

On Political Opinion
Surveys

SONDAJELE DE OPINIE POLITICĂ

VOICU BOȘCAIU

The complex problem of the implementation of public opinion surveys is considered in the Romanian context and probabilistic framework.

The probabilistic definitions of survey's tools, the main sources of errors, the survey's precision and validation and also, the quality of management are shortly discussed.

Retrospectiva sondajelor naționale cu tematică electorală organizate în România în perioada 1992-1995 surprinde două aspecte anormale. Pe de-o parte, sondajele preelectorale din '92 (cu o singură excepție) au fost contrazise de alegeri și, pe de altă parte, sondajele din perioada '93-'95 au oferit adesea rezultate (aparent) contradictorii.

Să alegem, de exemplu, două sondaje politice similare, nu întâmplător botezate de presă "Un cadou de Anul Nou pentru putere" (decembrie '94) și respectiv "Cheia pe primul loc!" (februarie '95). Tehnici de rutină ale statisticii matematice ar permite compararea lor. S-ar putea decide astfel dacă diferențele între rezultatele celor două sondaje sunt întâmplătoare sau sunt statistic semnificative. Având în vedere că în cele șase săptămâni care despart sondajele nu au avut loc evenimente politice de natură să modifice în mod esențial opinia electoratului, eventuala prezență a diferențelor semnificative ar trebui pusă pe seama erorilor de sondaj.

O astfel de analiză ar fi acum tardivă: impactul unui sondaj asupra electoratului și clasei politice nu mai poate fi schimbat ulterior.

Pe de altă parte, publicitatea în jurul unor erori de sondaj este deocamdată riscantă: ar putea fi acreditată astfel opinia falsă că sondajele nu pot fi relevante în România (extrapolările abuzive sunt obișnuite la noi!). Totuși, o discuție generală privind credibilitatea sondajelor pare să fie necesară.

Câtă încredere putem avea într-un sondaj?

Executantul unui sondaj are obligația profesională să comunice, iar beneficiarul are posibilitatea să



ceară (de vreme ce plătește) toate informațiile privind tehnica de sondaj utilizată, metodologia implementării acestei tehnici, precizia rezultatelor obținute, precum și metoda de calcul a acestei precizii. Absența sau incoerența acestor informații transformă un sondaj dintr-un demers științific într-o tehnică artizanală discutabilă.

Prezentarea rezultatelor unui sondaj empiric drept concluzii fundamentate științific discreditează instituția care a efectuat sondajul. Responsabilitatea mediatizării acestor rezultate aparține însă în special celor care au comandat sondajul. Pentru consumatorul mass-media problema încrederii într-un sondaj este, deci, simplă: ea se reduce la încrederea în instituția care a solicitat sondajul.

Clasificarea sondajelor

Tehnicile de sondaj cu aplicabilitate practică pot fi clasificate astfel: tehnici de cotare, tehnici probabiliste și tehnici mixte.

Tehnicile de cotare impun alegerea (uni - sau multistadială) unor obiecte "reprezentative", în proporții (cote) bine determinate, caracteristice populației studiate.

Tehnicile probabiliste pretind selecția aleatoare în toate etapele generării eșantionului: (O selecție este aleatoare dacă probabilitatea apriorică de selectare este nenulă și cunoscută pentru toate entitățile populației. Entitățile ar putea fi indivizi sau grupuri de indivizi din populație.) Este suficient ca o singură etapă a generării să nu fie aleatoare, pentru ca sondajul să-și piardă caracterul probabilist.

Tehnicile mixte sunt acelea care utilizează succesiv proceduri de cotare, proceduri probabiliste și raționamente de tip "expert".

Determinarea preciziei indicilor estimați este imposibilă, sau, în cel mai bun caz, complicată și discutabilă pentru

sondajele neprobabiliste. Descrierea preciziei unui sondaj mixt prin precizia unui sondaj probabilist "asemănător" poate fi uneori utilă, dar nu are nici o justificare teoretică. Altfel spus, singurele tehnici de sondaj măsurabile sunt cele probabiliste.

Clasificarea sondajelor probabiliste

În funcție de procedurile probabiliste, folosite, un sondaj probabilist aparține uneia dintre următoarele patru categorii.

1) Sondajul probabilist simplu este caracterizat de două proprietăți: respondenții eșantionului au opinii independente și probabilitățile apriorice de a aparține eșantionului sunt egale pentru toți indivizii populației supusă sondajului.

2) Sondajul multistadial (cluster) se poate folosi în populații care se structurează în mod natural într-un număr mare de entități numite nuclee (de exemplu: locatarii unei clădiri, locuitorii unei localități). În acest caz, se selectează aleator un număr de nuclee și ulterior, din fiecare nucleu, un număr de obiecte. Obiectele pot fi indivizi sau nuclee de rang inferior, și, de asemenea, trebuie selectate aleator. Este deci vorba de o selecție cu cel puțin două trepte probabiliste.

3) Sondajul stratificat presupune partiționarea populației inițiale în straturi. Fiecare strat va fi tratat ulterior ca o populație distinctă (tehnica de sondaj ar putea fi diferită de la strat la strat).

4) Sondajul probabilist complex este un sondaj uni - sau multistratificat; în fiecare strat se execută un sondaj probabilist simplu sau un sondaj multistadial.

Pentru cazul României, un eșantion probabilist simplu este nerealizabil din cauza prețului prohibitiv. (De exemplu, un eșantion de 1.000 respondenți ar trebui să fie răspândit în aproximativ 500 de localități!) În schimb, sunt intensiv folosite tehnicile complexe de sondaj.

Despre precizia sondajelor probabiliste

Pentru un sondaj probabilist, determinarea preciziei indicilor estimați se poate face prin proceduri standard ale teoriei probabilităților. (Este vorba de determinarea intervalului de încredere a fiecărui indice studiat, pentru un risc dat.) Metodele de calcul sunt însă diverse, în funcție de tehnica probabilistă de sondaj folosită.

Singura tehnică de sondaj a cărei precizie poate fi calculată și tabelată a priori este sondajul probabilist simplu. Afirmatia binecunoscută "pentru un eșantion de 1.000 de indivizi eroarea maximă este 3%" se referă la acest tip de sondaj și un risc de 5%.

Precizia sondajelor probabiliste complexe depinde nemijlocit de frecvențele variantelor de răspuns în straturi și nuclee. În consecință, determinarea preciziei se poate face exclusiv a posteriori, separat, pentru fiecare variantă de răspuns.

În general, precizia sondajelor multistadiale este inferioară preciziei sondajelor probabiliste simple. Pierderea de precizie poate fi însă compensată prin stratificarea după criterii sociologice adecvate. Alegerea avantajoasă a acestor criterii mărește precizia, dar pretinde o analiză sociologică profesionistă anterioară sondajului (eventual, dublată de analiză cluster). Aplicarea corectă a procedurilor de stratificare cere stăpânirea tehnicilor de ponderare și de calcul al preciziei sondajelor ponderate.

Practica curentă de-a comunica pentru toți indicii calculați ai unui sondaj o valoare globală a erorii maxime reduce relevanța procentelor mici. (De exemplu, o eroare de 3% pentru un partid care a obținut în eșantionul de sondaj 5% voturi, "garantează" la alegeri un rezultat cuprins între 2% și 8%, ceea ce nu spune mare lucru.) Calculând eroarea pentru fiecare procent în parte, precizia crește.

Aproximarea preciziei unor

sondaje probabiliste complexe folosind tabelele sau formulele elaborate pentru sondaje probabiliste simple s-a făcut frecvent, dar este o eroare: nu are nici justificare probabilistă și face imposibilă analiză completă a sondajului.

Despre erori

Într-un sondaj de opinie, sursele de eroare ar putea fi trei:

1) tehnica de sondaj defectuoasă (erori de eșantionare, legate de structurarea necorespunzătoare a eșantionului; greșeli ale operatorilor de sondaj; calcul eronat al preciziei);

2) imprecizia inerentă oricărei metode (corectă sau nu) de măsurare (un sondaj este și el o metodă de măsurare!);

3) instabilitatea, incoerența opțiunilor electoratului.

Eventualele erori ale tehnicii de sondaj afectează în special procentele mici. Concret, în prezența erorilor de eșantionare sau a greșelilor de operare, interpretarea procentelor mici este imposibilă.

Din punctul de vedere al specialistului, profesionalismul unui sondaj înseamnă capacitatea de-a lua în considerare și de-a elimina sau diminua toate sursele previzibile de eroare. Un sondaj poate fi compromis și de decizii care - cel puțin în aparență - sunt corecte, decizii care păreau să fi fost validate de o experiență anterioară relativ îndelungată. Literatura de specialitate consemnează alături de nenumărate sondaje electorale de mare acuratețe și eșecuri răsunătoare "atribuite" unor instituții a căror profesionalism nu poate fi pus la îndoială. Executantul unui sondaj trebuie să știe că un sondaj preelectoral poate fi compromis de neglijențe minore.

Despre analiza complexă a sondajului

Găsirea eventualelor erori de eșantionare nu este o problemă simplă. Nu există o listă exhaustivă a erorilor posibile



(există erori imprezibile!). Nu există criterii suficiente pentru a verifica corectitudinea unui sondaj. Este deci evident faptul că interpretarea răspunsurilor unui sondaj nu poate fi redusă la aritmetica elementară a calculării și comparării unor medii și frecvențe. Analiza unui sondaj este un demers complex, care apelează un șir de proceduri cu funcționalitate bine determinată:

- 1) validarea sondajului;
- 2) ameliorarea (corectarea) sondajului;
- 3) estimarea parametrilor și preciziei lor de calcul;
- 4) interpretarea rezultatelor.

Obiectivele validării sunt descoperirea (recunoașterea), și elaborarea metodologiei de corectare a eventualelor erori de eșantionare. Validarea unui sondaj începe cu studierea fișelor de lucru ale operatorilor de sondaj, continuă cu studiul reprezentativității eșantionului și se încheie cu verificarea unor ipoteze sociologice de control.

Despre eșantioane reprezentative

Prin definiție, un eșantion este reprezentativ în raport cu o variabilă de control specificată, dacă repartițiile variabilei de control în eșantion și în populație sunt identice. Trebuie menționat însă faptul că reprezentativitatea nu este suficientă pentru validarea unui sondaj. (Un eșantion poate fi reprezentativ în raport cu un set de variabile de control fără a fi reprezentativ în raport cu variabilele a căror studiu face obiectul sondajului!) Pentru cazul sondajelor probabiliste, reprezentativitatea nu este nici măcar necesară pentru validare. Raportarea la un eșantion reprezentativ oferă, însă, o metodă elementară și intuitivă de abordare a calcului ponderilor, ponderi necesare estimării parametrilor și preciziei lor de calcul.

Despre operatorul de sondaj

Operatorul de sondaj este persoana care intră în contact nemijlocit cu respondenții, pune întrebările și înregistrează răspunsurile.

Pentru un sondaj probabilist, selectare respondenților trebuie să se facă aleator, operatorului revenindu-i exclusiv sarcina de-a identifica (nu de-a alege!) respondentul, aplicând un algoritm de randomizare impus. Algoritmul trebuie definit astfel încât să nu-i lase operatorului nici un grad de libertate.

Un operator se poate aștepta să fie "constrâns de șansă" (obligat de algoritm) să caute un anume respondent într-o zonă nefamiliară și greu accesibilă, să nu-l găsească sau să constate că respondentul este necooperant, agresiv, neîncrezător sau incapabil să răspundă. Este deci lesne de înțeles tendința operatorului de-a cenzura algoritmul pentru a evita situațiile neplăcute și pierderea "inutilă" de timp. Tentația operatorului de-a se substitui algoritmului crește odată cu complexitatea chestionarului de sondaj.

Cenzura operatorului poate compromite caracterul probabilist al sondajului și implicit, corectitudinea parametrilor calculați. Situația operatorului este deci paradoxală: el trebuie să acționeze împotriva interesului propriu. (Are oare întotdeauna operatorul de sondaj o motivație suficient de puternică pentru asta?)

Operatorul ar putea, de asemenea, influența involuntar opiniile unor respondenți prin comportamentul și vestimentația proprie, persoanele care (eventual) îl însoțesc, un ziar sau o carte pe care ar putea s-o păstreze la vedere. Mecanismul agresiunii operatorului asupra respondentului este complex. Pentru al înțelege, operatorul trebuie să fie familiarizat cu unele noțiuni elementare de psihologia sondajului de opinie.

Despre corectarea erorilor de eșantionare

În mod obișnuit, pentru corectarea erorilor unui sondaj este suficientă utilizarea tehnicilor uzuale ale teoriei sondajului probabilist (de exemplu, metoda post-stratificării). În această fază, folosirea pentru estimarea parametrilor a tehnicilor de ponderare este inevitabilă.

Corectarea erorilor unui sondaj, deși posibilă în principiu, este uneori complicată. Iată un exemplu. Să presupunem că într-un sondaj prin poștă, procentul non-răspunsurilor este de 30%. Nu există nici un temei pentru presupunerea că nonrespondenții și respondenții ar avea opțiuni similare. Interpretarea sondajului exclusiv pe baza răspunsurilor primite este deci inacceptabilă. Cel puțin o parte a nonrespondenților va trebui contactată direct în vederea interviului. Dacă evidența nonrespondenților nu a fost păstrată, validarea sondajului este imposibilă sau, în cel mai bun caz, dacă este disponibilă o teorie care descrie nonrespondenții, sondajul poate fi validat, dar precizia sa este mult diminuată. Un raționament asemănător se poate face și pentru cazul unor sondaje telefonice, relativ la subpopulația celor fără telefon.

Despre particularitățile sondajelor electorale

Obiectivul unui sondaj de opinie cu tematică electorală este determinarea intențiilor de vot ale electoratului, în ansamblul său. În consecință, efortul (timpul consumat, dificultatea alegerii) necesar pentru a răspunde chestionatului unui sondaj electoral trebuie să fie similar cu cel necesar completării unui buletin de vot. Să presupunem, de exemplu, că întrebările cu tematică electorală au fost incluse într-un sondaj complex, cu multe întrebări. Un astfel de sondaj înregistrează exclusiv opțiunile celor care au timpul, bunăvoința și capacitatea intelectuală de a

răspunde sondajului complex. Această subpopulație nu poate descrie electoratul, în ansamblul său.

Pe de altă parte, există întrebări care prin simpla lor formulare pot influența opțiunile unor alegători. Dacă un chestionar conține astfel de întrebări (pe care nu orice alegător le-ar formula pentru sine înainte de a vota), atunci ele nu trebuie să le precedă pe cele cu caracter electoral. Ordinea întrebărilor unui sondaj nu este, deci, lipsită de importanță.

Includerea unui chestionar pre-electoral într-un sondaj de opinie cu tematică largă comportă riscuri a căror asumare nu poate fi justificată întotdeauna. (Obiectivele unui sondaj electoral sunt exclusiv cantitative, în timp ce obiectivele unei anchete sociologice complexe sunt preponderent calitative!) O soluție ar putea fi întreruperea interviului după epuizarea primei părți a chestionarului, cea corespunzătoare întrebărilor preelectorale, dacă respondentul nu este în măsură să răspundă.

Informații despre respondenți

Fiecare chestionar trebuie să conțină toate informațiile necesare analizei complexe a sondajului, inclusiv cele relative la respondenți. Nu pot lipsi din baza de date codul respondentului, codul localității de domiciliu, sexul, vârsta, nivelul de instruire, ocupația, codurile nucleelor de selecție din care respondentul face parte, stratul respondentului. Localitatea trebuie identificată printr-un cod de circulație publică (de exemplu, codul poștal), care poate fi recunoscut de orice beneficiar. Se asigură astfel compatibilizarea sondajelor cu problematica comună (compararea unor sondaje nu pretinde identitatea chestionarelor!)

Despre know-how

Unii consideră că lista localităților de sondaj trebuie protejată de concurență. Este ridicol, pentru că: 1) tehnica generării



acestei liste este o problemă de rutină, la îndemâna oricărui specialist în statistică matematică; 2) "furtul" listei nu poate fi mascat și deci este echivalent cu asumarea incompetenței; 3) precizia sondajului nu poate fi calculată fără a ști cu exactitate cum a fost stratificată populația: aceste informații nu fac însă parte din baza de date. Altfel spus, "hoțul" nu deține informațiile necesare analizei complexe a sondajului, deci nu-și poate valifica "marfa". (Nimeni nu plătește un sondaj dacă nu poate obține certificatul său de calitate, dovada corectitudinii.)

Pe de altă parte, prezența în baza de date a localității de domiciliu a fiecărui respondent permite o analiză sociologică zonală foarte interesantă, dincolo de obiectivele imediate ale sondajului.

Despre omologarea unui sondaj

Omologarea unui sondaj este ceea ce în industrie se numește "controlul de recepție". Această procedură de control final se execută de către o comisie neutră multidisciplinară de specialiști, singura în măsură să elibereze sondajului un pașaport pentru mass-media. În principiu, procedura de omologare reia analiza complexă a sondajului, așa cum a fost ea descrisă mai sus.

Piese necesare omologării sondajului vor fi descrise în continuare.

1) Chestionarul sondajului. (Va trebui să suporte analiza sociologului și psihologului.)

2) Documentația care descrie modul de generare a eșantionului și algoritmul de identificare al respondenților.

3) Dosarul cu fișele de lucru ale operatorilor de sondaj. (Va trebui comparat cu textul de instruire a operatorilor de sondaj în vederea selectării respondenților.)

4) Baza de date, într-un format standard, etichetată sau însoțită de toate indicațiile necesare utilizării sale de orice persoană abilitată. (Prin contract, baza de date ar trebui să devină proprietatea benefi-

ciarului. Ponderile atașate fiecărui respondent ar putea fi considerate, eventual, know-how al executantului.)

5) Studiul reprezentativității eșantionului (o analiză statistică, logică și sociologică, pe baza unor variabile de control specifice respondenților).

6) Descrierea procedurii de determinare a ponderilor. Deși acest ultim material ar putea fi considerat confidențial, el trebuie pus la dispoziția comisiei de omologare. În cazul sondajelor proporționale, determinarea ponderilor se reduce la dimensionarea straturilor și nucleelor de sondaj.

7) Dosarul prelucrărilor statistice solicitate, preciziile indicilor calculați, precum și modul de calcul al acestor precizii.

Concluzie

Realizarea unui sondaj de opinie cu relevanță națională (proiectare, implementare, omologare) este un demers multidisciplinar de mare complexitate. Succesul depinde de munca și coordonarea unui număr mare de specialiști. Într-o măsură mai mare sau mai mică, fiecare poate induce erori. În etapa de realizare a interviurilor, soarta eșantionului se află în mâinile operatorilor de sondaj; erorile acestora sunt foarte greu de sesizat. În etapele de proiectare și de analiză complexă, este indispensabilă conlucrarea cvartetului celor patru profesioniști: informaticianul, matematicianul, psihologul și sociologul. (Situțiile în care unul dintre cei patru poate fi suplinit de ceilalți, fără a prejudicia nici calitatea, nici credibilitatea sondajului, sunt fericite excepții).

Absența de pe piață, performanțele și contraperformanțele instituțiilor specializate în domeniul sondajului statistic vor primi semnificații noi cu ocazia prognozelor preelectorale ale anului 1996. Anul electoral 1996 este acela care va stabili ierarhia valorilor pe piața românească a sondajelor.