

Prof.univ.dr. Carmen RADU
Conf.univ.dr. Costel Marian IONAȘCU
Lect.univ.dr. Ilie MURĂRIȚA

STATISTICĂ

TEORETICĂ

EDIȚIA A II-A REVIZUITĂ



EDITURA UNIVERSITARIA
Craiova, 2013

Referenți științifici:

Prof.univ.dr. Nicolae VASILESCU

Prof.univ.dr. Lucian BUȘE

Copyright © 2013 Universitaria

Toate drepturile sunt rezervate Editurii Universitaria

ISBN 978-606-510-452-5

Apărut: 2013

TIPOGRAFIA UNIVERSITATII DIN CRAIOVA

Str. Brestei, nr. 156A, Craiova, Dolj, România

Tel.: +40 251 598054

Tipărit în România

INTRODUCERE

Cu un secol în urmă, celebrul romancier, jurnalist și istoric englez Herbert George Wells spunea: *modul de gândire statistic va deveni, într-o zi, la fel de necesar precum abilitatea de a citi ori de a scrie*. Calitatea acestei predicții a fost validată în timp de încrederea unui număr tot mai mare de oameni în această știință. Neîncrederea altora, pentru că mai există, se datorează în mare parte, în opinia noastră, necunoașterii acestei științe uimitoare. Și lor li se adresează lucrarea noastră, nu pentru a-i cinsti memoria lui H. G. Wells, ci pentru a onora eforturile multor alți specialiști în acest domeniu, pentru a pune la locul cuvenit această știință fără de care viața contemporană ar arăta altfel.

Conștient sau nu, toți oamenii folosesc statistica sau statisticile. Vrem să îi ajutăm să o folosească mai bine: să transmită și să recepționeze corect informația a cărei cantitate imensă este o caracteristică a noului mileniu, dar și să ia decizii cât mai bine fundamentate. Astăzi, mai mult ca oricând, este în interesul nostru, al tuturor, asumarea sprijinului statisticii în procesul decizional la orice nivel.

Pentru majoritatea oamenilor, statistica înseamnă *descrieri numerice*. Random House College Dictionary definește astfel statistica: *știința care se ocupă cu colectarea, clasificarea, analiza și interpretarea informațiilor sau datelor*. Statisticianul profesionist este pregătit în știința statisticii: în colectarea informațiilor numerice sub formă de date, în evaluarea lor și formularea concluziilor pe baza acestora; mai departe, statisticianul stabilește ce informație este relevantă într-o anumită problemă și dacă concluzia desprinsă dintr-un studiu este de încredere.

Am realizat lucrarea de față pornind de la două idei fundamentale pentru noi: statistica este o știință semnificativă și folositoare, a cărei aplicabilitate în afaceri, administrație, științele naturii și științele sociale este aproape nelimitată; statistica joacă un rol cheie în gândirea critică – ea poate minți doar dacă este incorect aplicată.

Pornind de la diferitele maniere de definire a acestei științe, am încercat să cuprindem în această monografie tot ceea ce este esențial pentru înțelegerea și aplicarea statisticii; de aceea am structurat-o într-un mod original, dar logic, în opinia noastră, în trei părți: statistică descriptivă, elemente de statistică matematică și inferență statistică.

În prima parte a lucrării, de *statistică descriptivă*, sunt prezentate, într-o succesiune logică și amplu exemplificate, conceptele, metodele și instrumentele statisticii specifice unei faze exploratorii (descriptive) a cercetării științifice:

observarea, sistematizarea și prezentarea datelor statistice, indicatorii numerici ai seriilor de distribuție și teoria indicilor.

Elementele de statistică matematică, calculul probabilităților și distribuții clasice de probabilitate, prezentate în partea a doua a lucrării, au rolul de a furniza teoriei statisticii bazele logice și rigoarea necesare, conferind abilități calculatorii și criterii de validitate pentru orizontul de competență propriu sectorului din realitatea concretă supus investigației. Cunoașterea lumii reale – a fenomenelor social-economice în special -, deși ghidată de teorie, este întotdeauna factuală. De aceea, în opinia noastră, statistica teoretică are menirea să realizeze o punte de legătură între statistica matematică (cu statutul său de știință logico-formală) și statisticile aplicate, destinate (asemenea oricăror științe factuale) investigării materialului factual dintr-o zonă bine delimitată a realității.

A treia parte a lucrării, de *statistică inferențială*, prezintă conceptele, metodele și instrumentele statistice specifice fazei decizionale (inferențiale) a demersului statistic: introducere în inferența statistică, sondajul statistic, regresie și corelație, ajustarea și predicția seriilor de timp.

Prezentarea corelată și într-o succesiune logică a celor trei părți ale monografiei, ca și riguroasa fundamentare științifică a conceptelor, teoriilor, tehnicilor și metodelor prezentate contribuie, în opinia noastră, la o adresabilitate și accesibilitate sporite, pentru categorii foarte diverse de utilizatori: studenți și masteranzi în domeniul economic, specialiști în alte domenii științifice, oameni de afaceri.

Considerăm utilă detalierea argumentației privind necesitatea și importanța cunoașterii științei și metodologiei statistice pentru viață, în general și pentru mediul de afaceri, în special.

Oamenii de afaceri își doresc să ia decizii bune și să le implementeze corect. Când realizează acest lucru, afacerile lor înfloresc, rezolvă probleme, fac bani, au succes în dezvoltarea de noi oportunități. În munca de implementare - executarea planurilor de afaceri – statistica nu poate contribui decât într-o anumită măsură. Dar în luarea unor decizii bune – planificarea, alegerea dintre mai multe opțiuni, descoperirea a ceea ce își doresc clienții, a ceea ce gândesc și fac angajații, controlul muncii oamenilor și al mașinilor – oamenii de afaceri au nevoie de tot ajutorul pe care îl pot obține. Și statistica îi poate ajuta într-o mare măsură.

Matematica nu poate ajuta afacerile direct, deoarece este o pură abstracție și oamenii de afaceri vor să înțeleagă, să ia decizii, să lucreze în și să schimbe lumea reală. Statistica aduce puterea matematicii în lumea reală, prin colectarea datelor despre lumea reală și aplicarea matematicii asupra lor. Partea bună este că, în timp ce statistica utilizează des matematica, statisticienii adesea nu au nevoie de ea. În lumea practică a statisticii pentru afaceri, matematica este lăsată (sau cel puțin calculele) în seama calculatoarelor. Dar trebuie să știm suficientă matematică pentru a înțelege ecuațiile din instrumentele statistice, pentru a ști care ecuație este folositoare și când.

Statistica nu este o parte a matematicii, a probabilităților, a afacerilor, a științelor naturii sau ingineriei. Ea există independent de alte științe. În același timp, statistica trebuie să facă uz, să se asocieze, să stabilească raporturi cu acestea; numai astfel îi poate ajuta pe oameni să ia decizii corecte.

O problemă fundamentală a afacerilor, și poate a vieții în general, este aceea că am dori să știm cu exactitate cum merge lumea și să știm tot ceea ce se întâmplă, dar nu putem. Mai degrabă, avem doar informații parțiale – și prea adesea incorecte – despre ceea ce se întâmplă în lumea reală. Avem de asemenea un mănunchi de presupuneri – uneori numite teorii, altele, modele – despre cum merge lumea reală. Presupunerile pe care le folosim în afaceri provin adesea de la experți în științele naturii, ingineriei, științele sociale și teoria afacerilor.

Oamenii de afaceri cer adesea ajutorul experților în luarea deciziilor. Înțelegem că experții cunosc domeniul lor de studiu. Dar dacă întregul lor model este greșit? Tot ceea ce putem primi de la cineva printr-un model despre lumea reală – model de afaceri, model științific, social sau ingineresc – se poate rezuma astfel: dacă modelul este bun atunci sfatul expertului va spori șansele pentru luarea deciziei corecte; dar dacă modelul este greșit?

Statistica are un rol aparte față de alte domenii. Ingineria, științele naturii, științele sociale și modelele de afaceri, toate se bazează pe faptul că au dreptate în ceea ce privește mersul lumii. Statistica nu pleacă de la aceeași bază. Ea presupune un singur lucru: viitorul este în general asemănător cu trecutul. Dacă acceptăm acest principiu, putem utiliza statistica pentru a înțelege lumea, chiar dacă nu avem modele, în care să avem suficientă încredere, despre lumea reală. O parte însemnată a procesului de luare a unei decizii corecte este evitarea presupunerilor care ar putea fi greșite. În utilizarea statisticii, evităm presupunerea că o idee particulară despre mersul lumii este adevărată. Ne bazăm mai mult pe general, pe principii dovedite.

Statistica ajută afacerile în două moduri de bază: primul este numit *statistică descriptivă* și ne învață câteva lucruri utile despre ce se întâmplă în jurul nostru, cu ajutorul datelor pe care le avem despre lume; al doilea este numit *statistică inferențială* și ne ajută să cunoaștem lucrurile pe care nu le putem măsura sau număra și ce se poate probabil întâmpla dacă luăm o anumită decizie.

Statistica este un excelent mod de a obține informații despre lume și de a furniza, cu ajutorul acestora, descrieri generale ale lumii. Este foarte important să recunoaștem când, unde și cum statistica ne poate ajuta în luarea unei decizii de afaceri (și nu numai).

Sperăm că pledoaria noastră, pe care o considerăm argumentată prin conținutul acestei lucrări, va convinge tot mai mulți oameni să creadă în cuvintele lui H. G. Wells; pe toți cei convingși de importanța și utilitatea acestei științe uimitoare, statistica, îi invităm să citească această lucrare și le mulțumim anticipat pentru observațiile critice pe care ni le vor transmite.

Autorii

PARTEA I

**STATISTICĂ
DESCRIPTIVĂ**

1 ISTORICUL ȘI DEFINIREA STATISTICII

1.1. Etimologie

Rădăcinile statisticii sunt fixate în antichitate odată cu primele însemnări cantitative legate de populație¹.

Originea termenului statistică se regăsește în cuvântul latin **status** (stare, poziție, situație), respectiv termenul italian **stato** (stat). Mai târziu, pornind de la acesta, a apărut un nou termen **statista** care semnifică persoană ce desfășura afaceri cu statul. Astfel, cuvântul statistică, semnifică la început o colecție de fapte utile unui *statista*. Cu acest înțeles statistica a fost utilizată în Italia în secolul 16, apoi a fost preluată în Franța, Olanda și Germania. Sensul său a fost generalizat ajungând să semnifice *colecție de date numerice referitoare la anumite caracteristici ale unor mulțimi de elemente (persoane, obiecte etc.)* din aproape oricare domeniu.

Statistica, în accepțiunea sa de disciplină științifică, ne oferă principiile pentru a domina materialul factual, a identifica în masa amorfă a datelor structuri, regularități, tendințe, relații, ierarhii și de asemenea, definește regulile de prezentare și interpretare corectă și adecvată a acestora.

Aceste caracteristici definesc imaginea tradițională a statisticii. Progresele realizate, dezvoltările contemporane, adaugă acestei științe noi dimensiuni, redefinindu-i continuu limitele.

Mai important poate decât atingerea scopului său imediat (extragerea semnificației din date) este *impactul cultural* al statisticii, și anume, impunerea unui nou mod de gândire care permite identificarea și recunoașterea, alături de conexiunea proprie **determinismului**² **cauzal**, a unei conexiuni specifice, mai puțin stricte, **determinismul statistic**.

Principalele diferențe existente între cele două concepte: determinism cauzal și determinism statistic sunt prezentate sintetic în continuare:

¹ Vezi A. Hald – *A history of Probability and Statistics and Their Applications before 1750*, New York, Willey, 1990.

² determinism – teorie, concepție potrivit căreia fenomenele sunt generate de înălțări de cauze și efecte, prin condiționări și legități, prin interacțiuni necesare și repetitive.

Determinismul causal	Determinismul statistic
<ul style="list-style-type: none">• Studiază fenomene certe, univoc determinate, ca efect al unei cauze unice, care ori de câte ori se manifestă reproduce invariabil același efect, dacă celelalte condiții rămân neschimbate.	<ul style="list-style-type: none">• Studiază fenomene sociale și economice de masă ce apar ca efect al influenței combinate a unui mare număr de cauze și condiții:<ul style="list-style-type: none">○ în cadrul raporturilor de cauzalitate, nu toate cauzele au același grad de esențialitate și influențare.○ acțiunea unor factori se poate anihila sau compensa reciproc (influența factorilor obiectivi esențiali se asociază cu influența factorilor subiectivi, cu caracter întâmplător).
<ul style="list-style-type: none">• Metoda de descoperire a legităților de apariție a fenomenelor este bazată pe experimente.	<ul style="list-style-type: none">• Metoda utilizată în descoperirea legităților este analitică și se bazează pe generalizări și abstractizări succesive (prin care se înlătură ceea ce este neesențial și întâmplător în dezvoltarea fenomenelor).
<ul style="list-style-type: none">• Fenomenele se produc pe baza legilor dinamicii și apar ca fenomene identice.	<ul style="list-style-type: none">• Fenomenele se produc pe baza legilor statisticii, ce se manifestă ca tendință³ (fenomen de masă) asemănătoare dar nu neapărat identice.
<ul style="list-style-type: none">• Caracteristicile legilor dinamicii: obiectivitate, necesitate, generalitate, repetabilitate.	<ul style="list-style-type: none">• Caracteristicile legilor statisticii sunt aceleași ca și în cazul celor ale dinamicii cu condiția plasării lor la nivel de masă (nu individual).

Cu un secol în urmă, Herbert George Wels⁴ spunea: „Modul de gândire statistic va deveni, într-o zi, la fel de necesar precum abilitatea de a citi ori de a scrie.”

Necesitățile practice în continuă creștere au condus la crearea unei profesii noi - statistician - ale cărei baze teoretice și practice sunt furnizate de statistică. Ele nu se opresc aici; astăzi principiile, noțiunile, metodele, tehnicile furnizate de statistică au devenit fundamentale pentru mai multe domenii de activitate: economic, medical, biologic, psiho-sociologic, agricol, chimic, filologic, meteorologic, sportiv, politic etc.

³ Cunoscută prin studierea unui număr mare de cazuri pentru a reliefa ceea ce este comun și tipic pentru întreaga masă și pentru a elimina abaterile întâmplătoare.

⁴ Profesor, jurnalist, istoric și romancier englez – literatură științifico-fantastică. A publicat *Mașina timpului*, *Războiul lumilor*, *Omul invizibil*, *Insula doctorului Moreau*.

Istoricul și definirea statisticii

Statistica reprezintă arta și știința⁵ colectării și înțelegerii datelor ce caracterizează fenomenele social-economice de masă.

În prezent, în domeniul social-economic, obținerea unor informații care să caracterizeze corect și concret situația existentă, realizarea feed-back-urilor rapide, constituie condiții pentru validitatea și viabilitatea oricăror acțiuni și decizii.

1.2. Istoric

Statistica a cunoscut în dezvoltarea sa mai multe etape:

STATISTICA PRACTICĂ (*evidența statistică*)

Apariție: aproximativ 5 milenii în urmă;

Condiții de apariție:

- creșterea numerică a populației

Perioada	Rata medie anuală de creștere a populației (%)
1950-2000	1,63
1965-1970	2,06

Rata de creștere din a doua jumătate a secolului este dublă față de cea din prima jumătate a secolului.

Perioada	Populația (mil. loc.)
1 mil. ani î.e.n.	2,5
1 e.n.	300
1800	1000
1927	2000
1960	3000
1975	4000
1986	5000
1998	6000
2005	6500
2010	7000 ^{*)}
2050	9000 ^{*)}
2100	10000 ^{*)} stabilizare

^{*)} estimări

- diversificarea activităților oamenilor și apariția necesității unei evidențe a tuturor bunurilor statului pentru atingerea anumitor scopuri

Exemple:

Egiptul antic

- recensăminte regulate ale populației, aurului și pământului (faraonul Amenopolis III numărătoare atestată a populației în anul 1500 î.e.n.)

⁵ Orice știință parcurge trei stadii: preștiință, știință empirică (stadiu în care se bazează numai pe experiența senzorială), știință teoretică (stadiu marcat de descoperirea unor adevăruri cu caracter general, ce descriu legități specifice domeniului investigat, structurate într-un cadru conceptual unitar și coerent).

Statistică teoretică

- *lectura hieroglifelor*⁶
- *nilometrul*⁷ - *primul barometru economic*

China antică

- *informații despre populație și terenuri*

Grecia antică

- *evidențe în scop militar și fiscal, evidențe despre repartizarea terenurilor, despre fixarea condițiilor sociale ale locuitorilor etc.*

Roma antică

- *69 de recensăminte (census)*
- *Cezar – primul recensământ pe imobile (pe răspunderea proprietarului), metodă folosită și astăzi*

Dacia

- *birouri speciale de evidență (tabularium)*

Caracteristici:

- nu erau fundamentate științific
- erau statistici de stare

STATISTICA DESCRIPTIVĂ

Apariție: sec. 13-14 și a atins apogeul în sec. 17-18.

Condiții:

- consolidarea statelor
- apariția de noi ramuri economice
- extinderea raporturilor comerciale dintre state

Caracteristici:

- se trece de la simpla consemnare a faptelor la descrierea, analiza și corelarea acestora;
- statistica este introdusă ca disciplină în învățământul superior și ridicată la rang de știință.

Exemple:

- Rapoartele Republicii Venețiene (*relazioni*): informații despre partenerii comerciali (mici anuare de comerț exterior)
- Epoca Renașterii – apare o enciclopedie a statelor europene
- Școala germană de statistică descriptivă prin reprezentanții săi de seamă: Hermann Conring (fondator al școlii statisticii descriptive, a înființat primul curs de statistică); Martin Schmeitzel; Gottfried Achenwall (cel care a consacrat numele de statistică acestei disciplinei)
- Moldova, realizarea lucrării *Descriptio Moldaviae* de către Dimitrie Cantemir.

⁶ Semn sau caracter din scrierea vechilor egipteni, care reprezenta noțiunile prin figuri de ființe și de obiecte. În limbajul hieroglific nu existau expresii numerice corespunzătoare anilor, aceștia fiind reprezentați cu ajutorul evenimentelor importante care îi marcau (cuceriri militare, ridicarea unui templu etc.)

⁷ Nivelul Nilului era folosit ca un indice pentru nivelul de fertilitate al terenurilor agricole și pentru determinarea mărimii izvoarelor.