⊙ GRANDI DEVIAZIONI PER LA MISURA EMPIRICA: TEOREMA DI SANOV

Contesto $\quad \mathcal{X}$ SPAZIO POLACCO $\quad (\mathcal{X} = \mathbb{R}^n)$

per semplicità di discussione $\mathcal{X}$ compatto
($\mathcal{X}$ finito)

μ PROB. su $\mathcal{X}$ che volete "scoprire"

campionamento

$X_1, X_2, \dots, X_n$ iid con legge μ

la situazione di Cramer era $\mathcal{X} = \mathbb{R}^n$
e media empirica (media campionaria) $\frac{1}{n}(X_1 + \cdots + X_n)$

ora vogliamo tenerci tutte le informazioni del campione

Misura empirica

$$\Pi_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

Π_n V.A. a valori $\mathcal{P}(\mathcal{X})$

A parte scambiare gli indici, Π_n ha tutte le informazioni sul campione $X_1, \dots, X_n$

Prop (Legge dei grandi numeri) $\quad \Pi_n \xrightarrow{n \to \infty} \mu$

in probabilità (o, equivalentemente, in legge)

ovvero:

$\forall \varepsilon > 0$

$$\mathbb{P}(d(\Pi_n, \mu) > \varepsilon) \underset{n \to \infty}{\longrightarrow} 0$$

distanza per la convergenza debole in $\mathcal{P}(\mathcal{X})$

dim

traggo vantaggio dalla compattezza di $\mathcal{X}$, ma l'affermazione è vera in generale senza ipotesi aggiuntive.

$$\mathcal{X} \text{ compatto} \Rightarrow \mathcal{P}(\mathcal{X}) \text{ compatto} \underset{(\text{w.-v. des.})}{} \Rightarrow \mathcal{P}(\mathcal{P}(\mathcal{X})) \text{ compatto}$$

quindi per Prohorov

Legge (Π_n) è tight come prob. su $\mathcal{P}(\mathcal{X})$

Basta identificare i punti ~~limite~~ di accumulazione di tale ~~seq~~ successione

(il limite è non random quindi conv in prob ⟺ conv in legge)

Fisso $\phi \in C_b(\mathcal{X})$ e capisco a chi converge
, non serve poiché $\mathcal{X}$ è compatto.

$$\Pi_n(\phi) = \frac{1}{n}\sum_{i=1}^{n} \phi(X_i)$$

ora $Y_i = \phi(X_i)$ sono v.a. iid (a valori $\mathbb{R}$) limitate

per la legge debole dei grandi numeri

$$\Pi_n(\phi) \xrightarrow{\mathbb{P}} \mathbb{E}(\phi(X_1)) = \mu(\phi)$$

- Sia ora $\pi^\star$ punto di accumulazione di $\{\Pi_n\}$ ~~che~~ ricavo

$$\pi^\star(\phi) = \mu(\phi) \qquad \forall \phi \in C_b(\mathcal{X})$$

$$\Rightarrow \quad \pi^\star = \mu$$

OSS Enunciato equivalente, nel linguaggio delle probabilità invece che variabili aleatorie:

Legge (Π_n) $\longrightarrow \delta_\mu$ conv. debole in $\mathcal{P}(\mathcal{P}(\mathcal{X}))$

Sanov stabilisce il principio di grandi deviazioni associato alla legge dei grandi numeri $\pi_n \to \mu$
mentre Cramer riguarda la media empirica
$\frac{1}{n} \sum_{i=1}^{n} X_i$

<u>Sottoprodotto</u> Fisso $\phi : \mathcal{X} \to \mathbb{R}$ continua (e limitata)

$$S_n^{\phi} = \sum_{i=1}^{n} \phi(X_i)$$ successione di v.a. a valori in $\mathbb{R}$

Vogliamo ricavare Cramer da Sanov via principio delle contrazioni

Sia $\Phi : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$
definita da $\Phi(\pi) = \cancel{\pi(\phi)} \; \pi(\phi)$

per definizione della topologia su $\mathcal{P}(\mathcal{X})$ è continua (e limitata)
$$|\Phi(\pi)| = |\pi(\phi)| \leq |\phi|_\infty$$

ora
$$\frac{S_n^{\phi}}{n} = \Phi(\pi_n)$$

da contrazioni

$$I_{Cramer}(x) = \inf_{\pi \in \Phi^{-1}(\{x\})} I_{Sanov}(\pi)$$

<u>Ex</u> Ricavare I_{Cramer} da I_{Sanov}
(vedi enunciato seguente)

Slogan: " Sanov è il modo di fare grandi deviazioni per $\frac{S_n^\phi}{n}$ contemporaneamente per tutte le ϕ continue e limitate "

TEO (SANOV, linguaggio di v.a.)

π_n (succ. di v.a. a valori $\mathcal{P}(\mathcal{X})$)

soddisfa LDP con funzionale di tasso

$$I(\pi) = \mathrm{Ent}(\pi \mid \mu)$$

ovvero [con $\mathbb{P}$ = legge $(X_1, \dots, X_n, \dots)$ $= \mu^{\otimes \mathbb{N}}$]

$\forall$ C chiuso in $\mathcal{P}(\mathcal{X})$

$$\overline{\lim_n} \frac{1}{n} \log \mathbb{P}(\pi_n \in C) \leq - \inf_C \mathrm{Ent}(\cdot \mid \mu)$$

$\forall$ A aperto in $\mathcal{P}(\mathcal{X})$

$$\underline{\lim_n} \frac{1}{n} \log \mathbb{P}(\pi_n \in A) \geq - \inf_A \mathrm{Ent}(\cdot \mid \mu)$$

OSS Già sappiamo che $\mathrm{Ent}(\cdot \mid \mu)$ è lsc su $\mathcal{P}(\Omega)$ senza ipotesi di $\mathcal{X}$ compatto mostrare che ha insiemi di livello compatti (Ex)

OSS Sanov è ~~[illegible]~~ vero in generale per $\mathcal{X}$ polacco senza ipotesi aggiuntive su μ mentre Cramer necessita di condizioni di integrabilità

[si capisce media empirica è più delicata di misura empirica] dim

- Stima dall'alto

Disuguaglianza di Chebyshev esponenziale

esempio per Cramer $\qquad S_n = \sum_{i=1}^n X_i \qquad$ X_i v.a. a valori in $\mathbb{R}$

chiuso $= [a, +\infty)$

$$\mathbb{P}\left(\frac{S_n}{n} \geq a \right) \leq \dots$$

Fisso $\lambda > 0$

$$\mathbb{P}\left(\frac{S_n}{n} \geq a \right) = \mathbb{P}\left(\lambda S_n \geq \lambda a n \right)$$

chebyshev exp

$$\leq e^{-\lambda a n} \, \mathbb{E}\left(e^{\lambda S_n} \right) = e^{-\lambda a n} \, \mathbb{E}\left(e^{\lambda X_1 + \dots + \lambda X_n} \right)$$

$$= e^{-n \{ \lambda a - \log \mathbb{E}(e^{\lambda X_1}) \}}$$

$\uparrow$ fattorizza

ottimizza su λ e viene.

Sanov $\qquad$ Fisso $\phi \in C_b(\mathcal{X})$

$\mu^\phi \in \mathcal{P}(\mathcal{X})$ definita da $\qquad d\mu^\phi = \dfrac{e^\phi \, d\mu}{\mu(e^\phi)}$

$$\mathbb{P}_n = \mu^{\otimes n}$$

$$\mathbb{P}_n^\phi = \mu^{\phi \otimes n} \qquad n \text{ copie iid con legge } \mu^\phi$$

evidentemente

$$\frac{d\mathbb{P}_n^\phi}{d\mathbb{P}_n} = \frac{e^{\sum_{i=1}^n \phi(x_i)}}{[\mu(e^\phi)]^n} = e^{n \{ \pi_n(\phi) - \log \mu(e^\phi) \}}$$

Per $B \subset \mathcal{B}(\mathcal{X})$ Borelino

$$\mathbb{P}(\pi_n \in B) = \mathbb{P}_n(\pi_n \in B)$$

$$= \mathbb{E}_n\left(\mathbb{1}_B(\pi_n)\right) = \mathbb{E}_n^{\phi}\left(\frac{d\mathbb{P}_n}{d\mathbb{P}_n^{\phi}} \mathbb{1}_B(\pi_n)\right)$$

$$= \mathbb{E}_n^{\phi}\left(e^{-n\{\pi_n(\phi) - \log \mu(e^{\phi})\}} \mathbb{1}_B(\pi_n)\right)$$

$$\leq \sup_{\pi \in B} e^{-n\{\pi(\phi) - \log \mu(e^{\phi})\}} \underbrace{\mathbb{E}_n^{\phi}(\mathbb{1}_B(\pi_n))}_{\leq 1}$$

quindi

$$\overline{\lim_n} \frac{1}{n} \log \mathbb{P}(\pi_n \in B) \leq - \inf_{\pi \in B} \{\pi(\phi) - \log \mu(e^{\phi})\}$$

ora ottimizzo su $\phi \in C_b(\mathcal{X})$

$$\overline{\lim_n} \frac{1}{n} \log \mathbb{P}(\pi_n \in B) \leq - \sup_{\phi} \inf_{\pi \in B} \{\pi(\phi) - \log \mu(e^{\phi})\}$$

ricordo che

$$\mathrm{Ent}(\pi|\mu) = \sup_{\phi \in C_b(\mathcal{X})} \{\pi(\phi) - \log \mu(e^{\phi})\}$$

se potessi scambiare $\sup_{\phi}$ con $\inf_{\pi \in B}$

verrebbe "giusto"

$$\ldots \leq - \inf_{\pi \in B} \mathrm{Ent}(\pi|\mu)$$

Non è proprio giusto, ma quasi

$B \to K$ compatto

lo ricopro con un numero finito di palle piccole

...

ecco i dettagli

[minimax]

<u>Lemma</u> $\forall$ $\mathcal{X}$ Polacco $K \subset\subset \mathcal{X}$ compatto

J_α $\alpha \in A$ (arbitrario) famiglia di funz. USC

Allora

$$\inf_{\substack{\mathcal{O}_1,\dots,\mathcal{O}_n \\ K \subset \mathcal{O}_1 \cup \dots \cup \mathcal{O}_n}} \max_{i=1,\dots,n} \inf_{\alpha \in A} \sup_{y \in \mathcal{O}_i} J_\alpha(y)$$

$$\leq \sup_{x \in K} \inf_{\alpha \in A} J_\alpha(x)$$

<u>dim</u> omessa

Finisco stima dall'alto per Sanov

$K \subset\subset \mathcal{P}(\mathcal{X})$ compatto

lo ricopro $K \subset \mathcal{O}_1 \cup \dots \cup \mathcal{O}_k$

$$\mathbb{P}(\pi_n \in K) \leq \mathbb{P}(\pi_n \in \mathcal{O}_1 \cup \dots \cup \mathcal{O}_k)$$

$$\leq k \max_{i=1,\dots,k} \mathbb{P}(\pi_n \in \mathcal{O}_i)$$

Uso ~~il~~ ~~Hölder~~ la stima trovata via Chebyshev

$$\frac{1}{n} \log \mathbb{P}(\pi_n \in K) \le \underbrace{\frac{1}{n} \log k}_{\downarrow \atop 0} + \max_{i=1..k} \frac{1}{n} \log \mathbb{P}(\pi_n \in \theta_i)$$

$$\overline{\lim_n} \frac{1}{n} \log \mathbb{P}(\pi_n \in K) \le \max_{i=1,..,k} \inf_{\phi \in C_b(X)} \sup_{\pi \in \theta_i} J_\phi(\pi)$$

con $J_\phi(\pi) = -\{ \pi(\phi) - \log \mu(e^\phi) \}$

che è continuo rispetto a π
(ϕ è fissato)

ottimizzo sui ricoprimenti aperti e dal lemma ricavo

$$\ldots \le \sup_{\pi \in K} \inf_{\phi \in C_b(X)} \{ -(\pi(\phi) - \log \mu(e^\phi)) \}$$

$$= - \inf_{\pi \in K} \underbrace{\sup_{\phi \in C_b(X)} \{ \pi(\phi) - \log \mu(e^\phi) \}}_{\shortparallel \atop Ent(\pi | \mu)}$$

OSS Senza l'ipotesi di X compatto serve un po' di lavoro per ricondurre la stima dall'alto da $\forall C$ chiuso a $\forall K$ compatto

[tightness esponenziale]