



UMF
UNIVERSITATEA DE
MEDICINĂ ȘI FARMACIE
IULIU HAȚIEGANU
CLUJ-NAPOCA

NOȚIUNI FUNDAMENTALE DE STATISTICĂ MEDICALĂ

”

Plan de curs ...

Terminologia statistică

Statistica descriptivă versus statistica inferențială

Variabila, Data, Constanta

Tipuri de variabile

Dicționar statistic

- **Statistica:**

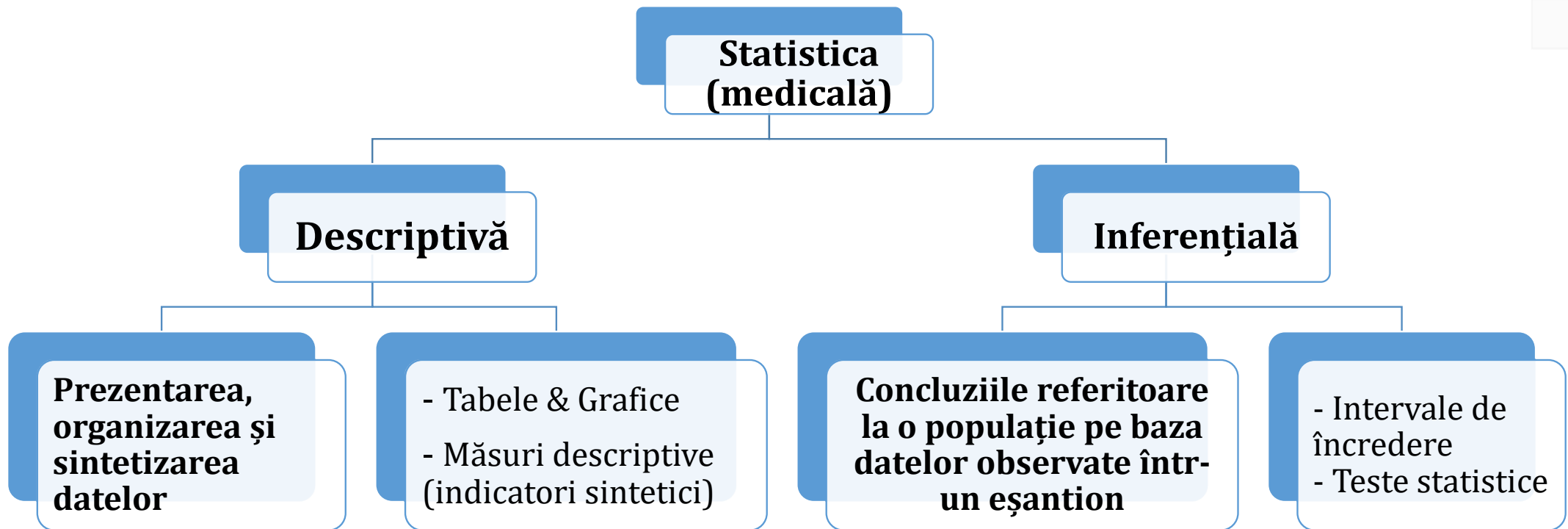
- ✓ știință care culege, sintetizează, descrie și interpretează date referitoare la fenomene generale.
- ✓ domeniu științific care permite studiul fenomenelor a căror proprietate fundamentală este variabilitatea
- ✓ știința colectării datelor, a prezentării lor într-o formă sistematică, a analizei acestora și a interpretării rezultatelor obținute

- **Statistica Medicală:**

- ✓ se ocupă cu culegerea, sintetizarea, descrierea și interpretarea datelor medicale
- ✓ stabilește concluzii despre o populație plecând de la o analiză/studiu efectuat asupra unei părți a populației

Definiții

- Statistica medicală se ocupă de aplicațiile biostatisticii în medicină și științele sănătății, inclusiv epidemiologia și sănătatea publică



Undergraduate nursing students' perceptions on nursing education during the 2020 COVID-19 pandemic: A national sample

Alexandra Michel ¹, Nicole Ryan ¹, Deborah Mattheus ¹, Amelia Knopf ², Nadia N Abuelezzam ³, Kelly Stamp ⁴, Sandra Branson ⁵, Barbara Hekel ⁵, Holly B Fontenot ⁶



Statistica **descriptivă** vs. Statistica **inferențială**

- “..., Purpose: To explore nursing students' perspectives about the effects of the pandemic on their education and intention to join the nursing workforce.
- **Methods:** Undergraduate nursing students from 5 universities across 5 United States regions were invited to participate in an online survey to elicit both quantitative and qualitative data.
- **Findings:** The final sample included quantitative data on 772 students and qualitative data on 540 students. Largely (**65.1%**), students reported that the pandemic strengthened their desire to become a nurse; only **11%** had considered withdrawing from school. Qualitatively, students described the effect of the pandemic on their psychosocial wellbeing, adjustment to online learning, and challenges to clinical experiences.....”
- link articol: <https://pubmed.ncbi.nlm.nih.gov/34183191/>
- “...**Study design:** The aim of InSOMNIA study was to determine the prevalence of sleep disorders among nursing students of the University of Perugia and, therefore, to evaluate how lifestyle, eating habits, health status and academics performance are linked to night-time and daytime symptoms of the interrupted sleep.
- **Results:** We found a statistical **significant association** between PREDIMED score and BMI (**p-value=0.0127**), smoking habit (p-value = 0.0125), quality of life (**p-value = 0.0480**) and academic progress (**p-value = 0.0092**).
- Conclusions: We found a high prevalence of sleep disturbances statistically associated with diet and poor academic progress. “
- link articol: <https://pubmed.ncbi.nlm.nih.gov/30614496/>

Terminologie statistică

- **Unitatea statistică** = elementul constitutiv al populației statistice sau eșantionului
- **Variabila** = caracteristică care poate varia de la un element la altul al unei colectivități;
- **Datele statistice** = valorile observate sau măsurate ale unei variabile statistice
- **Informațiile statistice** = rezultă din prelucrarea și analiza datelor statistice
- **Estimarea** = procedeul utilizat pentru determinarea valorii unui parametru statistic asociat unei populații
- **Estimatorul** = funcție statistică aplicată asupra eșantionului pentru a estima un parametru necunoscut al populației (Valoarea obținută este o estimare a valorii populației)

Populația

POPULAȚIA = mulțime de elemente (obiecte sau subiecți) care au anumite însușiri (atribute sau caracteristici) comune, care formează obiectul unei analize statistice

- **Talia sau volumul** populației = numărul elementelor populației
- **Unitatea statistică** = un element al populației

Exemple:

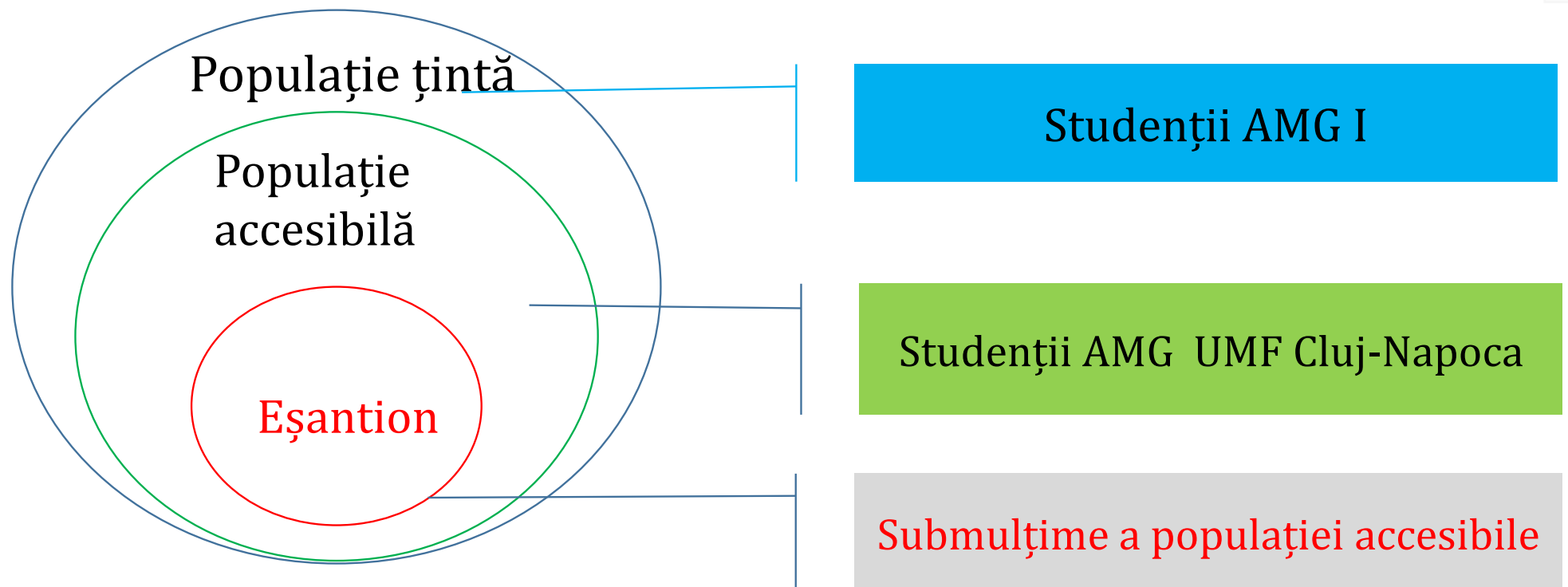
- Studenții AMG din ciclul de licență din România
- Subiecții adulți cu diabet zaharat de tip II din Ro
- Femeile peste 50 ani din Ro cu cancer de sân

Terminologie statistică

- **EȘANTIONUL** = o colecție (submulțime) de elemente din populație
- **Eșantioane perechi** =
 - două eșantioane în care elementele sunt clar perechi (determinarea TAS la **aceiași pacienți** la includerea în studiu și respectiv la 6 luni de la inițierea tratamentului)
 - Elementele unui eșantion sunt selectate pentru a se potrivi cu elementele celui de-al doilea eșantion
- **Eșantioane independente** = probabilitatea ca un element să fie inclus în eșantion nu depinde de selecția unui alt element în cel de-al doilea eșantionul

Populație vs. Eșantion

Obiectiv studiu: Evaluarea statusului ponderal la studenții AMG I



Eșantionul: De ce?

1. Eșantionul se poate evalua mai repede decât populația.
2. Evaluarea unui eșantion este mai ieftină comparativ cu studierea întregii populații.
3. Studiul întregii populații poate determina distrugerea ei.
4. Rezultatele obținute pe eșantion sunt adesea mai precise comparativ cu rezultatele obținute pe populație.
5. Prin extragerea corectă a participanților la studiu dintr-o populație specifică, cercetătorul poate analiza eșantionul și realiza inferențe despre caracteristica studiată a populației.

Caracteristicile eșantionului

- **Reprezentativ pentru populație (vezi cursul cu metode de eșantionare):**

- ✓ Talie
- ✓ Caracteristici măsurate



empiric

- **Calculul taliei eșantionului:**

- ✓ Riscul de a respinge ipoteza nulă dacă ea e adevărată (alfa, $\alpha = 5\% = 0,05$)
- ✓ Puterea studiului (probabilitatea de a respinge ipoteza nulă când nu este adevărată)

Talia populației (N)	Talia eșantionului ($n = \% \times N$)
0 – 100	100
101 – 1000	10
1001 – 5000	5
5001 – 10000	3
> 10000	1

Volumul eşantionului

- **Cost** (cu cât volumul eşantionului e mai mare cu atât costul cercetării e mai mare)
- **Omogenitatea** populației (membrii populației sunt similari în ceea ce privește caracteristica de studiat): cu cât variabilitatea în populație e mai mare cu atât volumul eşantionului trebuie să fie mai mare
- Alți **factori** care pot influența volumul (mărimea) eşantionului:
 - (a) există variabile pe care nu le putem controla
 - (b) se dorește împărțirea eşantionului în subgrupe
 - (c) se prevede un număr mare de pierduți din vedere
 - (d) se dorește o putere statistică mare.

Variabila & Data

- **UNITATEA STATISTICĂ (observația)** = elementul constitutiv al populației statistice sau eșantionului
- **VARIABILA** o caracteristică de interes a cărei valoare nu se cunoaște înainte de examinare / măsurare
- **DATA:** o valoare măsurată/numărată care corespunde unei caracteristici/variabile

Exemplu: observație vs. variabilă vs. dată

VARIABILA

DATĂ

ID	Gen (F/M)	Fumat (Da/Nu)	Greutate (kg)	Domiciliu	Stadiul patologiei maligne
001	F	Da	56	...	Cluj	1
002	M	Da	85	...	Cluj	0
003	M	Da	120	...	Sălaj	4
...
199	M	Nu	75	...	Maramureș	2

Observație / Caz/ Unitate statistică

Date: Example

- Înălțimea dvs.: 165 cm
- Numele unei persoane: Lavinia
- Valoarea pulsului unei persoane: 76 bătăi/min
- Culoarea ochilor a colegului din dreapta dvs.: albaștrii
- Presiunea arterială a unui pacient: 120/60 mmHg
- Valoarea temperaturii unui pacient cu septicemie: 41,6°C
- Dimensiunea inciziei unei apendicectomii: 5,5 cm

Caracteristici ale datelor:

- » Datele sunt materia primă a informației
- » Data e o parte distinctă a informației
- » Datele nu sunt organizate sau prelucrate

VARIABILA VS. CONSTANTĂ

- **CONSTANTA** = un număr specific care rămâne același indiferent de condiții
 - Punctul de înghețare al apei: în orice condiții va fi 0°C.
- **VARIABILA** = o caracteristică care poate lua diferite “valori”
 - Greutatea: dacă măsurăm greutatea fiecărui student din încăperea vom obține o serie de numere care vor varia de la un student la altul
- **VARIABILA** = caracteristică care poate prezenta variație de la un element la altul al unei colectivități
- **DATA** = valoarea pe care o ia variabila pentru un subiect dat

Tipuri de Variabile

1. Variabile **CANTITATIVE**

- Măsurabile
- pot avea ca valori numere întregi sau valori cuprinse într-un interval de numere reale

2. Variabile **CALITATIVE** (Atribut)

- Nemăsurabile
- pot avea lua un număr finit de categorii/clase/modalități/ valori

Tipuri de **VARIABLE CANTITATIVE**

a) Discrete: Valori numere întregi

b) Continue: Exisă un număr infinit de valori posibile care se găsesc în intervalul a două valori observate

Tipuri de **VARIABILE CALITATIVE**

a) Dihotomiale (binare)



b) Nominale



c) Ordinale



TIPURI DE VARIABILE MEDICALE

<p>Culoarea părului (blond;saten,negru)</p> <p>Starea civilă (căsătorit, divorțat, văduv, necăsătorit) ...</p> <p>Tipul anemiei: anemie microcitica (def. de fier);megaloblastica (def. vit. B12); normocitica (asociata cu o boala cronica)</p>	<p>Gen (M, F)</p> <p>Hipertensiune arterială (prezenta/absenta)</p>
<p>Hipertensiune arterială (stadiul I/II, III)</p> <p>Educația (fara studii, ciclu primar, gimnazial, liceal, studii superioare);</p> <p>Starea de sănătate: precara, buna, foarte buna</p> <p>Severitatea durerii: ușoară, moderată, severă</p>	<p>Greutatea corporală (kg)</p> <p>Temperatura corporală (°C)</p> <p>Tensiunea arterială (mmHg)</p> <p>Numărul de copii dintr-o familie;</p> <p>Numărul de intervenții chirurgicale anterioare</p> <p>Numărul de recidive la pacienții cu cancer</p>

Transformarea variabilelor

- Este posibilă transformarea continue în variabile ordinale sau nominale dar întotdeauna transformarea se face cu pierdere de informație
 - Vârstă (ani) în “Grupe (clase) de vârstă” (18-29 ani; 30-39 ani; 40-49 ani; 50-59 ani; ≥ 60 ani)
- Nu este posibilă transformarea variabilelor nominale sau ordinale în variabile cantitative chiar dacă atribuim valori diferitelor clase
 - Sex: M = 1, F = 0

ID	Vârsta (ani)	Clasa de vârstă	IMC (kg/m ²)	Clasa IMC
1	89	>80 ani	28,0	supraponderal
2	58	51-60	23,2	normal
3	64	61-70	31,2	obez-clasa I
...				
100	63	61-70	36,3	obez-clasa II

VARIABLE PRIMARE VS. SECUNDARE

- **Clasificarea variabilelor după modul de obținere**
 - **VARIABLE PRIMARE:** observate, numărate sau măsurate pe unitatea statistică (ex. greutatea (kg), culoarea tegumentelor: paloare / roșeață / cianoză / icter , etc.)
 - **VARIABLE SECUNDARE:** se obțin în urma aplicării unei formule de calcul pe două sau mai multe variabile primare.

Denumire variabilei	Formula de calcul
Indicele de masă corporală	$= \text{greutatea (kg)} / (\text{înălțimea (m)}^2)$
Presiunea arterială medie	$= \text{TAD (mmHg)} + 1/3 * (\text{TAS (mmHg)} - \text{TAD (mmHg)})$ [88] ¶ unde TAD = tensiunea arterială diastolică, TAS = tensiunea arterială sistolică

VARIABLE DEPENDENTE VS. INDEPENDENTE

- **După relația de dependență:**

- Variabile **independente**: variabilele care se pot utiliza pentru a descrie sau explica alte variabile
- Variabile **dependente**: variabile care pot fi explicate de alte variabile independente.

Exemplul 1: Timpul de triaj în serviciile de urgență depinde de factori clinici (experiența, personalitatea, rolul asistentului medical) și instituționali (ghiduri clinice, educația medicală continuă, aparatura de urgență, resursele umane).

În acest exemplu:

- Variabila dependentă: timpul de triaj
- Variabile independente: factorii clinici și instituționali

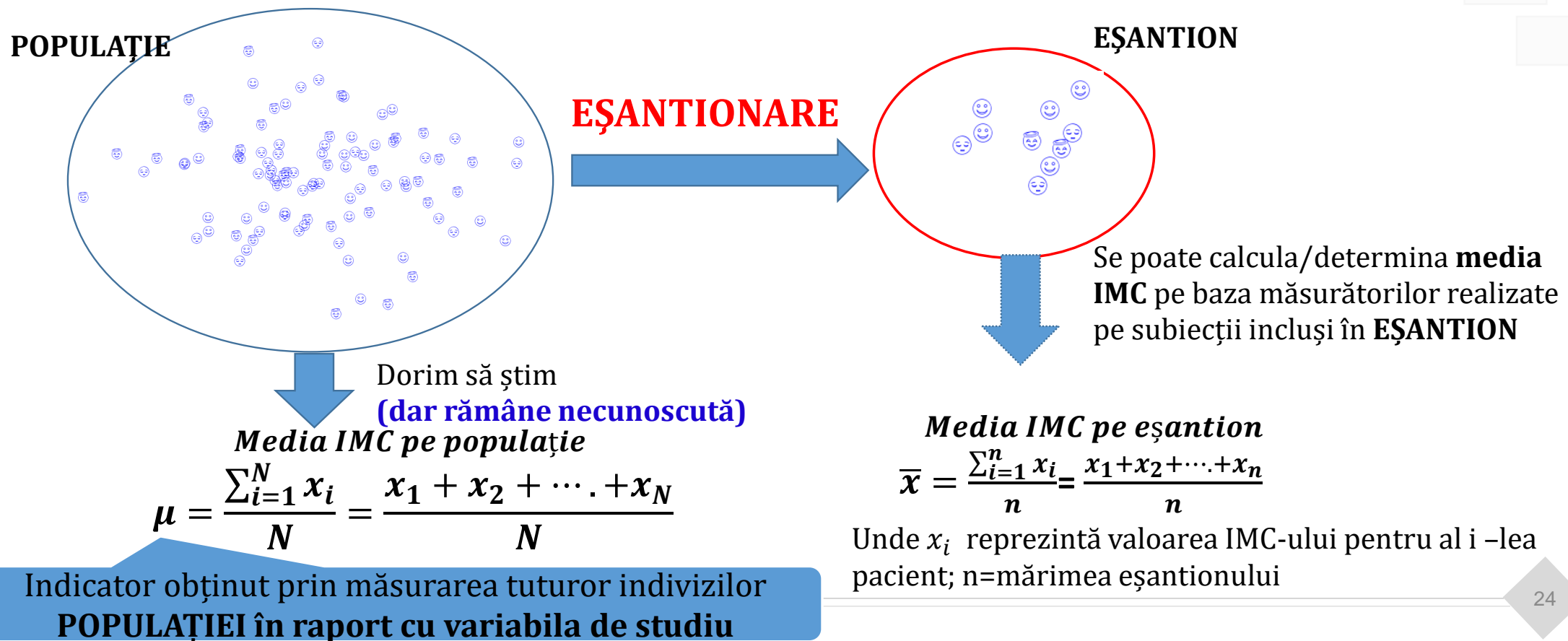
Exemplul 2: Satisfacția pacienților (comunicarea centrata pe pacient) cu cancer de sân depinde de factori socio-demografici și clinici (vârsta, mărimea tumorii, durata consultăției).

În acest exemplu:

- Variabila dependentă: comunicarea centrata pe pacient (scor)
- Variabile independente: vârsta, mărimea tumorii, durata consultăției

Parametru vs. statistică

Evaluarea **statusului ponderal** (IMC, indice de masă corporală, kg/m²)



Parametru vs. statistică: exemplu

Parametru

*"..... Specifically, for Italy, we obtain a **prevalence of COVID-19** of 9.22%, a number that, reported to the total population of Italy, reveals that 5,985,537 people could have been affected by Covid-19 in the country as of 26 May 2020..."*

(link articol:

<https://www.sciencedirect.com/science/article/pii/S0048969720363282>)

Statistica

*Eșantion de 1000 adulți;
Frecvența infecției COVID-19*

De reținut.....

- În cercetarea medicală evaluăm eșantionul pentru a identifica ce se întâmplă în populația de interes
- Doar eșantionul corect selectat din populație (**eșantion reprezentativ**) permite extrapolarea rezultatelor obținute pe eșantion la nivelul populației.
- Scopul **statisticii descriptive** este de a prezenta/sumariza caracteristicile sau variabilele măsurate pe eșantion
- Scopul **statisticii inferențiale** este de a concluziona existența sau nu a unei diferențe semnificative statistic între două sau mai multe populații în raport cu una sau mai multe variabile de interes
- Variabilele pot fi: **CANTITATIVE** (discrete, continue) sau **CALITATIVE** (dihotomiale, ordinale, nominale).
- **Parametru** = măsură (indicator) ce descrie o populație în raport cu o variabilă de studiu
- **Statistica** = măsură (indicator) ce descrie un eșantion în raport cu o variabilă de studiu

Vă mulțumesc!



Contact: Conf.univ.dr. Mihaela Iancu
Email: miancu@umfcluj.ro