Generarea variabilelor neuniforme

Curs 3

December 4, 2023

Introducere

- Fie X o variabilă aleatoare.
- Generarea v.a X = găsirea unui numar n (mare) de valori pe care le poate lua X.
- Cum se gaseste o astfel de valoare?
- Presupunem că $S_1, S_2, ..., S_n$ sunt v.a. pentru care se cunosc metode de generare (de exemplu variabilele uniforme pe $[0, 1]$ se generează cu ajutorul generatorilor de numere aleatoare, generatori care sunt deja implementati în majoritatea limbajelor de programare).
- Atunci o valoare a lui X se poate determina găsind o relație între X și $S_1, S_2, ..., S_n$.
- Algoritm efectiv de generare a lui X = aplicarea de n ori a metodei furnizate de relația dintre X și $S_1, S_2, ..., S_n$.
- Fiecare dintre algoritmii prezentați în curs se referă la o astfel de metodă.

Metoda inversă - Cazul continuu

Fie U o variabilă aleatoare uniformă pe $[0,1]$, cu densitatea de repartiție $f(x)$ și funcția de repartiție $F(x)$.

$$f(x) = \begin{cases} 1, & \text{dacă } x \in [0,1] \\ 0, & \text{în rest} \end{cases} \quad , \quad F(x) = \begin{cases} 0, & \text{dacă } x < 0 \\ x, & \text{dacă } x \in [0,1] \\ 1, & \text{dacă } x > 1 \end{cases} \quad .$$

Teoremă

(Hincin) Fie X o variabilă aleatoare continuă cu funcția de repartiție F, și U o variabila uniformă pe $[0,1]$. Atunci variabila aleatoare $F(X)$ este uniformă pe intervalul $[0,1]$, iar $F^{-1}(U)$ are funcția de repartiție F.

Dem (a doua parte):
Fie $x \in \mathbb{R}$. Atunci:

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

Algoritm M-inversă

Intrare: `Inversa funcţiei de repartiţie` F: F^{-1}.

`P1:` `Se generează` U `variabilă uniformă pe` $[0,1]$;

`P2:` $X = F^{-1}(U)$;

Ieșire: X `cu funcţia de repartiţie` $F(x)$

Propoziție

Variabila aleatoare U este uniformă pe $[0,1]$ dacă și numai dacă variabila $1-U$ este uniformă pe $[0,1]$.

Dem:

Fie $x < 0$. Atunci:

$$P\{1-U \leq x\} = P\{U \geq 1-x\} = 1 - P\{U < 1-x\} = 1 - 1 = 0$$

Fie $x \in [0,1]$. Atunci:

$$P\{1-U \leq x\} = P\{U \geq 1-x\} = 1-P\{U < 1-x\} = 1-(1-x) = x$$

Fie $x > 1$. Atunci:

$$P\{1 - U \leq x\} = P\{U \geq 1 - x\} = 1 - P\{U < 1 - x\} = 1 - 0 = 1$$

Observăm că dacă în expresia care definește funcția F^{-1} din algoritmul M-Inversă, apare $1 - U$, atunci, datorită Propoziției 1, $1 - U$ poate fi inlocuită direct cu U.

Exemplu

Fie X o variabilă Exp(λ) cu densitatea și funcția de repartiție:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \textit{dacă } x \geq 0; \\ 0, & \textit{în rest} \end{cases}, \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & \textit{dacă } x \geq 0; \\ 0, & \textit{în rest} \end{cases}$$

Atunci:

$$F^{-1}(U) = -\frac{1}{\lambda} \ln(1 - U)$$

iar algoritmul de generare prin metoda inversă este:

Algoritm M-inversă-Exp

Intrare: `Parametrul` λ.

`P1:` `Se generează` U `variabilă uniformă pe` $[0, 1]$;

`P2:` $X = -\frac{1}{\lambda} \ln(U)$;

Ieșire: X `cu funcţia de repartiţie` $F(x)$

Cazul discret

Fie X o variabilă aleatoare discretă cu repartiția:

$$X : \begin{pmatrix} a_1 & a_2 & \dots & a_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix} \quad \text{cu} \quad \sum_{i=1}^{m} p_i = 1$$

Funcția de repartiție a lui X va lua valorile:

$$F(x) = \begin{cases} 0, \text{ dacă } x < a_1 \\ p_1, \text{ dacă } a_1 \leq x < a_2 \\ p_1 + p_2, \text{ dacă } a_2 \leq x < a_3 \\ \dots \\ p_1 + p_2 + \dots + p_k, \text{ dacă } a_k \leq x < a_{k+1} \\ \dots \\ 1, \text{ dacă } a_m \leq x \end{cases}$$

Algoritmul constă în găsirea valorii a_i astfel încât $F(a_i) = U$, unde U este o variabilă uniformă pe $[0,1]$.

Fie $s_i = \sum_{j=1}^{i} p_j$. Observăm că:

$$P\{s_{i-1} < U \leq s_i\} = F_U(s_i) - F_U(s_{i-1}) = p_i = P\{X = a_i\}$$

Algoritm M-inversă-Discret
Intrare: $s_i = \sum_{j=1}^{i} p_j$ `şi` a_i, $i = 1, 2, ..., m$.
`P1:` `Se generează` U `variabilă uniformă pe` $[0,1]$;
`P2:` $i = 1$;
`P3:` `Dacă` $U \leq s_i$ $X = a_i$ `STOP. Altfel mergi la P4.`
`P4:` $i := i + 1$, `mergi la P3;`
Ieșire: X `cu funcţia de repartiţie` $F(x)$

Exemplu

Simularea unei variabile aleatoare Bernoulli Z:

$$Z : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix} \quad cu\ p + q = 1$$

Z are funcția de repartiție:

$$F(x) = P(Z \leq x) = \begin{cases} 0, & dacă\ x < 0 \\ q, & dacă\ x \in [0, 1) \\ 1 & dacă\ x \geq 1 \end{cases}$$

Un algoritm de generare a unei variabile Bernoulli este:

Intrare: `Parametrul` p, $q = 1 - p$.

`P1:` `Se generează` U `variabilă uniformă pe` $[0, 1]$;

`P2:` `Dacă` $U \leq q$ $Z = 0$. `Altfel` $Z = 1$.

Ieșire: Z `cu funcţia de repartiţie` $F(x)$.

Metoda compunerii sau a amestecării

Cazul discret

Definiție

Funcția de repartiție este o amestecare (sau compunere sau mixtură) discretă a mulțimii de funcții de repartiție $\{F_i(x)\}_{1\le i\le m}$ *după repartiția discretă*

$$J : \begin{pmatrix} 1 & 2 & \dots & m \\ p_1 & p_2 & \dots & p_m \end{pmatrix} \quad cu \quad \sum_{i=1}^{m} p_i = 1$$

dacă

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

Relația precedenta poate fi scrisă și în funcție de densitățile de repartiție:

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

Fie X variabila aleatoare cu funcția de repartiție $F(x)$ și X_i variabila aleatoare cu funcția de repartiție $F_i(x)$.

Algoritm compunere discretă
Intrare: Repartiţia lui J, familia de funcţii $\{F_i(x)\}_{1 \leq i \leq m}$;
P1: Generează J;
P2: Generează X_J cu funcţia de repartiţie $F_J(x)$;
P3: $X = X_J$.
Ieșire: X cu funcţia de repartiţie $F(x)$

Exemplu

Presupunem că la o stație de benzină sosesc m tipuri de mașini și se cunoaște p_i probabilitatea să sosească un automobil de tipul i, $1 \leq i \leq m$. Presupunem că timpul X_i între sosirile autoturismelor de tipul i este distribuit exponențial de parametru λ_i. Atunci timpul dintre două sosiri oarecare, X, are o repartiție mixt exponențială. Variabila X este o amestecare discretă, cu densitatea:

$$f(x) = \begin{cases} 0, & \text{dacă } x < 0 \\ \sum_{i=1}^{m} p_i \lambda_i e^{-\lambda_i x_i}, & \text{dacă } x \geq 0 \end{cases}$$

Prin urmare, o variabila mixt exponențiala poate fi generată cu ajutorul metodei compunerii cazul discret, în care

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

si $F_i(x) = 1 - e^{-\lambda_i x}$.

Exemplu

Fie X variabila aleatoare cu repartiția Laplace(λ) a cărei densitate este:

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

Atunci, putem să scriem

$$f(x) = p_1 f_1(x) + p_2 f_2(x)$$

cu

$$p_1 = p_2 = \frac{1}{2}$$

și

$$f_1(x) = \begin{cases} \lambda e^{\lambda x} \ \textit{dacă} \ x \leq 0 \\ 0 \ \textit{dacă} \ x > 0 \end{cases} \qquad f_2(x) = \begin{cases} 0 \ \textit{dacă} \ x \leq 0 \\ \lambda e^{-\lambda x} \ \textit{dacă} \ x > 0 \end{cases}$$

Prin urmare un algoritm de generare al variabilei Laplace(λ) poate fi:

Algoritm Laplace

Intrare: `parametrul` λ.

`P1:` `Se generează` U `variabilă uniformă pe` $[0, 1]$;

`P2:` `Dacă` $U \leq 0.5$ `atunci` $s := -1$, `altfel` $s = 1$;

`P3:` `Generează` $Y \sim Exp(\lambda)$;

`P4:` $X := sY$.

Ieșire: X `cu funcţia de repartiţie` $F(x)$

S se numește semn aleator.

Variabila Lapalace se poate simula ușor și cu metoda inversă.

Metoda compunerii discrete se poate aplica pentru orice densitate de repartiție:

Teoremă

Fie X o variabilă aleatoare cu densitatea de repartiție $f(x)$, $x \in \Delta \subseteq \mathbb{R}$. Fie o diviziune a lui Δ de forma $\Delta = \cup_{i=1}^{m} \Delta_i$, cu $\Delta_i \cap \Delta_j = \varnothing, \forall i \neq j$. Notând cu $p_i = P(X \in \Delta_i) > 0$, există densitățile $f_i(x)$, care iau valoarea 0 pentru $x \notin \Delta_i$ astfel încât

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

Dem:

Fie funcțiile f_i, definite astfel:

$$f_i(x) = \begin{cases} \frac{f(x)}{p_i} \text{ dacă } x \in \Delta_i \\ 0, \text{ dacă } x \notin \Delta_i \end{cases}$$

Observăm că:

$$\int_{-\infty}^{\infty} f_i(t)dt = \int_{\Delta_i} f_i(t)dt = \int_{\Delta_i} \frac{f(t)}{p_i}dt = \frac{1}{p_i}\int_{\Delta_i} f(t)dt = 1.$$

Deci funcțiile f_i astfel sunt densități de repartiție.

Fie un $x \in \Delta$, oarecare, cu $f(x) \neq 0$. Atunci există un i, $1 \leq i \leq m$ astfel încât $x \in \Delta_i$. Atunci avem:

$$f(x) = \frac{f(x)}{p_i}p_i = f_i(x)p_i = \sum_{j=1}^{m} p_j f_j(x)$$

pentru că $f_j(x) = 0, \forall j \neq i$.

Definiție

Funcția de repartiție $F(x)$ este o amestecare continuă a familiei de funcții de repartiție $\{G(x, Y)\}_{Y\in\mathbb{R}}$ după funcția de repartiție continuă $H(y)$ a lui Y dacă ea este de forma:

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

unde ultima integrală este integrala Stieltjes.

Relația precedenta poate fi scrisă și în funcție de densitățile de repartiție:

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

Intrare: Funcţiile de repartiţie H şi G.
P1: Se generează Y cu funcţia de repartiţie $H(y)$;
P2: Se generează Z_Y cu funcţia de repartiţie $G(x, Y)$;
P3: $X = Z_Y$
Ieșire: X cu funcţia de repartiţie $F(x)$

Exemplu

Fie $X > 0$ o v.a. care reprezintă durata în funcționare a unui aparat. Presupunem că X este o variabilă exponențială de parametru $\eta\lambda$, unde $\lambda > 0$ este un parametru care reprezintă o caracteristică aparatului, iar η este un parametru aleator care indică influența mediului în care lucrează aparatul. Presupunem că η este la rândul ei o variabilă aleatoare și că are densitatea de repartiție Gama$(0, b, a)$:

$$h(\eta) = \begin{cases} \frac{b^a}{\Gamma(a)}\eta^{a-1}e^{-b\eta}, \ \text{dacă } x \geq 0; \\ 0 \ \text{dacă } x < 0 \end{cases} .$$

Observăm că X se obține ca o amestecare continuă a unei familii de variabile exponențiale după o distribuție Gama. Densitatea de repartiție a variabilei X are forma:

$$f(x) = \int_0^\infty \eta\lambda e^{-\lambda\eta x} \frac{b^a}{\Gamma(a)} \eta^{a-1} e^{-b\eta} d\eta = \frac{\lambda b^a}{\Gamma(a)} \int_0^\infty \eta^a e^{-\eta(\lambda x + b)} d\eta \tag{2}$$

Facând schimbarea de variabila

$$\eta(\lambda x + b) = t$$

și ținând cont de faptul ca funcția Gama este:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt,$$

având proprietatea că $\Gamma(a+1) = a\Gamma(a)$, funcția f devine:

$$f(x) = \frac{\lambda b^a \Gamma(a+1)}{\Gamma(a)(\lambda x + b)^{a+1}} = \frac{\lambda a}{b} \frac{b^{a+1}}{(\lambda x + b)^{a+1}} = \frac{a\theta}{(\theta x + 1)^{a+1}}$$

cu

$$\theta = \lambda / b$$

Deci densitatea lui X este:

$$f(x) = \begin{cases} \frac{a\theta}{(\theta x+1)^{a+1}}, \text{ dacă } x \geq 0 \\ 0 \text{ dacă } x < 0 \end{cases}$$

Variabila cu densitatea f se numește variabilă Lomax, iar algoritmul ei de generare prin metoda compunerii continue (presupunând ca se cunoaște o metodă de generare a unei variabile Gama) se poate scrie astfel:

Algoritm Lomax
Intrare: `Parametrii` λ, a `şi` b.
`P1: Se generează` Y `cu repartiţia Gama(0,b,a);`
`P2: Se generează` Z_Y `cu funcţia de repartiţie` $Exp(x, Y)$;
`P3:` $X = Z_Y$
Ieșire: X `cu densitatea de repartiţie Lomax.`

Metoda respingerii

Mai poate fi numită metoda acceptării-respingerii.
Fie X o variabilă aleatoare pe care vrem să o generăm cu metoda respingerii și fie următoarele elemente cunoscute:

- Un procedeu de generare a unei variabile aleatoare N cu valori întregi pozitive;
- Procedee de generare a unor variabile aleatoare $S_i \in \mathcal{S}$, $i \geq 1$, unde $\mathcal{S}$ este o familie de variabile aleatoare dată;
- Un predicat $\mathcal{P}(S_1, S_2, ..., S_n)$ care se poate calcula simplu;
- Funcția Ψ, astfel încât
  $X = \Psi(\{S_1, S_2, ..., S_n\}, \mathcal{P}(S_1, S_2, ..., S_n) = \text{true})$

Atunci forma generală a unui algoritm de respingere este:

Algoritm RESPINGERE

Intrare: N, $\mathcal{S}$, $\mathcal{P}(S_1, S_2, ..., S_n)$, Ψ.

P1: Se generează N;

P2: Se generează $S_1, S_2, ..., S_n$ din $\mathcal{S}$;

P3: Dacă $\mathcal{P}(S_1, S_2, ..., S_n)$ = true atunci

$X = \Psi(S_1, S_2, ..., S_n)$ şi STOP, altfel mergi la P1;

Ieșire: Variabila aleatoare X.

Observăm că

- Dacă $\mathcal{P}(S_1, S_2, ..., S_n)$=false atunci mulțimea de variabile aleatoare $\{S_1, S_2, ..., S_n\}$ se respinge, de aici provenind numele de “metoda respingerii”.
- Dacă $p_a = P(\mathcal{P}(S_1, S_2, ..., S_n) = \text{true})$, numită și probabilitate de acceptare, este mare, atunci algoritmul este “bun”, altfel algoritmul este prea lent

Trei algoritmi de respingere bazați pe trei teoreme:

Prima teoremă de respingere

Teoremă

Fie X o variabilă aleatoare cu densitatea de repartiție f pentru $x \in \mathbb{R}$. Fie Y o altă variabilă aleatoare pentru care este cunoscută o metodă de generare și a cărei densitate de repartiție este h, astfel încât densitățile f și h iau valori diferite de 0 pe aceeași submulțime $A \subseteq \mathbb{R}$. Presupunem că există o constantă α, cu $0 < \alpha < \infty$ astfel încât $f(x) \leq \alpha h(x)$ pentru $\forall x \in A$. Atunci dacă U este o variabilă aleatoare $U(0,1)$, independentă de Y, densitatea de repartiție a variabilei Y, condiționată de

$$0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}$$

este f.

Dem:

Vom arăta că funcția de repartiție a lui Y condiționată de $0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}$ este F. Adică:

$$P\left(Y < x | 0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}\right) = F(x) = \int_{-\infty}^{x} f(v)dv$$

Fie evenimentele A și B definite astfel:

$$A = \{Y < x\}, \quad B = \left\{0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}\right\}$$

Atunci trebuie să arătăm că:

$$P(A|B) = F(x)$$

Conform definiției

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Calculam întâi $P(B)$:

$$P(B) = P\left(0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}\right) = \int_{-\infty}^{+\infty} \left[\int_0^{\frac{f(v)}{\alpha h(v)}} du\right] h(v)dv =$$

$$= \int_{-\infty}^{+\infty} \frac{f(v)}{\alpha h(v)} h(v)dv = \frac{1}{\alpha}.$$

deci

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \alpha \int_{-\infty}^{x} \left[\int_0^{\frac{f(v)}{\alpha h(v)}} du\right] h(v)dv =$$

$$= \alpha \int_{-\infty}^{x} \frac{f(v)}{\alpha h(v)} h(v)dv = \int_{-\infty}^{x} f(v)dv = F(x)$$

Observăm că:

- Această teoremă este cunoscută ca "teorema înfășurătoarei" pentru că graficul densității $f(x)$ se poate "înfășura" cu $\alpha h(x)$.
- Din demonstrație rezultă că probabilitatea de acceptare este $p_a = 1/\alpha$. De aici rezultă că pentru a avea o metodă a înfășurătoarei nebanală, trebuie ca $\alpha > 1$.
- Procedura de respingere este formată din următoarele elemente:
  - $N = 2$ variabilă aleatoare constantă;
  - $\mathcal{S} = \{U, Y\}$;
  - $\mathcal{P}(U, Y)$=true dacă $0 \leq U \leq \frac{f(Y)}{\alpha h(Y)}$;
  - $\Psi(U, Y) = Y$.

Algoritm Respingere1

Intrare: h, α.

P1: Se generează Y cu densitatea h;

P2: Se generează $U \sim U(0,1)$;

P3: Dacă $U \leq \frac{f(Y)}{\alpha h(Y)}$, mergi la P4. Altfel, mergi la P1;

P4: X=Y.

Ieșire: Variabila aleatoare X.

Exemplu

Fie X o variabilă Gama$(0,1,\nu)$ (adică Gama standard) cu $0<\nu<1$. Variabila X are densitatea de repartiție:

$$f(x)=\begin{cases}\frac{1}{\Gamma(\nu)}x^{\nu-1}e^{-x}, \ dac\breve{a}\ x\geq 0\\ 0\ dac\breve{a}\ x<0.\end{cases};$$

Vom aplica metoda înfășurătoarei, folosind o densitate Weibull$(0,1,\nu)$:

$$h(x)=\begin{cases}\nu x^{\nu-1}e^{-x^{\nu}}, \ dac\breve{a}\ x\geq 0\\ 0\ dac\breve{a}\ x<0\end{cases}$$

Pentru a determina constanta α de înfășurare analizăm raportul:

$$r(x)=\frac{f(x)}{h(x)}=\frac{1}{\nu\Gamma(\nu)}e^{-x+x^{\nu}},$$

pentru care trebuie să determinam valoarea maximă.

Rezolvăm ecuația
$$r'(x) = 0.$$
Soluția, care este și punct de maxim al funcției $r(x)$, este $x_{\max} = \nu^{-\frac{1}{\nu-1}}$ de unde rezultă:
$$\alpha = \frac{e^{\zeta(1-\nu)}}{\Gamma(\nu+1)} \ cu \ \zeta = \nu^{\frac{\nu}{1-\nu}}$$
Algoritmul pentru generarea variabilei X prin metoda respingerii este:

Algoritm Gama-Resp

Intrare: ν, $c := 1/\nu$, $\zeta = \nu^{\frac{\nu}{1-\nu}}$, $a = e^{\zeta(\nu-1)}$.

`P1:` `Se generează` $Y \sim Weib(0,1,\nu)$ `(metoda inversă);`

`P1.1:` $U \sim U(0,1)$;

`P1.2:` $Y := [-\ln(U)]^c$

`P2:` `Se generează` $U \sim U(0,1)$;

`P3:` `Dacă` $U \leq ae^{Y^\nu - Y}$, `mergi la P4. Altfel, mergi la P1;`

`P4:` `X=Y.`

Ieșire: `Variabila aleatoare` X.

A doua teoremă de respingere

Teoremă

Fie X o variabilă aleatoare cu funcția de repartiție F, de forma:

$$F(x) = c\int_{-\infty}^{x} Q(\phi(t))dR(t) \tag{3}$$

unde Q este funcția de repartiție a unei variabile aleatoare Z, $Z \in [0, M]$, ϕ este o funcție care ia valori în $[0, M]$ (cu M putând lua și valoarea ∞), iar R este funcția de repartiție a unei variabile aleatoare $Y \in \mathbb{R}$, independente de Z. În aceste condiții funcția de repartiție a variabilei Y condiționată de $Z \leq \phi(Y)$ este F.

Dem:
Mai întâi observam că F fiind funcție de repartiție, rezultă că $\lim_{x\to\infty} F(x)dx = 1$, adică:

$$c\int_{-\infty}^{\infty} Q(\phi(t))dR(t) = 1.$$

Deci c, constanta de normare, este:

$$c = \left[\int_{-\infty}^{+\infty} Q(\phi(x))dR(x)\right]^{-1}.$$

Trebuie să arătăm că:

$$P(Y < x | Z \leq \phi(Y)) = F(x).$$

Fie evenimentele A și B definite astfel:

$$A = \{Y < x\}; \quad B = \{Z \leq \phi(Y)\}$$

pentru a demonstra teorema trebuie să arătăm că:

$$P(A|B) = F(x).$$

Din definiție

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Probabilitatea de realizare a evenimentului B este:

$$P(B) = \int_{-\infty}^{+\infty} \left(\int_0^{\phi(x)} dQ(y) \right) dR(x) = \int_{-\infty}^{+\infty} Q(\phi(x)) dR(x) = \frac{1}{c}.$$

Prin urmare:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = cP(A \cap B) = c \int_{-\infty}^{x} \left(\int_0^{\phi(y)} dQ(z) \right) dR(y) =$$

$$= c \int_{-\infty}^{x} Q(\phi(y)) dR(y) = F(x).$$

Observăm că:

- Probabilitatea de acceptare este $p_a = P(B) = \frac{1}{c}$;
- Elementele algoritmului de respingere sunt
  - $N = 2$;
  - $\mathcal{S} = \{Z, Y\}$;
  - $\mathcal{P}(Z, Y)$=true dacă $Z \leq \phi(Y)$;
  - $\Psi(Z, Y) = Y$.
- teorema se verifică și dacă relația (3) se scrie în termeni de densități de repartiție:

$$f(x) = cQ(\phi(x))r(x), \text{ cu } r(x) = R'(x).$$

- O formă duală a teoremei se obține dacă $F(x)$ este de forma:

$$F(x) = c \int_{-\infty}^{x} (1 - Q(\phi(x)))dR(x) \tag{4}$$

cu

$$c = \left[\int_{-\infty}^{+\infty} (1 - Q(\phi(x)))dR(x) \right]^{-1}$$

în acest caz evenimentul B este: $B = \{Z \geq \phi(Y)\}$.

- Relația (4) se poate scrie în funcție de densități astfel:

$$f(x) = c(1 - Q(\phi(x)))r(x)$$

iar probabilitatea de acceptare pentru varianta duală este:

$$p_a = P(Z \geq \phi(Y)) = \frac{1}{c}$$

Algoritm Respingere2

Intrare: Q, ϕ, R.

P1: Se generează Z cu funcţia de repartiţie Q;

P2: Se generează Y cu funcţia de repartiţie R;

P3: Dacă $Z \leq \phi(Y)$, mergi la P4. Altfel, mergi la P1;

P4: X=Y.

Ieșire: Variabila aleatoare X.

Exemplu

Fie X o varibilă aleatoare cu densitatea:

$$f(x) = c(1 - e^{-\lambda x})\mu e^{-\mu x}, \quad x \geq 0 \tag{5}$$

unde c este o constantă de normare. Atunci un algoritm de generare a variabilei aleatoare X se poate scrie folosind a doua teoremă de respingere. Avem:

$$\phi(x) = x, \quad Q(z) = 1 - e^{-\lambda z}, \quad z > 0, \quad r(x) = \mu e^{-\mu x}.$$

și

$$c = \left[\int_0^{+\infty} (1 - e^{-\lambda x})\mu e^{-\mu x} dx\right]^{-1} = \left[\frac{\lambda}{\lambda + \mu}\right]^{-1}$$

iar un algoritm pentru generarea lui X este următorul:

Algoritm ExResp2

Intrare: `Parametrii` λ, μ.

`P1:` `Se generează` $Z \sim Exp(\lambda)$;

`P2:` `Se generează` $Y \sim Exp(\mu)$;

`P3:` `Dacă` $Z \leq Y$, `mergi la P4. Altfel, mergi la P1;`

`P4:` $X := Y$.

Ieșire: `Variabila aleatoare` X.

- *Algoritmul ExResp2 este rapid dacă* $\mu << \lambda$.
- *Metoda inversă nu este recomandabilă pentru că determinarea inversei funcției* F *nu este imediată.*

A treia teoremă de respingere

Teorema șirului descendent

Teoremă

Fie variabilele $Z_i \sim G(x)$, $i = 1, 2, , ...,$ $Z_0 \sim G_0(z)$ *independente. Atunci următoarele afirmații sunt adevărate:*

1. *Dacă* x *și* k *sunt fixate atunci:*

$$P(x \geq Z_1 \geq Z_2 \geq ... \geq Z_{k-1} < Z_k) = \frac{[G(x)]^{k-1}}{(k-1)!} - \frac{[G(x)]^k}{k!}. \tag{6}$$

2. *Dacă* x *este fixat și* K *este indicele aleator la care se "rupe" șirul descendent (ca la punctul 1), atunci*

$$P(K = nr.impar) = P(K \mod 2 = 1) = e^{-G(x)}. \tag{7}$$

3. *Dacă subșirul descendent este* $Z_0 \geq Z_1 \geq ... \geq Z_{K-1} < Z_K$ *(adică se rupe la* K *aleator și incepe cu* $Z_0 \sim G_0(x)$*), atunci:*

$$P(Z_0 < x | K \mod 2 = 1) = \frac{1}{p_a} \int_{-\infty}^{x} e^{-G(t)} dG_0(t), \qquad (8)$$

unde p_a este constanta de normare:

$$p_a = \int_{-\infty}^{+\infty} e^{-G(x)} dG_0(x). \qquad (9)$$

Dem:

1. Fie evenimentele:

$$A = \{x \geq Z_1 \geq Z_2 \geq ... \geq Z_{k-1}\}; \quad B = \{x \geq Z_1 \geq Z_2 \geq ... \geq Z_k\}$$

Pentru a demonstra (6) observăm că probabilitatea din membrul stâng se scrie $P(A \setminus B)$ și pentru că $B \subseteq A$ avem:

$$P(A \setminus B) = P(A) - P(B)$$

Acum îl vom calcula pe $P(A)$.
Observăm că $P(Z_i \leq x) = G(x)$ și

$P(Z_1 \leq x, Z_2 \leq x, ..., Z_{k-1} \leq x) = [G(x)]^{k-1}$, pentru că variabilele $Z_1, Z_2, ...$ sunt independente.
Deoarece subșirul care definește evenimentul A conține numai una din cele $(k-1)!$ ordini în care se pot afla cele $k-1$ variabile aleatoare Z_i, $1 \leq i \leq k-1$, rezultă că:

$$P(A) = \frac{[G(x)]^{k-1}}{(k-1)!}$$

Deci:

$$P(A \setminus B) = P(A) - P(B) = \frac{[G(x)]^{k-1}}{(k-1)!} - \frac{[G(x)]^{k}}{k!}$$

iar afirmația 1. este adevărată.
2. Probabilitatea ca indicele aleator K să fie impar este:

$$P(K = \text{nr. impar}) = P(K = 1) + P(K = 3) + ... =$$

$$= 1 - \frac{G(x)}{1!} + \frac{[G(x)]^2}{2!} - \frac{[G(x)]^3}{3!} + ... = e^{-G(x)}$$

3. Observăm că atunci când Z_0 este aleator avem:

$$P(K = \text{nr. impar}) = P(K \mod 2 = 1) = \int_{-\infty}^{+\infty} e^{-G(t)} dG_0(t)$$

Adică, ținând cont de forma probabilității p_a din (9) avem:

$$P(K = \text{nr. impar}) = p_a$$

Conform definiției probabilității condiționate:

$$P(Z_0 < x | K \mod 2 = 1) = \frac{P(\{Z_0 < x\} \cap \{K \mod 2 = 1\})}{P(K \mod 2 = 1)}.$$

Prin urmare:

$$P(Z_0 < x | K \mod 2 = 1) = \frac{1}{p_a} \int_{-\infty}^{x} e^{-G(t)} dG_0(t),$$

ceea ce demonstrează punctul 3 al teoremei.

Variabilele aleatoare X care pot fi generate cu a treia teoremă de respingere sunt acele variabile aleatoare care au funcția de repartiție F, de forma:

$$F(x) = \frac{1}{p_a} \int_{-\infty}^{x} e^{-G(t)} dG_0(t).$$

Algoritmul rezultat din a treia teoremă de respingere este următorul:

Algoritm Respingere3

Intrare: Funcţiile de repartiţie G_0, G.

P1: Se generează $Z_0 \sim G_0(x)$;

P2: $Z^* := Z_0$, $K = 1$;

P3: Se generează $Z_1 \sim G(x)$;

P4: Dacă $Z_0 \geq Z_1$ mergi la P5, altfel mergi P6;

P5: $K := K + 1$, $Z_0 := Z_1$, mergi pa P3;

P6: Dacă $K \mod 2 = 1$, mergi la P7. Altfel mergi la P1.

P7: $X = Z^*$.

Ieșire: Variabila aleatoare X.

Exemplu

Presupunem că variabilele Z_i, $i \geq 0$ sunt uniforme pe intervalul $[0,1]$, $Z_i = U_i$, $i \geq 0$. Atunci conform celei de-a treia teoreme de respingere avem:

$$P(U_0 \leq x | K = nr.impar) = \frac{1}{p_a} \int_0^x e^{-t} dt$$

cu

$$p_a = \int_0^1 e^{-x} dx = 1 - e^{-1}.$$

Cu alte cuvinte U_0 acceptat are funcția de repartiție

$$F(x) = \begin{cases} 0, \ \textit{dacă } x < 0 \\ \frac{1-e^{-x}}{1-e^{-1}}, \ \textit{dacă } 0 \leq x \leq 1 \\ 1, \ \textit{dacă } x > 1 \end{cases}$$

care reprezintă funcția de repartiție a unei variabile $Exp(1)$ trunchiată pe intervalul $[0,1]$.