

Lp 0* Rezumat

Partea I. INTRODUCERE ȘI CONCEPTE STATISTICE DE BAZĂ

Capitolul 1. INTRODUCERE

Statistica = instrument de cunoaștere științifică, complex, bazat pe reguli empirice, teoria probabilităților, teoria (matematică a) informației, geometrie euclidiană, algebră liniară, topologie, logică matematică și formală etc.

1.1. Biometrie, biomatematică, biostatistică

Biometria = „știința aplicării metodelor matematice (statistice) la studiul multiplelor aspecte ale vieții“ (K. Pearson). **Biostatistică** = știința aplicării în biologie a statisticii. **Biomatematică** = știința aplicării în biologie a matematicii diferită de statistica matematică [8].

Demers **interdisciplinar**:

- strict necesar pentru **verificarea ipotezelor biologice**, prin *statistica inductivă*, singurul instrument științific de generalizare a observațiilor;
- util pentru **sugerarea de ipoteze biologice**, în special prin *statistica descriptivă multivariată* considerată unul dintre extrem de rarele instrumente de descoperire.

1.2. Schema metodologică a apelului la biostatistică

Regulă fundamentală: Metodele de prelucrare a datelor se stabilesc *înaintea* efectuării observației sau experimentului, deși se aplică după acestea.

1.3. Observații cu privire la softul de biostatistică

Nici un program de calculator, oricât de performant va fi, nu va putea suplini nici gândirea statistică a biologului și nici consultanța unui biometrician.

* Lucrarea Practică 0.

Capitolul 2. CONCEPTE STATISTICE DE BAZĂ

Unitate statistică, caracteristică variabilă (sau, mai simplu, *variabilă*), *serie statistică* (*șir statistic*), respectiv *distribuție de frecvențe, populație statistică, eșantion, eșantioane prelevate independent și eșantioane de observații perechi*.

Scopul statisticii = studiul **seriilor statistice** (mulțimi de observații efectuate asupra unor obiecte de aceeași natură, denumite **unități statistice**, care prezintă **caracteristici variabile** susceptibile de a fi *clasate, ordonate sau măsurate*)

2.1. Clasificări ale variabilelor

2.1.1. Clasificare grosieră

- *Variabile calitative* (având *variante*),
- *Variabile cantitative* (având *valori*), care includ și *variabilele semicantitative* (având *ranguri*).

2.1.2. Clasificarea duală a lui Anderberg

1° După mulțimea de reprezentare:

- variabile *discrete* (*finite* - în particular *binare* - ori *infinite*) și
- variabile *continue* (întotdeauna *infinite*).

2° După proprietățile matematice exprimate în scala de reprezentare,

variabilele sunt reprezentabile pe următoarele *scales*: **a. nominală**, **b. ordinală**, **c. interval** și **d. raport**.

2a° Proprietățile scalelor de reprezentare

Scalele au proprietăți din ce în ce mai bogate. *a* și *b* sn* **scale nonmetrice**, *c* și *d* sn **scale metrice**, iar variabilele respective sn **parametri** (în sensul larg al cuvântului).

+2b° Transformări permise în cadrul fiecărei scale

a. Redenumirea, permutarea. **b.** Orice funcție $f(x)$ strict crescătoare. **c.** $f(x) = mx + n$, cu $m \neq 0$. **d.** $f(x) = mx$, cu $m \neq 0$. ($m > 0$ păstrează ordinea, $m < 0$ inversează ordinea.)

* **sn** = se numesc sau se numește (după caz).

2.1.3. Clasificarea de lucru

- *Variabile calitative* (reprezentabile pe a),
- *Variabile tip rang* (reprezentabile pe b) și
- *Variabile tip măsurătoare sau parametri* în sensul larg al cuvântului sau *dimensiuni* (reprezentabile pe c sau d).

2.2. Clasificări ale seriilor statistice

2.2.1. În funcție de ordinea elementelor în șir

Șiruri în care ordinea nu contează, respectiv contează (șiruri temporale, șiruri spațiale).

2.2.2. După numărul de variabile luate în considerație simultan

Serii statistice: *univariate*, *bivariate* și *multivariate* (cu n variabile, $n > 2$). În cazul dimensiunilor se pot numi *unidimensionale*, *bidimensionale* sau *multidimensionale*.

2.3. Clasificarea mulțimilor de unități statistice și structura statisticii clasice

Dacă orizontul interesului nostru se limitează la mulțimea respectivă o denumim *populație statistică (reală sau teoretică)*. Dacă suntem interesați de o supramulțime, mulțimea se numește *eșantion (statistic)*, iar supramulțimea - populație statistică.

Statistica descriptivă studiază complet *populațiile statistice reale*. **Teoria probabilităților** studiază *populațiile statistice teoretice*. **Statistica inductivă** studiază *populațiile statistice reale* incomplet, prin intermediul *eșantioanelor* extrase prin randomizare.

Orice populație biologică poate fi considerată drept o populație statistică, dar invers nu.

2.4. Eșantioane prelevate independent și eşantioane de observații perechi

Eșantioanele de observații perechi au același volum și unitățile dintr-un eşantion sunt în corespondență unu la unu, prin modul de prelevare, cu unitățile celui alt eşantion.

Cazuri generale: (1) aceleași unități măsurate la un singur caracter de două ori ((a) de către doi operatori, (b) cu două aparate sau (c) la o diferență de timp, eventual (c1) după aplicarea unui tratament); (2) aceleași unități pentru care s-au măsurat două caractere cu aceeași unitate de măsură. (3) două mulțimi de câte n unități în corespondență unu la unu în sensul exprimat în formularea comparației. Poate fi măsurat (a) un singur caracter sau (b) câte unul la fiecare eșantion, dar cu aceeași unitate de măsură.

Exemple: (1a) atestarea unui operator nou în raport cu unul atestat, (1b) etalonarea unui nou aparat în raport cu unul etalonat, (1c) studiul creșterii copiilor într-un an, (1c1) studiul eficacității unui tratament, (2) compararea circumferinței abdominale cu cea toracică la bărbați, (3a) compararea coeficientului de inteligență la soți, (3b) compararea lărgimilor umerilor bărbaților cu lărgimile transversale ale bazinelor soțiilor acestora [9].

Eșantioanele independent prelevate pot avea atât volume egale cât și inegale.

Observație: Atunci când un eșantion este format din n unități, pentru care s-a notat un anumit caracter, și un alt eșantion este format din aceleași n unități, studiat după alt caracter, perechile de valori sau variante formează un singur eșantion bivariat. Dacă cele două caractere sunt exprimate în aceeași unitate de măsură, atunci seria bivariată poate fi „văzută” și univariat sub forma a două eșantioane și anume, de observații perechi (vezi ex.(2)).

Lp 0 Teste, exerciții și probleme

TG0*. Durata 140" pe calculator.

Alegeti afirmatia ERONATA:

1. prin statistica inductiva se realizeaza verificarea ipotezelor biologice de specialitate
2. prin statistica descriptiva multivariata se realizeaza sugerarea de ipoteze biologice de specialitate
3. prin statistica inductiva se realizeaza sugerarea de ipoteze biologice de specialitate

* Test Grilă 0.

Schema metodologica a utilizării biostatisticii este următoarea:

1. formularea problemei bio, alegerea programelor de calculator necesare, a tipului și configurației de calculator necesare, stabilirea metodelor biostatistice de prelucrare a datelor, efectuarea experimentului, prelucrarea datelor și interpretarea rezultatelor
2. formularea problemei bio, stabilirea metodelor biostatistice de prelucrare a datelor, alegerea programelor de calculator necesare, a tipului și configurației de calculator necesare, efectuarea experimentului, culegerea datelor, prelucrarea lor și interpretarea rezultatelor

Alegeti afirmatia corecta:

1. studiul populatiilor statistice reale este obiectivul teoriei probabilitatilor
2. studiul populatiilor statistice reale este obiectivul statisticii descriptive
3. studiul populatiilor statistice teoretice este obiectivul statisticii descriptive

Seria statistica este:

1. o multime de obiecte de aceeași natura
2. o multime de obiecte de natura diferită
3. o multime de caracteristici variabile
4. o multime de observatii

Funcțiile de forma $f(x)=mx+n$, sunt transformările permise pentru:

1. scalele raport
2. scalele ordinale
3. scalele interval

Studiul populatiilor statistice reale este obiectivul:

1. statisticii inductive
2. statisticii descriptive
3. teoriei probabilitatilor

Esantioanele prelevate independent pot avea:

1. atât volume egale, cât și inegale
2. numai volume egale
3. numai volume inegale

TC0*. Durata 3'.

1. Concentrația este o variabilă _____ ce se poate reprezenta pe scala _____, iar după mulțimea de reprezentare este o variabilă _____.
2. Anii calendaristici sunt reprezentați pe scala _____, după mulțimea de reprezentare fiind o variabilă _____, iar după clasificarea de lucru o variabilă tip _____.
3. Eșantioanele care pot avea volume _____ sau _____ și nu au corespondență unu la unu între unitățile fiecărui eșantion, sunt eșantioane _____.

Exerciții sau probleme rezolvate

1.

Pentru testarea unui nou preparat de insulină în tratamentul diabetului, s-au luat în studiu două loturi de câte 31 șobolani:

Lotul 1 = indivizi sănătoși; Lotul 2 = indivizi bolnavi de diabet.

S-au determinat glicemia și timpul de coagulare pentru fiecare individ, înainte și după tratamentul cu preparatul testat (timp 0, respectiv timp 1), eșantioanele de observații obținute notându-se astfel:

- | | |
|--------------------------------|---|
| A. Glicemie lot 1 la timpul 0. | B. Timp de coagulare lot 1 la timpul 0. |
| C. Glicemie lot 2 la timpul 0. | D. Timp de coagulare lot 2 la timpul 0. |
| E. Glicemie lot 1 la timpul 1. | F. Timp de coagulare lot 1 la timpul 1. |
| G. Glicemie lot 2 la timpul 1. | H. Timp de coagulare lot 2 la timpul 1. |

Precizați cum sunt următoarele perechi de eșantioane: A și C, E și G, F și H, A și E, D și H, C și G, precum și A și B ?

Rezolvare:

A și C, E și G și F și H sunt eșantioane independent prelevate.

A și E, D și H, C și G sunt eșantioane de observații perechi.

A și B sunt aparent eșantioane de observații perechi dar nu pot fi tratate univariat astfel, deoarece caracterele nu sunt exprimate în aceeași unitate de măsură. Formează doar o serie bivariată.

* Test de Completare 0.

Exerciții sau probleme propuse

2.

Pentru a studia efectul cofeinei asupra tensiunii arteriale la om, mai multe persoane au fost repartizate aleator în două loturi. Cei din primul lot au primit o doză activă de cofeină, iar ceilalți au primit placebo. S-a măsurat tensiunea arterială la toți subiecții din ambele loturi, înainte și după administrarea de cofeină, respectiv placebo. Eșantioanele de observații realizate s-au notat astfel:

- | | |
|--|--|
| A. Lotul 1 înainte de administrarea
cofeinei. | B. Lotul 2 înainte de administrarea de
placebo. |
| C. Lotul 1 după administrarea
cofeinei. | D. Lotul 2 după administrarea de
placebo. |

cum sunt cuplurile de eșantioane: A și B, A și C, C și D, A și D, respectiv B și D ?