⊙ GRANDI DEVIAZIONI PER IL PROCESSO EMPIRICO

Contesto del teorema di Sanov con notazioni diverse

$\omega_1, \dots, \omega_n$ iid a valori $\mathcal{X}$

pensiamo $\mathcal{X} = \{a, \dots, z\}$ = alfabeto finito (anche $\mathcal{X} = \{0, 1\}$)

$\omega = (\omega_1, \dots, \omega_n)$ = parola di n lettere : $\omega_1 \cdots \omega_n$

$$\pi_n = \text{misura empirica} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\omega_i}$$

$$\pi_n(\{a\}) = \frac{1}{n} \left| \{ i : \omega_i = a \} \right| = \frac{\#\text{ DI } a \text{ in } \omega}{n}$$

$$= \text{frazione delle } a$$

π_n (misura empirica) individua la statistica delle varie lettere

Vogliamo fare la statistica dei dittonghi, . . .

ad esempio

$$\pi_n^{(2)}(\{(a,b)\}) = \frac{\#\text{ coppie (consecutive) con } ab}{n-1}$$

su n lettere osservo $n-1$ dittonghi

$$1 \; 2 \; 3 \; \dots \; n-1 \; n$$

$$\pi_n^{(3)}(\{aca\}) = \frac{\#\text{ terne (consecutive) con } aca}{n-2}$$

⋮

$\pi^{(2)}$ probabilità su $\mathcal{X} \times \mathcal{X}$, $\pi^{(3)}$ prob. su $\mathcal{X} \times \mathcal{X} \times \mathcal{X}$, ...

è un po' scomodo che la normalizzazione cambia

conviene definire in modo leggermente diverso

$\omega = \omega_1 \cdots \omega_n$ parola di n lettere

$\omega^{(n)} = \cdots \omega_1 \cdots \omega_n \omega_1 \cdots \omega_n \omega_1 \cdots \omega_n \omega_1 \cdots \omega_n \omega_1 \cdots \omega_n$

= parola di ∞ lettere (ambo i lati) ottenuta ripetendo $\omega_1 \cdots \omega_n$ in modo periodico

$\omega^{(n)} \in \mathcal{X}^{\mathbb{Z}}$

ora

$$\pi_n^{(2)}(\{(a,b)\}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{(\omega_i^{(n)}, \omega_{i+1}^{(n)})}((a,b))$$

" è come avere aggiunto la prima lettera alla fine della parola e contare tra gli n dittonghi "

$$\pi_n^{(3)}(\{acb\}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{(\omega_i^{(n)}, \omega_{i+1}^{(n)}, \omega_{i+2}^{n})}(a\,c\,a)$$

Il processo empirico è il modo per ricordarsi (con una sola probabilità) la statistica delle lettere, dei dittonghi, ...

Per le applicazioni ai campi di Gibbs $\mathbb{Z}^d$ invece che $\mathbb{Z}$

$\Omega = \mathcal{X}^{\mathbb{Z}^d}$ le traslazioni come al solito

$x \in \mathbb{Z}^d \qquad \vartheta_x : \Omega \circlearrowleft$

$(\vartheta_x \omega)_y = \omega_{y-x}$

e $\quad \mathcal{P}_\theta(\Omega) =$ probabilità su Ω invarianti per traslazione : $\mu = \mu \circ \vartheta_x^{-1}$ $\forall x \in \mathbb{Z}^d$

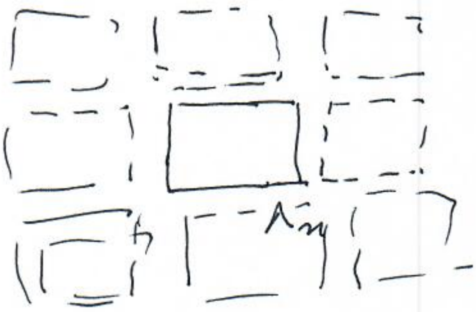
∘ Definizione Processo empirico

$\Lambda_n = \{1,\dots,n\}^d \subset\subset \mathbb{Z}^d$ cubo di lato n in $\mathbb{Z}^d$

campione $\{\omega_x,\ x\in\Lambda_n\} \in \mathcal{X}^{\Lambda_n}$

lo ripeto in modo periodico

$\omega^{(n)} \in \mathcal{X}^{\mathbb{Z}^d}$



$$R_n = \underbrace{\frac{1}{|\Lambda_n|}\sum_{x\in\Lambda_n} \underbrace{\delta_{\theta_x\omega^{(n)}}}_{\text{è una prob. su } \Omega}}_{\text{per costruzione (ho replicato in modo periodico) è prob. su } \Omega \text{ invariante per traslazioni}}$$

$R_n \in \mathcal{P}_\theta(\Omega)$

se campione aleatorio, tipo generato da v.a. i.i.d.
R_n prob. random invariante per traslazioni su Ω

Vediamo che effettivamente R_n contiene le informazioni sulla statistica delle lettere, dei dittonghi, ...
[caso $d=1$]

∘ Scegli $F:\Omega\to\mathbb{R}$ così

$$F(\omega) = \begin{cases} 1 & \text{se } \omega_0 = a \\ 0 & \text{altrimenti} \end{cases}$$

$$R_n(F) = \frac{1}{n}\sum_{x=1}^{n} \delta_{\theta_x\omega^n}(F)$$

$$= \frac{1}{n} \sum_{x=1}^{n} 1\!\!1_{(\theta_x \omega^n)_0 = a}$$

$$= \frac{1}{n} \sum_{x=1}^{n} 1\!\!1_{\omega_x = a} = \frac{\# a \text{ in } \omega_1 \dots \omega_n}{n}$$

- scegli

$$f(\omega) = \begin{cases} 1 & \text{se} \quad \omega_0 = a,\ \omega_1 = b \\ 0 & \text{altrimenti} \end{cases}$$

$$R_n(f) = \dots = \frac{1}{n} \sum_{x=1}^{n} 1\!\!1_{\left((\theta_x \omega^n)_0, (\theta_x \omega^n)_1 \right) = (a,b)}$$

$= \frac{1}{n}$ # dittonghi ab per la parola $\omega_1 \dots \omega_n \omega_1$
[n dittonghi]

OSS Π_n = misura empirica

$$= \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \delta_{\omega_x} \qquad \text{probab. su } \mathcal{X}$$

R_n = processo empirico

$$= \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \delta_{\theta_x \omega^{(n)}} \qquad \text{prob. su } \mathcal{X}^{\mathbb{Z}^d}$$
inv. per traslazioni

il marginale di R_n su un sito fissato [tutti uguali per inv. per trasl.] è proprio la misura empirica

Scopo: grandi deviazioni per R_n, prima nel caso iid poi per campi di Gibbs.

• **Legge dei grandi numeri per il processo empirico**

$\omega_x \quad x \in \mathbb{Z}^d$ iid con legge μ

(a valori in $\mathcal{X}$ compatto)

$\Omega = \mathcal{X}^{\mathbb{Z}^d}$ $\quad \mathcal{P}(\Omega)$ con topologia della conv. debole

Prop Sia $\mathbb{P} = \mu^{\otimes \mathbb{Z}^d}$ probabilità prodotto su Ω (è invariante per traslazioni)

$R_n \xrightarrow{\mathbb{P}rob} \mathbb{P}$ conv. in probabilità

dim per compattezza di $\mathcal{P}(\Omega)$ basta identificare i punti limite

ovvero $\forall f \in C(\Omega)$ $\quad R_n(f) \xrightarrow{\mathbb{P}rob} \mathbb{P}(f)$

per densità basta considerare il caso f locale

- f dipende solo da ω_0

$$R_n(f) = \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} f(\omega_x)$$

$\longrightarrow \mu(f)$ legge grandi numeri variabili iid

- f dipende solo da ω_0, ω_{e_1} $\quad (0 \quad e_1)$

è una situazione analoga a questa

$X_1, \dots X_n$ iid (a valori $\mathbb{R}$) con legge μ

$f : \mathbb{R}^2 \to \mathbb{R}$ continua e limitata

$$\frac{1}{n-1} \sum_{i=1}^{n-1} f(X_i, X_{i+1}) \xrightarrow{\mathbb{P}} \int_{\mathbb{R}^2} \mu(dx)\,\mu(dy) \quad f(x,y)$$

Poiché (X_i, X_{i+1}) e (X_j, X_{j+1})

sono indipendenti non appena $|i-j| > 1$

viene facile

[usare Chebyshev come legge debole dei grandi numeri]

Se f è locale dipende solo da un numero finito di ω_x, $x \in \mathbb{Z}^d$ è lo stesso

Grandi deviazioni per il processo empirico nel caso iid

Ambientazione topologica

$\mathcal{X}$ polacco compatto (per noi $\mathcal{X} = \{0,1\}$)

$\Omega = \mathcal{X}^{\mathbb{Z}^d}$ con topologia prodotto : polacco compatto

$\mathcal{P}(\Omega)$: probabilità su Ω . " "

con topologia indotta

convergenza debole

$\mathcal{P}_\theta(\Omega)$: - — -

.. invarianti per traslazioni

è un sottoinsieme chiuso di $\mathcal{P}(\Omega)$

$\rightsquigarrow$ topologia relativa ... polacco compatto

$\mu \in \mathcal{P}(\mathcal{X})$ legge di ω_x

$\mathbb{P} = \mu^{\otimes \mathbb{Z}^d} \in \mathcal{P}_\theta(\Omega)$

R_n processo empirico: variabile aleatoria a valori $\mathcal{P}_\theta(\Omega)$

- Vi ricordo $\mathrm{ent}(\cdot|\cdot)$

[discussa nel caso $\Omega = \{0,1\}$ e μ uniforme, ma è lo stesso in questa situazione]

$$\mathrm{ent}(\cdot \mid \mathbb{P}) : \mathcal{P}_\theta(\Omega) \longrightarrow [0, +\infty]$$

lsc e affine

se $Q \in \mathcal{P}_\theta(\Omega)$

marginale su $\Omega_\Lambda = X^\Lambda$

$$\mathrm{ent}(Q|P) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \mathrm{Ent}(\pi_\Lambda Q \mid \underbrace{\pi_\Lambda \mathbb{P}}_{= \mu^{\otimes \Lambda}})$$

superadditività

$$= \sup_{\Lambda \subset\subset \mathbb{Z}^d} \frac{1}{|\Lambda|} \mathrm{Ent}(\pi_\Lambda Q \mid \pi_\Lambda \mathbb{P})$$

[P è prodotto]

<u>TEO</u> (Principio di grandi deviazioni per il processo empirico)

R_n soddisfa un principio di grandi deviazioni con tasso $I : \mathcal{P}_\theta(\Omega) \to [0,+\infty]$ dato da

$$I(Q) = \mathrm{ent}(Q|\mathbb{P})$$

ovvero

- $\forall\, C$ chiuso in $\mathcal{P}_\theta(\Omega)$

$$\overline{\lim_n} \frac{1}{|\Lambda_n|} \lg \mathbb{P}(R_n \in C) \leq - \inf_{Q \in C} \mathrm{ent}(Q|\mathbb{P})$$

- $\forall\, A$ aperto in $\mathcal{P}_\theta(\Omega)$

$$\underline{\lim_n} \frac{1}{|\Lambda_n|} \lg \mathbb{P}(R_n \in A) \geq - \inf_{Q \in A} \mathrm{ent}(Q|\mathbb{P})$$

OSS Sanov è un corollario.

Infatti se $\mathbb{Q} = \pi^{\otimes \mathbb{Z}^d}$ ($\mathbb{P} = \mu^{\otimes \mathbb{Z}^d}$)

allora

$$ent(\mathbb{Q} | \mathbb{P}) = Ent(\pi | \mu)$$

per contrazioni si tratta di verificare che se fisso ~~di~~ il marginale di $\mathbb{Q}$ su un sito ($= \pi$) la scelta più conveniente è $\mathbb{Q} = \pi^{\otimes \mathbb{Z}^d}$

Esercizio: su $\mathcal{X} \times \mathcal{X}$ sia $P = \mu \times \mu$

sia ora Q pos. su $\mathcal{X} \times \mathcal{X}$ con gli stessi marginali ($= \pi$)

Allora
$$Ent(Q \mid \mu \times \mu) \geq Ent(\pi \times \pi \mid \mu \times \mu)$$
$$\parallel$$
$$2\, Ent(\pi | \mu)$$

dim ~~...~~ Per il Teorema di Sanov sceglievamo perturbazioni rimanendo nel mondo di v.a. i.i.d

Qui non basterà più.

- Stima dall'alto

La strategia è come per Sanov, ci organizziamo una famiglia di "cambi esponenziali" e poi ottimizziamo sul parametro

$$\mathbb{P}_\Lambda = \mu^{\otimes \Lambda} \qquad \Lambda \subset\subset \mathbb{Z}^d \qquad \text{legge originaria}$$

fissa $f : \Omega \to \mathbb{R}$ continua e locale (automaticamente limitata)

ed introduci $\mathbb{P}_\Lambda^f$ come segue

$$\frac{d\mathbb{P}_\Lambda^f}{d\mathbb{P}_\Lambda} = \frac{1}{Z_\Lambda^f} \exp\{|\Lambda|\, R_n(f)\}$$

costante di normalizzazione

<u>oss</u> se $f(\omega) = \phi(\omega_0)$ [f dipende solo da ω in 0]

$\mathbb{P}_\Lambda^f$ è ancora prodotto, $Z_\Lambda^f = \mu(e^\phi)^{|\Lambda|}$, ...

... è come Sanov

Ripeti la stessa strategia che per Sanov.
Con gli ingredienti dei due lemmi seguenti concludi

<u>Lemma 1</u> Sia f locale (continua e limitata)

$$p(f) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log Z_{\Lambda_n}^f \quad \text{esiste}$$

<u>Lemma 2</u> Sia $\mathbb{Q} \in \mathcal{P}_\theta(\Omega)$.

$$\mathrm{ent}(\mathbb{Q} \mid \mathbb{P}) = \sup_{f \text{ locali}} \{\mathbb{Q}(f) - p(f)\}$$

<u>dim lemma 1</u>

In effetti è $\exists$ del limite termodinamico per la pressione nel caso (prima non considerato) di condizioni al bordo periodiche

infatti

$$|\Lambda|\, R_\Lambda(f) = \sum_{x\in\Lambda} f(\theta_x \omega^{(n)})$$

ω ripetuta periodicamente

$$= H_\Lambda(\omega \mid \text{periodiche}) = \sum_{X\cap\Lambda\neq\emptyset} \phi_X(\omega^{(n)})$$

per un opportuna interazione Φ di portata finita

(4)

dim lemma 2

(a) $$\sup_{F \text{ locali}} \{ Q(F) - P(F) \} \geq ent(Q \mid \mathbb{P})$$

Se $f \in \mathcal{F}_\Lambda$ [f è $\mathcal{F}_\Lambda$ mis : dipende solo da ω_x con x in Λ]

$$P(f) \leq \frac{1}{|\Lambda|} \lg \mathbb{E}(e^{|\Lambda| f})$$

valore di attesa rispetto a $\mathbb{P}$ (che è prodotto)

lo dimostro dopo

Uso questa disug. con $f \in \mathcal{F}_\Lambda$ rimpiazzata da $\frac{f}{|\Lambda|}$

$$Q\left(\frac{f}{|\Lambda|}\right) - P\left(\frac{f}{|\Lambda|}\right) \geq \frac{1}{|\Lambda|} \Big[\underbrace{Q}_{= \pi_\Lambda Q}(f) - \lg \underbrace{\mathbb{E}}_{= \pi_\Lambda \mathbb{P}}(e^{f}) \Big]$$

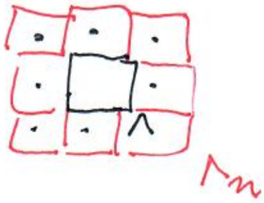
ottimizzo su
$f \in \mathcal{F}_\Lambda$ via formula variazionale di $Ent(\cdot|\cdot)$

$$\sup_{f \in \mathcal{F}_\Lambda} \{ Q(f) - P(f) \} \geq \frac{1}{|\Lambda|} Ent(\pi_\Lambda Q \mid \pi_\Lambda \mathbb{P})$$

ora ottimizzo su $\Lambda \subset\subset \mathbb{Z}^d$ e concludo

$$\left[\quad ent(Q \mid \mathbb{P}) = \sup_{\Lambda \subset\subset \mathbb{Z}^d} \frac{1}{|\Lambda|} Ent(\pi_\Lambda Q \mid \pi_\Lambda \mathbb{P}) \quad \right]$$

- Dimostro la stima $\quad F \in \mathfrak{F}_\Lambda$



$$\Lambda_n = \bigcup_k \{x_k + \Lambda\} \qquad [\text{unione disgiunta}]$$

$$Z^F_{\Lambda_n} = \mathbb{E}\left(e^{|\Lambda_n| R_n(F)} \right) = \mathbb{E}\left(e^{\sum_{x \in \Lambda_n} \theta_x f} \right)$$

$$= \mathbb{E}\left(\prod_{y \in \Lambda} e^{\sum_k \theta_{x_k + y} f} \right)$$

Holder $\qquad$ non ci sono più sovrapposizioni

$$\leq \prod_{y \in \Lambda} \mathbb{E}\left(e^{|\Lambda| \sum_k \theta_{x_k+y} f} \right)^{1/|\Lambda|}$$

$$= \prod_{y \in \Lambda} \prod_k \mathbb{E}\left(e^{|\Lambda| f} \right)^{1/|\Lambda|} = \mathbb{E}\left(e^{|\Lambda| f} \right)^{\frac{|\Lambda_n|}{|\Lambda|}}$$

ricavo

$$\frac{1}{|\Lambda_n|} \log Z^F_{\Lambda_n} \leq \frac{1}{|\Lambda|} \log \mathbb{E}\left(e^{|\Lambda| f} \right)$$

ora $n \to \infty$ e finisco

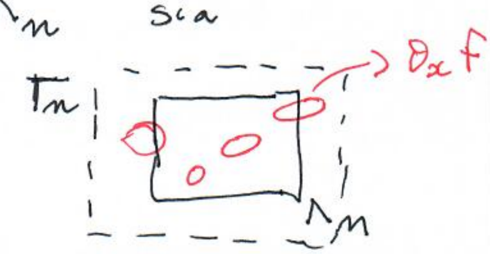
(b) $\forall$ F locale

$$\mathrm{ent}(Q \| P) \geq Q(f) - p(f)$$

qui uso $p(f)$ come limite ottenuto con condizioni al bordo ~~[illegible]~~ arbitrarie.

Fisso $f \in \mathfrak{F}_\Lambda$. ~~sia~~ Dato Λ_n sia

$$F = \sum_{x \in \Lambda_n} \theta_x f$$



è $\mathfrak{F}_{T_n}$ mis per un T_n un po' più grande

$$\frac{|\Lambda_n|}{|\Gamma_n|} \to 1$$

Per la formula variazionale di $Ent(\,|\cdot\,)$

$$Ent(\pi_{\Gamma_n} Q \mid \pi_{\Gamma_n} \mathbb{P}) \geq Q(F) - \log E(e^{F})$$

divido per $|\Lambda_n|$ e passo al limite

$$ent(Q \mid \mathbb{P}) = \lim_n \frac{1}{|\Gamma_n|} Ent(\pi_{\Gamma_n} Q \mid \pi_{\Gamma_n} \mathbb{P})$$

$$Q(f) = \lim_n \frac{1}{|\Lambda_n|} Q(F)$$

$$P(f) = \lim_n \frac{1}{|\Lambda_n|} \log E(e^{F})$$

∎

• <u>Stima dal basso</u> (Alcuni dettagli omessi)

Stessa strategia che nel caso di Sanov. ovvero

Fisso $Q \in \mathcal{P}_\vartheta(\Omega)$

devi cambiare la legge del campione $\{\omega_x, x \in \Lambda_n\}$ in modo che

- la legge dei grandi numeri per R_n sia Q
- ~~$\overline{\lim} \frac{1}{|\Lambda_n|} Ent(\text{modifica} \mid \ldots) \leq ent(Q \mid \mathbb{P})$~~
- $$\overline{\lim_n} \frac{1}{|\Lambda_n|} Ent(\text{legge modificata} \mid \text{legge originaria}) \leq ent(Q \mid \mathbb{P})$$

Facciamo questo nel caso in cui Q è ergodica

f invariante per traslazioni $\Rightarrow$ f costante Q q.s.

Esempio $\mathbb{Q}$ non ergodica $\Omega = \{0,1\}^{\mathbb{Z}}$

tira una moneta $\omega_0 = \begin{cases} 0 & \text{prob } 1-p \\ 1 & p \end{cases}$

poi $\omega_x = \omega_0 \quad \forall x \in \mathbb{Z}^d \setminus \{0\}$

evidentemente, per questa $\mathbb{Q}$,

$$\mathbb{Q} = (1-p)\,\underbrace{\delta_0}_{\omega_x = 0\ \forall x} + p\,\underbrace{\delta_1}_{\omega_x = 1\ \forall x}$$

e vero in generale ogni $\mathbb{Q} \in \mathcal{P}_\theta(\Omega)$ si può scrivere come combinazione convessa (infinita in generale) di probabilità ~~...~~ ergodiche

$$\left(\mathcal{P}_\theta(\Omega)\right)_e = \{\mathbb{Q} \in \mathcal{P}_\theta(\Omega) \text{ ergodiche}\}$$

estremali

- Se $\mathbb{Q}$ è ergodica, per il Teorema ergodico di Birkhoff

se $f \in C(\Omega)$ ~~...~~

$$\frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \theta_x f \xrightarrow{\mathbb{Q}\text{ q.s.}} \mathbb{Q}(f)$$

Data $\mathbb{Q}$ dichiaro la nuova legge del campione ω_x, $x \in \Lambda_n$ come il marginale di $\mathbb{Q}$ su Ω_{Λ_n}

- Per il Teorema ergodico ricavo

$$R_n \xrightarrow{\mathbb{Q}\text{ prob}} \mathbb{Q}$$

[non basta più legge dei grandi numeri per v.c. i.i.d.]

- Per definizione di $\text{ent}(\cdot|\cdot)$

$$\text{ent}(\mathbb{Q} \mid \mathbb{P}) = \lim_n \frac{1}{|\Lambda_n|} \text{Ent}(\underbrace{\pi_{\Lambda_n} \mathbb{Q}} \mid \pi_{\Lambda_n} \mathbb{P})$$

scelgo il campione con questa legge

Devo fare la stima dal basso $\forall\, \mathbb{Q}$ t.c. $\text{ent}(\mathbb{Q}|\mathbb{P}) < +\infty$ e ho fatto solo $\mathbb{Q}$ ergodiche.

Argomento di densità

Finora ho dimostrato stime dal basso con I_0 fatto così

$$I_0(\mathbb{Q}) = \begin{cases} \text{ent}(\mathbb{Q}|\mathbb{P}) & \text{se } \mathbb{Q} \text{ è ergodica} \\ \underline{+\infty}, & \text{altrimenti} \end{cases}$$

attenzione vuota: $\text{Prob}(-) \geq 0$

Per l'osservazione dell'altra volta posso migliorare tale stima facendo l'inviluppo esc di I_0,

$$I = \text{sc}^- I_0$$

è ... miracolo ... $I(\mathbb{Q}) = \text{ent}(\mathbb{Q}|\mathbb{P})$

<u>Lemma</u> [lemma di densità]

Sia $\mathbb{Q} \in \mathcal{P}_\theta(\Omega)$ $\exists$ successione $\mathbb{Q}_j$ con $\mathbb{Q}_j$ ergodiche t.c.

- $\mathbb{Q}_j \to \mathbb{Q}$
- $\text{ent}(\mathbb{Q}_j | \mathbb{P}) \longrightarrow \text{ent}(\mathbb{Q}|\mathbb{P})$

OSS $\mathcal{P}_\theta(\Omega)$ è un convesso peculiare: gli elementi estremali sono pure densi

Vedi lemma 6.9 libro Seppalainen per la dimostrazione.

Discuto invece come costruire $\mathbb{Q}_j$ per l'esempio di prima

$$\mathbb{Q} = (1-p)\,\delta_0 + p\,\delta_1$$

$$000 \cdots 0\ 111 \cdots 11\ \overbrace{000 \cdots 0}^{(1-p)j}\ \overbrace{111 \cdots 11}^{pj}\ 0000\ 11 \cdots 111 = \omega^j$$

ripeti periodicamente

ω è periodico $\rightsquigarrow$ vale il teorema ergodico

così non è però invariante per traslazioni

devi fare la media sulle traslazioni

$$\mathbb{Q}_j = \frac{1}{j} \sum_{i=0}^{j-1} \delta_{\theta_i \omega^j}$$

evidentemente $\mathbb{Q}_j \longrightarrow (1-p)\,\delta_0 + p\,\delta_1$

poiché $\mathrm{ent}(\cdot|\cdot)$ è affine

$$\mathrm{ent}(\mathbb{Q}\,|\,\mathbb{P}) = (1-p)\,\mathrm{ent}(\delta_0|\mathbb{P}) + p\,\mathrm{ent}(\delta_1|\mathbb{P})$$

si capisce (?) che anche $\mathrm{ent}(\cdot)$ si aggiusta