

UTILIZAREA SONDAJULUI STATISTIC IN ECONOMIE

Delia, Teselios¹
Mihaela, Albici²

Rezumat: Sondajul statistic reprezintă o metodă eficientă de investigare statistică, care implică colectarea datelor cantitative, acesta fiind de multe ori preferat rapoartelor statistice datorită informațiilor care se pot obține despre întreaga colectivitate cercetată, prin observarea unei părți a acesteia. Datorită informațiilor oferite, sondajele sunt folosite în foarte multe domenii de cercetare. În economie, statistica este utilizată în procesul de luare a deciziilor, în alegerea unor strategii competitive, în analiza anumitor fenomene economice, în formularea de previziuni. Studiul economic prezentat în această lucrare are rolul de a ilustra modalitatea prin care un sondaj aleator simplu este folosit în analiza privind situația locurilor de parcare existente într-o anumită localitate.

Cuvinte cheie: sondaj, medie, deviație standard, scor standard.

Clasificarea JEL: C10, C13.

1. Introducere

Progresul din ultimii 20 de ani, când conceptul de „societate informațională” a fost acceptat treptat atât în mediul universitar, cât și în cel politic și economic, a contribuit la dezvoltarea și aprofundarea tehnicilor utilizate în statistica economică. Fenomenele care au produs schimbări radicale (încă în curs de desfășurare) în ceea ce privește creșterea cererii de informații statistice economice sunt [14]:

- transformările survenite în procesele de producție datorită noilor tehnologii de informații și comunicații;
- globalizarea sistemelor economice și sociale;
- conceptele de „economie a cunoașterii” și „economie digitală”.

Între sursele de obținere a datelor statistice, alături de procedee precum recensămintele, rapoartele statistice, sondajele statistice constituie o variantă aflată în plină dezvoltare. Datorită operativității, a economicității obținerii datelor, metoda sondajului se află printre procedeele preferate în obținerea datelor. Într-o economie de piață bazată pe proprietatea privată, sondajul este forma preponderentă de obținere a datelor statistice în dauna rapoartelor statistice – formă specifică economiilor centralizate [13].

Necesitatea obținerii de informații, cu maximă operativitate, a condus la o extindere a utilizării metodei sondajului statistic ca formă a observării parțiale. În unele situații practice sondajul statistic este singura metodă de obținere a informațiilor. Este vorba despre cazul în care cercetarea exhaustivă este fie nerațională, datorită caracterului distructiv al metodei de cercetare – de exemplu, în controlul calității produselor, fie neeconomică din cauza costurilor mari antrenate de cuprinderea tuturor unităților colectivității.

2. Cercetarea prin sondaj

Cercetarea selectivă presupune culegerea și prelucrarea datelor referitoare la o parte a colectivității generale, iar rezultatele obținute se extind asupra întregului obiect al cunoașterii pentru a obține o caracterizare satisfăcătoare. Este necesar ca partea observată (eșantion, mostră, probă, selecție) să fie reprezentativă, adică să reproducă la scară redusă trăsăturile esențiale ale întregii colectivități [1].

Sondajul este o formă a cercetării statistice realizată pe baza unei părți reprezentative din populația (colectivitatea generală) studiată.

¹ Lector univ. dr., Universitatea „Constantin Brâncoveanu” din Pitești, delia_teselios@yahoo.com.

² Lector univ. dr., Universitatea „Constantin Brâncoveanu” din Pitești, mturmacu@yahoo.com.

Toate afirmațiile, concluziile stabilite pe baza datelor provenite dintr-un sondaj nu sunt de tip determinist, ele având caracterul unor afirmații statistice, fiind deci efectuate în condițiile unei anumite probabilități, ale unui anumit nivel de încredere. Avantajul metodei sondajului constă în faptul că nivelul acestei probabilități și al erorilor este controlabil și posibil de estimat statistic.

Metoda sondajului oferă variante și tehnici de prelevare diferențiate și adaptate diferitelor tipuri de populații, astfel încât să se asigure caracterul aleator al prelevării unităților și, în final, reprezentativitatea eșantionului.

Într-o cercetare prin sondaj, confruntându-se două tipuri de colectivități – colectivitatea totală pe care vrem să o cunoaștem și eșantionul pe care îl înregistrăm –, se întâlnesc o serie de termeni perechi care au același conținut metodologic, dar diferă din punctul de vedere al informației cuprinse.

Colectivitatea totală mai este numită și populație mamă sau populație statistică de referință, populație originală, populație generală, și reprezintă ansamblul unităților componente ale unei colectivități care face obiectul cercetării statistice. Definirea populației (colectivității totale) se face prin enumerarea tuturor unităților componente. Ansamblul acestora formează volumul colectivității și se notează cu N .

Colectivitatea parțială extrasă din colectivitatea totală în scopul observării este cunoscută sub diferite denumiri, în funcție de domeniul de aplicabilitate al cercetării prin sondaj, și anume: selecție, eșantion, mostră, probă, populație observată. Eșantionul reprezintă un subansamblu de unități extrase dintr-o colectivitate pe care dorim să o cunoaștem.

Definirea eșantionului se face prin enumerarea unităților componente ale subansamblului extras și observat din populația de volum N . Volumul eșantionului se notează cu n . Volumul eșantionului, din motive de cost în principal, este mult mai mic decât volumul populației din care este extras, adică în cazul observării parțiale are loc relația: $n < N$. Dacă $n = N$, este cazul unei observări exhaustive.

Variabila observată poate fi *cantitativă* (venitul indivizilor, numărul angajaților unei firme) sau *calitativă*. Valorile unei caracteristici X sunt x_i .

Definirea populației statistice cere, pe de o parte, delimitarea statistică a populației, iar, pe de altă parte, verificarea gradului de omogenitate a populației.

Verificarea gradului de omogenitate a populației constă în analiza indicatorilor variației caracteristicilor de definire a acesteia, operație ce se poate efectua pe baza datelor oferite de o observare totală anterioară sau a datelor rezultate din mai multe sondaje organizate succesiv. Verificarea omogenității populației este o condiție necesară alegerii corecte a procedurilor de eșantionare.

Rezultatele obținute pe baza datelor de sondaj se extrapolează la dimensiunea întregii populații. Extinderea rezultatelor de la „parte” la „întreg” nu are însă caracter determinist, ci probabilist, deci, sunt supuse unui risc de a fi eronate [3]. Avantajul metodei sondajului constă nu atât în eliminarea erorilor, cât, mai ales, în predimensionarea lor și a „controlului” probabilității afirmațiilor. Avantajele cercetării prin sondaj pot fi privite din punctul de vedere al costului, al rapidității obținerii rezultatelor, al cunoașterii unor colectivități greu accesibile etc.

Cercetarea selectivă este mai ieftină, mai operativă și mai exactă pentru că se face pe un număr redus de elemente, iar rezultatele ei pot fi verificate dacă este necesar prin altă cercetare selectivă sau printr-o cercetare totală. Avantajul din punctul de vedere al costului derivă din faptul că într-o cercetare prin sondaj cheltuielile materiale, umane și de timp sunt mai reduse comparativ cu o observare totală. Referitor la cheltuielile umane, observându-se numai o parte din colectivitate – un eșantion reprezentativ, volumul de

muncă se reduce și poate fi acoperit de un număr mic de persoane, de regulă, un personal de specialitate, sporind prin aceasta și calitatea înregistrării [1].

3. Studiu economic

În vederea realizării unei analize privind situația reală a locurilor de parcare, primăria unei localități efectuează un sondaj aleator simplu, alegând un eșantion reprezentativ de 30 de parcări din totalul parcărilor amenajate existente în număr de 150. Datele rezultate sunt prezentate în tabelul 1.

Observație: Datele din tabelul 1 au doar rol ilustrativ.

Tabelul nr. 1. Datele obținute în urma sondajului

Parcare	Capacitate de parcare (locuri/parcare)	Parcare	Capacitate de parcare (locuri/parcare)
1	30	16	30
2	32	17	65
3	35	18	74
4	40	19	71
5	48	20	45
6	71	21	89
7	65	22	80
8	90	23	35
9	87	24	30
10	52	25	50
11	85	26	70
12	50	27	32
13	72	28	67
14	83	29	60
15	50	30	64

Sursa: Creat de autor

În prima fază, datele disponibile urmează să fie sistematizate.

Pentru realizarea acestui lucru, se notează cu x caracteristica de grupare (capacitatea de parcare) și se calculează amplitudinea variației caracteristicii.

$$A_x = x_{\max} - x_{\min} = 90 - 30 = 60 \text{ locuri} \quad (1)$$

Cu ajutorul formulei lui Sturges determinăm mai întâi numărul intervalelor de grupare (de variație) (notat M) și apoi mărimea intervalului de grupare (de variație) (notată d).

$$M = 1 + 3.22 * \log n, \quad n = \text{numărul măsurătorilor efectuate} \quad (2)$$

$$\Rightarrow M = 1 + 3.22 * \log 30 = 5.756 \approx 6$$

$$d = \frac{A_x}{M} = \frac{60}{6} = 10 \text{ locuri} \quad (3)$$

Între valorile x_{\max} și x_{\min} , ținând cont de valorile calculate M și d , se stabilesc intervalele de grupare.

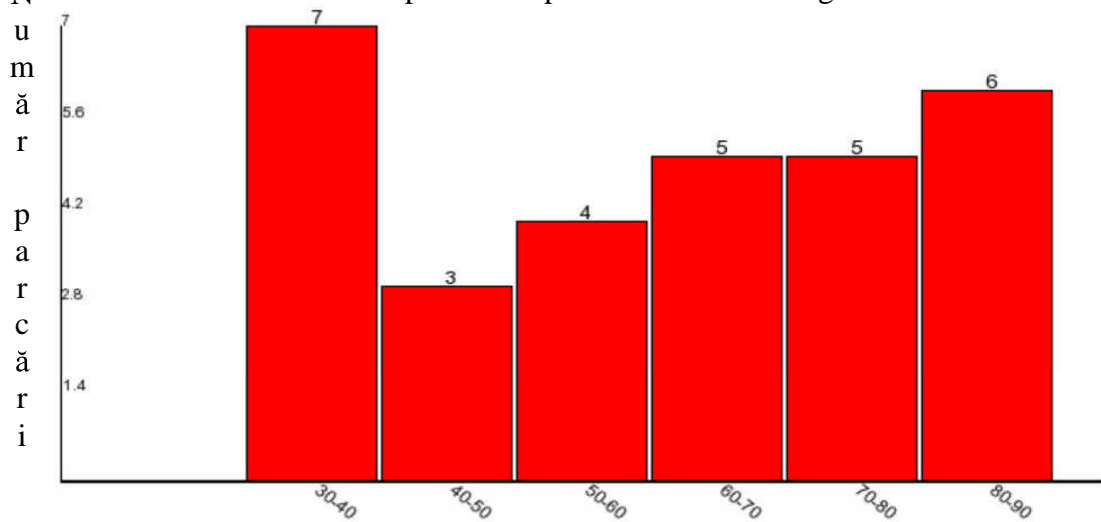
Centralizăm datele obținute în tabelul 2, unde considerăm intervale închise la stânga și deschise la dreapta.

Tabelul nr. 2. Datele sistematizate

Intervale de variație a capacității de parcare (locuri/parcare)	Număr de parcări amenajate
30-40	7
40-50	3
50-60	4
60-70	5
70-80	5
80-90	6

Sursa: Creat de autor

N Datele sistematizate sunt reprezentate prin următoarea histogramă.



Sursa: Creat de autor

Capacitate de parcare

Figura 1. Reprezentarea grafică a datelor sistematizate

Calculăm valoarea medie în eşantion, folosind formula

$$\bar{x} = \frac{\sum_{i=1}^6 x_i n_i}{\sum_{i=1}^6 n_i} \quad (4)$$

și stabilim dacă aceasta este reprezentativă pentru colectivitatea de selecție analizată.

Tabelul nr. 3. Valori utilizate în formula mediei în eşantion

	x_i	n_i	$x_i * n_i$
	35	7	245
	45	3	135
	55	4	220
	65	5	325
	75	5	375
	85	6	510
Total	360	30	1810

Sursa: Calcule efectuate de autor

Conform calculelor prezentate în tabelul 3,

$$\bar{x} = \frac{1810}{30} = 60, (3) \text{ locuri/parcare} \quad (5)$$

Pentru a determina gradul de împrăștiere al datelor în jurul valorii centrale, calculăm deviația standard, măsurată aici în locuri/parcare.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^6 (x_i - \bar{x})^2 \cdot n_i}{\sum_{i=1}^6 n_i}} = 18.39 \text{ locuri/parcare, unde } \sigma^2 = \text{dispersia.} \quad (6)$$

Cu alte cuvinte, limita minimă a variației normale a datelor este 41.94, iar limita maximă 78.72. Aproximând, putem afirma că există între 42 și 79 locuri/parcare.

Coeficientul de variație este

$$v = \frac{\sigma}{\bar{x}} \cdot 100 \approx 30\% \quad (7)$$

Prin urmare, colectivitatea de selecție analizată este omogenă și media $\bar{x} \approx 60$ locuri/parcare este reprezentativă pentru colectivitate.

În măsura în care doar un eșantion din populație a fost analizat, se impune determinarea intervalului de încredere, folosind proprietățile legii normale (legea lui Gauss).

Pentru a calcula intervalul de încredere al mediei aplicăm formula:

$$\bar{x} - z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \quad (8)$$

unde μ = media reală și z_{α} = pragul de semnificație

Pentru un test simetric, utilizat în general pentru calculul marjei de eroare, valorile legii normale sunt diferite pentru praguri de semnificație diferite.

Tabelul nr. 4. Scoruri standard

	1%	5%	10%	20%
z_{α}	2.57	1.96	1.64	1.28

Sursa: Realizat de autor folosind calculatorul z-score disponibil la <http://www.measuringusability.com/pcalcz.php>

$$60.3 - 1.96 \cdot \frac{18.39}{\sqrt{30}} \leq \mu \leq 60.3 + 1.96 \cdot \frac{18.39}{\sqrt{30}}$$

$$\Leftrightarrow 53.71 \leq \mu \leq 66.88 \text{ locuri/parcare.} \quad (9)$$

4. Concluzii

Rezultatul obținut semnifică faptul că media reală a locurilor de parcare pe ansamblul parcărilor are 95% șanse să se afle în intervalul [54,67] locuri.

Dacă se dorește ca eroarea maximă admisibilă să fie redusă la trei sferturi (de la 6.58 la 4.93) iar rezultatele obținute să fie garantate cu o probabilitate de 99% ($z_{\alpha} = 2.57$), eșantionul selectat trebuie să aibă un volum mai mare.

Calculăm acest nou volum cu ajutorul relației:

$$n^* = \frac{(2.57)^2 \cdot \sigma^2}{(4.93)^2} = \frac{6.60 \cdot 338.22}{24.35} = 91.74 \approx 92 \text{ parcări} \quad (10)$$

Bibliografie

1. Albici M. (2010), *Statistică economică. Matematici aplicate în economie*, Editura Universitaria, Craiova
2. Biji E., Baron T. (1991), ș.a, *Statistică teoretică și economică*, Editura Didactică și Pedagogică, București
3. Isaic-Maniu A. (2001), *Tehnica sondajelor și anchetelor*, Editura Independența Economică, Pitești
4. Isaic-Maniu A., Korca M., Mitruț C., Voineagu V. (1998), *Statistică*, Editura Independența Economică, Brăila
5. Isaic-Maniu A., Mitruț C., Voineagu V., *Statistică generală*, disponibilă la <http://www.biblioteca-digitala.ase.ro/>
6. Popa, M., *Scoruri standard. Curba normala (Gauss)*, disponibilă la http://www.mpopa.ro/statistica_licenta/St1_05_z_gauss.pdf
7. Porojan D. (1993), *Statistică și teoria sondajului*, Editura Șansa
8. Rotariu T., Iluț P. (1997), *Ancheta sociologică și sondajul de opinie*, Editura Polirom, Iași
9. Voineagu, V., Mitrut, C., ș.a, *Statistică*, disponibilă la <http://www.biblioteca-digitala.ase.ro/>
10. <http://math.tutorvista.com/statistics/z-score-table.html#>
11. <https://statistics.laerd.com/statistical-guides/standard-score.php>
12. <http://www.measuringusability.com/pcalcz.php>
13. <http://simionmija.wordpress.com/2009/10/02/sondajul-statistic/>
14. <http://www.oecd.org/std/41746710.pdf>