

Véletlenszám-generátorok

4. rész

– egy MATLAB[®] alapú megközelítés –

Róth Ágoston, Vas Orsolya

Matematika és Informatika Intézet, Babeş-Bolyai Tudományegyetem, Kolozsvár, Románia

(agoston.roth@gmail.com, vas.orsolya@yahoo.com)

4. labor / 2018. október 22–25.



- Megjegyezzük, hogy az ebben a pontban leírt eljárást még az **elfogadás módszerének** is nevezik a szakirodalomban – a továbbiakban viszont mi csak az elutasítás módszereként említjük. A módszer lényege a következő tulajdonságokra épül.

1. Tulajdonság

Adott az $f : \mathbb{R}^d \rightarrow [0, 1]$ sűrűségfüggvényű X valószínűségi vektor és az ettől független $[0, 1]$ intervallumon értelmezett egyenletes eloszlású U valószínűségi változó.

- Ekkor az $(X, cUf(X))$ valószínűségi vektor egyenletes eloszlású az

$$A = \left\{ (x, u) \in \mathbb{R}^{d+1} : x \in \mathbb{R}^d, 0 \leq u \leq cf(x) \right\}$$

tartományon, ahol $c > 0$ egy tetszőleges konstans.

- Fordítva, ha (X, U) olyan valószínűségi vektor az \mathbb{R}^{d+1} térben, mely egyenletes eloszlású az A tartományon, akkor az \mathbb{R}^d térbeli X valószínűségi vektor sűrűségfüggvénye éppen f .



Bizonyítás

- Először a **direkt állítást** bizonyítjuk.
- Tekintsünk egy tetszőleges Borel-mérhető $B \subseteq A$ részhalmazt és ennek az $x \in \mathbb{R}^d$ ponthoz tartozó

$$B_x = \{u : (x, u) \in B\}$$

szekcióját!

- Ekkor

$$P((X, cUf(X)) \in B) = \int_{\mathbb{R}^d} \left(\int_{B_x} \frac{1}{cf(x)} du \right) f(x) dx = \frac{1}{c} \int_B du dx,$$

ahol felhasználtuk, hogy az $(X, cUf(X))$ valószínűségi vektor X és $cUf(X)$ komponense független, valamint, hogy az X valószínűségi változó bármely $x \in \mathbb{R}^d$ értékére a $cf(x)$ U valószínűségi változó egyenletes eloszlása a $[0, cf(x)]$ intervallumon, azaz sűrűségfüggvénye azonosan megegyezik az $\frac{1}{cf(x)}$ mennyiséggel ezen az intervallumon.

- Mivel az A tartomány térfogata éppen

$$\int_A du dx = \int_{\mathbb{R}^d} \int_0^{cf(x)} du dx = c \int_{\mathbb{R}^d} f(x) dx = c,$$

a fentiekből következik, hogy az $(X, cUf(X))$ valószínűségi vektor valóban egyenletes eloszlású az A tartományon.



Bizonyítás – folytatás

- A valószínűségi vektorok sűrűségfüggvényének értelmezése alapján a **fordított állítás** igazolásához be kellene látnunk, hogy bármely $B \subseteq \mathbb{R}^d$ Borel-mérhető halmazra teljesül a

$$P(X \in B) = \int_B f(x) dx$$

egyenlőség.

- Tetszőlegesen rögzített $B \subseteq \mathbb{R}^d$ Borel-mérhető halmaz esetén vezessük be a

$$B^* = \{(x, u) : x \in B, 0 \leq u \leq cf(x)\} \subset \mathbb{R}^{d+1}$$

jelölést!

- Ekkor a következőket írhatjuk:

$$P(X \in B) = P((X, U) \in B^*) = \frac{\int_{B^*} du dx}{\int_A du dx} = \frac{1}{c} \int_B cf(x) dx = \int_B f(x) dx,$$

amit éppen igazolni kellett.



2. Tulajdonság

Tekintsük az $\{X_i\}_{i \geq 1}$ független, azonos eloszlású, \mathbb{R}^d -értékkészletű valószínűségi vektorokból álló sorozatot, továbbá legyen $A \subseteq \mathbb{R}^d$ egy Borel-mérhető halmaz, amelyre $P(X_1 \in A) = p > 0$. Az Y valószínűségi vektor egyezzen meg az első olyan X_i valószínűségi vektorral, amely értékeit az A halmazból veszi fel!

- Ekkor bármely $B \subseteq \mathbb{R}^d$ Borel-mérhető halmazra az Y eloszlását a

$$P(Y \in B) = \frac{1}{p} P(X_1 \in A \cap B)$$

képlet adja.

- Sajátságosan, ha X_1 egyenletes eloszlású az A_0 halmazon (ahol $A \subseteq A_0$), akkor Y egyenletes eloszlású az A halmazon.



Bizonyítás

- Vegyük észre, hogy tetszőleges $B \subseteq \mathbb{R}^d$ Borel-mérhető halmaz esetén

$$\begin{aligned} P(Y \in B) &= \sum_{i \geq 1} P(X_1 \notin A, X_2 \notin A, \dots, X_{i-1} \notin A, X_i \in B \cap A) \\ &= \sum_{i \geq 1} (1-p)^{i-1} P(X_1 \in A \cap B) \\ &= \frac{1}{1-(1-p)} P(X_1 \in A \cap B) \\ &= \frac{1}{p} P(X_1 \in A \cap B), \end{aligned}$$

amit éppen igazolni kellett!

- Ha X_1 egyenletes eloszlású az A_0 halmazon, akkor

$$P(Y \in B) = \frac{P(X_1 \in A \cap B)}{P(X_1 \in A)} = \frac{\int_{A_0 \cap A \cap B} dx}{\int_{A_0} dx} \frac{\int_{A_0} dx}{\int_{A_0 \cap A} dx} = \frac{\int_{A \cap B} dx}{\int_A dx},$$

amiből következik, hogy Y egyenletes eloszlású az A halmazon.



- Az elutasítás módszerének alapértelmezett változata feltételezi, hogy az $f : \mathbb{R}^d \rightarrow [0, 1]$ sűrűségfüggvényű X valószínűségi vektor esetén létezik egy olyan $g : \mathbb{R}^d \rightarrow [0, 1]$ sűrűségfüggvény és egy olyan $c \geq 1$ konstans, amelyek esetén teljesül az

$$f(x) \leq cg(x), \forall x \in \mathbb{R}^d \quad (1)$$

egyenlőtlenség.

- Ekkor f sűrűségfüggvényű valószínűségi vektorokat az alábbi algoritmussal generálhatunk.



1. Algoritmus (Az elutasítás módszerének alapértelmezett változata)

- **Bemenet:** az (1)-es feltételt kielégítő $f, g : \mathbb{R}^d \rightarrow [0, 1]$ sűrűségfüggvények, a $c \geq 1$ „elutasítási” konstans, valamint az $n \geq 1$ természetes szám, amely a kimeneti mintavétel nagyságát határozza meg.
- **Kimenet:** az f sűrűségfüggvénnyel jellemzett X valószínűségi vektornak egy n -elemű független $\{X_i\}_{i=1}^n$ mintavétele.

① Minden $i = 1, 2, \dots, n$ indexre **végezd el:**

② **ismételd:**

③ generáld az \mathbb{R}^d térbeli g sűrűségfüggvényű Y valószínűségi vektort, valamint az ettől független $U \sim \mathcal{U}([0, 1])$ valószínűségi változót

④ **ameddig** $U \leq \frac{f(Y)}{cg(Y)}$

⑤ $X_i = Y$.

Az elutasítás módszerének helyessége

- A 1. algoritmus helyességét vagy az 1. és a 2. tulajdonságok együttes alkalmazásával, vagy akár direkt számolással is beláthatjuk.
- Minden $i = 1, 2, \dots, n$ index és $B \subset \mathbb{R}^d$ Borel-mérhető halmaz esetén a következőket írhatjuk:

$$\begin{aligned}
 P(X_i \in B) &= P\left(Y \in B \mid U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{P\left(\left(U \leq \frac{f(Y)}{cg(Y)}\right) \cap (Y \in B)\right)}{P\left(U \leq \frac{f(Y)}{cg(Y)}\right)} \\
 &= \frac{\int_{y \in B} P\left(U \leq \frac{f(y)}{cg(y)} \mid Y = y\right) g(y) dy}{\int_{y \in \mathbb{R}^d} P\left(U \leq \frac{f(y)}{cg(y)} \mid Y = y\right) g(y) dy} = \frac{\int_{y \in B} \frac{f(y)}{cg(y)} g(y) dy}{\int_{y \in \mathbb{R}^d} \frac{f(y)}{cg(y)} g(y) dy} \\
 &= \frac{\int_{y \in B} f(y) dy}{\int_{y \in \mathbb{R}^d} f(y) dy} = \int_{y \in B} f(y) dy,
 \end{aligned}$$

vagyis az X_i mintavételbeli valószínűségi vektor sűrűségfüggvénye éppen f .

Az elutasítás módszerének tanulmányozása

- Jelölje N azt a valószínűségi változót, mely azt számolja, hogy az 1. algoritmusbeli **ismételd–ameddig** ciklus hány (Y, U) párt generál valamely $i = 1, 2, \dots, n$ index esetén az $U \leq \frac{f(Y)}{cg(Y)}$ leállási feltétel teljesüléséig.
- Ugyanakkor vegyük észre, hogy a leállási feltétel

$$\begin{aligned} p &= P\left(U \leq \frac{f(Y)}{cg(Y)}\right) = \int_{\mathbb{R}^d} P\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) dy = \int_{\mathbb{R}^d} \frac{f(y)}{cg(y)} g(y) dy \\ &= \frac{1}{c} \int_{\mathbb{R}^d} f(y) dy = \frac{1}{c} \end{aligned}$$

valószínűséggel következik be!

- Ekkor az N valószínűségi változó p -paraméterű geometriai eloszlást követ, hiszen $P(N = j) = (1 - p)^{j-1} p$, $j \geq 1$. Ezért egyetlen f sűrűségfüggvényű X_i minta generálása várhatóan $E(N) = \frac{1}{p} = c$ iteráció alatt történik, és ennek a várható iterációszámnak a szórásnégyzete $D^2(N) = \frac{1-p}{p^2} = c^2 - c$. Mindezekből következik, hogy a c konstans értékét a lehető legkisebbnek célszerű megválasztani. Az N geometriai eloszlása miatt a $P(N = j) = (1 - p)^{j-1} p$ valószínűségek exponenciális sebességgel csökkenek: vegyük észre, hogy $P(N > j) = (1 - p)^j \leq e^{-pj}$, $j \geq 1$.

- Jelölje $\mathcal{C}_{M,a,b}$ az $f : [a, b] \rightarrow [0, M]$ ($M > 0$) sűrűségfüggvények osztályát. Ekkor minden $f \in \mathcal{C}_{M,a,b}$ sűrűségfüggvényt az $[a, b]$ intervallumon egyenletes eloszlás $g(x) = \frac{1}{b-a}$ sűrűségfüggvényével majorálhatunk úgy, hogy a c elutasítási konstanst $M(b-a)$ alakban választjuk meg. az 1. algoritmus ekkor a következőképpen módosul.

2. Algoritmus (Kompakt tartományú és korlátos sűrűségfüggvények esete)

- Bemenet:** az $f \in \mathcal{C}_{M,a,b}$ sűrűségfüggvény és az $n \geq 1$ természetes szám, amely a kimeneti mintavétel nagyságát határozza meg.
- Kimenet:** az f sűrűségfüggvényű X valószínűségi változónak egy n -elemű független $\{X_i\}_{i=1}^n$ mintavétele.

1 Minden $i = 1, 2, \dots, n$ indexre **végezd el**:

2 **ismételd**:

3 generáld a független $U \sim \mathcal{U}([0, 1])$ és $V \sim \mathcal{U}([0, 1])$ valószínűségi változókat

4 $Y \leftarrow a + (b - a)V$

5 **ameddig** $UM \leq f(Y)$

6 $X_i = Y$.

- Elégséges csak az $\mathcal{N}(0, 1)$ standard normális eloszlású valószínűségi változók generálását megoldanunk, hiszen egy alkalmas lineáris transzformációval általános $\mathcal{N}(\mu, \sigma)$ normális eloszlású számokat mintavételezhetünk ($\mu \in \mathbb{R}$, $\sigma > 0$).
- Ismert tény, hogy az

$$F_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, x \in \mathbb{R} \quad (2)$$

eloszlásfüggvényt nem tudjuk elemi függvényekkel kifejezni, ezért az egzakt inverziós módszer nem alkalmas, a közelítő numerikus inverziós módszer pedig – a (2)-es képletbeli improprius integrál kvadratúra képletekkel történő ismételt approximálása miatt – túlságosan költséges lenne standard normális eloszlású valószínűségi változók generálására. Ezért más módszerekhez kell folyamodnunk.

- Az egyik lehetőség az alábbi, Laplace-féle sűrűségfüggvényre épülő, elutasítás típusú módszer.



- Tekintsük a standard normális eloszlás

$$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

sűrűségfüggvényét!

- Ha az $f_{\mathcal{N}(0,1)}$ függvényre keresünk felső határt, akkor valójában az $\frac{x^2}{2}$ függvényre kell egy alsó határt találnunk.
- Vegyük észre, hogy

$$\frac{1}{2} (|x| - 1)^2 = \frac{x^2}{2} + \frac{1}{2} - |x| \geq 0,$$

ezért

$$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2} - |x|} = cg(x),$$

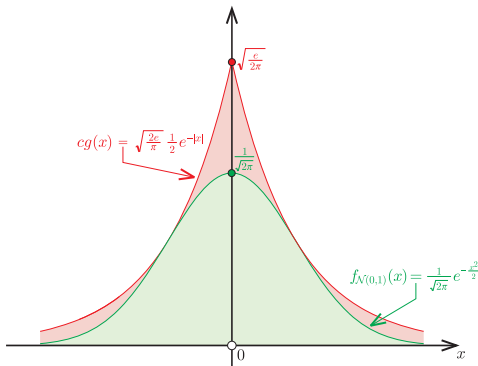
ahol

$$g(x) = \frac{1}{2} e^{-|x|}, x \in \mathbb{R}$$

a Laplace-féle (másképpen dupla-exponenciális) sűrűségfüggvényt, $c = \sqrt{\frac{2e}{\pi}}$ pedig az elutasítási konstanst jelöli.

- Az 1. ábra szemlélteti ezeket a függvényeket.





ábra. Az elutasítás módszerére épülő 3. algoritmus a standard normális eloszlás $f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$ sűrűségfüggvényét, valamint az azt határoló $cg(x)$ függvényt használja fel, ahol $g(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$ a Laplace-féle (másképpen dupla-exponenciális) sűrűségfüggvényt, $c = \sqrt{\frac{2e}{\pi}}$ pedig az optimális elutasítási konstanst jelöli.

3. Algoritmus ($\mathcal{N}(0, 1)$ -eloszlású számgenerátor a Laplace-féle sűrűségfüggvény alapján)

- **Bemenet:** az $n \geq 1$ természetes szám, amely a kimeneti mintavétel méretét rögzíti.
- **Kimenet:** az $X \sim \mathcal{N}(0, 1)$ valószínűségi változónak egy n -elemű független $\{X_i\}_{i=1}^n$ mintavétele.

① Minden $i = 1, 2, \dots, n$ indexre **végezd el:**

② **ismételd:**

③ generáld az $Y \sim \text{Exp}(1)$, $U \sim \mathcal{U}([0, 1])$ és $V \sim \mathcal{U}([0, 1])$ független valószínűségi változókat

④ ha $U < \frac{1}{2}$, akkor $Y \leftarrow -Y$

⑤ ameddig $\text{Ve}^{\frac{1}{2}-|Y|} \leq e^{-\frac{Y^2}{2}}$

⑥ $X_i \leftarrow Y$.



- Vegyük észre, hogy a 3. algoritmus 4. sorában feleslegesen generálunk véletlenszerű előjelet olyan Y exponenciális eloszlású valószínűségi változók esetén, amelyeket elutasítunk! Ugyanakkor numerikus szempontból jobb, ha az algoritmus **ismételd-ameddig** ciklusának 5. sorában megfogalmazott leállási feltételben az exponenciális függvény helyett természetes alapú logaritmust használunk. Ezért a 3. algoritmus az alábbi trükkel hatékonyabbá tehető.

4. Algoritmus (A 3. algoritmus optimálisabb változata)

- Bemenet:** egy $n \geq 1$ természetes szám, amely a kimeneti mintavétel méretét szabja meg.
 - Kimenet:** az $X \sim \mathcal{N}(0, 1)$ valószínűségi változónak egy n -elemű független $\{X_i\}_{i=1}^n$ mintavétele.
- Minden $i = 1, 2, \dots, n$ indexre **végezd el:**
 - ismételd:**
 - generáld a független $Y \sim \text{Exp}(1)$ és $V \sim \mathcal{U}([-1, 1])$ valószínűségi változókat
 - ameddig** $(Y - 1)^2 \leq -2 \ln(|V|)$
 - $X_i \leftarrow Y \text{ sign}(V)$.

- Adott $f : \mathbb{R}^d \rightarrow [0, 1]$ és az azt majoráló $g : \mathbb{R}^d \rightarrow [0, 1]$ sűrűségfüggvények esetén a c elutasítási konstans értéke legalább a

$$\sup_{x \in \mathbb{R}^d} \frac{f(x)}{g(x)}$$

mennyiséggel kell megegyezzen.

- Semmit sem veszítünk ezzel a beállítással, hiszen ekkor az f és cg függvények érintik valahol egymást.
- A 3. algoritmus esetén a g majoráló sűrűségfüggvényt egy speciális egyenlőtlenség szerkesztésével határoztuk meg, sokszor viszont bölcsebb, ha az ismeretlen majoráló sűrűségfüggvényt egy $\theta \in \mathbb{R}^k$ ($k \geq 1$) paramétervektortól függő

$$\{g_\theta : \mathbb{R}^d \rightarrow \mathbb{R} : \theta \in \mathbb{R}^k\}$$

sűrűségfüggvény-család legjobb g_θ elemeként választjuk ki úgy, hogy az optimális elutasítási konstans a

$$c_\theta = \sup_{x \in \mathbb{R}^d} \frac{f(x)}{g_\theta(x)}$$

alakban választjuk meg, ahol azt a θ paramétervektort tekintjük optimálisnak, amelyre c_θ minimális, azaz amelyre c_θ értéke a lehető legközelebb esik az 1-hez.



- Ezt az optimalizálási folyamatot egy példán keresztül mutatjuk be.
- Az egyszerűség kedvéért maradjunk a standard normális eloszlás

$$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

sűrűségfüggvényénél, valamint válasszuk ki az ezt majoráló sűrűségfüggvényt a

$$\left\{ g_{\theta}(x) = \frac{\theta}{\pi} \frac{1}{x^2 + \theta^2}, x \in \mathbb{R} : \theta > 0 \right\}$$

Cauchy-féle sűrűségfüggvények családjából, ahol θ egy skálázási paraméter.

- Mivel bármely $\theta > 0$ paraméter esetén a g_{θ} sűrűségfüggvény az $f_{\mathcal{N}(0,1)}$ sűrűségfüggvényhez hasonlóan szigorúan növekszik a $(-\infty, 0]$ intervallumon, illetve szigorúan csökken a $[0, \infty)$ tartományon, a θ skálázási paraméter mellett nincs szükség egyéb, például translációs, paraméter bevezetésére.



- A továbbiakban kimutatjuk, hogy ebben a sajátos esetben az optimális elutasítási konstans

$$c_{\theta} = \begin{cases} \frac{\sqrt{2\pi}}{e^{\theta}} e^{\frac{\theta^2}{2}}, & 0 < \theta < \sqrt{2}, \\ \theta \sqrt{\frac{\pi}{2}}, & \theta \geq \sqrt{2} \end{cases} \quad (3)$$

alakú.



Nemegyenletes eloszlású véletlenszám-generátorok

Elutasítás módszere: standard normális eloszlású számgenerátor a Cauchy-féle sűrűségfüggvény alapján

- A következőképpen gondolkodhatunk: az $\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}$ függvény maximális, ha a $\ln\left(\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}\right)$ függvény maximális.
- Szüksőérték hely érdekében a

$$\frac{d}{dx} \ln\left(\frac{f_{\mathcal{N}(0,1)}(x)}{g_{\theta}(x)}\right) = 0$$

egyenletet kell megoldanunk, amely a

$$-x + \frac{2x}{x^2 + \theta^2} = 0$$

egyszerűbb alakra hozható, ahonnan az $x = 0$ és az $x = \pm\sqrt{2 - \theta^2}$ szélsőérték helyeket kapjuk eredményül, ahol az utóbbi esetnek csak a $\theta^2 \leq 2$ feltétel teljesülésekor van értelme.

- Az $x = 0$ pontra az $\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}$ függvény a $\theta\sqrt{\frac{\pi}{2}}$ értéket veszi fel, míg az $x = \pm\sqrt{2 - \theta^2}$ esetén az $\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}$ értéke éppen $\frac{\sqrt{2\pi}}{e\theta} e^{\frac{\theta^2}{2}}$.
- Könnyű belátni, hogy a $0 < \theta < \sqrt{2}$ esetben az $\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}$ függvény a maximumát az $x = \pm\sqrt{2 - \theta^2}$, a minimumát pedig az $x = 0$ pontokra veszi fel.
- A $\theta \geq \sqrt{2}$ esetben az $\frac{f_{\mathcal{N}(0,1)}}{g_{\theta}}$ függvény a maximumát az $x = 0$ pontban éri el. Ezzel az optimális c_{θ} elutasítási konstans (3)-as alakjának helyességét beláttuk.



- Az optimalizálási feladat fennmaradt része már egyszerűbb.
- A c_θ függvénynek egyetlen minimuma van, amelyet a $\theta = 1$ paraméter esetén ér el.
- Ez a minimális érték

$$c_1 = \sqrt{\frac{2\pi}{e}},$$

amivel az

$$U_{c_\theta g_\theta}(Y) \leq f_{\mathcal{N}(0,1)}(Y)$$

leállási/elfogadási feltételt az

$$U \sqrt{\frac{2\pi}{e}} \frac{1}{\pi} \frac{1}{1+Y^2} \leq \frac{1}{\sqrt{2\pi}} e^{-\frac{Y^2}{2}}$$

formára, vagy az annál egyszerűbb

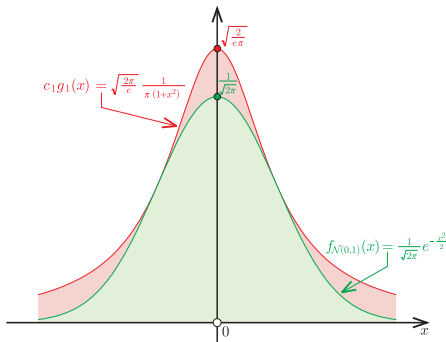
$$U \leq (1+Y^2) \frac{\sqrt{e}}{2} e^{-\frac{Y^2}{2}}$$

alakra hozható, ahol az U valószínűségi változó egyenletes eloszlású a $[0, 1]$ intervallumon, az Y valószínűségi változó pedig Cauchy-eloszlású $\theta = 1$ paraméterrel (lásd 3. labor 1. táblázatát).



Nemegyenletes eloszlású véletlenszám-generátorok

Elutasítás módszere: standard normális eloszlású számgenerátor a Cauchy-féle sűrűségfüggvény alapján



ábra. Az elutasítás módszerére épülő 5. algoritmus a standard normális eloszlás

$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$ sűrűségfüggvényét, valamint az azt majoráló $c_1 g_1(x)$ függvényt használja fel, ahol $g_1(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$ a $\theta = 1$ paraméterű Cauchy-féle

sűrűségfüggvényt, $c_1 = \sqrt{\frac{2\pi}{e}}$ pedig a Cauchy-féle sűrűségfüggvények

$\{g_\theta(x) = \frac{\theta}{\pi} \frac{1}{x^2 + \theta^2}, x \in \mathbb{R} : \theta > 0\}$ családjára nézve az optimális elutasítási konstans jelöli.



5. Algoritmus ($\mathcal{N}(0, 1)$ -eloszlású számgenerátor a Cauchy-féle sűrűségfüggvény alapján)

- **Bemenet:** az $n \geq 1$ természetes szám, amely a kimeneti mintavétel méretét határozza meg.
- **Kimenet:** az $X \sim \mathcal{N}(0, 1)$ valószínűségi változónak egy n -elemű független $\{X_i\}_{i=1}^n$ mintavétele.

1 $a \leftarrow \frac{\sqrt{e}}{2}$

2 Minden $i = 1, 2, \dots, n$ indexre **végezd el:**

3 **ismételd:**

4 generáld a független $U \sim \mathcal{U}([0, 1])$ és $V \sim \mathcal{U}([0, 1])$ valószínűségi változókat

5 $Y \leftarrow \tan(\pi V)$

6 $S \leftarrow Y^2$

7 **ameddig** $U \leq a(1 + S)e^{-\frac{S}{2}}$

8 $X_i \leftarrow Y$.

1. feladat

Kódoljátok 2. algoritmust az $f(x) = 168 \cdot x^5 \cdot (1 - x)^2$, $x \in [0, 1]$ sűrűségfüggvényű folytonos valószínűségi változó mintavételezésére!

2. feladat

Standard normális eloszlású valószínűségi változók mintavételezése a 4. és 5. algoritmusok segítségével. **Hogyan lehetne általános $\mathcal{N}(\mu, \sigma)$ normális eloszlású valószínűségi változókat ($\mu \in \mathbb{R}$, $\sigma > 0$) generálni ezekkel az algoritmusokkal? Implementáljátok is a válaszokat!**

3. feladat (+1 pont elméleti megoldással együtt)

- Mutassátok ki, hogy az $f_X(x) = 9xe^{-3x}$, $x \geq 0$ leképezés egy folytonos X valószínűségi változó sűrűségfüggvénye, majd keressétek meg a λ -paraméterű

$$\{g_\lambda(x) = \lambda e^{-\lambda x} : x \geq 0, \lambda > 0\}$$

exponenciális sűrűségfüggvények családjában azt a g_λ függvényt, amelyre teljesül az $f_X(x) \leq c_\lambda g_\lambda(x)$, $x \geq 0$ egyenlőtlenség úgy, hogy a $c_\lambda = \sup_{x \geq 0} \frac{f_X(x)}{g_\lambda(x)}$ elutasítási konstans optimális!

- Ábrázoljátok az f_X és $c_\lambda g_\lambda$ függvényeket, majd implementáljátok az elutasítás módszerét az X valószínűségi változó mintavételezésére!

