

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Sipos Nikolett

A Monte Carlo szimulációk gyakorlati alkalmazásai

BSc szakdolgozat
Alkalmazott matematikus szakirány

Témavezető:

Kovács Péter
Numerikus Analízis Tanszék



Budapest, 2016.

Köszönetnyilvánítás

Elsősorban szeretném megköszönni témavezetőmnek, Kovács Péternek, hogy elvállalta a konzulensi teendőket. Köszönöm, hogy a szakdolgozat írása során számos észrevétellel és tanáccsal segítette a munkámat, mindig türelmesen elmagyarázta a fogalmakat, tételeket és nagyon sok időt fordított a dolgozatom alapos átnézésére.

Továbbá szeretném megköszönni a barátaimnak a sok támogatást és érdeklődést a témával kapcsolatban. Külön köszönettel tartozom a sok biztatásért Réti Attilának, aki az első oldaltól az utolsóig nyomon követte a dolgozatom alakulását.

Tartalomjegyzék

1. Bevezetés	4
1.1. Előszó	4
1.2. Szimulációk	5
2. Numerikus integrálás	6
2.1. Motiváció	6
2.2. Kvadratúra formulák	8
2.3. Általános konvergenciatételek	8
3. Alkalmazás numerikus integrálásra	12
3.1. Valószínűségszámítási áttekintés	13
3.2. Monte Carlo integrálok kiszámítása	17
3.3. Példák Monte Carlo integrálásra	20
3.4. A Monte Carlo integrálás hibája	26
4. Szóráscsökkentő eljárások	29
4.1. A főrészt leválasztása	29
4.2. Az integrációs tartomány részekre bontása	30
4.3. Dimenziócsökkentés	31
4.4. A sűrűségfüggvény optimális megválasztása	32
4.5. Az integrandus szimmetrikussá tétele	33
5. Kitekintés	37
5.1. Véletlen szám generálási technikák	37
5.2. Egyéb alkalmazások	41

Jelölések

Jelölés	Magyarázat
dP	$dx dy$
$\ f\ _\infty$	$\sup_{x \in [a, b]} f(x) $
ξ, X, Y, Z	Valószínűségi változók
s_G	A G tartomány területe
\mathbb{P}_n	A legfeljebb n -edfokú polinomok tere
$f \in C[a, b]$	$f: [a, b] \rightarrow \mathbb{R}$ folytonos függvény
$A \in B(X, Y)$	$A: X \rightarrow Y$ folytonos lineáris operátor
$a_j^{(n)}$	$\int_a^b \ell_j^{(n)}$, ahol $\ell_j^{(n)}$ n alapontra illesztett Lagrange interpolációs polinom
Θ	paramétertér, legtöbbször véges dimenziós euklideszi tér részhalmaza
$X_N \xrightarrow{p} a$	X_N sztochasztikus értelemben konvergál a -hoz
$r(x)$	hibatag
$D(f)$	f értelmezési tartománya
S_n	n . részletösszeg (sor, bolyongás)

1. fejezet

Bevezetés

1.1. Előszó

A dolgozatomban a Monte Carlo szimulációk különféle gyakorlati alkalmazásait mutatom be. Gyakorlatban elterjedt, hogy a szimulációt használják egyes matematikai, fizikai, illetve gazdasági számítások modellezésére. A dolgozatomban ezek valószínűségszámítási, statisztikai, valamint numerikus analízisbeli háttérével foglalkozunk. A második fejezetben integrálszámítással kapcsolatos alkalmazhatóságát vizsgáljuk, összehasonlítva az analízisből, illetve numerikus analízisből tanult módszerekkel. A második fejezetben példákon keresztül röviden áttekintjük a numerikus integrálás konvencionális módszereit. Ezt fogjuk összevetni a harmadik fejezetben a Monte Carlo módszer eredményeivel és a hibaképleteivel. Majd a 4. fejezetben a Monte Carlo integrálás hatékonyságát fogjuk vizsgálni.

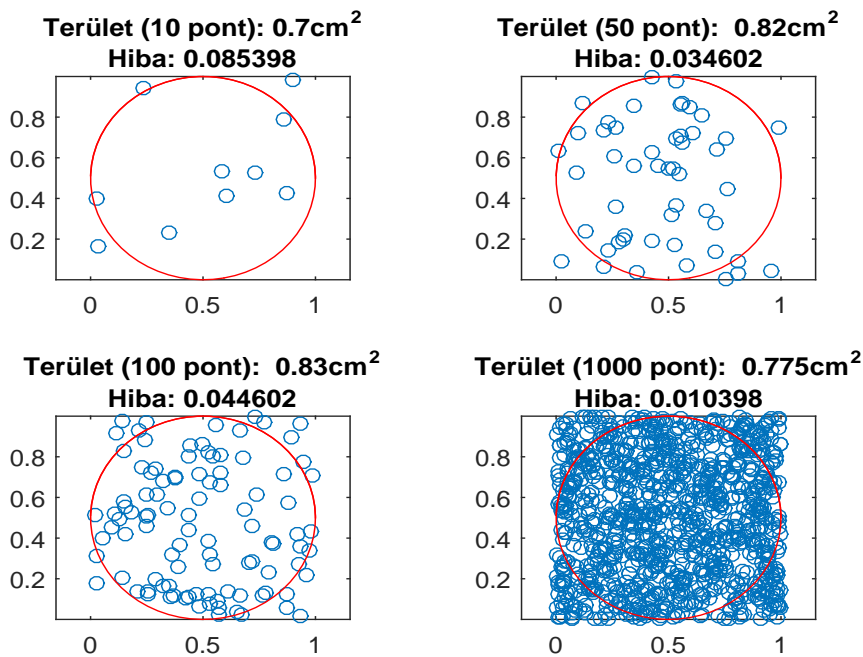
Fontos megemlíteni, hogy a Monte Carlo módszer véletlenszerű mintavételen alapul, azaz pl. véletlenszerűen kell kiválasztanunk számokat egy megadott tartományból, amikor egy integrálási feladatot szeretnénk elvégezni. Azonban tudjuk, hogy a véletlen számok számítógépes generálása sem annyira véletlenszerű, mindegyik mögött felfedezhető egy-egy algoritmus. Ezzel kapcsolatosan szó lesz az ötödik fejezetben a véletlen szám generátorokról, azok működéséről. A Monte Carlo szimulációban összehasonlítjuk az eredményeket, amiket az egyes véletlen szám generátorok által kapott mintából megismerhetünk. Végül kitérünk a Monte Carlo szimuláció gazdasági folyamatokban történő alkalmazására.

1.2. Szimulációk

A szimulációk olyan vizsgálati módszert jelentenek, melyek egy folyamat, illetve rendszer viselkedését, várható kimenetelét vizsgálják. A szimuláció egy absztrakt, matematikailag definiált modellt használ, hogy vizsgálja a rendszer működését. Azaz egy algoritmus lépéseit követve szolgáltat adatokat, illetve mintát. A szimulációk célja, hogy a folyamatokat valósághűen modellezzük és ki tudjuk értékelni az állapotváltozásokat, illetve a mintákat statisztikailag össze tudjuk hasonlítani.

A Monte Carlo módszer (röviden MC módszer) egy speciális szimulációs módszer, mellyel a valószínűségszámítás és a statisztika elemeit hívjuk segítségül, majd numerikusan értékeljük ki a kapott eredményeket. A módszer lényegében véletlenszerű mintavételen alapul, mellyel elég nagy elemű minta esetén meg tudunk becsülni határozott integrálokat, egyes kockázati faktorok (pl. kockázatosított érték: VaR) becslésére is alkalmazható a gazdasági életben, valamint számos becsléshez (pl. a π közelítésére) is felhasználható a szimuláció.

Az 1.1 ábrán 10, 50, 100 és 1000 pont beszórásával futattam a szimulációt, hogy közelítse az $(0,5; 0,5)$ középpontú $0,5$ cm sugarú kör területét és a hibát is számolja.



1.1. ábra. Monte Carlo szimuláció a $0,5$ cm sugarú kör területének kiszámítására

2. fejezet

Numerikus integrálás

A gyakorlati életben sokszor előfordul, hogy egy határozott integrál esetén nem tudjuk az analízisből tanult módon kiszámolni az eredményt. Gondoljunk pl. egy olyan függvényre, amelynek nem létezik primitív függvénye. Ekkor az integrált numerikus módszerekkel próbáljuk közelíteni.

2.1. Motiváció

Olyan határozott integrálok kiszámítását fogjuk megnézni, ahol a számítógép által számolt numerikus közelítés nem konvergens vagy nem optimális.

Első példa

$\int_0^1 e^{-x^2} dx$ integrált szeretnénk kiszámítani, 10^{-6} pontossággal.

Mivel a függvénynek nem létezik primitív függvénye, nem tudjuk alkalmazni a Newton-Leibniz formulát, ezért szeretnénk egy másik módszert alkalmazni. Analízisből ismeretes, hogy e^x hatványsorba fejthető a 0 körül, konvergenciatartománya végtelen és $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ alakban felírható. Ehhez hasonlóan hatványsorba tudjuk fejteni az e^{-x^2} függvényt is. Azaz felírva az első öt tagot, kapjuk, hogy:

$$e^{-x^2} = \sum_{k=0}^{\infty} \frac{(-x^2)^k}{k!} = 1 - \frac{x^2}{1} + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} + r(x),$$

ahol $r(x)$ a hibatag.

Ekkor az integrált a következőképp közelítjük:

$$\begin{aligned} I &= \int_0^1 e^{-x^2} dx = \int_0^1 1 dx - \int_0^1 \frac{x^2}{1} dx + \int_0^1 \frac{x^4}{2} dx - \int_0^1 \frac{x^6}{6} dx + \int_0^1 r(x) dx = \\ &= 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + R(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2 * k + 1) * k!} \end{aligned}$$

Szintén ismeretes, hogy ez a hatványsor egy Leibniz típusú sor, így konvergens, mivel $a_n = \frac{1}{(2 * k + 1) * k!}$ és $\lim_{n \rightarrow \infty} a_n = 0$. Ekkor úgy kell megválasztanunk n -et, hogy $\left| I - \sum_{k=0}^n (-1)^k * \frac{1}{(2 * k + 1) * k!} \right| \leq 10^{-6}$ teljesüljön.

A Leibniz típusú hatványsorok konvergenciasebességét le tudjuk vezetni a következőképpen:

$$S_{2 * n - 1} \leq I \leq S_{2 * n}.$$

Ebből következik, hogy:

$$|S_{2 * n - 1} - I| \leq S_{2 * n} - S_{2 * n - 1} = a_{2 * n} \quad \text{és} \quad |S_{2 * n} - I| \leq S_{2 * n} - S_{2 * n - 1} = a_{2 * n}.$$

Így a közelítés hibájára az alábbi becslést kapjuk:

$$|S_n - I| \leq a_n = \frac{1}{(2 * k + 1) * k!} \leq 10^{-6}$$

Második példa

Az a feladatunk, hogy kiszámoljuk az $\int_0^1 \cos(x) dx$ integrált. A Newton-Leibniz formulából azonnal adódik a megoldás:

$$\int_0^1 \cos(x) dx = \sin(1)$$

Ezt viszont a számítógép szintén hatványsorból fogja közelíteni. A hatványsorba fejtett függvény azonban nem mindig konvergens és sokszor a konvergenciasebesség sem optimális számunkra.

A numerikus integrálás alapötlete, hogy a függvények közelítésére használt interpolációs polinomokat (Lagrange vagy Hermite interpolációs polinom) használjuk az integrálandó függvény közelítésére. Kérdés, hogy ez az integrálnak is közelítése-e?

2.2. Kvadratúra formulák

2.2.1. Definíció (Kvadratúra formula). Legyen $f \in C[a, b] \rightarrow \mathbb{R}$. Tegyük fel, hogy $L_n \in \mathbb{P}_n$, ahol L_n az f függvény Lagrange interpolációs polinomja az $\mathbf{a} = x_0 < x_1 < \dots < x_n = \mathbf{b}$ alappontokon, ahol $\mathbf{a}_i = \int_a^b \ell_i(x) dx$. Ekkor az alábbi formulát interpolációs kvadratúra formulának nevezzük:

$$I(f) = \int_a^b f(x) dx \approx \int_a^b L_n(x) dx = \int_a^b \sum_{i=0}^n f(x_i) * \ell_i(x) dx = \sum_{i=0}^n f(x_i) * \mathbf{a}_i = I_n(f) \quad (2.1)$$

2.3. Általános konvergenciatételek

Ebben a részben meg fogjuk nézni, hogy az interpolációs típusú kvadratúra formulával kapott közelítés a valóságban is a függvény integráljához konvergál-e. Tehát az a kérdés, hogyha növeljük az illesztett polinom fokszámát, akkor a kvadratúra formulával kapott becslés tartani fog-e az eredeti integrálhoz.

Megsejthetjük a polinomillesztés kapcsán, hogy nem mindig lesz igaz a fenti állítás. Gondoljunk pl. arra az esetre, amikor az interpolációs polinom sem tart az eredeti függvényhez (Faber, Marcinkiewicz tétel), ekkor a kvadratúra formulával felírt közelítés sem fog a függvény integráljához tartani. Azonban bizonyos feltételek mellett garantálni lehet, hogy a polinom integrálja is tartson az interpolált függvény integráljához.

Ahhoz, hogy belássuk a konvergenciát, szükségünk lesz néhány tételre, amelyek alkalmazásával el fogunk jutni addig, hogy a kvadratúra formulák a legfeljebb n -edfokú polinomokra pontosak. Ennek megfelelően az [1] jegyzet alapján áttekintjük a kapcsolódó elméletet.

2.3.1. Definíció (Norma). Legyen X vektortér \mathbb{K} felett, ahol $\mathbb{K} = \mathbb{C}$ vagy $\mathbb{K} = \mathbb{R}$. Egy $\|\cdot\|: X \rightarrow \mathbb{R}^+$ függvényt normának nevezünk, ha teljesíti az alábbi normaaxiómákat:

- i.) minden $x \in X$ esetén $\|x\| \geq 0$, és $\|x\| = 0 \Leftrightarrow x = 0$,
- ii.) minden $\lambda \in \mathbb{K}$ és $x \in X$ esetén $\|\lambda * x\| = |\lambda| * \|x\|$,
- iii.) minden $x, y \in X$ esetén $\|x + y\| \leq \|x\| + \|y\|$ (háromszög egyenlőtlenség).

Ekkor $(X, \|\cdot\|)$ párt normált térnek nevezzük.

2.3.2. Definíció (Banach-tér). Egy normált teret Banach-térnek, nevezünk, ha teljes, azaz ha minden Cauchy sorozat konvergens.

2.3.3. Definíció (Lineáris funkcionál). Az $A : X \rightarrow \mathbb{K}$ lineáris leképezéseket lineáris funkcionáloknak nevezzük.

2.3.4. Definíció (Korlátos leképezés). Egy $A : X \rightarrow Y$ lineáris leképezést korlátosnak nevezünk, ha létezik olyan $M \geq 0$ állandó, hogy:

$$\|Ax\| \leq M * \|x\| \quad \forall x \in D(A). \quad (2.2)$$

2.3.5. Tétel. Egy lineáris leképezés pontosan akkor folytonos, ha korlátos.

2.3.6. Definíció (Operátornorma). Ha $A \in B(X, Y)$, akkor legyen

$$\|A\| := \sup\{\|Ax\| : x \in X, \|x\| \leq 1\} \quad (2.3)$$

az A úgynevezett operátornormája, vagy normája.

2.3.7. Tétel. Ha $A \in B(X, Y)$, ahol X és Y normált terek és $\|Ax\| \leq C * \|x\| \quad \forall x \in X$, akkor $\|A\| \leq C$. Sőt, $\|A\|$ épp az ilyen C -k infimumával egyezik meg, azaz $\|A\| = \inf C$. Ha tehát C olyan, hogy alkalmas $x_0 \in X$ -re $\|Ax_0\| = C * \|x_0\|$ és fennáll, hogy $\|Ax\| \leq C * \|x\|$, akkor $\|A\| = C$.

Ebből már speciálisan ki tudjuk számítani az integrál és a kvadratúra formula operátornormáját is.

2.3.8. Tétel (Integrál- és kvadratúra formula operátornormája). I és I_n folytonos lineáris funkcionálok a $C[a, b]$ Banach-téren. Ekkor $\|I\| = b - a$ és $\|I_n\| = \sum_{j=0}^n |a_j|$.

Bizonyítás. Az 2.3.7 tételt felhasználva kapjuk, hogy:

$$|I(f)| = \left| \int_a^b f(x) dx \right| \leq |b - a| * \|f\|_\infty \quad \Rightarrow \quad \|I\| \leq b - a,$$

és $f \equiv 1$ függvényt választva:

$$\|I\| = b - a.$$

Hasonlóan belátható, hogy:

$$|I_n(f)| = \left| \sum_{j=0}^n a_j * f(x_j) \right| \leq \sum_{j=0}^n |a_j| * |f(x_j)| \leq \sum_{j=0}^n |a_j| * \|f\|_\infty,$$

a háromszög egyenlőtlenség miatt.

Alkalmas f -re (az x_0, x_1, \dots, x_n alappontokon a $\text{sign}(a_0), \text{sign}(a_1), \dots, \text{sign}(a_n)$) értékekre illeszkedő, szakaszonként elsőfokú polinomra az egyenlőség igaz, mert $\|f\|_\infty = 1$ nyilvánvalóan teljesül (eltekintve a triviális $a_0 = a_1 = \dots = a_n$ választástól). Ekkor:

$$|I_n(f)| = \left| \sum_{j=0}^n a_j * \text{sign}(a_j) \right| = \sum_{j=0}^n |a_j| \quad \Rightarrow \quad \|I_n\| = \sum_{j=0}^n |a_j|.$$

□

2.3.9. Definíció (Pontonkénti- és egyenletes korlátosság). *Legyenek X, Y normált terek. Egy $(A_n) \subset B(X, Y)$ operátorsorozat pontonként korlátos, ha minden $x \in X$ esetén $(A_n x) \subset Y$ korlátos, illetve egyenletesen korlátos, ha $(\|A_n\|) \subset \mathbb{R}$ korlátos számsorozat.*

2.3.10. Tétel (Banach-Steinhaus tétel). *Legyen X Banach, Y normált tér, $(A_n) \subset B(X, Y)$. Ekkor $\sup_n \|A_n x\| < \infty \quad \forall x \in X \Leftrightarrow \sup_n \|A_n\| < \infty$. Azaz (A_n) pontosan akkor korlátos pontonként, ha egyenletesen is.*

2.3.11. Tétel. *Legyen X Banach-tér, Y normált tér, valamint A_n és $A \subset B(X, Y)$. $A_n x \rightarrow Ax \quad \forall x \in X$, ha:*

- i.) A_n pontonként (és egyenletesen is) korlátos,
- ii.) A_n pontonként konvergens egy X -ben sűrű halmazon.

2.3.12. Tétel (Weierstrass 1. approximációs tétele). *Legyen $[a, b]$ tetszőleges, zárt intervallum, ahol $a, b \in \mathbb{R}$. Legyen f tetszőleges, $[a, b]$ intervallumon folytonos, valós függvény. Ekkor $\forall \epsilon > 0$ -hoz megadható egy olyan $p \in \mathbb{P}$ polinom, hogy $|f(x) - p(x)| < \epsilon \quad \forall x \in [a, b]$ -re.*

2.3.13. Következmény (Pólya-Szegő tétel). *Speciálisan $X := C[a, b]$, $Y := \mathbb{R}$, $A_n := I_n$, $M := \mathbb{P}$ (Polinomok tere). Tudjuk, hogy \mathbb{P} sűrű $C[a, b]$ -ben, Weierstrass 1. approximációs tétele szerint. Így azt kapjuk, hogy az I_n kvadratúraszorozat pontosan akkor tart az I integrálhoz, ha $\forall f \in C[a, b]$ folytonos függvény esetén:*

- i.) $\sum_{j=0}^n |a_j^{(n)}| \leq C$, alkalmas $0 \leq C < \infty$ mellett, $\forall n \in \mathbb{N}$ -re,
- ii.) $I_n(p) \rightarrow I(p) \quad \forall p \in \mathbb{P}$ polinomra.

2.3.14. Megjegyzés. *A \mathbb{P} trigonometrikus polinomokat is jelenthet (Weierstrass 2. approximációs tétele alapján.)*

Így jutunk el az egyik legfontosabb következményig, miszerint a kvadratúra formulák legfeljebb n -edfokú polinomokra pontosak.

2.3.15. Következmény (Pólya-Szegő tételének speciális esete). *Ha*

i.) $\sum_{j=0}^n |a_j^{(n)}| \leq C$, azaz I_n egyenletesen korlátos,

akkor I_n pontos minden legfeljebb n -edfokú polinomra: $I_n(p) = I(p) \quad \forall p \in \mathbb{P}_n$.

Bizonyítás. Ekkor ugyanis tetszőleges $p \in \mathbb{P}$ polinomra, minden elég nagy n mellett $p \in \mathbb{P}_n$, így $I_n(p) = I(p) \Rightarrow I_n(p) \rightarrow I(p)$ pontonként \mathbb{P} -ben. \square

Másik fontos következmény, ami kimondja, hogy a legfeljebb n -edfokú polinomok esetén a kvadratúra formula operátora tart az integráloperátorhoz pontonként.

2.3.16. Következmény (Sztyeklov tétele). *Ha*

i.) $a_j^{(n)} > 0 \quad \forall j, n \in \mathbb{N}$,

ii.) I_n pontos minden legfeljebb n -edfokú polinomra,

akkor $I_n \rightarrow I$ pontonként $C[a, b]$ -n.

Bizonyítás. Ekkor ugyanis I_n egyenletesen korlátos:

$$\|I_n\| = \sum_{j=0}^n |a_j^{(n)}| = \sum_{j=0}^n a_j^{(n)} = \sum_{j=0}^n a_j^{(n)} * 1 = \int_a^b 1 dx = b - a,$$

mert I_n pontos az azonosan 1 függvényre, tetszőleges $n \in \mathbb{N}$ -re. \square

2.3.17. Megjegyzés. *A tétel a konvergenciasebességről semmit sem mond, így ezek a tételek elméleti jelentőségűek. A numerikus integrálás jól használható alacsony dimenzióban és kevés kiértékelés esetén. Azonban amikor nagyobb dimenzióban keressük az integrált, ugyanahhoz a pontossághoz kevesebb függvénykiértékelésre van szükség, mint a kvadratúra formuláknál. Ekkor térünk át a Monte Carlo integrálás alkalmazására. A Monte Carlo integrálás sikeresen alkalmazható olyan esetekben is, amikor egy többdimenziós integrált szeretnénk számolni, de a tartomány nem szabályos. A második fejezetben ennek az összehasonlítása következik.*

3. fejezet

Alkalmazás numerikus integrálásra

A Monte Carlo integrálás (röviden MC integrálás) egy olyan eljárás, mely során határozott integrálokat tudunk numerikus módszerek segítségével közelíteni. A numerikus analízis egyik alapproblémája a numerikus integrálás, mely szinte minden tudományterületen megjelenik. A klasszikus numerikus integrálási módszerek, az integrációs típusú kvadratúra formulák eredményesen használhatóak alacsony dimenzióban és az alappontok számának növelésével a közelítés hibáját is tetszőlegesen kicsire csökkenthetjük. Ebben a fejezetben az [3] és a [4] jegyzet alapján átismételjük a Monte Carlo módszert, majd alkalmazzuk függvények határozott integráljának kiszámítására, majd összevetjük a két módszer pontosságát.

3.0.1. Tétel (A közelítés hibája).

$$|e_n(f)| = \left| \int_a^b f(x) dx - \sum_{i=1}^n f(x_i) * a_i \right|.$$

3.0.2. Megjegyzés (Speciális, interpolációs kvadratúra formulák esetén a hiba). Az alábbi felső becslést írhatjuk fel:

$$|e_n(f)| \leq \frac{M_{n+1}}{(n+1)!} * \int_a^b |\omega_{n+1}(x)| dx,$$

ahol $M_{n+1} = \sup_{x \in [a,b]} |f^{(n+1)}(x)| = \|f^{(n+1)}\|_\infty$, ha $f \in C^{n+1}[a, b]$.

3.0.3. Tétel. $\lim_{n \rightarrow \infty} e_n(f) = 0$, $\forall f \in C[a, b]$, ha $\sup_{n \in \mathbb{N}} \sum_{i=1}^n |a_i^{(n)}| < \infty$.

3.0.4. Tétel (Elemi kvadratúra formulák és hibaformuláik). *Nézzük az alábbi formulákat:*

*Érintőformula: ha $f \in C^2[a, b]$ és $I_0f = f\left(\frac{a+b}{2}\right) * (b - a)$*

$$\left| \int_a^b f(x) - I_0f \right| \leq \frac{1}{24} * \|f^{(2)}\|_\infty * (b - a)^3$$

*Trapéz formula: ha $f \in C^2[a, b]$ és $I_1f = \frac{b-a}{2} * (f(a) + f(b))$*

$$\left| \int_a^b f(x) - I_1f \right| \leq \frac{1}{12} * \|f^{(2)}\|_\infty * (b - a)^3$$

*Simpson formula: ha $f \in C^4[a, b]$ és $I_2f = \frac{b-a}{6} * (f(a) + 4 * f\left(\frac{a+b}{2}\right) + f(b))$*

$$\left| \int_a^b f(x) - I_2f \right| \leq \frac{1}{720} * \|f^{(4)}\|_\infty * (b - a)^5$$

3.0.5. Megjegyzés. *A kvadratúra formulák azonban magasabb dimenzióban nem használhatók eredményesen, mert a kiértékelések számával exponenciálisan nő a lépések száma a futtatás során. Pl. ahhoz, hogy 10 kiértékelést el tudjunk végezni 100 dimenzióban, szükséges 10^{100} lépés. Ez a futtatást lelassítja egy idő után. Szintén problémát okoz, ha a tartomány, nem egymásba ágyazott integrálokból áll.*

3.1. Valószínűségszámítási áttekintés

A Monte Carlo integrálás alap gondolata az, hogy egymástól független véletlenszerűen megválasztott mintából közelítjük az integrált, amihez általában egyenletes eloszlás szerint választunk pontokat, de léteznek más véletlenszám generátorok is, (Mersenne-Twister, lineáris kongruencia generátorok) melyekről szó lesz a későbbi fejezetekben.

Először vezessünk be néhány alapvető fogalmat, amire szükségünk lesz a várható érték becsléséhez a [7] és [8] jegyzetek alapján.

3.1.1. Definíció (Statisztikai mező). *Az $(\Omega, \mathcal{A}, \mathbb{P})$ hármast statisztikai mezőnek hívjuk, ahol Ω nemüres halmaz (eseménytér), \mathcal{A} egy σ -algebra (események családja), \mathbb{P} pedig a szóba jöhető valószínűségi mértékek családja. Azaz $\mathbb{P} = \{\mathbb{P}_\vartheta \mid \vartheta \in \Theta\}$, ahol \mathbb{P}_ϑ valószínűségi mértékek és Θ a paramétertér.*

3.1.2. Megjegyzés. *A Monte Carlo integrálás bevezetése során Θ egy véges dimenziós euklideszi tér részhalmaza, pl. $\Theta \subset \mathbb{R}^k$.*

3.1.3. Definíció (N elemű minta). Egy $X=(X_1, \dots, X_N): \Omega \rightarrow \mathcal{X} \subset \mathbb{R}$ valószínűségi változót (N elemű) mintának nevezzük, ahol \mathcal{X} a mintatér, N pedig a minta nagysága vagy elemszáma, a X_i koordináták pedig a minta elemei.

3.1.4. Tétel (Hincsin tétele). Legyen X_1, X_2, \dots, X_N páronként független, azonos eloszlású, melyekre igaz, hogy $E(X_i) = \mathbf{a} < +\infty$ és legyen $\bar{X}_N = \frac{1}{N} * \sum_{i=1}^N X_i$. Ekkor fennáll, hogy $\bar{X}_N \xrightarrow{P} \mathbf{a}$, ha $N \rightarrow \infty$ azaz $\lim_{N \rightarrow \infty} P(|\bar{X}_N - \mathbf{a}| \geq \epsilon) = 0, \forall \epsilon > 0$. Tehát Hincsin tétele kimondja a gyenge konvergenciát, ha X_1, \dots, X_N -nek véges a várható értéke, ami gyengébb feltétel, mint a nagy számok gyenge törvényénél, ugyanis annak a feltétele a második momentum végeessége.

3.1.5. Tétel (A várható érték becslése). Legyen X tetszőleges valószínűségi változó, véges várható értékkel: $E(X) = \mathbf{a}$. Az \mathbf{a} paraméter becslésére válasszuk a X_1, X_2, \dots, X_N független mintát és tekintsük a $\bar{X}_N = \frac{1}{N} * \sum_{i=1}^N X_i$ átlagot. Ekkor Hincsin tétele szerint $\bar{X}_N \xrightarrow{P} \mathbf{a}$, ha $N \rightarrow \infty$. Ha feltesszük, hogy a szórás is létezik, azaz, hogy:

$$\sigma^2 = \sigma^2(X) = E(X^2) - (E(X))^2 < +\infty, \quad (3.1)$$

akkor érvényes a centrális határeloszlás tétel:

$$\lim_{N \rightarrow \infty} P\left(\sqrt{N} * \frac{\bar{X}_N - \mathbf{a}}{\sigma} \in (x_1, x_2)\right) = \Phi(x_2) - \Phi(x_1),$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2} * \pi} * \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (3.2)$$

Ebből következik, hogy:

$$\begin{aligned} \lim_{N \rightarrow \infty} P\left(|\bar{X}_N - \mathbf{a}| < x * \frac{\sigma}{\sqrt{N}}\right) &=: H(x), \\ H(x) &= \frac{2}{\sqrt{2} * \pi} \int_0^x e^{-\frac{t^2}{2}} dt = 2 * \Phi(x) - 1. \end{aligned} \quad (3.3)$$

Legyen $x = x_\beta$ a $H(x) = \beta$ megoldása. $H(x) = \beta = \lim_{N \rightarrow \infty} P(|\bar{X}_N - \mathbf{a}| < x * \frac{\sigma}{\sqrt{N}})$ miatt fennáll, hogy:

$$|\bar{X}_N - \mathbf{a}| < x_\beta * \frac{\sigma}{\sqrt{N}} \quad (3.4)$$

teljesülésének valószínűsége β .

3.1.6. Megjegyzés. Gyakran alkalmazzák a következő megbízhatósági szinteket: $\beta = 0,997$ és $x_\beta = 3$ vagy $\beta = 0,95$ és $x_\beta = 1,96$. Ezek a megbízhatósági szintek szintén

megjelennek a Monte Carlo szimuláció egy másik alkalmazásában, a kockázatos érték számításában, ahol 99%-os és 97%-os megbízhatósági szintekkel fogunk dolgozni. Erről az 5. fejezetben lesz szó bővebben.

3.1.7. Megjegyzés. Az a probléma merül fel ebben az esetben, hogy legtöbbször nem ismerjük az X szórását, amikor a várható értéket szeretnénk számolni. Szóval a speciális eseteken kívül nem tudjuk meghatározni a szórást. Azonban a várható értéket tudjuk becsülni is, amihez a $\sum_{i=1}^N X_i^2$ összeget kell kiszámolnunk. Ugyanis statisztikai megfontolások alapján tudjuk, hogy:

$$E(X_i^2) \approx \frac{1}{N} \sum_{i=1}^N X_i^2,$$

ezért

$$\sigma^2(X) \approx \frac{1}{N} \sum_{i=1}^N X_i^2 - \bar{X}^2$$

képlettel becsülhető a szórásnégyzet. Ezt továbbra sem tudjuk hatékonyan használni. Ezért fontos, hogy definiáljunk egy fogalmat, a torzítatlan becslés fogalmát.

3.1.8. Definíció (k dimenziós statisztika). A mintatéren megadott $T : X \rightarrow \mathbb{R}^k$ függvényt, illetve magát a $T = T(X)$ valószínűségi változót k dimenziós statisztikának nevezzük.

3.1.9. Megjegyzés (Gyakran használt statisztikák). Nézzük az alábbi statisztikákat:

- 1) $T(X) = \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ a minta tapasztalati mintaátlaga.
- 2) $T(X) = S_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$ a minta tapasztalati szórásnégyzete.
- 3) $T(X) = (X_1^{(n)}, X_2^{(n)}, \dots, X_n^{(n)})$ a minta rendezett mintája, ahol $X_1^{(n)} < \dots < X_n^{(n)}$.
- 4) $T(X) = X_n^{(n)} - X_1^{(n)}$ a minta terjedelme.

3.1.10. Definíció (Torzítatlan becslés). Legyen az X eloszlásának egy függvénye $\Psi(\vartheta)$, ahol ϑ az X paramétere. Azt mondjuk, hogy $\Psi(\vartheta)$ függvény torzítatlan becslése a $T(X)$ statisztika, ha

$$E_{\vartheta}(T(X)) = \Psi(\vartheta) \quad \forall \vartheta \in \Theta.$$

3.1.11. Megjegyzés. Belátható, hogy a $\sigma^2(X)$ fenti becslése helyett jobban alkalmazható az $(s_N^*)^2 = \frac{1}{N-1} * \sum_{i=1}^N (X_i - \bar{X}_N)^2$ becslés, mivel ez torzítatlan becslése a σ^2 -nek (ennek részletes levezetése [2] cikkben megtalálható). Így meg tudjuk becsülni a a közelítés hibáját a szórás közelítő kiszámítása nélkül.

Tegyük fel, hogy σ^2 létezik. Legyen $N = m * N_1$, $m_1 \geq 3$ nem nagy szám, N_1 viszont olyan nagy, hogy az

$$X_j = \frac{1}{N_1} \sum_{i=1}^{N_1} X_{i+(j-1)*N_1} \quad (j = 1, \dots, m)$$

átlagokról joggal feltehető, hogy közel normális eloszlásúak.

3.1.12. Tétel (Fisher tétele). Ha X_1, X_2, \dots, X_m független, azonos eloszlású valószínűségi változók és $E(X_j) = a$ ($j = 1, \dots, m$), ekkor nézzük a mintaátlagot:

$$\bar{X}_m = \frac{1}{m} * \sum_{k=1}^m X_k,$$

és a tapasztalati szórásnégyzetet, amit az alábbi módon definiáltunk:

$$(s_m^*)^2 = \frac{1}{m} \sum_{k=1}^m (X_k - \bar{X})^2.$$

Nézzük ekkor az alábbi változót:

$$t = \sqrt{m-1} * \frac{\bar{X}_{m-1} - a}{s_{m-1}^*}.$$

Ez valójában egy valószínűségi változó: $m-1$ szabadságfokú Student-eloszlású.

3.1.13. Tétel (m szabadságfokú Student-eloszlás sűrűségfüggvénye). Az alábbi módon határozzuk meg, ahol $\Gamma(\xi)$ a Gamma eloszlás sűrűségfüggvénye.

$$s_m(x) = c_m \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}} \quad -\infty < x < \infty,$$

$$c_m = \frac{1}{\sqrt{\pi * m}} * \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}. \quad (3.5)$$

Ezt pedig, ha visszahelyettesítjük, megkapjuk, hogy

$$P\left(|\bar{X}_N - a| < x * \sqrt{\frac{s_{m-1}^2}{m-1}}\right) \approx 2 * \int_0^x s_{m-1}(y) dy. \quad (3.6)$$

3.2. Monte Carlo integrálok kiszámítása

Ebben a részben egyszerűbb Monte Carlo integrálokat fogunk kiszámítani az alábbi alakból: $\int_G f(\mathbf{P})d\mathbf{P}$. Ehhez fel fogjuk használni a [5] és [6] jegyzetek tételeit.

3.2.1. Tétel (Megfelelő alak keresése). *Legyen G a sík egy tetszőleges tartománya. Jelölje \mathbf{P} a sík egy tetszőleges pontját és legyen a $\mathbf{p}(\mathbf{P})$ a G tartományon értelmezett valószínűségi eloszláshoz tartozó sűrűségfüggvény. Mivel \mathbf{p} sűrűségfüggvény, ezért fennáll rá az alábbi két tulajdonság:*

- $\mathbf{p}(\mathbf{P}) \geq 0$
- $\int_G \mathbf{p}(\mathbf{P})d\mathbf{P} = 1$.

3.2.2. Tétel. *Minden integrál felírható szorzatalakban.*

Bizonyítás. Vezessük be az alábbi integrált:

$$I(f) = \int_G f(\mathbf{P}) * \mathbf{p}(\mathbf{P})d\mathbf{P}. \quad (3.7)$$

Erre az alakra fogunk minden integrált visszavezetni, amivel a későbbiekben foglalkozunk. Ezt azért tehetjük meg, mert minden integrál felírható az alábbi szorzatalakban:

$$\int_G g(\mathbf{P})d\mathbf{P} = \int_G h(\mathbf{P}) * \mathbf{p}(\mathbf{P})d\mathbf{P}, \quad (3.8)$$

ha a h függvényt megfelelően választjuk meg, azaz: $h(\mathbf{P}) = \frac{g(\mathbf{P})}{\mathbf{p}(\mathbf{P})}$. □

Visszatérünk az alapproblémához, a $\int_G f(\mathbf{P})d\mathbf{P}$ integrált szeretnénk kiszámítani. Feltesszük, hogy G területe (s_G) véges. Vezessük be a $\mathbf{p}_1(\mathbf{P}) = \frac{1}{s_G}$ függvényt. Ez a G -n egyenletes eloszlású valószínűségi változó sűrűségfüggvénye.

(Emlékeztető: 1 dimenzióban az $[a, b]$ intervallumon egyenletes eloszlású valószínűségi változó sűrűségfüggvényét a következőképp definiáltuk: $f(x) = \frac{1}{b-a}$ ha $x \in [a, b]$.)

(3.7) alakba írva az integrált, megkapjuk, hogy:

$$\int_G f(\mathbf{P})d\mathbf{P} = \int_G f_1(\mathbf{P}) * \mathbf{p}_1(\mathbf{P}), \quad (3.9)$$

ahol $f_1(\mathbf{P}) = s_G * f(\mathbf{P})$. Ezzel az átírással továbbra sem fogjuk megváltoztatni a feladat megoldását.

Vezessük be az X valószínűségi változót úgy, hogy az a G tartományon legyen definiálva. Legyen X sűrűségfüggvénye a $p(P)$. Legyen továbbá $Y = f(X)$ és X_1, \dots, X_N legyenek X független realizációi. Az $Y_i = f(X_i)$.

Tekintsük a $\Theta_N = \frac{1}{N} \sum_{i=1}^N Y_i$ összeget. $E(|Y|) < \infty$ esetén, 3.1.4 tételt felhasználva kapjuk, hogy $\forall \epsilon > 0$ -ra:

$$\lim_{N \rightarrow \infty} P(|\Theta_N - I(f)| \geq \epsilon) = 0. \quad (3.10)$$

Azaz a fenti integrált becsülhetjük az alábbi alakban:

$$E(|Y|) = \int_G |f(P)| * p(P) dP. \quad (3.11)$$

Az integrált máshogyan is becsülhetjük:

$$0 \leq f(x, y) \leq c \text{ és } P = (x, y) \in G. \quad (3.12)$$

Legyen $\tilde{G} := G \times (0, c)$ és legyen (X, Y) olyan eloszlás a G -n, ami $p(x, y)$ sűrűségfüggvénnyel rendelkezik, Z pedig a $[0, c]$ intervallumon egyenletes eloszlású. Feltehetjük, hogy Z és (X, Y) függetlenek, hiszen mindig tudunk így választani valószínűségi változókat. Ekkor a függetlenségből adódóan $\rho = (X, Y, Z)$ vektorváltozó sűrűségfüggvénye az alábbi módon fejezhető ki:

$$\tilde{p}(x, y, z) = \frac{1}{c} * p(x, y) \quad (x, y, z) \in \tilde{G}. \quad (3.13)$$

Most nézzük az előbbi vektorváltozónkat $\rho = (X, Y, Z)$ és vegyük ennek N darab független realizációját: $\rho_1, \rho_2, \dots, \rho_N$ -et. Vezessük be ν -t azon ρ_i vektorok számának tárolására, melyekre fennáll az alábbi tulajdonság:

$$Z_i < f(X_i, Y_i) \quad \rho_i = (X_i, Y_i, Z_i). \quad (3.14)$$

Most tekintsük a $Z < f(X, Y)$ esemény valószínűségét:

$$P(Z < f(X, Y)) = \int_G \int_0^{f(x, y)} \tilde{p}(x, y, z) dx dy dz = \frac{1}{c} \int_G f(x, y) * p(x, y) dx dy = \frac{1}{c} I(f). \quad (3.15)$$

Vezessük be az alábbi jelölést erre az értékre:

$$p = \frac{1}{c} * I(f), \quad (3.16)$$

ekkor azt kaptuk, hogy $\nu \sim \text{Binom}(p)$, tehát:

$$P(\nu = m) = \binom{N}{m} * p^m * (1 - p)^{N-m} \quad (m = 0, 1, \dots, N). \quad (3.17)$$

Ahhoz, hogy ki tudjuk számítani az integrált, a nagy számok törvényének Bernoulli-féle alakját fogjuk felhasználni. A tétel szerint egy esemény bekövetkeztének elméleti valószínűsége p és az esemény tapasztalati relatív gyakorisága kicsi, sőt tetszőleges ϵ számnál kisebb lehet közel 1 valószínűséggel, azaz a nagy eltérés esélye kicsi.

3.2.3. Tétel (Nagy számok törvénye - Bernoulli-féle alak). *Legyen A egy tetszőleges esemény, melyre $P(A) = p$. Végezzünk N darab független kísérletet és jelölje ezek között az A esemény bekövetkezésének számát ξ_N . Ekkor a relatív gyakoriság: $\frac{\xi_N}{N}$ -nel egyenlő. Tetszőleges $\epsilon > 0$ és $\delta > 0$ esetén $\exists N_0$, hogy $\forall N > N_0$ esetén:*

$$\lim_{N \rightarrow \infty} P \left(\left| \frac{\xi_N}{N} - p \right| \geq \epsilon \right) \leq \delta. \quad (3.18)$$

Ebből az alakból adódik, hogy a $\tilde{\Theta}_N = c * \frac{\nu}{N}$ változóra teljesül, hogy:

$$\lim_{N \rightarrow \infty} P \left(\left| \tilde{\Theta}_N - I(f) \right| \geq \epsilon \right) = 0 \quad (\forall \epsilon > 0). \quad (3.19)$$

Tegyük fel, hogy a $Z = f(X)$ valószínűségi változó σ szórása létezik. Ekkor fennáll, hogy:

$$\sigma^2 = \sigma^2(Z) = \int_G f^2(P) * p(P) dP - (I(f))^2. \quad (3.20)$$

3.2.4. Tétel. *Ha elég kísérletet elvégzünk, akkor $\tilde{\Theta}_N$ közelítőleg normális eloszlású, $I(f)$ várható értékkel és $\frac{\sigma}{\sqrt{N}}$ szórással.*

Felhasználva (3.2), (3.3) és (3.4) már korábban levezetett formulákat, megkapjuk a következő egyenlőtlenséget:

$$\left| \tilde{\Theta}_N - I(f) \right| < \chi_\beta * \frac{\sigma}{\sqrt{N}}. \quad (3.21)$$

(3.21) valószínűsége β .

3.2.5. Megjegyzés. *Tegyük fel, hogy rögzítettük β megbízhatósági szintet, pl. 95%-ra. Ekkor (3.21)-et a következőféleképp értelmezhetjük:*

$I(f)$ β valószínűséggel a $\left[\tilde{\Theta}_N - \chi_\beta * \frac{\sigma}{\sqrt{N}}, \tilde{\Theta}_N + \chi_\beta * \frac{\sigma}{\sqrt{N}} \right]$ intervallumban helyezkedik el.

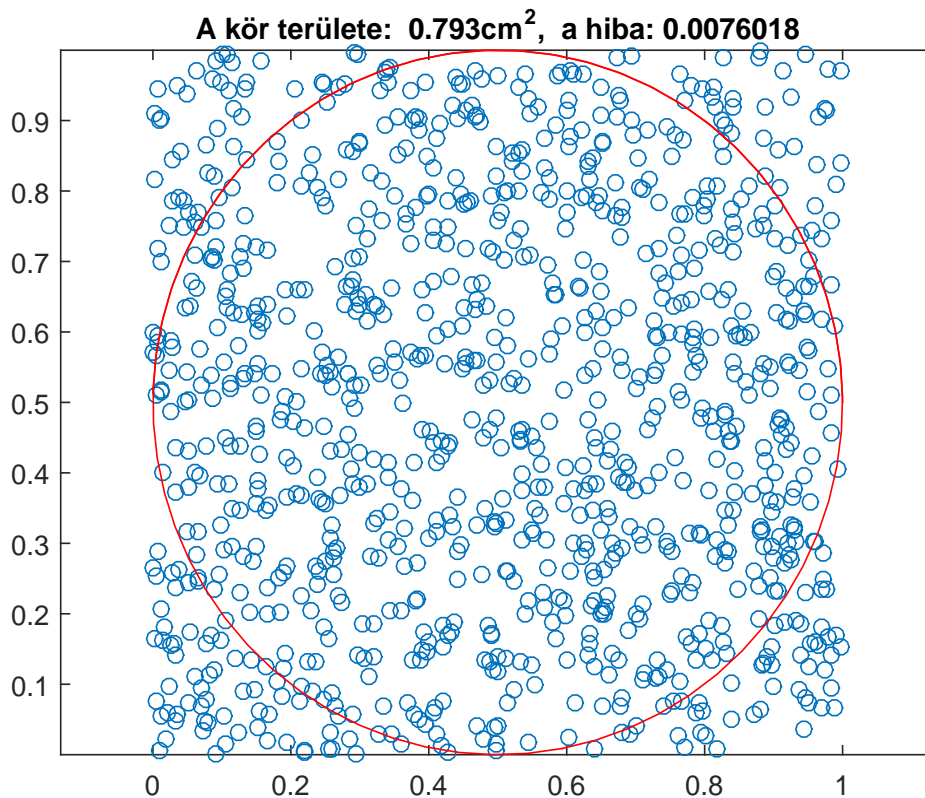
Ennek az intervallumnak a hossza σ -val arányos. Ha rögzítjük az intervallum hosszát, akkor $x_\beta * \frac{\sigma}{\sqrt{N}} = \epsilon$ és $N = x_\beta^2 * \frac{\sigma^2}{\epsilon^2}$ miatt N pedig σ^2 -tel arányos.

3.2.6. Megjegyzés. Ebből látható, hogy a becslésünk hatékonysága annál jobb, minél kisebb a szórás. Ezért kell olyan valószínűségi változókkal dolgozni, amiknek kicsi a szórása.

3.3. Példák Monte Carlo integrálásra

Ebben a részben néhány alkalmazást fogunk megnézni a Monte Carlo integrálásra. A határozott integrálok kiszámításához a Monte Carlo integrálás Matlab implementációját alkalmaztuk (Hit and Miss módszer). A becsült integrálok mellett fel lesznek tüntetve a pontos értékektől való eltérések.

A kör területének kiszámítása



3.1. ábra. Monte Carlo szimuláció a kör területének kiszámítására

A legelső számítás a legegyszerűbb példa a Monte Carlo integrálás alkalmazására. Az a feladatunk, hogy a $(0,5;0,5)$ középpontú $0,5$ cm sugarú kör területét kiszámítsuk a Monte Carlo módszerrel.

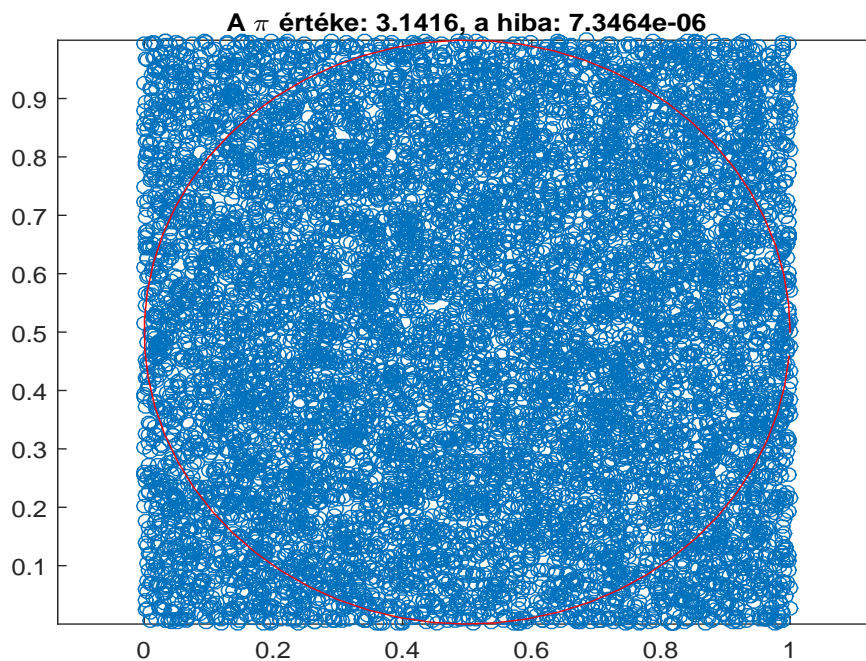
A kör területét a következőképp kapjuk:

$$T = r^2 * \pi \approx 0,7854\text{cm}^2. \quad (3.22)$$

Látható, hogy a Hit and Miss módszerrel kapott közelítés már 500 pontnál jó becslést ad. Itt a hiba: $h < 10^2$, ahogy 3.1 ábrán is látszik.

A pi értékének kiszámítása

A Monte Carlo integrálással tudjuk közelíteni pl. a π értékét. Ha véletlenszerűen egy egységnégyzetbe és egy $0,5$ cm sugarú, $(0,5;0,5)$ középpontú körbe n darab pontot szórunk, akkor a körbe eső és a körön kívül eső pontok aránya éppen a kör területe lesz. Ha ezt leosztjuk a sugár négyzetével (jelen esetben $0,5^2$ -tel), akkor a π közelítését fogjuk kapni. 10000 pont beszórásával már viszonylag jó becslést kapunk a π értékére, ahogy a 3.2 ábra is mutatja.



3.2. ábra. Monte Carlo szimuláció π értékének kiszámítására

A gömb térfogatának kiszámítása

Szeretnénk a gömb térfogatát kiszámítani analitikusan és a Monte Carlo integrálással is. Tudjuk, hogy a gömb térfogata felírható az alábbi képlettel, ahol r a kör sugara:

$$V_{\text{gömb}} = \frac{4 * r^3 * \pi}{3}. \quad (3.23)$$

Analitikusan

Forgástestek térfogatát legtöbbször integrálással vagy Cavalieri-elv alkalmazásával tudjuk kiszámítani. [9] és [10] jegyzetek alapján át fogjuk nézni a kapcsolódó tételeket. Most tekintsük az integrálással kiszámított térfogatot. Speciális esetben, ha az $[a, b]$ intervallumon folytonos, nemnegatív f függvény grafikonját x tengely körül forgatjuk, akkor a kapott test térfogata az alábbi integrállal számítható ki:

$$V = \pi * \int_a^b f^2(x) dx. \quad (3.24)$$

Tekintsük az $f : [-r, r] \rightarrow \mathbb{R}$, $f(x) = \sqrt{r^2 - x^2}$ nemnegatív, folytonos függvényt. Ez egy origó középpontú, r sugarú félkör hozzárendelési utasítása. Ha megforgatjuk az x tengely körül, akkor egy origó középpontú, r sugarú gömböt fogunk kapni. Számoljuk ki a térfogatát:

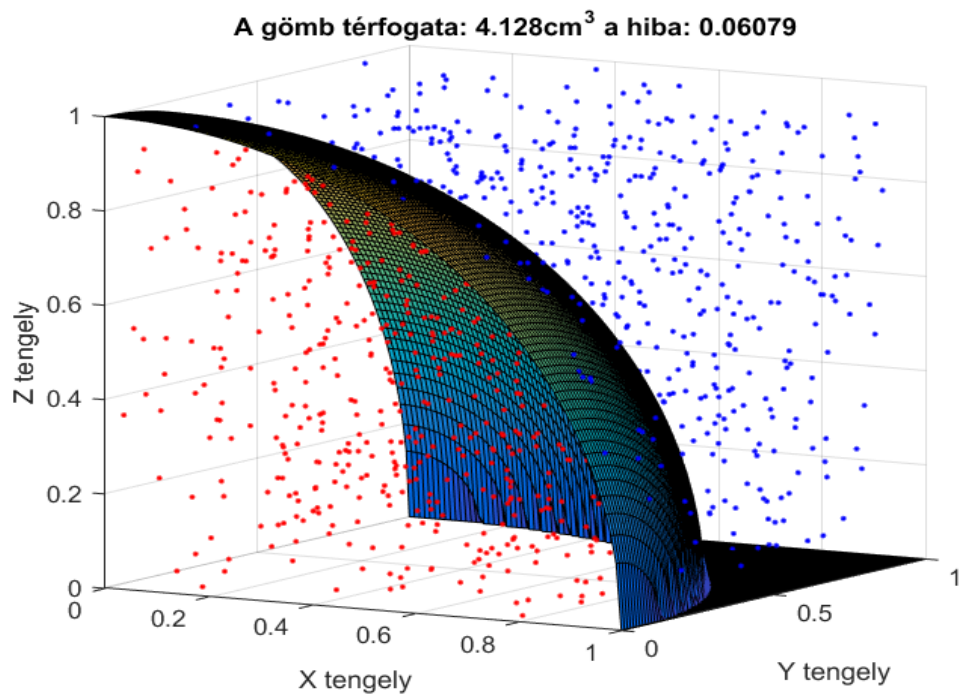
$$V = \pi * \int_{-r}^r f^2(x) dx = \pi * \int_{-r}^r (r^2 - x^2) dx = \pi * \left[r^2 * x - \frac{x^3}{3} \right]_{-r}^r = \frac{4 * r^3 * \pi}{3}. \quad (3.25)$$

Monte Carlo szimulációval

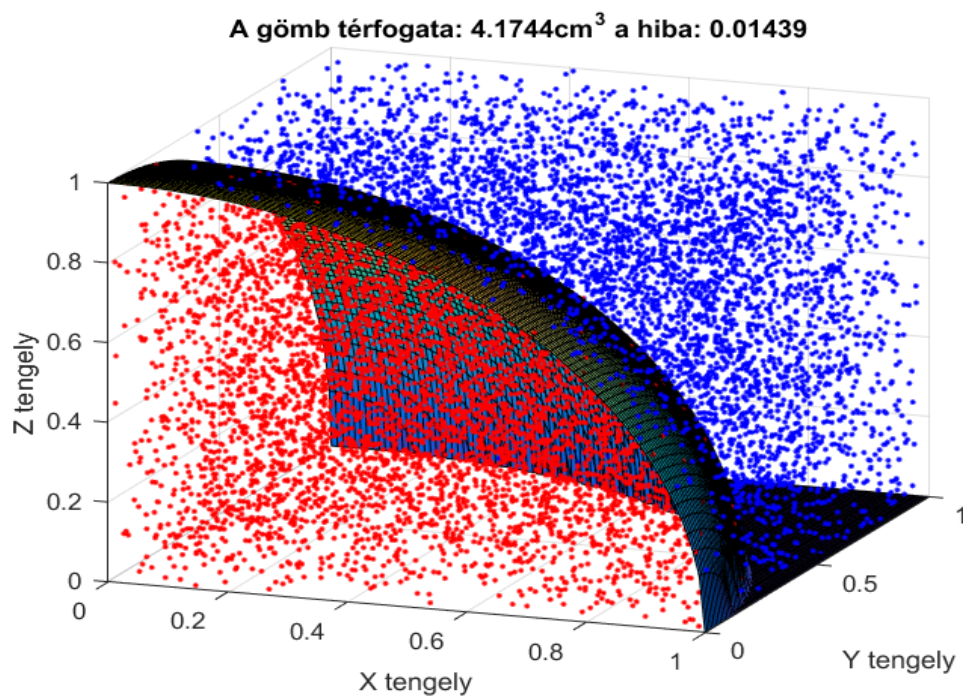
Egy egység sugarú negyedgömbön szimuláltuk a Monte Carlo integrálást, 1000 és 10000 pont beszórásával. 3.3 és 3.4 ábrán megtalálható a számolás hibája és a számolt térfogat is. Az egységgömb térfogata:

$$V_{\text{gömb}} = \frac{4 * \pi}{3} \approx 4,189 \text{cm}^3. \quad (3.26)$$

3.3.1. Megjegyzés. *Látszik, hogy több pont beszórásával csökken a hiba, ami az előző fejezet tételeiből következik.*



3.3. ábra. Monte Carlo szimuláció a gömb térfogatának kiszámítására



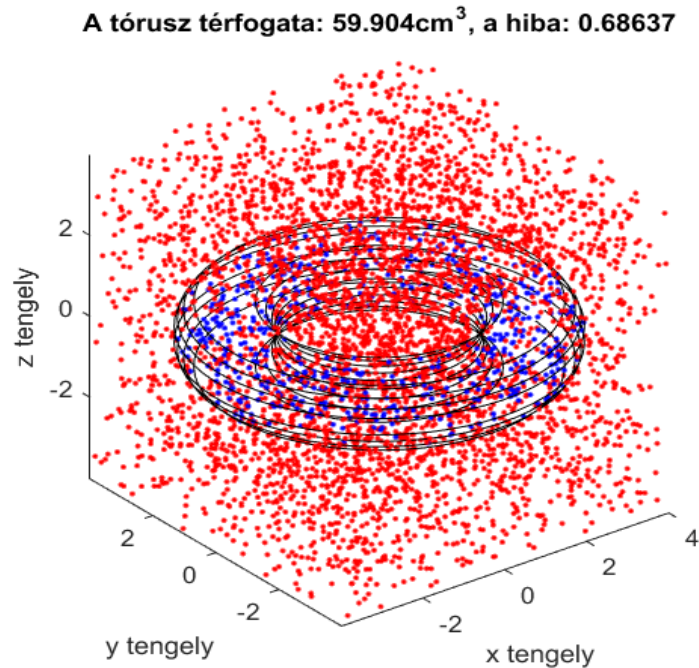
3.4. ábra. Monte Carlo szimuláció a gömb térfogatának kiszámítására

A tórusz térfogatának kiszámítása

A tórusz térfogatát a gömbhöz hasonlóan lehet kiszámítani: itt az y tengely körül forgatunk meg egy kört, aminek középpontja az z tengelytől R távolságra van és r sugarú. Ekkor a tórusz térfogata a következő:

$$V_{\text{tórusz}} = 2 * \pi^2 * R * r^2 \quad (3.27)$$

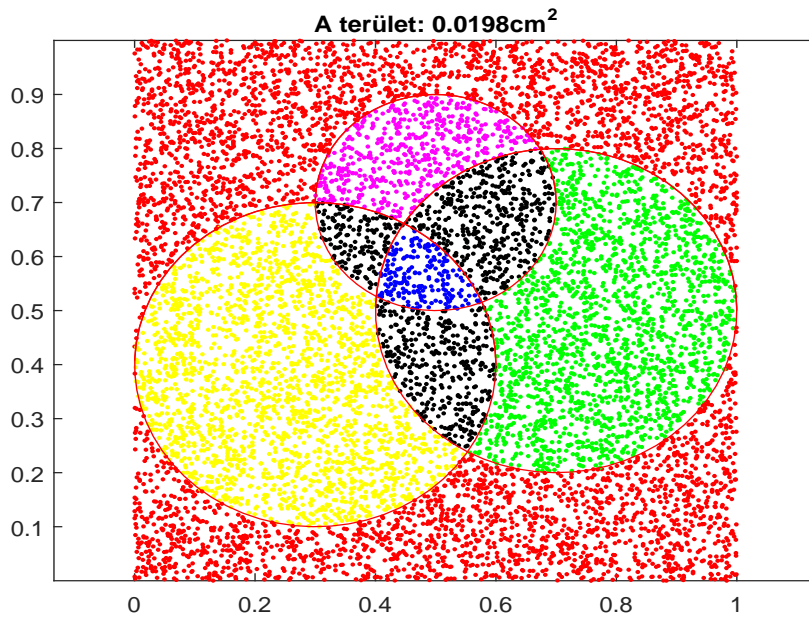
A Monte Carlo szimulációt futtatva $r = 1$, $R = 3$ paraméterekre, 5000 pont beszórásával a 3.5 ábrán található közelítést kaptuk:



3.5. ábra. Monte Carlo szimuláció a tórusz térfogatának kiszámítására

Alakzatok metszetének kiszámítása

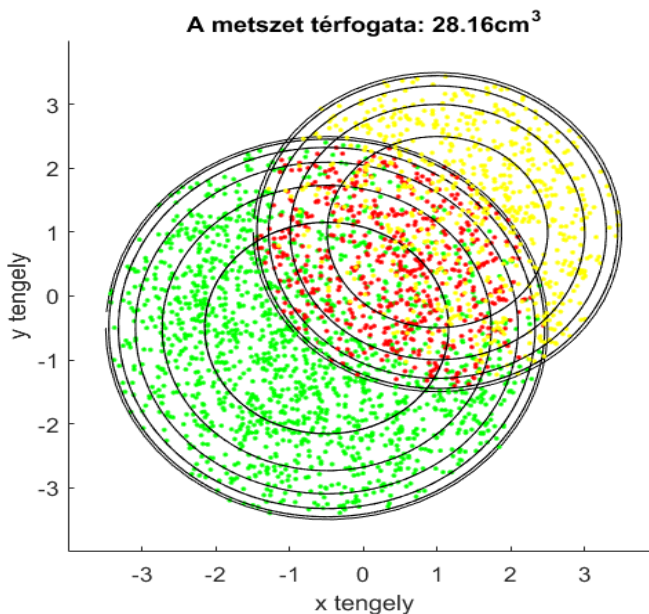
Feladat: Szeretnénk a $K_1 (0, 7; 0, 5)$ középpontú, $0, 3$ cm sugarú, a $K_2 (0, 5; 0, 7)$ középpontú, $0, 2$ cm sugarú és a $K_3 (0, 3; 0, 4)$ középpontú, $0, 3$ cm sugarú körök metszetével keletkező területet. Ehhez használjuk a Monte Carlo integrálást, 10000 pont beszórásával.



3.6. ábra. Monte Carlo szimuláció a körök metszetének kiszámítására

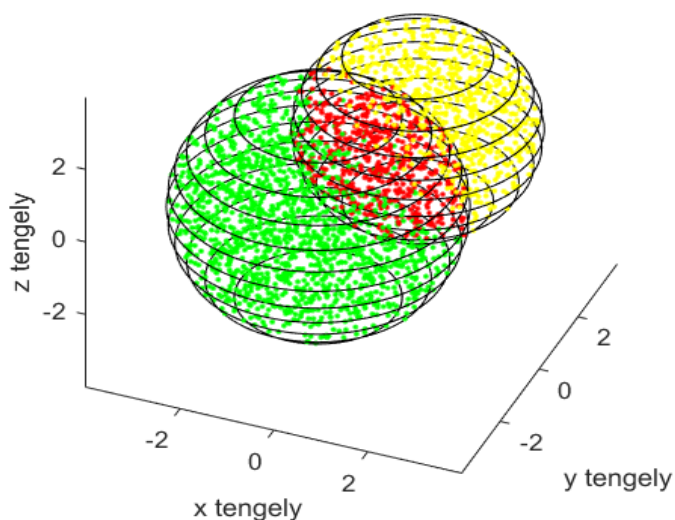
Egymást metsző gömbök térfogatának kiszámítása

A MC integrálás hatékonyan alkalmazható 3 dimenziós testek metszetének kiszámítására. Ki fogjuk számítani a $(1; 1)$ középpontú, $2,5$ cm sugarú és a $(-0, 5; -0, 5)$ középpontú, 3 cm sugarú gömbök metszetét, ahogy 3.7 és 3.8 ábrán látható.



3.7. ábra. Monte Carlo szimuláció gömbök metszetének kiszámítására (felülnézet)

A metszet térfogata: 28.4672cm³



3.8. ábra. Monte Carlo szimuláció a gömbök metszetének kiszámítására (oldalnézet)

3.4. A Monte Carlo integrálás hibája

Ebben a részben össze fogjuk hasonlítani a Monte Carlo integrálás és a kvadratúra formulák hibáját, Meg fogjuk nézni, hogy mikor és miért jobb a Monte Carlo integrálás a kvadratúra formuláknál. Ehhez fel fogjuk használni a [11] jegyzet eredményeit.

2.1-ben már definiáltuk 1 dimenzióban az integrációs kvadratúra formulákat. Most definiáljuk M dimenzióban is, az $([a_1, b_1] \times [a_2, b_2] \times \dots \times [a_M, b_M])$ M dimenziós téglán, ahol $x_i \in [a_i, b_i]$. Ekkor kapunk egy $M + 1$ dimenziós térfogatot, amit a következőképp írhatunk fel:

$$V^{(M+1)} = \frac{V^{(M)}}{N} \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} \dots \sum_{j_M=0}^{N_M} a_{j_1} * \dots * a_{j_M} * f(x_{j_1}^{(N_1)}, \dots, x_{j_M}^{(N_M)}) \quad (3.28)$$

Definiáljuk M dimenzióban a Monte Carlo integrált is:

$$V^{(M+1)} \approx \frac{V^{(M)}}{N} \sum_{i=0}^N f(x_i) \quad (3.29)$$

3.4.1. Következmény. Innen látszik a különbség: a kvadratúra formulák kiszámításához M darab összegre van szükség, amíg a MC integráláshoz csupán 1 is elegendő.

- 1 – 2 dimenzióban még a kvadratúra formulákkal pontosabban és hatékonyabban tudjuk számolni, mivel csak az alappontokon kell kiértékelni a formulát, míg a Monte Carlo integrálás során akár 10000 pontot is be kell szórniuk hasonló pontosság eléréséhez.
- Ahogy a dimenziószámot (M) növeljük, a kvadratúra formulákhoz M összeget kell kiszámítanunk, ami egyre bonyolultabb lesz. Viszont a Monte Carlo integráláshoz továbbra is csak 1 összeget kell számolni.
- Nézzünk egy példát. Legyen $M = 10$ és válasszunk minden intervallumból $N_i = 5$ alappontot. Így ha ki szeretnénk számítani a kvadratúra formulát szükség van $N_i^M = 5^{10} \approx 10$ millió pontra. Mivel az alappontok száma nagyon kevés, így a kiértékelés sem lesz túl pontos. Viszont a Monte Carlo integráláshoz elég összesen N pontot beszórni.

3.4.2. Példa. Nézzünk meg egy tesztet, ami a 3.4.1 következményt támasztja alá. Egy szimulációt futtattunk, ami egy M dimenziós gömb térfogatát közelíti numerikus integrálással (érintőformulával) és MC módszerrel. Ennek eredménye a 3.1 ábrán látható.

- A numerikus integráláshoz minden dimenzióban 20 alappontot vettünk.
- A Monte Carlo integrálást végig 10^5 db ponttal végeztük.

3.1. táblázat. Érintőformula és Monte Carlo integrálás hibájának összehasonlítása¹

Analitikus		Numerikus			Monte Carlo		
Dimenzió	Pontos érték	Idő	Eredmény	Hiba	Idő	Eredmény	Hiba
2	3,1415	0,00	3,1524	$1,09 * 10^{-2}$	0,01	3,1435	$2,00 * 10^{-3}$
3	4,1887	0,00	4,1737	$1,50 * 10^{-2}$	0,07	4,1896	$9,00 * 10^{-4}$
4	4,9348	0,00	4,9023	$3,25 * 10^{-2}$	0,08	4,9330	$1,80 * 10^{-3}$
5	5,2637	0,02	5,2381	$2,56 * 10^{-2}$	0,10	5,2787	$1,50 * 10^{-2}$
6	5,1677	0,30	5,1451	$2,26 * 10^{-2}$	0,13	5,1748	$7,10 * 10^{-3}$
7	4,7247	5,02	4,6704	$5,43 * 10^{-2}$	0,15	4,7098	$1,49 * 10^{-2}$
8	4,0587	89,9	3,9595	$9,92 * 10^{-2}$	0,17	4,0479	$1,08 * 10^{-2}$
9	3,2985	1320	3,3998	$1,01 * 10^{-1}$	0,20	3,3191	$2,06 * 10^{-2}$

3.4.3. Megjegyzés. Látható, hogy 6 dimenzió alatt az érintőformula gyorsabb és pontosabb, viszont 6 dimenzió felett már a Monte Carlo integrálás válik hatékonyabbá. Ez mutatja a Monte Carlo integrálás gyakorlati hasznát.

¹A szimuláció és a 3.1 ábra a [11] cikk a 12. oldalán szerepel.

3.4.4. Tétel. *A Monte Carlo integrálás hibája egyenesen arányos a szórással, ami pedig fordítottan arányos a felvett véletlen pontok gyökének számával.*

$$|I - I_{MC}| = V^2 * \frac{\sigma_N}{\sqrt{N}} \quad (3.30)$$

A hibaképlet levezetésével nem foglalkozunk, a témakör részletes kifejtése a [13] és [14] jegyzetekben megtalálható.

A Monte Carlo integrál nem determinisztikus, mivel véletlen számokat használunk a becslésre. (3.30) alapján láttuk, hogy a hiba a szórásnégyzettől függ, ami pedig a véletlen számok darabszámának növelésével csökkenthető. Ez viszont több számítást igényel. Belátható, hogy a konvergencia lassabb, mint determinisztikus esetekben (főleg a trapéz és a Simpson módszerhez képest), viszont magasabb dimenzióban is megmarad a konvergencia sebessége, míg a determinisztikus módszerek egyre idő- és számításigényesebbé válnak. Tehát a Monte Carlo integrálásnál fontos, hogy olyan módszereket alkalmazzunk, melyek csökkentik a szórást, viszont a számítási időt nem, vagy nem jelentősen növelik. A következő fejezetben ezekről a szóráscsökkentő eljárásokról lesz szó bővebben.

4. fejezet

Szóráscsökkentő eljárások

Az előző fejezetben láttuk, hogy a becslés hatékonysága a szórás csökkentésével vagy a pontok számának emelésével nő. Ebben a fejezetben szórást csökkentő eljárásokkal fogunk megismerkedni. Ezeket az eljárásokat felhasználva a Monte Carlo integrál alkalmazásánál a becslés pontosabb lesz. A fejezet részletes kifejtéséhez fel fogjuk használni a [5] és [6] jegyzetek eljárásait.

4.1. A főrészt leválasztása

4.1.1. Tétel. *Nézzük ismét az alábbi integrált:*

$$I(f) = \int_G f(P) * p(P) dP. \quad (4.1)$$

Ha f függvényt egy olyan h függvénnyel közelítjük, amire $I(h)$ integrált könnyen ki tudjuk számolni, akkor a Monte Carlo módszert az $g = f - h$ függvényre alkalmazva a szórás csökkenthető.

Bizonyítás. Közelítsük f -et egy ilyen h függvénnyel. Ekkor a szórásnégyzet a következőképpen becsülhető:

$$\sigma^2(f - h) = \sigma^2(f) + \sigma^2(h) - 2 * \text{Cov}(f, h) < \sigma^2(f). \quad (4.2)$$

Az utolsó egyenlőtlenség fennáll, ha a kovariancia elég nagy, azaz a h függvény hasonlít f -hez. Azaz megkaptuk, hogy a szórásnégyzet valóban csökken. \square

4.1.2. Példa. *Szeretnénk kiszámítani $\int_0^1 e^x dx$ határozott integrált. Legyen h a következő: $h(x) = 1 + x$, mivel $e^x \approx 1 + x$ a 0 egy kis környezetében, $\int_0^1 (1 + x) dx = 1,5$*

könnyen kiszámítható. Ekkor a főrészt leválasztva pontosabb becslést tudunk adni.

4.1. táblázat. MC szimuláció a főrészt leválasztásával¹

Pontok száma	Becsült integrál	Szórás
10	1,6450	$5,01 * 10^{-2}$
100	1,7190	$2,29 * 10^{-2}$
1000	1,7250	$6,89 * 10^{-3}$
10000	1,7213	$2,10 * 10^{-3}$
100000	1,7198	$6,60 * 10^{-4}$
1000000	1,7184	$2,00 * 10^{-4}$

4.1.3. Megjegyzés. $I(h)$ integrált könnyen meg tudjuk határozni pl. abban az esetben, ha h egy analitikus függvény. A 2.2 fejezetben bemutatott interpolációs kvadratúra formulák használhatók erre célra. Ebben az esetben h egy polinom, melynek integrálja könnyen kiszámolható.

4.2. Az integrációs tartomány részekre bontása

Ebben az esetben nem egy függvényt fogunk keresni, aminek könnyen ki tudjuk számítani az integrálját, hanem G -nek egy olyan részhalmazára szűkítjük az integrálást, melyen már meg tudjuk határozni a (4.1)-ben felírt integrált.

4.2.1. Tétel. *Tegyük fel, hogy egy $B \subseteq G$ tartományon meg tudjuk határozni az alábbi integrálokat.*

$$\int_B f(P) * p(P) dP = A, \quad \int_B p(P) dP = a. \quad (4.3)$$

A Monte Carlo módszert $G \setminus B$ -re alkalmazva a szórásnégyzet csökkenthető.

Bizonyítás. Tegyük fel, hogy (4.3) képlettel ki tudjuk számítani az integrálokat B -n. Legyen $D = G \setminus B$. Ekkor $\int_D f(P) * p(P)$ integrált kell meghatároznunk. Erre alkalmazzuk a Monte Carlo módszert.

Nézzük az alábbi függvényt:

$$p_1(P) = \begin{cases} \frac{p(P)}{1-a} & \text{ha } P \in D \\ 0 & \text{ha } P \notin D. \end{cases}$$

Bebizonyítható, hogy $p_1(P)$ sűrűségfüggvény.

¹A szimuláció és a 4.1 ábra a [12] könyvben szerepel.

$$\int_{\mathbb{D}} f(\mathbf{P}) * \mathbf{p}(\mathbf{P}) d\mathbf{P} = (1 - \alpha) * \int_{\mathbb{D}} f(\mathbf{P}) * \mathbf{p}_1(\mathbf{P}) d\mathbf{P} \quad (4.4)$$

A fenti (4.4) egyenlőség jobb oldalán álló integrál meghatározásához nézzük a $\mathbf{p}_1(\mathbf{P})$ sűrűségfüggvényű \mathbf{X} valószínűségi változót. Az integrál kiszámításához vegyük ennek N db független realizációját, azaz X_i ($i = 1, \dots, N$) mintát. Legyen $Y = f(\mathbf{X})$ és $Y_i = f(\mathbf{X}_i)$. Ekkor:

$$\sigma^2(Y) = \int_{\mathbb{D}} f^2(\mathbf{P}) * \mathbf{p}_1(\mathbf{P}) d\mathbf{P} - \left(\int_{\mathbb{D}} f(\mathbf{P}) * \mathbf{p}_1(\mathbf{P}) d\mathbf{P} \right)^2 < \frac{1}{1 - \alpha} \int_{\mathbb{D}} f^2(\mathbf{P}) * \mathbf{p}(\mathbf{P}) d\mathbf{P}. \quad (4.5)$$

Ha megnézzük a (4.5) egyenletet, akkor látni fogjuk, hogy a kapott szórásnégyzet kisebb lett mint az eredeti szórásnégyzet. \square

4.3. Dimenziócsökkentés

Vegyünk egy olyan 3 dimenziós esetet, ahol az integrálnak az egyik változójára meg tudunk adni primitív függvényt. Ekkor ennek kiszámításával, azaz ha a pontos értéket adjuk erre az integrálra, a hibát ez is csökkenteni tudjuk.

Példa

Legyen $f(x, y, z) = z * e^{(x+y)^2} : \mathbb{R}^3 \rightarrow \mathbb{R}$ függvény. Ekkor a határozott integrál a $[0, 1]$ intervallumon a következőképp írható fel:

$$\begin{aligned} \int_0^1 \int_0^1 \int_0^1 f(x, y, z) dx dy dz &= \int_0^1 \int_0^1 \int_0^1 z * e^{(x+y)^2} dx dy dz = \\ \int_0^1 \int_0^1 \left(\int_0^1 z * e^{(x+y)^2} dz \right) dx dy &= \frac{1}{2} * \int_0^1 \int_0^1 e^{(x+y)^2} dx dy. \end{aligned} \quad (4.6)$$

Látszik, hogy az 3. változó szerinti integrál kiszámításával csökkentettük a dimenziószámot, amivel nyilvánvalóan könnyebbé vált a háromváltozós integrál kiszámítása.

4.4. A sűrűségfüggvény optimális megválasztása

Ebben a részben azt fogjuk megnézni, ha a sűrűségfüggvényünket megfelelően választjuk meg, akkor az csökkenti a szórást.

4.4.1. Tétel. *Ha a sűrűségfüggvényt az alábbi függvénynek választjuk, akkor csökken a Monte Carlo integrálás hibája.*

$$p^*(P) = \frac{|f(P)|}{\int_G |f(P)| dP} \quad (4.7)$$

Bizonyítás. Szeretnénk az alábbi integrált meghatározni:

$$I = \int_G f(P) dP. \quad (4.8)$$

Tegyük fel, hogy $\int_G |f^2(P)| dP < \infty$. Legyen a G_0 azoknak a $P \in G$ pontoknak a halmaza, amelyekre $f(P) = 0$ fennáll és legyen $G_1 = G \setminus G_0$. Olyan p sűrűségfüggvényeket fogunk nézni, amelyekre $p(P) > 0$ ($P \in G_1$) teljesül.

Legyen

$$g(P) = \begin{cases} \frac{f(P)}{p(P)} & \text{ha } P \in G_1, \\ 0 & \text{ha } P \in G_0. \end{cases}$$

Ekkor a (4.8)-ben szereplő integrálra: $I = I(g)$.

Most írjuk fel a szórást:

$$\sigma_p^2 = \int_G g^2(P) * p(P) dP - I^2 = \int_G \frac{f^2(P)}{p(P)} dP - I^2. \quad (4.9)$$

Olyan sűrűségfüggvényt keresünk, amire a szórás minimális:

Legyen:

$$p^*(P) = \frac{|f(P)|}{\int_G |f(P)| dP} \quad (4.10)$$

Nézzük meg ennek a szórását:

$$\sigma_{p^*}^2 = \left(\int_G |f(P)| dP \right)^2 - I^2. \quad (4.11)$$

Meg fogjuk mutatni, hogy erre a sűrűségfüggvényre a legkisebb a szórás. Ehhez írjuk

fel a Cauchy-Bunyakovszkij-Schwarz egyenlőtlenséget a bal oldalra, azaz:

$$\begin{aligned} \left(\int_{\mathbf{G}} |f(\mathbf{P})| d\mathbf{P} \right)^2 &\leq \left(\int_{\mathbf{G}_1} |f(\mathbf{P})| d\mathbf{P} \right)^2 = \left(\int_{\mathbf{G}_1} |f(\mathbf{P})| * p(\mathbf{P})^{-\frac{1}{2}} * p(\mathbf{P})^{\frac{1}{2}} d\mathbf{P} \right)^2 \\ &\leq \left(\int_{\mathbf{G}_1} \frac{f(\mathbf{P})^2}{p(\mathbf{P})} d\mathbf{P} \right) * \left(\int_{\mathbf{G}_1} p(\mathbf{P}) d\mathbf{P} \right) \leq \int_{\mathbf{G}_1} \frac{f(\mathbf{P})^2}{p(\mathbf{P})} d\mathbf{P}. \end{aligned} \quad (4.12)$$

Ha f nem vált előjelet \mathbf{G} -n, akkor $\sigma_{p^*} = 0$. □

Ha a sűrűségfüggvény választását jobban szemügyre vesszük, akkor feltűnhet, hogy ennek a kiszámításához ismernünk kellene $\int_{\mathbf{G}} |f(\mathbf{P})| d\mathbf{P}$ integrált. Így valójában nem lesz egyszerűbb a feladat, viszont azt megkaptuk, hogy érdemes a sűrűségfüggvényt $|f(\mathbf{P})|$ -vel arányosnak választani. Ezt az eljárást lényeg szerinti mintavételnek nevezzük.

4.5. Az integrandus szimmetrikussá tétele

Ebben a részben a szimmetrizálás alkalmazásával javított Monte Carlo integrálást fogjuk vizsgálni. Itt is be tudjuk bizonyítani, hogy a szimmetrikussá tétel csökkenti a szórást.

4.5.1. Tétel. *Ha szimmetrikussá tesszük az integrandust, akkor a szórásnégyzet csökkenni fog.*

Bizonyítás. Szeretnénk kiszámítani az alábbi integrált:

$$I = \int_a^b f(x) dx, \quad -\infty < a \leq b < \infty. \quad (4.13)$$

Legyen X egyenletes eloszlású $[a, b]$ intervallumon. Továbbá legyen $Y = (b - a) * f(X)$, $Y_i = (b - a) * f(X_i)$, ahol X_1, X_2, \dots, X_N az X független realizációi. Az integrált a Y_1, Y_2, \dots, Y_N minta átlagával fogjuk becsülni:

$$\Theta_N = \frac{1}{N} \sum_{i=1}^N Y_i = \frac{b - a}{N} \sum_{i=1}^N f(X_i). \quad (4.14)$$

Nézzük az alábbi függvényeket:

$$f^{(1)}(x) = \frac{1}{2} [f(x) + f(a + b - x)],$$

$$Y_i^{(1)} = (b - a) * f^{(1)}(X_i), \quad Y^{(1)} = (b - a) * f^{(1)}(X),$$

$$\Theta_N^{(1)} = \frac{b-a}{2 * N} \sum_{i=1}^N (f(X_i) - f(a+b-X_i)). \quad (4.15)$$

Írjuk fel az eredeti szórásnégyzetet és a fenti Y és $Y^{(1)}$ valószínűségi változók sűrűségfüggvényét:

$$\begin{aligned} \sigma_1^2 &= \sigma^2(Y^{(1)}), \\ \sigma^2 &= \sigma^2(Y). \end{aligned} \quad (4.16)$$

Ekkor a 2. momentumokat fel tudjuk írni a következőképpen:

$$\begin{aligned} E(Y^2) &= (b-a) * \int_a^b f^2(x) dx, \\ E((Y^{(1)})^2) &= \frac{b-a}{4} \int_a^b [f^2(x) + 2 * f(x) * f(a+b-x) + f^2(a+b-x)] dx = \\ &= \frac{b-a}{2} * \left[\int_a^b f^2(x) dx + \int_a^b f(x) * f(a+b-x) dx \right]. \end{aligned} \quad (4.17)$$

A Cauchy-Bunyakovszkij-Schwarz egyenlőtlenség felhasználásával a második integrálra az alábbi felső becslés adható:

$$\begin{aligned} \int_a^b f(x) * f(a+b-x) dx &= \sqrt{\left(\int_a^b f(x) * f(a+b-x) \right)^2} \\ &\leq \left(\int_a^b f^2(x) dx \right)^{\frac{1}{2}} * \left(\int_a^b f^2(a+b-x) dx \right)^{\frac{1}{2}} = \int_a^b f^2(x) dx. \end{aligned} \quad (4.18)$$

Ebből következik, hogy:

$$E((Y^{(1)})^2) \leq E(Y^2) \quad \Rightarrow \quad \sigma_1^2 \leq \sigma^2. \quad (4.19)$$

□

4.5.2. Megjegyzés. Ha $f(x)$ monoton és szakaszonként folytonos függvény az $[a, b]$ intervallumon, akkor élesebb becslés is adható σ_1^2 -re:

$$\sigma_1^2 \leq \frac{1}{2} * \sigma^2. \quad (4.20)$$

Bizonyítás. Írjuk fel $2 * \sigma_1^2$ -et az (4.16) és (4.17) egyenletek alapján:

$$2 * \sigma_1^2 = (b-a) * \int_a^b f^2(x) dx + (b-a) * \int_a^b f(x) * f(a+b-x) dx - 2 * I^2,$$

$$\sigma^2 = (b - a) * \int_a^b f^2(x) dx - I^2. \quad (4.21)$$

Be kell bizonyítanunk, hogy:

$$(b - a) * \int_a^b f(x) * f(a + b - x) dx \leq I^2. \quad (4.22)$$

Tegyük fel, hogy $f(x)$ monoton növekvő függvény, azaz $f(b) > f(a)$. Definiáljuk a $v(x)$ függvényt a következőképpen:

$$v(x) = (b - a) * \int_a^x f(a + b - t) dt - (x - a) * I. \quad (4.23)$$

Ekkor a $v(x)$ függvény az $[a, b]$ intervallum két végpontjában 0 értéket vesz fel. Ha deriváljuk a függvényt, akkor:

$$v'(x) = (b - a) * f(a + b - x) - I. \quad (4.24)$$

A $v'(x)$ monoton csökkenő függvény lesz. Ha behelyettesítjük a végpontokat, akkor azt kapjuk, hogy $v'(a) > 0$ és $v'(b) < 0$, ezért $v(x) \geq 0$ is fenn kell hogy álljon $\forall x \in [a, b]$. Mivel $f(x)$ monoton növekvő függvény, ezért $f'(x) \geq 0$. Ebből pedig következik, hogy:

$$\int_a^b v(x) * f'(x) dx \geq 0. \quad (4.25)$$

Ha ezt parciálisan integráljuk, az alábbi egyenlőtlenséget kapjuk:

$$\begin{aligned} \int_a^b v(x) * f'(x) dx &= [v(x) * f(x)]_a^b - \int_a^b f(x) * v'(x) dx = - \int_a^b f(x) * v'(x) dx \geq 0, \\ \int_a^b f(x) * v'(x) dx &\leq 0. \end{aligned} \quad (4.26)$$

Most helyettesítsük vissza a $v'(x) = (b - a) * f(a + b - x) - I$ értéket az egyenlőtlenségbe:

$$\begin{aligned} \int_a^b f(x) * ((b - a) * f(a + b - x) - I) dx &= (b - a) * \int_a^b f(x) * f(a + b - x) dx - I * \int_a^b f(x) dx = \\ &= (b - a) * \int_a^b f(x) * f(a + b - x) dx - I^2 \leq 0, \\ (b - a) * \int_a^b f(x) * f(a + b - x) dx &\leq I * \int_a^b f(x) dx. \end{aligned} \quad (4.27)$$

Azaz visszakaptuk (4.22) egyenlőtlenséget. \square

4.5.3. Megjegyzés. *Világos, hogy a szimmetrikussá tétel egyszerű és gyors hibacsökkentéssel alkalmazható egydimenziós integrálok esetén. A többdimenziós eset azonban már több és bonyolultabb számítást igényel. Nézzük azt példát, amikor $f(x, y, z)$ 3 dimenziós függvényt szeretnénk szimmetrizálni az $[0, 1] \times [0, 1] \times [0, 1]$ egységkockán. Ekkor az új függvényünket a következőképp írhatjuk fel:*

$$f^{(1)} = \frac{1}{8} * [f(x, y, z) + f(1 - x, y, z) + f(x, 1 - y, z) + f(x, y, 1 - z) + f(1 - x, 1 - y, z) + f(1 - x, y, 1 - z) + f(x, 1 - y, 1 - z) + f(1 - x, 1 - y, 1 - z)]. \quad (4.28)$$

Azaz itt már 8 taggal kell számolnunk minden lépés során.

4.5.4. Megjegyzés. *Gyakorlatban legtöbbször először a dimenziócsökkentést használják (ha lehetséges), utána pedig az integrációs tartomány részekre bontásával csökkentik a szórást. Belátható, hogy ezek alkalmazása akár 90%-kal csökkenti a szórást. Hátrányuk azonban, hogy sokszor nehezen vagy egyáltalán nem implementálhatóak a gyakorlatban.*

5. fejezet

Kitekintés

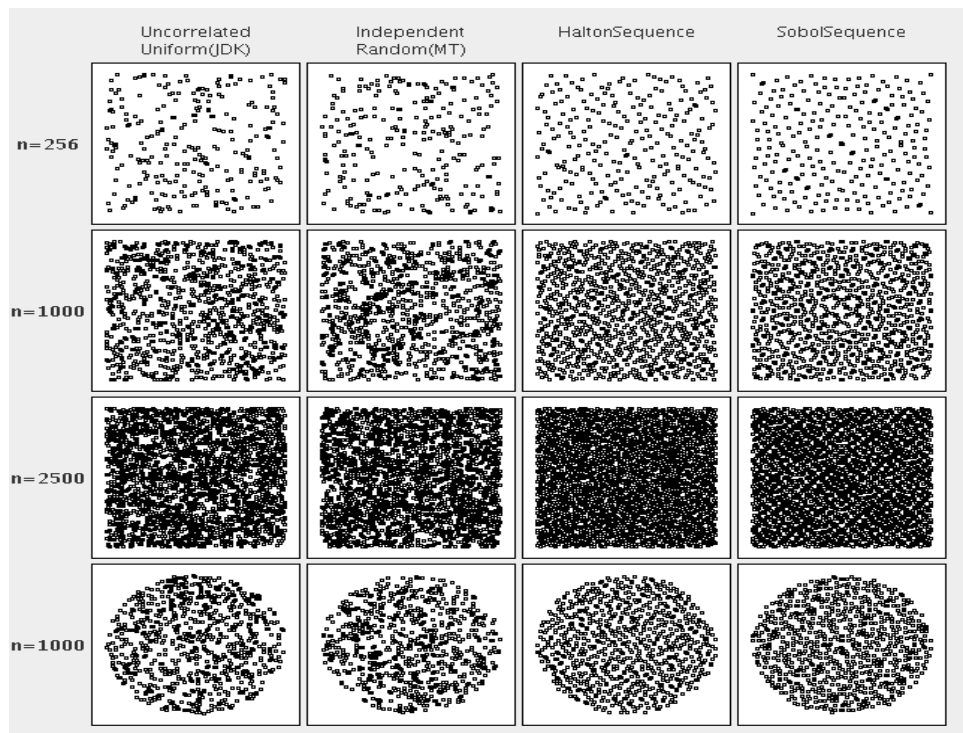
5.1. Véletlen szám generálási technikák

Az előző fejezetek során láttuk, hogy a Monte Carlo integrálás középpontjában a véletlen számok állnak. Ezért lényeges, hogy a szimulációkhoz rendelkezünk ilyen számokkal. Valójában ezeket nem olyan egyszerű előállítani számítógépek segítségével, bár meg kell említenünk, hogy szinte minden számítógépes program része egy véletlen szám táblázat (Matlabban `rand()`-ot használunk). Ezekkel a számítógépes véletlen szám generátorokkal fogunk foglalkozni ebben a részben.

A véletlen számok három kategóriába sorolhatóak:

- Igazi véletlen számok: ezeknek a számoknak az a lényege, hogy nem tudjuk megjósolni, mi lesz a szám. Pl. a lottósorsolásnál, amikor valaki kihúz egy számot a dobozból, akkor az $\{1, 2, \dots, 90\}$ halmazból kapunk egy véletlenszerűen kiválasztott számot.
- Pszeudovéletlen számok: ezek a számok egy algoritmussal előállított számok, amelyek számítógépes implementációját fel tudjuk használni pl. szimulációk során. Ezek a számok valójában nem tekinthetők véletlennek, ugyanis ha ismerjük az algoritmust, akkor vissza tudjuk adni az összes előállított számot.
- Kvázirandom számok. Ezek célja az N dimenziós tér egyenletes kitöltése (Halton, Hammersley, Sobol).

A pszeudovéletlen és kvázivéletlen számok közti különbséget a 5.1 ábrán szemléletesebben is láthatjuk.



5.1. ábra. Pszeudorandom és kvázirandom számok 2 dimenzióban¹

A Monte Carlo integrálás implementálásához használhatók pszeudovéletlen számok és kvázirandom számok is. Utóbbit kvázi Monte Carlo módszernek nevezik. A 3. fejezet Matlab implementációiban pszeudovéletlen számokat használtunk és a gyakorlati életben is ez az elterjedtebb, ugyanis az összes véletlen szám generátor a programokban ilyen számokat állít elő. Erre ad példát a dolgozat CD mellékletén található `lcg()` program. Ilyen pszeudovéletlen szám generátorok pl. a lineáris kongruenciagenerátorok, a Mersenne Twister és a Fibonacci generátorok.

A pszeudovéletlen számok közös jellemzői:

- Kezdeti érték: egy megadott kezdeti értékre a generátor ugyanazt a számsort fogja visszaadni.
- Periódus: a sorozat egy bizonyos idő után ismétlődni fog, ezt a számot nevezzük periódusnak. Ennek nagyobbnak kell lennie, mint amilyen hosszú számsorozatot akarunk generálni, ugyanis ha kisebb lenne, akkor biztosan lenne egy ismétlődő rész a sorozatban.

¹A szimuláció és a 5.1 ábra Common Math: The Apache Mathematics Library csomagban található.

- Intervallum: $(0, 1)$, $(0, 1]$, $[0, 1)$ vagy $[0, 1]$.

A periódust leggyakrabban egy 4-byte-os számként reprezentálják, $2^{32} \approx 4 \cdot 10^9$. Ezért ugyanennyi véletlen szám is előállítható.

Lineáris kongruenciagenerátorok

A lineáris kongruenciagenerátorok a legegyszerűbb generátorok egyike, amik segítségével előállíthatunk pszeudovéletlen $[0, 1)$ sorozatokat. Az alábbi implementáció [15] cikkben szerepel. Egy N hosszú sorozatot fogunk készíteni. Ehhez 4 paraméterre van szükségünk: $m > N$, A , a és b számokra. A generátor algoritmus a következő:

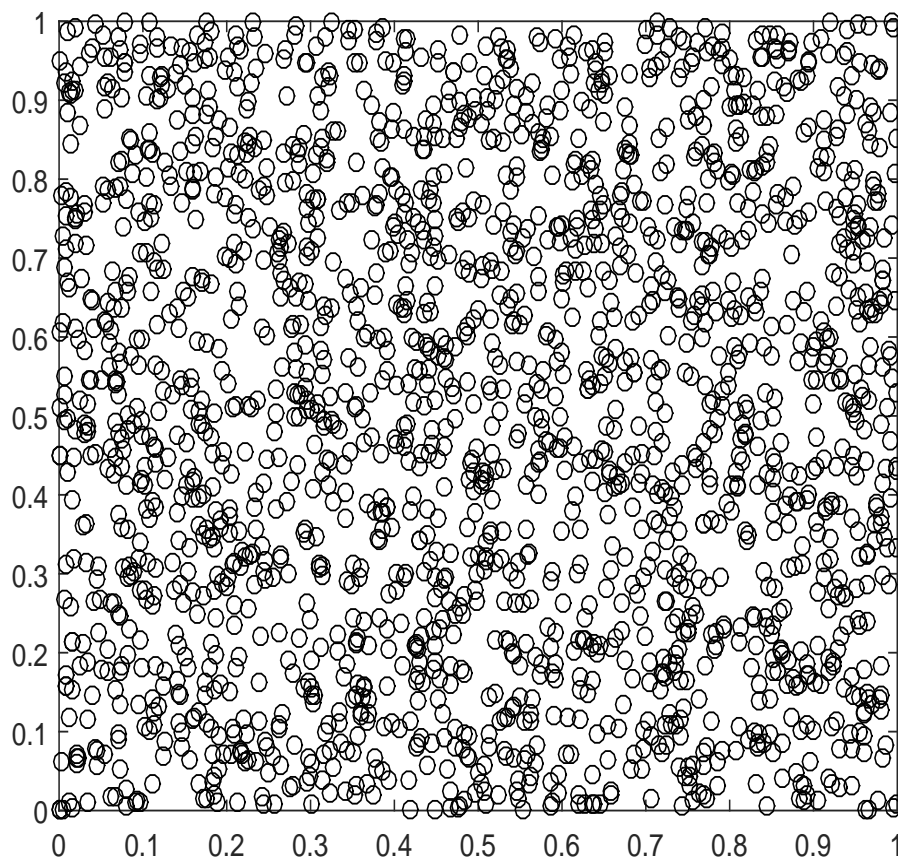
$$\begin{aligned} y_1 &\equiv A \pmod{m} & 0 \leq y_1 < m, \\ y_{n+1} &\equiv a * y_n + b \pmod{m} & 0 \leq y_{n+1} < m \\ n &= 1, 2, \dots, N - 1. \end{aligned} \tag{5.1}$$

Legyen továbbá $i = 1, 2, \dots, N$ esetén $x_i = \frac{y_i}{m}$, ez pedig bizonyítottan tekinthető pszeudovéletlen sorozatnak a $[0, 1)$ intervallumon.

Nézzük az alábbi paraméterekkel a lineáris kongruencia generátort:

- $A = 5$,
- $a = 2^{16} + 3$,
- $b = 0$,
- $m = 2^{31}$,
- $N = 2000$.

A $[0, 1] \times [0, 1]$ egységnégyzeten nézzük a következő konstrukciót: generáljuk le 1 dimenzióban a lineáris kongruencia generátorral x_i számokat, majd rajzoltassuk ki (x_i, x_{i-1}) pontpárokat. Ekkor a 5.2 ábrán láthatjuk a lineáris kongruencia generátor által generált számokat az egységnégyzeten. Láthatóan ezekkel a paraméterekkel a pontok eloszlása teljesen véletlenszerűnek tűnik.



5.2. ábra. A lineáris kongruencia generátor által generált véletlen pontok.

Mersenne Twister

A Mersenne Twister algoritmus a leggyakrabban használt véletlen szám generátor, ugyanis ezt használja jelenleg a legtöbb program beépített függvényként (pl. Matlab: `rand()`, `randi()`, `randn()`). A generátor periódusa $2^{19937} - 1$, ami kb. 6000 számjegyből áll, azaz ha másodpercenként egymilliárd számot generálunk, akkor 5985 évig tart, mire előlről kezdődik a számsor.

5.1.1. Megjegyzés. *Említettük az előző részben, hogy a teljes számsorozat visszaadható, ha kezdeti érték ugyanaz. Észrevehetjük, ha a Matlabban elindítunk két véletlen szám generátort, akkor ugyanazt az eredményt adja vissza, ha nem 'reseedelünk' (`seed=kezdeti érték`), ugyanis ilyenkor ugyanarról a kezdőértékről indítja a számítást (ez konkrétan a 0 szám). Ahhoz, hogy mindig független számsorozatokat kapjunk, a 'shuffer' parancsot használhatjuk, ami mindig egy új kezdőértéket állít be a pillanatnyi*

idő alapján. Bár ez véletlenszerűnek tűnik, nem célszerű mindig 'reseedelni' a generátort, ugyanis ez hatással lehet a véletlen számaink statisztikai tulajdonságaira. A 'default' beállítás annyiban előnyös, hogy a szimulációkat ugyanazokkal a véletlen számokkal újra tudjuk futtatni. Fontos megemlíteni, hogy a Matlabban lehet használni másik véletlen szám generátort, pl. a *Combined Multiplicative Recursive* generátort az `rng(0, 'combRecursive')` paranccsal, viszont ez lassabban tud számokat generálni.

5.1.2. Megjegyzés. A *Mersenne Twister* generátor onnan kapta a nevét, hogy a periódusideje egy *Mersenne-prím* ($2^{19937} - 1$).

Generátorok hasznossága

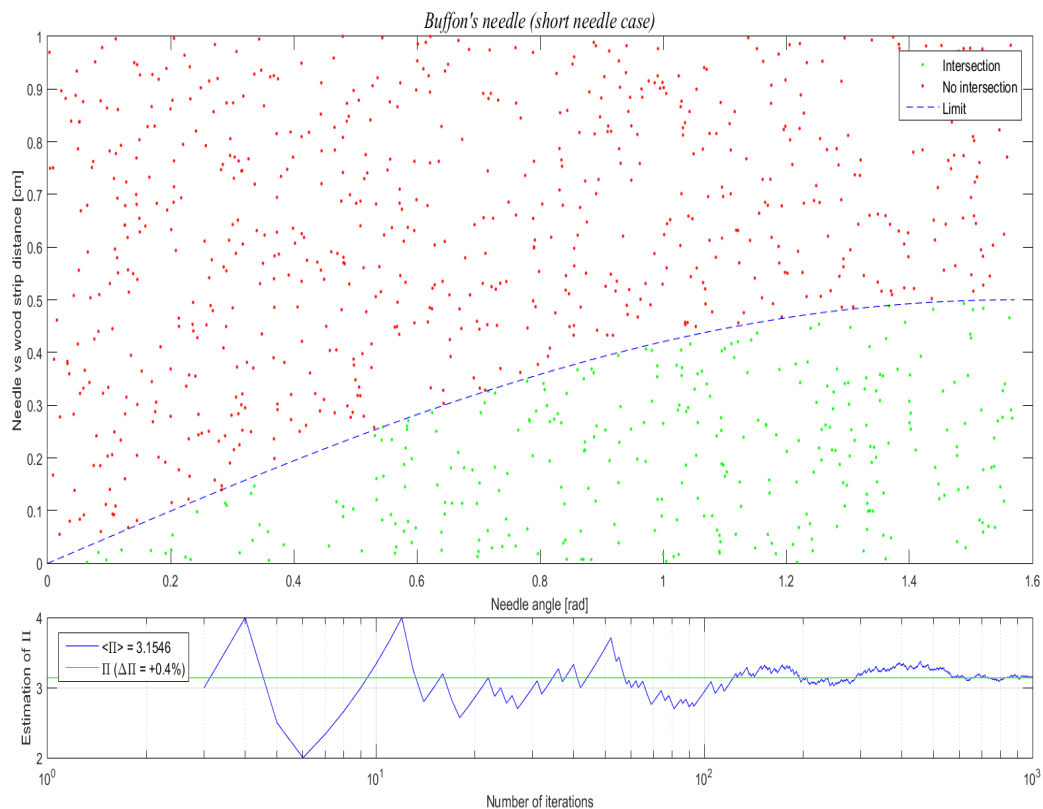
Azokat a generátorokat nevezhetjük jó generátoroknak, amik bizonyos statisztikai tesztekkel teljesítenek. Ezek közül a leggyakrabban használt tesztek egy csomagban vannak (TestU01). A csomag számos empirikus statisztikai tesztet tartalmaz, aminek részletes leírása a [16] cikkben szerepel. A *Mersenne Twister* és a lineáris kongruencia generátorok is ennek elvégzése során hatásosnak bizonyultak.

5.2. Egyéb alkalmazások

A Monte Carlo módszert egyre gyakrabban alkalmazzák a tudomány egyes területein (főleg modellezés terén, pl. a fizikában, a matematikában, a gazdasági életben, biológiában és kémiában is). Ezekből fogunk néhányat áttekinteni, főleg matematikai vonatkozásban.

Buffon-féle tűprobléma

A legismertebb probléma, amire alkalmazták a Monte Carlo módszert, a Buffon-féle tűprobléma. A történet szerint George L. Leclerc 1777-ben végzett egy kísérletsorozatot, hogy megnézze, mekkora annak a valószínűsége, hogy az asztallapra d távolságban felrajzolt vonalak egyikét metszeni fogja a feldobott ℓ hosszúságú tű, ahol $d > \ell$. Ezt végül megoldotta analitikusan és a kísérletsorozatot N -szer végrehajtotta, majd megsámolta, hogy az esemény n -szer következett be és arra jutott, hogy elég nagy N esetén $\frac{n}{N}$ jó közelítést ad a valószínűségekre. A kísérletet a π kísérleti meghatározására használták. Annak a valószínűsége, hogy a tű metszi a padló vonalát: $p = \frac{2}{\pi} * \frac{\ell}{d}$. Innen $\pi = \frac{2*\ell}{d*p}$. Az 5.3 ábrán a Matlab beépített szimulációja látható.

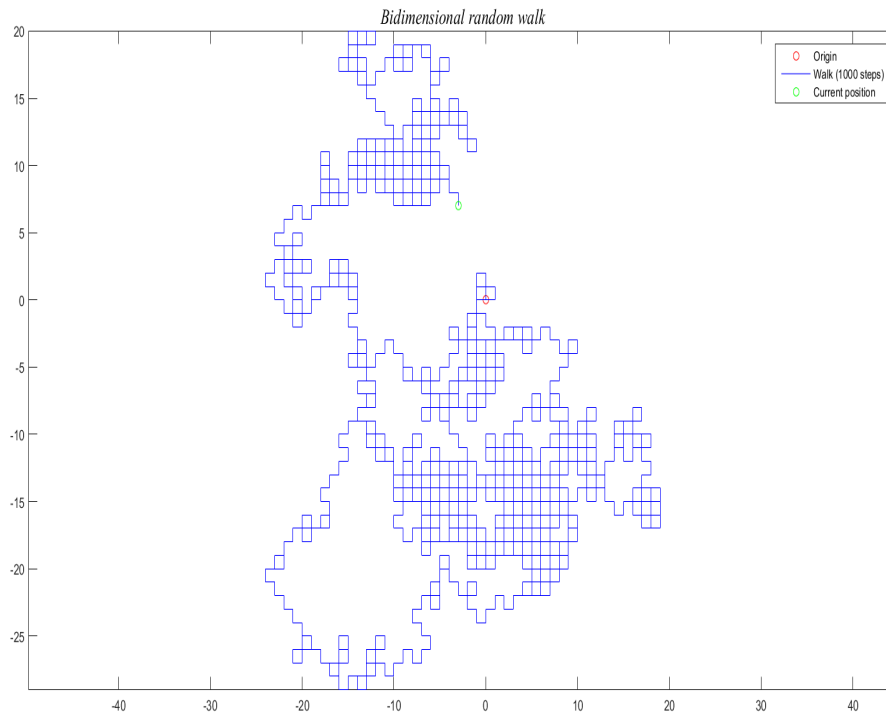


5.3. ábra. Buffon tű probléma a π közelítésére

2 dimenziós véletlen bolyongás

A véletlen bolyongás problémájának szimulálására is használhatunk Monte Carlo módszert. Legyen $S_n = X_1 + X_2 + \dots + X_n$ a bolyongást végző részecske helyzete n lépés után. A lépések egymástól függetlenek. Az 5.4 ábrán 1000 lépést tettük meg. A piros kör a kezdőpont a zöld pedig az utolsó állás.

5.2.1. Megjegyzés. *Érdekesség, hogy a szimmetrikus véletlen bolyongások témakörében elért eredményeket először hadifoglyok szökésénél alkalmazták, hogy adott idő alatt milyen messzire jutnak. Napjainkban alkalmazható a fizikában, pl. gáz és folyadékrészecskék véletlenszerű mozgásának szimulációjára, biológiában pedig pl. a populációdinamika modellezésére használják.*



5.4. ábra. 2 dimenziós véletlen bolyongás szimulálása

Egyéb alkalmazások

A Monte Carlo módszert gyakran alkalmazzák a gazdasági életben is. Két fontos felhasználási területe a kockázatos érték számítás és az opcióárazás. A módszer használható fizikai, biológiai területen is (genetikai modellezésnél, részecskék mozgásának modellezésénél). Ezekre már nem fogunk részletesen kitérni. Ezek is a Monte Carlo módszer sokrétű alkalmazhatóságára adnak bizonyítékot.

Irodalomjegyzék

- [1] **Karátson János**, *Numerikus funkcionálanalízis*, egyetemi jegyzet, Budapest, 2014.
- [2] **Christopher M. Bishop**, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag, New York, 2016.
- [3] **Günther Hämmerlin, Karl-Heinz Hoffmann**, *Numerical Mathematics*, Springer-Verlag, 1989.
- [4] **Gergó Lajos**, *Numerikus módszerek*, ELTE Eötvös Kiadó, 2010.
- [5] **Káta Imre**, *Szimulációs módszerek*, Tankönyvkiadó, Budapest, 1981.
- [6] **James E. Gentle**, *Random Number Generation and Monte Carlo Methods*, Springer-Verlag, New York, 2003.
- [7] **Bolla Marianna, Krámlí András**, *Statisztikai következtetések elmélete*, Typotex, Budapest, 2005.
- [8] **Rényi Alfréd**, *Valószínűségszámítás*, Tankönyvkiadó, Budapest, 1981.
- [9] **Ron Larson, Bruce Edwards**, *Calculus*, Brooks Cole, 2005.
- [10] **Simon Péter**, *Bevezetés az analízisbe I*, egyetemi jegyzet, Budapest, 2013.
- [11] **Kai Nordlund**, *Basics of Monte Carlo simulations*, egyetemi jegyzet, Helsinki, 2006.
- [12] **Rüdiger Seydel**, *Tools for Computational Finance*, Springer Verlag, Berlin, 2008.
- [13] **George Casella, Roger L. Berger**, *Statistical Inference (2.nd edition)*, Duxbury, 2002.

- [14] **George Casella, Roger L. Berger**, *A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data*, Statistic Science, 2014.
- [15] **Harald Niederreiter**, *On the distribution of pseudo-random numbers generated by the linear congruential method I-II-III.*, Springer Verlag, 1985.
- [16] **Pierre L'Ecuyer, Richard Simard**, *TestU01: A C Library for Empirical Testing of Random Number Generators*, *ACM Transactions on Mathematical Software*, 2007.