

# Sandsynlighedsbaserede metoder

## – Monte Carlo-metoden

Daniel Kjær

I sidste udgave af FAMØS kunne læseren finde første halvdel af en todelte artikelserie om sandsynlighedsbaserede metoder under afsnittet med titlen »*Sandsynlighedsbaserede metoder*« med undertitlen »*Pseudotilfældige tal*«. I denne artikel præsenterede forfatteren flere fundamentale begreber for pseudotilfældige tal-generatorer fra et ergodeteoretisk synspunkt.

Den foreliggende artikel er den anden halvdel i serien og denne artikel omhandler den egentlige Monte Carlo-metode. De problemer vi i det følgende vil betragte er kendetegnet ved, at de er meget enkle og fremstillingen af emnet kræver kun et indledende kendskab til målteori.

## Monte Carlo-metodens generelle idé og oprindelse

Inspiration til numeriske metoder kan springe fra alle mulige uforventede kilder. *Simuleret udglødning* har taget sit navn fra metalurgien (læren om metalleres fremstilling, egenskaber og bearbejdning) og *splines* blev først taget i anvendelse af tekniske skitsetegnere. På en tilsvarende facon er Monte Carlo-metoden rodfæstet i hazardspillet. Og hvem føler ikke en slags fjernt ekko af noget glamourøst ved navnet Monte Carlo-metoden?

Selve navnet er afledt af byen Monte Carlo i fyrstendømmet Monaco. Byen er berømt verden over for sit grand casino, »*Casino de Monte-Carlo*«, et ikon for industrien for hazardspil. Det korte af det lange er, at Monte Carlo-metoden ikke kan hjælpe én med at vinde i roulette; metoden er end ikke et forsøg på det.

Monte Carlo-metoden er en numerisk metode til løsningen af matematiske problemer ved hjælp af simulation.

Monte Carlo-metodens teoretiske fundament, i særdeleshed de store tals lov, har længe været kendt. Men siden simulation af stokastiske variable per håndkraft er en svært møjsommeligt proces, har Monte Carlo-metoder været praktisk umulige at udføre uden hjælp fra en computer. Det er derfor ikke overraskende, at Monte Carlo-metoden som universelt anvendt numerisk teknik er snævert forbundet med udbredelsen af computeren.

Monte Carlo-metodens alment accepterede oprindelsesdag er 1949, hvor en artikel under overskriften »*The Monte Carlo Method*« dukkede op. De amerikanske matematikere Jersey Neyman og Stanislaw Ulam betragtes som værende ophavsmændene til denne artikel. I Sovjetunionen publiceredes de første artikler om Monte Carlo-metoden i 1955 og 1956 af V. V. Chavchanidze, Yu. A. Schreider og V. S. Vladimirov.

For hver ny generation af computere finder Monte Carlo-metoden nye anvendelsesområder.

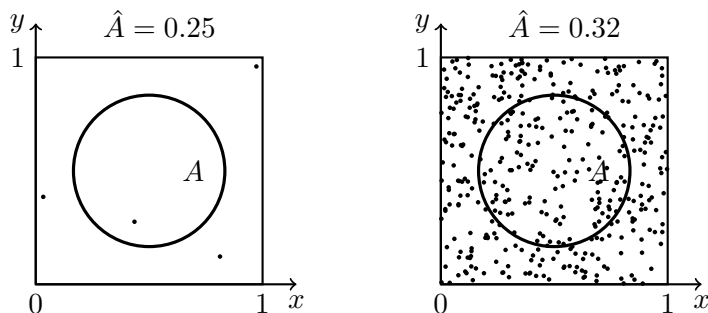
## Differentia specifica og et simpelt eksempel

Vores udgave af Monte Carlo-metoden er kendetegnet ved, for det første, den beregningsmæssige enkle struktur af algoritmen. Algoritmen består i almindelighed af en proces til produktionen af en tilfældig hændelse. Processen gentages  $n \in \mathbb{N}$  gange, med hvert forsøg uafhængigt af de forrige og resultaterne bruges til dannelsen af et gennemsnit.

Et andet særkende ved metoden er, som hovedregel, at fejlen fra udregningen er proportional med  $\sqrt{1/n}$ , hvor  $n$  er antallet af forsøg. For at opnå høj præcision må man altså have et stort antal gentagelser (for at at formindske fejlen med en faktor 10 er det

nødvendigt at øge  $n$  med en faktor 100). Lad os illustrere metoden med et simpelt eksempel.

**Eksempel 1 (Hit-or-Miss Integration)** Forestil dig, at vi har brug for at udregne arealet,  $A$ , af en figur i planen begrænset af enhedskvadratet. Det skal principielt være en kompliceret figur, før det kommer til sin ret at sætte Monte Carlo-metoden i anvendelse, men af hensyn til bekvemmelighed lader vi figuren være en cirkel som vist i figur 1.



**Figur 1** Approksimation af arealet  $A = 0.35$ .

Vælg tilfældigt  $n$  punkter i kvadratet og betegn ved  $n'$  antallet af punkterne, der falder indenfor cirklen. Det er fra figuren klart, at cirkelns areal approksimativt er lig med  $n'/n$ . Jo større  $n$ , desto bedre approksimation.

I skemaet til venstre ser vi, at 1 af 4 punkter er faldet indenfor cirklen, så vores estimat for cirkelns areal er  $n'/n = 1/4 = 0.25$ . I skemaet til højre er 128 af 400 punkter faldet indenfor cirklen, så vores estimat for cirkelns areal er  $n'/n = 128/400 = 0.32$ . Cirkelns faktiske areal er lig med  $A = 0.35$ .

Observer, at metoden til estimation af arealet  $A$  kun er gyldig, når de tilfældige punkter ikke blot er ‘tilfældige punkter’ i enhedskvadratet, men faktisk er ‘ligeligt fordelt’ på hele enhedskvadratet.

For at kunne replikere eksemplet kræves altså en metode til produktionen af ligefordelte stokastiske variable i enhedskvadratet. Dette skridt består i virkeligheden i at kalde en subrutine i et computerprogram, en s.k. *pseudotilfældig tal-generator*, som udfører den bestemte opgave at producere visse følger af tal, der til ethvert praktisk formål er uskelnelige fra udfald af ligefordelte stokastiske variable. Denne frembringelsesproces blev beskrevet i detaljer i sidste udgave af FAMØS i artiklen »*Pseudotilfældige tal*«.

## Simple sandsynlighedsbaserede modeller

### Fraktiltransformation metoden

Vi beskriver i dette afsnit en simpel model til frembringelsen af værdier af en stokastisk variabel med en given sandsynlighedsfordeling.

**Definition 2** Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt. Lad  $X$  være en reel stokastisk variabel på  $\Omega$  og lad  $P_X$  være fordelingen af  $X$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ . Ved  $p$ -fraktilerne hørende til  $P_X$  for et  $p \in (0, 1)$  forstås mængden

$$\left\{ x \in \mathbb{R} : \lim_{y \uparrow x} P_X((-\infty, y]) \leq p \leq P_X((-\infty, x]) \right\}.$$

Et element i mængden kaldes en  $p$ -fraktil for  $P_X$ . Ved en fraktilfunktion hørende til  $P_X$  mener vi en funktion  $q$  på  $(0, 1)$  ind i  $\mathbb{R}$  således at for alle  $p \in (0, 1)$ , så er  $q(p)$  en  $p$ -fraktil for  $P_X$ .

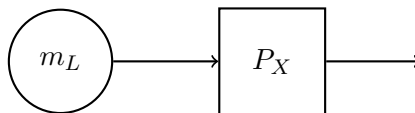
Vigtigheden af følgende sætning er ikke til at tage fejl af. Sætningen understreger den centrale rolle, som ligefordelingen spiller i simulationsmodellering.

**Sætning 3** *Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt. Lad  $X$  være en reel stokastisk variabel på  $\Omega$ , lad  $P_X$  være fordelingen af  $X$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$  og lad  $q$  være en fraktilfunktion hørende til  $P_X$ . Betragt sandsynlighedsfeltet  $((0, 1), \mathfrak{B}_{(0,1)}, m_L)$ , hvor  $m_L$  er Lebesgue målet på Borel  $\sigma$ -algebraen  $\mathfrak{B}_{(0,1)}$  af delmængder af  $(0, 1)$ , og lad  $U$  være en stokastisk variabel på  $\Omega$  ind i  $(0, 1)$  med fordeling  $m_L$ . Da gælder, at*

$$P_X = (q \circ U)(P).$$

*Bevis.* Se [3]. □

Lad  $X$  være en reel stokastisk variabel på et bagvedliggende sandsynlighedsfelt  $(\Omega, \mathfrak{F}, P)$  og lad  $q$  være en fraktilfunktion hørende til fordelingen  $P_X$ . Frembringelsesprocessen af stokastiske variable med denne fordeling er sammensat som en serieforbindelse af to delprocesser, som i figur 2.



**Figur 2** Frembringelse af stokastiske variable.

Den første delproces i udarbejdelsen af input til en simulation består i, at frembringe en ligefordelt stokastisk variabel  $U$  på det ovennævnte sandsynlighedsfelt. Dernæst fokuseres på at lade den ligefordelte stokastiske variabel undergå metamorfosen til  $q \circ U$ .

Sætningen klargør, at denne produktionsproces faktisk giver et udfald af en stokastisk variabel med den ønskede fordeling  $P_X$ . Lad os illustrere metoden med et simpelt eksempel.

**Eksempel 4** Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt. Lad  $X$  være positiv reel stokastisk variabel og lad for et  $\lambda > 0$ ,  $\exp(\lambda)$  være fordelingen af  $X$  på  $(\mathbb{R}_+, \mathfrak{B}_{\mathbb{R}_+})$ . Fraktilfunktionen  $q$  hørende til  $P_X$  er givet ved

$$q(p) = -\frac{\log(1-p)}{\lambda}, \quad p \in (0, 1).$$

1. Lad  $U$  være en ligefordelt stokastisk variabel defineret på  $(\Omega, \mathfrak{F}, P)$ .
2. Den stokastiske variabel  $-\log(1-U)/\lambda$  på  $(\Omega, \mathfrak{F}, P)$  er fordelt  $\exp(\lambda)$  på  $(\mathbb{R}_+, \mathfrak{B}_{\mathbb{R}_+})$ .

Der findes situationer, hvor man gerne vil simulere stokastiske variable med en fordeling, hvor fordelingen er så kompliceret, at det ikke er helt enkelt at gå direkte til fraktilfunktionen.

**Eksempel 5** Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt. Lad  $X$  være en reel stokastisk variabel og lad  $N(0, 1)$  være fordelingen af  $X$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ . Genkald, at fordelingsfunktionen hørende til  $P_X$ , betegnet ved  $\Phi$ , er givet ved

$$\Phi(y) = \int_{(-\infty, y]} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} m_L(dx), \quad y \in \mathbb{R}.$$

Da det er klart, at  $\Phi \in C^1(\mathbb{R})$  med positiv differentialkvotient, så er funktionen strengt voksende på  $\mathbb{R}$  og en fraktilfunktion  $q$  hørende til  $P_X$  eksisterer entydigt, som den inverse af  $\Phi$ .

For så vidt som fordelingsfunktionen  $\Phi$  ikke kan gives et analytisk udtryk (uden brug af specielle funktioner), så kan fraktiltransformationsmetoden ikke direkte anvendes.

Eksemplet klargør, at alternative metoder til frembringelse af normalfordelte variable må tages i brug. Vi nævner i flæng: *Numerisk approksimation* af  $q$ , *Box-Muller*, *Acceptance-Rejection*, etc.

**Sætning 6** (Den centrale grænseværdisætning) *Lad  $\{X_n: n \in \mathbb{N}\}$  være en familie af reelle stokastiske variable på et sandsynlighedsfelt  $(\Omega, \mathfrak{F}, P)$ . Antag, at følgende tre betingelser er opfyldt.*

1. *Familien  $\{X_n: n \in \mathbb{N}\}$  er et uafhængigt system af stokastiske variable på  $(\Omega, \mathfrak{F}, P)$ .*
2. *Der findes et sandsynlighedsmål  $Q$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$  således at  $Q$  er fordelingen af  $X_n$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$  for ethvert  $n \in \mathbb{N}$ .*
3. *Følgende integraler eksisterer og er endelige:  $\mu \equiv \mathbf{E}(X_n)$ ,  $\sigma^2 \equiv \mathbf{E}[(X_n - \mu)^2] > 0$ ,  $n \in \mathbb{N}$ .*

*Lad  $\{Z_n: n \in \mathbb{N}\}$  være systemet af stokastiske variable på  $(\Omega, \mathfrak{F}, P)$  defineret ved*

$$Z_n = \frac{\sum_{k=1}^n X_k - n\mu}{\sigma\sqrt{n}}.$$

*Da gælder, for alle  $a, b \in \mathbb{R}$  hvor  $a < b$ , at*

$$\lim_{n \rightarrow \infty} P_{Z_n}((a, b]) = \Phi(b) - \Phi(a),$$

*hvor  $\Phi$  er fordelingsfunktionen hørende til fordelingen  $N(0, 1)$ .*

*Bevis.* Se [1]. □

**Eksempel 7** Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt. Lad  $X$  være en reel stokastisk variabel og lad  $N(0, 1)$  være fordelingen af  $X$  på  $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ . Fraktilfunktionen  $q$  hørende til  $P_X$  eksisterer ikke på eksplicit form.

1. Lad  $\{U_n: n = 1, \dots, 12\}$  være et uafhængigt system af ligefordelte stokastiske variable defineret på  $(\Omega, \mathfrak{F}, P)$ . Observér, at  $\mu \equiv \mathbf{E}(U_n) = 1/2$ ,  $\sigma = \mathbf{E}[(U_n - \mu)^2]^{1/2} = 1/\sqrt{12}$ .
2. Den stokastiske variabel  $\sum_{k=1}^{12} U_k - 6$  på  $(\Omega, \mathfrak{F}, P)$  er fordelt approksimativt  $N(0, 1)$ .

## Monte Carlo-integration

Monte Carlo-integration er en numerisk metode til approksimation af integraler ved hjælp af sandsynlighedsbaserede metoder og næsten enhver Monte Carlo udregning kan betragtes som et forsøg på at approksimere integralet

$$\int_D f \, dm_L,$$

for et  $f \in L_1(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, m_L)$  og  $D \in \mathfrak{B}_{\mathbb{R}}$ , hvor  $m_L$  er Lebesgue målet på Borel  $\sigma$ -algebraen  $\mathfrak{B}_{\mathbb{R}}$  af delmængder af  $\mathbb{R}$ . Vi ser, at integralet kan repræsenteres som en middelværdi af en stokastisk variabel.

**Sætning 8** (Monte Carlo-integration) *Lad  $f \in \mathcal{M}(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$  og lad  $D \in \mathfrak{B}_{\mathbb{R}}$ . Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt og lad  $X$  være en stokastisk variabel på  $\Omega$  ind i  $D$  og lad  $P_X$  være fordelingen af  $X$  på  $(D, \mathfrak{B}_D)$ . Antag, at  $P_X$  er absolut kontinuert med hensyn til  $m_L$  og lad  $p \equiv \frac{dP_X}{dm_L}$  være den Radon-Nikodym afledede af  $P_X$  med hensyn til  $m_L$ .*

1. *Der gælder, at*

$$\mathbf{E} \left[ \frac{f(X)}{p(X)} \right] = \int_D f \, dm_L,$$

*i den forstand, at eksistensen af den ene side medfører eksistensen af den anden og lighedstegnet.*



2. Dersom integralerne i 1. eksisterer og dersom  $\{X_n: n \in \mathbb{N}\}$  er et uafhængigt system af stokastiske variable på  $\Omega$  ind i  $D$  med fælles fordeling  $P_X$  af  $X_n$  for ethvert  $n \in \mathbb{N}$ , så findes en nulmængde  $\Lambda$  i  $(\Omega, \mathfrak{F}, P)$  således, at

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \frac{f(X_k)(\omega)}{p(X_k)(\omega)} = \int_D f \, dm_L, \quad \omega \in \Lambda^c.$$

*Bevis.* 1. er oplagt og 2. følger af STORE TALS STÆRKE LOV.  $\square$

**Eksempel 9** (Simpel Monte Carlo-integration) Lad  $\{U_n: n \in \mathbb{N}\}$  være et uafhængigt system af stokastiske variable på et sandsynlighedsfelt  $(\Omega, \mathfrak{F}, P)$  ind i  $(a, b)$  for  $a, b \in \mathbb{R}$  med  $a < b$ . Lad ligefordelingen på  $(a, b)$  være fordelingen af  $U_n$  for ethvert  $n \in \mathbb{N}$ . Lad  $f \in L_1(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, m_L)$ .

1. Der gælder, at

$$(b - a)\mathbf{E}[f(U_1)] = \int_{(a,b)} f \, dm_L, \quad n \in \mathbb{N}.$$

2. Der findes en nulmængde  $\Lambda$  i  $(\Omega, \mathfrak{F}, P)$  således, at

$$(b - a) \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(U_k)(\omega) = \int_{(a,b)} f \, dm_L, \quad \omega \in \Lambda^c.$$

Det er klart, at den ovenstående procedure for Monte Carlo-integration let lader sig generalisere til flere dimensioner. Dette generalisationsproblem overlades til læseren.

## Monte Carlo-optimering

Simulation af stokastiske variable i numerisk optimering er en anden stærk og meget populær anvendelse af Monte Carlo-metoden.

Problemet er at minimere (eller maksimere) funktioner, som ofte har et stort antal dimensioner.

Lad  $\mathbf{W}(\mathbb{R}^d)$ , for et  $d \in \mathbb{N}$ , betegne samlingen af alle kontinuerte funktioner  $w$  på  $\mathbb{R}^d$  ind i  $\mathbb{R}$ . Vores ambition i dette afsnit er at beskrive metoder til løsning af (evt. multiektremale) optimeringsproblemer:

$$\begin{cases} \min_{x \in D} w(x), & w \in \mathbf{W}(\mathbb{R}^d), \\ D \text{ kompakt}, & D \subset \mathbb{R}^d. \end{cases}$$

Mange teoretiske og praktiske problemstillinger kan modelleres indenfor denne generelle ramme.

**Lemma 10** (Borel-Cantelli) *Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt og lad  $F_n \in \mathfrak{F}$  for  $n \in \mathbb{N}$ . Hvis*

$$\sum_{n \in \mathbb{N}} P(F_n) < \infty,$$

så er

$$P(\cap_{n \in \mathbb{N}} \cup_{k=n}^{\infty} F_k) = 0.$$

*Bevis.* Se [4]

□

**Sætning 11** (Simpel Monte Carlo-optimering) *Lad  $w \in \mathbf{W}(\mathbb{R}^d)$ , for et  $d \in \mathbb{N}$ , være en kontinuert funktion på  $\mathbb{R}^d$  ind i  $\mathbb{R}$  og lad  $D$  være en kompakt delmængde af  $\mathbb{R}^d$ . Antag, at der findes et entydigt punkt  $x^*$  i  $\mathbb{R}^d$  således, at*

$$x^* = \arg \min_{x \in D} w(x).$$

*Lad  $(\Omega, \mathfrak{F}, P)$  være et sandsynlighedsfelt og lad  $\{X_n: n \in \mathbb{Z}_+\}$  være et uafhængigt system af  $d$ -dimensionale stokastiske variable på  $\Omega$*

ind i  $D$  med fælles fordeling  $P_X$  af  $X_n$  på  $(D, \mathfrak{B}_D)$  for ethvert  $n \in \mathbb{Z}_+$  og antag at  $\{x^*\}$  er en nulmængde i  $(D, \mathfrak{B}_D, P_X)$ . Lad  $\{Y_n: n \in \mathbb{Z}_+\}$  være et system af reelle stokastiske variable på  $(\Omega, \mathfrak{F}, P)$  givet ved

$$Y_n = w(X_n), \quad n \in \mathbb{Z}_+.$$

Hvis der for ethvert  $\delta > 0$  gælder, at

$$P_{Y_n}((w(x^*), w(x^*) + \delta]) > 0,$$

så findes en nulmængde  $\Lambda$  i  $(\Omega, \mathfrak{F}, P)$  således, at

$$\lim_{n \rightarrow \infty} \min\{Y_1, \dots, Y_n\}(\omega) = w(x^*), \quad \omega \in \Lambda^c.$$

*Bevis.* Siden  $\{x^*\}$  er en nulmængde i  $(D, \mathfrak{B}_D, P_X)$ , så er  $(X_n)^{-1}(\{x^*\})$  en nulmængde i  $(\Omega, \mathfrak{F}, P)$  for ethvert  $n \in \mathbb{Z}_+$ . Men  $(X_n)^{-1}(\{x^*\}) = (Y_n)^{-1}((-\infty, w(x^*)])$ . Således har vi etableret, at

$$P_{Y_n}((-\infty, w(x^*)]) = 0. \quad (1)$$

Lad  $\delta > 0$  være givet. Jævnfør vor antagelse har vi, at

$$P_{Y_n}((-\infty, w(x^*) + \delta]) = P_{Y_n}((w(x^*), w(x^*) + \delta]) > 0,$$

hvor lighedstegnet følger af (1). Fra dette og fra uafhængigheden af systemet  $\{Y_n: n \in \mathbb{Z}_+\}$  har vi, at

$$P_{\min\{Y_1, \dots, Y_n\}}((w(x^*) + \delta, \infty)) = \{1 - P_{Y_n}((w(x^*), w(x^*) + \delta])\}^n$$

hvilket giver anledning til følgende geometriske række:

$$\sum_{n \in \mathbb{N}} P_{\min\{Y_1, \dots, Y_n\}}((w(x^*) + \delta, \infty)) = \frac{1 - P_{Y_n}((w(x^*), w(x^*) + \delta])}{P_{Y_n}((w(x^*), w(x^*) + \delta])}.$$

Den højre side af lighedstegnet er oplagt endelig. Fra Lemma 10 (Borel-Cantelli) har vi for ethvert  $\delta > 0$ , at

$$P(\cap_{n \in \mathbb{N}} \cup_{k=n}^{\infty} \{\min\{Y_1, \dots, Y_k\} \geq w(x^*) + \delta\}) = 0,$$

eller ækvivalent hermed: for ethvert  $\delta > 0$ , så er

$$P(\cup_{n \in \mathbb{N}} \cap_{k=n}^{\infty} \{\min\{Y_1, \dots, Y_k\} < w(x^*) + \delta\}) = 1.$$

Heraf sluttet det ønskede resultat. □

## Litteratur

- [1] Billingsley, P., *Convergence of Probability Measures*, John Wiley & Sons, Inc., 1968.
- [2] Hammersley, J. M., & Handscomb, D. C., *Monte Carlo Methods*, Methuen's Monographs on Applied Probability and Statistics, London, 1964.
- [3] Hansen, E., *Measure Theory*, Institut for Matematiske Fag, Københavns Universitet, 4. udgave, 2009.
- [4] Jacobsen, M., *Videregående Sandsynlighedsregning*, Institut for Matematiske Fag, Københavns Universitet, 3. udgave, 2003.
- [5] Metropolis, N. & Ulam, S., *The Monte Carlo method*, J. Am. Stat. Assoc., **44**, N 247, 335–341, 1949
- [6] Sobol', I. M., *A Primer for the Monte Carlo Method*, Boca Raton: CRC Press, Inc., 1994