# Laboratorul 11 - 2017

**I.** *Elemente de statistică descriptivă*

```
mean(x) %media
var(x) % dispersia normalizata cu n-1
var(x,1) %dispersia normalizata cu n
std(x) %abaterea standard
median(x) %mediana
prctile(x,[25, 75]) %cuartila inferioara, cuartila superioara
```

Notaţii: ▶ $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ media de selecţie
▶ $\tilde{S}_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2$ varianţa de selecţie
▶ $\hat{F}_n : \mathbb{R} \to [0,1]$ este funcţia de repartiţie empirică, corespunzătoare variabilelor de selecţie $X_1, \ldots, X_n$

$$\hat{F}_n(x) = \frac{\#\{i \in \{1,...,n\} : X_i \leq x\}}{n}, x \in \mathbb{R}.$$

**Proprietăţi:** Fie $X_1, \ldots, X_n$ variabile aleatoare independente, care au aceeaşi distribuţie, şi au media $m = E(X_n)$ şi varianţa $\sigma^2 = Var(X_n)$ pentru orice $n \geq 1$. Dacă
▶ $X_1, \ldots, X_n \sim N(m, \sigma^2)$
sau
▶ $n > 30$ şi $X_1, \ldots, X_n$ urmează o altă distribuţie decât cea normală,

$$\Longrightarrow \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \text{ şi } \frac{\bar{X}_n - m}{\frac{\tilde{S}_n}{\sqrt{n}}} \sim Student(n-1).$$

**Aplicaţie:** Se dau $m \in \mathbb{R}$ şi $\sigma > 0$. Să se genereze $n$ (=100,1000...) variabile aleatoare independente care urmează
a) legea $N(m, \sigma^2)$
b) o altă lege de distribuţie cu media $m$ şi varianţa $\sigma^2$.
1) Să se deseneze histograma frecvenţelor relative corespunzătoare datelor generate pentru variabilele aleatoare $Z_n = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}}$. Pe aceeaşi figura să se reprezinte grafic funcţia de densitate a distribuţiei normale standard (*normpdf*).
2) Să se deseneze funcţia de repartiţie empirică corespunzătoare variabilelor aleatoare $Z_n$, adică

$$\hat{F}_n(z) = \frac{\#\{i \in \{1,...,n\} : Z_i \leq z\}}{n}, z \in \mathbb{R}.$$

Pe acelaşi desen să se reprezinte grafic funcţia de repartiţie a distribuţiei normale standard (*normcdf*).
3) Să se deseneze histograma frecvenţelor relative corespunzătoare datelor generate pentru variabilele aleatoare $V_n = \frac{\bar{X}_n - m}{\frac{\tilde{S}_n}{\sqrt{n}}}$. Pe aceeaşi figură să se reprezinte grafic funcţia de densitate a distribuţiei $Student(n-1)$ (*tpdf*).
4) Să se deseneze funcţia de repartiţie empirică corespunzătoare variabilelor aleatoare $V_n$, adică

$$\hat{F}_n(v) = \frac{\#\{i \in \{1,...,n\} : V_i \leq v\}}{n}, v \in \mathbb{R}.$$

Pe aceeaşi figură să se reprezinte grafic funcţia de repartiţie a distribuţiei $Student(n-1)$ ($tcdf$).
5) Să se aplice elementele de statistică descriptivă datelor generate şi să se compare cu parametrii teoretici ai distribuţiilor considerate ($N(0,1)$, respectiv $Student(n-1)$).

**II. Metode Monte-Carlo pentru estimarea unor integrale**

Se consideră următoarele funcţii:

**a)** $g:[0,1]\to[0,\infty)$, $g(x)=x^3$, $x\in[0,1]$.

**b)** $g:[2,5]\to[0,\infty)$, $g(x)=\frac{1}{\sqrt{x-1}}$, $x\in[2,5]$.

**c)** $g:[-1,2]\to[0,\infty)$, $g(x)=\begin{cases}\frac{1}{1+x^2}, & x\in[-1,0]\\ \frac{1}{x^2}\sin^2\frac{x}{1-x}, & x\in(0,1)\\ \sqrt{2x-x^2}, & x\in[1,2].\end{cases}$

**1)** Realizaţi pentru fiecare funcţie de mai sus un program în Octave/Matlab care returnează valoarea $g(x)$ pentru $x$ dat din domeniul de definiţie al funcţiei $g$.

**2)** Implementaţi în Octave/Matlab fiecare din metodele descrise mai jos pentru funcţiile date. Comparaţi rezultatele obţinute cu metode diferite pentru aceeaşi funcţie.

BONUS: Folosind metodele de tip Monte-Carlo, să se estimeze $\Phi(1)$ şi $\Phi(2)$ dacă $\Phi:[0,\infty)\to[0,\infty)$, $\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_0^x e^{-\frac{t^2}{2}}\,dt$.

*Integrare Monte-Carlo - versiunea I*

Fie $g:[a,b]\to[0,\infty)$ o funcţie integrabilă dată şi $M>0$ astfel încât $g(x)\le M$, oricare ar fi $x\in[a,b]$. Considerăm următorii paşi pentru aproximarea integralei $\int_a^b g(x)\,dx$:

- se generează $N$ numere aleatoare uniform distribuite pe intervalul $[a,b]$:

$$x_1,x_2,\ldots,x_N\in[a,b],$$

unde $N\in\mathbb{N}$ este dat ($N=100,1000,\ldots$).

- se generează $N$ numere aleatoare uniform distribuite pe intervalul $[0,M]$:

$$y_1,y_2,\ldots,y_N\in[0,M].$$

- se calculează numărul $P$ de perechi $(x_i,y_i)$ care verifică inegalitatea: $y_i\le g(x_i)$, pentru $i=1,2,\ldots,N$.
- se calculează valoarea aproximativă $\mathcal{A}$ a integralei $\int_a^b g(x)\,dx$: $\mathcal{A}=M(b-a)\dfrac{P}{N}$.

*Integrare Monte-Carlo - versiunea II*

Fie $g:[a,b]\to[0,\infty)$ o funcţie integrabilă dată. Considerăm următorii paşi pentru aproximarea integralei $\int_a^b g(x)\,dx$:

- se generează $N$ numere aleatoare uniform distribuite pe intervalul $[a,b]$:

$$x_1,x_2,\ldots,x_N\in[a,b],$$

unde $N\in\mathbb{N}$ este dat ($N=100,1000,\ldots$).

- se notează: $y_i=(b-a)g(x_i)$, pentru $i=1,2,\ldots,N$.
- se calculează valoarea aproximativă $\mathcal{A}$ a integralei $\int_a^b g(x)\,dx$: $\mathcal{A}=\dfrac{1}{N}\big(y_1+y_2+\ldots+y_N\big)$.

*Integrare numerică - regula trapezului*

Fie $g:[a,b]\to[0,\infty)$ o funcţie integrabilă dată. Considerăm următorii paşi pentru aproximarea integralei $\int_a^b g(x)\,dx$:

• se consideră o diviziune echidistantă a intervalului $[a, b]$:

$$a = x_1 < x_2 < \cdots < x_N < x_{N+1} = b,$$

unde $N \in \mathbb{N}$ este dat ($N = 100, 1000, \ldots$).

• se notează: $y_i = g(x_i)$, pentru $i = 1, 2, \ldots, N + 1$.

• se calculează aria $\mathcal{A}_i$ a trapezului cu vârfurile în punctele de coordonate $(x_i, 0)$, $(x_{i+1}, 0)$, $(x_{i+1}, y_{i+1})$ şi $(x_i, y_i)$, pentru $i = 1, 2, \ldots, N$.

• se calculează valoarea aproximativă $\mathcal{A}$ a integralei $\int_a^b g(x)\, dx$: $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 + \ldots + \mathcal{A}_N$.