FIŞA DISCIPLINEI

Bazele statisticii

Anul universitar 2025-2026

1. Date despre program

1.1 Instituţia de învăţământ superior	**Universitatea Babeș-Bolyai Cluj-Napoca**
1.2 Facultatea	**Facultatea de Matematică şi Informatică**
1.3 Departamentul	**Departamentul de Informatică**
1.4 Domeniul de studii	**Informatică**
1.5 Ciclul de studii	**Master**
1.6 Programul de studiu / Calificarea	**Ştiinţa datelor în industrie şi societate (in limba engleză)**
1.7. Forma de învățământ	**Cu frecvenţă**

2. Date despre disciplină

2.1 Denumirea disciplinei (Ro) / (En)	**Bazele statisticii** **Basics of Statistics**			Codul disciplinei		**MME8180**	
2.2 Titularul activităţilor de curs				**Prof. dr. habil. Sanda Micula**			
2.3 Titularul activităţilor de seminar				**Prof. dr. habil. Sanda Micula**			
2.4 Anul de studiu	**1**	2.5 Semestrul	**1**	2.6. Tipul de evaluare	**E**	2.7 Regimul disciplinei	**Obligatorie Fundamentală**

3. Timpul total estimat (ore pe semestru al activităţilor didactice)

3.1 Număr de ore pe săptămână	**4**	Din care: 3.2 curs	**2**	3.3 seminar/laborator	**1 + 1 pr.**
3.4 Total ore din planul de învăţământ	**56**	Din care: 3.5 curs	**28**	3.6 seminar/laborator	**28**
Distribuţia fondului de timp pentru studiul individual (SI) și activități de autoinstruire (AI)					ore
Studiul după manual, suport de curs, bibliografie şi notiţe (AI)					32
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate şi pe teren					12
Pregătire seminarii/laboratoare/proiecte, teme, referate, portofolii şi eseuri					26
Tutoriat					10
Examinări					18
Alte activităţi:					-

3.7 Total ore studiu individual	98
3.8 Total ore pe semestru	154
3.9 Numărul de credite	6

4. Precondiţii (acolo unde este cazul)

4.1 de curriculum	• Noţiuni de bază de algebră • Noţiuni de bază de analiză matematică
4.2 de competenţe	• Gândire logică • Noţiuni de bază de programare

5. Condiţii (acolo unde este cazul)

5.1 De desfăşurare a cursului	• Sală de curs cu tablă mare şi video proiector
5.2 De desfăşurare a seminarului/laboratorului	• Laborator cu calculatoare cu Office şi Matlab instalat

6. Competenţele specifice acumulate

Competenţe profesionale	C 4.1 Definirea conceptelor şi principiilor de bază ale informaticii, precum şi a teoriilor şi modelelor matematice C 4.2 Interpretarea de modele matematice şi informatice (formale) C 4.3 Identificarea modelelor si metodelor adecvate pentru rezolvarea unor probleme reale C 4.4 Încorporarea de modele formale în aplicaţii specifice din diverse domenii
Competenţe transversale	CT1 Aplicarea regulilor de muncă organizată şi eficientă, a unor atitudini responsabile faţă de domeniul didactic-ştiinţific, pentru valorificarea creativă a propriului potenţial, cu respectarea principiilor şi a normelor de etică profesională CT3 Utilizarea unor metode şi tehnici eficiente de învăţare, informare, cercetare şi dezvoltare a capacităţilor de valorificare a cunoştinţelor, de adaptare la cerinţele unei societăţi dinamice și de comunicare în limba română și într-o limbă de circulație internațională

7. Obiectivele disciplinei (reieşind din grila competenţelor acumulate)

7.1 Obiectivul general al disciplinei	• Inţelegerea direcţiilor de bază ale statisticii descriptive şi inferenţiale pentru prelucrarea statistică a datelor • Abilitatea de a folosi facilităţile statistice ale diferitelor softuri
7.2 Obiectivele specifice	• Abilitatea de colectare, organizare şi afişare a datelor pentru a putea fi prelucrate statistic • Dobândirea capacităţii de a lucra cu diferite modele şi algoritmi statistici • Abilitatea de a folosi softuri statistice şi de a interpreta rezultatele

8. Conţinuturi

8.1 **Curs**	Metode de predare	Observaţii
1. **Noţiuni de teoria probabilităţilor.** Variabile aleatoare. Distribuţie de probabilitate. Funcţie de repartiţie. Distribuţii uzuale de tip discret şi de tip continuu.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
2. Caracteristici numerice ale variabilelor aleatoare. Medie, mediană, mod, dispersie, deviaţie standard, covarianţă, coeficient de corelaţie. Proprietăţi. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
3. **Introducere.** Populaţii şi eşantioane, terminologie. Stabilirea obiectivelor. Colectarea datelor. Tehnici statistice de data mining. Reprezentarea grafică a datelor, pie charts, bar graphs. Tabele statistice şi tabele centralizate de frecvenţă. Histograme, poligonul frecvenţelor, grafice stem-and-leaf.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
4. **Statistică descriptivă.** Parametri de distribuţie statistică. Indicatori de poziţie centrală şi indicatori de variabilitate. Varianţă şi deviaţie standard, interpretare. Coeficient de variaţie.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
5. Percentile, decile, cuartile, interval intercuartilic. Outliers, detectarea outliers, regula 3σ. Boxplots.	• Expunere interactivă • Explicaţie • Conversaţie	

Noţiuni de teoria selecţiei. Funcţii de selecţie. Eroare standard. Exemple şi aplicaţii. Interpretarea datelor.	• Demonstratie didactică	
6. Statistică inferenţială. Noţiuni de teoria estimaţiei. Estimatori punctuali, proprietăţi de bază.. Metoda momentelor. Exemple. Estimarea erorii standard.	• Expunere interactivă • Explicaţie • Conversaţie • Demonstratie didactică	
7. Distribuţia normală şi normală standard, cuantile Z. Distribuţia Student T şi cuantile T. Intervale de încredere uni- şi bilaterale. Estimarea mediei şi proporţiei prin intervale de încredere. Exemple. Determinarea volumului de selecţie.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
8. Statistici bivariate, proporţie centralizată. Estimarea diferenţei proporţiilor. Intervale de încredere pentru date pereche. Varianţa centralizată a două selecţii. Estimarea diferenţei mediilor prin intervale de încredere. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
9. Testarea ipotezelor statistice. Noţiuni de bază, cadru general. Regiune critică. Erori de tipul I şi de tipul II. Testarea semnificaţiei şi valoarea P a testului.	• Expunere interactivă • Explicaţie • Conversaţie • Demonstratie didactică	
10. Teste Z pentru medie. Determinarea volumului de selecţie. Exemple. Teste T (Student) pentru medie. Teste pentru proporţii.	• Expunere interactivă • Explicaţie • Conversaţie	
11. Distribuţia Fisher F şi cuantile F. Teste F pentru raportul varianţelor. Teste pentru diferenţa mediilor. Teste pentru date-pereche. Calculul probabilităţilor erorilor de tipul II. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie • Demonstratie didactică	
12. Corelaţie şi regresie. Statistici pentru date bidimensionale. Grafice scatter plots şi time plots. Medie condiţionată. Curbe de regresie. Estimare prin metoda celor mai mici pătrate. Regresie liniară. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie	
13. Modelare overfitting. Regresie polinomială. Exemple şi aplicaţii. Modele de fitting. Modele ANOVA cu un singur factor şi R-pătrat. Predicţii. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie • Descriere	
14. Modele ANOVA cu mai mulţi factori şi testul F. Coeficient de determinare, R-pătrat ajustat. Predictori categoriali şi variabile fictive. Corelaţie semnificativă. Exemple.	• Expunere interactivă • Explicaţie • Conversaţie • Demonstratie didactică	

Bibliografie

1. Micula, S., Probability and Statistics for Computational Sciences, Cluj University Press, 2009.
2. Miller, J.D., Statistics for Data Science, Packt Publishing, Birmingham, UK, 2017.
3. Bruce P., Bruce A., Practical Statistics for Data Scientists, 50 Essential Concepts, O'Reilly Media, CA, USA, 2017.
4. Baron, M., Probability and Statistics for Computer Scientists, 3rd edition, CRC Press, Taylor and Francis, Boca Raton, FL, USA, 2019.

5. Milton, J.S., Arnold, J. C., Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences, 3rd Edition. McGraw-Hill, New York, 1995.
6. Gentle, J. E., Elements of Computational Statistics, Springer-Verlag, New York, 2002.

8.2 **Seminar/Laborator**	Metode de predare	Observaţii
1. Introducere în Matlab (Octave).	• Expunere interactivă • Explicaţie • Conversaţie	Seminarul este structurat astfel: 2 ore din două în două săptămâni
2. Variabile aleatoare şi caracteristicile lor. Toolbox-ul de Statistică şi Machine Learning din Matlab.	• Expunere interactivă • Explicaţie • Conversaţie • Studiu individual şi în grup	
3. Statistică descriptivă. Tabelul centralizat al frecvenţelor. Calcularea parametrilor de distribuţie statistică. Reprezentarea grafică a datelor. Histograma, poligonul frecvenţelor, boxplot.	• Expunere interactivă • Conversaţie • Sinteză • Studiu individual şi în grup	
4. Intervale de încredere. Interpretarea rezultatelor.	• Expunere interactivă • Explicaţie • Conversaţie • Studiu individual şi în grup	
5. Testarea ipotezelor statistice. Interpretarea rezultatelor.	• Expunere interactivă • Explicaţie • Conversaţie • Studiu individual şi în grup	
6. Corelaţie şi regresie.	• Expunere interactivă • Explicaţie • Conversaţie • Studiu individual şi în grup	
7. Prezentare. Proiect de analiză statistică a datelor.	• Conversaţie • Studiu individual şi în grup	

Bibliografie

1. Micula, S., Probability and Statistics for Computational Sciences, Cluj University Press, 2009.
2. Miller, J.D., Statistics for Data Science, Packt Publishing, Birmingham, UK, 2017.
3. Bruce P., Bruce A., Practical Statistics for Data Scientists, 50 Essential Concepts, O'Reilly Media, CA, USA, 2017.
4. Baron, M., Probability and Statistics for Computer Scientists, 3rd edition, CRC Press, Taylor and Francis, Boca Raton, FL, 2014.
5. Milton, J.S., Arnold, J. C., Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences, 3rd Edition. McGraw-Hill, New York, 1995.
6. Gentle, J. E., Elements of Computational Statistics, Springer-Verlag, New York, 2002.

9. Coroborarea conţinuturilor disciplinei cu aşteptările reprezentanţilor comunităţii epistemice, asociaţiilor profesionale şi angajatori reprezentativi din domeniul aferent programului

- Cursul respectă recomandările curicullare IEEE şi ACM pentru studiile în informatică
- Cursul există în programa de studiu a numeroase facultăţi de profil din ţară şi străinătate
- Cunoștințele şi abilitățile acumulate oferă studenților masteranzi o bază pentru lansarea în cercetarea ştiinţifică.
- Capacitatea de a face analize şi prelucrări statistice dobândită la acest curs e folositoare în orice carieră pe care studenții masteranzi o vor alege pe viitor.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	- însușirea principiilor de bază din Statistică, cu accent pe aplicaţii din viaţa reală - aplicarea corectă a noțiunilor predate la curs la diferite probleme - utilizarea instrumentelor de analiză statistică exploratorie	**Examen scris**	**70%**
10.5 Seminar/laborator	- aplicarea conceptelor şi algoritmilor de la curs la probleme practice - rezolvarea unor probleme statistice numerice în Excel şi Matlab - prelucrarea statistică a datelor	- participarea la discuții, rezolvarea şi implementarea problemelor de-a lungul semestrului - prezentarea individuală a unor soluții - **prezentarea unui proiect de analiză statistică a datelor**	**30%**
10.6 Standard minim de performanţă			
O notă minimă de 5 la **fiecare** din activitățile menționate mai sus (examen scris, evaluare la seminar/ laborator)			

11. Etichete ODD (Obiective de Dezvoltare Durabilă / Sustainable Development Goals)[1]

	Eticheta generală pentru Dezvoltare durabilă							

Data completării	Semnătura titularului de curs	Semnătura titularului de seminar
29.04.2025	Prof. dr. habil. Sanda Micula	Prof. dr. habil. Sanda Micula

Data avizării în departament

Semnătura directorului de departament

Prof. dr. Andrei Mărcuş

[1] Păstrați doar etichetele care, în conformitate cu *Procedura de aplicare a etichetelor ODD în procesul academic*, se potrivesc disciplinei și ștergeți-le pe celelalte, inclusiv eticheta generală pentru *Dezvoltare durabilă* - dacă nu se aplică. Dacă nicio etichetă nu descrie disciplina, ștergeți-le pe toate și scrieți "*Nu se aplică.*".