Thema's uit de Mathematische Fysica

Deel 1: Fluctuatietheorie - Large Deviations (Maes)

1. Inleiding

a - Macroscopische gemiddelden

- Bvb: N keer een munt opgooien. „Fractie kop" $\to \frac{1}{2}$ als $N \uparrow \infty$ voor een goed muntstuk.
  Als $p_{kop} \neq \frac{1}{2}$ zijn er fluctuaties
- Maxwellverdeling vind je typisch voor snelheden van deeltjes in een doos. Als er een niet-Maxwelliaanse verdeling gevonden wordt, zijn er fluctuaties.

b - Tijdsgemiddelden

Stel X_t een irreducibel Markovproces in een eindige toestandsruimte K met een unieke stationaire verdeling ρ. ρ geeft dan een kansmaat op K weer.
Voor $x \in K$ zal $\underbrace{\frac{1}{\tau}\int_0^\tau dt\, \delta_{X_t,x}}_{\to \rho}$ de fractie tijd weergeven dat $X_t = x$, met $\tau \to \infty$.

Analoog voor ergodische systemen, met het Birkhoff-Khinchin Theorem.

c - Klassieke Limiet

$\hbar \to 0$, $c \to \infty$.

Het onderwerp van dit deel is het berekenen, bekijken, bespreken van de plausibiliteit van afwijkingen ten opzichte van normaal gedrag. Een motivatie om dit te onderzoeken is te vinden in het variatieprincipe. We zullen aantonen dat voor een dichtheid μ geldt $\text{Prob}\left[\frac{1}{\tau}\int_0^\tau dt\, \delta_{X_t,x} = \mu(x),\ \forall x \in K\right] \simeq e^{-\tau D(\mu)}$
met D een functionaal op kansruimte K.

2. Independent and Identically Distributed Random Variables (IID RV)

Neem $X_1, X_2, \dots, X_n$ IID RV's en stel dat $\langle |X_i| \rangle < \infty$

i tje is arbitrair want alle i's zijn gelijk verdeeld.

a - Sterke wet van de grote getallen.

Voor $N \to \infty$ zal $\underbrace{\frac{1}{N}(X_1 + \dots + X_n)}_{:= m_N \text{ (arithmetic mean)}} \longrightarrow \underbrace{\langle X_i \rangle}_{:= \mu}$ Dit gebeurt met kans 1.

Er is ook een zwakke wet:

Stel dat $Var(X_i)$ eindig is, dan is $Var\left(\frac{1}{n} \sum_i X_i\right) \to 0$ voor $n \to \infty$.

Bewijs: Stel dat $Var(X_i) = \sigma^2$ (voor alle i), dan is

$$Var\left(\frac{1}{n}\sum_i X_i\right) = \frac{1}{n^2} Var\left(\sum X_i\right) = \frac{\sum Var(X_i)}{n^2} = \frac{n\sigma^2}{n^2} \to 0 \text{ voor } n \uparrow \infty.$$

en zo zal $\bar{X}_n \to \langle X_i \rangle$

De wet van de grote getallen impliceert ook een universaliteit die enkel afhangt van het eerste moment.

b - Centrale limietstelling

Er zijn een aantal correcties gemaakt op de sterke wet, zoals de centrale limietstelling. Deze centrale limietstelling zegt dat:

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (X_i - \langle X_i \rangle) \longrightarrow N(0,1) \quad (\text{als } Var\, X_i = 1).$$

$$= \sqrt{N}\left(\frac{1}{N}\sum_1^N X_i - \frac{1}{N}\underbrace{\sum_1^N \langle X_i \rangle}_{N \cdot \mu}\right)$$

$$= \sqrt{N}\left(\frac{1}{N}\sum_1^N X_i - \mu\right) \longrightarrow N(0,1)$$

m.a.w., in termen van de verdeling: ($a \leq b$ en $N \uparrow \infty$)

$$Prob\left[a \leq \sqrt{N}\, \underset{\substack{\uparrow \\ \mu = 0}}{m_N} \leq b\right] = Prob\left[\frac{a}{\sqrt{N}} \leq m_N \leq \frac{b}{\sqrt{N}}\right] \longrightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}$$

De centrale limietstelling geeft een kleine afwijkingen rond het gemiddelde. De theorie van de grote afwijkingen moet in geval van kleine afwijkingen terug de CLS kunnen reproduceren.

c- Grote afwijkingen

We bekijken een voorbeeld. $X_i = 0$ of 1, elk met kans $\frac{1}{2}$. Wat is dan

$0 \leq a, b \leq 1$

$P_N(a,b) := \text{Prob}\left[a < \underbrace{\frac{X_1 + X_2 + \dots + X_N}{N}}_{=: m_N} < b \right]$? We tonen hieronder aan dat

$$\lim_{N \to \infty} \frac{1}{N} \log P_N(a,b) = -\inf_{a \leq x \leq b} \Big\{ \log 2 - \underbrace{\left(-x \log x - (1-x) \log(1-x)\right)}_{=: s(x)} \Big\} \quad \text{met } x \in [0,1]$$

Om dit te bewijzen maken we gebruik van de Stirlingbenadering $\ln(N!) = N \ln N - N$, wat eenvoudig te bewijzen is:

$$\ln(N!) = \sum_{k=1}^{N} \ln k \underset{N \to \infty}{\simeq} \int_1^N \ln k \, dk = k \ln k - k \Big|_{k=1}^{k=N} = N \ln N - N$$

Vanuit de kanstheorie weten we dat $\text{Prob}_N[X = k] = \binom{N}{k} \frac{1}{2}^k \frac{1}{2}^{N-k} = 2^{-N} \binom{N}{k}$ en

$$\text{Prob}[X \leq x, N] = \sum_{i=1}^{\lfloor x \rfloor} \binom{N}{i} \frac{1}{2}^i \cdot \frac{1}{2}^{N-i} = 2^{-N} \sum_{i=1}^{\lfloor x \rfloor} \binom{N}{i} = F_X(x).$$

Dan is

$$\text{Prob}[a < X < b, N] = F_X(b) - F_X(a) = 2^{-N} \sum_{i=1}^{\lfloor b \rfloor} \binom{N}{i} - 2^{-N} \sum_{i=1}^{\lfloor a \rfloor} \binom{N}{i}$$
$$= 2^{-N} \sum_{\lfloor a \rfloor < i \leq \lfloor b \rfloor} \binom{N}{i}$$

De kans $P_N(a,b)$ is dus exact uit te rekenen als $P_N(a,b) = 2^{-N} \sum_{aN < j < bN} \binom{N}{j}$, maar dat is nogal veel werk. Definieer de maximale waarde uit deze som als $q_N(a,b) := \max_{aN < j < bN} \frac{N!}{(N-j)!\,j!}$. Er zal steeds gelden dat $2^{-N} q_N \leq P_N$, met gelijkheid in geval van een, som met één term. Ook zal steeds $P_N \leq (b-a) N q_N$ met gelijkheid in hetzelfde geval.

We vinden zo een onder- en bovengrens op P_N:

$2^{-N} q_N(a,b) \le P_N(a,b) \le 2^{-N}(b-a)N\, q_N(a,b)$. Omdat we op zoek zijn naar $\lim\limits_{N\to\infty} \frac{1}{N} \log P_N(a,b)$ vermenigvuldigen we de hele ongelijkheid met $\frac{1}{N}$ nadat we de logaritme genomen hebben. Omwille van het strikt stijgend-zijn van de log blijft de ongelijkheid gelden:

$$\frac{1}{N}\log\left(2^{-N} q_N\right) \le \frac{1}{N}\log P_N \le \frac{1}{N}\log\left(2^{-N}(b-a)N\, q_N\right)$$

$$\Rightarrow \frac{-N}{N}\log 2 + \frac{1}{N}\log q_N \le \frac{1}{N}\log P_N \le \frac{-N}{N}\log 2 + \frac{1}{N}\log\left((b-a)N\right) + \frac{1}{N}\log q_N$$

$$\Rightarrow -\log 2 + \frac{1}{N}\log q_N \le \frac{1}{N}\log P_N \le -\log 2 + \frac{1}{N}\log\left((b-a)N\right) + \frac{1}{N}\log q_N$$

$$N \to \infty \Rightarrow \frac{1}{N}\log\left((b-a)N\right) \to 0$$

$$\Rightarrow -\log 2 + \frac{1}{N}\log q_N \le \frac{1}{N}\log P_N \le -\log 2 + \frac{1}{N}\log q_N$$

Dus $\lim\limits_{N\to\infty} \frac{1}{N}\log P_N(a,b) = -\log 2 + \frac{1}{N}\log q_N(a,b)$

Nu berekenen we $\log q_N(a,b) = \log\left(\max\limits_{\dots}\left\{\frac{N!}{(N-j)!\,j!}\right\}\right)$

omwille van de strikt stijgendheid van log

$$= \max_{aN<j<bN}\left\{\log\left(\frac{N!}{(N-j)!\,j!}\right)\right\}$$

$$= \max\left\{(N\log N - \cancel{N}) - \left((N-j)\log(N-j) - \cancel{(N-j)}\right) - (j\log j - \cancel{j})\right\}$$

$$= \max\left\{N\log\left(\frac{N}{N-j}\right) + j\log\left(\frac{N-j}{j}\right)\right\}$$

$$= N\max\left\{\log\left(\frac{j}{N}\cdot\frac{N/j}{1-j/N}\right) + j/N\,\log\left(\frac{N}{j}\cdot\frac{1-j/N}{1}\right)\right\}$$

$$= N\max\left\{\underbrace{\log j/N + \log N/j}_{=0} - \log(1-j/N) + j/N\,\log(N/j) + j/N\,\log(1-j/N)\right\}$$

$$= N\max_{aN<j<bN}\left\{-j/N\,\log j/N - (1-j/N)\log(1-j/N)\right\}$$

$\begin{cases} aN<j<bN \Rightarrow a < j/N < b \Rightarrow \text{stel } m = j/N. \\ \text{omdat } N \nearrow \infty \text{ wordt het supremum van de (continue) verzameling genomen.}\end{cases}$

$$= N\sup_{a<m<b}\underbrace{\left\{-m\log m - (1-m)\log(1-m)\right\}}_{=s(m)} = N\sup_{a<m<b} s(m)$$

Zo wordt $\lim_{N\to\infty} \frac{1}{N}\log P_N(a,b) = -\log 2 + \frac{1}{N}\log q_N(a,b)$

$$= -\log 2 + \frac{1}{N}\cdot N \sup_{a<m<b} s(m) = -\log 2 + \sup_{a<m<b} s(m)$$

$\sup(-f) = -\inf(f)$

$\sup(f) = -\inf(-f)$

$$= -\log 2 - \inf_{a<m<b}\{-s(m)\}$$

$$= -\inf_{a<m<b}\{\log 2 - s(m)\} \qquad (*)$$

We kunnen $P_N(a,b)$ ook direct benaderen via Stirling $\theta! = \sqrt{2\pi\theta}\left(\frac{\theta}{e}\right)^{\theta}$, zodat

$$\frac{N!}{(N-j)!\,j!} \simeq \frac{\sqrt{2\pi N}\,(N/e)^N}{\sqrt{2\pi(N-j)}\left(\frac{N-j}{e}\right)^{N-j}\sqrt{2\pi j}\,(j/e)^j} = \sqrt{\frac{N}{2\pi(N-j)j}}\; N^N j^{-j}(N-j)^{-(N-j)}$$

$\underbrace{e^{-N+j+N-j}}_{=1}$

$$= \sqrt{\frac{N}{2\pi(N-j)j}}\; e^{N\log N - j\log j - (N-j)\log(N-j)}$$

$$= \sqrt{\frac{N}{2\pi(N-j)j}}\; e^{N\left\{\log\left(\frac{N}{N-j}\right) + \frac{j}{N}\log\left(\frac{N-j}{j}\right)\right\}}$$

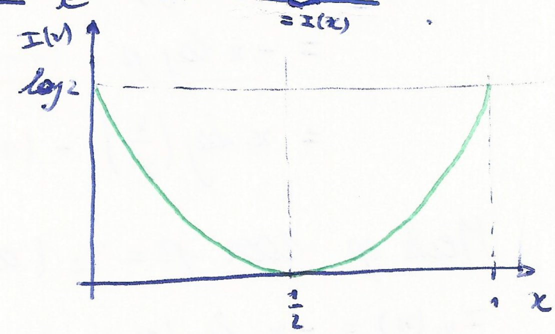
$$= \sqrt{\frac{N}{2\pi(N-j)j}}\; e^{N\left\{\underbrace{\log(j/N)+\log(N/j)}_{0} - \log(1-j/N) + j/N\log(N/j) + j/N\log(1-j/N)\right\}}$$

$$= \sqrt{\frac{N}{2\pi(N-j)j}}\; e^{N\left\{-j/N\log(j/N) - (1-j/N)\log(1-j/N)\right\}}$$

We concluderen (informeel) dat $\text{Prob}[m_N \simeq m] \simeq e^{-NI(m)}$, met $I(m) = \log 2 - s(m)$ de fluctuatiefunctionaal. Merk op dat we ook vanuit (*) kunnen vinden dat $P_N(a,b) \simeq e^{-N\inf_{a<x<b}\{\underbrace{\log 2 - s(x)}_{=I(x)}\}}$.

We besluiten:

$$\text{Prob}\left[\frac{1}{N}(X_1+\ldots+X_N) \simeq x\right] = e^{-NI(x)}.$$



Een volgende natuurlijke vraag is wat er gebeurt bij een oneerlijke munt, bij voorbeeld $X_i = 0$ met kans $1-p$ en $X_i = 1$ met kans p. Dan is de kans om in N worpen k keer „1" te tellen gelijk aan $\text{Prob}[X=k,N] = \binom{N}{k} p^k (1-p)^{N-k}$ en

$\tilde{P}_N(a,b) = \text{Prob}\left[a < \underbrace{\frac{X_1+\dots+X_N}{N}}_{=:m_N} < b\right] = \sum_{aN<j<bN} \binom{N}{j} p^j (1-p)^{N-j}$. De redenering verloopt op exact dezelfde manier, dus $\frac{1}{N}\log \tilde{P}_N(a,b) \longrightarrow \frac{1}{N}\log \tilde{q}_N(a,b)$, met $\tilde{q}_N(a,b) := \max_{aN<j<bN} \binom{N}{j} p^j (1-p)^{N-j}$.

$$\Rightarrow \frac{1}{N}\log\left\{\max_{aN<j<bN}\binom{N}{j} p^j (1-p)^{N-j}\right\}$$

$$= \max\left\{\underbrace{\frac{1}{N}\left(\log\left(\frac{N!}{(N-j)!\,j!}\right)}\ + j\log p + (N-j)\log(1-p)\right)\right\}$$

$$= \max_{a<j/N<b}\left\{s(j/N) + j/N \log p + (1-j/N)\log(1-p)\right\}$$

$$= \sup_{a<m<b}\left\{s(m) + m\log p + (1-m)\log(1-p)\right\}$$

$m = j/N$

$N \to \infty \Rightarrow \max \to \sup.$

$$= -\inf_{a<m<b}\left\{-m\log p - (1-m)\log(1-p) - s(m)\right\}$$

Dan is (ongeveer) $\text{Prob}[m_N \simeq x] = e^{-N I_p(x)}$, met

$$I_p(x) = -x\log p - (1-x)\log(1-p) - s(x)$$

$$= -x\log p - (1-x)\log(1-p) - (-x\log x - (1-x)\log(1-x))$$

$$= x\log\left(\frac{x}{p}\right) - (1-x)\log\left(\frac{1-x}{1-p}\right)$$

Merk op, voor $p = 1/2$ (eerlijk muntstuk) is

$$I_{1/2}(x) = -x\log(1/2) - (1-x)\log(1-1/2) - s(x)$$

$$= x\log 2 - (\log 1/2 - x\log 1/2) - s(x) = \cancel{x\log 2} + \log 2 - \cancel{x\log 2} - s(x)$$

$= \log 2 - s(x)$, zoals gevonden op de achterkant van dit blad!

d- n verschillende waarden (multinomiaal)

We bekijken nu het geval dat de toevalsvariabelen geen 2 (zoals bij het opwerpen van een munt) mogelijkheden heeft, maar n. We nemen een eindige toestandsruimte K, met $|K| = n$.

We hebben N onafhankelijke en identiek verdeelde variabelen met waarden in K en distributie ρ (op de waarden in K).

Noteer N_1 als het aantal keer „1" voorkomt, N_2 …, tot N_n (de relatieve frequentie, zoals we zeggen). Dan is $N_1 + \ldots + N_n = N$, want N i.i.d.'s.

We bekijken $\dfrac{N!}{N_1! \ldots N_n!}$, meer bepaald

$$\lim_{N\to\infty} \frac{1}{N} \log\left(\frac{N!}{N_1! \ldots N_n!}\right) = \lim_{\infty}\left\{\frac{1}{N}\log N! - \frac{1}{N}\log N_1! - \ldots - \frac{1}{N}\log N_n!\right\}$$

$$= \lim_{\infty}\left\{\frac{1}{N}(N\cdot\log N - N) - \frac{1}{N}(N_1 \log N_1 - N_1) - \ldots - \frac{1}{N}(N_n \log N_n - N_n)\right\}$$

$$= \lim_{\infty}\left\{\log N \cancel{- 1} - \sum_{i=1}^{n} \frac{N_i}{N} \log N_i + \underbrace{\sum_{i=1}^{n} \frac{N_i}{N}}_{=1}\right\}$$

$$= \lim_{\infty}\left\{\frac{\overbrace{N_1 + \ldots + N_n}^{=N}}{N} \log N - \sum \frac{N_i}{N} \log N_i\right\} = \lim_{\infty}\left\{-\sum_{i=1}^{n} \frac{N_i}{N} \log \frac{N_i}{N}\right\}$$

$\frac{N_i}{N} \longrightarrow p_i$ voor $N \to \infty$

$$= -\sum_{i=1}^{n} p_i \log p_i$$

Op die manier zal, berekeningen op vorige blz indachtig

$$P(X \simeq x) \simeq e^{-N\sum_i x_i \log x_i} \quad \text{voor } N \to \infty$$

Nu zoeken we naar $\text{Prob}_\rho\left[\frac{n_i}{N} = x_i \mid i = 1, \ldots, n\right] = e^{-N I_\rho(\{x_i\})}$, met fluctuatiefunctionaal

$$I_\rho(\{x_i\}) = -\sum_{i=1}^{n} x_i \log p_i + \sum_{i=1}^{n} x_i \log x_i$$

= relatieve entropie tn twee kansmaten. $S(\{x_i\} \mid \{p_i\})$

De relatieve entropie $S(\mu|\rho)$ is gedefinieerd als

$$S(\mu|\rho) = \sum_i \mu_i \log(\mu_i/\rho_i)$$

e- Sanov Theorem

Zij $X_1, \ldots, X_N$ onafhankelijke, identiek verdeelde variabelen op eindige kansruimte K met gekende kansverdeling ρ.

Nu definiëren we een nieuwe kansverdeling op K door $m^N(\cdot) = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_i,\cdot}$, met $\cdot \in K$. Merk op, $m^N(\cdot) \to \rho(\cdot)$ als $N \to \infty$. Deze kansverdeling is de „random probability distribution".

Voorbeeld: dobbelsteen 10 keer gooien en de relatieve frequenties als kansverdeling op $\{1 \ldots 6\}$ bekomen.

Wat is bijvoorbeeld kans op $x \in K$: $\mathrm{Prob}[x]$ = proportie van veranderlijken die x aannemen.

Merk op dat voor m^N geldt dat het eerste moment a gelijk is aan

$$a = \sum_{x\in K} x\, m^N(x) = \sum_{x\in K} x\left(\sum_{i=1} \frac{1}{N}\delta_{X_i,x}\right) = \frac{1}{N}\sum_{i=1} \underbrace{\sum_{x\in K} x\cdot\delta_{X_i,x}}_{\text{„="}\ X_i\ \leadsto\ \text{meet wat } X_i \text{ is}} = \frac{1}{N}\sum X_i = \langle X\rangle,$$

het gewone gemiddelde dus.

Nu, wat is de kans om proporties $\mu(x)$ te observeren wanneer we de a priori-verdeling ρ kennen: $\mathrm{Prob}_\rho[m^N(x) \simeq \mu(x),\ x\in K] \simeq e^{-N I_\rho(\mu)}$,

met $I_\rho(\mu) = \sum_x \mu(x)\log\left(\frac{\mu(x)}{\rho(x)}\right)$, wat wei de formule voor de relatieve entropie $S(\mu|\rho)$. Dit resultaat heet het Sanov Theorem:

$$\mathrm{Prob}_\rho[\rho_N \simeq \mu] = e^{-N S(\mu|\rho)}$$

(1) We checken of het Sanov Theorem compatibel is met „b", voor $\rho(0) = 1-p$ en $\rho(1) = p$.

We gaan na dat $\text{Prob}_\rho\left[\frac{1}{N}\sum X_i = \langle X\rangle_\nu = a\right] = e^{-N \inf_\nu S(\nu|\rho)}$

$$\langle X\rangle_\nu = \sum_{x\in K} x\,\nu(x) \underset{K=\{0,1\}}{=} 0\cdot\nu(0) + 1\cdot\nu(1) = \nu(1) \quad (\circ)$$

De relatieve entropie van „een" ν is gedefinieerd als $S(\nu|\rho) = \sum_i \nu_i \log \nu_i - \nu_i \log \rho_i$ of in dit geval $S(\nu|\rho) = \nu(0)\log(\nu(0)) - \nu(0)\log\rho(0) + \nu(1)\log\nu(1) - \nu(1)\log\rho(1)$, of omdat $\nu(1) = 1-\nu(0)$ en $\rho(0) = 1-p$, $\rho(1) = p$:

$$S(\nu|\rho) = \nu(0)\log\nu(0) - \nu(0)\log(1-p) + (1-\nu(0))\log(1-\nu(0)) - (1-\nu(0))\log p.$$

Nu nemen we het infimum over alle ν, wat echter betekent dat we $S(\nu|\rho)$ gi $\nu(0)$ moeten minimaliseren en de waarde waarbij het minimum optreedt, nemen we als $\nu(0)$. $\nu(1)$ is dan $\nu(1) = 1-\nu(0)$. Noem voor het gemak $\nu(0) = t$.

$$S(t) = t\log t - t\log(1-p) + (1-t)\log(1-t) - (1-t)\log p.$$

$$\Rightarrow \frac{\partial S}{\partial t} = \log t + 1 - \log(1-p) - \log(1-t) + 1 + \log p \quad (\Delta)$$

$$= \log\left(\frac{t}{1-t}\right) + \underbrace{\log\left(\frac{p}{1-p}\right) + 2}_{=\log\left(e^2 \frac{p}{1-p}\right)} = 0 \Rightarrow \log\left(\frac{t}{1-t}\right) = \log\left(\frac{1-p}{e^2 p}\right)$$

$$\rightarrow \frac{t}{1-t} = \frac{1-p}{e^2 p}$$

$$\Rightarrow e^2 p t = 1-p-(1-p)t \Rightarrow t^* = \frac{1-p}{1+(e^2-1)p} = \nu(0).$$

Dit zullen we invullen in (Δ) en meteen gebruik maken van $(\circ)$: $\nu(1) = 1-\nu(0) = a$; of $\nu(0) = 1-a$

$$\inf_\nu S(\nu|\rho) = (1-a)\log(1-a) - (1-a)\log(1-p) + a\log a - a\log p$$

$$= (1-a)\log\left(\frac{1-a}{1-p}\right) + a\log\left(\frac{a}{p}\right),$$

zoals reeds gevonden!

(2) Als nu $\rho(x) = e^{-V(x)}/Z$ $\quad (Z = \sum_x e^{-V(x)})$ een kansverdeling is. (zie blz 6) We tonen aan dat dan $S(\nu|\rho) = \sum_x \nu(x)V(x) - s(\nu) + \log Z$, met $s(\nu) = -\sum_x \nu(x)\log\nu(x)$ de Shannonentropie. We rekenen gewoon uit:

Zij ν een kansverdeling op K, dan is

$$S(\nu|\rho) = \sum_x \nu(x)\log\nu(x) - \nu(x)\log\rho(x)$$

$$\rho(x) = \frac{e^{-V(x)}}{Z}$$

$$= -\Big(\underbrace{-\sum_x \nu(x)\log\nu(x)}_{s(\nu)}\Big) - \sum_x \nu(x)\underbrace{\log(e^{-V(x)})}_{-V(x)} - \sum_x \nu(x)\cdot(-\log Z)$$

$$= \sum_x \nu(x)V(x) + \underbrace{\sum_x \nu(x)\log Z}_{=\log Z \underbrace{\sum_x \nu(x)}_{=1} = \log Z} - s(\nu)$$

$$= \sum_x \nu(x)V(x) + \log Z - s(\nu).$$

f- De vrije-energiemethode

Voor willekeurige onafhankelijke a priori distributies is bovenstaande methode niet zo eenvoudig, moeilijk zelfs. Een methode om de fluctuatiefunctionaal $I(x)$ in $\mathrm{Prob}_\rho[m_N \simeq x] = e^{-NI(x)}$ te berekenen voor „algemene" verdelingen is gebaseerd op de Laplace-methode. Voor alle functies $F: K \to \mathbb{R}$ geldt dat $\frac{1}{N}\log\left\langle e^{NF\left(\frac{1}{N}(X_1+\ldots+X_N)\right)}\right\rangle = \frac{1}{N}\log\left\{\sum_x e^{NF(x)}\,\mu_N(x)\right\}$

kans om x te meten

$$= \frac{1}{N}\log\left\{\sum_x e^{NF(x)}e^{-NI(x)}\right\} = \frac{1}{N}\log\left\{\sum_x e^{N(F(x)-I(x))}\right\}$$

$$\simeq \frac{1}{N}\log\left\{e^{N\sup_x\{F(x)-I(x)\}}\right\} = \frac{1}{N}\cdot N\sup_x\{F(x)-I(x)\} = \sup_x\{F(x)-I(x)\}$$

Als we nu $F(m) = \theta m$ kiezen, voor „een" $\theta \in \mathbb{R}$. Dan is $\left\langle e^{NF\left(\frac{X_1+\ldots+X_N}{N}\right)}\right\rangle$ waarschijnlijk gemakkelijk, want $NF\left(\frac{X_1+\ldots+X_N}{N}\right) = \theta(X_1+\ldots+X_N)$ en zo

$$\left\langle e^{\theta(X_1+\ldots+X_N)}\right\rangle = \left\langle e^{\theta X_1}e^{\cdots}\ldots e^{\cdots}e^{\theta X_N}\right\rangle = \left\langle e^{\theta X}\right\rangle^N \simeq e^{N\sup_m\{\theta m - I(m)\}}$$

Als we nu ψ definiëren op $\mathbb{R}$ als $\psi(\theta) := \frac{1}{N} \log\langle e^{\theta(X_1+\ldots+X_N)}\rangle$

X_i zijn onafhankelijke, identiek verdeelde variabelen

X is verdeeld volgens ρ.

$$= \frac{1}{N} \log\langle e^{\theta X}\rangle^N = \frac{N}{N} \log\left\{\sum_x \rho(x) e^{\theta x}\right\}$$

$$= \log\left\{\sum_x \rho(x) e^{\theta x}\right\}.$$

ψ is dan de Legendretransformatie van fluctuatiefunctionaal $I(m)$:

$$\psi(\theta) = \sup_m \{\theta m - I(m)\}.$$

Interludium: Legendretransformatie

Stel dat $f: \mathbb{R} \to \mathbb{R}$ een concave ($f''(x) \geq 0$), continu afleidbare functie, dan is de Legendretransformatie gedefinieerd als,

$$I: \mathbb{R} \to \mathbb{R}: \theta \mapsto I(\theta) = \sup_x \{\theta x - f(x)\}.$$

We veronderstellen dat f braaf is, we berekenen de eerste afgeleide van $\theta x - f(x)$ en stellen nul (want we zoeken het extremum):

$$\frac{\partial}{\partial x}\left(\theta x - f(x)\right)\Big|_{x^*(\theta)} = \theta - f(x^*(\theta)) = 0 \Rightarrow \theta = f(x^*(\theta)) \text{ en zo is}$$

$$I(\theta) = x^*(\theta)\, f(x^*(\theta)) - f(x^*(\theta))$$

Nu tonen we nog aan dat de Legendretransformatie van I terug f oplevert. Legendretransformatie is dus haar eigen inverse.

Merk nu op, omdat Legendretransformatie haar eigen inverse is, en we $\psi(\theta)$ gedefinieerd hadden als de Legendretransformatie van de fluctuatiefunctionaal $I(m)$: $\psi(\theta) = \sup_m \{\theta m - I(m)\}$, kan de fluctuatiefunctionaal gedefinieerd worden als Legendregetransformeerde van ψ: $I(m) = \sup_\theta \{\theta m - \psi(\theta)\}$.

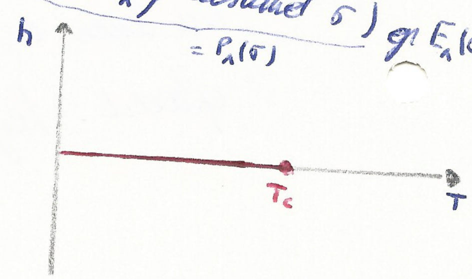
Nu kan het interessant zijn uit te leggen waarom dit de vrije-energiemethode heet. We zullen dit doen aan de hand van het Isingmodel in 2D. We nemen een vierkant $\Lambda \subset \mathbb{Z}^2$ rond de oorsprong. Voor $i \in \Lambda$ zal $\sigma_i = \pm 1$ en zal de totale energie van Λ gelijk zijn aan $E_\Lambda(\sigma) = -J \sum_{\boxed{i\sim j}} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i$. We leggen ook randvoorwaarden

naaste buren

(periodische, vrije (zonder spins aan de rand), vaste (alles + of $-$) ...) op. We definiëren nu de magnetisatie als $M_\Lambda(\sigma) = \sum_{i\in\Lambda} \sigma_i$. Experimenteel zal $\frac{1}{\Lambda} M_\Lambda(\sigma) \xrightarrow[\Lambda \to \mathbb{Z}^2]{} m(T, h, J, \dots)$ (de verwachte magnetisatie). We kunnen $\frac{1}{\Lambda} M_\Lambda(\sigma)$ ook zien als een geïnduceerde Gibbsverdeling $\underbrace{\frac{e^{-\beta E_\Lambda(\sigma)}}{Z} \text{ (= kans}_\Lambda \text{ op toestand } \sigma)}_{= P_\Lambda(\sigma)}$ op $E_\Lambda(\sigma)$

Vanuit hier is dus m^* en niet m de verwachte magnetisatie?

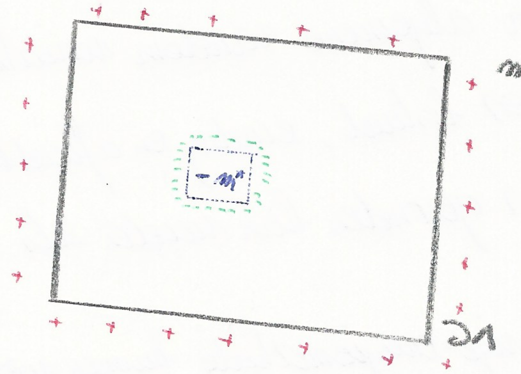
Herinner je, er is een faseovergang.



Stel een $m^* = \lim_\Lambda \frac{M_\Lambda(\sigma)}{\Lambda}$ onder $P_\Lambda(\sigma)$ en de RVW.
We weten dat $m^* > 0$ en uniform zal blijven voor $h \to 0$ als $T < T_c$ en $m^* = 0$ als $h = 0$ in $T = T_c$. Voor $T = T_c$ is er een spin(flip-)symmetrie.
Merk op dat in de thermodynamische limiet en $\Lambda \to \mathbb{Z}^2$ de oorsprongspin de rand niet meer voelt, maar nog wel de effecten ervan draagt (zijnde haar spin) (Homeopathisch effect! :)). $\langle \sigma_0 \rangle_\Lambda \to m^*$.

Nu zijn we geïnteresseerd in $\mathrm{Prob}\left[\frac{1}{|\Lambda|} M_\Lambda(\sigma) \simeq m\right]$. Naïef zouden we kunnen denken aan $e^{-|\Lambda| I(m)}$, maar dit zal fout blijken te zijn. Als $\frac{J}{k_BT}$ nu klein zou zijn en laten de interactieterm vallen, is $E_\Lambda(\sigma) = -h\sum_i \sigma_i$, wat uiteindelijk gewoon het opgooien van een muntstuk is, met $\mathrm{Prob}[\sigma_i = +1] = \frac{e^{h/k_BT}}{2\cosh\left(\frac{h}{k_BT}\right)}$. Als we nu de interactieterm niet verwaarlozen

Kunnen we in een gebiedje „$-m^*$" creëren als voor Λ m^* geldt? Ja, door rond het systeem een rand van spin aan te leggen. De kans op $-m^*$ is dus de kans op het kunnen maken van de rand.



Nu; vanuit de statistische mechanica volgt:

$$\mathrm{Prob}\left[\frac{1}{|\Lambda|} M_\Lambda(\sigma) \simeq m\right] = \frac{1}{Z}\cdot \underbrace{\sum_\sigma \delta_{M_\Lambda(\sigma)\simeq|\Lambda| m} \exp\Big\{-\beta\Big(-J\sum_\Lambda \sigma_i\sigma_j - h\underbrace{\sum_\Lambda \sigma_i}_{=M_\Lambda(\sigma)}\Big)\Big\}}_{\text{telt het aantal toestanden met } M_\Lambda(\sigma)\simeq|\Lambda| m \text{ volgens hun gewicht (= waarschijnlijkheid van voorkomen).}}$$

$$= \frac{1}{Z}\sum_\sigma \delta_{M\simeq|\Lambda| m}\, e^{\theta|\Lambda|M_\Lambda(\sigma)} \exp\Big\{\beta J\sum_\Lambda \sigma_i\sigma_j + \beta h M_\Lambda(\sigma)\Big\}\, e^{-\theta|\Lambda|M_\Lambda(\sigma)}$$

$$= \frac{1}{Z}\sum_\sigma \delta_{M_\Lambda(\sigma)\simeq|\Lambda| m} \exp\Big\{\beta J\sum_\Lambda \sigma_i\sigma_j + \theta|\Lambda|M_\Lambda(\sigma)\Big\}\, e^{\beta h M_\Lambda(\sigma) - \theta|\Lambda|M_\Lambda(\sigma)}$$

$$Z_\theta = \sum_\sigma \exp\Big\{\beta J\sum_\Lambda \sigma_i\sigma_j + \theta|\Lambda|M_\Lambda(\sigma)\Big\}\text{, beschouw } \theta \text{ als magnetisch veld}$$

$$= \frac{Z_\theta}{Z}\underbrace{\left(\frac{1}{Z_\theta}\sum_\sigma \delta_{M_\Lambda(\sigma)\simeq|\Lambda| m} \exp\Big\{\beta J\sum_\Lambda \sigma_i\sigma_j + \theta|\Lambda|M_\Lambda(\sigma)\Big\}\right)}_{\text{Kies } \theta=\theta^* \text{ zodat dit 1 is, } \mathrm{Prob}_\theta\left[\frac{M_\Lambda(\sigma)}{\Lambda}\simeq m\right]\simeq 1} e^{\beta h M_\Lambda(\sigma) - \theta|\Lambda|M_\Lambda(\sigma)}$$

$$= \frac{Z_{\theta^*}}{Z} e^{(\beta h - \theta^*)M_\Lambda(\sigma)} = \frac{Z_{\theta^*}}{Z} e^{(\beta h-\theta^*)|\Lambda| m}$$

$$= e^{-|\Lambda|(\theta^*-\beta h)m + \frac{|\Lambda|}{|\Lambda|}\log\left(\frac{Z_\Lambda(\theta^*)}{Z_\Lambda(h)}\right)} = e^{-|\Lambda|\left[(\theta^*-\beta h)m - \frac{1}{|\Lambda|}\log\left(\frac{Z_\Lambda(\theta^*)}{Z_\Lambda(h)}\right)\right]}$$

De fluctuatiefunctionaal is dus

$$I(m) = (\theta^*-\beta h)m - \lim_{\Lambda\to\mathbb{Z}^2}\frac{1}{|\Lambda|}\log\frac{Z_\Lambda(\theta^*)}{Z_\Lambda(h)}\text{, met } \theta^* \text{ zodat } m(T,\theta^*) = m.$$

Merk op dat voor $\lim_{\Lambda} \frac{1}{|\Lambda|} \underbrace{\log \frac{Z_\Lambda(\theta^*)}{Z_\Lambda(h)}}_{(\beta)}$ geldt dat deze limiet bestaat en dat

(α) een verschil in Helmholtz vrije energie ($F = -k_B T \log Z$) uitdrukt. Dit verschil drukt de arbeid uit die je moet leveren met een magneet (= de θ^*) om een magnetisatie m te krijgen. We begrijpen nu waar de naam vrije-energiemethode vandaan komt.

Voor algemene random variabelen stelt het Gärtner-Ellis Theorema dat als $\psi(\theta)$ bestaat, eindig en afleidbaar is voor alle θ, de fluctuatie functionaal $I(m)$ gevonden kan worden als Legendre-getransformeerde van ψ.

De vrije-energiemethode kunnen we gebruiken bij het berekenen van $I(m) = \sup_\theta \{m\theta - \psi(\theta)\}$. Dit supremum wordt bereikt in een bepaalde $\theta^* = \theta^*(m)$ en zo zodat (argument afleiden en nulstellen) $\psi'(\theta^*) = m$. Merk nu op dat door gebruik te maken van de vorm $\psi(\theta) = \log\left\{\sum_x \rho(x) e^{\theta x}\right\}$ en af te leiden naar θ, θ^* in te vullen en m te stellen dat

$$\psi'(\theta) = \frac{1}{\underbrace{\sum_x \rho(x) e^{\theta x}}_{= e^{\psi(\theta^*)}}} \cdot \left\{\sum_x x \rho(x) e^{\theta x}\right\} = e^{-\psi(\theta^*)} \sum_x x \rho(x) e^{\theta^* x} =$$

$$\sum_x x \underbrace{\rho(x) e^{\theta^* x - \psi(\theta^*)}}_{:= \rho^*(x)} = m.$$

Als we nu een nieuwe „tilted" distributie ρ^* definiëren als $\rho^*(x) := \rho(x) e^{\theta^* x - \psi(\theta^*)}$, heeft deze meteen verwachte waarde m, wat handig kan zijn.

y- Contractieprincipe

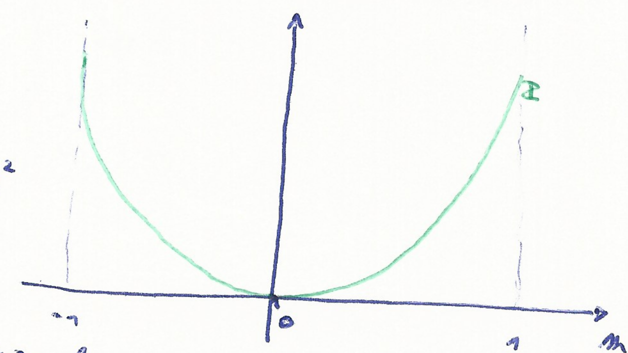
We hebben reeds grote afwijkingen op $m_N = \frac{1}{N}(x_1 + \dots + x_N)$ bestudeerd, maar dit kan ook op andere niveau's, bijvoorbeeld grote fluctuaties op de empirische dichtheidsfunctie $m^N(\cdot) = \frac{1}{N}\sum \delta_{x_i}(\cdot \in K)$.

We hebben op bladzijde 9 reeds bewezen dat voor een eenvoudig muntstuk $\rho(0) = 1-p$, $\rho(1) = p$ geldt dat $I_p(m) = \inf_{\nu : \langle X \rangle_\nu = m} S(\nu | \rho)$

- In de motivering van de large deviation theory is gezegd dat vanuit de theorie voor grote afwijkingen ook de theorie voor kleine afwijkingen (= CLT) gevonden moet kunnen worden.

→ Voorbeeld: het eerlijk muntstuk met $\sigma_i = \pm 1$. We hebben al bewezen dat dan $\text{Prob}\left[\frac{1}{N}\sum_{i=1}^{N}\sigma_i \simeq m\right] = e^{-N I(m)}$, met $I(m)$ de fluctuatie-functionaal $I(m) = \log 2 + \frac{m+1}{2}\log\left(\frac{m+1}{2}\right) + \frac{1-m}{2}\log\left(\frac{1-m}{2}\right)$

Omdat we geïnteresseerd zijn in kleine afwijkingen maken we een Taylorbenadering rond $m_0 = 0$.



$$I(m) \simeq I(m_0) + I'(m_0)(m - m_0) + \frac{I''(m_0)}{2}(m-m_0)^2$$

$$= I(0) + I'(0)\cdot m + \frac{I''(0)}{2} m^2$$

- $I(0) = \log 2 + \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2} = \log 2 - \log 2 = 0$
- $I'(m) = \frac{1}{2}\log\left(\frac{m+1}{1-m}\right) \Rightarrow I'(0) = \frac{1}{2}\log 1 = 0$
- $I''(m) = \frac{1}{m+1} \Rightarrow I''(0) = 1$

$$= \frac{1}{2}m^2$$

Dan is $\text{Prob}\left[\frac{1}{\sqrt{N}}\sum \sigma_i \simeq m\right] = \text{Prob}\left[\frac{1}{N}\sum \sigma_i = \frac{m}{\sqrt{N}}\right] = e^{-N I(m/\sqrt{N})}$

4. Dynamische fluctuatietheorie

Als we nu een irreducibel Markovproces X_t beschouwen op een eindige toestandsruimte K, met stationaire verdeling ρ en transitieprobabiliteiten $k(x,y)$.

Het is interessant om de empirische occupaties $\rho^T(x) = \frac{1}{T}\int_0^T dt\, \delta_{X_t,x}$ $(x\in K)$ te bekijken. Als $T\to\infty$ zal bijvoorbeeld $\rho^T(x) \to \rho(x)$.

In een dynamische fluctuatiefunctionaaltheorie zoeken we een functionaal D op verdelingen naar $\mathbb{R}$, zodat

$$\mathrm{Prob}_\rho\left[\frac{1}{T}\int_0^T dt\, \delta_{X_t,x} \simeq \mu(x) : x\in K\right] = e^{-TD(\mu)} \qquad (*)$$

De stationaire verdeling wordt gekarakteriseerd als minimiser van D.

Om dit te begrijpen, denken we aan het Gibbsvariatieprincipe: de Gibbs-verdeling minimaliseert de vrije-energiefunctionaal

$$F(\mu) := \beta \sum_{x\in K} \mu(x) E(x) + \sum_{x\in K} \mu(x) \log(\mu(x))$$

in de zin dat voor de Gibbsverdeling $\rho(x) = \dfrac{e^{-\beta E(x)}}{Z}$ $(x\in K)$ geldt dat $F(\rho) = -\log Z \leq F(\mu)$ voor elke kansverdeling μ op K.

Om dit te bewijzen, maken we gebruik van het niet-negatief zijn van de relatieve entropie. Immers, voor elke μ,ρ geldt dat $-S(\mu|\rho) = -\sum_{x\in K} \mu(x)\log\frac{\mu(x)}{\rho(x)} =$

$$\sum_{x\in K}\mu(x)\log\frac{\rho(x)}{\mu(x)} \leq \log\left(\sum_x \cancel{\mu(x)}\frac{\rho(x)}{\cancel{\mu(x)}}\right) = \log\left(\sum_x \rho(x)\right) = \log(1) = 0$$

Jensen ongelijkheid

$$\boxed{E[f(x)] = \sum_x p(x) f(x) \leq f\left(\sum x\, p(x)\right) = f(E[X])}$$

Dus voor alle kansmaten μ,ρ geldt $S(\mu|\rho) \geq 0$, met gelijkheid voor $\mu=\rho$.

Zij ρ nu de Gibbsverdeling. Dan is

$$S(\nu|\rho) = \sum_{x\in K} \nu(x)\log\frac{\nu(x)}{\rho(x)} = \sum_{x\in K} \nu(x)\log(\nu(x)) - \nu(x)\log\rho(x)$$

$$= \sum_{x\in K}\nu(x)\log\nu(x) - \sum_x (-\beta E(x))\nu(x) + \sum_x \nu(x)\log Z$$

$$= \sum_{x\in K}\nu(x)\log\nu(x) + \beta\sum_x \nu(x)E(x) + \log Z$$

$$F(\rho) = \beta\sum_x \rho(x)E(x) + \sum_x \rho(x)\underbrace{\log\rho(x)}_{= -\beta E(x) - \log Z}$$

$$= \cancel{\beta\sum_x \rho(x)E(x)} \cancel{- \beta\sum_x \rho(x)E(x)} - \sum \rho(x)\log Z = -\log Z$$

$$= F(\nu) - F(\rho)$$

Omdat $S(\nu|\rho) \geq 0$ is $F(\nu) - F(\rho) \geq 0$, of $F(\nu) \geq F(\rho) = -\log Z$. Voor alle kansverdelingen ν op K geldt dus $F(\rho) = -\log Z \leq F(\nu)$.

Merk dus op, vanuit Sanov's Theorem weten we dat de relatieve entropie gelijk is aan de fluctuatiefunctionaal. Zoals hierboven getoond bepaalt het verschil in vrije energie de relatieve entropie en zo de fluctuatiefunctionaal.

4.2 Dynamische fluctuatiefunctionaal en de fysica erachter.