

# Pràctiques de Fenòmens col·lectius i transicions de fase

Mètodes de simulació en Física Estadística

---

Jordi Baró (*material adaptat de notes Eduard Vives*)

Tardor 2024

Departament de Física de la Matèria Condensada

Facultat de Física

Universitat de Barcelona

## Calendari ( 6+2 hores )

- Sessió introducció (2h): avui 18/11 16:00-17:30
- Sessions 1,2 i 3 (3 × 2h) **Assistència Obligatoria**

(comencen aquest  
divendres!)

22/11 – 28/11    P1  
29/11 – 5/12     P2  
9/12 – 13/12    P3

|             | dll   | dt | dc | dj | dv    |
|-------------|-------|----|----|----|-------|
| 8:30-10:30  |       |    |    |    | E1-E2 |
| 16:00-18:00 | A2-A3 |    |    |    |       |
| 17:00-19:00 |       |    |    |    | E3-E4 |
| 18:00-20:00 |       |    |    | D2 |       |

- Entrega passat Nadal:
- Codis + Petits exercicis emprant els resultats.
- Estricta política de **plagi**.

### Professorat

|               |  |
|---------------|--|
| <b>A2, A3</b> | Mozota, Álvaro<br>Benavent, Andreu<br>Calero, Carles |
|               |  |
| <b>D2</b>     | Benavent, Andreu<br>Baró, Jordi                      |
|               |  |
| <b>E1</b>     | Dosil, Miguel<br>Levis, Demian                       |
| <b>E2</b>     | Mozota, Álvaro<br>Calero, Carles                     |
|               |  |
| <b>E3</b>     | Benavent, Andreu<br>Baró, Jordi                      |
| <b>E4</b>     | Mozota, Álvaro<br>Rodriguez, Víctor                  |

### Avaluació:

- Assistència obligatòria a les 6 hores (3 sessions)
- Entrega obligatòria de l'exercici escrit **individual i original** i codi. Passat Nadal (no apureu!).
- Entrega codi + exercicis: 1 punt de la nota final de l'assignatura
- Examen final: conté una qüestió relacionada amb les pràctiques

### Com es fan les pràctiques?

- FORTRAN (gfortran), GNUPLOT, (reviseu el software de F. Comp.)
- editor de textos: gedit, geany, notepad++, ...
- Sistema operatiu: LINUX, WINDOWS, OSX
- Es treballa a l'aula d'informàtica. Es pot treballar amb portàtil.

1. Simulació a la Física Estadística.
  - 1.1 Repàs de F.E.: models i mesura (integral fonamental de la F.E.).
  - 1.2 Problemes amb la resolució de la integral.
  - 1.3 Mètodes deterministes vs. estocàstics.
2. Monte-Carlo en cadenes de Markov (MC-MC) i Metropolis.
  - 2.1 Cadenes de Markov.
  - 2.2 Equació de balanç.
  - 2.3 Probabilitats de transició.
  - 2.4 Algorisme de Metropolis.
3. Annex: Repàs Física Computacional.

# 1 Simulación a la Física Estadística

---

- Models de la F.E.
  - Microestat: configuració  $X = (x_1, x_2, x_3, \dots, x_{\mathcal{N}})$
  - Espai fàsic  $\Omega = \{X\}$ , amb  $\mathcal{N}$  dimensions. (Per exemple  $\mathbb{R}^{\mathcal{N}}$  o  $2^{\mathcal{N}}$  )
  - Hamiltonià  $\mathcal{H}(X) = \mathcal{H}(x_1, x_2, \dots, x_{\mathcal{N}})$
  - Variables d'interès:  $A(X) = A(x_1, x_2, \dots, x_{\mathcal{N}})$ : exemples  $\mathcal{H}$ ,  $K$ ,  $U$ ,  $g(r)$  ...
  - Límit termodinàmic  $\mathcal{N} \rightarrow \infty$
- Exemples models:
  - Models continus: Flúid clàssic o semiclàssic
  - Models reticulars: Model d'Ising

Sistema de  $N$  partícules clàssiques que representa un sòlid, líquid o gas.

$$\mathcal{H}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, \vec{p}_1, \vec{p}_2, \dots, \vec{p}_N) = K + U = \sum_{i=1}^N \frac{(\vec{p}_i)^2}{2m_i} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$$

L'espai de les fases té dimensió  $\mathcal{N} = 6N$

E.g., potencials a parelles de Lennard-Jones

$$U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \sum_{i,j}^{N(N-1)/2} V(|\vec{r}_i - \vec{r}_j|) = \sum_{i,j}^{N(N-1)/2} 4\epsilon \left[ \left( \frac{\sigma}{|\vec{r}_i - \vec{r}_j|} \right)^{12} - \left( \frac{\sigma}{|\vec{r}_i - \vec{r}_j|} \right)^6 \right]$$

Equacions dinàmiques:

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial \mathcal{H}}{\partial \vec{r}_i} = -\frac{\partial U}{\partial \vec{r}_i} \quad 3N \text{ equacions}$$

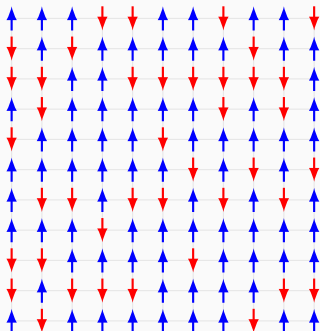
$$\frac{d\vec{r}_i}{dt} = \frac{\partial \mathcal{H}}{\partial \vec{p}_i} = \frac{\vec{p}_i}{m_i} \quad 3N \text{ equacions}$$



## Model d'Ising (I)

Xarxa d' $N$  nusos (1d, 2d, 3d,...). A cada nus es defineix una variable de spin:

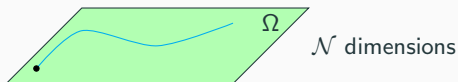
$$S_i = \pm 1 \quad i = 1 \dots N$$
$$\mathcal{H}(S_1, S_2, \dots, S_N) = -J \sum_{i,j}^{n.n.} S_i S_j - H \sum_{i=1}^N S_i$$



Aplicació multidisciplinària a transicions de  $1r/2n$  ordre: magnetisme, opinions, infeccions, borsa,... [ book: Mark Buchanan, *Ubiquity* (2002) ]

Per resoldre els models en el marc de la mecànica clàssica, farem servir les equacions de Newton (o Hamilton) per trobar trajectòries:

- Equacions dinàmiques
  - Condicions inicials  $X(0)$
- } trajectòries  
 $X = X(t)$



Promitjos temporals:

$$\bar{A} = \lim_{t \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_t^{t+T} A(X(t)) dt$$

Aquests tipus de promitjos ens donaria el comportament de les quantitats mesurables en un sistema en equilibri

La mida del sistema d'equacions a resoldre ( $\mathcal{N} \rightarrow \infty$ ) ho fa impossible i, a més a més, en alguns casos (Ising) no tenim equacions dinàmiques.

La F.E. d'equilibri es basa en el **principi ergòdic**, si la dinàmica és tal que el sistema assoleix l'equilibri termodinàmic, els promitjos temporals es poden substituir per promitjos fàscs:

$$\langle A \rangle = \int_{\Omega} p(X) A(X) dX = \sum_{\Omega} p(X) A(X)$$

on  $p(X) = f(X)/\mathcal{Z}$  és una probabilitat o densitat de probabilitat convenient, que depèn de les característiques de les parets del sistema que es troba en equilibri i la funció de partició  $\mathcal{Z}$  és el factor que normalitza la probabilitat sobre  $\Omega$ .

- Sistema totalment aïllat: col·lectivitat microcanònica

$$f(X) = 1 \quad \mathcal{Z} = \int_{\Omega} dX = \text{Vol}_{\Omega} \quad p(X) = \frac{1}{\text{Vol}_{\Omega}}$$

- Sistema en contacte tèrmic amb un bany a  $T \propto \frac{1}{\beta}$ : col·lectivitat canònica

$$f(X) = e^{-\beta \mathcal{H}(X)} \quad \mathcal{Z} = \int_{\Omega} e^{-\beta \mathcal{H}(X)} dX \quad p(X) = \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}}$$

- Sistema obert en massa: col·lectivitat gran canònica

Els càlculs en **F.E. d'equilibri**, per tant, es redueixen a resoldre l'anomenada **integral fonamental**:

$$\langle A \rangle = \int_{\Omega} A(X) \frac{f(X)}{\mathcal{Z}} dX \quad \langle A \rangle = \sum_{\Omega} \frac{f(X)}{\mathcal{Z}} A(X)$$

Es tracta d'una integral multidimensional, en un espai  $\Omega$  que en el límit termodinàmic té infinites dimensions.

Noteu també que als denominadors hi ha una integral (o suma)  $\mathcal{Z}$  que també té la mateixa dificultat.

$$\mathcal{Z} = \int_{\Omega} f(X) dX \quad \mathcal{Z} = \sum_{\Omega} f(X)$$

# Problemes en la resolució de la integral (1)

Les mesures a F.E. es redueixen a calcular integrals del tipus:

$$\langle A \rangle = \int_{\Omega} A(X) \frac{f(X)}{\mathcal{Z}} dX$$

## 1r PROBLEMA: valor de $\mathcal{Z}$

- Ex. Model d'Ising, a la canònica.

$$\mathcal{Z} = \sum_{S_1, S_2, S_3, S_N}^{2^N} e^{\beta \sum_{\langle i,j \rangle}^{n,n} S_i S_j}$$

- Generalment  $\mathcal{Z}$  no es pot calcular analíticament. (Si es pogués calcular, ja no caldria resoldre el model amb mètodes de simulació numèrica).

### 2n PROBLEMA: Dimensió $\mathcal{N}$ de l'espai de fases

- Mètodes lineals (e.g. trapezis, Simpson) creixen ràpidament en complexitat.

$$\text{error trapezis : } \varepsilon(h) \propto h^2 \propto M^{-2/d}$$

$$\text{error Simpson : } \varepsilon(h) \propto h^4 \propto M^{-4/d}$$

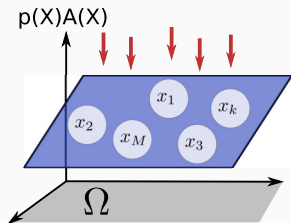
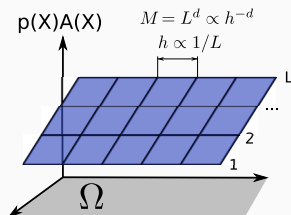
- Alternativa: **Integració Monte-Carlo cru**, mostreig aleatori de valors  $\{x_1, x_2, \dots, x_M\}$ , independents. Estimem una integral  $I$  com:

$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M f(x_i) \quad \rightarrow \quad \langle \hat{A} \rangle = \frac{1}{M} \sum_{i=1}^M A(x_i) p(x_i)$$

$$\text{error M.C. : } \varepsilon(h) \propto M^{-1/2}$$

Per  $d > 8$  els mètodes M.C. poden ser més eficients que els lineals.

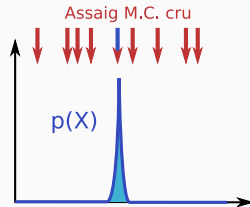
(E.g., Ising 2x2 lattice  $d = \mathcal{N} = L \times L$ )



## Problemes en la resolució de la integral (3)

### 3r PROBLEMA: Integrem molts valors $p(x_i) \approx 0$

- El nombre de possibles estats és molt gran  
E.g., Model d'Ising:  $\#(\Omega) = 2^N$
- En equilibri termodinàmic (canònica i gran canònica), els estats amb  $p(x_i)$  probables són una petita part de  $\Omega$ .



- Podem corregir-ho si els estats que seleccionem ja són distribuïts segons una funció (no-normalitzada)  $g(x_i) \propto \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}}$  (Importance sampling).  
Llavors:

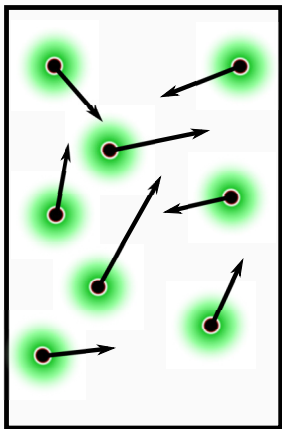
$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M \frac{f(x_i)}{g(x_i)} \quad \rightarrow \quad \langle \hat{A} \rangle = \frac{1}{M} \sum_{i=1}^M A(x_i) \frac{p(x_i)}{p(x_i)} = \boxed{\frac{1}{M} \sum_{i=1}^M A(x_i)}$$

. L'error decau més depressa que en MC cru (appendix).

- Cal trobar un **mètode** per generar un conjunt d'estats independents  $\{X_1, X_2, \dots, X_I, \dots, X_M\}$  (configuracions del sistema) en l'espai de les fases  $\Omega$ , distribuïts d'acord amb  $f(X)/\mathcal{Z}$  sense conèixer  $\mathcal{Z}$ .

# Mètodes classics de simulació

E.g., fluid classic  
(potencial L-J)



Mètodes deterministes o Dinàmica Molecular

*Problema originalment microcanònic ( $N, E$ ).*

Problema del valor inicial per EDOs. Integrem seqüències temporals  $X_I(t)$  :

$$\begin{matrix} m_i \\ \vec{r}_i \end{matrix} \nearrow \vec{p}_i$$

$$X_I(t) = \begin{cases} \dot{\vec{p}}_i = -\frac{\partial U(\{r_j\})}{\partial r_i} \\ \dot{\vec{q}}_i = -\frac{\vec{p}_i}{m_i} \end{cases}$$

*Resolem per Euler, Verlet, Runge-Kutta ...*

**Tornem a l'inici. Però ara busquem en  $p(x_i)$**

Mètodes estocàstics o Monte Carlo

*Problema originalment canònic ( $T$  constant).*

Generem configuracions (sense dinàmica) amb el pes adequat:

$$p(X) = \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}}$$

*Resolem amb mètodes lliures de  $\mathcal{Z}$*

*Markov-chain Monte-Carlo (MC-MC).*



## 2 MC-MC i Metropolis

*Instead of choosing configurations randomly, then weighting them with  $\exp(-E/kT)$ , we choose configurations with a probability  $\exp(-E/kT)$  and weight them evenly.*

– Metropolis et al., 1953

## Mètode de Monte Carlo a la collectivitat canònica

Ens centrem en el cas canònic i amb un espai de les fases  $\Omega$  discret.

L'objectiu és generar una seqüència d'estats  $\{X_1, X_2, \dots, X_I, \dots, X_M\}$  a l'espai  $\Omega$ , independents i distribuïts d'acord amb la probabilitat canònica

$$p(X) = \frac{f(X)}{\mathcal{Z}} = \frac{e^{-\beta\mathcal{H}(X)}}{\sum_{Y \in \Omega} e^{-\beta\mathcal{H}(Y)}}$$

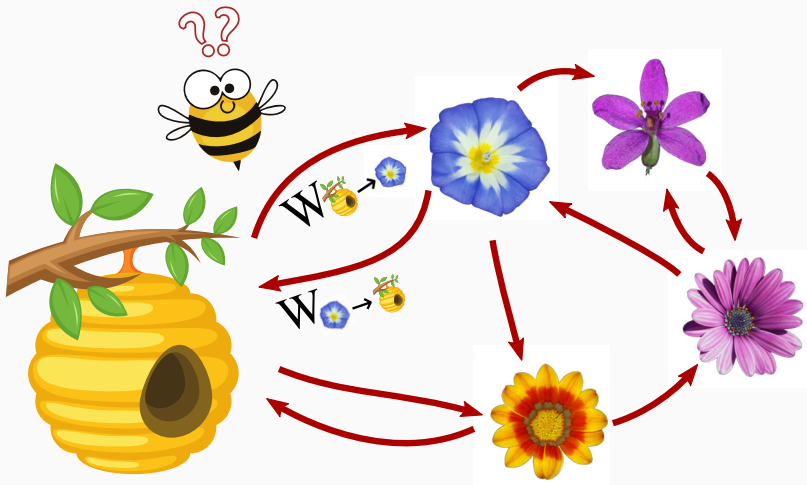
Aleshores, els promitjos de les quantitats d'interès els podrem estimar com

$$\langle A \rangle = \sum_{X \in \Omega} A(X) \frac{f(X)}{\mathcal{Z}} \sim \langle A \rangle_M = \frac{1}{M} \sum_{I=1}^M A(X_I)$$

Per aconseguir-ho, ens cal desenvolupar un concepte matemàtic anomenat cadena de Markov (correspon a la versió discreta dels anomenats processos de Markov).

**Què és una cadena de Markov?**

# Introducció a les cadenes de Markov (C.M.) (I)



## Introducció a les cadenes de Markov (C.M.) (II)

[ D.Revuz, *Markov Chains*, North Holland (1975). ]

[ V.Petrov & E.Mordecki, *Teoria de Probabilidades*, URSS (2002). ]

Es defineixen sobre un espai de configuracions  $\Omega = \{X\}$ . (Per nosaltres és l'espai de les fases)

Una C.M. és una seqüència de v.a.  $X_1, X_2 \dots X_I, \dots$  tals que les transicions  $X_I \rightarrow X_{I+1}$  són estadísticament independents. Per ternir-les ben definides calen dos elements:

- Distribució de probabilitats inicial

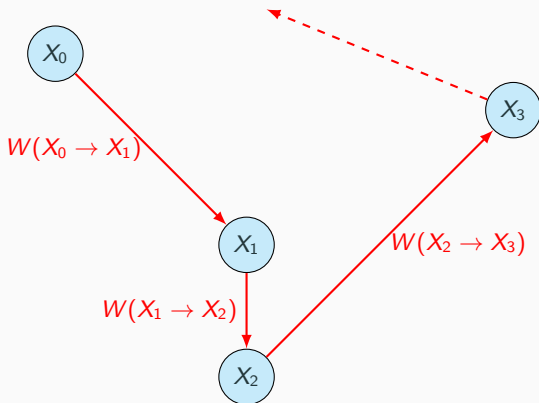
$$p_0(X) : \Omega \rightarrow [0, 1] \quad \sum_{X \in \Omega} p_0(X) = 1$$

- Matriu de probabilitats de transició

$$W(X \rightarrow Y) : \Omega \times \Omega \rightarrow [0, 1] \quad \forall X \quad \sum_{Y \in \Omega} W(X \rightarrow Y) = 1$$

$\Omega$ 

$$p_0(X_0 = x)$$



La probabilitat de transició  $X_l = x \rightarrow X_{l+1} = y$ :  $W(x \rightarrow y)$  no depèn de els valors de  $X_k$  per passos  $k$  anteriors a  $l$ .

El procés només depèn del pas anterior. No té memòria.

- Probabilitat d'una seqüència determinada  $p(X_0, X_1, X_2, \dots, X_l, \dots)$   
Donat que els salts son independents, tenim:

$$p(X_0, X_1, X_2, \dots, X_l, \dots) = p(X_0)W(X_0 \rightarrow X_1)W(X_1 \rightarrow X_2) \dots$$

**Experiment mental:** Imaginen que generem  $n$  seqüències

$$\text{seq.1 : } X_0^1 \rightarrow X_1^1 \rightarrow X_2^1 \rightarrow X_3^1 \rightarrow \dots \rightarrow X_l^1 \rightarrow \dots$$

$$\text{seq.2 : } X_0^2 \rightarrow X_1^2 \rightarrow X_2^2 \rightarrow X_3^2 \rightarrow \dots \rightarrow X_l^2 \rightarrow \dots$$

$$\text{seq.3 : } X_0^3 \rightarrow X_1^3 \rightarrow X_2^3 \rightarrow X_3^3 \rightarrow \dots \rightarrow X_l^3 \rightarrow \dots$$

.....

$$\text{seq.n : } X_0^n \rightarrow X_1^n \rightarrow X_2^n \rightarrow X_3^n \rightarrow \dots \rightarrow X_l^n \rightarrow \dots$$

- Distribució de configuracions després de  $l$  passes  $p_l(X)$

$$p_l(X_l = y) = \sum_x p_{l-1}(X_{l-1} = x)W(x \rightarrow y)$$

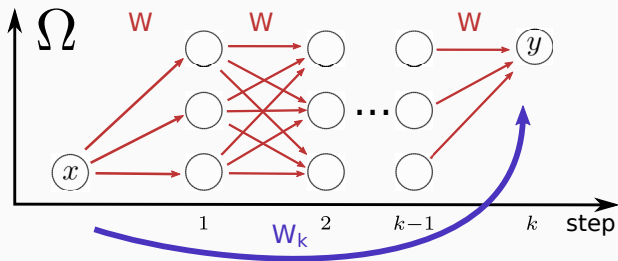
Noteu que la suma en principi pot tenir infinits termes ( $\sum \rightarrow \int$ ) i que, en general, existeix  $W(X \rightarrow X) \geq 0$ .

## Definicions II

- Probabilitat de transició en  $k$  passes  $W_k(X \rightarrow Y)$

Espai  $\Omega \in \mathbb{R}$ :  $W_k(X_I = x \rightarrow X_{I+k} = y) =$

$$= \int_{x_1} \int_{x_2} \cdots \int_{x_{k-1}} W(x \rightarrow x_1) W(x_1 \rightarrow x_2) \cdots W(x_{k-1} \rightarrow y) dx_1 \cdots dx_{k-1}$$



En el cas que l'espai de configuracions sigui finit  $\Omega = \{x_1, x_2, x_3, x_4, \dots x_N\}$  (e.g. *Ising amb  $L$  finit*), les sumes anteriors son finites i podem fer servir una notació matricial:

$$p_0 \equiv \begin{pmatrix} p_0(x_1) \\ p_0(x_2) \\ p_0(x_3) \\ \dots \\ p_0(x_N) \end{pmatrix}$$
$$W(X \rightarrow Y) \equiv \begin{pmatrix} W(x_1 \rightarrow x_1) & W(x_2 \rightarrow x_1) & \dots & W(x_N \rightarrow x_1) \\ W(x_1 \rightarrow x_2) & W(x_2 \rightarrow x_2) & \dots & W(x_N \rightarrow x_2) \\ \dots & \dots & \dots & \dots \\ W(x_1 \rightarrow x_N) & W(x_2 \rightarrow x_N) & \dots & W(x_N \rightarrow x_N) \end{pmatrix}$$



## Notació matricial (II)

$$\begin{pmatrix} p_I(x_1) \\ p_I(x_2) \\ \dots \\ p_I(x_N) \end{pmatrix} = \begin{pmatrix} W(x_1 \rightarrow x_1) & W(x_2 \rightarrow x_1) & \dots & W(x_N \rightarrow x_1) \\ W(x_1 \rightarrow x_2) & W(x_2 \rightarrow x_2) & \dots & W(x_N \rightarrow x_2) \\ \dots & \dots & \dots & \dots \\ W(x_1 \rightarrow x_N) & W(x_2 \rightarrow x_N) & \dots & W(x_N \rightarrow x_N) \end{pmatrix} \begin{pmatrix} p_{I-1}(x_1) \\ p_{I-1}(x_2) \\ \dots \\ p_{I-1}(x_N) \end{pmatrix}$$

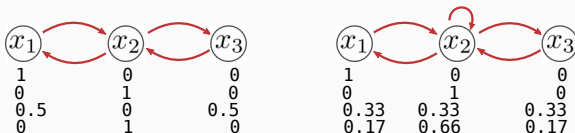
$$W_k(X \rightarrow Y) = \begin{pmatrix} W(x_1 \rightarrow x_1) & W(x_2 \rightarrow x_1) & \dots & W(x_N \rightarrow x_1) \\ W(x_1 \rightarrow x_2) & W(x_2 \rightarrow x_2) & \dots & W(x_N \rightarrow x_2) \\ \dots & \dots & \dots & \dots \\ W(x_1 \rightarrow x_N) & W(x_2 \rightarrow x_N) & \dots & W(x_N \rightarrow x_N) \end{pmatrix}^k$$

## Definicions i propietats

- Una C.M. és irreduïble si de qualsevol estat  $X$  es pot aconseguir arribar a qualsevol altre estat  $Y$  (el graf no està "trecat" ni té regions *inaccessibles*):

$$\forall X, Y \in \Omega \exists k ; W_k(X \rightarrow Y) > 0$$

- Donada una cadena de Markov, anomenem període  $d_X$  d'un estat  $X$  al m.c.d. de tots els  $m > 0$  pels que  $W_m(X \rightarrow X) > 0$ .



- Un estat  $X$  s'anomena aperiòdic si  $d_X = 1$ .  
(E.g. trivial, si  $W(X \rightarrow X) > 0$ ).

- **Teorema 1:** En una cadena irreduïble tots els estats de la cadena tenen el mateix període (que s'anomena període de la cadena).
- **Definició:** cadena aperiòdica  $\forall X, d_X = 1$
- **Teorema 2:** Una C.M. és irreduïble i aperiòdica si i només si

$$\forall X, Y \in \Omega \exists \text{ pas mínim } m; \forall k > m; W_k(X \rightarrow Y) > 0$$

- **Definició:** una distribució  $\Pi(X)$  és estacionària per a una C.M. quan

$$p(Y) \equiv \sum_X W(X \rightarrow Y) \Pi(X) = \Pi(Y)$$

En el cas que  $\Omega$  sigui finit això correspon a dir que  $\Pi$  és vector propi de  $W$  amb valor propi 1.

- **Teorema 3:** En una C.M. irreduïble i aperiòdica sempre existeix  $\Pi(Y)$  estacionària tal que  $\Pi(Y)$  és estrictament positiva i única, i compleix:

$$p(X) \lim_{k \rightarrow \infty} W_k(X \rightarrow Y) = \Pi(Y) \quad \forall p(X)$$

**TOTES** les configuracions convergeixen a  $\Pi(Y)$ .

## Exemple 1, exercici

- a)** Simuleu l'evolució de la C.M. següent i estudeu estadísticament l'evolució de les  $p_I$ .
- b)** Resoleu l'evolució de la C.M. analíticament i compareu el resultat.

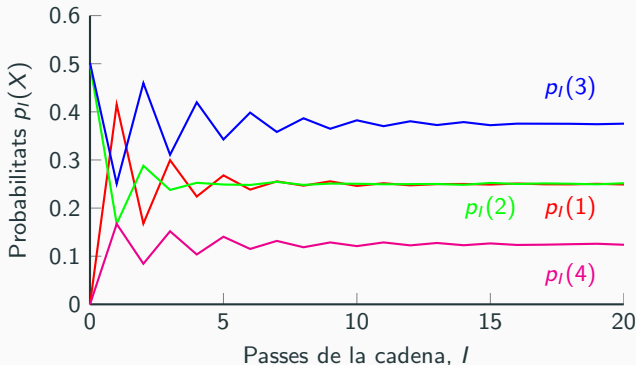
$$p_0 \equiv \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$$

$$W(X \rightarrow Y) \equiv \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \\ 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

## Exemple 1: (a) simulació MC-MC

Fent 100000 simulacions de la cadena (amb 100000 llavors diferents del generador de nombres aleatoris), podem estimar les probabilitats de cada un dels 4 estats a cada passa  $l$ , mesurant les freqüències amb que es visita cada estat.

(codi Markov.f)



## Exemple 1: (b) càlcul numèric i analític de les $p_l(X)$

Numeric: Si multipliquem el vector  $p_0$  repetides vegades per  $W$ , obtenim:

$$p_0 \equiv \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, p_1 \equiv \begin{pmatrix} \frac{5}{12} \\ \frac{1}{6} \\ \frac{1}{4} \\ \frac{1}{6} \end{pmatrix}, p_2 \equiv \begin{pmatrix} \frac{1}{6} \\ \frac{7}{24} \\ \frac{11}{24} \\ \frac{1}{12} \end{pmatrix}, p_3 \equiv \begin{pmatrix} \frac{43}{144} \\ \frac{17}{72} \\ \frac{5}{16} \\ \frac{11}{72} \end{pmatrix}, \dots, p_{20} \equiv \begin{pmatrix} 0.249922 \\ 0.249923 \\ 0.375286 \\ 0.124869 \end{pmatrix}$$

Analític: Si calculeu els vectors propis de la matriu  $W$  trobareu un vector propi de valor propi 1

$$\Pi = A \begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix},$$

Si el normalitzeu  $A$  tal que  $\sum \Pi = 1$ , tenim:

$$\Pi = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{3}{8} \\ \frac{1}{8} \end{pmatrix},$$

## Exemple 2, exercici

Modelitzeu el comportament d'un borratxo a casa seva, suposant que a cada passa de temps canvia a l'atzar d'habitació. Proposeu canvis en les portes o en la dinàmica per aconseguir que sigui aperiòdica.

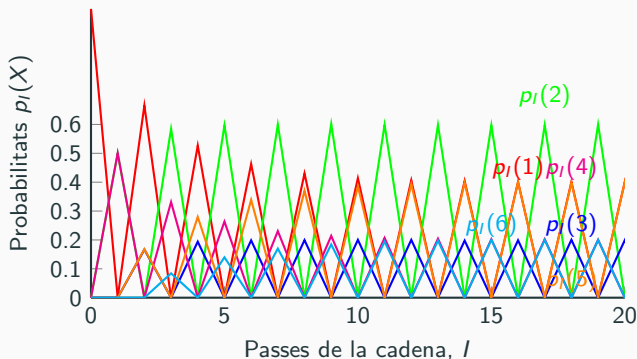
|   |   |   |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |

$$p_0 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad W(X \rightarrow Y) \equiv \begin{pmatrix} 0 & \frac{1}{3} & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

## Exemple 2: simulació

Fent 1000000 simulacions de la cadena (amb 100000 llavors diferents del generador de nombres aleatoris), podem estimar les probabilitats de cada un dels 6 estats a cada passa  $l$ , mesurant les freqüències amb que es visita cada estat.

(codi Markov.f)

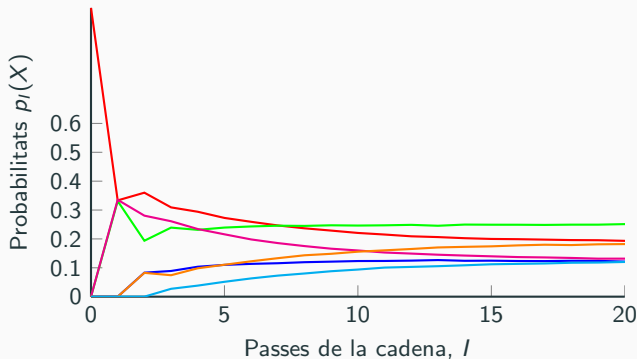


No convergeix!. A les passes parell tenim unes  $p_l(X)$  i a les senars unes altres.



## Exemple 2: simulació variant

Si permeteu que el borratxo pugui triar entre les opcions canviar d'habitació i quedar-se, es trenca la periodicitat i aleshores la cadena sí que convergeix.



Podem escriure la variació de probabilitats a cada transició:

$$\begin{aligned}\Delta p &= p_I(X) - p_{I-1}(X) = \sum_Y W(Y \rightarrow X)p_{I-1}(Y) - 1p_{I-1}(X) = \\ &= \sum_Y W(Y \rightarrow X)p_{I-1}(Y) - \sum_Y W(X \rightarrow Y)p_{I-1}(X)\end{aligned}$$

Si la cadena és estacionària (no varia en el temps),  $p_I(X) - p_{I-1}(X) = 0$  i, per tant cal demanar que 'el  $\Delta p(\rightarrow X)$  que entra a cada  $X$  és el mateix que el  $\Delta p(X \rightarrow)$  que en surt', el que anomenem condició de balanç global:

$$\boxed{\sum_Y W(Y \rightarrow X)p_{I-1}(Y) = \sum_Y W(X \rightarrow Y)p_{I-1}(X)}$$

(Nota: F.E. fora d'equilibri: Associant MC-MC a una evolució temporal, aquesta condició s'escriu, normalment dividida per  $\Delta t$  i s'enten com una integració numèrica d'una equació per  $dp/dt$ .)

Per a fer una simulació de MC en cadena de Markov ens cal trobar, en el nostre espai de les fases  $\Omega$ , una probabilitat de transició  $W(X \rightarrow Y)$  que compleixi:

- Sigui irreduïble i aperiòdica (permeti anar de qualsevol estat a qualsevol altre amb un nombre finit de passes). Això garanteix que existirà una probabilitat estacionària  $\Pi(X)$ .
- La probabilitat estacionària sigui  $\Pi(X) = f(X)/Z$

Aleshores fent realitzacions de la cadena (simulacions), després d'un nombre prou gran de passes, les configuracions  $X$  que obtindrem es distribuïran com  $f(X)/Z$  i podrem calcular els promitjos de la F.E..

*(Nota: Tant MC-MC com Metropolis en particular són mètodes d'integració MC per importance sampling. Com a tals es poden emprar en altres problemes d'integració.)*

Com ha de ser  $W(X \rightarrow Y)$  per tal que la probabilitat estacionaria sigui  $\Pi(X) = f(X)/\mathcal{Z}$ ?

Condicció de balanç global:

$$\sum_Y W(Y \rightarrow X) \frac{f(Y)}{\mathcal{Z}} = \sum_Y W(X \rightarrow Y) \frac{f(X)}{\mathcal{Z}}$$

La  $\mathcal{Z}$  desapareix de l'expressió!!!. No cal conèixer  $\mathcal{Z}$  per dissenyar una  $W$  adequada per als nostres propòsits.

A la pràctica s'imposa una condició més restrictiva (*condició suficient de balanç global*): el balanç a l'intercanvi entre cada parella d'estats  $X, Y$ , anomenada balanç detallat:

$$W(Y \rightarrow X)f(Y) = W(X \rightarrow Y)f(X)$$

Centrem-nos en el cas del factor de Boltzmann:

$$f(X) = e^{-\beta \mathcal{H}(X)}$$

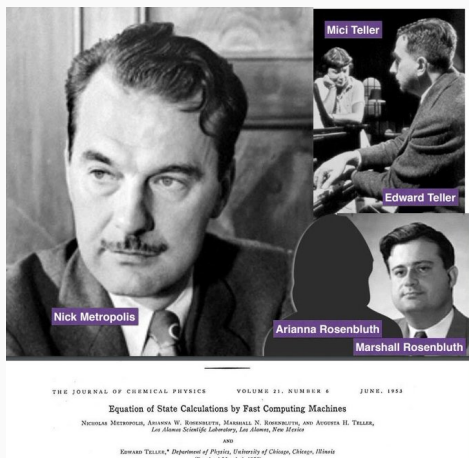
$$W(Y \rightarrow X)e^{-\beta \mathcal{H}(Y)} = W(X \rightarrow Y)e^{-\beta \mathcal{H}(X)}$$

$$\frac{W(Y \rightarrow X)}{W(X \rightarrow Y)} = \frac{e^{-\beta \mathcal{H}(X)}}{e^{-\beta \mathcal{H}(Y)}} = e^{-\beta [\mathcal{H}(X) - \mathcal{H}(Y)]}$$

Aquesta equació té moltes solucions: cadascuna es correspon amb un algorisme de simulació diferent

- **Algorisme de Metropolis:** proposta i rebuig
- Algorisme de heat bath
- ...

# Algorithme de Metropolis-Hastings ( A. W. Rosenbluth )



Arianna W. Rosenbluth  
(1927 - 2020)



“She actually did all the coding, which at that time was a new art for these new machines,”

## Algorisme de Metropolis (I)

La transició  $X \rightarrow Y$  a cada pas es decideix en dues passes

- Es proposa un possible canvi de configuració  $X \rightarrow Y$  d'acord amb una probabilitat de proposta simètrica

$$T(X \rightarrow Y) = T(Y \rightarrow X)$$

Es requereix que  $T$  tingui bones propietats per tal de garantir que la cadena sigui irreduïble i aperiòdica.: bàsicament que les propostes siguin canvis molt elementals.

- El canvi s'accepta o es rebutja d'acord amb una probabilitat d'acceptació  $Q(X \rightarrow Y)$

Per tant:

$$\begin{aligned} W(X \rightarrow Y) &= \\ &= \begin{cases} T(X \rightarrow Y)Q(X \rightarrow Y) & \text{if } X \neq Y \\ \cancel{T(X \rightarrow X)Q(X \rightarrow X)} + \sum_{Y \neq X} T(X \rightarrow Y) [1 - Q(X \rightarrow Y)] & \text{if } X = Y \end{cases} \end{aligned}$$

## Algorisme de Metropolis (II)

Donada la simetria de  $T$ , el balanç detallat només afecta a  $Q$

$$\frac{W(Y \rightarrow X)}{W(X \rightarrow Y)} = \frac{\cancel{T(Y \rightarrow X)}}{\cancel{T(X \rightarrow Y)}} \frac{Q(Y \rightarrow X)}{Q(X \rightarrow Y)} = \frac{e^{-\beta \mathcal{H}(X)}}{e^{-\beta \mathcal{H}(Y)}} = e^{-\beta[\mathcal{H}(X) - \mathcal{H}(Y)]}$$

Per tant hem de dissenyar una  $Q$  que compleixi

$$\frac{Q(Y \rightarrow X)}{Q(X \rightarrow Y)} = e^{-\beta[\mathcal{H}(X) - \mathcal{H}(Y)]}$$

Si aquesta condició es compleix per  $Q(X \rightarrow Y)$  quan  $X \neq Y$ , es trivial demostrar que també es satisfarà balanç detallat per  $W(X \rightarrow X)$ .

Però  $\mathcal{H}(X) - \mathcal{H}(Y)$  en general no és trivial!

Tampoc ens cal una CM amb transicions connectant "gaire" estats de  $\Omega$  a cada pas.

Quines transicions proposem?

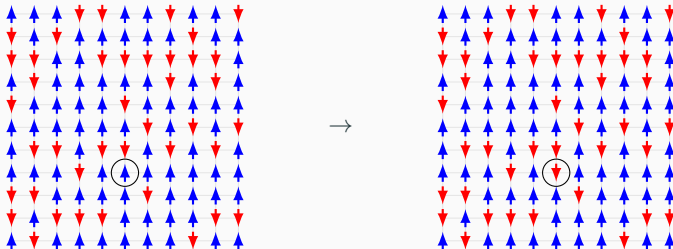


## Probabilitat de proposta de Metropolis o Glauber

Per al cas del model d'Ising, en una certa configuració  $X = \{S_k\}$ , proposarem invertir un spí a l'atzar:  $\mathcal{H}(Y) - \mathcal{H}(X) = 2S_{k,l} \left( \sum_j^{n.n.} S_{j,l} \right)$

Noteu que en aquesta definició de  $T(X \rightarrow Y)$ :

- $X$  i  $Y$  difereixen només en el valor d'un spí.
- Des d'una configuració només es pot proposar anar a  $N$  configuracions diferents (canviar un dels  $N$  spins). La resta de configuracions no són accessibles en un pas
- $T(X \rightarrow Y) = T(Y \rightarrow X)$



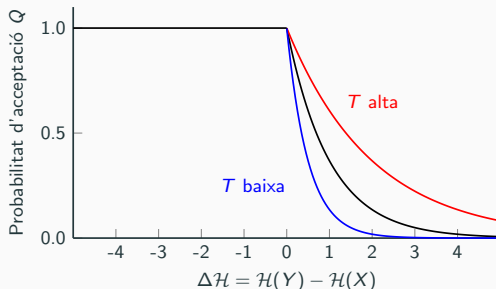
Per explorar tot  $\Omega$  calen, almenys,  $N$  canvis  $\equiv$  un pas Monte-Carlo.

# Probabilitat d'acceptació de Metropolis

Metropolis va proposar fer servir

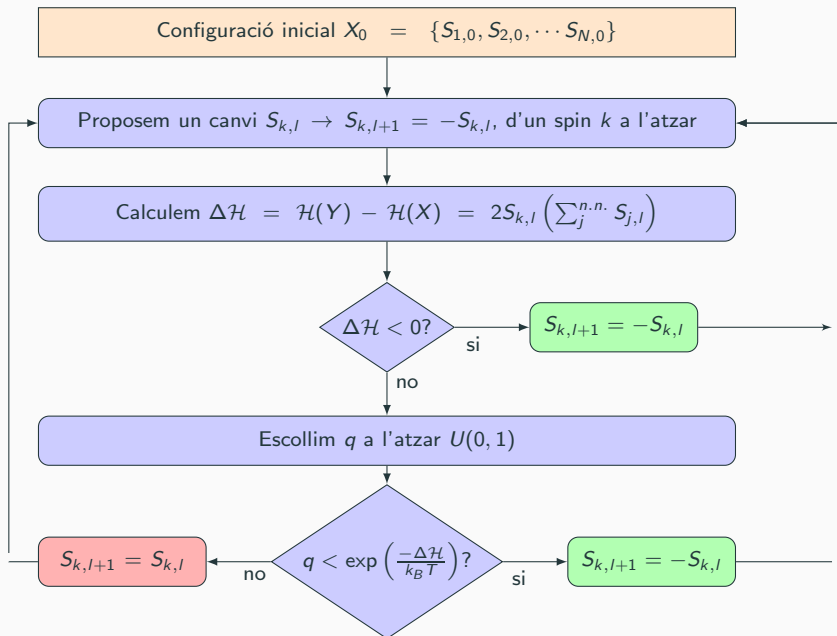
$$Q(X \rightarrow Y) = \min \left\{ 1, e^{-\beta(\mathcal{H}(Y) - \mathcal{H}(X))} \right\}$$

Podeu comprovar que compleix balanç detallat, tractant per separat els casos  $\mathcal{H}(X) > \mathcal{H}(Y)$  i  $\mathcal{H}(Y) > \mathcal{H}(X)$



- Els canvis que fan disminuir l'energia s'accepten sempre.
- Els que la fan augmentar, s'accepten amb una probabilitat que decau amb  $\Delta\mathcal{H}$  exponencialment.

## Esquema de l'algorisme de Metropolis



- P1: Configuracions inicials. Mesura de magnetització i energia.
- P2: Condicions de contorn. Implementar Metropolis. Termalització. Inspecció en funció de  $T$ .
- P3: Mesures promitjades. Susceptibilitat i capacitat calorífica. Dependències en  $L$  i  $T$ . Efectes de mida finita.
- A casa: Simulacions. Extreure resultats. Resoldre qüestions.

- N.Metropolis, A.Rosenbluth, M.Rosenbluth,A.Teller & E.Teller,  
*Equation of State Calculations by Fast Computing Machines*  
J. Chem. Phys.**21**, 1087 (1953).  
<http://dx.doi.org/10.1063/1.1699114>
- B.J.Alder & T.E.Wainwright,  
*Phase Transition for a Hard Sphere System*  
J.Chem. Phys. **27**, 1208 (1957).  
<https://doi.org/10.1063/1.1743957>

- M.Plischke & B.Bergensen,  
*Equilibrium Statistical Physics*,  
World Scientific 3rd edition (2006).
- J.J.Binney, N.J.Dowrick, A.J.Fisher & M.E.J.Newman,  
*The theory of critical phenomena*,  
Oxford Science Publications (1992).
- M.LeBellac, F.Mortessagne, & G.G.Batrouni,  
*Equilibrium and non-equilibrium statistical thermodynamics*,  
Cambridge University Press (2004).

## **Annex: Repàs mètodes d'integració MC**

---

En el marc general dels mètodes numèrics s'anomena mètode de MC a qualsevol mètode que faci servir números a l'atzar per resoldre un problema matemàtic consistent en calcular una quantitat  $I$ . Tots aquets mètodes es basen en dissenyar un algorisme que, a partir d'una mostra de  $M$  nombres a l'atzar (distribuïts d'acord amb alguna densitat de probabilitat  $p(X)$ ) permeti trobar un estimador  $\hat{I}_M$  de la quantitat  $I$ .

Considerem  $X$ , v.a. i sigui  $\{x_1, x_2, x_3, \dots, x_M\}$  una mostra d'aquesta variable.

$$\hat{I}_M = f(\{x_1, x_2, x_3, \dots, x_M\}) \rightarrow I$$

En ésser un mètode estadístic, aquesta convergència és molt feble. Els errors disminueixen molt lentament amb la mida de la mostra

$$\varepsilon(\hat{I}_M) \propto \frac{1}{\sqrt{M}}$$

Són mètodes molt pitjors que la majoria de mètodes numèrics. Només es fan servir quan no tenim res millor:

- Integració multidimensional



## Monte carlo cru (I)

Volem resoldre una integral (que pot ser d-dimensional)

$$I = \frac{1}{b-a} \int_a^b f(x) dx$$

Generem una mostra de  $M$  nombres a l'atzar  $\{x_1, x_2, \dots, x_M\}$ , independents i amb densitat uniforme  $U(a, b)$ . Proposem el següent estimador de  $I$ :

$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M f(x_i)$$

Noteu que  $\hat{I}_M$  és, de fet, una v.a., suma de moltes variables aleatòries. Podem intentar trobar-ne algunes propietats:

- Esperança

$$\langle \hat{I}_M \rangle = \frac{1}{M} \langle \sum_{i=1}^M f(x_i) \rangle = \frac{1}{M} \sum_{i=1}^M \langle f(x_i) \rangle = \langle f(x_i) \rangle = \int_a^b f(x) \frac{1}{b-a} dx = I$$

Diem que l'estimador és no-esbiaixat, ja que la seva esperança coincideix amb el valor exacte de la quantitat que volem estimar  $\langle \hat{I}_M \rangle = I$ .

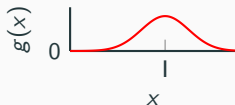
- Variància:

$$\begin{aligned} V(\hat{I}_M) &= \frac{1}{M^2} V\left(\sum_{i=1}^M f(x_i)\right) = \frac{1}{M^2} \sum_{i=1}^M V(f(x_i)) = \frac{1}{M} V(f(x_i)) = \\ &= \frac{1}{M} \left[ \int_a^b f^2(x) \frac{1}{b-a} dx - \left( \int_a^b f(x) \frac{1}{b-a} dx \right)^2 \right] \end{aligned}$$

Diem que l'eficiència de l'estimador es proporcional a  $1/M$ .

Fixeu-vos que les desviacions o error de l'estimador son proporcionals a  $\sqrt{V(\hat{I}_M)}$ , i per tant decauen com  $1/\sqrt{M}$ .

- Aplicant el TCL, quan  $M \gg 1$ , la variable aleatòria  $\hat{I}_M$  es distribuirà segons una densitat gaussiana  $N(I, \sqrt{\frac{V(f)}{M}})$



Com més gran és  $M$ , més estreta és la gaussiana. Quan  $M \rightarrow \infty$ , la distribució esdevindrà una  $\delta$  de Dirac centrada a  $I$ .

$$\lim_{N \rightarrow \infty} \hat{I}_M = I$$

Diem que l'estimador  $\hat{I}_M$  és consistent. En el límit  $\hat{I}_M$  deixa de ser una v.a.

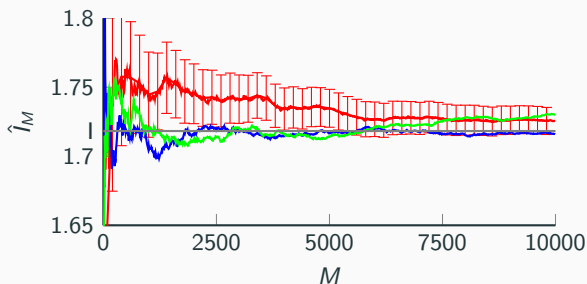
## Exemple: MC cru

Calcular amb el mètode de MC cru la integral

$$I = \int_0^1 e^x dx = e^1 - e^0 = e - 1 = 1.71828182$$

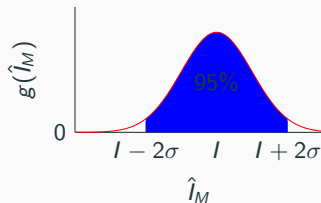
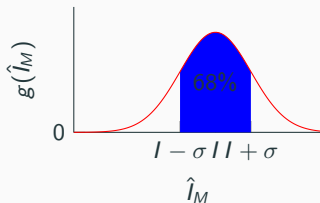
Generen seqüències de  $M$  nombres a l'atzar  $U(0, 1)$   $\{x_1, x_2, \dots, x_M\}$  i calculem:

$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M e^{x_i}$$



## Error estadístic en un estimador MC

Per  $M$  prou gran,  $\hat{I}_M$  es comporta com una v.a. gaussiana  $N(I, \sigma = \sqrt{\frac{V(f)}{M}})$



Per tant, com a resultat del nostre càlcul podem donar el millor valor i una barra d'error amb el 68% de confiança o el 95% de confiança:

$$I = \hat{I}_M \pm \sqrt{\frac{V(f)}{M}} \quad I = \hat{I}_M \pm 2\sqrt{\frac{V(f)}{M}}$$

on

$$V(f) = \left[ \frac{1}{b-a} \int_a^b f^2(x) dx - \left( \frac{1}{b-a} \int_a^b f(x) dx \right)^2 \right]$$

Per trobar l'error, doncs, ens cal estimar  $V(f)$ .

## Estimadors estadístics de la variància

Hi ha dos possibles estimadors de la variància:

- La desviació típica, que és un estimador esbiaixat

$$\hat{s}_M^2 = \left[ \frac{1}{M} \sum_{i=1}^M f^2(x_i) - \left( \frac{1}{M} \sum_{i=1}^M f(x_i) \right)^2 \right]$$

- La desviació típica amb correcció de Bessel, que és no esbiaixat:

$$\hat{\sigma}_M^2 = \frac{M}{M-1} \left[ \frac{1}{M} \sum_{i=1}^M f^2(x_i) - \left( \frac{1}{M} \sum_{i=1}^M f(x_i) \right)^2 \right]$$

En resum, quan fem una estimació MC, a mesura que generem els  $x_i$ , a l'atzar, cal calcular els estimadors de moments

$$E(x) = \frac{1}{M} \sum_{i=1}^M f(x_i) \quad \text{i} \quad E(x^2) = \frac{1}{M} \sum_{i=1}^M f^2(x_i)$$

I, donar com a resultat

$$I = E(x) \pm \sqrt{\frac{E(x^2) - (E(x))^2}{M-1}}$$

La comparació entre l'error associat a un dels mètodes d'integració usuals i el mètode de MC posa de manifest que per integrals multidimensionals el mètode MC serà millor

| Mètode   | error                                       | Augment computacional necessari per reduir l'error un factor 1/2                           |
|----------|---|--|
| MC       | $\varepsilon(N) \propto \frac{1}{\sqrt{N}}$ | $\varepsilon' = \varepsilon/2 \Rightarrow N' = 4N$   |
| Trapezis | $\varepsilon(h) \propto h^2$                | $\varepsilon' = \varepsilon/2 \Rightarrow h' = h/\sqrt{2} \Rightarrow N' = (\sqrt{2})^d N$ |

Per  $d > 4$  el mètode MC començarà a ésser competitiu !

## Reducció de la variància: importance sampling (I)

L'error associat a l'estimador  $\hat{I}_M$  és  $\propto \sqrt{\frac{V(f)}{M}}$ . Com poder aconseguir que  $V(f)$  sigui petita?

El mètode del mostreig d'acord amb la importància intenta minimitzar  $V(f)$ .

Suposem que volem estimar

$$I = \int_a^b f(x) dx$$

Escollim una funció  $g(x)$  tal que  $g(x) > 0 \forall x$  i que compleixi  $\int_a^b g(x) dx = 1$ .  
Aleshores

$$I = \int_a^b f(x) dx = \int_a^b \frac{f(x)}{g(x)} g(x) dx = \int_a^b h(x) g(x) dx$$

Considerem ara una mostra  $\{x_i\}$  amb  $i = 1 \dots M$  de v.a. independents i idènticament distribuïdes segons la densitat  $g(x)$ . Estudiem les propietats de l'estimador

$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M \frac{f(x_i)}{g(x_i)}$$

## Reducció de la variància: importance sampling (II)

Busquem les propietats de l'estimador  $\hat{I}_M$ :

- Esperança

$$\langle \hat{I}_M \rangle = \left\langle \frac{1}{M} \sum_{i=1}^M \frac{f(x_i)}{g(x_i)} \right\rangle = \frac{1}{M} \sum_{i=1}^M \left\langle \frac{f(x_i)}{g(x_i)} \right\rangle = \left\langle \frac{f(x)}{g(x)} \right\rangle = \int_a^b \frac{f(x)}{g(x)} g(x) dx = I$$

Per tant, és no-esbiaixat.

- Variança

$$V(\hat{I}_M) = \frac{1}{M^2} MV\left(\frac{f(x_i)}{g(x_i)}\right) = \frac{1}{M} V\left(\frac{f(x)}{g(x)}\right)$$

Donat que tenim molta llibertat per escollir  $g(x)$ , si escollim una  $g(x)$  tal que  $\frac{f(x)}{g(x)} \sim K$ , aconseguirem una variància molt petita. El problema rau en que cal saber generar nombres aleatoris d'acord amb  $g(x)$ .

(A F.C. vàreu estudiar mètodes de generació de nombres aleatoris)



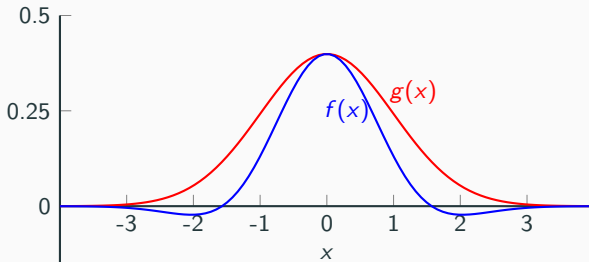
## Exemple: mètode importance sampling(I)

Calcular amb el mètode importance sampling la integral

$$I = \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cos(x) e^{-\frac{x^2}{2}} dx$$

Noteu que si multipliquem i dividim per  $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  (que és una  $N(0, 1)$ ), tenim

$$I = \int_{-\infty}^{\infty} \cos(x) g(x) dx \Rightarrow \hat{I}_N = \frac{1}{M} \sum_{i=1}^M \cos(x_i) \text{ amb } \{x_i\} \text{ v.a } N(0, 1)$$



## Exemple: mètode importance sampling (II)

Generen seqüències de  $M$  nombres a l'atzar  $N(0, 1)$   $\{x_1, x_2, \dots, x_M\}$  i calculem:

$$\hat{I}_M = \sum_{i=1}^M \cos x_i$$

