

## CONCEPTUL DE REPREZENTATIVITATE ÎN CERCETAREA STATISTICĂ PE BAZĂ DE SONDAJ

DĂNUȚ-VASILE JEMNA\*

### Le concept de représentativité dans la recherche statistique par sondage

#### Resumé

La représentativité statistique a enregistré une évolution conceptuelle dans le même sens avec le développement des problèmes théoriques et pratiques du sondage statistique. La représentativité statistique d'un sondage a été définie en principal sur deux positions: de façon empirique, intuitive, suggérée par le mode de manifestation de la réalité observée et de façon scientifique, fondamentée par la théorie des probabilités et de la statistique mathématique.

Ces deux possibilités de définir le concept de représentativité ont une correspondance en deux méthodes de sondage spécifique: le sondage empirique ou par choix raisonné (méthode des quotas, méthode des unités-types etc.) et le sondage aléatoire.

**Mot cles:** représentativité, sondage, empirique, aléatoire, précision

#### Introducere

În elaborarea unei cercetări pe bază de sondaj statistic, una dintre cele mai importante probleme teoretice și metodologice este reprezentativitatea eșantionului. În esență, conceptul de reprezentativitate statistică pune următoarea problemă: în ce condiții partea, eșantionul, poate reprezenta, înlocui întregul, populația totală? La această problemă s-au dat diferite soluții, conceptul de reprezentativitate suportând o evoluție semantică în strânsă legătură cu dezvoltarea problemelor teoretice și practice ale cercetării statistice.

Reprezentativitatea a fost definită în principal de pe două poziții. O primă abordare este de tip empiric, intuitiv, conceptul fiind definit în acord cu *modul* de manifestare a realității ca întreg, pe baza informațiilor disponibile despre acest întreg. Cea de-a doua abordare a conceptului este de tip teoretic, pe baza unui set de *principii și legi* care permit definirea și extragerea unui eșantion din populație. Această poziție s-a dezvoltat grație aparatului oferit de teoria probabilităților și de statistica matematică.

---

\*Lector doctor la Universitatea „Alexandru Ioan Cuza” Iași, Facultatea de Economie și Administrarea Afacerilor

În lucrare, se realizează o analiză a principalelor semnificații ale conceptului de reprezentativitate statistică, punctând elementele specifice care au permis elaborarea acestora, împreună cu problematica pe care o ridică fundamentarea teoretică și posibilitățile de aplicare ale conceptului în cercetarea statistică pe bază de sondaj.

### **Elaborările empirice, intuitive ale conceptului de reprezentativitate**

Primele tentative de a înlocui cunoașterea pe baza cercetărilor exhaustive cu ajutorul rezultatelor parțiale datează încă din secolul al XIV-lea, prin încercările de estimare a numărului populației pornind de la conceptul de "foc", care face referire la o locuință, familie sau gospodărie [J.J. Droesbeke, P. Tassi, 1990, p. 41]. Din această perioadă datează primele înregistrări ale fenomenelor demografice în acte de stare civilă, activitate care asigură principala sursă de date statistice. Rezultatele cele mai semnificative însă apar în secolul al XVII-lea, odată cu formarea *Școlii Aritmeticii Politice Engleze*. Prin folosirea *tehnicii multiplicatorului*, W. Petty (1623-1687) și mai apoi Laplace (1749-1827), ajung la rezultate spectaculoase în estimarea a numeroși indicatori demografici și economici.

Aceste rezultate însă vor cunoaște o dezvoltare de amploare și pe baze științifice riguroase abia la începutul secolului al XX-lea, într-o perioadă determinată de factori politici, economici și științifici favorabili. Până atunci însă, cercetările prin sondaj se realizează într-un cadru restrâns și pe baze empirice, intuitive. Două modalități de definire a reprezentativității sau de justificare a cercetărilor prin sondaj rețin atenția în perioada secolelor XVIII-XIX: reprezentativ înseamnă tipic și reprezentativ înseamnă a păstra structura întregului la nivelul părții sau a eșantionului.

#### **a. Reprezentativ - tipic**

În cea mai mare parte a anchetelor parțiale și a monografiilor secolelor XVIII-XIX, reprezentativ înseamnă tipic [A. Desrosieres, 2000, p. 266]. În această situație, reprezentativitatea este asigurată prin găsirea uneia sau a mai multor unități tip din populație. Unitățile tip reprezintă acele unități care au valorile variabilelor studiate foarte aproape de medie. Justificarea acestei abordări este una intuitivă. Dacă întregul este descris cu ajutorul mărimilor medii, atunci e suficient să se studieze acele unități din populație ale căror caracteristici se apropie de medie.

Unitățile tip sunt alese cu grijă pe baza datelor obținute în anchetele exhaustive. Din datele statistice oficiale, se obțin informații cu privire la populația de referință, se obțin mărimile medii pentru caracteristicile de interes, iar pe baza acestor informații se identifică unitățile care vor fi supuse studiului monografic, în profunzime. Specificul anchetei monografice este calitatea, studiul complex, multilateral al unităților tip alese din populație.

Critica adusă acestui mod de a defini reprezentativitatea constă în principal în aceea că prin analiza unităților tip nu se ține cont de diversitatea populației, de structura acesteia după diferite criterii [A. Desrosieres, 2000, p. 279].

#### **b. Reprezentativ – miniatură a întregului**

Condiția de reprezentativitate impune să se obțină un eșantion care să respecte structura populației de referință și deci diversitatea existentă la nivelul acesteia. Când această structură este cunoscută, eșantionarea se poate realiza utilizând această

informație, iar eșantionul poate apare ca o miniatură a populației totale. Aceasta este viziunea despre reprezentativitate promovată de statisticianul norvegian N.A. Kiaer.

La Congresul Institutului Internațional de Statistică din 1895 de la Berna, Kiaer readuce în lumea statisticii internaționale discuția asupra problemei sondajului [N.A. Kiaer, 1895, p.p. 176-178]. Prin Kiaer, problema reprezentativității devine un filon de cercetare important pentru statisticieni, miezul problemei fiind acesta: cum să obții un număr de unități dintr-un întreg astfel încât acesta să (re)prezinte în miniatură, cât mai fidel, colectivitatea totală.

Criticând viziunea holistă a anchetelor din secolele XVII-XIX, Kiaer consideră că o anchetă reprezentativă trebuie să ia în considerare nu doar cazurile tipice, ci varietatea care se găsește într-un anumit spațiu social, propunând ca în anchetă să se realizeze o miniatura a aceluși spațiu sau țări. Organizarea de anchete reprezentative în Norvegia (mobilul practic) a fost motivată de interesul statului cu privire la politicile de protecție socială și cu privire la necesitatea creării de organisme în acest sens.

Prin Kiaer, se realizează o trecere de la statistica întregului, caracterizată prin medie (Quetelet), la statistica analizei distribuției indivizilor, o statistică bazată pe varianță, corelație, comparație. Kiaer afirmă că totalul nu este dat de cazul tip, de medie, ci de toată varietatea cazurilor care se întâlnesc în viață. Abordarea lui Kiaer este mai mult intuitivă și empirică. El este un bun cunoscător al muncii de teren, a realizării practice a anchetelor, fiind mai puțin preocupat de formalizarea matematică și teoretică a metodei sale. În ciuda acestei lipse de preocupări pentru formalizare, discuțiile deschise de Kiaer sunt fundamentale, de pionierat. Kiaer nu a pus problema modului de extragere a unui eșantion (aleatoare sau de alt tip), ci a sugerat modul în care se poate verifica reprezentativitatea, după metoda variabilelor de control.

La Kiaer, verificarea reprezentativității cercetărilor parțiale se realizează prin intermediul variabilelor de control, variabile care au aceleași valori la nivelul părții cât și la nivelul ansamblului. Rezultatele obținute pentru aceste variabile sunt comparate cu cele obținute prin recensământ. Dacă abaterile nu sunt semnificative, atunci rezultatele anchetei prin sondaj sunt considerate valide.

Puterea cognitivă a cercetării prin sondaj propusă de Kiaer se poate evidenția în comparație cu metodele clasice: recensământul și monografia. Față de recensământ, sondajul se impune prin două elemente specifice: posibilitățile de cunoaștere și precizia. Dacă prin recensământ realitatea este analizată la suprafață, metoda reprezentativă permite un studiu în profunzime al realității. În primul caz, informațiile obținute sunt descriptive, pe când în al doilea caz permit explicarea realității supusă cercetării. În plus, metoda reprezentativă este mai precisă, rezultatele fiind afectate de mai puține erori. Atât sondajul, cât și monografia analizează în profunzime realitatea, însă monografia se rezumă la cazurile tipice, pe când sondajul studiază întreaga diversitate a populației.

Valoarea științifică a cercetării prin sondaj depinde de caracterul reprezentativ al eșantionului și mai puțin de volumul acestuia. Deseori se întâmplă că datele disponibile sau practic ușor de obținut nu se referă decât la o parte a societății. A realiza o cercetare doar pe baza acestor date înseamnă a comite o eroare de reprezentativitate. Ori, o eroare de acest tip au comis-o americanii în anchetele de opinie prin mass-media, mizând pe un volum mare de date, dar fără a cunoaște gradul lor de reprezentativitate [D.V. Jemna, 2005, p.52].

### Elaborări științifice ale conceptului de reprezentativitate

Cu toate realizările lui Kiaer, metoda reprezentativă va pătrunde cu greu, atât în teoria, cât și în practica statistică. O primă cauză este de ordin teoretic: metoda lui Kiaer rămâne una intuitivă și empirică, nu are un fundament științific, matematic. Deschiderea se realizează câțiva ani mai târziu prin utilizarea distribuțiilor de probabilitate la extragerea unui eșantion și la compararea rezultatelor de sondaj cu cele de la nivelul întregii populații. O altă cauză a demarajului greoi al metodei sondajului statistic este dată de reținerea comunității științifice cu privire la sondaj, sub impactul și influența administrației și statisticii oficiale, bine înrădăcinate în tradiția recensămintelor.

După anul 1925, când are loc Congresul Institutului Internațional de Statistică de la Roma, se apreciază că *metoda reprezentativității* este validă, iar problemele care se deschid țin de modalitățile de obținere a unui eșantion din populația totală. Pe baza rezultatelor obținute de specialiști în prima parte a secolului al XX-lea, conceptul de reprezentativitate a fost definit în sens restrâns în legătură cu procedeele aleatoare sau mixte de extragere a unui eșantion. În acest cadru, se remarcă următoarele corelații conceptuale: reprezentativ-aleator, reprezentativ-precis, reprezentativ-stratificat.

#### a. Reprezentativ - aleator

Abordarea, în plan teoretic, a relației eșantion-populație, cu ajutorul instrumentelor oferite de teoria selecției, determină o altă accepțiune asupra reprezentativității. În acest caz, reprezentativ înseamnă să extragi eșantionul după modelul unei experiențe aleatoare, după o schemă probabilistă, caz în care fiecare unitate din populație are aceeași șansă de a intra în eșantion.

#### b. Reprezentativ - precis

Dezvoltările din teoria probabilităților și statistica matematică au permis determinarea unui tip de erori care însoțesc cercetarea prin sondaj statistic, și anume erorile de reprezentativitate. Aceste erori reprezintă diferențe între valorile tipice de sondaj calculate la nivelul unui eșantion și parametrii populației din care a fost extras eșantionul [E. Jaba, 2002, p.287]. Cu cât erorile de reprezentativitate sunt mai mari, cu atât este afectată reprezentativitatea rezultatelor.

În acest caz, reprezentativitatea statistică înseamnă obținerea celor mai bune estimări sau cu erorile cele mai mici. Un eșantion este considerat reprezentativ dacă pentru o variabilă dată estimează valoarea adevărată sau parametrul cu o eroare minimă, adică cu eroare zero [P. Ardilly, 1994, p. 388]).

#### c. Reprezentativ - stratificat

Fundamentarea reprezentativității statistice a suportat o nouă clarificare prin combinarea celor două posibilități: cea empirică și cea teoretică. Utilizând informațiile despre existența unor categorii diferite de unități în populația totală, pe de o parte, și fundamentul matematic al extragerii aleatoare, pe de altă parte, J. Neyman elaborează metoda sondajului stratificat. Autorul arată superioritatea unui sondaj aleator față de unul raționat, nealeator, precum și posibilitatea de a ameliora rezultatele printr-o stratificare a priori a populației totale [J. Neyman, 1934, p. 558-606].

Obținerea reprezentativității în cazul unui sondaj stratificat reclamă două exigențe:

- construirea de nomenclatoare, clase care să asigure o stratificare a priori a populației totale după variabile care asigură diferențe semnificative între indivizii din clase diferite și diferențe cât mai mici între indivizii din aceeași clasă;
- construirea schemelor de extragere aleatoare din aceste clase a indivizilor și a instrumentelor care să verifice gradul de reprezentativitate a rezultatelor (intervalele de încredere).

În cazul sondajului stratificat, există mai multe posibilități de construire a eșantionului: din fiecare strat să se aloce același număr de unități, un număr proporțional de unități cu ponderea stratului din total sau un număr de unități care să țină cont nu numai de pondere, ci și de gradul de omogenitate a fiecărui strat, numită alocare optimală. Eșantionul obținut prin stratificare, cu o alocare proporțională a unităților din fiecare strat, este numit eșantion reprezentativ, iar alocarea optimală mai poartă și numele de alocare Neyman [J. Desabie, 1966, p. 144].

### Concluzie

Conceptul de eșantion reprezentativ se referă la proprietatea fundamentală a unui eșantion de a permite descrierea și explicarea populației din care a fost extras. Aceasta este condiția de bază a valabilității științifice a unei colectivități de selecție [Al. Bărbat, 1971, p. 436]. Celor două modalități de abordare a conceptului de reprezentativitate le corespund două metode de cercetări prin sondaj distincte: metode raționate sau empirice (sondajul pe cote, metoda unităților tip etc.) și metode aleatoare. Prin combinarea celor două metode fundamentale, a rezultat un tip de sondaj mixt, care utilizează avantajele ambelor metode de sondaj (cazul sondajului stratificat, de exemplu).

Definirea reprezentativității unui eșantion se poate face atât în sens restrâns, cât și în sens larg.

- În sens restrâns, un eșantion este reprezentativ dacă fiecare unitate din populație are posibilitatea de a intra în eșantion cu o probabilitate cunoscută. Cu alte cuvinte, este reprezentativ eșantionul care este extras după o metodă aleatoare.
- În sens larg, un eșantion este reprezentativ dacă permite descrierea și explicarea populației de referință cu erori cât mai mici. Potrivit acestei accepțiuni, reprezentativitatea poate fi extinsă și la eșantioane obținute după metode de sondaj nealeatoare și mixte.

Problema reprezentativității eșantionului este esențială într-un sondaj statistic, însă rămâne o problemă sensibilă și dificil de rezolvat.

În practică, de regulă, reprezentativitatea unui eșantion se definește în raport cu o singură variabilă, cu o variabilă fundamentală pentru tema de cercetare. Problema practică cea mai dificilă este că nimic nu garantează faptul că eșantionul este reprezentativ și în raport cu o altă variabilă, mai ales în cazul în care prin sondaj se cercetează o problemă complexă. O soluție, pentru asigurarea reprezentativității eșantionului, este să se organizeze cercetări pe teme bine conturate, iar variabila prin care se definește reprezentativitatea să fie puternic corelată cu variabilele care apar în studiu. În cazul problemelor complexe se recomandă descompunerea problemei pe componente care pot face obiectul unor cercetări separate.

În fundamentarea reprezentativității unui eșantion, pe lângă problemele de ordin practic, esențiale sunt: definirea populației de referință sau a întregului, definirea unităților individuale de selecție și de observare, definirea caracteristicilor populației după care se realizează cercetarea, alegerea metodelor de extragere a eșantionului etc.

### **Bibliografie**

1. Ardilly, P. - *Les techniques de sondage*, Editions Technip, Paris, 1994
2. Bărbat, Al. - *Teoria statisticii sociale*, Universitatea "Al.I. Cuza" Iași, 1971
3. Desrosieres, A. - *La politique des grands nombre*, La Decouverte, Paris, 2000
4. Drosbeke, J.J., Tassi, P. - *Histoire de la statistique*, Presses Universitaires de France, Paris, 1990
5. Desabie, J. - *Theorie et pratique des sondages*, Dunod, Paris, 1966
6. Kiaer, N.A. - *Observations et experiences concernant des denombrements representatifs*, Buletin de l'IIS, vol. 9, 1895, pp. 176-178
7. Jaba, E. - *Statistica*, ediția a treia, Editura Economică, București, 2002
8. Jemna, D.V. - *Eficiența sondajului statistic*, Editura Sedcom Libris, Iași, 2005
9. Neyman, J. - *On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selectio*, Journal of the Royal Statistical Society, 97, p. 558-606, 1934