

Introducere în analiza datelor economico-financiare

Aplicații și exemple rezolvate prin: R, Python, Excel, PowerBI

Ciprian Alexandru-Caragea

9/8/24

Cuprins

Lista figurilor	v
Lista tabelelor	vii
Despre autor	3
Introducere	5
Quarto	7
Informații despre software	7
1 Concepte de bază în statistică	9
1.1 Ce este statistica?	9
1.2 Începutul demersului statistic	10
1.3 Mărimi relative	16
Summary	25
Bibliografie	27

Lista figurilor

Lista tabelelor

1.1	Exemple de variabile calitative	12
1.2	Exemple de variabile cantitative discrete	14
1.3	Exemple de variabile cantitative continue	14
1.4	Producția industrială în anul t	18
1.5	Structura producției industriale în anul t	19
1.6	Distribuția salariaților după nivelul salariului brut, în anul t	19

Despre autor

Ciprian Alexandru-Caragea este conferențiar universitar la Facultatea de Management Financiar, Universitatea Ecologică din București și Analist de Date la diverse instituții internaționale.

Titlul de doctor în Economie l-a obținut sub egida Academiei Române, Institutul de Economie Națională. A participat la un program de studii postdoctorale în care a implementat utilizarea software-ului R ca instrument de analiză a evoluției indicilor bursieri.

Activitatea sa didactică se concentrează, în principal, în domeniul burselor de valori, prin cursuri și seminarii la programele de licență și masterat (Piețe de capital, Managementul Portofoliului, Piețe internaționale de capital).

A participat la diverse proiecte de cercetare, workshop-uri, conferințe naționale și internaționale. Activitatea de cercetare a fost pusă în valoare prin publicarea studiilor în reviste din țară și din Europa, precum și în baze de date internaționale recunoscute (RePEC, DOAJ, EBSCO).

În prezent, în cadrul Institutului Național de Statistică, participă ca expert în proiecte BigData și utilizează software-ul de analiză statistică R pentru Data cleaning, Data Matching, Web Scraping, analize de date și vizualizare, Data mining, Data integration, data processing, data validation, dar și utilizarea datelor din sursele administrative pentru realizarea de statistici oficiale.



ResearcherID: V-2168-2017

<https://orcid.org/0000-0001-8215-6671>

Introducere

“...acum nu mai e nimic nou de descoperit;
tot ce rămâne e doar măsurătoarea din ce în ce mai precisă”

— Lord Kelvin (1894)

Cartea tipărită merită răsfoită. Trăim în vremea în care internetul facilitează comunicarea globală, informația fiind disponibilă oricând și oricum. Toată lumea, de la oameni de știință și până la copii de vârstă școlară primesc și oferă informații și propagă idei pe calea internetului. Tirajele publicațiilor, cărților și manualelor tipărite sunt în scădere în întreaga lume, în timp ce postările online captează atenția omenirii.

Obiectivul principal al cărții pe care o propun este de a fi un ghid cuprinzător, în termeni de concepte și tehnici, reprezentativ și, mai ales, practic, în ceea ce privește utilizarea instrumentelor software de analiză statistică, R fiind principalul software utilizat pentru aplicațiile propuse. Ca abordare generală, cartea prezintă principalele concepte utilizate în statistică, cu exemple și explicații descriptive. Exemplele din viața economică - cele mai multe dintre ele bazate pe date statistice reale - problemele rezolvate, dar și cele propuse, acoperă o arie cuprinzătoare de tematici, cititorul având șansa de a fi introdus în sfera aplicativă a conceptelor teoretice parcurse.

Cartea este destinată tuturor celor care doresc să înțeleagă, prin mijloace științifice, fenomenele economice și sociale, sub aspectul măsurării cantitative și din perspectiva determinării cauzale. Deși se adresează, în principal, studenților care se pregătesc să devină specialiști în științele economice, lucrarea este utilă și celor care își propun să cunoască un domeniu atât de frumos și de captivant. Tocmai nevoia de informații, din ce în ce mai complexe, dar și posibilitățile de calcul avansat cu ajutorul soft-urilor tot mai performante, au condus la crearea unui bazin imens de date care pot fi cu ușurință exploatare pe baza analizei statistice. Poate că acesta este și motivul pentru care statistica rămâne o disciplină percepută ca fiind adesea prea matematizată, destinată specialiștilor. Pentru mulți cititori, mai ales dintre cei care nu au o formare bazată pe un aparat matematic, studiul fenomenelor economice prin metode statistice și matematice, presupune

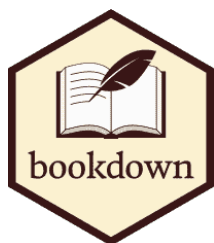
un efort deosebit. Din acest motiv, am încercat să tratez aspectele teoretice, dar și problemele cu aplicație practică din sfera economică, într-o manieră simplă, accesibilă. Așadar, lucrarea are menirea de a facilita înțelegerea conceptelor fundamentale cu care operează statistica, utilizarea adecvată a metodelor de analiză statistică, precum și interpretarea corectă a rezultatelor, în vederea cunoașterii modului de manifestare a fenomenelor.

Nicoleta Caragea
Septembrie, 2018

Quarto

Această carte a fost editată cu ajutorul pachetului R **bookdown** (xie2015?). Cartea are la bază manualul *Statistică - concepte și metode de analiză a datelor*(caragea2015?).

Pachetul R **bookdown** este integrat R Markdown (<http://rmarkdown.rstudio.com>). Documentele elaborate pe baza acestui tip de instrumentar de editare sunt pe deplin reproductibile și dau posibilitatea creării unor formate de ieșire diverse (PDF/HTML/Word/...). Informații suplimentare referitoare la utilizarea pachetului **bookdown** se pot găsi la adresa: <https://bookdown.org>.



Informații despre software

Software-ul R a devenit în prezent unul dintre cele mai utilizate instrumente de analiză statistică, fiind utilizat în statisticile oficiale, în mediile universitare și de cercetare academică, dar și în mediul de afaceri. Acest manual este destinat tuturor celor care doresc să învețe statistica, fiind un material introductiv de studiu, care prezintă un spectru larg de exemple, prezentări grafice și analiză a datelor, dezvoltate cu ajutorul R.

Aplicațiile din această carte utilizează R, ceea ce înseamnă că pentru reproducerea acestora va fi nevoie de instalarea R pe calculatorul pe care lucrați.

R este un sistem pentru analize statistice și reprezentare grafică creat de către Ross Ihaka și Robert Gentleman, profesori de statistică la Universitatea Auckland din Noua Zeelandă¹.

R este considerat un dialect al limbajului S creat de AT&T Bell Laboratories. S este disponibil sub forma software-ului S-PLUS, comercializat de compania Insightful. Există diferențe importante între cele două limbaje, R și S: acestea sunt documentate de către Ihaka & Gentleman (1996) sau se regăsesc în R-FAQ².

Astfel, numele limbajului R provine de la inițiala prenumelui creatorilor, dar este totodată și un omagiu adus limbajului S.

În primul rând, R este open-source, fiind distribuit în mod gratuit sub licență

¹Ihaka R. & Gentleman R. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 299–314.

²R-FAQ

*GNU - General Public Licence*³; dezvoltarea și distribuirea sunt în grija câtorva profesori și statisticieni, afiliați companiilor și universităților, cunoscuți sub denumirea generică de *R Development Core Team*.

Conform filosofiei *GNU*⁴, software-ul open-source este caracterizat de libertatea acordată utilizatorilor săi de a-l utiliza, copia, distribui, studia, modifica și îmbunătăți. Mai exact, este vorba de patru forme de libertate acordate utilizatorilor(**dusa2015?**):

- Libertatea de a utiliza programul, în orice scop (libertatea 0);
- Libertatea de a studia modul de funcționare a programului, și de a-l adapta nevoilor proprii (libertatea 1). Accesul la codul-sursă este o precondiție pentru aceasta;
- Libertatea de a redistribui copii, în scopul ajutorării aproapelui tău (libertatea 2);
- Libertatea de a îmbunătăți programul, și de a pune îmbunătățirile la dispoziția publicului, în folosul întregii societăți (libertatea 3). Accesul la codul-sursă este o precondiție pentru aceasta.

Faptul că este gratuit atrage automat avantajul competitiv în fața altor software-uri de analiză statistică, precum Stata, SAS și SPSS. Astfel, costurile alocate licenței de software dispar. R este denumit de către Norman Nie, unul dintre fondatorii SPSS și CEO al Revolution Analytics, “cel mai puternic și flexibil limbaj de programare statistică din lume” (în engleză “*the most powerful and flexible statistical programming language in the world*”).⁵ Dovadă a succesului pe care R îl are în știința datelor, s-au dezvoltat medii de integrare a acestuia în SAS și chiar SPSS. Este vorba despre modulul SAS/IML⁶, care integrează limbajul R în SAS, și despre *translate2R*, un serviciu de traducere a codului SPSS direct în R dezvoltat de compania *eoda*⁷. R are susținerea comunității științifice, dar și a multor companii internaționale. Dintre acestea, menționăm: Google, Facebook, Mozilla, Twitter, The New York Times, The Economist, NewScientist, Lloyd’s, Bing, Johnson&Johnson, Pfizer, Shell, Bank of America, Ford.⁸ R este susținut și de mediul academic. Marile universități din lume sprijină R, la fel cum sprijină și alte inițiative sau software-uri open-source, precum sistemul de operare Linux sau sistemul de preparare a documentelor L^AT_EX.

³GNU

⁴GNU Philosophy

⁵Smith, D., 2010, “R is Hot”, Revolution Analytics

⁶SAS/IML Module

⁷translate2R - eoda

⁸Revolution Analytics, “Companies Using R”

Capitolul 1

Concepte de bază în statistică

1.1 Ce este statistica?

Termenul de “statistică” este unul familiar, pe care îl auzim sau folosim în fiecare zi. Este asociat adesea cu calcule, cu un rezultat exprimat procentual sau cu un tabel de date. Cu toate acestea, termenul trebuie privit într-o accepțiune mai largă.

Poate v-ați întrebat deseori care este valoarea medie a cifrei de afaceri sau care a fost creșterea medie a acesteia într-un interval de timp, ce valoare a vânzărilor ar putea avea o firmă de telefonie mobilă care ar lansa pe piață un nou produs sau serviciu, sau care ar fi cel mai potrivit loc de muncă disponibil când veți absolvi facultatea. Cine va fi ales noul președinte al țării la următoarele alegeri prezidențiale? Sau poate vă întrebați care ar fi contribuția creșterii nivelului de educație al populației ocupate la dezvoltarea economică a țării.

Răspunsurile la astfel de întrebări, dar și la multe altele, vin din înțelegerea corectă a valorilor numerice înregistrate de o variabilă căreia îi vom spune generic “de interes”, a fluctuațiilor sau a tendinței de creștere sau de scădere a acestor valori într-o anumită perioadă de timp.

Conform dicționarelor publicate de institutele de cercetare din România, **statistica** este o *știință care, folosind calculul probabilităților, studiază fenomenele și proceselor de tip colectiv (din societate, natură etc.) din punct de vedere cantitativ în scopul descrierii acestora și a descoperirii legilor care guvernează manifestarea lor*¹. Altfel definită, statistica este *știința care se ocupă cu descrierea și analiza numerică a fenomenelor de masă, dezvăluind particularitățile lor de volum, struc-*

¹Academia RPR *Dicționar Enciclopedic Român*, București: Editura Politică, 1962-1966; Academia Română, Institutul de Lingvistică, Iorgu Iordan, *Dicționarul explicativ al limbii române* (DEX), București: Editura Univers Enciclopedic, 1998

*tură, dinamică, conexiune, precum și regularitățile sau legile ce le guvernează*².

În practică, metodele statistice furnizează multor domenii ale științei un puternic set de instrumente pentru analiza datelor și pentru interpretarea semnificației rezultatelor cercetărilor statistice.

Sensul termenului statistică poate avea mai multe conotații. Una dintre ele este aceea de *date* sau *colecție de date*, referitoare la un anumit fenomen sau domeniu. Într-o altă accepțiune, statistica înseamnă activitatea de *producere a datelor statistice*, de culegere, prelucrare și prezentare a informației statistice sub diverse forme (diseminare), la dispoziția utilizatorilor. Mai mult, termenul de statistică se utilizează și pentru a desemna *metodele statistice aplicate* în cunoașterea fenomenelor de masă, sau teoria statistică, implicând expunerea sistematică a conceptelor și metodelor de cercetare statistică.

Statistica are ca *obiect* studierea aspectelor cantitativ-numerice ale *fenomenelor de masă*, a dimensiunii, dinamicii și structurii acestora, a raporturilor de interdependență și a altor aspecte care pot fi caracterizate numeric. Fenomenele de masă sunt fenomene care se produc sub acțiunea comună și repetată a unui număr mare de factori, cu caracter sistematic sau întâmplător. Fenomenele de masă sunt caracterizate printr-o mare diversitate, formele individuale de manifestare diferind de la o unitate la alta în funcție de modul în care se combină acțiunea acestor factori. Esența acestor fenomene poate fi pusă în evidență numai prin studierea unui număr mare de cazuri. De exemplu, productivitatea muncii, la nivelul economiei naționale, poate fi cunoscută numai prin analiza informațiilor privind rezultatele activității tuturor unităților din economia națională, respectiv cantitatea de forță de muncă utilizată (dintr-o ramură de activitate sau dintr-o unitate teritorială) sau a unei părți semnificative a acestora.

Metoda statistică este constituită dintr-un ansamblu de operații, tehnici, procedee și metode de cercetare statistică. Specificitatea metodei statistice constă în faptul că se bazează pe observarea unui număr mare de cazuri și utilizează instrumente derivate din teoria probabilităților, dezvoltate în cadrul statisticii matematice.

1.2 Începutul demersului statistic

Înainte de a începe studiul statisticii este necesară o privire de ansamblu asupra termenilor cu care operează statistica. Este cunoscut faptul că, pentru organizarea unei cercetări corecte și cuprinzătoare, este necesar să se folosească un limbaj științific specific disciplinei. Acest lucru este valabil și pentru statistică, care a reușit să-și elaboreze propriile noțiuni, concepte de bază pe care să le folosească pe parcursul întregului demers statistic, de la stabilirea obiectivelor cercetării și până la analiza și interpretarea rezultatelor. În acest sens, cercetarea statistică operează cu câteva concepte specifice: colectivitatea sau populația statistică,

²*Mică enciclopedie de statistică*, Editura Științifică și Enciclopedică, București, 1985

unitatea statistică, caracteristicile sau variabilele statistice, datele statistice și indicatorii statistici.

Colectivitatea sau populația statistică reprezintă totalitatea elementelor sau a manifestărilor de aceeași natură, asupra cărora se efectuează cercetarea. În funcție de obiectivul și particularitățile cercetării, populația statistică poate fi formată din persoane sau grupuri de persoane, din instituții, din obiecte, din tranzacții economice, din evenimente etc.

De exemplu:

- O cercetare privind veniturile populației se realizează pe o colectivitate formată din gospodării;
- O cercetare referitoare la nivelul de educație a populației – se realizează pe o colectivitate formată din persoane;
- Colectivitatea care face obiectul unei investigații privind producția și eficiența economică este alcătuită din unități economice;
- Colectivitatea cercetată în cadrul unui studiu privind activitatea sistemului de sănătate cuprinde unități sanitare (toate spitalele, policlinicile, centrele de tratament etc.);
- Studiul statistic al calității producției se realizează pe colectivități formate din obiecte (produse finite, semifabricate etc.);
- Studiul statistic al natalității sau mortalității se realizează pe colectivități compuse din evenimente (nașteri sau decese).

Colectivitatea supusă cercetării poate fi totală (cercetare exhaustivă), atunci când se fac înregistrări referitoare la toate elementele care formează obiectul studiului, sau parțială (eșantion), atunci când se fac înregistrări referitoare numai la o parte din această colectivitate.

Unitățile statistice reprezintă mulțimea numărabilă de elemente care compun colectivitatea statistică. De exemplu, unitatea statistică a unei cercetări referitoare la veniturile populației poate fi gospodăria sau persoana. Unitatea statistică a unei cercetări privind calitatea producției este produsul finit, semifabricatul sau piesa a căreia i se testează caracteristicile.

Unitățile statistice pot fi simple sau complexe. Acestea din urmă sunt formate din mai multe unități simple. O astfel de unitate este gospodăria³. Unitățile pe care se realizează cercetarea salariilor sunt, de asemenea unități complexe: unitățile economice la care se face înregistrarea datelor. Fiecare element al colectivității este

³Prin **gospodărie** se înțelege grupul de două sau mai multe persoane care locuiesc împreună în mod obișnuit, având, în general, legături de rudenie și care se gospodăresc (fac menajul) în comun, participând în totalitate sau parțial la formarea veniturilor și la cheltuirea lor. Persoana care nu aparține de o gospodărie și care declară că locuiește și se gospodărește singură se consideră gospodărie formată dintr-o singură persoană. Se consideră membri ai gospodăriei și persoanele plecate din localitate pentru o perioadă mai mare de 6 luni, care se află în țară sau străinătate, dacă acestea păstrează legături familiale cu gospodăria

purătorul unei *caracteristici* supuse observării statistice. Caracteristica statistică reprezintă acea proprietate/însușire care este comună tuturor unităților unei colectivități statistice cercetate. Formele sau nivelurile concrete ale acestora, denumite *variante* sau *valori*, diferă de la o unitate la alta (sau în timp, în cazul aceleiași unități) sub influența unui complex de factori. Numărul unităților la care se înregistrează aceeași variantă sau valoare poartă denumirea de *frecvență* a variantei/valorii respective.

Caracteristicile statistice poartă numele și de variabile. Acestea pot fi calitative sau cantitative.

Iată câteva exemple de variabile calitative în tabelele de mai jos:

Tabel 1.1: Exemple de variabile calitative

<i>Variabilă</i>	<i>Categorie</i>
Gen	masculin, feminin
Mediu de rezidență	urban, rural
Formă juridică	societate comercială pe acțiuni, societate cu răspundere limitată, regie autonomă
Activitate economică	industrie, agricultură, comerț
Statut ocupațional	salariat, agricultor, șomer, pensionar
Nivel de educație	scăzut, mediu, superior

R Code

```
date_educatie <- read.csv("date/date_educatie.csv")
head(date_educatie)
```

```
id varsta      gen mediu_rezidenta nivel_educatie
1  1      63 masculin      urban      mediu
2  2      55 masculin      rural      scazut
3  3      31 masculin      rural      scazut
4  4      56 masculin      urban      mediu
```

```
5 5      31  feminin      urban      mediu
6 6      63  feminin      rural      superior
```

```
summary(date_educatie[2:4])
```

```
      varsta      gen      mediu_rezidenta
Min.   :16.00  Length:70      Length:70
1st Qu.:33.00  Class  :character  Class  :character
Median :44.50  Mode   :character  Mode   :character
Mean   :43.79
3rd Qu.:55.00
Max.    :65.00
```

Python Code

```
import pandas as pd
date_educatie = pd.read_csv("date/date_educatie.csv")
print(date_educatie.head())
```

```
   id  varsta      gen mediu_rezidenta nivel_educatie
0   1      63  masculin      urban      mediu
1   2      55  masculin      rural      scazut
2   3      31  masculin      rural      scazut
3   4      56  masculin      urban      mediu
4   5      31  feminin      urban      mediu
```

```
print(date_educatie.iloc[:, 1].describe())
```

```
count      70.000000
mean       43.785714
std        14.613764
min        16.000000
25%        33.000000
50%        44.500000
75%        55.000000
max        65.000000
Name: varsta, dtype: float64
```

```
print(date_educatie.iloc[:, 2:5].describe())
```

```
      gen mediu_rezidenta nivel_educatie
count      70           70           70
unique      2           2           3
```

top	feminin	rural	scazut
freq	39	41	25

Variabilele cantitative sunt măsurabile și se exprimă numeric. Acestea pot fi discrete sau continue.
De exemplu, nota la statistică, este o variabilă cantitativă, discretă.

Tabel 1.2: Exemple de variabile cantitative discrete

<i>Variabilă</i>	<i>Valoare</i>
Nota la statistică	10
	9
	8
	7
	6
	5

Un exemplu de variabilă cantitativă continuă este venitul net al salariaților unei firme.
Considerăm că se cunosc veniturile tuturor celor n=200 salariați.

Tabel 1.3: Exemple de variabile cantitative continue

<i>Variabilă</i>	<i>Valoare</i>
Venit net	3500
	1800
	...
	6300
	5450

Statistica operează, în principal, cu frecvențele diferitelor variante ale caracteristicilor calitative, iar analiza statistică a acestor caracteristici se realizează prin metode ce diferă de cele utilizate în studiul variabilelor cantitative.
În cazul unei cercetări referitoare la activitatea economică, de exemplu, se vor înregistra informații privind volumul producției, cheltuielile de producție și profitul, numărul de salariați și dotarea cu active corporale etc. din fiecare unitate economică.
Evaluarea situației existente pe piața forței de muncă, evoluția fenomenelor de ocupare, șomaj și inactivitate va înregistra informații referitoare la populația în vârstă de muncă, populația activă, populația ocupată, numărul șomerilor etc. Un alt exemplu este cel al unei cercetări referitoare la venituri, care va cuprinde,

în principal, informații numerice despre nivelul veniturilor de care beneficiază gospodăriile (din activități salariale, din agricultură, din activități pe cont propriu, din prestații sociale etc.), despre impozitele și contribuțiile de asigurări sociale plătite de gospodării, despre mărimea gospodăriei, vârsta membrilor acesteia etc.

Prin natura lor, variabilele cantitative se grupează în:

- variabile discrete - pot lua numai anumite valori (exprimate prin numere întregi);
- variabile continue - pot lua orice valori într-un interval finit sau infinit.

Numărul salariaților unei firme sau numărul de persoane care compun o gospodărie sunt variabile discrete, în timp ce valoarea producției, nivelul salariilor sau nivelul veniturilor sunt variabile continue (chiar dacă nivelurile înregistrate ale acestora apar ca valori ale unor variabile discrete). În practică, variabilele discrete sunt tratate ca fiind continue, dacă valorile lor posibile sunt numeroase (de ex., numărul salariaților unităților economice).

Datele statistice exprimă valori ale unor caracteristici cantitative ale unităților statistice, precum și ale unor grupe ale colectivității sau ale colectivității în ansamblul ei. Informațiile numerice sunt date statistice numai dacă sunt definite sub aspectul conținutului, al unității de măsură, al identificării unității/grupului/colectivității la care se referă, precum și al localizării în spațiu și în timp. Datele se obțin fie direct prin observare, fie sunt rezultate ale prelucrării datelor primare. De exemplu, date statistice sunt nivelurile cifrei de afaceri înregistrate de firmele X și Y din București, în luna septembrie 2014.

Indicatorii statistici reprezintă expresia numerică ce exprimă manifestările unor fenomene economice sau sociale, caracterizând nivelul și structura, modificările în timp și variația spațială, precum și interdependențele dintre ele.

Indicatorii statistici sunt rezultate ale prelucrării informațiilor primare înregistrate în cursul unei observări statistice sau a unor informații cuprinse în publicații statistice sau baze de date. De exemplu, modificarea cifrei de afaceri a firmei X în luna septembrie 2014 față de aceeași lună a anului precedent este un indicator statistic.

Pentru cunoașterea fenomenelor de masă, indicatorii statistici îndeplinesc mai multe funcții și anume: de măsurare, de comparare, de analiză, de sinteză; de estimare, de verificare a ipotezelor și/sau de testare a semnificației parametrilor utilizați.

Simpla enumerare a principalelor funcții⁴ ale indicatorilor statistici pune în

⁴Pentru elaborarea și utilizarea corectă a indicatorilor statistici este esențială îndeplinirea unor cerințe de principiu, generale. În acest sens, Yule (1945) precizează condițiile care ar trebui să le îndeplinească un astfel de indicator și anume:

- să fie definit în mod obiectiv, independent de dorința utilizatorului;
- să depindă determinarea sa de toate valorile individuale înregistrate;

evidență o multitudine de aspecte care trebuie avute în vedere la elaborarea și folosirea acestora în analiză, inclusiv stabilirea condițiilor și limitelor în care pot fi utilizați indicatorii statistici în raport cu conținutul specific al fenomenelor, al surselor de informație de care se dispune, cu scopul cercetării.

Indicatorii statistici se pot grupa în indicatori primari și derivați:

1. Indicatori primari (mărimi absolute) – exprimă direct, general nivelul caracteristicii cercetate. Se pot obține prin înregistrarea directă, centralizarea datelor sau prin însumarea parțială sau totală a datelor individuale; prezintă o capacitate relativ limitată de descriere a fenomenului/procesului analizat, și nu permite realizarea unor aprecieri calitative, însă reprezintă punctul de plecare al analizei statistice;
2. Indicatori derivați – se obțin prin prelucrarea indicatorilor primari (absoluți) și fac posibilă analiza aspectelor calitative ale fenomenelor și proceselor analizate (de exemplu: mărimi relative, mărimi medii, indicatori ai variației, indici, indicatori ai corelației, etc.).

1.3 Mărimi relative

Mărimile relative constituie una dintre principalele categorii de indicatori statistici utilizați în analiza economică. Indicatorii statistici reprezintă expresii numerice ale unor fenomene, procese sau categorii economico-sociale, definite în timp, în spațiu și instituțional-organizatoric.

Statistica operează cu indicatori primari și derivați.

Indicatorii primari sunt mărimi absolute care exprimă nivelul caracteristicii cercetate, obținându-se prin înregistrare directă sau prin însumarea datelor individuale. Indicatorii primari se exprimă în unități de măsură, naturale sau monetare, specifice caracteristicii observate: bucăți, tone, kilograme, metri, metri pătrați, dolari, Euro, lei, mii lei, milioane lei etc. Un astfel de indicator este, de exemplu: cifra de afaceri (în milioane lei) înregistrată de o firmă având ca principală activitate comercializarea carburanților și cea rezultată prin însumare la nivelul întregii țări; volumul producției la nivelul unei firme sau la nivel național, numărul populației unei localități, a unui județ și a țării; numărul salariaților unei unități economice (societate comercială sau instituție publică), al celor dintr-o anumită ramură de activitate sau dintr-o unitate teritorială (localitate sau județ) și al tuturor salariaților din economia națională; veniturile și cheltuielile bugetelor

-
- să aibă o semnificație concretă, ușor de înțeles chiar și de nespecialiști;
 - să fie simplu și rapid de calculat;
 - să fie puțin sensibil la fluctuațiile de selecție (să nu prezinte valori puternic diferite, dacă se calculează pe baza mai multor eșantioane, de același volum, extrase prin același procedeu din aceeași colectivitate);
 - să se preteze la calcule algebrice (să poată fi utilizat în operații de comparare a mai multor serii statistice sau în operații de agregare/dezagregare).

de asigurări sociale (în miliarde lei); volumul exportului sau importului realizat la o anumită grupă de produse, cu o anumită țară sau cu toate țările (în milioane dolari sau Euro).

Indicatorii derivați se obțin în urma prelucrării datelor primare, prin aplicarea diferitelor metode și procedee de calcul statistic. Există o mare diversitate de indicatorii derivați, cu metode variate de calcul, de la operații simple de împărțire sau scădere la calcule bazate pe formule complexe. În categoria indicatorilor derivați intră mediile, quantilele, indicatorii variației și ai inegalității, indicatorii corelației, indicii etc. O parte a acestora au caracter absolut și se exprimă în unități de măsură naturale sau monetare specifice caracteristicilor care au stat la baza calculului lor. Altele de exprimă sub formă de coeficienți sau procente. Mărimile relative formează una dintre categoriile de indicatori derivați cele mai frecvent utilizate în analiză.

Mărimile relative rezultă din compararea a doi indicatori, prin raportarea unuia, numit termen de comparat, la altul numit bază de comparație. Rezultatul obținut este un număr întreg sau fracționar, care arată câte unități din indicatorul de comparat revin la o unitate din indicatorul bază de raportare. În cazul celor mai multe mărimi relative, rezultatul raportului capătă relevanță în analiză prin multiplicarea cu 10k (respectiv cu 10, 100, 1000, 10000, 100000,, corespunzător lui $k = 1, 2, 3, 4, 5, \dots$). Mărimile relative calculate ca raport a doi indicatori absoluți similari, exprimați în aceeași unitate de măsură, iau forma unor coeficienți rezultați direct din împărțire sau se exprimă sub formă de procente (% , dacă rezultatul raportului este înmulțit cu 100), promile (‰, dacă rezultatul este multiplicat cu 1000), prodecimile (în cazul înmulțirii cu 10000) sau procentimile (în cazul înmulțirii cu 100000). Mărimile relative calculate ca raport a doi indicatori diferiți, exprimați în unități de măsură diferite, sunt exprimate în unități de măsură compuse din unitățile de măsură ale celor doi indicatori. Dacă asigurarea relevanței rezultatului impune multiplicarea raportului cu 100, 1000 etc., aceasta este reflectată în unitatea de măsură a indicatorului rezultat. Principalele tipuri de mărimi relative sunt: mărimile relative de structură, mărimile relative de coordonare, mărimile relative de intensitate, mărimile relative ale dinamicii, mărimile relative ale planului.

1.3.1 Mărimi relative de structură

Mărimile relative de structură reprezintă raportul dintre un element sau un grup de elemente ale unei colectivități și întreaga colectivitate, calculat pe baza unei caracteristici de volum sau a frecvenței. Mărimea relativă de structură este raportul dintre o parte și întregul de care aparține această parte. Mărimile relative de structură sunt exprimate, în principal, sub formă procentuală, indicând câte unități revin părții analizate la fiecare 100 de unități ale întregului. Dacă sunt exprimate sub formă de coeficienți, acestea arată ce fracțiune din întreg revine

părții analizate (0,25, adică 25 de sutimi un sfert; 0,10 sau o zecime; 0,33(3) sau o treime etc).

Dacă se calculează pe baza unei caracteristici de volum (volumul producției, al cheltuielilor de producție, al profitului, al importului sau exportului, al veniturilor etc), mărimea relativă de structură poartă denumirea de pondere sau greutate specifică. Aceasta exprimă raportul în care se află volumul aferent părții analizate (x_i) față de volumul însumat al tuturor părților care compun întregul:

$$g = \frac{x_i}{\sum_{i=1}^n x_i} 100$$

Observații:

* Suma ponderilor tuturor componentelor este egală cu 100%.

** Diferența dintre doi indicatori calculați sub formă procentuală este exprimată în puncte procentuale (p.p.).

Exemple: 1. Producția industrială realizată în anul t, pe principalele grupe de activități, a înregistrat următoarele valori, conform tabelului de mai jos:

Tabel 1.4: Producția industrială în anul t

	<i>Q (mil.lei)</i>
Industria extractivă	55403,4
Industria prelucrătoare	769938,6
Energie electrică și termică, gaze și apă	141103,0
<i>Total</i>	966445,0

În anul t, ponderea industriei extractive în volumul total al producției industriale a fost de:

$$g = \frac{x_i}{\sum_{i=1}^n x_i} 100 = \frac{55403,4}{966445} 100 = 5,7\%$$

Ponderile celorlalte două grupe de activități se calculează în același mod, astfel încât structura producției industriale are următoarea configurație:

Mărimile relative de structură calculate pentru seriile de distribuție de frecvențe, numite **frecvențe relative** sau **ponderi**, exprimă raportul în care se află numărul unităților din fiecare grupă (n_i) față de numărul unităților din întreaga colectivitate.

2. În anul t, s-a înregistrat următoarea distribuție a salariaților după nivelul salariilor:

Tabel 1.5: Structura producției industriale în anul t

	<i>g (%)</i>
Industria extractivă	<i>5,7</i>
Industria prelucrătoare	<i>79,7</i>
Energie electrică și termică, gaze și apă	<i>14,6</i>
<i>Total</i>	<i>100,0</i>

Tabel 1.6: Distribuția salariaților după nivelul salariului brut, în anul t

	Grupe de salarii brute realizate (lei)	Pondere în numărul total de salariați n' (%)
1	Până la 350	<i>23,6</i>
2	350-400	<i>8,1</i>
3	400-500	<i>12,6</i>
4	500-700	<i>17,9</i>
5	700-1000	<i>16,1</i>
6	1000-1500	<i>12,5</i>
7	1500-2000	<i>4,6</i>
8	2000-3000	<i>2,7</i>
9	3000-4000	<i>1,0</i>
10	4000-5000	<i>0,4</i>
11	Peste 5000	<i>0,5</i>
	<i>Total</i>	<i>100,0</i>

Dacă numărul total al salariaților a fost de 3814 mii în anul t, poate fi estimat și numărul celor cuprinși în prima grupă (în mii):

$$n_1 = \frac{n'_1}{100} N = \frac{23,6}{100} 3814 = 900$$

În mod similar se determină și numărul salariaților cuprinși în celelalte grupe. De exemplu, pentru ultima grupă, a unsprezecea:

$$n_{11} = \frac{n'_{11}}{100} N = \frac{0,5}{100} 3814 = 19,07$$

Așadar, numărul salariaților cu un salariu mai mare de 5000 este de 19 mii.

1.3.2 Mărimi relative de intensitate

Mărimile relative de coordonare se utilizează pentru a compara două niveluri ale aceluiași indicator, niveluri care privesc două grupe/părți (A și B) ale aceleiași populații statistice sau două populații (A și B) de același tip situate în spații diferite. Calculul mărimilor relative de coordonare se poate face raportând nivelul înregistrat în cazul populației/grupei A la cel înregistrat în cazul populației/grupei B sau, invers, raportând nivelul aferent populației B la cel corespunzător populației/grupei A:

$$k_{A/B} = \frac{x_A}{x_B}$$

$$k_{B/A} = \frac{x_B}{x_A}$$

Rezultatul este supraunitar într-una dintre variantele de calcul și subunitar în cealaltă. Dacă $x_A > x_B$, raportul $k_{A/B}$ este supraunitar și arată de câte ori este mai mare nivelul corespunzător populației A, comparativ cu B. În acest caz, raportul invers $k_{B/A}$ este subunitar, indicând ce fracțiune reprezintă nivelul înregistrat de populația B în raport cu nivelul A. Indicatorul subunitar este mai relevant dacă este exprimat procentual.

Exemple:

1. În anul t, raportul dintre produsul intern brut pe locuitor al Italiei și cel al României (calculate în dolari la paritatea puterii de cumpărare) a fost de 3,3.

$$k_{Italia/Romnia} = \frac{PIB_{Italia}}{PIB_{Romnia}}$$

$$k_{Italia/Romnia} = \frac{23300}{7100} = 3,3$$

Rezultatul indică un nivel de bunăstare de 3,3 ori mai înalt în Italia decât în România.

Raportul invers:

$$k_{Romnia/Italia} = \frac{PIB_{Romnia}}{PIB_{Italia}} = 0,304$$

arată că nivelul de bunăstare înregistrat în România reprezintă numai 30,4% în raport cu cel atins de Italia.

1.3.3 Mărimi relative de coordonare

Mărimile relative de intensitate se determină prin raportarea a doi indicatori absoluți (x și y) de natură diferită, dar care se află într-o relație de interdependență. În calculul indicatorului relativ, rezultatul raportului este multiplicat cu 10^k , dacă această operație este necesară pentru asigurarea relevanței:

$$r = \frac{x}{y} 10^k$$

Indicatorii utilizați în calculul mărimilor relative de intensitate sunt exprimați, în general, în unități de măsură diferite, iar rezultatul este exprimat într-o unitate de măsură compusă. Din categoria mărimilor relative de intensitate fac parte următorii indicatori:

- a) **productivitatea muncii** (w), calculată ca raport între volumul fizic sau valoarea producției (Q) și consumul de muncă (T , dacă este exprimat în unități de timp de muncă; N , dacă este exprimat în număr de persoane ocupate/salariați):

$$w = \frac{Q}{N}$$

$$w = \frac{Q}{T}$$

sau ca raport între consumul de muncă (T) și volumul producției (Q):

$$w = \frac{T}{Q}$$

- b) **consumul specific** (c), calculat ca raport între consumul total de materii prime, materiale și energie (C) și volumul producției (Q):

$$c = \frac{C}{Q}$$

- c) **randamentele în agricultură**, calculate ca raport între producția vegetală realizată și suprafețele cultivate cu diferite culturi (kilograme de grâu sau de porumb boabe pe hectar, de exemplu) sau între producția animală și numărul animalelor;
- d) **indicatorii demografici**: ratele natalității și mortalității, nupțialității și divorțialității, calculate ca raport (multiplicat cu 1000) între numărul copiilor născuți vii, al deceselor, al căsătoriilor și divorțurilor înregistrate în

cursul unui an și numărul populației; rata mortalității infantile și juvenile, calculată ca raport (multiplicat cu 1000) între numărul copiilor sub un an, respectiv sub 5 ani, decedați și numărul copiilor născuți vii în cursul unui an etc.

O altă categorie de mărimi relative de intensitate este formată din cele calculate pe baza a doi indicatori diferiți sub aspectul conținutului economic, dar exprimați în aceeași unitate de măsură. Indicatorul relativ rezultat din calcul are forma unei rate procentuale:

- rata de activitate și rata de ocupare, calculate ca raport între populația activă, respectiv ocupată și populația în vârstă de muncă;
- rata șomajului, calculată ca raport între numărul șomerilor și numărul populației active;
- rata sărăciei, calculată ca raport între populația săracă și populația totală;
- rata rentabilității;
- rata dobânzii etc.

1.3.4 Mărimi relative ale dinamicii și ale planului

Mărimile relative ale dinamicii se utilizează pentru caracterizarea evoluției în timp a fenomenelor social-economice. Cele mai importante componente ale acestei categorii de mărimi relative sunt indicii și ritmurile. Indicii dinamicii se determină ca raport între nivelurile înregistrate de o variabilă într-o anumită perioadă sau la un moment dat (x_t) și într-o perioadă sau la un moment anterior, considerat bază de comparație (x_0):

$$i_{t/0} = \frac{x_t}{x_0}$$

Rezultatul raportului arată de câte ori a crescut nivelul variabilei în perioada t față de perioada de bază (dacă sunt supraunitari) sau la ce fracțiune din nivelul de bază a scăzut nivelul variabilei (dacă sunt subunitari). Forma de prezentare a indicilor este cea procentuală.

Ritmul reprezintă raportul dintre sporul absolut al variabilei în perioada t față de perioada de bază și nivelul înregistrat în perioada de bază:

$$r_{t/0} = \frac{x_t - x_0}{x_0} 100$$

și arată cu cât a crescut sau a scăzut nivelul variabilei în raport cu nivelul de bază. Mărimile relative ale planului sunt utilizate pentru analiza relației dintre nivelul înregistrat de un indicator în perioada de bază (x_0), cel programat (x_{pl}) și cel realizat în perioada curentă (x_1). Se utilizează doi indici ai planului, ambii calculați ca raport între nivelurile indicatorului analizat:

- indicele sarcinii de plan, determinat prin raportarea nivelului planificat la cel realizat în perioada de bază:

$$i_{plan/0} = \frac{x_{plan}}{x_0} 100$$

- indicele realizării planului, determinat prin raportarea nivelului realizat la cel planificat pentru perioada curentă:

$$i_{1/plan} = \frac{x_1}{x_{plan}} 100$$

Cei doi indici arată proporția modificării programate a fenomenului analizat, respectiv proporția în care a fost realizat nivelul programat. Ambii indici ai planului se calculează sub formă procentuală.

Summary

In summary, this book has no content whatsoever.

`1 + 1`

[1] 2

Bibliografie

