inferențe la nivel de țară utilizând date dintr-o singură companie, însă faptul este că descoperirile acestuia nu pot fi neglijate. Între timp, Hofstede și-a mai rafinat teoria și a emis o a șasea dimensiune culturală, însă în această lucrare voi trata doar cele 5 dimensiuni originale care au fost luate în considerare în lucrarea scrisă de Pralong.

Un aspect deosebit de interesant (și crucial totodată) al teoriei lui Hofstede este faptul că se referă numai la caracteristici culturale *agregate*. Deși a utilizat date la nivel individual (chestionare cu angajații IBM), indicii calculați de el au sens doar la nivel agregat. Cu alte cuvinte, nu există caracteristici culturale la nivel de individ, ci doar la un nivel de agregare superior (în acest caz țara).

Întrucât intenția lui Hofstede a fost de a compara țări (mai bine spus, culturi), el nu calculează decât măsuri la acest nivel de agregare și nu analizează niciun detaliu la nivel de individ.

Urmând un traseu metodologic similar, am încercat să testez ipoteza emisă de Pralong utilizând datele European Values Survey (valul 3 din 1999 și valul 4 din 2008) pentru a compara diferențele în ce privește implicarea în societatea civilă între cele 18 regiuni socio-culturale definite de Sandu (1998). Datele EVS se pot utiliza în același mod pentru comparații (inter)naționale, însă lucrarea scrisă de Pralong se referă la diferențierea culturală în interiorul României.

Deși este o cercetare care se referă la valori (care țin într-o oarecare măsură de cultura națională), chestionarul EVS nu a fost conceput special pentru a calcula indicii pentru dimensiunile lui Hofstede. În mod normal, acele dimensiuni și indicatorii asociați ar trebui adaptate special pentru o astfel de cercetare, însă pentru scopul acestui exemplu am considerat că o analiză secundară pe datele EVS este suficientă.

Pentru unii indicatori (de exemplu ai dimensiunii Masculinitate / Feminitate), întrebările din EVS se referă chiar la același cadru instituțional utilizat de Hofstede, anume locul de muncă. De aceea, în acord cu itemii utilizați în lucrarea originală a acestuia, precum și în acord cu explicațiile oferite de Hofstede asupra naturii dimensiunii respective, am utilizat trei itemi proveniți de la întrebarea: “Care dintre

următoarele lucruri le considerați importante pentru o slujbă?”

	Menționat	Nemenționat
v69. Să fie bine plătită	1	2
v74. Posibilitatea de a avea inițiativă	1	2
v78. O slujbă în care să simți că poți realiza ceva	1	2

Valorile acestor trei itemi pot fi ușor transformate în valori binare, prin simpla diferență din cifra 2. Rămâne cifra 1 (din  $2 - 1 = 1$ ) pentru categoria “Menționat” și se obține cifra 0 (din  $2 - 2 = 0$ ) pentru categoria “Nemenționat”.

Întrucât aceste două răspunsuri posibile au devenit binare (de tip Da sau Nu), poate fi construit un scor metric individual pentru “Masculinitate” prin însumarea valorilor la cei trei itemi, care variază între un minim egal cu 0 (minim de masculinitate, sau maxim de feminitate), și un maxim egal cu 3 (maxim de masculinitate, sau minim de feminitate). Pentru regiunile socio-culturale, scorul a fost agregat prin calcularea mediei scorurilor individuale.

Pentru dimensiunea IND - individualism, au fost selectați patru itemi măsurați pe o scală de răspuns diferită, de la 1 la 10:

v194.	Fiecare individ ar trebui să-și asume mai multă responsabilitate pentru propria bunăstare	Statul ar trebui să își asume mai multă responsabilitate pentru bunăstarea individuală
	1 2 3 ... 8 9	10
v197.	Statul ar trebui să acorde mai multă libertate firmelor	Statul ar trebui să controleze mai mult firmele
	1 2 3 ... 8 9	10
v198.	Diferențele între venituri ar trebui să se micșoreze	Diferențele între venituri ar trebui să fie și mai mari
	1 2 3 ... 8 9	10
v199.	Proprietatea privată ar trebui să se extindă și să se dezvolte	Proprietatea de stat ar trebui să se extindă și să se dezvolte
	1 2 3 ... 8 9	10



Trei dintre acești indicatori (v194, v197 și v199) au fost inversați întrucât individualismul se găsește pe polul din stânga al dualității, iar ulterior toți cei patru itemi au fost aduși la o scală de 0 ... 9 prin scăderea unei unități. În final, scorul total individual a fost calculat prin suma valorilor de la cei patru itemi, iar scorul agregat pe regiuni socio-culturale a fost calculat prin media scorurilor sumative individuale. Scăderea unei unități pentru fiecare item a fost necesară întrucât calcularea mediei presupune în mod implicit (matematic) prezența valorii 0 în cadrul valorilor, iar scala de răspuns originală începe abia de la cifra 1. Dintr-o variabilă cvasi-ordinală, se obține astfel o variabilă metrică pentru care putem aplica operațiunii matematice de tipul sumei și mediei.

În ce privește dimensiunea PDI - distanța față de putere, nu am identificat decât un singur item în chestionarul EVS:

	Menționat	Nemenționat
v84. Să ai un cuvânt de spus în luarea deciziilor importante	1	2

Acest item a fost inversat prin scăderea unei unități, obținând în același timp și valori binare cu cifra 0 ( $1 - 1 = 0$ ) pentru categoria “Menționat” și cifra 1 ( $2 - 1 = 1$ ) pentru categoria “Nemenționat”. Inversarea a fost necesară întrucât indivizii care au un cuvânt de spus în luarea deciziilor reprezintă categoria celor cu distanță “mică” față de putere (cifra 0), iar indivizii care nu au nici un cuvânt de spus reprezintă categoria celor cu distanță “mare” față de putere (cifra 1). Scorul agregat pe regiuni socio-culturale a fost calculat prin simpla medie a scorurilor individuale.

În cazul dimensiunii LTO - orientarea pe termen lung, au fost utilizați patru itemi proxy pe baza întrebării Q52: “Am să vă citesc o listă cu lucruri pe care copiii le pot învăța acasă. Considerați că există calități importante în mod deosebit? Care?”

	Menționat	Nemenționat
v175. Toleranța și respectul pentru alte persoane	1	2
v176. Spiritul de economie	1	2
v177. Perseverența	1	2
v178. Credința religioasă	1	2

Este adevărat că întrebarea nu se referă la respondenți, ci la copii în general, însă putem asuma că dacă acele lucruri sunt importante pentru a fi transmise copiilor atunci reprezintă calități importante și pentru respondenții înșiși. Cei patru itemi selectați sunt în acord cu descrierea

dimensiunii de către Hofstede, așa încât i-am utilizat pe post de variabile proxy care au fost transformate în valori binare prin diferența din cifra 2 (pe un procedeu similar cu cel descris la dimensiunea masculinitate). Scorurile individuale au fost calculate, ca de obicei, prin însumarea celor patru itemi (obținându-se un scor cuprins între un minim de 0 pentru orientare pe termen scurt și un maxim de 4 pentru orientare pe termen lung), iar scorurile agregate au fost calculate la fel ca la ceilalți itemi, prin calcularea mediei răspunsurilor individuale pentru fiecare regiune socio-culturală în parte.

Singura dimensiune pentru care nu am găsit itemi în valul din 2008 este LUA - evitarea incertitudinii, întrucât întrebările care se refereau la aceasta au fost scoase din chestionar între momentul 1999 (valul 3) și 2008 (valul 4). Din acest motiv, am apelat la setul de date din valul 3, presupunând că scorurile agregate la nivel de regiuni nu s-au schimbat prea mult între cele două valuri. Există desigur posibilitatea ca unele schimbări sociale majore să aibă loc în orice societate, însă pe termen scurt (ex. 10 ani) este foarte puțin probabil ca o societate să se schimbe radical. Din acest motiv, utilizarea datelor din valul 3 poate fi considerată o soluție rezonabilă în condițiile în care nu există date în valul 4. Am selectat itemi de la următoarele întrebări: “În general, în viața de zi cu zi, preferați:”

s5.	1. ce este obișnuit	<i>sau</i>	2. ce este nou				
s6.	1. să planificați modul de cheltuire a banilor	<i>sau</i>	2. să îi cheltuiți fără nici un plan				
	“Credeți că ...”		În foarte mare măsură			În foarte mică măsură	
s8.	În viață omul e bine să se călăuzească după obișnuință		1	2	3	4	
o22.			1	2	3	... 8	9 10
	Oamenii ar trebui să fie mai precauți înainte de a face schimbări importante						Niciodată nu vei realiza ceva important în viață dacă nu acționezi cu îndrăzneală

Întrucât orientarea pe termen lung se găsește în partea stângă la toate aceste întrebări, toți itemii au fost inversați și transformați într-un interval comparabil de tip  $[0, 1]$ . Primii doi itemi s5 și s6 au fost transformați prin diferență din cifra 2, itemul s8 a fost adus la un minim de zero prin

scăderea unei unități, apoi împărțit la noul maxim de 3, iar pentru ultimul item o22 a fost aplicat un procedeu similar cu aducerea la un minim de zero prin scăderea unei unități, apoi transformat în intervalul  $[0, 1]$  prin împărțirea la noul maxim de 9.

În acest fel, itemi măsurați pe scale de răspun diferite (1...2, 1...4 și 1...10) au fost transformați la un set de valori comparabile, în acest fel fiind posibilă însumarea acestora pentru calcularea scorului individual (cu un minim de 0 și un maxim de 4), iar în final calcularea scorului agregat pe regiuni socio-culturale prin derivarea mediei răspunsurilor indivizilor.

Am lăsat la final variabila de interes SOCIV - implicarea în societatea civilă, măsurată cu ajutorul variabilelor referitoare la apartenența în diverse organizații (servicii sociale, educaționale, drepturile omului etc.), sub formă de voluntariat, utilizându-se întrebarea:

“Q5. V-aș ruga în continuare să vă uitați atent la următoarea listă de organizații și activități voluntare și să-mi spuneți (...) pentru care faceți muncă voluntară?”

Au fost utilizați toți itemii de la v28 până la v42 (servicii sociale, educaționale, drepturile omului etc.), cu excepția a doi itemi care se referă la organizațiile politice (D. sindicate, E. grupuri sau partide politice). Toți itemii selectați au fost transformați în valori binare prin diferență din cifra 2, iar apoi s-a utilizat procedura obișnuită de calcul al scorului sumativ individual și de agregare prin media acestor scoruri la nivel de regiune socio-culturală. Scorurile agregate finale sunt prezentate în tabelul 5.21

Acest tabel prezintă scorurile brute, obținute în urma procedurilor descrise mai sus. Datele de intrare în analiza calitativ comparativă însă trebuie să se afle în intervalul  $[0, 1]$  pentru toate condițiile cauzale PDI, MASC, IND, LTO, LUA și deasemenea pentru fenomenul de interes studiat SOCIV.

În acest scop, se utilizează calibrarea acestor condiții cauzale. Pentru toate acestea, se va utiliza o funcție crescătoare logistică pentru care este nevoie de stabilirea celor trei praguri: cel de excludere completă din mulțime, cel de mijloc și cel de includere completă în mulțime. Ideală ar fi utilizarea unui corp teoretic care să ghideze stabilirea acestor praguri, însă, pe de o parte, dimensiunile lui Hofstede nu au fost studiate în detaliu pentru regiunile socio-istorice din România, iar pe de altă parte, scorurile brute obținute în tabelul 5.21 sunt obținute într-un mod relativ ad-hoc, pe baza analizei secundare și a disponibilității variabilelor dintr-o bază de date deja existentă.

Tabelul 5.21: Scorurile medii agregate pentru cele 18 regiuni socio-culturale

	PDI	MASC	IND	LUA	LTO	SOCIV
R <sub>1</sub>	0.30	2.54	16.81	2.39	3.20	0.26
R <sub>2</sub>	0.13	2.60	19.05	2.30	2.77	0.09
R <sub>3</sub>	0.33	2.43	18.53	2.31	3.50	0.56
R <sub>4</sub>	0.63	1.91	17.28	2.38	2.72	0.19
R <sub>5</sub>	0.41	1.82	19.43	2.06	1.36	0.35
R <sub>6</sub>	0.28	2.55	20.23	2.38	3.36	0.09
R <sub>7</sub>	0.63	1.82	17.55	2.20	1.80	0.05
R <sub>8</sub>	0.37	1.96	17.93	2.44	2.15	0.43
R <sub>9</sub>	0.43	2.13	16.83	2.14	3.35	0.09
R <sub>10</sub>	0.22	2.15	19.83	2.46	3.21	0.03
R <sub>11</sub>	0.26	2.66	17.04	2.28	2.19	0.32
R <sub>12</sub>	0.59	2.47	20.23	2.22	2.82	0.10
R <sub>13</sub>	0.82	1.79	18.00	2.04	2.64	0.98
R <sub>14</sub>	0.69	1.52	19.33	2.36	2.74	0.45
R <sub>15</sub>	0.38	2.16	16.25	2.31	2.00	0.34
R <sub>16</sub>	0.29	2.30	20.17	2.33	2.15	0.49
R <sub>17</sub>	0.29	2.52	18.82	2.58	2.22	0.06
R <sub>18</sub>	0.40	2.27	18.40	2.35	3.02	0.20

Se poate spune însă cu certitudine faptul că scorurile pornesc de la un minim care înseamnă o prezență redusă a proprietății respective, până la un maxim care înseamnă o prezență ridicată a proprietății. În ce privește minimul absolut, pentru toate condițiile utilizate este întotdeauna 0. Maximul existent și cel absolut sunt diferite de la o condiție cauzală la alta, în funcție de metoda de calcul și operaționalizarea utilizată pentru fiecare în parte. Stabilirea pragurilor, în absența corpului teoretic, va fi realizată în funcție de datele existente.

Să luăm de exemplu masculinitatea (MASC), pentru care scorurile medii brute au un minim de 0 și un maxim de 3. Care este pragul pentru care o regiune este complet în afara acestei mulțimi, și care este pragul de la care putem spune că o regiune este complet înăuntrul mulțimii? Având în vedere distribuția scorurilor existente, am ales următoarele praguri: 1.8 pentru excludere completă, 2.2 pentru punctul de mijloc și 2.6 pentru pragul de includere completă în mulțime.

Tabelul 5.22: Pragurile de calibrare pentru fiecare condiție cauzală

	PDI	MASC	IND	LUA	LTO	SOCIV
I	0.8	2.6	20	2.5	3.3	0.50
M	0.4	2.2	18	2.3	2.7	0.25
E	0.1	1.8	16	2.1	1.4	0.05

Altfel spus, orice regiune cu o medie mai mică de 2.2 va fi considerată complet în afara mulțimii, iar orice regiune cu o medie mai mare decât 2.6 va fi considerată complet înăuntrul mulțimii. Toate pragurile de calibrare (E pentru excludere, M pentru mijloc și I pentru includere) sunt prezentate în tabelul 5.22, iar datele în mulțimi vagi, obținute în urma procedurii de calibrare sunt prezentate în tabelul 5.23, cu cele 18 regiuni socio-culturale aranjate (ca și la tabelul cu scorurile medii agregate) pe prima coloană.

Tabelul 5.23: Scorurile fuzzy obținute după calibrare

	PDI	MASC	IND	LUA	LTO	SOCIV
R <sub>1</sub>	0.27	0.93	0.15	0.80	0.92	0.54
R <sub>2</sub>	0.07	0.95	0.82	0.51	0.58	0.09
R <sub>3</sub>	0.34	0.85	0.69	0.55	0.98	0.98
R <sub>4</sub>	0.84	0.11	0.26	0.77	0.52	0.29
R <sub>5</sub>	0.52	0.06	0.89	0.03	0.05	0.77
R <sub>6</sub>	0.23	0.93	0.96	0.76	0.96	0.09
R <sub>7</sub>	0.85	0.06	0.34	0.19	0.12	0.05
R <sub>8</sub>	0.43	0.14	0.48	0.88	0.22	0.89
R <sub>9</sub>	0.55	0.37	0.15	0.09	0.96	0.09
R <sub>10</sub>	0.15	0.41	0.94	0.92	0.93	0.04
R <sub>11</sub>	0.21	0.97	0.20	0.42	0.24	0.70
R <sub>12</sub>	0.80	0.88	0.96	0.24	0.64	0.10
R <sub>13</sub>	0.96	0.05	0.50	0.02	0.47	1.00
R <sub>14</sub>	0.89	0.01	0.88	0.70	0.55	0.92
R <sub>15</sub>	0.45	0.42	0.07	0.53	0.17	0.75
R <sub>16</sub>	0.26	0.68	0.96	0.60	0.22	0.95
R <sub>17</sub>	0.25	0.92	0.77	0.98	0.25	0.06
R <sub>18</sub>	0.49	0.62	0.64	0.67	0.83	0.32

În legătură cu alegerea valorilor pentru aceste praguri se pot scrie, probabil, lucrări întregi. Nu se poate spune că aceste praguri sunt perfecte, după cum nici procesul de operaționalizare și alegerea indicatorilor nu este perfect. Scopul acestei secțiuni însă nu este acela de a oferi concluzii cu valoare științifică, ci doar de a demonstra aplicarea unei proceduri. Se atinge obiectivul de a obține pentru toate condițiile cauzale, valori fuzzy cuprinse între un minim de 0 și un maxim de 1.

Există o mică, dar importantă diferență între calibrarea prin atribuire directă și calibrarea prin utilizarea distribuției logistice. Utilizată aici pentru calibrarea dimensiunilor lui Hofstede, distribuția logistică presupune calcularea logaritmului șanselor de succes (în engleză *log odds*) asociat cu ancorele de excludere și includere completă ca diferență simetrică față de punctul de mijloc. Rezultatul este că toate cazurile care depășesc un scor de apartenență de 0.95 (log odds aproximativ egal cu 3) sunt considerate complet înăuntrul mulțimii și toate cele cu un scor de apartenență sub 0.05 (log odds aproximativ egal cu -3) sunt considerate complet în afara mulțimii.

După obținerea scorurilor calibrate, următorul pas este construirea tabelii de adevăr, pe modelul prezentat la secțiunea 5.1.5. Toate cazurile vor fi asociate unui colț al spațiului vectorial, iar pentru fiecare dintre acestea va fi calculat câte un scor de incluziune în mulțimea SOCIV, pe ipoteza suficienței în producerea acesteia. Toate scorurile rezultate sunt prezentate în tabelul 5.24

Tabelul 5.24: Tabela de adevăr pentru dimensiunile lui Hofstede

	PDI	MASC	IND	LUA	LTO	OUT	n	incl	cazuri
3	0	0	0	1	0	1	2	0.926	R <sub>8,15</sub>
8	0	0	1	1	1	0	1	0.718	R <sub>10</sub>
9	0	1	0	0	0	1	1	0.941	R <sub>11</sub>
12	0	1	0	1	1	0	1	0.815	R <sub>1</sub>
15	0	1	1	1	0	0	2	0.614	R <sub>16,17</sub>
16	0	1	1	1	1	0	4	0.543	R <sub>2,3,6,18</sub>
17	1	0	0	0	0	0	1	0.769	R <sub>7</sub>
18	1	0	0	0	1	0	1	0.788	R <sub>9</sub>
20	1	0	0	1	1	0	1	0.823	R <sub>4</sub>
21	1	0	1	0	0	0	2	0.884	R <sub>5,13</sub>
24	1	0	1	1	1	0	1	0.893	R <sub>14</sub>
30	1	1	1	0	1	0	1	0.721	R <sub>12</sub>

Există  $2^5 = 32$  combinații cauzale posibile pentru cele cinci condiții cauzale utilizate. Pentru economie de spațiu, tabelul nu prezintă decât acele combinații cauzale pentru care a putut fi calculat un scor de incluziune. Restul lor reprezintă așa numitele combinații rămase, colțuri ale spațiului vectorial fără niciun caz în apropiere care să ofere evidențe empirice dacă fenomenul se poate produce sau nu.

Cele 18 regiuni socio-culturale sunt prezentate pe ultima coloană “cazuri”, de exemplu combinația cauzală (00010) este asociată cu două regiuni, 8 și 15. Primele două combinații cauzale (00000 și 00001) nu au asociate niciun caz, de aceea tabelul începe cu prima combinație pentru care există dovezi empirice, adică linia 3 în tabela de adevăr completă (cu toate cele 32 de combinații posibile).

Ultima linie din acest tabel este a 30a combinație cauzală din tabela de adevăr completă (11101), în timp ce ultimele două combinații cauzale posibile (11110 și 11111) nu au niciun caz empiric asociat.

Pentru toate combinațiile empiric observate a fost calculat câte un scor de includere, iar coloana “OUT” a fost codificată în funcție de un prag de includere stabilit aici la valoarea de 0.9 (pentru ca o combinație cauzală să fie declarată pozitivă, trebuie să aibă un scor de incluziune cel puțin egal cu 0.9). Toate combinațiile pentru care scorul de incluziune este mai mic de 0.9 au fost codificate cu valoarea 0 pentru coloana OUT(come).

După cum se poate vedea, doar două combinații cauzale au suficientă forță empirică pentru a fi asociate cu prezența fenomenului de interes (unde OUT este egal cu 1). Acestea combinații, scrise cu ajutorul obișnuitei notații cu litere mici pentru absență și litere mari pentru prezență, sunt:

(00010) pdi·masc·ind·LUA·lto  
 (01000) pdi·MASC·ind·lua·lto

Algoritmul Quine McCluskey de minimizare booleană, prezentat în secțiunea 5.1.6, presupune compararea tuturor perechilor posibile de combinații cauzale pozitive și minimizarea acelor care se deosebesc printr-o singură condiție cauzală. În acest caz, sunt doar două combinații cauzale observate, pozitive, care se deosebesc prin două condiții cauzale (MASC și LUA), așadar ele nu mai pot fi minimizate mai departe.

Aceste combinații minime reprezintă ceea ce în QCA se numesc “soluții complexe”. Prin includerea combinațiilor rămase în algoritmul de minimizare booleană, se obțin ceea ce se numesc “soluții restrânse” (în engl. *parsimonious solutions*). În cazul de față, adăugarea celor 20 de combinații rămase la cele două combinații empiric observate conduce la

Tabelul 5.25: Implicanții primari

	a. rezultați din minimizare					b. acoperă combinațiile pozitive	
	PDI	MASC	IND	LUA	LTO	(00010)	(01000)
17	–	–	0	1	0	×	
35	–	0	–	1	0	×	
59	–	1	–	0	0		×
65	–	1	0	–	0		×
67	–	1	0	0	–		×
85	0	–	–	0	–		×
92	0	–	0	–	0	×	×
110	0	0	–	–	0	×	
118	0	0	0	–	–		×

un număr de 22 de combinații inițiale care intră în procesul de minimizare booleană.

În prima iterație, din compararea celor  $C_{20}^2 = 190$  de perechi posibile de combinații inițiale rezultă un număr de 38 de implicații reduși.

În a doua iterație, din compararea celor  $C_{38}^2 = 703$  de perechi posibile de implicații rezultă un număr de alți 19 implicații reduși mai mult.

În a treia și ultima iterație, din analiza celor  $C_{19}^2 = 171$  de perechi posibile de implicații rezultă un număr de 14 implicații primari (în engleză *prime implicants*), adică implicații care nu pot fi reduși mai mult întrucât toate perechile de implicații primari diferă prin mai mult de o condiție cauzală.

Dintre aceștia, doar 9 acoperă cele două combinații inițiale empiric observate, pentru care coloana rezultat OUT a fost codificată cu cifra 1 pentru prezență, iar cei 9 implicații primari valizi sunt prezentați în tabelul 5.25.a. Există un număr total de  $3^5 = 243$  de linii în matricea tuturor implicațiilor, cei 9 implicații primari au numerele cuprinse între 17 și 118. De exemplu, primul implicant primar rezultat după procesul de minimizare se regăsește pe a 17a linie din matricea tuturor implicațiilor.

În etapa finală de obținere a soluției, se rezolvă ceea ce se numește tabelul implicațiilor primari (în engleză *PI chart*). În tabelul 5.25.b este așezat semnul × pe fiecare celulă unde implicantul primar acoperă combinația observată respectivă (este o supramulțime a acesteia).

De exemplu, implicantul primar cu numărul 17 de pe prima linie este o supramulțime a combinației observate (00010), însă nu este o



supramulțime și a celeilalte combinații inițiale observate, din cauză că a doua combinație observată nu arată prezența LUA în penultima poziție.

Din cei 9 implicați primari prezenți în tabel, doar unul singur acoperă ambele combinații inițiale. Sigur, există mai multe perechi de implicați primari care luați împreună (prin reuniune) acoperă ambele combinații, însă principiul minimizării absolute primează, scopul final în QCA fiind acela de a găsi combinațiile absolut minime care explică (acoperă) toate combinațiile pozitive empiric observate.

Un singur implicant primar supraviețuiește acestui proces dur, acesta fiind campionul absolut al procesului de minimizare booleană. Este cel de pe linia cu numărul 92 din matricea tuturor implicanților (0–0–0), iar soluția finală este:

$$\text{pdi}\cdot\text{ind}\cdot\text{lto}$$

Conform datelor din EVS, implicarea în societatea civilă (destul de slabă în România, comparat cu alte țări europene) este explicată de o combinație simultană între o distanță mică față de putere, absența individualismului și absența orientării pe termen lung.

Cunoscând faptul că Hofstede a introdus dimensiunea LTO doar pentru a diferenția între cultura europeană și cea asiatică, am putea spune că în cazul României (o țară europeană), implicarea în societatea civilă ar putea fi explicată simplu doar de combinația simultană dintre o distanță mică față de putere și absența individualismului (pdi·ind).

În cuvinte simple, implicarea în societatea civilă este mai mare în acele regiuni în care oamenii simt că pot schimba ceva (distanță mică față de putere) și care nu sunt individualiști, ci orientați către valori comune.

Interesant este și faptul că, dacă am scoate complet din analiză dimensiunea LTO, soluția finală în urma procesului de minimizare booleană se modifică în:

$$\text{pdi}\cdot\text{ind}\cdot\text{masc}$$

Se menține combinația dintre distanța mică față de putere și absența individualismului (este o relație care apare în mod consistent odată cu implicarea în societatea civilă), de această dată întrând în combinație și absența masculinității.

Aceste două caracteristici (pdi·ind) combinate fie cu orientarea pe termen scurt dacă se analizează toate dimensiunile lui Hofstede, sau cu absența masculinității (dacă se elimină dimensiunea LTO), conduc la un portret destul de realist al explicației cu privire la implicarea în societatea civilă.

Această demonstrație trebuie privită ca atare, întrucât analiza secundară a unor date create pentru alte scopuri de cercetare permite mai greu identificarea unor itemi care să corespundă teoriei lui Hofstede. Probabil că itemii utilizați nici nu se suprapun perfect peste operaționalizarea conceptelor lui Hofstede.

Pe de altă parte, Hofstede utilizează dimensiuni bipolare, un lucru nepotrivit din punctul de vedere al teoriei mulțimilor întrucât este puțin probabil ca dimensiunile respective să fie perfect bipolare. De exemplu, este ușor de imaginat o persoană care să fie în același timp și masculină și feminină. Acest fapt este întru totul coerent cu teoria mulțimilor vagi, a căror limite nu sunt bipolare, ci există câte un scor de apartenență de la 0 (complet în afară) la 1 (complet înăuntru) pentru fiecare mulțime în parte. Una și aceeași persoană (indiferent dacă vorbim de un bărbat sau o femeie) poate să aibă câte un scor pentru ambele mulțimi, atât a persoanelor caracterizate de masculinitate cât și a persoanelor caracterizate de feminitate.

Exemplul prezentat trebuie privit cu scepticismul unui proces imperfect de operaționalizare, cu implicații directe asupra concluziilor și soluției finale. Este corectă sublinierea acestei limite și caracterizarea exemplului ca o încercare limitată în mare măsură de disponibilitatea itemilor în proiectul EVS care nu a fost conceput în mod special pentru măsurarea dimensiunilor lui Hofstede. O cercetare dedicată, cu o operaționalizare special concepută, ar putea să genereze rezultate diferite, într-o măsură mai mare sau mai mică.

# Glosar

- abducție, 64, 66
- algebră booleană, 102, 147, 162
- algoritmi, 23, 120, 143
- Anglia, 55
- asumpții, 23, 27, 29, 34, 52, 59, 70, 100, 136, 197
  - simplificatoare, 196
- baze de date, 26
- cadre analitice, 47, 68, 69, 72
- calibrare, 167, 169, 171–173, 187
- canoane inductive, 60, 137, 146
- capital, 40, 42, 84
- cauzalitate, 46, 47, 51, 53, 59
  - aparentă, 61
  - multiplă, 61
- cazuri
  - constituire, 30
  - lipsă, 196
  - negative, 32, 83
  - nonconforme, 33
  - selecție, 19, 21, 46
  - uniforme, 31
- centralizare, 42, 43
- China, 72, 73, 91
- comunism, 6, 15, 26, 37, 42, 73, 94, 95
- concepte, 3
  - dense, 35
  - măsurare, 36
  - operaționalizare, 13, 15
  - vagi, 162
- condiții cauzale, 55
  - concordante, 60, 140
  - contradictorii, 87, 140
  - contrafactice, 55, 196
  - macrosociale, 85
  - necesare, 54
  - prezență/absență, 137, 143, 147
  - suficiente, 56
- conjuncție cauzală, 132–134, 136, 180
- corelație, 31, 53, 55, 63, 64, 71, 75, 78, 121, 122, 146, 155, 178, 180, 190
- cuantificare, 12, 24, 36
- cultură, 16, 35, 40, 49, 78, 79, 90, 197
  - difuziune, 80
- cutia neagră, 17, 54, 75
- deducție, 64, 66, 68, 69, 81
- democrație, 43, 159, 163
- determinism, 33, 83, 85
- diagrame Venn, 114–118, 125, 130, 155, 158
- disjuncție cauzală, 132, 134, 136, 137, 180
- diversitate limitată, 20, 86–88, 192, 193, 196
- eroare, 17, 19, 33, 37, 50, 52, 53, 75
- eterogenitate, 35, 90–92
- etnie, 163, 164
- experiment, 59, 61, 70
- eșantion, 9, 11, 13, 15, 17–21, 41, 52, 63, 69, 75, 79, 96, 169

- fidelitate, 36, 96  
 Franța, 31, 55, 72, 73  
 Galton, 17, 78  
 Germania, 26, 55, 163, 198  
 indicatori, 3, 15, 35, 36, 51, 76, 78, 79, 159, 168  
 individualism, 40, 41, 72, 198  
 inducție, 65, 81  
     analitică, 67  
     neo-analitică, 69  
 industrializare, 41  
 inferență, 55, 146  
 Japonia, 16, 40, 41, 44, 72, 90  
 longitudinal, 53, 72, 77, 81  
 matematică, 8, 16, 29, 101, 109, 114, 162, 167  
     binară, 102  
 migrație, 25, 67  
 minimizare booleană, 140  
 modernitate, 40  
 mulțimi, 173  
     distincte, 114, 159  
     intersecție, 112, 116, 141, 164  
     negație, 113, 163  
     reuniune, 111, 165  
     vagi, 158, 159, 162  
 neutralitate axiologică, 39  
 PIB, 24, 37, 57  
 planificare, 42, 44  
 plauzibilitate, 37  
 Polonia, 55, 72  
 post-modernism, 39  
 pozitivism, 38, 39  
 probabilitate, 13, 33, 37, 54, 57, 64, 83–85, 110  
     condiționată, 53, 129  
 QCA, 101, 119, 146  
 mvQCA, 150  
 regresie, 15, 27, 33, 50, 75, 79  
 replicabilitate, 12, 16, 23, 24, 46, 146, 160  
 reprezentativitate, 15, 17, 18, 41, 52, 74, 96  
 Republica Moldova, 91  
 revoluție, 6, 21, 31, 56, 72, 73, 86, 94  
 România, 2, 24, 26, 37, 63, 94, 149, 159, 163  
 Rusia, 72, 73  
 spațiu de proprietăți, 82, 181  
 spațiu vectorial, 183–188, 191, 192, 206  
 Statele Unite, 6, 18, 21, 40, 72, 78  
 statistică, 25, 29, 76, 82  
 stratificare, 72, 91, 196  
 structuralism, 39, 91, 99  
 tabela de adevăr, 137, 142, 148, 151, 181–183, 191, 195, 207  
 teorie universală, 11  
 tip-ideal, 186  
 trasarea proceselor, 81  
 Turcia, 43, 44  
 validitate, 36, 74, 78, 96, 97, 168  
     aparentă, 97  
 variabile, 6, 7, 12, 13, 16, 27, 31, 34, 36, 41, 55, 70, 75, 76, 78, 80, 82, 96, 101, 108, 109, 121, 122, 140, 155  
     dependente, 14, 16, 32–34, 50, 53, 58, 193  
     dummy, 20, 151  
     independente, 13, 20, 22, 27, 33, 50, 52, 53, 58, 136, 193  
 variație, 13, 50, 64, 75

# Bibliografie

- Aickin, M. (2004). Millsian causation. *Quality & Quantity* 38, 533–545.
- Aiton, E. J. (1985). *Leibniz: a biography*. UK: Hilger.
- Bartolini, S. (1993). On time and comparative research. *Journal of Theoretical Politics* 5(2), 131–167.
- Berg-Schlosser, D. (2002). Macro-quantitative vs macro-qualitative methods in the social sciences - testing empirical theories of democracy. COMPASS Working Paper.
- Boole, G. (1854). *An Investigation of the Laws of Thought*. Walton & Maberly. Disponibilă pe adresa: <http://www.gutenberg.org/files/15115/15115-pdf.pdf>, în august 2014.
- Borkowski, L. (ed.) (1970). *Jan Łukasiewicz: Selected Works*. Amsterdam-North Holland.
- Chiot, D. și C. Ragin (1975). The market, tradition and peasant rebellion: the case of Romania in 1907. *American Sociological Review* 40(4), 428–444.
- Coppedge, M. (1999). Thickening thin concepts and theories: combining large N and small in comparative politics. *Comparative Politics* 31(4), 465–476.
- Cronqvist, L. (2003). Using multi-value logic synthesis in social science. Articol pregătit pentru cea de a 2a Conferință generală a ECPR - European Consortium For Political Research, Marburg, septembrie 2003.
- Cronqvist, L. și D. Berg-Schlosser (2009). Multi-Value QCA (mvQCA). În *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*, Volume 51 of *Applied social research methods*, pp. 69–86. SAGE Publications Ltd.

- Dauben, J. W. (1979). *Georg Cantor. His Mathematics and Philosophy of the Infinite*. New Jersey: Princeton University Press.
- De Meur, G. și B. Rihoux (2002). *L'Analyse Quali-Quantitative Comparée. Approche, techniques et applications en science humaines*. Louvain-La-Neuve: Academia-Bruylant.
- Dion, D. (1998). Evidence and inference in the comparative case study. *Comparative Politics* 30(2), 127–145.
- Dogan, M. (1994). Limits to quantification in comparative politics: the gap between substance and method. În M. Dogan și A. Kazancigil (ed.), *Comparing nations: concepts, strategies, substance*, pp. 35–71. Oxford UK și Cambridge SUA: Blackwell.
- Dogan, M. și D. Pelassy (1990). *How to compare nations. Strategies in comparative politics*. New Jersey: Chatham House Publishers, Inc. (prima ediție tipărită în 1984).
- Durkheim, E. (1993). *Despre sinucidere*. Iasi: Institutul European. (originalul tipărit în 1897 cu titlul “Le Suicide: Étude de sociologie”, ed. Félix Alcan, Paris).
- Durkheim, E. (1995). *Formele elementare ale vieții religioase*. Iasi: Polirom.
- Durkheim, E. (2002). *Regulile metodei sociologice*. București: Ed. Științifică, Polirom. (originalul tipărit în 1999 cu titlul Les Règles de la Méthode Sociologique, ediția a X-a, Presses Universitaires de France, Paris).
- Dușa, A. (2010). A mathematical approach to the boolean minimization problem. *Quality & Quantity* 44, 99–113.
- Ebbinghaus, B. (2005). When less is more. Selection problems in large-N and small-N cross-national comparisons. *International Sociology* 20(2), 133–152.
- Eff, E. A. (2001). Does Mr. Galton still have a problem?: Autocorrelation in the Standard Cross-Cultural Sample. Prezentare făcută la întâlnirea anuală a Southern Anthropological Association, Nashville Tennessee, 5-8 aprilie.
- Elman, C. (2005). Explanatory typologies in qualitative studies of international politics. *International Organization* 59, 293–326.
- Faure, A. M. (1994). Some methodological problems in comparative politics. *Journal of Theoretical Politics* 6, 307–322.

- Fuhrmann, G. (1998). Fuzziness of concepts and concepts of fuzziness. *Synthese* 75, 349–372.
- Galilei, G. (1953[1632]). *Dialogue Concerning the Two Chief World Systems, Ptolemaic and Copernican*. University Of California Press. Traducere de Stillman Drake.
- Goertz, G. și H. Starr (2003). *Necessary Conditions: Theory, Methodology, and Applications*. Lanham, Md., Boulder: Rowman & Littlefield.
- Goldstone, J. A. (1997). Methodological issues in comparative macrosociology. *Comparative Social Research* 16, 107–120.
- Goldthorpe, J. H. (1997a). Current issues in comparative macrosociology: a debate on methodological issues. *Comparative Social Research* 16, 1–26.
- Goldthorpe, J. H. (1997b). A response to the commentaries. *Comparative Social Research* 16, 121–132.
- Harman, G. (1965). The inference to the best explanation. *The Philosophical Review* 74(1), 88–95.
- Hicks, A. (1994). Qualitative Comparative Analysis and analytical induction. *Sociological Methods & Research* 23(1), 86–113.
- Hofstede, G. H. (1997). *Cultures and organizations. Software of the mind - Intercultural cooperation and its importance for survival*. McGraw-Hill.
- Hume, D. (1748). *An enquiry concerning human understanding*. Text disponibil pe adresa web [http://infidels.org/library/historical/david\\_hume/human\\_understanding.html](http://infidels.org/library/historical/david_hume/human_understanding.html) în septembrie 2014.
- Kazancigil, A. (1994). High stateness in a muslim society: the case of Turkey. În M. Dogan și A. Kazancigil (ed.), *Comparing nations: concepts, strategies, substance*, pp. 213–238. Oxford UK și Cambridge SUA: Blackwell.
- King, G.; R. Keohane și S. Verba (2000). *Fundamentele cercetării sociale*. Iași: Ed. Polirom. (originalul tipărit în anul 1994 cu titlul “Designing Social Inquiry. Scientific Inference in Qualitative Research”. Princeton University Press).

- King, G.; R. O. Keohane și S. Verba (1994). *Designing social inquiry: scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Lazarsfeld, P. (1937). Some remarks on typological procedures in social research. *Zeitschrift für sozialforschung* 6, 119–139. Text disponibil pe adresa web <http://ia600608.us.archive.org/32/items/ZeitschriftFrSozialforschung6.Jg/ZeitschriftFrSozialforschung61937.pdf> în martie 2014.
- Lieberson, S. (1991). Small N's and big conclusions: an examination of the reasoning in comparative studies based on a small number of cases. *Social Forces* 70(2), 307–320.
- Lieberson, S. (1994). More on the uneasy case for using Mill-type methods in small-N comparative studies. *Social Forces* 72(4), 1225–1237.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review* LXV(3), 682–693.
- Lipset, S. M. (1994). American exceptionalism - Japanese uniqueness. În M. Dogan și A. Kazancigil (ed.), *Comparing nations: concepts, strategies, substance*, pp. 153–212. Oxford UK și Cambridge SUA: Blackwell.
- Mahoney, J. (2000). Strategies of causal inference in small-N analysis. *Sociological Methods & Research* 28(4), 387–424.
- Mahoney, J. (2004). Comparative-historical methodology. *Annual Review Of Sociology* 30, 81–101.
- Mahoney, J. și G. Goertz (2004). The possibility principle: choosing negative cases in comparative research. *American Political Science Review* 98(4), 653–669.
- Mahoney, J. și G. Goertz (2006). A tale of two cultures: contrasting quantitative and qualitative research. *Political Analysis* 14(3), 227–249.
- McCluskey, E. J. (1956). Minimization of boolean functions. *The Bell System Technical Journal* 5, 1417–1444.
- Mill, J. S. (1843). *A System of Logic* (a 8a ed.). London: Longmans, Freen, Reader, and Dyer, 1872.
- Mill, J. S. (1970). Two methods of comparison. În A. Etzioni și F. DuBow (ed.), *Comparative perspectives: theories and methods*, pp. 205–212.



- Boston: Little Brown. Text selectat din 'A system of logic' de John Stuart Mill (New York: Harper & Row, Publishers, 1888), pp. 278-283 și 610-612.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63(2), 81–97.
- Moore, B. (1966). *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of Modern World*. Boston: Beacon Press.
- Murdock, G. P.; și D. R. White (2006). Standard Cross-Cultural Sample: on-line edition. Disponibil pe adresa web <http://repositories.cdlib.org/imbs/socdyn/wp/> în martie 2014.  
Originalul publicat în revista *Ethnology* 9, pp.329-369 în anul 1969.
- Naoi, A. și C. Schooler (1985). Occupational conditions and psychological functioning in Japan. *The American Journal of Sociology* 90(4), 729–752.
- Newman, K. S. (2002). Qualitative research on the frontline of controversy. *Sociological Methods & Research* 31(2), 123–130.
- Nichols, E. (1986). Skocpol on revolution: comparative analysis vs. historical conjuncture. *Comparative Social Research* 9, 163–186.
- Platt, J. (1992). Cases of cases... of cases. În C. C. Ragin și H. S. Becker (ed.), *What is a case? Exploring the foundations of social inquiry*, pp. 21–52. New York and Cambridge: Cambridge University Press.
- Popper, K. (2000). *Filosofie socială și filosofia științei*. Iași: Editura Trei. Traducere după Miller, D. (ed.) "Popper Selections", Princeton University Press, 1985.
- Pralong, S. (2008). Cultural influences in the development of civil society. The case of Romania. (teză de doctorat nepublicată).
- Przeworski, A. și H. Teune (1970). *The logic of comparative social inquiry*. New York: Wiley.
- Qnwuegbuzie, A. J. și N. L. Leech (2005). Taking the "Q" out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity* 39, 267–296.
- Quine, W. V. O. (1952). The problem of simplifying truth functions. *The American Mathematical Monthly* 59(8), 521–531.

- Quine, W. V. O. (1955). A way to simplify truth functions. *The American Mathematical Monthly* 62(9), 627–631.
- Ragin, C. (1987). *The comparative method. Moving beyond qualitative and quantitative strategies*. Berkeley, Los Angeles & London: University Of California Press.
- Ragin, C. (1994). *Constructing social research: the unity and diversity of method*. Thousand Oaks: Pine Forge Press.
- Ragin, C. (1997). Turning the tables: how case-oriented research challenges variable-oriented research. *Comparative Social Research* 16, 27–42.
- Ragin, C. (2000). *Fuzzy-set social science*. Chicago: The University of Chicago Press.
- Ragin, C. (2003). Making comparative analysis count. COMPASSS Working Paper.
- Ragin, C. (2005). From fuzzy sets to crisp truth tables. COMPASSS Working Paper.
- Ragin, C. (2008a). Fuzzy sets: calibration versus measurement. În J. Box-Steffensmeier, H. E. Brady, și D. Collier (ed.), *The Oxford Handbook of Political Methodology*, pp. 87–121. Oxford: Oxford University Press.
- Ragin, C. (2008b). *Redesigning Social Inquiry. Fuzzy Sets And Beyond*. Chicago and London: University of Chicago Press.
- Ragin, C. și H. S. Becker (1992). *What is a case? Exploring the foundations of social inquiry*. New York and Cambridge: Cambridge University Press.
- Rihoux, B. (2003). Bridging the gap between the qualitative and quantitative worlds? A retrospective and prospective view on qualitative comparative analysis. *Field Methods* 15(4), 351–365.
- Rihoux, B. și C. Ragin (ed.) (2009). *Configurational Comparative Methods*, Volume 51 of *Applied social research methods*. Thousand Oaks, California: SAGE Publications Ltd.
- Rueschmeyer, D. și J. D. Stephens (1997). Comparing historical sequences - a powerful tool for causal analysis. *Comparative Social Research* 16, 55–72. A reply to John Goldthorpe's 'Current issues in comparative macrosociology'.

- Rughiniș, C. (2003). Mize și strategii ale cercetării comparative. *Sociologie Românească* 1(1-2), 129–143.
- Ruskey, F. și M. Weston (2005). A survey of venn diagrams. *Electronic Journal of Combinatorics*. Disponibil pe adresa web <http://www.combinatorics.org/files/Surveys/ds5/VennEJC.html> în martie 2014.
- Sandu, D. (1998). Eșantionare și analiză regională: aplicație pentru barometrul 1998, resurse socio-umane ale reformei. Fundația pentru o Societate Deschisă, București.
- Sandu, D. (2011). Construcția și validarea tipologiilor sociale: analize în spațiile de atribute prin tabele de contingență, “elaborare” și regresie. Disponibil online pe adresa web [http://doctorat.sas.unibuc.ro/wp-content/uploads/2011/11/Sandu\\_Tipurisociale\\_nov2011.pdf](http://doctorat.sas.unibuc.ro/wp-content/uploads/2011/11/Sandu_Tipurisociale_nov2011.pdf) în martie 2014.
- Savolainen, J. (1994). The rationality of drawing big conclusions based on small samples: in defence of Mill’s methods. *Social Forces* 72(4), 1217–1224.
- Schmitt, C. (1996). *The concept of the political* (a 3a ed.). Chicago: University of Chicago Press. Ediția princeps publicată în 1927.
- Schneider, C. Q. și C. Wagemann (2012). *Set-theoretic methods for the social sciences*. Cambridge: Cambridge University Press.
- Shannon, C. E. (1938). A symbolic analysis of relay and switching circuits. Teză de master, Massachusetts Institute of Technology, Dept. of Electrical Engineering. Disponibilă pe adresa: <http://theses.mit.edu> în octombrie 2005.
- Skocpol, T. (1979). *States and social revolutions. A comparative analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
- Skocpol, T. și M. Somers (1980). The uses of comparative history in macrosocial inquiry. *Comparative Studies in Society and History* 22(2), 174–197.
- Slomczynski, K. M.; J. Miller și M. L. Kohn (1981). Stratification, work, and values: A Polish-United States comparison. *American Sociological Review* 46(6), 720–744.
- Smelser, N. J. (1976). *Comparative methods in the social sciences*. Englewood Cliffs, N.J.: Prentice Hall, Inc.

- Stahl, H. H. (1939). *Nerej, un village d'une région archaïque: monographie sociologique*. Bucarest: Institut de sciences sociales de Roumanie.
- Taylor, S. J. și R. Bogdan (1998). *Introduction to qualitative research methods* (a 3a ed.). New York: John Wiley.
- Teune, H. (1997). Stories, observations, systems, theories. *Comparative Social Research* 16, 73–84.
- Thiem, A. și A. Dușa (2013). *Qualitative Comparative Analysis with R. A user's guide*. Springer briefs in political science. Springer.
- Tilly, C. (1997). Micro, macro or megrim? Articol disponibil pe adresa: <http://www.asu.edu/clas/polisci/cqrm/papers/Tilly/> în octombrie 2005.
- Weber, M. (1978). *Economy and society: an outline of interpretive sociology*. Berkeley ; Los Angeles ; London: University Of California Press.
- Weber, M. (2003). *Etica protestantă și spiritul capitalismului*. București: Humanitas. Ediția originală apărută în 1934, cu titlul 'Die protestantische ethik und der geist des kapitalismus'. Tübingen: Verlag von J.C.B. Mohr.
- White, D. R. (2004). Using SPSS: Analysis and comparison in the social sciences. Disponibil la adresa: <http://eclectic.ss.uci.edu/~drwhite/xc/book.htm> în august 2007.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control* 8, 338–353.
- Zamfir, C. (2005[1990]). *Spre o paradigmă a gândirii sociologice* (a 2a ed.). Iași: Polirom.
- Zamfir, C. și E. Zamfir (ed.) (1993). *Țigani: între ignorare și îngrijorare*. București: Alternative.