$E = \{a_1 \_ a_q\} \simeq \{1, 2, \_ q\}$

$(E, \mathcal{P}(E), \mu)$ $\quad \mu(\{i\}) = p_i \quad p_1 \_ p_q \geq 0 \quad \sum_{i=1}^{q} p_i = 1$

$\{X_n\}_{n\in\mathbb{N}}$ i.i.d $X_n(\Omega) \subseteq E \quad \mathbb{P}_{X_n} = \mu \quad \mathbb{P}(X_n = i) = p_i \quad \forall i = 1 \_ q$
$\forall n \in \mathbb{N}$

$$\mathbb{P}((X_1 \_ X_n) = (i_1 \_ i_n)) = \prod_{j=1}^{n} p_{i_j}$$

$$Y_n(\omega) = -\ln p_{X_n(\omega)} \quad \text{i.i.d.}$$

$$\mathbb{E}[Y_n] = \mathbb{E}[-\ln p_{X_n}] = \int_{\mathbb{R}} -\ln p(x)\,\mu(dx) =$$

$$= \sum_{i=1}^{q} (-\ln p_i)\,\mu(\{i\}) = -\sum_{i=1}^{q} p_i \ln(p_i)$$

$$\lim_{x \to 0^+} x \ln(x) = 0$$

$$\mathbb{E}[Y_n^2] = \mathbb{E}[\ln^2 p_{X_n}] = \int_{\Omega} \ln^2 p_{X_n(\omega)}\,\mathbb{P}(d\omega) =$$

$$= \int_{\mathbb{R}} \ln^2 p(x)\,\mu(dx) = \sum_{i=1}^{q} \ln^2 p_i\,\mu(\{i\}) = \sum_{i=1}^{q} p_i \ln^2 p_i$$

$$\mathrm{Var}[Y_n] = \sum_{i=1}^{q} p_i \ln^2 p_i - \left(-\sum_{i=1}^{q} p_i \ln(p_i)\right)^2$$

$\frac{1}{n}\sum_{k=1}^{n} Y_k$ converge in probabilità, L^2 e q.c. alla v.a. costante

$$H(p_1 \_ p_q) = -\sum_{i=1}^{q} p_i \ln(p_i)$$

Ben definita

$$\Delta = \left\{(p_1 \_ p_q) \in \mathbb{R}^q : p_i \geq 0 \quad \sum_{i=1}^{q} p_i = 1\right\}$$

$$H : (p_1, \ldots p_q) \in D \longmapsto -\sum_{i=1}^{q} p_i \ln(p_i) \in \mathbb{R}$$
è detta FUNZIONE ENTROPIA

$$\frac{1}{n}\sum_{k=1}^{n} Y_k(\omega) = \frac{1}{n}\sum_{k=1}^{n} -\ln P_{X_k(\omega)} = \frac{-1}{n} \ln \prod_{k=1}^{n} P_{X_k(\omega)} =$$
$$= \frac{1}{n}\left(-\ln \prod_{k=1}^{n} P_{X_k(\omega)}\right) \longrightarrow H(p_1 \ldots p_q)$$

$$H(p_1, \ldots p_q) = -\sum_{i=1}^{q} p_i \ln(p_i)$$

$$D = \left\{(p_1 \ldots p_q) : p_i \geq 0 \quad \sum_{i=1}^{q} p_i = 1\right\}$$

$\mathbb{R}^n$
$(x_1 \ldots x_n)$
$x_1 + x_2 \ldots + x_n = 1$
$(1, \ldots, 1)$

$$p_q = 1 - \sum_{i=1}^{q-1} p_i$$

$$H(p_1 \ldots p_q) = \tilde{H}(p_1 \ldots p_{q-1}) = -\sum_{i=1}^{q-1} p_i \ln(p_i) -$$
$$-\left(\left(1 - \sum_{i=1}^{q-1} p_i\right) \ln\left(1 - \sum_{i=1}^{q-1} p_i\right)\right)$$

$$j = 1 \ldots q-1 \quad \frac{\partial \tilde{H}}{\partial p_j} = -\ln(p_j) - \cancel{p_j}\,\frac{1}{\cancel{p_j}} - \left(-\ln\left(1 - \sum_{i=1}^{q-1} p_i\right) + \cancel{\left(1 - \sum_{i=1}^{q-1} p_i\right)}\,\frac{-1}{\cancel{1 - \sum_{i=1}^{q-1} p_i}}\right)$$
$$= -\ln(p_j) + \ln(p_q) = 0$$

$$\forall j = 1 \ldots q-1 \quad \ln(p_j) = \ln(p_q)$$

$$p_j = p_q \quad \forall j = 1 \ldots q-1$$

L'unico pto stazionario è $\left(\frac{1}{q}, \frac{1}{q}, \ldots \frac{1}{q}\right)$ che è interno a D

$$H\left(\frac{1}{q}, \frac{1}{q}, \dots, \frac{1}{q}\right) = -\sum_{i=1}^{q} \frac{1}{q} \ln\left(\frac{1}{q}\right) = -\cancel{q} \frac{1}{\cancel{q}_{1}} \ln\left(\frac{1}{q}\right)$$

$$= \ln(q)$$

$$\Delta = \left\{ (p_1 \dots p_q) \in \mathbb{R}^q : p_i \geq 0 \quad \sum_{i=1}^{q} p_i = 1 \right\}$$

$q = 2$ $\quad \Delta = \left\{ (p_1, p_2) \in \mathbb{R}^2 : p_1 \geq 0 \quad p_2 \geq 0 \quad p_1 + p_2 = 1 \right\}$

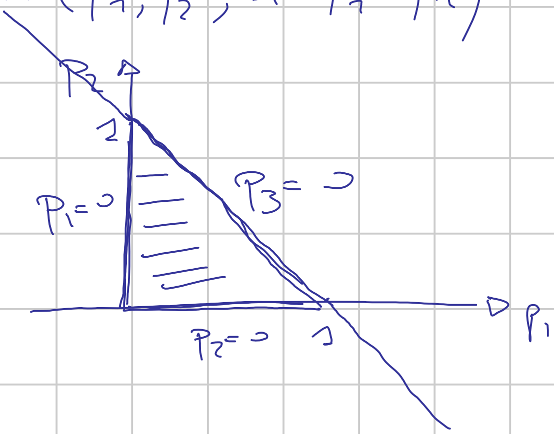
$p_2 = 1 - p_1$ $\quad \tilde{H}(p_1) = H(p_1, 1 - p_1)$

$p_1 \geq 0 \quad 1 - p_1 \geq 0 \quad p_1 \in [0, 1]$

$q = 3$ $\quad \Delta = \left\{ (p_1, p_2, p_3) \in \mathbb{R}^3 \quad p_1 \geq 0 \quad p_2 \geq 0 \quad p_1 + p_2 + p_3 = 1 \right\}$

$p_3 = 1 - p_1 - p_2$ $\quad \tilde{H}(p_1, p_2) = H(p_1, p_2, 1 - p_1 - p_2)$

$$\begin{cases} p_1 \geq 0 \\ p_2 \geq 0 \\ 1 - p_1 - p_2 \geq 0 \end{cases} \qquad \begin{array}{l} p_1 \geq 0 \\ p_2 \geq 0 \\ p_1 + p_2 \leq 1 \end{array}$$



$$0 \leq H(p_1 \dots p_q) \leq \ln(q)$$

— o —

$$\frac{1}{n} \sum_{k=1}^{n} Y_k(\omega) \longrightarrow H(p_1 \dots p_q)$$

$$\exp\left(-\frac{1}{n} \sum_{k=1}^{n} Y_k\right) = \left(\exp \sum_{k=1}^{n} (-Y_k)\right)^{1/n} =$$

$$= \left(\prod_{k=1}^{n} \exp(\underbrace{-Y_k}_{\ln P_{X_k}})\right)^{1/n} = \left(\prod_{k=1}^{n} \exp \ln(P_{X_k})\right)^{1/n} =$$

~

$$\left(\prod_{k=1}^{n} p_{k_k}\right)^{1/n} = \sqrt[n]{\prod_{k=1}^{n} p_{X_k}}$$

$$\Rightarrow \quad \sqrt[n]{\prod_{k=1}^{n} p_{X_k}} \longrightarrow e^{-H(p_1 - p_q)}$$

$$0 \le H(p_1, - p_q) \le \ln(q)$$

$$\ln\frac{1}{q} = -\ln q \le -H(p_1 - p_q) \le 0$$

$$\frac{1}{q} \le e^{-H(p_1 - p_q)} \le 1$$

— o —

TEMPI DI ATTESA

$\{X_n\}_{n\in\mathbb{N}}$ successione di v.a. su $(\Omega, \mathcal{E}, \mathbb{P})$

$X_n \ge 0$ – Supponiamo che valga uno dei seguenti set di ipotesi:

1) Hanno tutte lo stesso valore finito, varianze equilimitate e sono a due a due scorrelate

2) Sono i.i.d.

Sia $T_n(\omega) := \sum_{k=0}^{n} X_k(\omega) \qquad \omega \in \Omega \quad n \in \mathbb{N}$

$(\Rightarrow T_n(\omega) \le T_{n+1}(\omega))$

Per $t > 0 \qquad N_t(\omega) := \sup\{n \in \mathbb{N} : T_n(\omega) \le t\}$

Sia $E = \mathbb{E}[X_n]$ (se sono nel caso 2 può anche darsi che sia $E = +\infty$)

Allora $\displaystyle \lim_{t\to+\infty} \frac{N_t(\omega)}{t} = \frac{1}{E}$ – $\mathbb{P}$-q.c.

$$\underbrace{\qquad\qquad\qquad}_{X_k(\omega)}$$

$$T_n(\omega) = \sum_{k=0}^{n} X_k(\omega)$$

$$N_t(\omega) = \sup\{n \in \mathbb{N} : T_n(\omega) \leq t\}$$

$$[0, t] \qquad \frac{N_t(\omega)}{t}$$

<u>Bozza di dimostrazione</u> $\qquad T_n = \sum_{k=0}^{n} X_k$

$\frac{1}{n}\sum_{k=0}^{n} X_k = \frac{1}{n} T_n$ converge ad E in probabilità, L^2 e $\mathbb{P}$-q.c.

$$f : t \in [0, +\infty) \longmapsto \sum_{k=0}^{\lfloor t \rfloor} X_k(\omega)$$

- f non negativa
- f monotona non decrescente
- $\lim_{t \to +\infty} \frac{f(t)}{t} = E$

Sia $\psi(s) := \sup\{t \in [0, +\infty) : f(t) \leq s\}$

Allora $\lim_{s \to \infty} \frac{\psi(s)}{s} = \frac{1}{E}$

$$\psi(s) = \sup\{t \geq 0 : f(t) \leq s\} = \sup\{n \in \mathbb{N} : f(n) \leq s\}$$
$$= \sup\{n \in \mathbb{N} : T_n(\omega) \leq s\} = N_s(\omega)$$

TEOREMA CENTRALE DEL LIMITE

Sia $(\Omega, \mathcal{E}, \mathbb{P})$

$X : \Omega \to \mathbb{R}$ v.a. con legge F_X

Per $n \in \mathbb{N}$ $\quad X_n(\omega) := \frac{1}{n} + X(\omega)$

$\forall \omega \in \Omega \quad \lim_{n \to \infty} X_n(\omega) = X(\omega)$

$$\sup_{\omega \in \Omega} |X_n(\omega) - X(\omega)| = \frac{1}{n} \longrightarrow 0$$

$$F_{X_n}(t) = \mathbb{P}(X_n \le t) = \mathbb{P}\left(\frac{1}{n} + X \le t\right) = \mathbb{P}\left(X \le t - \frac{1}{n}\right) \overset{\text{"}}{=} F_X\left(t - \frac{1}{n}\right)$$

$$\lim_{n \to \infty} F_{X_n}(t) = \lim_{s \to t^-} F_X(s) = F_X(t^-)$$

$\lim_{n \to \infty} F_{X_n}(t) = F_X(t)$ SSE F_X è continua in t

CONVERGENZA IN LEGGE

Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di v.a. su $(\Omega, \mathcal{E}, \mathbb{P})$ e sia X un'ulteriore v.a. su tale spazio.

Dico che X_n CONVERGE IN LEGGE AD X

se

$$\lim_{n \to \infty} F_{X_n}(t) = F_X(t)$$

per ogni $t \in \mathbb{R}$ in cui F_X è continua.

— o —

DISTRIBUZIONE GAUSSIANA DI PARAMETRI (NORMALE)

$\mu \in \mathbb{R}$ e $\sigma > 0$

Si indica $N(\mu, \sigma^2)$ è la distribuzione A.C. associata alla densità

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad x \in \mathbb{R}$$

Se X è una v.a. con $\mathbb{P}_X = N(\mu, \sigma^2)$, allora

$$\mathbb{E}[X] = \mu \qquad e \qquad Var[X] = \sigma^2$$

Se $\mu = 0$ e $\sigma = 1$, la distribuzione $N(0,1)$ si dice DISTRIBUZIONE GAUSSIANA STANDARD; la legge associata si indica Φ e la densità è

$$f_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \qquad x \in \mathbb{R}$$

Se X_0 è una v.a. t.c. $\mathbb{P}_{X_0} = N(0,1)$, allora la v.a.

$X := \mu + \sigma X_0$ ha distribuzione $N(\mu, \sigma^2)$

<u>RICORDINO</u> Sia X una v.a. con distribuzione A.C. e densità $g(x)$

Siano $a, b \in \mathbb{R}$ con $a \neq 0$

Allora la v.a. $Y := aX + b$ è ancora A.C. con densità

$$h(x) = \frac{1}{|a|} g\left(\frac{x-a}{b}\right)$$

La legge Φ avendo densità $f_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$ che è una funzione pari, ha la proprietà

$$\Phi(t) + \Phi(-t) = 1$$

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx$$

$$\Phi(-t) = 1 - \Phi(t)$$

TEOREMA CENTRALE DEL LIMITE

Sia $\{X_n\}_{n\in\mathbb{N}}$ successione di v.a. i.i.d su uno spazio probabilizzato $(\Omega, \mathcal{E}, \mathbb{P})$ t.c.

$$\mathbb{E}[X_n] = E \quad , \quad Var[X_n] = \sigma^2$$

siano entrambe finite.

Sia $S_n := \sum_{k=1}^{n} X_k$.

Allora la v.a $\frac{S_n - nE}{\sigma\sqrt{n}}$ converge in legge ad una v.a. X t.c. $\mathbb{P}_X = N(0,1)$

cioè

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - nE}{\sigma\sqrt{n}} \leq t\right) = \Phi(t) \qquad \forall t \in \mathbb{R}$$

Inoltre il limite è uniforme ovvero

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \left| \mathbb{P}\left(\frac{S_n - nE}{\sigma\sqrt{n}} \leq t\right) - \Phi(t) \right| = 0$$

—— o ——

$$\lim_{n\to\infty} f_n(t) = f(t) \quad \forall t \in \mathbb{R} \qquad (1)$$

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} |f_n(t) - f(t)| = 0 \qquad (2)$$

(1) $\forall t \in \mathbb{R} \quad \forall \varepsilon > 0 \quad \exists \bar{n} = \bar{n}(\varepsilon, t)$ t.c.

$\forall n > \bar{n} \qquad |f_n(t) - f(t)| < \varepsilon$

(2) $\forall \varepsilon > 0 \quad \exists \bar{n} \quad$ t.c. $\quad \forall n > \bar{n}$

$$\sup_{t \in \mathbb{R}} |f_n(t) - f(t)| < \varepsilon$$

(2) $\forall \varepsilon > 0 \quad \exists \bar{n} = \bar{n}(\varepsilon) \quad$ t.c. $\quad \forall n > \bar{n}$

$$|f_n(t) - f(t)| < \varepsilon \quad \forall t$$

$f_n(t) = t^n \quad t \in [0,1] \quad$ converge (2) $\quad \begin{cases} 0 & t \in [0,1) \\ 1 & t = 1 \end{cases}$

ma il limite non è uniforme $\cancel{(2)}$

$$\frac{S_n - n\bar{\varepsilon}}{\sigma\sqrt{n}} \qquad \mathbb{E}[X_n] = \varepsilon \qquad Var[X_n] = \sigma^2$$

$$\mathbb{E}[S_n] = n\bar{\varepsilon} \qquad Var[S_n] = n\sigma^2$$

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var[S_n]}}$$

$$\mathbb{E}\left[\frac{S_n - n\bar{\varepsilon}}{\sigma\sqrt{n}}\right] = \frac{1}{\sigma\sqrt{n}} \mathbb{E}[S_n - n\bar{\varepsilon}] =$$

$$= \frac{1}{\sigma\sqrt{n}}(n\bar{\varepsilon} - n\bar{\varepsilon}) = 0$$ – valore stesso se $\mathbb{P}_X = N(0,1)$

$$Var\left[\frac{S_n - n\bar{\varepsilon}}{\sigma\sqrt{n}}\right] = \frac{1}{\sigma^2 n} Var[S_n - n\bar{\varepsilon}] = \frac{1}{\sigma^2 n} Var[S_n] =$$

$$= \frac{1}{\sigma^2 n} n\sigma^2 = 1$$

— o —

$$\lim_{n \to \infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \varepsilon\right| \geq \delta\right) = 0 \qquad \forall \delta > 0$$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \varepsilon\right| \geq \delta\right) = \mathbb{P}\left(\frac{S_n}{n} - \varepsilon \geq \delta\right) + \mathbb{P}\left(\frac{S_n}{n} - \varepsilon \leq -\delta\right)$$

$\to \mathbb{P}\left(\frac{S_n}{n} - \varepsilon \leq -\delta\right) \longrightarrow 0$

$\mathbb{P}\left(\frac{S_n}{n} - \varepsilon \geq \delta\right) \longrightarrow 0$

$$\mathbb{P}\left(\frac{S_n}{n}-E>\delta\right)\to 0$$
$$\left(=1-\mathbb{P}\left(\frac{S_n}{n}-E\le\delta\right)\right.$$

$$\begin{cases}\mathbb{P}\left(\frac{S_n}{n}-E\le-\delta\right)\to 0 \\ \mathbb{P}\left(\frac{S_n}{n}-E\le\delta\right)\to 1\end{cases} \quad \forall\delta>0$$

$$F_{\frac{S_n}{n}-E}(t)\to\begin{cases}0 & t<0\\ 1 & t>0\end{cases}$$

$$\frac{S_n}{n}-E=\frac{S_n-nE}{n}$$

$$\mathbb{P}\left(\frac{S_n-nE}{\sigma\sqrt{n}}\in(a,b]\right)=\mathbb{P}\left(a<\frac{S_n-nE}{\sigma\sqrt{n}}\le b\right)=$$
$$=\mathbb{P}\left(\frac{S_n-nE}{\sigma\sqrt{n}}\le b\right)-\mathbb{P}\left(\frac{S_n-nE}{\sigma\sqrt{n}}\le a\right)$$
$$\to\Phi(b)-\Phi(a)$$

TEOREMA DI BERRY-ESSEEN

Sia $\{X_n\}$ successione v.a. i.i.d. con
$E[X_n]=0$, $Var[X_n]=\sigma^2$ e $E[|X_n|^3]=\gamma<+\infty$

Allora
$$\left|\mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}}\le t\right)-\Phi(t)\right|\le\frac{C}{\sqrt{n}} \qquad \forall t\in\mathbb{R}$$

$$C=0.8\,\frac{\gamma}{\sigma^3}$$

$\{X_n\}$ successione di v.a. i.i.d.
con $E[X_n]=E$ $Var[X_n]=\sigma^2$ $E[|X_n|^3]$ finito
$Y_n:=X_n-E$ soddisfa le ipotesi del Teorema.

$$\gamma := \mathbb{E}\left[|Y_n|^3\right] = \mathbb{E}\left[|X_n - E|^3\right]$$

$$|X_n - E| \le |X_n| + |E|$$

$$|X_n - E|^3 \le \underbrace{|X_n|^3} + 3\underbrace{|X_n|^2}\,\underbrace{|E|} + 3\underbrace{|X_n|}\,\underbrace{|E|^2} + \underbrace{|E|^3}$$

$$\gamma = \mathbb{E}\left[|X_n - E|^3\right] < +\infty$$

$$\tilde{S}_n = \sum_{k=1}^{n} Y_k = \sum_{k=1}^{n} (X_k - E) = \sum_{k=1}^{n} X_k - nE = S_n - nE$$

$$\left| \mathbb{P}\left(\frac{S_n - nE}{\sigma\sqrt{n}} \le t \right) - \Phi(t) \right| \le \frac{C}{\sqrt{n}} \qquad \forall t \in \mathbb{R}$$

$$\frac{S_n - nE}{\sigma\sqrt{n}} \le t \qquad\qquad S_n \le nE + t\sigma\sqrt{n}$$

$$\left| \mathbb{P}\big(S_n \le \underbrace{nE + t\sigma\sqrt{n}}_{s \in \mathbb{R}}\big) - \Phi(t) \right| \le \frac{C}{\sqrt{n}} \qquad \forall t \in \mathbb{R}$$

$$s = nE + t\sigma\sqrt{n} \qquad\qquad t = \frac{s - nE}{\sigma\sqrt{n}}$$

$$\left| \underbrace{\mathbb{P}(S_n \le s)}_{F_{S_n}(s)} - \Phi\left(\frac{s - nE}{\sigma\sqrt{n}}\right) \right| \le \frac{C}{\sqrt{n}} \qquad \forall s \in \mathbb{R}$$

$$F_{S_n}(s) \sim \Phi\left(\frac{s - nE}{\sigma\sqrt{n}}\right)$$

, n sufficientemente grande

VETTORI E MATRICI STOCASTICHE

$p_1 \ldots, p_q \qquad p_i \geq 0 \qquad \sum_{i=1}^{q} p_i = 1$

$(p_1 \ldots p_q) \qquad p_i \geq 0 \qquad \sum_{i=1}^{q} p_i = 1$

Indico con $\mathbb{R}^{N*}$ i vettori riga a N componenti reali e indico con $\mathcal{T}$ l'insieme dei vettori riga a componenti non negative e a somma 1

$$\mathcal{T} := \left\{ (x_1 \ldots x_N) \in \mathbb{R}^{N*} : x_i \geq 0 \quad \sum_{i=1}^{N} x_i = 1 \right\}$$

$\mathcal{T}$ si dice INSIEME DEI VETTORI STOCASTICI

Sia $P \in M_{N \times N}(\mathbb{R})$ una matrice $N \times N$.
Considero l'applicazione lineare

$$\underline{P} : x \in \mathbb{R}^{N*} \mapsto xP \in \mathbb{R}^{N*}$$

$$x = (x_1 \ldots x_N) \qquad P = (P^i_j)_{i,j = 1 \ldots N}$$

$$(xP)_j = \sum_{i=1}^{N} x_i P^i_j \qquad \forall j = 1 \ldots N$$

$e^1 \ldots \ldots e^N \qquad e^1 = (1, 0, \ldots 0)$

$\qquad e^2 = (0, 1, 0, \ldots 0)$

$$\underline{P}(e^k) = e^k P$$

$$(e^k P)_j = \sum_{i=1}^{N} (e^k)_i P^i_j = P^k_j \Rightarrow e^k P \text{ è la } k\text{-esima riga di } P$$

$P \in M_{N \times N}(\mathbb{R})$ si dice MATRICE STOCASTICA

se ogni sua riga è un vettore stocastico cioè

$$P^i_j \geq 0 \quad \forall i, j = 1 \ldots N$$

$$\sum_{j=1}^{N} P^i_j = 1 \quad \forall i = 1 \ldots N$$