Universitatea din Bucureşti
Facultatea de Matematică şi Informatică

METODE MODERNE DE CALCUL ŞI SIMULARE

Cursuri nr. 5-6 – Anul III – Tehnologia Informaţiei

Asist. dr. Bianca Mogoş

Bucureşti, Februarie 2014

Cuprins

Capitolul 1

Generarea variabilelor aleatoare

1.1 Metode generale pentru generarea variabilelor aleatoare

1.1.1 Numere aleatoare continue

Majoritatea metodelor de generare a v.a. se bazează pe generarea unor numere aleatoare uniform distribuite pe intervalul (0,1). Notăm un număr aleator uniform distribuit cu $U \sim \mathcal{U}(0, 1)$. Datorită calculatorului avem posibilitatea de a genera foarte uşor numere aleatoare uniforme. Totuşi, trebuie cunoscut faptul că numerele generate de calculator sunt pseudo-aleatoare, deoarece acestea sunt generate cu un algoritm determinist.

Observaţie 1.1.1.

- 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.

- Folosind sintaxa **rand('state',0)** Matlab-ul resetează generatorul la starea iniţială.

- Se foloseşte sintaxa **rand('state',**j**)** pentru a seta generatorul la starea j.

- Pentru a obţine vectorul de stări se apelează **S = rand('state')**, S va reprezenta vectorul conţinând cele 35 de stări posibile.

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

$$X = (b-a) \cdot U + a, \tag{1.1.1}$$

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

Exemplul 1.1.2.

În acest exemplu se va ilustra utilizarea funcţiei **rand**.

% Generam un vector de numere aleatoare pe intervalul (0,1).
x = rand(1,1000);
% Histograma eşantionului generat in 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 1.1.

1.1.2 Funcţii de o variabilă aleatoare

Fie X o v.a. $X : \Omega \to \mathbb{R}$ care ia valori în $D \subset \mathbb{R}$ şi $\phi : D \to \mathbb{R}$, ϕ continuă. Atunci v.a. $Y = \phi(X) : \Omega \to \phi(X) \subset \mathbb{R}$ are repartiţia dată de

$$P(\{\omega \in \Omega | Y(\omega) \in A\} = P(\big\{\omega \in \Omega | X(\omega) \in \phi^{-1}(A)\big\})) \tag{1.1.2}$$

pentru orice $A \subset \phi(D)$ şi $\phi^{-1}(A) = \{z | \phi(z) \in A\}$.

Proprietăţi 1.1.3.

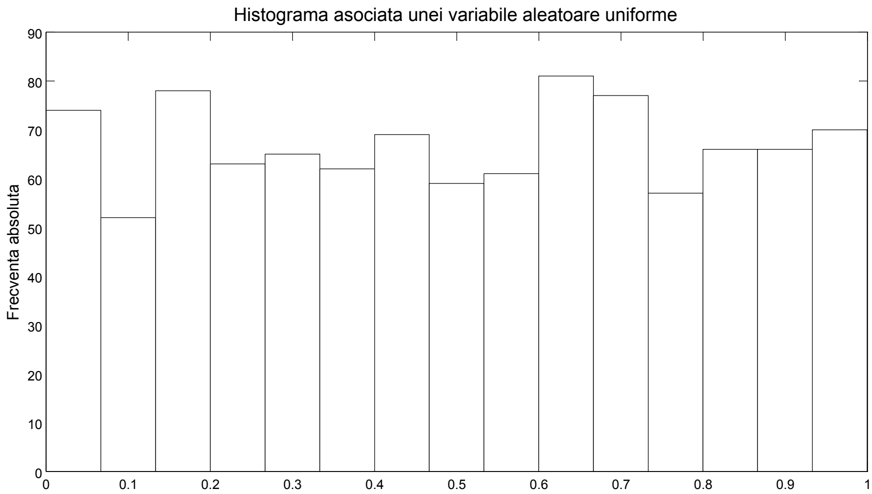

Figura 1.1: Histograma asociată unui eşantion de numere aleatoare uniform distribuite

Dacă $X \sim N(m, \sigma^2)$ atunci $Y = e^X$ are densitatea de repartiţie:

$$f_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - m)^2}{2\sigma^2}}, y > 0 \tag{1.1.3}$$

Demonstraţie

$$F_Y(y) = P(\{\omega \in \Omega | Y(\omega) \leq y\}) = P(e^X \leq y) = P(X \leq \ln y) = F_X(\ln y) =$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\ln y} e^{-\frac{(t-m)^2}{2\sigma^2}} dt.$$

Definiţia 1.1.4. *(Variabilă aleatoare lognormală)*

Variabila aleatoare Y având densitatea de probabilitate (1.1.3) se numeşte lognormală.

Proprietăţi 1.1.5.

Dacă $X \sim U(0, 1)$ atunci variabila aleatoare

$$Y = \left(-\frac{1}{a} \ln(1 - X)\right)^{1/b}$$

cu $a, b > 0$ are densitatea de probabilitate

$$f(y) = aby^{b-1}e^{-ay^b}, \quad 0 < y < \infty.$$

Definiţia 1.1.6. *(Variabilă aleatoare Weibull)*

O variabilă aleatoare cu densitatea de probabilitate (1.1.5) se numeşte Weibull cu parametrii a şi b şi este notată $W(a, b)$.

Proprietăţi 1.1.7.

Fie X o variabilă aleatoare cu funcţia de repartiţie inversabilă $F(X)$ de tip continuu care este strict crescătoare. Atunci variabila aleatoare $Y = F(X)$ este repartizată uniform pe $(0, 1)$.

Demonstraţie
Funcţia de repartiţie a variabilei Y este

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y), 0 < y < 1) = F(F^{-1}(y)) = y,$$

$y \in (0, 1)$, ceea ce arată că $Y \sim U(0, 1)$.
Un rezultat pentru reciproca propoziţiei (1.1.7) este dat în următoarea propoziţie.

Teorema 1.1.8. *(Teorema lui Hincin)*

Fie $Y \sim U(0, 1)$ şi $F(x)$ o funcţie de repartiţie inversabilă, de tip continuu cu $F(a) = 0$ şi $F(b) = 1$ şi F strict crescătoare pe (a, b). Atunci v.a.

$$X = F^{-1}(Y) \tag{1.1.4}$$

este o v.a. continuă cu funcţia de repartiţie F.

Demonstraţie
Funcţia de repartiţie a v.a. X este

$$P(X \leq x) = P(F^{-1}(Y) \leq x) = P(Y \leq F(x))$$

deoarece F este strict crescătoare.
Cum $Y \sim U(0, 1)$ rezultă $F_Y(y) = y$ pentru $y \in (0, 1)$. Obţinem astfel

$$P(X \leq x) = F(x).$$

1.1.3 Metoda inversă

Această metodă este introdusă ca o consecinţă directă a teoremei 1.1.8 (Teorema lui Hincin). Ea se aplică în cazul în care funcţia de repartiţie se poate inversa uşor. Astfel, 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 $x_i = F^{-1}(u_i), 1 \leq$

Algoritm 1.1: Metoda inversă pentru variabile aleatoare continue

Intrare	F: 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

$i \leq n$. Dacă funcţia F^{-1} se poate calcula atunci valorile de selecţie x_i pot fi generate folosind Algoritmul 1.1.

Spunem că acest algoritm simulează o selecţie de volum n asupra v.a. X. Presupunem că putem genera valori de selecţie $u_i, 1 \leq i \leq n$ uniforme pe (0,1) şi independente atunci şi valorile de selecţie $x_i = F^{-1}(u_i), 1 \leq i \leq n$ sunt independente. De aceea, este suficient să dăm algoritmul pentru simularea unei singure valori de selecţie, urmând să repetăm Paşii 1–3 de n ori.

În Tabelul 1.1 se prezintă câteva repartiţii de probabilitate care se pot simula uşor folosind metoda inversă. În formulele lui f din tabel se precizează numai valorile lui x pentru care $f(x) > 0$, presupunându-se că $f(x) = 0$ în rest.

Tabel 1.1: Variabile aleatoare care pot fi simulate folosind metoda inversă

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

Metoda inversă se poate adapta pentru a dezvolta un algoritm pentru simularea unei variabile discrete definite prin repartiţia

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix}, \sum_{i=1}^{m} p_i = 1, x_1 < x_2 < \dots < x_m. \tag{1.1.5}$$

Funcţia de repartiţie a variabilei X este dată de

$$F(x) = \begin{cases} 0 & \text{dacă} \quad x < x_1 \\ p_1 & \text{dacă} \quad x_1 \leq x < x_2 \\ p_1 + p_2 & \text{dacă} \quad x_2 \leq x < x_3 \\ \dots & \dots \quad \dots \\ p_1 + p_2 + \dots + p_k & \text{dacă} \quad x_k \leq x < x_{k+1} \\ \dots & \dots \quad \dots \\ 1 & \text{dacă} \quad x \geq x_m \end{cases} \tag{1.1.6}$$

Se obţine o valoare de selecţie a v.a. discrete X generând un număr aleator uniform U şi alegând numărul aleator X conform regulii

$$X = x_i \quad \text{dacă } F(x_{i-1}) < U \leq F(x_i). \tag{1.1.7}$$

Exemplul 1.1.9.

Vrem să generăm o v.a. discretă X cu repartiţia

$$X : \begin{pmatrix} 0 & 1 & 2 \\ 0.3 & 0.2 & 0.5 \end{pmatrix} \tag{1.1.8}$$

Funcţia de repartiţie este dată de

$$F(x) = \begin{cases} 0 & \text{dacă} \quad x < 0 \\ 0.3 & \text{dacă} \quad 0 \leq x < 1 \\ 0.5 & \text{dacă} \quad 1 \leq x < 2 \\ 1 & \text{dacă} \quad x \geq 2 \end{cases} \tag{1.1.9}$$

Se generează valori de selecţie asupra v.a. X conform regulilor

$$X = \begin{cases} 0 & U \leq 0.3 \\ 1 & 0.3 < U \leq 0.5 \\ 2 & 0.5 < U \leq 1 \end{cases} \tag{1.1.10}$$

De exemplu, dacă variabila aleatoare uniformă $U = 0.78$ atunci obţinem valoare de selecţie $x = 2$.

Algoritm 1.2: Metoda inversă pentru variabile aleatoare discrete

Intrare	Repartiţia variabilei X pe care ne propunem să o simulăm $P(X = x_i) = p_i, \sum_{i=1}^{m} p_i = 1, x_1 < x_2 < \ldots < x_m.$
Pas 1	Se generează o valoare de selecţie U uniformă pe (0,1).
Pas 2	Dacă $U \leq p_1$ atunci $X = x_1$
Pas 3	Altfel dacă $U \leq p_1 + p_2$ atunci $X = x_2$
Pas 4	Altfel dacă $U \leq p_1 + p_2 + p_3$ atunci $X = x_3$
Pas 5	… Altfel dacă $U \leq p_1 + p_2 + \ldots + p_m$ atunci $X = x_m$
Ieşire	Valoarea de selecţie, X, a v.a. X

1.1.4 Metoda compunerii sau amestecării

Această metodă se aplică variabilelor aleatoare X a căror repartiţie de probabilitate satisface următoarea definiţie.

Definiţia 1.1.10. *(Amestecare discretă şi amestecare continuă)*

Funcţia de repartiţie $F(x)$ este o amestecare (compunere sau mixtură) discretă a mulţimii de funcţii de repartiţie $\{F_i(x)\}_{1\leq i\leq m}$ cu repartiţia discretă

$$J : \begin{pmatrix} 1 & 2 & \ldots & m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}, \sum_{i=1}^{m} p_i = 1, \tag{1.1.11}$$

dacă

$$F(x) = \sum_{i=1}^{m} p_i F_i(x). \tag{1.1.12}$$

Funcţia de repartiţie $F(x)$ este o amestecare continuă a familiei de funcţii de repartiţie $\{G(x, Y)\}_{Y\in\mathbb{R}}$, cu funcţia de repartiţie continuă $H(y)$ a lui Y dacă este de forma

$$F(x) = \int_{\mathbb{R}} G(x, y)dH(y). \tag{1.1.13}$$

Observaţie 1.1.11.

În definiţia amestecării se pot considera în loc de funcţii de repartiţie, densităţi de probabilitate şi formulele devin

$$f(x) = \sum_{i=1}^{m} p_i f_i(x) \tag{1.1.14}$$

şi respectiv

$$f(x) = \int_{\mathbb{R}} g(x,y)h(y)dy. \tag{1.1.15}$$

Pentru cazul amestecării discrete notăm cu $X_i, i = 1, \ldots, m$ v.a. având funcţiile de repartiţie $F_i(x), i = 1 \ldots, m$ şi definim v.a.

$$X : \begin{pmatrix} X_1 & X_2 & \ldots & X_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}, \tag{1.1.16}$$

cu proprietăţile:

$$X_i \cap X_j = \emptyset, \forall i \neq j, \bigcup_i \{X = X_i\} = \Omega \tag{1.1.17}$$

şi

$$X_i(\{X_i \neq X\}) \subset A_i, A_i \cap X_i(\{X_i = X\}) = \emptyset \text{ cu } P(X_i^{-1}(A_i)) = 0. \tag{1.1.18}$$

Arătăm că v.a. X, definită în (1.1.16), are funcţia de repartiţie $F(x) = \sum_{i=1}^{m} p_i F_i(x)$. Calculăm $P(X \leq x)$ precum urmează:

$$\begin{aligned} P(\{X \leq x\}) &= P(\{X \leq x\} \cap (\{X = X_1\} \cup \ldots \cup \{X = X_m\})) \\ &= P(\bigcup_{i=1}^{m} (\{X \leq x\} \cap \{X = X_i\})) \\ &= \sum_{i=1}^{m} P\left(\{X \leq x\} \cap \{X = X_i\}\right) \\ &= \sum_{i=1}^{m} P(\{X_i \leq x\}) \cdot P(\{X = X_i\}) \\ &= \sum_{i=1}^{m} p_i F_i(x). \end{aligned} \tag{1.1.19}$$

Am folosit faptul că

$$\{\{X \leq x\} \cap \{X = X_i\}\} \cap \{\{X \leq x\} \cap \{X = X_j\}\} = \emptyset, \forall i \neq j$$

şi că evenimentele

$$\{X_i \leq x\} \text{ şi } \{X = X_i\}$$

sunt independente.

Procedura de generare a unei valori de selecţie a v.a. X este prezentată în Algoritm 1.3.

Algoritm 1.3: Metoda amestecării discrete

Intrare	Funcţiile de repartiţie $F_i(x)$ ale v.a. X_i
	Repartiţia v.a. discrete J: $P(X = i) = p_i, \sum_{i=1}^{m} p_i = 1$
Pas 1	Se generează un indice j având repartiţia J
Pas 2	Se generează X_j cu funcţia de repartiţie $F_j(x)$
Pas 3	Se defineşte $X = X_j$
Ieşire	Valoarea de selecţie, X, a v.a. X

Exemplul 1.1.12.

Fie X variabila având repartiţia Laplace a cărei densitate de probabilitate este

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|}, x \in \mathbb{R}, \lambda > 0 \tag{1.1.20}$$

Observăm că

$$f(x) = p_1 f_1 x + p_2 f_2(x), p_1 = p_2 = \frac{1}{2}$$

unde

$$f_1(x) = \begin{cases} \lambda e^{\lambda x} & \text{dacă } x \leq 0 \\ 0 & \text{dacă } x > 0 \end{cases}, f_2(x) = \begin{cases} 0 & \text{dacă } x \leq 0 \\ \lambda e^{-\lambda x} & \text{dacă } x > 0 \end{cases}.$$

Generarea variabilei X este realizată folosind Algoritm 1.4.

Algoritm 1.4: Metoda amestecării discrete

Intrare	V.a. $Y \sim Exp(\lambda)$ cu densitatea de probabilitate $f_2(x)$ şi V.a. $-Y$ cu densitatea de probabilitate $f_1(x)$
Pas 1	Se generează $U \sim \mathcal{U}(0,1)$
Pas 2	Dacă $U \leq 0.5$ atunci mergi la Pas 3; altfel mergi la Pas 4
Pas 3	s = 1; mergi la Pas 5
Pas 4	s = -1
Pas 5	Se generează $Y \sim Exp(\lambda)$
Pas 6	Se defineşte $X = sY$
Ieşire	Valoarea de selecţie, X, a v.a. X.

Bibliografie

[Craiu (1998)] M. Craiu (1998), *Statistică matematică: teorie şi probleme*, Editura Matrix Rom, Bucureşti

[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