

# STATISTICĂ I



# Cine sunt eu?

- Adrian Gorbănescu
  - Formare
  - Specializare
  - Pasiuni
- 
- E-mail: [adrian.gorbanescu@unibuc.ro](mailto:adrian.gorbanescu@unibuc.ro) (Subject – Statistica)
  - Consultații: prin e-mail



# Nota

- Examen – 5 puncte
  - a) 23 ianuarie 2023 (ora 10.00)
  - b) Examen grilă – 40 itemi
  - c) 4/5 itemi cu cel puțin două variante corecte de răspuns
- Parțiale teoretice – 2 puncte
  - a) 15 itemi cu răspuns scurt / parțial
  - b) Parțial 1 – Curs 7
  - c) Parțial 2 - Curs 13
- Seminar – 3 puncte
  - a) Teme – 1 punct
  - b) Prezențe – 1 punct
  - c) Test practic – 1 punct

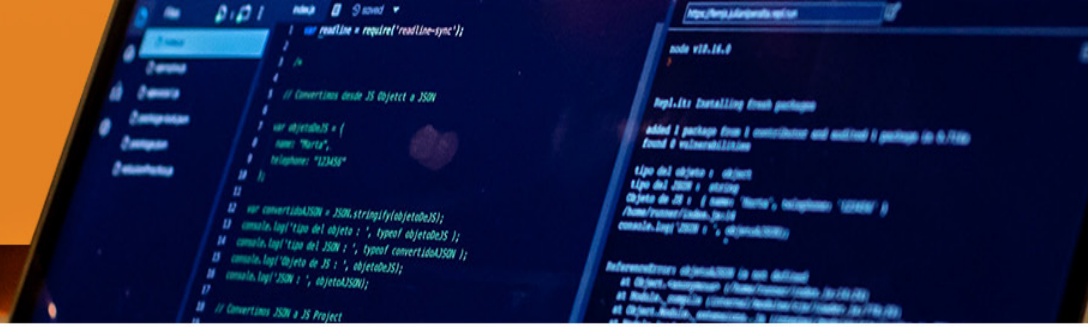


# Bibliografie

- Statistică pentru psihologie. Teorie și aplicații SPSS, M. Popa, Editura Polirom, 2008
- Discovering Statistics Using R, Andy Field, 2013.
- Suportul de curs

# Seminar

- Teodora Georgescu
- Tiana Voicu



# Tematica

1. Noțiuni introductive – 9 octombrie
2. R & R Studio – 16 octombrie
3. Analiza de frecvențe – 23 octombrie
4. Indicatorii sintetici ai distribuției – 30 octombrie
5. Scorurile standardizate – 6 noiembrie
6. Verificarea datelor parametrice – 13 noiembrie
7. Parțial 1; Testul t pentru un singur eșantion - 20 noiembrie
8. Testul t pentru două eșantioane independente – 27 noiembrie
9. Testul t pentru două eșantioane dependente – 4 decembrie
10. Erorile statistice, mărimea efectului și puterea testului – 11 decembrie
11. Coeficientul de corelație liniară Pearson – 18 decembrie
12. ANOVA One-Way – 8 ianuarie
13. Parțial 2; Teste neparametrice pentru date ordinale - 15 ianuarie



# Știați că?

- Folosim statistica în cele mai multe dintre activitățile pe care le desfășurăm într-o zi

# Știați că?

- Aproximativ 80% din incidența cancerului pulmonar la bărbați și 35% la femei este asociată cu fumatul.
- 65% din persoanele private de libertate erau consumatoare de droguri în momentul comiterii infracțiunii
- În procent de 74% proba de maraton de la Jocurile Olimpice a fost câștigată de sportivi din Africa.
- Studentul A are un IQ egal cu 115. Astfel, este mai inteligent decât aproximativ 84% din populația celor de vârstă sa.



# Concepte fundamentale

- **Ce este statistica?**
  - Statistica este știința care se ocupă cu recoltarea, descrierea și analiza datelor în vederea extragerii unor concluzii pe baza acestora (Popa, 2008).
  - Statistica se referă la o paletă de tehnici și proceduri de analiză, interpretare și luare a deciziilor pornind de la un set de date.
  - Statistica face ceva complex să devină simplu.

# Concepte fundamentale

- Exemplu – Anxietate





Q & A



# Concepte fundamentale

- Benjamin Disraeli (fost premier al Marii Britanii) a spus: „Există trei tipuri de minciuni: minciuni, minciuni grosolane și statistica” ([Miles & Banyard, 2007, p. 3](#)).



Benjamin Disraeli



# Concepte fundamentale

Poți  
utiliza  
statistica  
pentru a  
explica  
fenomenele  
care au loc  
în lume



Poți  
spune tot  
ce dorești  
folosindu-  
te de  
statistică,  
chiar și  
minciuni

# Concepte fundamentale

- **Exemplu**

*Producătorul de înghețată „Statistics Icecream” introduce în luna mai o campanie publicitară foarte intensă în mass-media. La începutul lunii septembrie managerul „Statistics Icecream” constată că vânzările au crescut cu 30% și este foarte mulțumit de acțiunea sa. Putem considera campania publicitară ca fiind eficientă?*

# Concepte fundamentale

- **Variabila**

- reprezintă proprietatea, caracteristica unui eveniment, obiect sau persoană de a lua *valori diferite*.

- Exemple:

- a) Genul (1 = masculin, 2 = feminin);

- b) Temperamentul (sanguinic, coleric, flegmatic, melancolic).

- c) Vârsta

- d) Anxietatea

# Concepte fundamentale

- **Variabile independente vs Variabile dependente**

- Atunci când o variabilă este **manipulată** de cercetător este denumită **variabilă independentă**.
- Variabila care **primește efectul** variabilei independente poartă numele de **variabilă dependentă**.
- *Variabila independentă este cea care generează efecte asupra variabilei dependente.*



# Concepte fundamentale

- Într-o cercetare dorim să studiem influența consumului de alcool asupra timpului de reacție.
- În acest caz, timpul de reacție variază în funcție de consumul de alcool.
- Astfel, consumul de alcool este variabila independentă, iar timpul de reacție, cel care suportă efectul, este variabila dependentă.

# Concepte fundamentale

- **Exemplu**
- S-a observat că fructele de pădure au un efect pozitiv asupra memoriei. Într-un studiu, studenții participanți au primit timp de 4 săptămâni trei tipuri de diete: cu fructe de pădure, portocale și morcov. După această dietă studenții au intrat în sesiune și s-a observat că studenții care au ținut dieta cu fructe de pădure au avut note mai mari decât studenții care au consumat portocale sau morcovi.

# Concepte fundamentale

- *Variabile continue și variabile discrete*

- O variabilă este continuă atunci când poate lua orice valoare numerică.
- Greutatea, înălțimea sau timpul de reacție sunt exemple de variabile continue.
- Prin variabilă discretă ne referim la o variabilă care poate primi un număr finit de valori, acestea fiind numere întregi.
- Numărul copiilor dintr-o familie, numărul răspunsurilor corecte la un test, numărul de repetări în vederea memorării unui material sunt exemple de variabile discrete.

# Concepte fundamentale

- ***Variabile observabile vs variabile latentă***

- Agresivitatea – variabilă latentă

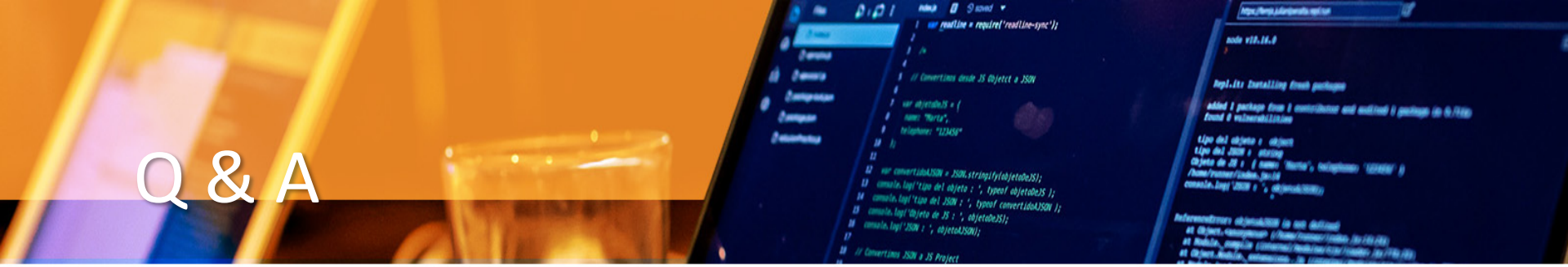
- Lovește, țipă, înjură, trânteste sunt variabile observabile ale agresivității



# Concepte fundamentale

- ***Variabile calitative vs variabile cantitative***
- **Variabilele calitative** sunt acele variabile care exprimă caracteristici non numerice cum ar fi culoarea părului, culoarea ochilor, religia, mașina preferată, genul etc.
- **Variabilele cantitative** sunt acele variabile care sunt exprimate în numere. De exemplu, greutatea, înălțimea, numărul de pagini citite pe zi etc.

Q & A



# Concepte fundamentale

- Tipuri de statistică

**Statistica descriptivă** - reprezintă organizarea, prezentarea și descrierea datelor. Statistica descriptivă ne oferă o expresie (un număr) care descrie setul de date



**Statistica inferențială** se referă la extinderea rezultatelor obținute la nivel de eșantion la nivelul populației din care acesta a fost extras. Rezultatele majorității experimentelor se bazează pe

statistica inferențială.

		Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Current Salary	Equal variances assumed	29.2	.000	-9.36	109	.000	-\$33,038	\$3,530	-\$40,034	-\$26,044
	Equal variances not assumed			-16.3	89.6	.000	-\$33,038	\$2,032	-\$37,076	-\$29,002

# Procesul de măsurare

- **A măsura** înseamnă a atribui numere sau simboluri unui aspect al realității în funcție de anumite aspecte cantitative sau calitative care le caracterizează.
- Modul în care sunt atribuite anumite numere sau simboluri pentru a măsura ceva se numește **scală de măsurare**.



# Procesul de măsurare

- **Scala nominală** permite etichetarea valorilor și plasarea lor în anumite categorii, dar fără a realiza comparații.
- Exemple: starea civilă, genul, temperamentul, trăsăturile de personalitate, CNP-ul etc.
- Acest tip de scală nu permite operații matematice (adunări, scăderi, calcularea mediei) și nu precizează dacă o valoare este mai mare sau mai mică decât alta.

# Procesul de măsurare

- **Scala ordinală** este utilizată în situația în care dorim să stabilim o ierarhie, dar fără a preciza cu cât o valoare este mai mare sau mai mică decât alta.
- Exemple de variabile măsurate pe scală de tip ordinal: ierarhia într-un concurs de admitere, nivelul de școlarizare, categoria de vârstă.
- Acest tip de scală nu permite operații matematice.

# Procesul de măsurare

- **Scala de interval** are caracter cantitativ, este exprimată numeric, are intervale egale și ne exprimă ordinea valorilor.
- putem stabili intensitatea diferenței dintre valorile variabilei măsurate.
- O caracteristică a acestei scale este absența lui 0 (zero) absolut.
- Exemple: temperatura, nivelul de anxietate

# Procesul de măsurare

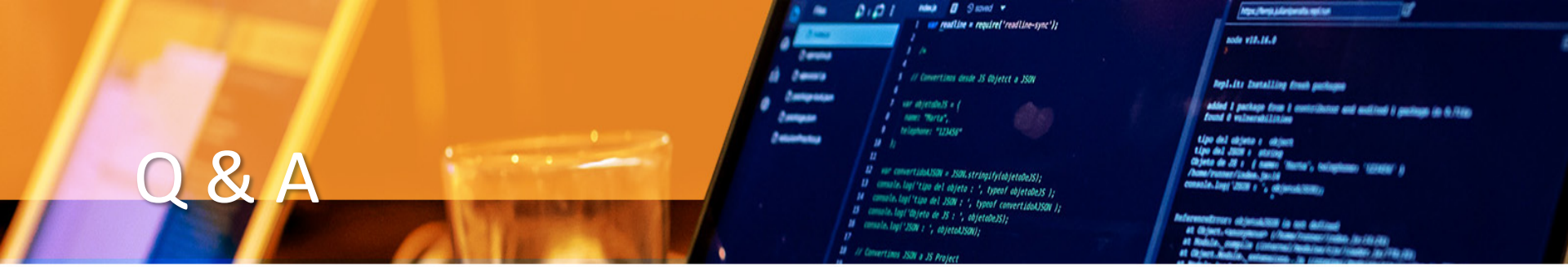
- **Scala de raport** deține cel mai înalt grad de măsurare. Pe lângă egalitatea intervalelor, specifică scalei de interval, cuprinde și valoarea 0 absolut.
- Exemple: numărul de erori la un examen, numărul de absențe la un seminar, viteza, numărul de răspunsuri corecte la un test etc.

# Populație și Eșantion

- **Populația** se referă la întregul grup pe care dorim să îl investigăm.
- Populația reprezintă totalitatea cazurilor care constituie obiectivul de interes al unei cercetări, investigații. Populația se referă la domeniul de interes pe care cercetătorul își propune să îl exploreze.
- Caracteristica numerică a unei populații se numește **parametru**.
- **Eșantionul** reprezintă o selecție de unități de informație extrase din populație cu scopul de a fi studiate.



Q & A



# Erori de măsurare

- Orice proces de măsurare implică o cantitate de eroare.
- $X = T + E$

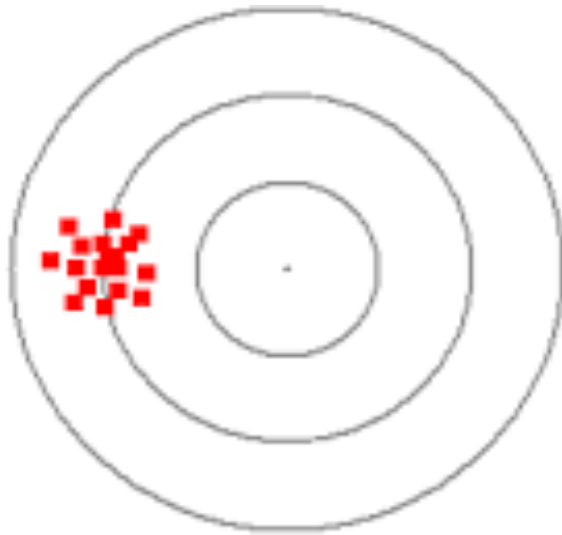
# Erori de măsurare

- **Eroarea aleatorie** este produsă de diferite surse care pot afecta valorile măsurate atât în sens crescător, cât și în sens descrescător.
- Exemplu
  - erorile provocate mediul ambiental la o testare.

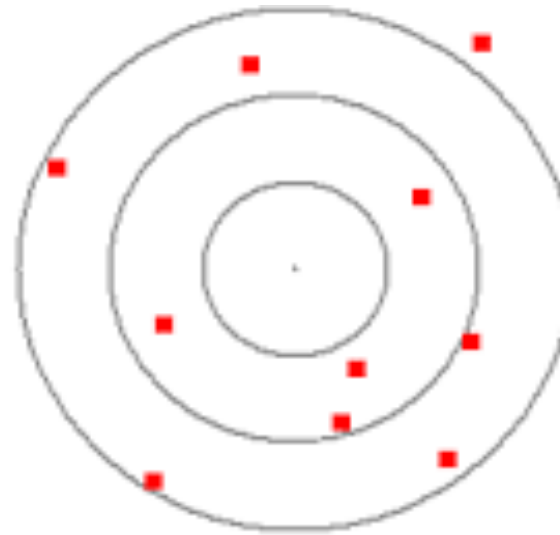
# Erori de măsurare

- **Eroarea sistematică**, spre deosebire de cea aleatorie care putea afecta valorile în orice sens, afectează scorurile într-un singur sens (mai mic sau mai mare) față de scorul adevărat.
- Exemplu
  - un item formulat greșit la examen.

# Erori de măsurare



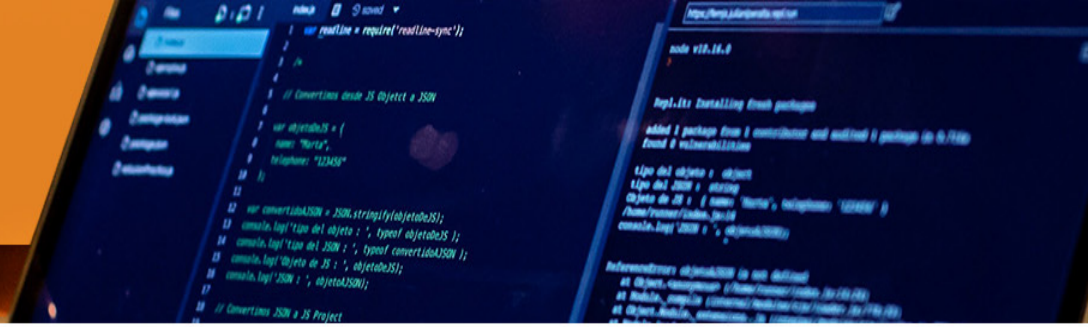
**Eroare sistematică**



**Eroare aleatorie**



# R & R Studio



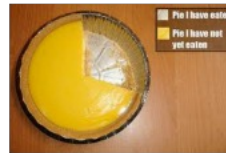
- R for Windows - <https://cran.r-project.org/bin/windows/base/>
- R for IOS - <https://cran.r-project.org/bin/macosx/>
- R Studio - <https://www.rstudio.com/products/rstudio/download/#download>

# Colțul vesel

## Ce face un statistician?



Ce cred prietenii ca fac



Ce crede mama ca fac



Ce crede lumea  
ca fac



Ce crede seful meu  
ca fac

	1	2	3	4	5	6	7	8	9	10
1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
10	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Ce cred eu ca fac



Ce fac de fapt