Laborator 4 – Metode moderne de calcul şi simulare
Autor: Lect. dr. Bianca Mogoş

1 Generarea numerelor uniforme în Matlab

Fie $(\Omega, \mathcal{B}, P)$ un câmp de probabilitate.

Definiţia 1.1. *(Variabilă aleatoare uniformă)*
Spunem că variabila aleatoare continuă $U : \Omega \to \mathbb{R}$ este repartizată uniform pe intervalul $[a, b]$ şi scriem $U \sim \mathcal{U}(a, b)$ dacă densitatea sa de repartiţie este de forma:

$$f(u) = \begin{cases} \frac{1}{b-a}, & \text{dacă } u \in [a, b] \\ 0, & \text{altfel} \end{cases} \tag{1.1}$$

Remarca 1.2. *(Observaţii referitoare la repartiţia uniformă şi v.a. uniforme)*

- Funcţia de repartiţie $F(u)$ a v.a. $U \sim \mathcal{U}(a, b)$ este definită prin

$$F(u) = \begin{cases} 0, & \text{dacă } u \leq a \\ \frac{u-a}{b-a}, & \text{dacă } u \in (a, b) \\ 1, & \text{dacă } u \geq b \end{cases} \tag{1.2}$$

- Probabilitatea ca v.a. uniformă $U \sim \mathcal{U}(a, b)$ să cadă într-un interval $[u_1, u_2]$, $a \leq u_1 < u_2 \leq b$, este proporţională cu lungimea intervalului:

$$P(u_1 < U \leq u_2) = \frac{u_2 - u_1}{b - a}. \tag{1.3}$$

- Pentru a genera numere aleatoare uniform distribuite pe un interval [a,b], şi scriem $X \sim \mathcal{U}(a, b)$, pornind de la numere generate uniform pe intervalul [0,1] se poate folosi transformarea

$$X = (b - a) \cdot U + a, \tag{1.4}$$

unde $U \sim \mathcal{U}(0, 1)$.

Remarca 1.3. *(Observaţii referitoare la generarea v.a. uniforme în Matlab)*

- În Matlab există funcţia **rand** pentru generarea variabilelor aleatoare uniforme.

- **rand**(n), unde n este un număr natural, returnează o matrice de dimensiune $n \times n$ având ca elemente numere aleatoare uniform distribuite între 0 şi 1.

- **rand**(m,n), unde m,n sunt numere naturale, returnează o matrice de dimensiune $m \times n$ având ca elemente numere aleatoare uniform distribuite între 0 şi 1.

- O secvenţă de numere aleatoare generată în Matlab depinde de "sămânţa" sau starea generatorului. Starea este resetată la valoarea implicită în momentul pornirii Matlab-ului, astfel aceeaşi secvenţă de variabile aleatoare este generată la o nouă pornire a Matlab-ului. Acesta poate fi un avantaj în situaţiile în care analistul are nevoie să reproducă rezultatele unei simulări pentru a verifica anumite concluzii.

- **rng(SD)** setează sămânţa generatorului pe baza întregului nenegativ SD.

- **rng('shuffle')** setează sămânţa generatorului pe baza momentului curent de timp.

- **rng('default')** resetează setările generatorului la valorile iniţiale, şi anume, se foloseşte generatorul Mersenne Twister cu sămânţa 0.

Exemplul 1.4. *(Exemplu care ilustrează utilizarea funcţiei **rand**)*

% Generăm un vector de numere aleatoare pe intervalul [0,1].
x = rand(1,1000);
% Histograma eşantionului generat în x.
[N,X] = hist(x,15);
% x: mulţimea eşantion
% 15 reprezintă numărul de dreptunghiuri ale histogramei
% N: vector conţinând numărul de elemente din fiecare dintre dreptunghiuri.
% X: vector conţinând centrele dreptunghiurilor
% Folosirea funcţiei "bar" pentru reprezentarea grafică a histogramei.
bar(X,N,1,'w')
title('Histograma asociata unei variabile aleatoare uniforme')
xlabel('X')
ylabel('Frecventa absoluta')

Histograma rezultată este prezentată în Figura 1.

Aplicaţii

Fie $U \sim \mathcal{U}(a,b)$, $a,b \in \mathbb{R}, a < b$.

1. Scrieţi o funcţie care generează o matrice A de numere aleatoare uniforme pe intervalul [a,b]. Antetul funcţiei va fi

$$\text{function } [A] = my\_rand(a,b,m,n)$$

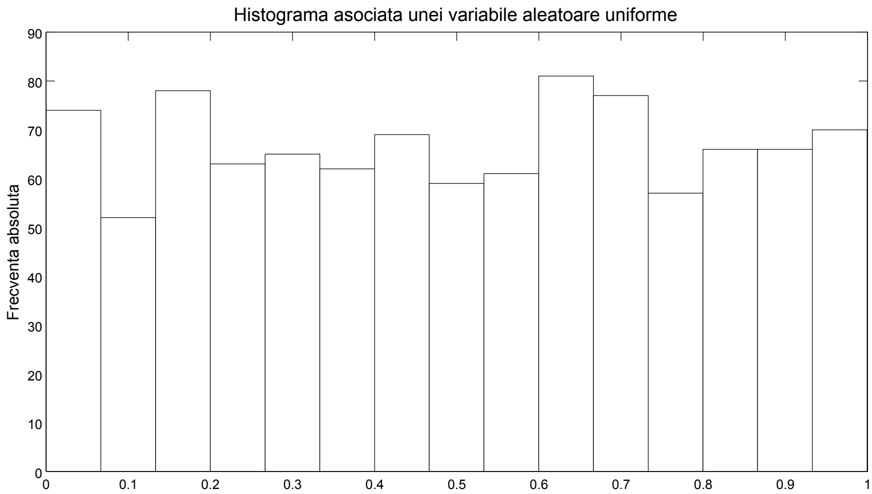

Figura 1: 1 – Histograma asociată unui eşantion de numere aleatoare uniform distribuite

unde m şi n reprezintă numărul de linii, respectiv numărul de coloane ale matricei A returnată de funcţia "my_rand".

2. Aplicaţi funcţia "my_rand" pentru a genera o mulţime de numere aleatoare uniforme pe $[-2, 2]$ având 1000 de elemente.

3. Reprezentaţi grafic histograma asociată mulţimii de valori de selecţie generate la punctul 2.

4. Calculaţi $P(U \in (i, i + 0.1))$ pentru $i = -2, -1.9, \ldots, 1.8, 1.9$. Interpretaţi rezultatele obţinute.

2 Metode generale de simulare a variabilelor aleatoare – Metoda inversă

Teorema 2.1. *(Teorema lui Hincin)*

Fie $U \sim \mathcal{U}([0, 1])$ şi $F(x)$ o funcţie de repartiţie inversabilă, de tip continuu. Atunci variabila aleatoare

$$X = F^{-1}(U) \tag{2.1}$$

este o variabilă aleatoare continuă cu funcţia de repartiţie $F(x)$.

Remarca 2.2. *(Observaţii referitoare la metoda inversă)*

- este introdusă ca o consecinţă directă a teoremei lui Hincin
- se aplică în cazul în care funcţia de repartiţie se poate inversa uşor
- dacă am putea produce valorile de selecţie $u_1, u_2, \ldots, u_n$ asupra v.a. $U \sim \mathcal{U}(0,1)$ şi am cunoaşte funcţia de repartiţie F a variabilei X atunci am putea produce valorile de selecţie $x_1, x_2, \ldots, x_n$ asupra lui X cu formula $x_i = F^{-1}(u_i), 1 \leq i \leq n$

Intrare	$F(x)$: funcţia de repartiţie a variabilei X pe care ne propunem să o simulăm
Pas 1	Se generează o valoare de selecţie u uniformă pe [0,1]
Pas 2	Se determină expresia inversei funcţiei de repartiţie: $F^{-1}(u)$
Pas 3	Se obţine valoarea de selecţie dorită $x = F^{-1}(u)$
Ieşire	Valoarea de selecţie, x, a v.a. X

Tabela 1: Algoritm pentru simularea v.a. continue folosind metoda inversă

Repartiţia	**Densitatea** f	**Inversa** F^{-1}
$Exp(\lambda)$, $\lambda > 0$	$f(x) = \lambda e^{-\lambda x}, x > 0$	$x = -\frac{1}{\lambda}\ln(u)$
$Weib(0,1,\nu)$, $\nu > 0$	$f(x) = \nu x^{\nu-1}e^{-x^{\nu}}$	$x = (-\ln(u))^{1/\nu}$
Cauchy	$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, x \in \mathbb{R}$	$x = \tan\pi(u - 1/2)$
Arcsin	$f(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}, x \in (-1,1)$	$x = \sin\pi(u - 1/2)$

Tabela 2: Variabile aleatoare continue care pot fi simulate folosind metoda inversă

Aplicaţii

1. Fie $X \sim Exp(\lambda)$.

   (a) Scrieţi o funcţie care generează n valori de selecţie asupra variabilei X, folosind metoda inversă.

   (b) Apelaţi funcţia definită la punctul (a) pentru $\lambda = 1$ şi $n = 1000$.

   (c) Realizaţi histograma asociată mulţimii de valori de selecţie generate la punctul (b). Scalaţi înălţimile dreptunghiurilor astfel încât suma ariilor tuturor dreptunghiurilor să fie 1.

   (d) Reprezentaţi grafic curba densităţii de repartiţie a variabilei X. Curba densităţii va fi reprezentată în aceeaşi fereastră cu graficul histogramei.

2. Fie X o variabilă aleatoare având repartiţia $Exp(\lambda)$. Cerinţe:

   (a) Arătaţi că funcţia de repartiţie a variabilei $Y = e^X$ este

$$F(x) = \begin{cases} 0, & \text{dacă } x \leq 1 \\ 1 - x^{-\lambda}, & \text{dacă } x > 1. \end{cases} \tag{2.2}$$

   (b) Scrieţi o funcţie pentru simularea variabilei Y folosind funcţia de repartiţie determinată la punctul (a) şi metoda inversă.

   Se rezolvă punctele de mai jos şi se interpretează rezultatele considerând pe rând parametrul λ cu valorile 2, 5, 7 şi 10.

   (c) Apelând funcţia implementată la punctul (b), generaţi mulţimi de valori de selecţie asupra variabilei Y conţinând 1000 de valori.

   (d) Realizaţi histograma asociată mulţimii generate la punctul (c). Adăugaţi în acest grafic curba densităţii de repartiţie a variabilei Y.

   (e) Scrieţi o funcţie pentru generarea a 1000 de valori de selecţie asupra variabilei Y folosind definiţia variabilei (adică, $Y = e^X$) şi mulţimea de valori de selecţie generată la punctul 1(b).

   (f) Realizaţi histograma asociată mulţimii generate la punctul (e). Adăugaţi în acest grafic curba densităţii de repartiţie a variabilei Y.

   (g) Comparaţi histogramele obţinute la punctele (d) şi (f).

Bibliografie

[Martinez, Martinez (2002)] W. L. Martinez, A. R. Martinez (2002), *Computational Statistics Handbook with MATLAB*, Chapman & Hall/CRC, Boca Raton London New York Washington, D.C.

[Văduva (2004)] I. Văduva (2004), *Modele de simulare: note de curs*, Editura Universităţii din Bucureşti, Bucureşti